



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년03월30일
(11) 등록번호 10-1843967
(24) 등록일자 2018년03월26일

(51) 국제특허분류(Int. Cl.)
G01N 33/574 (2006.01) C12Q 1/68 (2018.01)
G01N 33/15 (2006.01) G01N 33/68 (2006.01)
(21) 출원번호 10-2012-7022771
(22) 출원일자(국제) 2011년02월04일
심사청구일자 2016년02월02일
(85) 번역문제출일자 2012년08월30일
(65) 공개번호 10-2012-0130329
(43) 공개일자 2012년11월30일
(86) 국제출원번호 PCT/IB2011/000382
(87) 국제공개번호 WO 2011/095894
국제공개일자 2011년08월11일
(30) 우선권주장
61/337,465 2010년02월04일 미국(US)
(56) 선행기술조사문헌
W02009103061 A2*
(뒷면에 계속)

(73) 특허권자
각코우호우진 지치 이카다이가쿠
일본국 도쿄도 치요다쿠 히라카와초 2초메 6반 3
고
(72) 발명자
마노, 히로유키
일본 113-0033 도쿄 분쿄구 혼고 4-13-3
최, 영, 엘.
일본 111-0056 도쿄 타이토쿠 코지마 2-3-4-204
소다, 마나부
일본 329-0434 도치기 시모즈케시 이스트 기온
5-5-4
(74) 대리인
특허법인 남앤드남

전체 청구항 수 : 총 25 항

심사관 : 이수진

(54) 발명의 명칭 A L K 억제제에 대한 선천적 또는 후천적 내성을 갖는 암의 확인, 평가, 및 치료 방법

(57) 요약

본 출원은 ALK 돌연변이에 대해 양성인 암(들)을 갖는 피검체가 ALK 억제제를 이용한 치료에 반응할 지의 여부 및/또는 상기 암(들)을 갖는 환자가 비교적 느린 질병 진행을 가질 수 있는지의 여부를 결정하기 위한 조성물, 키트, 및 방법을 기재한다. 본 출원은 상기 암을 갖는 환자에서 질병의 시간 경과를 예측하기 위한 방법을 추가로 기재한다.

(56) 선행기술조사문헌

US20090156475 A1

US20090226443 A1

Lu L et al., Biochemistry(2009), Vol.48(16):
3600-3609.

Soda M et al., Nature(2007), Vol.448(7153):
561-566.

Choi et al., Cancer Resource(2008),
Vol.68(13): 4971-4976.

*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

a. 환자로부터 샘플을 수집하는 단계; 및

b. 하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자의 존재를 검출하기 위해 샘플을 분석하는 단계를 포함하는, ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 피검체를 확인하는 방법으로서,

하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자의 존재가 상기 피검체가 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된 것을 나타내고,

상기 돌연변이 ALK 폴리뉴클레오티드 분자는 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는 폴리펩티드를 엔코딩하고,

상기 ALK 억제제는 PF-02341066인, 방법.

청구항 2

a. 환자로부터 샘플을 수집하는 단계; 및

b. 하나 이상의 돌연변이 ALK 폴리펩티드의 발현 수준, 구조, 및 활성으로 구성된 군으로부터 선택된 하나 이상을 검출하기 위해 샘플을 분석하는 단계를 포함하는, ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 피검체를 확인하는 방법으로서,

하나 이상의 돌연변이 ALK 폴리펩티드의 존재가 상기 피검체가 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된 것을 나타내고,

상기 돌연변이 ALK 폴리펩티드는 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하고,

상기 ALK 억제제는 PF-02341066인, 방법.

청구항 3

제 1항 또는 제 2항에 있어서, 피검체가 ALK 억제제로 이전에 치료되지 않았거나, ALK 억제제로 이전에 치료되었고, ALK 억제제에 대해 적어도 부분적인 내성이 발달한 피검체인 방법.

청구항 4

제 1항 또는 제 2항에 있어서, 암이 역형성큰세포림프종(anaplastic large cell lymphoma), 신경모세포종, 유방암, 결장직장암, 염증성 근육섬유모세포 종양, 및 비소세포폐암으로 구성된 군으로부터 선택되는 방법.

청구항 5

삭제

청구항 6

제 1항 또는 제 2항에 있어서, 샘플이 가래, 기관지폐포 세척액(bronchoalveolar lavage), 흉막삼출액(pleural effusion), 조직, 전혈, 혈청, 혈장, 구강 채취물(buccal scrape), 타액, 뇌척수액, 소변, 대변, 순환 종양 세포, 순환 핵산, 및 골수로 구성된 군으로부터 선택되는 방법.

청구항 7

제 6항에 있어서, 샘플이 조직이고, 조직이 종양 또는 암 조직인 방법.

청구항 8

제 1항 또는 제 2항에 있어서, 샘플이 세포를 포함하는 방법.

청구항 9

삭제

청구항 10

제 1항에 있어서, 하나 이상의 ALK 돌연변이가 핵산 하이브리드화 검정에 의해 평가되는 방법.

청구항 11

제 1항에 있어서, 하나 이상의 ALK 돌연변이가 중합효소 연쇄반응에 의해 평가되는 방법.

청구항 12

제 2항에 있어서, 하나 이상의 ALK 폴리펩티드의 발현 수준이 하나 이상의 ALK 폴리펩티드에 특이적으로 결합하는 시약을 이용하여 검출되는 방법.

청구항 13

제 12항에 있어서, 시약이 항체, 항체 유도체, 및 항체 단편으로 구성된 군으로부터 선택되는 방법.

청구항 14

제 2항에 있어서, 하나 이상의 돌연변이 ALK 폴리펩티드의 양, 구조 및 활성으로 구성된 군으로부터 선택된 하나 이상이 대조군 샘플과 비교되는 방법.

청구항 15

제 1항 또는 제 2항에 있어서, 하나 이상의 ALK 돌연변이가 첫번째 시점 및 적어도 하나의 이후의 시점에서 평가되는 방법.

청구항 16

제 1항 또는 제 2항에 있어서, 샘플이 점라인(germline) 또는 체세포 유전체 DNA를 포함하는 방법.

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드에 특이적으로 결합하는 시약; 및 사용설명서를 포함하는, ALK 억제제를 이용한 치료에 대한 암 환자의 약제민감성(chemosensitivity)을 결정하기 위한 키트로서,

상기 돌연변이 ALK 폴리뉴클레오티드 분자는 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는 폴리펩티드를 엔코딩하고,

상기 돌연변이 ALK 폴리펩티드는 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하고,

상기 ALK 억제제는 PF-02341066인, 키트.

청구항 22

제 21항에 있어서, ALK 억제제를 추가로 포함하는 키트.

청구항 23

제 21항에 있어서, 시약이 하나 이상의 폴리뉴클레오티드 프로브를 포함하고, 이러한 프로브 각각이 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는 폴리펩티드를 엔코딩하는 뉴클레오티드 서열에 상보적인 폴리뉴클레오티드 서열을 포함하는 키트.

청구항 24

제 23항에 있어서, 프로브가 50 내지 10^7 개의 뉴클레오티드 길이의 폴리뉴클레오티드를 포함하는 키트.

청구항 25

제 23항에 있어서, 프로브가 올리고뉴클레오티드, cDNA 분자, RNA 분자, 및 핵염기를 포함하는 합성 유전자 프로브로 구성된 군으로부터 선택되는 키트.

청구항 26

제 21항에 있어서, 시약이 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는 폴리펩티드에 대한 항체, 항체 유도체, 또는 항체 단편을 포함하는 키트.

청구항 27

(a) 하나 이상의 돌연변이 ALK 폴리펩티드를 엔코딩하는 작제물로 트랜스팩션된 포유동물 세포를 시험 화합물과 접촉시키는 단계; 및

(b) 하나 이상의 돌연변이 ALK 폴리펩티드의 활성화에 대해 포유동물 세포를 평가하는 단계를 포함하는, 시험 화합물이 하나 이상의 돌연변이 ALK 폴리펩티드의 활성을 조절하는지의 여부를 결정하는 방법으로서,

대조군 실험에 비한 시험 화합물의 존재하에서 현저하게 조절된 활성이, 시험 화합물이 하나 이상의 돌연변이 ALK 폴리펩티드의 조절인자인 것으로 확인되고,

상기 돌연변이 ALK 폴리펩티드는 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는, 방법.

청구항 28

삭제

청구항 29

제 27항에 있어서, 대조군이 SEQ ID NO: 4에서 C1156Y 및 L1196M 중 하나 이상의 돌연변이를 포함하는 야생형 ALK 폴리펩티드를 발현하는 포유동물 세포를 포함하는 방법.

청구항 30

제 29항에 있어서, 하나 이상의 돌연변이 ALK 폴리펩티드의 활성이 ATP 결합, 티로신 키나제 활성화, 암세포 증식, 종양 성장, 종양 수, 아포토시스, 및 종양 전이로 구성된 군으로부터 선택되는 방법.

청구항 31

제 27항에 있어서, 대조군 실험이 시험 화합물의 부재하에서 하나 이상의 돌연변이 ALK 폴리펩티드를 발현하는 포유동물 세포를 포함하는 방법.

청구항 32

제 31항에 있어서, 하나 이상의 돌연변이 ALK 폴리펩티드의 활성이 ATP 결합, 티로신 키나제 활성화, 암세포 증식, 종양 성장, 종양 수, 아포토시스, 및 종양 전이로 구성된 군으로부터 선택되는 방법.

발명의 설명

발명의 내용

[0001] 관련 출원

[0002] 본 출원은 2010년 2월 4일에 출원된 미국 가특허출원 일련 번호 61/337,465호에 대해 우선권을 주장한다.

[0003] 발명의 배경

[0004] 티로신 키나제는 아데노신 트리포스페이트의 말단 포스페이트의 이동을 통해 단백질 기질의 티로신 잔기의 인산화를 촉매하는 효소 부류이다. 많은 상황에서, 티로신 키나제는 세포 증식, 암형성, 및 세포 분화를 포함하는 다수의 세포 기능에 대한 신호 전달에서 중요한 역할을 한다.

[0005] EML4-ALK는 비소세포폐암(NSCLC) 환자의 약 5%에서 존재하고, 인간 염색체 2의 짧은 아암(arm) 내에서의 작은 역전(inversion)의 결과로서 생성되는 융합형 단백질 티로신 키나제이다(Soda, M. *et al.* (2007) *Nature* 448:561-566; Mano, H. (2008) *Cancer Sci.* 99:2349-2355). EML4-ALK는 각각의 단량체의 EML4 영역 내의 코일드-코일(coiled-coil) 도메인 사이의 상호작용의 결과로서 항시성(constitutive) 이합체화를 겪으며, 이에 의해 현저한 종양발생 활성을 획득한다. 폐 상피세포에서 특이적으로 EML4-ALK를 발현하는 트랜스제닉 마우스는 출생 직후에 둘 모두의 폐에서 수백개의 샘암종 결절(adenocarcinoma nodule)이 발달하고, ALK 티로신 키나제 활성의 특정 억제제의 경구 투여는 폐로부터 상기 결절을 신속히 근절시킨다(Soda, M. *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105:19893-19897). 상기 관찰은 상기 융합 키나제를 갖는 NSCLC의 암형성에서 EML4-ALK의 필수적 역할을 나타내고, 이는 상기 암에 대해 ALK 억제제를 이용한 분자적으로 표적화된 치료의 실행가능성을 추가로 뒷받침한다. 예를 들어, ALK 및 MET 둘 모두의 티로신 키나제 활성의 억제제 PF-02341066의 임상 시험이 EML4-ALK-양성 NSCLC의 치료에 대해 수행중이고, 이의 중간 결과는 유망하다(Kwak, E.L. *et al.* (2009) *J. Clin. Oncol.* 27(suppl): 15s (abstract 3509)). 그러나, EML4-ALK-양성 종양의 서브셋은 억제제에 반응하지 않으며, 치료 실패의 분자적 기본 원리는 공지되어 있지 않다.

[0006] PF-02341066에 더하여, 다른 티로신 키나제 억제제(TKI)가 암 환자에서 현저한 치료 활성을 갖는 것으로 밝혀졌다. 예를 들어, ABL1 및 KIT에 대한 TKI인 이마티니브 메실레이트(Imatinib mesylate)는 BCR-ABL1 융합 키나제에 대해 양성인 골수성 백혈병 또는 활성화된 KIT에 대해 양성인 위장관기질종양을 갖는 개체의 결과를 현저히 개선시킨다(Druker, B.J. *et al.* (2001) *N. Engl. J. Med.* 344: 1031-1037; Heinrich, M.C. *et al.* (2008) *J. Clin. Oncol.* 26:5360-5367). 또한, 표피세포 성장인자 수용체(EGFR)에 대한 TKI인 제피티니브(gefitinib) 및 에를로티니브(erlotinib)는 EGFR 활성화와 관련된 NSCLC의 치료에서 효과적이다(Mok, T.S. *et al.* (2009) *J. Clin. Oncol.* 27:5080-2087; Mok, T.S. *et al.* (2009) *N. Engl. J. Med.* 361 :947-957). 불행히도, 표적 종양의 서브셋은 치료 개시로부터 해당 TKI에 대해 난치성이거나, 최초 반응 후에 내성이 된다. TKI 결합의 방해로 발생시키는 ATP-결합 포켓의 형태에 직접적 또는 알로스테리적으로 영향을 미치는 표적 키나제에서의 이차 돌연변이가 치료 실패의 일부 환자에서 검출되었다(Deininger, M. *et al.* (2005) *Blood* 105:2640-2653; Kobayashi, S. *et al.* (2005) *N. Engl. J. Med.* 352:786-792; Pao, W. *et al.* (2005) *PLoS Med.* 2:e73; Shah, N.P. *et al.* (2002) *Cancer Cell* 2: 117-125). 따라서, 이상 발현 및/또는 활성과 관련된 장애를 확인하고, 평가하고, 방지하고, 치료하기 위한 조성물, 키트 및 방법을 보다 잘 개발하기 위해 EML4-ALK와 같은 티로신 키나제에 내성을 부여하는 돌연변이를 확인하는 것이 즉시 필요하다.

[0007] 발명의 개요

[0008] 본 발명은 적어도 공지된 ALK 억제제에 대한 내성을 부여하는 신규한 역형성 림프종 키나제(ALK) 돌연변이(들)의 확인을 기초로 한 암의 확인, 평가 및 치료를 위한 조성물, 방법, 및 키트를 제공한다. 이러한 ALK 돌연변이는 또한 신규한 ALK 돌연변이(들)에 의해 생성된 비정상적 ATP-결합 포켓과 합치할 수 있으며, ALK 활성을 억제하는 약학적 조성물의 확인과 임상적으로 관련된다.

[0009] 한 양태에서, 본 발명은 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 환자로부터 샘플을 수집하는 단계, 및 하나 이상의 돌연변이 ALK 폴리뉴클레오타이드 분자의 존재를 검출하기 위해 샘플을 분석하는 단계를 포함하는, ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 환자를 확인하는 방법을 제공하며, 상기 하나 이상의 돌연변이 ALK 폴리뉴클레오타이드 분자의 존재는 상기 환자가 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된 것을 나타

낸다.

[0010] 또 다른 양상에서, 본 발명은 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 환자로부터 샘플을 수집하는 단계, 및 하나 이상의 돌연변이 ALK 폴리펩티드의 발현 수준, 구조, 및/또는 활성을 검출하기 위해 샘플을 분석하는 단계를 포함하는, ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된, 암을 갖거나 암이 발병할 위험이 있는 환자를 확인하는 방법을 제공하며, 상기 하나 이상의 돌연변이 ALK 폴리뉴클레오티드의 존재는 상기 환자가 ALK 억제제를 이용한 치료에 대한 무반응의 위험이 증가된 것을 나타낸다.

[0011] 본 발명의 임의의 양태의 일부 구체예에서, 환자는 ALK 억제제로 이전에 치료되지 않았거나, ALK 억제제로 이전에 치료되었고, ALK 억제제에 대해 적어도 부분적인 내성이 발달하였다(예를 들어, PF-02341066, PDD, 2-메틸-11-(2-메틸프로필)-4-옥소-4,5,6,11,12,13-헥사히드로-2H-인다졸로[5,4-a]피롤로[3,4-c]카르바졸-8-일[4-(디메틸아미노)벤질]카르바메이트, (1S,2S,3R,4R)-3-({5-클로로-2-[(1-에틸-2,3,4,5-테트라히드로-6-메톡시-2-옥소-1H-1-벤즈아제핀-7-일)아미노]-4-피리미디닐}아미노)바이시클로[2.2.1]헵트-5-엔-2-카르복사미드, 및 NVP-TAE684). 다른 구체예에서, 암은 역형성큰세포림프종(anaplastic large cell lymphoma), 신경모세포종, 유방암, 결장직장암, 염증성 근육섬유모세포 종양, 및 비소세포폐암으로 구성된 군으로부터 선택된다. 또 다른 구체예에서, 샘플은 가래, 기관지폐포 세척액(bronchoalveolar lavage), 흉막삼출액(pleural effusion), 조직, 전혈, 혈청, 혈장, 구강 채취물(buccal scrape), 타액, 뇌척수액, 소변, 대변, 순환 종양 세포, 순환 핵산, 및 골수로 구성된 군으로부터 선택된다. 또 다른 구체예에서, 샘플은 세포 또는 조직을 포함한다. 일부 구체예에서, 조직은 종양 또는 암 조직이다. 다른 구체예에서, 하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드는 표 1에 나열된 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드로 구성된 군으로부터 선택된다. 또 다른 구체예에서, 하나 이상의 ALK 돌연변이는 핵산 하이브리드화 검정에 의해 평가된다. 또 다른 구체예에서, 하나 이상의 ALK 돌연변이는 중합효소 연쇄반응에 의해 평가된다. 다른 구체예에서, 하나 이상의 ALK 폴리펩티드의 발현 수준은 하나 이상의 ALK 폴리펩티드에 특이적으로 결합하는 시약(예를 들어, 항체, 항체 유도체, 및 항체 단편)을 이용하여 검출된다. 또 다른 구체예에서, 하나 이상의 돌연변이 ALK 폴리펩티드의 양, 구조 및/또는 활성은 대조군 샘플과 비교된다. 또 다른 구체예에서, 하나 이상의 ALK 돌연변이는 첫번째 시점 및 적어도 하나의 이후의 시점에서 평가된다. 다른 구체예에서, 샘플은 점라인(germline) 또는 체세포 유전체 DNA를 포함한다.

[0012] 또 다른 양태에서, 본 발명은 암을 갖거나 암이 발병할 위험이 있는 환자로부터 샘플을 수집하는 단계, 표 1에 나열된 하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자의 존재를 검출하기 위해 샘플을 분석하는 단계, 및 치료적 유효량의 ALK 억제제를 상기 환자에 투여하는 단계를 포함하는, 암을 갖거나 암이 발병할 위험이 있는 환자를 치료하는 방법을 제공한다. 일부 구체예에서, ALK 억제제는 PF-02341066, PDD, 2-메틸-11-(2-메틸프로필)-4-옥소-4,5,6,11,12,13-헥사히드로-2H-인다졸로[5,4-a]피롤로[3,4-c]카르바졸-8-일[4-(디메틸아미노)벤질]카르바메이트, (1S,2S,3R,4R)-3-({5-클로로-2-[(1-에틸-2,3,4,5-테트라히드로-6-메톡시-2-옥소-1H-1-벤즈아제핀-7-일)아미노]-4-피리미디닐}아미노)바이시클로[2.2.1]헵트-5-엔-2-카르복사미드, 및 NVP-TAE684로 구성된 군으로부터 선택된다. 다른 구체예에서, 환자는 ALK 억제제로 이전에 치료되지 않았거나, ALK 억제제로 이전에 치료되었고, ALK 억제제에 대해 적어도 부분적인 내성이 발달하였다.

[0013] 또 다른 양태에서, 본 발명은 하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드에 특이적으로 결합하는 시약, 및 사용설명서를 포함하는, ALK 억제제를 이용한 치료에 대한 암 환자의 약제민감성(chemosensitivity)을 결정하기 위한 키트를 제공한다. 일부 구체예에서, 키트는 ALK 억제제를 추가로 포함한다. 다른 구체예에서, 시약은 하나 이상의 폴리뉴클레오티드 프로브를 포함하고, 이러한 프로브 각각은 표 1에 나열된 뉴클레오티드 서열에 상보적이거나 표 1에 나열된 폴리펩티드를 엔코딩하는 뉴클레오티드 서열에 상보적인 폴리뉴클레오티드 서열(예를 들어, 올리고뉴클레오티드, cDNA 분자, RNA 분자, 및 핵염기를 포함하는 합성 유전자 프로브)을 포함한다. 또 다른 구체예에서, 프로브는 약 50 내지 10^7 개 뉴클레오티드 길이의 폴리뉴클레오티드를 포함한다. 또 다른 구체예에서, 시약은 표 1에 나열된 하나 이상의 폴리뉴클레오티드 서열에 의해 엔코딩되는 폴리펩티드에 대한 항체, 및 항체 유도체, 및 항체 단편을 포함한다.

[0014] 또 다른 양태에서, 본 발명은 하나 이상의 돌연변이 ALK 폴리펩티드를 엔코딩하는 작체물로 트랜스펙션된 포유동물 세포를 시험 화합물과 접촉시키는 단계, 및 하나 이상의 돌연변이 ALK 폴리펩티드의 활성에 대해 포유동물 세포를 평가하는 단계를 포함하는, 시험 화합물이 하나 이상의 돌연변이 ALK 폴리펩티드의 활성을 조절하는지의 여부를 결정하는 방법을 제공하며, 대조군 실험에 비해 시험 화합물의 존재하에서 현저하게 조절된 활성은, 시

염 화합물이 하나 이상의 돌연변이 ALK 폴리펩티드의 조절인자인 것으로 확인된다. 일부 구체예에서, 하나 이상의 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드는 표 1에 나열된 돌연변이 ALK 폴리뉴클레오티드 분자 또는 폴리펩티드로 구성된 군으로부터 선택된다. 다른 구체예에서, 대조군은 표 1에 나열된 폴리펩티드로 구성된 군으로부터 선택된 야생형 ALK 폴리펩티드를 발현하는 포유동물 세포를 포함한다. 또 다른 구체예에서, 하나 이상의 돌연변이 ALK 폴리펩티드의 활성은 ATP 결합, 티로신 키나제 활성, 암 세포 증식, 종양 성장, 종양 수, 아포토시스, 및 종양 전이로 구성된 군으로부터 선택된다. 또 다른 구체예에서, 대조군 실험은, 예를 들어, 하나 이상의 돌연변이 ALK 폴리펩티드의 활성(예를 들어, ATP 결합, 티로신 키나제 활성, 암 세포 증식, 종양 성장, 종양 수, 아포토시스, 및 종양 전이)에 의해 결정되는 바와 같은 시험 화합물의 부재하에서의 하나 이상의 돌연변이 ALK 폴리펩티드를 발현하는 포유동물 세포를 포함한다.

[0015] **도면의 간단한 설명**

[0016] **도 1**은 ALK 티로신 키나제 억제제에 대한 내성과 관련된 본 발명의 신규한 ALK 돌연변이를 도시한다. 도 1A는 EML4-ALK 단백질의 개략적 대표도를 도시한다. 키나제 도메인 내의 2개의 새로운 돌연변이의 위치가 제시되며, 키나제-도메인 또는 융합 cDNA의 증폭을 위한 PCR 프라이머의 위치가 각각 닫힌 화살표 및 열린 화살표로 표시된다. 도 1B는 ALK 키나제-도메인 cDNA의 딥 시퀀싱(deep sequencing)의 결과를 도시한다. NSCLC 세포주 H2228 또는 검체 ID J-#1, J-#12, J-#113, J-#127 또는 LK-#33으로부터의 약 1000 bp의 PCR 생성물을 GAII 시스템으로 시퀀싱하였다. 전체 판독 범위(전체) 및 미스매치된 판독(미스매치)에 대한 번호가 각각 청색 및 적색 다이아몬드를 갖는 키나제-도메인 cDNA의 각각의 위치에 제시된다. 삽입물은 J-#1 및 J-#113에 대한 cDNA의 5' 영역에 대한 확대도를 도시한다(녹색 직사각형에 의해 도시됨). 도 1C는 G4374 및 C4493 위치 주위에서의 ALK cDNA 클론에 대한 전기영동도를 도시한다. 처리전에 수득(개시)된 가래 및 재발 후에 수득(재발)된 흉막삼출액 중의 세포로부터 제조된 cDNA로 PCR을 수행하였다. 치환된 A 뉴클레오티드는 적색으로 제시된다.

[0017] **도 2**는 ALK cDNA의 G4374 및 C4493에 해당하는 위치 주위의 유전체 서열을 도시한다. 환자의 흉막삼출액 중의 세포로부터 분리된 유전체 DNA를 Platinum Taq DNA 중합효소(Invitrogen, Carlsbad, CA) 및 프라이머(5'-GGTAAGAAGTGGCTCACTCTTGAG-3' 및 5'-CACAACTGCAGCAAGACTGG-3')를 이용하여 15초 동안 94℃, 30초 동안 60℃ 및 2분 동안 68℃의 35주기 PCR에 적용시키고, 생성물을 pT7Blue-2 플라스미드(Takara Bio)에 라이게이션시켰다. 이후, 플라스미드의 삽입물을 3130xl Genetic Analyzer로 시퀀싱하여, G4374A(좌측 패널) 또는 C4493A(우측 패널) 변화를 함유하는 PCR 클론을 확인하였다. 치환된 A 뉴클레오티드는 적색으로 제시된다.

[0018] **도 3**은 PF-02341066으로 처리된 BA/F3 세포의 결과를 도시한다. EML4-ALK(야생형), EML4-ALK(C1156Y), EML4-ALK(L1196M), 또는 이중 돌연변이 EML4-ALK(C1156Y/L1196M)를 발현하는 BA/F3 세포를 48시간 동안 지정된 농도의 PF-02341066의 존재하에서 인큐베이션시켰고, 이후 세포 형태를 위상차 현미경으로 검사하였다. 스케일 바, 20 μm.

[0019] **도 4**는 ALK 티로신 키나제 억제제에 대한 내성과 관련된 본 발명의 신규한 ALK 돌연변이의 특성을 도시한다. 도 4A는 지정된 농도의 PF-02341066과 함께 48시간 동안 5×10^5 개 세포의 인큐베이션 후에 계수된 EML4-ALK(야생형), EML4-ALK(C1156Y), EML4-ALK(L1196M), 또는 이중 돌연변이 EML4-ALK(C1156Y/L1196M)를 발현하는 BA/F3 세포의 수를 도시한다. 살아있는 세포의 백분율은 야생형 EML4-ALK를 발현하는 BA/F3 세포와 관련하여 제시된다. 데이터는 3개의 독립적 실험으로부터의 평균 \pm s.d.이다. 도 4B는 EML4-ALK의 야생형 또는 돌연변이 형태의 티로신 인산화에 대한 PF-02341066의 효과를 도시한다. FLAG-태깅된 야생형 EML4-ALK 또는 이의 돌연변이를 발현하는 BA/F3 세포가 15시간 동안 지정된 농도의 PF-02341066에 노출되었고, 이후 EML4-ALK가 세포 용해질로부터 면역침전되었고, Tyr¹⁶⁰⁴-인산화된 ALK 또는 FLAG 에피토프(ALK)에 특이적인 항체를 이용한 면역블롯 분석에 적용되었다. EML4-ALK(KM)의 비활성화 돌연변이를 발현하는 세포를 음성 대조군으로 시험하였다. 도 4C는 BA/F3 세포(ALK 억제제에 노출되지 않음)로부터 면역침전된 FLAG-태깅된 야생형 EML4-ALK 또는 이의 돌연변이에 대한 시험관내 키나제 검정을 도시한다. 면역침전물을 [γ -³²P]ATP, 합성 펩티드, 및 지정된 농도의 PF-02341066과 인큐베이션시켰다. 펩티드 기질 면역침전물의 인산화를 FLAG에 대한 항체를 이용한 면역블롯 분석에 독립적으로 적용시켰다(하부 패널).

[0020] **도 5**는 ALK의 키나제 도메인에 대한 3차원 구조 모델을 도시한다. ALK의 아미노산 위치는 결합된 ATP 유사체를 갖는 인슐린 수용체의 결정 구조와 겹쳐졌다(월드 와이드 웹 pdj.org/index.html에서 이용가능한 일본 단백질 데이터뱅크(Protein Data Bank of Japan)에서 ID "1ir3"). 우측 패널은 좌측 패널에서의 모델의 좌측으로부터 관찰되는 단백질 구조를 도시한다. 헬릭스(helix) 및 β 시트가 각각 자홍색 및 오렌지색으로 제시된다.

헬릭스 αC, Cys¹¹⁵⁶, 및 Leu¹¹⁹⁶의 위치가 또한 표시된다.

[0021] **발명의 상세한 설명**

[0022] 본 발명은, 적어도 부분적으로, 암을 치료하는데 있어서 ALK 억제제의 효능 예측과 관련된, 예를 들어, 역형성 림프종 키나제(Anaplastic Lymphoma Kinase, ALK) 돌연변이를 포함하는 유전체의 특정 영역의 확인을 기초로 한다. 특히, ALK 억제제를 이용한 요법에 대해 적어도 부분적으로 내성인 폴리펩티드를 야기시킬 수 있는 신규한 ALK 유전자 돌연변이(예를 들어, EML4-ALK 폴리펩티드 엔코딩 돌연변이)가 본원에서 확인되었다. 본 발명은 올리뉴클레오티드-기반 마이크로어레이(Brennan, *et al.* (2004) *Cancer Res.* 64(14):4744-8; Lucito, *et al.* (2003) *Genome Res.* 13:2291-2305; Bignell *et al.* (2004) *Genome Res.* 14:287-295; Zhao, *et al.* (2004) *Cancer Research*, 64(9):3060-71)을 포함하나 이에 제한되지는 않는 당 분야에 공지된 기술, 및, 예를 들어, 중합효소 연쇄반응(PCR) 및 직접 시퀀싱-기반 방법을 포함하는 본원에 기재된 바와 같은 다른 방법을 이용하여 상기 특이적 유전체 영역을 확인하는 방법을 추가로 제공한다. 본 발명은 상기 방법에서 사용하기 위한 진단 키트를 추가로 제공한다.

[0023] 본 발명의 다양한 양태가 하기 서브섹션에서 보다 상세히 기재된다.

[0024] **I. 정의**

[0025] 단수 형태의 관사는 하나 이상(즉, 적어도 하나)의 이러한 관사의 문법적 대상을 의미하는 것으로 본원에서 사용된다. 예를 들어, "구성요소"는 하나의 구성요소 또는 하나 이상의 구성요소를 의미한다.

[0026] 용어 마커의 "변경된 양" 또는 마커의 "변경된 수준"은 대조군 샘플에서의 마커의 발현 수준 또는 카피수에 비해 암 샘플에서의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)과 같은 마커 또는 염색체 영역의 증가되거나 감소된 카피수, 및/또는 특정 마커 유전자 또는 유전자들의 증가되거나 감소된 발현 수준을 의미한다. 용어 마커의 "변경된 양"은 또한 정상의 대조군 샘플 내의 마커의 단백질 수준과 비교시 샘플, 예를 들어, 암 샘플 내의 마커의 증가되거나 감소된 단백질 수준을 포함한다.

[0027] 용어 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 "변경된 발현 수준"은 발현 또는 카피수를 평가하는데 이용되는 검정의 표준 오차보다 크거나 작은 암에 걸린 환자로부터 유래된 샘플과 같은 시험 샘플에서의 마커의 발현 수준 또는 카피수를 의미하며, 이는 대조군 샘플(예를 들어, 관련 질병에 걸리지 않은 건강한 피검체로부터의 샘플) 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 발현 수준 또는 카피수, 또는 여러 대조군 샘플 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 평균 발현 수준 또는 카피수의 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배, 또는 적어도 10배 또는 이를 초과할 수 있다. 변경된 발현 수준은 발현 또는 카피수를 평가하는데 이용되는 검정의 표준 오차보다 크거나 작으며, 이는 대조군 샘플(예를 들어, 관련 질병에 걸리지 않은 건강한 피검체로부터의 샘플) 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 발현 수준 또는 카피수, 또는 여러 대조군 샘플 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 평균 발현 수준 또는 카피수의 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배, 적어도 10배 또는 이를 초과한다.

[0028] 용어 마커의 "변경된 활성"은 정상의 대조군 샘플의 활성과 비교시 질병 상태, 예를 들어, 암 샘플에서 증가되거나 감소되는 마커의 활성을 의미한다. 마커의 변경된 활성은, 예를 들어, 변경된 마커의 발현, 변경된 마커의 단백질 수준, 변경된 마커의 구조, 또는, 예를 들어, 마커로서의 동일하거나 상이한 경로에서 관련되는 다른 단백질과의 변경된 상호작용의 결과일 수 있다.

[0029] 용어 마커의 "변경된 구조"는 정상 또는 야생형 유전자 또는 단백질과 비교시 마커 유전자 또는 마커 단백질 내의 돌연변이 또는 돌연변이들, 예를 들어, 마커의 발현 또는 활성에 영향을 미치는 돌연변이의 존재를 의미한다. 예를 들어, 돌연변이는 염색체간 및 염색체내 재배열, 치환, 결실, 및 삽입 돌연변이를 포함하나, 이에 제한되지는 않는다. 돌연변이는 마커의 코딩 또는 비-코딩 영역에 존재할 수 있다.

[0030] "역형성림프종 키나제" 및 "ALK"는 본원에 상호교환적으로 사용되며, 이는 임의의 공급원(예를 들어, 설치류, 인간, 및 다른 포유동물)으로부터 유래되는 자연 역형성림프종 키나제, 및 이의 특정 변이체 및 돌연변이를 의미한다. 일부 구체예에서, ALK 단백질은 NCBI 참조 서열 확인 번호 NP_004295에 의해 표시된다. 달리 지정하지 않는 한, 상기 용어는 인간 단백질을 의미한다. ALK를 엔코딩하는 유전자도 또한 본원에서 "ALK"로 언급될 수 있다. 일부 구체예에서, ALK 뉴클레오티드 서열은 NCBI 참조 서열 확인 번호 NM_004304.3 및 GenBank 등록

번호 29029631이며, 이에 포함되는 관련 서열(예를 들어, 코딩, 5' UTR, 3' UTR, 전사 시작, 번역 시작, 전사 종료, 번역 종료 등의 서열)은 당업자에 의해 용이하게 확인될 수 있다.

[0031]

또한, "역형성림프종 키나제" 및 "ALK"는 ALK 융합 키나제 및 이의 변이체를 포함하는 것으로 본원에서 또한 사용되며, 이는 당업자에게 널리 공지되어 있다. 이러한 ALK 융합 키나제 및 이의 변이체는 ALK 키나제 활성을 포함하며, 이는 본원에 기재된 바와 같은 돌연변이를 가져 ALK 억제제에 대해 ALK 키나제 활성이 내성이 되도록 할 수 있다. 대표적인 예는 EML4-ALK 변이체 1(AB274722.1;BAF73611.1), EML4-ALK 변이체 2(AB275889.1;BAF73612.1), EML4-ALK 변이체 3a(AB374361.1;BAG55003.1), EML4-ALK 변이체 3b(AB374362.1;BAG55004.1), EML4-ALK 변이체 4(AB374363.1;BAG75147.1), EML4-ALK 변이체 5a(AB374364.1;BAG75148.1), EML4-ALK 변이체 5b(AB374365.1; BAG75149.1), EML4-ALK 변이체 6(AB462411.1;BAH57335.1), EML4-ALK 변이체 7(AB462412.1;BAH57336.1), KIF5B-ALK(AB462413.1;BAH57337.1), NPM-ALK, TPM3-ALK, TFGXL-ALK, TFGL-ALK, TFGS-ALK, ATIC-ALK, CLTC-ALK, MSN-ALK, TPM4-ALK, MYH9-ALK, RANBP2-ALK, ALO 17-ALK, 및 C ARS-ALK(예를 들어, 전체내용이 참조로서 본원에 포함되는 문헌[Pulford *et al.*, (2004) *J. Cell. Physiol.* 199:330-358] 참조)를 포함한다. 또한, 당업자는 ALK 키나제 변이체가 ALK 키나제와 이의 융합 파트너 사이의 특정 융합 사건에 따라 발생할 수 있는 것을 이해할 것이다(예를 들어, 전체내용이 참조로서 본원에 포함되는 문헌[Horn and Pao, (2009) *J. Clin. Oncol.* 27:4247-4253]에 기재된 바와 같이, EML4는 적어도 엑손 2, 6a, 6b, 13, 14, 및/또는 15와 융합할 수 있다). 예를 들어, 하기와 같은 대표적 ALK 서열이 본원에 제공된다:

[0032]

표 1

야생형 ALK cDNA 서열 (NM_004304.3; GI:29029631) :

```

1  gggggcgcca gcggtggtag cagctggtac ctcccgcgcg ctctgttcgg agggtcgcgg
61  ggcaccgagg tgcttttcgg ccgcctcttg gtgcggccacc caaagccgcg ggcgtgatg
121 atgggtgagg agggggcggc aagatttcgg gcgcccctgc cctgaacgcc ctacgtgct
181 gccgcggggg ccgctccagt gcttgcgaac tctgaggagc cgaggcgccg gtgagagcaa
241 ggacgtcgca aactttgcgc gcgcgggggc tgggattcac gccagaaagt tcagcaggca
301 gacagtcgca agccttcctc cagcggagag atagcttgag ggtgcgcaag acggcagcct
361 ccgcctctcg ttcccgccca gaccgggcag aagagcttgg aggagccaaa aggaacgcaa
421 aaggcgccca ggacagcgtg cagcagctgg gagccgccgt tctcagcctt aaaagttgca
481 gagattggag gctgccccga gaggggacag accccagctc cgactgcggg gggcaggaga
541 ggacgttacc caactgccac ctcccctcaa ccatagtagt tcctctgtac cgagcgcagc
601 gagctacaga cggggcgcg gcactcgcg cgagagcgcg gaggtcaag gtcccagcca
661 gtgagccca gttgcttgag tgtctctgga ctgcctctg agcttcacag tctgtttcat
721 ttgactcct gctgcctcc gtgcagtgg gggaaagcaa gagacttgcg cgcacgcaca
781 gtcctcttga gatcaggtg aaggagccgc tgggtacca ggactgttca gaccccttcc
841 ccatctcggg gagagcgaag ggtgaggctg ggcgcggaga gcagtgtaaa cggcctcctc
901 cggcggggat ggagccatcg ggctcctgtg gctcctgcgc ctgctgcttt ccacggcagc
961 tgtgggctcc gggatgggga ccggccagcg cgcgggctcc ccagctgcgg ggcgcgcgct
1021 gcagcccccg gagccactca gctactcgcg cctgcagagg aagagtctgg cagttgactt
1081 cgtggtgccc tgcctcttcc gtgtctacgc ccgggacctc ctgctgccac catcctcctc
1141 ggagctgaag gctggcaggc ccgaggcccc cggtctgcta gctctggact gcgccccgct
1201 gctcaggttg ctggggcgcg cgcgggggt ctcttgacc gcgggttca cagccccggc
1261 agaggcccg aegctgtcca ggtgtctgaa ggcgggctcc gtgcgcaagc tccggcggtg
1321 caagcagttg gtgctggagc tgggcgagga ggcgatcttg gaggtttgct tcgggcccc
1381 cggggaggcg gctgtggggc tgcctcagtt caatctcagc gagctgttca gttgtgggat
1441 tcgccaaggc gaaggcgac tgaggatccg cctgatgcc gagagaagg cgtcggaggt
1501 ggcagagag ggaaggctgt ccgcggcaat tgcgcctcc cagccccgc ttctcttcca
1561 gatcttcggg actggtcata gctccttga atcaccaaca aacatgcctt ctctctctcc
1621 tgattatttt acatggaatc tcacctggat aatgaaagac tccttccctt tctgtctca
1681 tcgagccga tatgtcttg agtgacgtt tgacttcccc tgtgagctgg agtatcccc
1741 tccactgcat gacctcagga accagagctg gtctggcgc cgcattccct ccgaggaggc
1801 ctcccagatg gacttgctg atgggcctgg ggcagagcgt tctaaggaga tcccagagg
1861 ctcttttctc ctctcaaca cctcagctga ctccaagcac accatcctga gtcctggat
1921 gaggagcagc agtgagcact gcacactggc cgtctcggtg cacaggcacc tgcagccctc
1981 tggaaagtac attgcccagc tgctgcccc caacgaggct gcaagagaga tctctctgat
2041 gcccaactca gggaagcatg gttggacagt gctccaggga agaactcggc gtccagacaa
2101 cccatttcga gtggccctgg aatacatctc cagtggaaac cgcaqcttgt ctgcagtgg
2161 ctcttttgcc ctgaagaact gcagtgaagg aacatcccca ggctccaaga tggccctgca
2221 gagctccttc acttggttga atgggacagt cctccagctt gggcaggcct gtgacttcca
2281 ccaggactgt gcccagggag aagatgagag ccagatgtgc cggaaaactg ctgtgggttt
2341 ttactgcaac tttgaagatg gcttctgtgg ctggacccaa ggcacactgt caccacacac
2401 tctccaatgg caggtcagga ccctaagga tgcccgttcc caggaccacc aagaccatgc
2461 tctattgtct agtaccactg atgtcccgc ttctgaaagt gctacagtga ccagtgtac
2521 gtttctctga ccgatcaaga gctctccatg tgagctccga atgtcctggc tcattcgtgg
2581 agtcttgagg ggaacgtgt ccttggtgct agtgagaac aaaaacggga aggagcaagg
2641 caggatggtc tggcatgtcg ccgcctatga aggcttgagc ctgtggcagt ggatggtgtt
2701 gcctctctct gatgtgtctg acaggttctg gctgcagatg gtgcagtggt ggggacaagg
2761 atccagagcc atcgtggctt ttgacaatat ctccatcagc ctggactgct acctcaccat
2821 tagcggagag gacaagatcc tgcagaatac agcacccaaa tcaagaaaac tgtttgagag
2881 aaacccaaac aaggagctga aaccgggga aaattcacca agacagacc ccatctttga
2941 cctacagttt cattggctgt tcaccacatg tggggccagc gggcccatg gccccacca
3001 ggcacagtgc aacaacgctt accagaactc caacctgagc gtggaggtgg ggagcgagg
3061 cccctgaaa ggcattccga tctggaaggt gccagccacc gacacctaca gcatctcggg
3121 ctacggagct gctggcggga aaggcgggaa gaacaccatg atgcggtccc acggcgtgtc
3181 tgtgtgggc atcttcaacc tggagaagga tgacatgctg tacatcctgg ttgggcagca
3241 gggagaggac gcctgcccc gtacaaacca gttaatccag aaagtctgca ttggagagaa
3301 caatgtgata gaagaagaaa tccgtgtgaa cagaagcgtg catgagtggg caggaggcgg
3361 aggaggaggg ggtggagcca cctacgtatt taagatgaag gatggagtgc cggtgccct

```

[0033]

```

3421 gatcattgca gccggagggtg gtggcagggc ctacggggcc aagacagaca cgttccaccc
3481 agagagactg gagaataact cctcggttct agggctaaac ggcaattccg gagccgcagg
3541 tgggtggagggt ggctggaatg ataacacttc cttgctcttg gccggaaaat ctttgcaggga
3601 ggggtgccacc ggagagacatt cctgccccca ggccatgaag aagtgggggt gggagacaag
3661 aggggggttgc ggaggggggtg gaggggggtg ctctcagggt ggaggaggcg gaggatatat
3721 aggcgggcaat gcagcctcaa acaatgacct cgaaatggat ggggaagatg gggtttctct
3781 catcagtcga ctgggcatcc tgtacacccc agctttaaaa gtgatggaa gcccacgggga
3841 agtgaatatt aagcattatc taaactgcag tcaactgtgag gtagacgaat gtcacatgga
3901 cccatgaaagc cacaagggtc tctgcttctg tgaccacggg acggtgcttg ctgaggatgg
3961 cgtctcctgc attgtgtcac ccaccccgga gccacacctg ccactctcgc tgatcctctc
4021 tgtggtgacc tctgcctcgc tggccgcctt ggtcctggct ttctccggca tcatgattgt
4081 gtaccgccgg aagcaccagg agctgcaagc catgcagatg gagctgcaga gccctgagta
4141 caagctgagc aagctccgca cctcgacct catgaccgac tacaacccca actactgctt
4201 tgcgtggcaag acctcctcca tcagtgcctt gaaggagggt cgcgggaaaa acatcacctt
4261 cattcgggggt ctgggcatat gcgccttttg ggaggtgtat gaaggccagg tgtccggaat
4321 gcccaacgac ccaagccccc tgcaagtggc tgtgaagacg ctgctgaag tgtgctctga
4381 acaggacgaa ctggatttcc tcatggaagc cctgatcacc agcaaatcca accaccagaa
4441 cattgttctg tgcattgggg tgagcctgca atccctgccc cggttcatcc tgcctggagct
4501 catgcccggg ggagacctca agtccttctc ccgagagacc cgcctctcgc cgagccagcc
4561 ctctcctctg gccatgctgg accttctgca cgtggctcgg gacattgctt gtggctgtca
4621 gtatttggag gaaaaccact tcatccaccg agacattgct gccagaaact gcccttgac
4681 ctgtccaggc cctggaagag tggccaagat tggagacttc gggatggccc gagacatcta
4741 caggcgagcg tactatagaa agggaggctg tgccatgctg ccagttaaat ggatgcccc
4801 agacggactc atggaaggaa tattcaactt taaaacagac acatggtcct ttggagtgtc
4861 gctatgggaa atcttttctc ttggatatat gccatacccc agcaaaagca accaggaagt
4921 tctgaggttt gtcaccagtg gaggccggat ggacccaccc aagaactgcc ctgggcctgt
4981 ataccggata atgactcagt gctggcaaca tcagcctgaa gacaggccca actttgccat
5041 cattttggag aggattgaat actgcaccca ggacccggat gtaatcaaca ccgctttgcc
5101 gatagaatat ggtccacttg tggaagagga agagaaaagt cctgtgaggc ccaaggacc
5161 tgaggggggt tctggtctct cctcctctcc ctaccacctc ctctggcaag gctgcaaaag aacccacagc
5221 agctgccccca ccacctctgc caccacctc ctctggcaag gctgcaaaag aacccacagc
5281 tgcagagatc tctgttctgag tccctagagg gccggccgtg gaagggggac acgtgaatat
5341 ggcattctct cagtcacaac ctccttcgga gttgcacaag gtccacggat ccagaaacaa
5401 gccaccagc ttgtggaaac caacgtacgg ctccctgggtt acagagaaac ccacaaaaaa
5461 gaataatcct atagcaaaag aggagccaca cgacagggtt aacctggggc tggaggggaag
5521 ctgtactgtc ccacctaaag ttgcaactgg gagacttccg ggggcctcac tgctcctaga
5581 gccctcttgc ctgactgcca atatgaagga ggtacactctg ttcagggtac gtcaactccc
5641 ttgtgggaaat gtcaattacg gctaccagca acagggtctg cccttagaag ccgctactgc
5701 ccctggagct ggtcattacg aggataccat tctgaaaagc aagaatagca tgaaccagcc
5761 tgggcccctga gctcggctgc aactcactt ctcttctctg ggatccctaa gaccgtggag
5821 gagagagagg caatggctcc ttcacaaaac agagacaaa tgtcacggtt tgttttgtgc
5881 caacctattt tgaagtacca ccaaaaaagc tgtattttga aaatgcttta gaaagggttt
5941 gagcatgggt tcactcattt ctttcgaaag aagaaaaat cataaaaatg agtgataaat
6001 acaaggccca gatgtggttg cataagggtt ttatgcatgt ttgtgtata cttccttatg
6061 cttctttcaa attgtgtgtg ctctgcttca atgtagtac aattagctgc ttctatgttt
6121 catagttggg gtcatagatg ttctcttgcc ttgttgatgt ggacatgagc catttgaggg
6181 gagagggaac ggaataaag gagttatttg taatgactaa aa

```

야생형의 동족체에서의 시스템이 아닌 아미노산을 엔코딩하는 야생형 cDNA 서열 TGC
(4373 내지 4375) 코돈 돌연변이(들) 또는 상응하는 돌연변이

야생형의 동족체에서의 루신이 아닌 아미노산을 엔코딩하는 야생형 cDNA 서열 CTG
(4493 내지 4495) 코돈 돌연변이(들) 또는 상응하는 돌연변이

야생형의 동족체에서의 야생형 cDNA 서열 G4374A 돌연변이 또는 상응하는 돌연변이

야생형의 동족체에서의 야생형 cDNA 서열 C4493A 돌연변이 또는 상응하는 돌연변이

야생형 ALK 단백질 서열 (NP_004295.2; GI:29029632) :

```

1  mgaigllwll plllstaavg sgmgtgqrag spaagpplqp replsysrlq rkslavdfvv
61  psifrvyard lllppsssel kagrpeargs laldcapllr llgpapgvsw tagspapaea
121 rtlsrvlkgg svrkirraq lvlelgeeeai legcvppge aavglqlfnl selfswwirq
181 gegrlrirlm pekkasevgr egrlsaaia sqprllfqif gtghsslesp tnmpspspdy
241 ftwnltwimk dsfpflshrs ryglecsfdf pceleyspl hdlrnqswsw rripseeasq
301 mdlldgpgae rskemprgsf llntsadsk htllspwmrs ssehctlavs vhrhlqpsgr
361 yiaqlphne aareillmpt pgkhgwtvlq grigrpdnpf rvaleyissg nrslsavdff
421 alkcsegts pgskmalqss ftcwngtvlq lgqacdfhqd caggedesqm crklpvgyfc
481 nfedgfcgwt ggtlsphptq wqvrtlkdar fqdhqdhall lsttdvpase satvtsatfp
541 apiksspcel rmswlrirgvl rgnvslvlve nktgkeggrm vwhvaayegl slwqwmvlpl
601 ldvsdrfwlq mvawwgqgsr aivafdnisi sldcyltisg edkilqntap ksrnlfernp
661 nkelkpgens prqtpifdpt vhwlfctcga sgphgptqag cnnayqnsnl svevgsegpl
721 kgigikwvpa tdtysisgyg aaggkggknt mmrshgvsvl gifnlekddm lyilvgqqge
781 dactpntqli qkvicgennv ieeeiirvnrs vhwagggggg gggatyvfkf kdgvpvplii
841 aaggggrayg aktdtfhper lennssvlgl ngnsaaggg ggwndntsll wagkslqega
901 tghscpqam kkwgwegrgg fggggggcgs ggggggyigg naasnndpem dgedgvsfis
961 plgilytpal kmeghgevn ikhylncshc evdechmdpe shkvicfcdh gtvlaedgvs
1021 civsptpeph lplslilsvv tsalvaalvl afsgimivyr rkhqelqamq melqspeykl
1081 sklrtstimt dynpnycfag ktssisdike vprknitlir glghgafgev yegqvsgmpn
1141 dpsplqvavk tlpvcseqd eldfmeali iskfnhqniv rcigvslqsl prfillelma
1201 ggdlksflre trprpsqps lamldllhva rdiacgcqyl eenhfihrdi aarnclltcp
1261 gpgrvakigd fgmardiya syyrkggcam lpvkwmppa fmegiftskt dtwsfgvllw
1321 eifslgymy pskanqeule fvtsggrmdp pknpcpgvyr imtqcwqhqp edrpnfail
1381 erieyctqdp dvintalpie ygplveeeek vpvprkdpeg vplllvsqqa kreeerspaa
1441 ppplpttssg kaakkptaee isvrprgpa vegghvnmf sqsnppsels kvhgsrnkpt
1501 slwnptygsf ftekttknn piakkephdr gnllegsct vppnvatgrl pgasilleps
1561 sltanmkevp lfrirhfcpg nvnygyqqg lpleaatap aghyedtilk sknsnmgpgp

```

야생형의 동족체에서의 Xaa 가 시스테인이 아닌 아미노산인 야생형 단백질 서열

Cys1156Xaa 돌연변이 또는 상응하는 돌연변이

야생형의 동족체에서의 Xaa 가 류신이 아닌 아미노산인 야생형 단백질 서열

Leu1196Xaa 돌연변이 또는 상응하는 돌연변이

야생형의 동족체에서의 야생형 단백질 서열 Cys1156Tyr 돌연변이 또는 상응하는

돌연변이

야생형의 동족체에서의 야생형 단백질 서열 Leu1196Met 돌연변이 또는 상응하는

돌연변이

EML4-ALK 변이체 1 cDNA 서열 (AB274722.1; GI:152002652)

```

1  ggcggcgcg ggcggcgctc gcggtgctg cctgggaggg aggcggggca ggcggctgag
61  cggcgcggt ctaacgtga cggggaagt gttcggcgcg ccggcggtta ctaccccagg
121 gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccggcgcgacc tagagaacga
181 ggcgggtcagg ctcagcgctg gccactctgt cggtcgctg aatgaagtgc ccgccctct
241 gagcccgagg cccggcgctt tccccgcaag atggacggtt tcgccggcag tctcgatgat
301 agtatctctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcaaga
361 gttcagcaac aagaagatga aatcactgtg ctaaaaggcg ctttggtgta tgttttgagg
421 cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaata aatcagctct aagtaaaggc
481 caaccaagcc ctcgagcagt tattcccatg tctgtataa ccaatggaag tggcgcaaac
541 agaaaaccaa gtcataccag tgotgtotca attgcaggaa aagaaactct tcatctgct
601 gctaaaagtg gtacagaaaa aaagaaagaa aaaccacaag gacagagaga aaaaaaagag
661 gaatctcatt ctaatgatca aagtcacaaa attcgagcat caccttctcc ccagccctct
721 tcacaacctc tccaaataca cagacaaact ccagaaagca agaattgtac tcccacaaa

```

[0035]

781 agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcagatgat
841 agccgttaata aatttgtgaa aataccttca acacccaaat taataccaaa agttaccaaa
901 actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt
961 atgcgcggtc ggccaattac catgtttcatt ccttcgatg ttgacaacta tgatgacatc
1021 agaacggaac tgccctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaaag
1081 gactgtagag ctaatgttta ccttcttcg accgggaaaa tagtttattt cattgcatca
1141 gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg ccatcacagac
1201 tgtgtgaaat gccttgcctat acatcctgac aaaattagga ttgcaactgg acagatagct
1261 ggcgtggata aagatggaag gcctctacaa cccacgtca gagtgtggga ttctgttact
1321 ctatccacac tgcagattat tggacttggc acttttgagc gtggagttagg atgcctggat
1381 ttttcaaaaag cagattcagg tgttcattta tgtgttattg atgactccaa tgagcatatg
1441 cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa
1501 gttgttttgg ctgtggagtt tcaccaaca gatgcaata ccataattac atgcggtaaa
1561 tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaa gggaaatttt
1621 gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttgggaa tggagatgtt
1681 cttactggag actcaggttg agtcatgctt atatggagca aaactactgt agagcccaca
1741 cctgggaaaag gacctaaagt gtaccgccg aagcaccagg agctgcaagc catgcagatg
1801 gagctgcaga gccctgagta caagctgagc aagctccgca cctcgaccat catgaccgac
1861 tacaaccoca actactgctt tgctggcaag acctcctcca tcagtgaact gaaggagtg
1921 ccgcgaaaaa acatcacccct cattcggggt ctgggccatg gagccttgg ggaggtgtat
1981 gaagggcagg tgtccggaat gcccacgac ccaagcccc tgcaagtggc tgtgaagacg
2041 ctgcctgaag tgtgctctga acaggacgaa ctggatttcc tcatggaagc cctgatcatc
2101 agcaaatcca accaccagaa cattgttcgc tgcattggg tgagcctgca atccctgcc
2161 cgggtcatcc tgctggagct catggcggg ggagacctca agtccctcct ccgagagacc
2221 cgccctcgcc cgagccagcc ctccctcctg gccatgctgg acctctgca cgtggctcgg
2281 gacattgcct gtggctgtca gtatttggag gaaaaccact tcattccacc agacattgct
2341 gccagaaaat gcctcctgac ctgtccaggc cctggaagag tggccaagt tggagaactc
2401 gggatggccc gagacatcta cagggcgagc tactatagaa agggaggctg tgccatgctg
2461 ccagttaagt ggtgcccc agaggccttc atggaaggaa tattcacttc taaaaacagac
2521 acatggctct ttggagtgtc gctatgggaa atcttttctc ttggatatat gccatacccc
2581 agcaaaagca accaggaagt tctggagtgt gtcaccagtg gaggcggat ggacccaccc
2641 aagaactgcc ctgggcctgt ataccggata atgactcagt gctggcaaca tcagcctgaa
2701 gacaggccca actttgccat cattttggag aggattgaat actgcacca ggacccggat
2761 gtaatacaac ccgctttgcc gatagaatat ggtccacttg tggaaagagga agagaaagtg
2821 cctgtgaggc ccaaggaccc tgagggggtt cctcctctcc tggctctcca acaggcaaaa
2881 cgggaggagg agcgcagccc agctgcccc ccacctctgc ctaccacctc ctctggcaag
2941 gctgcaaaag aaccacagc tgcagaggtc tctgttcgag tccctagagg gccggcgctg
3001 gaagggggac acgtgaatat ggcattctct cagtccaacc ctctctcgga gttgcacagg
3061 gtccacggat ccagaaacaa gcccaccagc ttgtggaacc caacgtacgg ctctgtgtt
3121 acagagaaac ccacaaaaa gaataatcct atagcaaaag aggagccaca cgagaggggt
3181 aacctggggc tggagggaag ctgtactgtc ccacctaacg ttgcaactgg gagacttccg
3241 ggggcctcac tgcctctaga gccctcttcg ctgactgcca atatgaagga ggtacctctg
3301 ttcaggctac gtcacttccc ttgtgggaat gtcaattacg gctaccagca acagggcttg
3361 cccttagaag ccgctactgc cctggagct ggtcattacg aggataccat tctgaaaagc
3421 aagaatagca tgaaccagcc tgggccctga gctcggtcac acactcactt ctcttcttg
3481 ggatccctaa gaccgtggag gagagagagg caatcaatgg ctccctcaca aaccagagac
3541 caaatgtcac gttttgtttt gtgccaacct attttgaagt accacaaaaa aagctgtatt
3601 ttgaaaatgc tttagaaagg ttttgagcat gggttcatcc tattcttctg aaagaagaaa
3661 atatcataaa aatgagtgat aaatacaagg ccagatgtg gttgcataag gtttttatgc
3721 atgtttgttg tatacttctt tatgcttctt ttaaatgtg tgtgctctgc ttcattgtag
3781 tcagaattag ctgcttctat gtttcatagt tggggtcata gatgttctct tgcttgttg
3841 atgtggacat gagccatttg aggggagagg gaacggaaat aaaggagtta tttgtaattg
3901 aaaaaaaaaa aaaaaaaaaa aaaaaa

EML4-ALK 변이체 1 단백질 서열 (BAF73611.1; GI:152002653)

1 mdgfagsldd sisaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva
61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
121 kpggqrekke eshndqspq iraspspps sqplqihrt pesknaptk sikrpspaek
181 shnswensdd srnklskipk tpklipkvtk tadkhkdvii ngegeyikmf mrgrpitmfi
241 psdvdnyddi rtelppekkl lewaygyrgk dcranvylp tgeivyfias vvvlfnyeer

[0036]

301 tqrhyllghtd cvkclaihpdi kiriatggia gvdkdgrplq phvrwvdsvt lstlqiiglg
 361 tfergvgcld fskadsgvhl cviddsnehm ltvwdwqkka kgaeikttne vvlavefhpt
 421 dantiitcgk shiffwtwsg nsitrkqgif gkyekpkfvq clafingndv ltgdsdggvml
 481 iwsktttvept pgkpgkpyrr khqelqamqm elqspeykls klrtsttimtd ynpnycfagk
 541 tssisdikev prknitlirg lghgafgevy egqvsgmpnd psplqvavkt lpevcseqde
 601 ldflmealii skfnhqniwr cigvslqslp rfillelmag gdlksflret rprpsqpsl
 661 amldllhvar diacgcqyle enhfihrdia arnclltcpq pgrvakigdf gmardiyas
 721 yyrkggcaml pvkwmppeaf megiftsktd twsfgvllwe ifslgympyp sksnqevlef
 781 vtsggrmdpp knpcgpyyri mtgcwqhpe drpnfaiile rleyctqdpd vintalpiey
 841 gplveeeekv pvrpkdpegv ppllvsgqak reeerspaap pplpttssgk aakkptaaev
 901 svrvprgpav egghvnmafs qsnppselhr vhgrrnkpts lwnptygswf tekptkknnp
 961 iakkepherg nllegscctv ppnvatgrlp gaslllepss ltanmkevpl frlrhfpccgn
 1021 vnygyqqqgl pleaatapga ghyedtilks knsmnqpgp

EML4-ALK 변이체 2 cDNA 서열 (AB275889.1; GI:152002654)

1 ggcggcgccg cgcggcgctc gcggctgctg cctgggaggg aggcggggca ggcggctgag
 61 cggcgcgccg ctcaacgtga cggggaagtg gttcggggcg cgcggctta ctaccccgag
 121 gcgaacggac ggacgacgga ggcgggagcc ggtagccgag cgcggcgacc tagagaacga
 181 ggcgggctcagg ctacgagctcg gccactctgt cgggtccgctg aatgaagtgc ccgcccctct
 241 gagcccgag cccggcgctt tcccgcgaag atggacgggt tcgcggcgag tctcgatgat
 301 agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtacaga
 361 gttcagcaac aagaagatga aatcactgtg cttaaggcgg ctttggtgta ggttttgag
 421 cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaata aatcagctct aagtaaggc
 481 caaccaagcc ctogagcagt tattcccatg tctgtataa ccaatggaag tgggtcaaac
 541 agaaaaacaa gtcataccag tgctgtctca attgcaggaa aagaaactct ttcatctgct
 601 gctaaaagtg gtacagaaa aaagaaaaga aaaccacaag gacagagaga aaaaaagag
 661 gaattctcatt ctaatgatca aagtcacaa attcgagcat caccttctcc ccagccctct
 721 tcacaacctc tccaaataca cagacaaact ccagaaagca agaattgtac tcccaccaa
 781 agcataaaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcatgatgat
 841 agccgtaata aattgtcgaa aataccttca acacccaaat taataccaaa agttaccaa
 901 actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt
 961 atgcgcggtc ggccaattac catgttcatt ccttcgatg ttgacaacta tgatgacatc
 1021 agaacgggaa tgccctctga gaagctcaaa ctggagtggt catatggtta tcgaggaaag
 1081 gactgttagag ctaattgtta ccttcttccg accggggaaa tagtttattt cattgcatca
 1141 gtatgtatgac tatttaatta tgaggagaga actcagcgac actacctggg ccatacagac
 1201 tgtgtgaaat cctctgctat acatcctgac aaaattagga ttgcaactgg acagatagct
 1261 ggcgtggata aagatggaag gcctctacaa ccccacgtca gagtgtggga ttctgttact
 1321 ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat
 1381 ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactccaa tgagcatatg
 1441 cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa
 1501 gttgttttgg ctgtggagtt tcacccaaca gatgcaata ccataattac atgcggtaaa
 1561 tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaa gggaaatttt
 1621 gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa tggagatgtt
 1681 cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagccaca
 1741 cctgggaaag gacctaaagg tgtatatcaa atcagcaaac aaatcaaggc tcatgatggc
 1801 agtgtgttca cactttgtca gatgagaaat gggatgttat taactggagg agggaaagac
 1861 agaaaaataa ttctgtggga tcatgatctg aatcctgaaa gagaaataga ggttctgat
 1921 cagtagtgca caatcagagc ttagcagaaa ggaaggcag atcaattttt agtaggcaca
 1981 tcacgaaact ttattttacg aggaacattt aatgatggct tccaaataga agtacagggt
 2041 catcacagat agctttgggg tcttgccaca catcccttca aagatttgc tttgacatgt
 2101 gctcaggaca ggcaggtgtg cctgtggaac tcaatggaac acaggctgga atggaccagg
 2161 ctggtagatg aaccaggaca ctgtgcagat tttcatccaa gtggcacagt ggtggccata
 2221 ggaacgcact caggcaggtg gtttgttctg gatgcagaaa ccagagatct agtttctatc
 2281 cacacagacg ggaatgaaca gctctctgtg atgcgctaact caatagatgg taccttctg
 2341 gctgtaggat ctcatgacaa ctttattttac ctctatgtag tctctgaaaa tggagaagaa
 2401 tatagcagat atggaagggt cactggacat tccagctaca tcacacacct tgactggtcc
 2461 ccagacaaca agtatataat gtctaactcg ggagactatg aaatattgta cttgtaccgc
 2521 cgggaagcacc aggaagctga agccatgcag atggagctgc agagccctga gtacaagctg
 2581 agcaagctcc gcacctcgac catcatgacc gactacaacc ccaactactg ctttgctggc
 2641 aagacctcct ccatcagtga cctgaaggag gtgcccgga aaaacatcac cctcattcgg

[0037]

2701 ggtctggggcc atggagcctt tggggaggtg tatgaaggcc aggtgtccgg aatgcccac
 2761 gaccccaagcc ccttgcaagt ggctgtgaag acgctgcctg aagtgtgctc tgaacaggac
 2821 gaactggatt tctcatgga agccctgac atcagcaaat tcaaccacca gaacattgtt
 2881 cgctgcattg gggtagcctt gcaatccctg ccccggttca tctgctgga gctcatggcg
 2941 gggggagacc tcaagtcctt cctcogagag acccgccctc gcccgagcca gccctcctcc
 3001 ctggccatgc tggaccttct gcacgtggct cgggacattg cctgtggctg tcagtatttg
 3061 gaggaaaacc acttcatcca ccgagacatt gctgccagaa actgcctctt gacctgtcca
 3121 ggccctggaa gaggggccaa gattggagac ttccggatgg cccgagacat ctacagggcg
 3181 agctactata gaaagggagg ctgtgccatg ctgccagtta agtggatgcc cccagaggcc
 3241 ttcattggaag gaattatcac ttctaaaaca gacacatggt cctttggagt gctgctatgg
 3301 gaaatctttt ctcttgata tatgccata cccagcaaaa gcaaccagga agtcttgagg
 3361 ttgttcacca gtggaggccg gatggaccca cccaagaact gccctggggc tgtataccgg
 3421 ataagtactc agtctggga acatcagcct gaagacaggc ccaactttgc catcattttg
 3481 gagaggattg aatctgcac ccaggaccgg gatgtaatca acaccgcttt gccgatagaa
 3541 tatgttccac ttgtggaaga ggaagagaaa gtgctgtga ggcccaagga ccttgagggg
 3601 gttcctctc tctcgtctc tcaacaggca aaacgggagg agggagcgag cccagctgcc
 3661 ccaccacctc tgctaccac ctctctggc aaggctgcaa agaaaccac agctgcagag
 3721 gtctctgttc gactcctag agggccggcc gtggaagggg gacacgtgaa tatggcattc
 3781 tctcagtcga accctccttc ggagttgcac aggtccacg gatccagaaa caagcccacc
 3841 agcttctgga acccaacgta cggctcctgg ttacagaga aaccaccaa aaagaataat
 3901 cctatagcaa agaaggagcc acacgagagg ggtaacctgg ggctggaggg aagctgtact
 3961 gtccaccta acgttgcaac tgggagactt ccgggggccc cactgctcct agagccctct
 4021 tcgctgactg ccaatatgaa ggaggtacct ctgttcaggc tacgtcaact ccttctggg
 4081 aatgtcaatt acggctacca gcaacagggc ttgcccttag aagccgctac tgcccctgga
 4141 gctggtcatt acgaggatac cattctgaaa agcaagaata gcatgaacca gccctggccc
 4201 tgagctcggc cacacactca cttctcttc ttgggatccc taagaccgtg gaggagagag
 4261 aggcaatcaa tggctccttc aaaaaccaga gacaaatgt cacgtttgtg ttgtgcca
 4321 cctattttga agtaccacca aaaaagctgt atttgaaaa tgctttagaa aggttttgag
 4381 catgggttca tctattctt tcgaaagaag aaaatatcat aaaaatgagt gataaatata
 4441 agggccagat gtggttgcac aaggttttta tgcatgtttg ttgtatactt ccttatgctt
 4501 ctttttaaat gtgtgtgctc tgcttcaatg tagtcagaat tagctgcttc tatgtttcat
 4561 agttggggtc atagatgttt ccttgccctg ttgatgtgga catgagccat ttgaggggag
 4621 agggaaacgga aataaaggag ttatttgtta tgaaaaaaa aaaaaaaaa aaaaaaaaa

EML4-ALK 변이체 2 단백질 서열 (BAF73612.1; GI:152002655)

1 mdgfagsldd sisaastsdv qdrlesalesr vqqgedeitv lkaaladvlr rlaisedhva
 61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
 121 kpqggqrekke eshsndqspq iraspspqp sqplqihrt pesknaptk sikrpspaek
 181 shnswensdd srnklskips tpklipkvtk tadkhkdvii ngegeyikmf mrgrpitmfi
 241 psdvndyddi rtelppeklk lewaygyrgk dcranvllp tgeivyfiav vvvlfnyeer
 301 tqrhylvghd cvkclaihp d kiriatgqia gvdkdgrplq phvrwvdsvt lstlqiiglg
 361 tfergvvcld fskadsgvhl cviddsnehm ltvdwqkka kgaeiktne vvlavefhpt
 421 dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq clafngndv ltgdsggvml
 481 iwsktttvept pgkpgkgyvq iskqikahdg svftlcqmrn gmltgggkd rkiilwdhdl
 541 npereievdp qygtiravae gkadqflvgt srnfilrgtf ndgffilevgg htdelwglat
 601 hpfkdlilltc aqdrqvcclwn smehrlewtr lvdepghead fhpsgtvva gthsgwfvf
 661 daetrdlvs i htdgneqlsv mrysidgtfl avgshdnfiy lyvvsengrk ysrygrctgh
 721 ssyithldws pdnkyimsns gdeilylyr rkghelqamq melqspaykl sklrtstint
 781 dynpnycfag ktssisdike vprknitlir glghgafgev yegqvsgmpn dpsplqvavk
 841 tlpevcseqd eldfilmeal i skfnhqniv rcigvslqsl prfilllelma ggdlsfllre
 901 trprpsqpss lamldllhva rdiacgcqyl eenhfihrdi aarnclltcp gpgrvakigd
 961 fgmardiya syyrkkgcam lpvkmppea fmegiftskt dtwsfgvllw eifslgympy
 1021 psksnqevle fvtsggrmdp pknccpgpvyr imtqcwqhqp edrpnfaiil erieyctqdp
 1081 dvintalpie ygplveeek vpvrpkdpeg vppllvsgqa kreeerspaa ppplpttssg
 1141 kaakkptaee vsrvrprgpa vegghvnmf sqsnppselh rvhgssrnkpt slwnptygsw
 1201 ftektptkkn piakkepher gnlglegsct vppnvatgrl pgaslllleps sltanmkevp
 1261 lfrlrhfpcc nvnygyqqqg lpleaatapg aghyedtilk sksnmngpgp

EML4-ALK 변이체 3a 핵산 서열 (AB374361.1; GI:194072592)

[0038]

1 actctgtcgg tccgctgaat gaagtgccg cccctctaag cccggagccc ggcgctttcc
61 ccgcaagatg gacgggtttcg ccgacagtct cgatgatagt atttctgtg caagtacttc
121 tgatgttcaa gatcgctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat
181 cactgtgcta aaggcggtt tggctgatgt tttagggcgt ctgcaatct ctgaagatca
241 tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gaggcagttat
301 tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc
361 tgtctcaatt cgaggaaaag aaactcttct atctgctgct aaaagtggta cagaaaaaaa
421 gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaaag
481 tccacaaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag
541 acaaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc
601 tgaaaagtca cataattctt gggaaaattc agatgatagc cgtataaat tgtcgaaaaat
661 accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt
721 catcatcaac caagtgtacc gccggaagca ccaggagctg caagccatgc agatggagct
781 gcagagccct gagtacaagc tgagcaagct ccgcacctcg accatcatga ccgactacaa
841 ccccaactac tgctttgtgt gcaagacctc ctccatcagt gacctgaagg aggtgccgcg
901 gaaaaacatc accctcattc ggggtctggg ccattggagcc ttgggggagg tgtatgaagg
961 ccagggtgtcc ggaatgccc aagacccaag cccctgcaa gtggctgtga agacgctgcc
1021 tgaagtgtgc totgaacagg acgaactgga tttctctatg gaagccctga tcatcagcaa
1081 attcaaccac cagaacattg ttcgctgcat tgggggtgagc ctgcaatccc tgccccggtt
1141 catcctgtcg gagctcatgg cggggggaga cctcaagtcc ttctccgag agacccgccc
1201 tcgcccagagc cagccctctc ccttggccat gctggacctt ctgcacgtgg ctcgggacat
1261 tgctgtggc tgtcagattt tggaggaaaa ccacttcatc caccgagaca ttgctgccag
1321 aaactgcctc ttgacctgtc caggccctgg aagagtggcc aagattggag acttcgggat
1381 ggcctcgagc atctacagg cgagctacta tagaaaggga ggctgtgcca tgcgtccagt
1441 taagtggatg ccccagagg ccttcatgga aggaatatc acttctaaaa cagacacatg
1501 gtcctttgga gtgctgctat gggaaatctt ttctcttga tatatgccat accccagcaa
1561 aagcaaccag gaagtcttgg agtttgcac cagtggagcg cggatggacc caccacaaga
1621 ctgcccgtgg cctgtatacc ggataatgac tcagtgtggt caacatcagc ctgaagacag
1681 gcccaacttt gccatcattt tggagaggat tgaatactgc acccaggacc cggatgtaat
1741 caacacgctg ttgcccagat aatatggctc acttgtggaa gaggaagaga aagtgcctgt
1801 gagggcccaag gacctgagg gggttctctc tctctgtgtc tctcaacagg caaaacggga
1861 ggaggagcgc agccagctg cccaccacc tctgctacc acctcctctg gcaaggctgc
1921 aaagaaaccc acagctgcag aggtctctgt tcgagtcctc agaggggccg ccgtgggaagg
1981 ggagacgctg aatatggcat tctctcagtc caacctctc tggagttgc acagggtcca
2041 cggatccaga aacaagccca ccagcttgtg gaacccaacg tacggctcct ggtttacaga
2101 gaaacccacc aaaaagaata atcctatagc aaagaaggag ccacacgaga ggggtaacct
2161 ggggctggag ggaagctgta ctgtccacc taacgttgca actgggagac ttccgggggc
2221 ctcaactgctc ctagagccct cttcgctgac tgccaatatg aaggagggtac ctctgttcag
2281 gctacgtcac ttcccttgtg ggaatgtcaa ttacggctac cagcaacagg gcttgcctt
2341 agaagccgct actgcccctg gagctgggtc ttacaggatg accattctga aaagcaagaa
2401 tagcatgaac cagcctgggc cctgagctcg gtcgcacact cacttctctt ccttgggac
2461 cctaagaccg tgg

EML4-ALK 변이체 3a 단백질 서열 (BAG55003.1; GI:194072593)

1 mdgfagsldd sisaastsdv qdrisalesr vqqgedeitv lkaaladvlr rlaisedhva
61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
121 kpqggqrekke eshsndqspq iraspspqs sqplqihrt pesknaptk sikrpspaek
181 shnswensdd srnklskips tpklipkvtk tadkhkdvii ngvyrrkhqe lqamqmelqs
241 peyklsklirt stimtdynpn ycfagktssi sdlkevprkn itlirglghg afgevyegqv
301 sgmpndpspl qvavktlpev cseqdeldfl mealiiskfn hqnivrcigv slqslprfil
361 lelmaggdlk sflretrprp sqpsslamld llhvardiac gcqyleenhf ihrdiaarnc
421 lltcpgpgrv akigdfgmar diyrasyrk ggcamlpvkw mppeafmegi ftsktdtwsf
481 gvllweifsl gypypsksn gevlefvtsg grmdppkncp gpvyrimtgc wqhqpdrpn
541 failerley ctqdpdvint alpieygplv eeeekvpvrp kdpegvppll vsqakreee
601 rspaappplp ttssgkaakk ptaaevsrv prgpaveggh vnmafsqsn pslhrvhgs
661 rnkptslwnp tygswftekp tkknnpiakk ephergnlgl egscvppnv atgrlpgasl
721 llepssltn mkevpflrlr hfpcgnvnyg yqqqglplea atapgaghye dtilkksnsm
781 nqpgp

[0039]

EML4-ALK 변이체 3b 핵산 서열 (AB374362.1; GI:194072594)

1 actctgtctcg tccgtctgaat gaagtgcctg cccctctaag cccgagagcc ggcgctttcc
61 ccgcaagatg gacggtttcg ccggcagctc cgatgatagt atttctgtcg caagtacttc
121 tgatgttcaa gatgcctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat
181 cactgtgtcta aaggcggctt tggctgatgt tttgaggcgt cttgcaatct ctgaagatca
241 tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat
301 tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc
361 tgtctcaatt gcaggaaaaa aaactcttcc atctgtctgt aaaagtggta cagaaaaaaa
421 gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag
481 tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag
541 acaaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc
601 tgaaaagtca cataattctt gggaaaattc agatgatagc cgtataaat tgtcgaaaat
661 acottcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt
721 catcatcaac caagcaaaaa tgtcaactcg cgaaaaaac agccaagtgt accgccggaa
781 gcaccaggag ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa
841 gctccgcacc tcgaccatca tgaccgacta caaccccaac tactgtcttg ctggcaagac
901 ctctctccatc agtgacctga aggaggtgcc gcggaaaaac atcacctcca ttcggggtct
961 ggcccatgga gcctttgggg aggtgtatga aggccaggtg tccggaatgc ccaacgaccc
1021 aagcccccctg caagtggctg tgaagacgct gcctgaagtg tgctctgaac aggacgaact
1081 ggatttctctc atggaagccc tgatcatcag caaattcaac caccagaaca ttgttcgctg
1141 cattgggggtg agcctgcaat ccctgccccg gttcatcctg ctggagctca tggcgggggg
1201 agacctcaag tcttctctcc gagagaccgg ccctgcgccg agccagccct cctccctggc
1261 catgctggac cttctgcacg tggctcgga cattgcctgt ggctgtcagt atttgagga
1321 aaaccacttc atccaccgag acattgctgc cagaaactgc ctcttgacct gtccaggccc
1381 tgggaagagt gccaagattg gagacttcgg gatggcccg gacatctaca gggcgagcta
1441 ctatagaaag ggaggctgtg ccatgtgcc agtlaagtgg atgccccag aggccttcat
1501 ggaaggaata ttcatctcta aaacagacac atggctcttt ggagtgtctg tatgggaaat
1561 cttttctctt ggatatatgc cataccocag caaaagcaac caggaaagttc tggagtttgt
1621 caccagtggg ggcgggatgg acccacccaa gaactgccct gggcctgtat accggataat
1681 gactcagtg cggcaacatc agcctgaaga caggcccaac ttgcatca ttttgagag
1741 gattgaatac tgcaccagg acccgatgt aatcaacacc gctttgccga tagaatatgg
1801 tccacttgtg gaagaggaag agaaagtgcc tgtgaggccc aaggaccctg agggggttcc
1861 tctctctctg gtctctcaac aggcacaaac ggaggaggag cgcagccag ctccccacc
1921 acctctgct accacctctc ctggcaaggc tgcaaaagaa cccacagctg cagaggtctc
1981 tgttcagagc cctagagggc cggccgtgga agggggacac gtgaatatgg cattctctca
2041 gtccaaacct ccttcggagt tgcacagggt ccacggatcc agaaacaagc ccaccagctt
2101 gtggaaacca acgtacggct cctggtttat agagaaaacc accaaaaaga ataactctat
2161 agcaaaagaag gagccacacg agagggttaa cctggggctg gagggaagct gtactgtccc
2221 acctaacgtt gcaactggga gaactccggg ggctcactg ctcctagagc cctcttcgct
2281 gactgccaat atgaaggagg tacctctgtt caggctacgt cacttccctt gtgggaatgt
2341 caattacggc taccagcaac agggcttgcc cttagaagcc gctactgcc ctggagctgg
2401 tcattacgag gataccatc tgaaaagcaa gaatagcatg aaccagcctg ggcctgagc
2461 tcggtcgcac actcacttct ctctctggg atccctaaga ccgtgg

EML4-ALK 변이체 3b 단백질 서열 (BAG55004.1; GI:194072595)

1 mdgfagsldd sisaastsdv qdralsalesr vqqqedeitv lkaaladvlr rlaisedhva
61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
121 kpqggqrekke eshsndqspq iraspspqs sgplqihrt pesknaptk sikrpspaek
181 shnswensdd srnklskips tpklipkvtk tadkhkdvii nqakmstrek nsqvyrkqh
241 elqamqmelq speyklslr tstimtdynp nycfagktss isdlkevprk nitlirglgh
301 gafgevyegq vsgmpndpsp lqvavktlpe vcseqdelf lmealiiskf nhqnivrcig
361 vslqslprfi llelmaggdl ksflretrpr psqpsslaml dlhvardia cgcqyleenh
421 fihrdiaarn cltctpggr vakigdfgma rdiyrasyr kggcamlpvk wmppeafmeg
481 iftsktdtws fgvlweifs lgympypsk ngevlvfts ggrmdppknc pgpvyrmtq
541 cwghqpedrp nfaiilerie yctqdpdvin talpieygpl veeekvpvr pkdpegvppl
601 lvsqqakree erspaapppl pttssgkaak kptaaevsvr vprgpavegg hvnmafsqsn
661 ppseihrvhg srnkptslwn tygswftek ptkknpia kphergnlq legscvppn
721 vatgrlpgas lllepsslta nmkevpflrl rhfpcgnvny gyqqqglple aatapgaghy
781 edtilkskns mnqggp

[0040]

EML4-ALK 변이체 4 핵산 서열 (AB374363.1; GI:209837703)

```

1  actctgtcgg tccgctgaat gaagtgccg cccctetaag cccggagccc ggcgctttcc
61  ccgcaagatg gacgggttcc cggcagttct cgatgatatg atttctgctg caagtacttc
121 tgatgttcaa gatcgctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat
181 cactgtgcta aaggcggtct tggtgatgt tttgaggcgt ctgcaatct ctgaagatca
241 tgtggcctca gtgaaaaaat cagttctcaag taaaggccaa ccaagccctc gagcagttat
301 tcccattgct tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc
361 tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa
421 gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaaag
481 tccacaaatt cgagcatcac cttctcccca gccctottca caacctctcc aaatacacag
541 acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc
601 tgaaaagtca cataattctt gggaaaattc agatgatagc cgtataaat tgtcgaaaaat
661 accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt
721 catcatcaac caagaaggag aatatattaa aatgtttatg cgcggtcggc caattacat
781 gttcattcct tccgatgttg acaactatga tgacatcaga acggaactgc ctctgagaa
841 gctcaaaactg gagtgggcat atgggtatcg aggaaggac ttagagcta atgtttacct
901 tcttccgacc ggggaatatg tttatttcat tgcacagta gtagtactat ttaattatga
961 ggagagaact cagcgacact acctgggcca tacagactgt gtgaaatgcc ttgctataca
1021 tcttgacaaa attaggattg caactggaca gatagctggc gtggataaag atggaaggcc
1081 tctacaaccc cactcgagag tgtgggattc tgttactcta tccacactgc agattatttg
1141 acttggcact tttgagcgtg gtagtaggat cctggatttt tcaaaagcag attcaggtgt
1201 tcatttatgt gttattgatg actccaatga gcatatgctt actgtatggg actggcagag
1261 gaaagcaaaa ggagcagaaa taaagacaac aaatgaagtt gttttggctg tggagtttca
1321 cccaacagat gcaaatacca taattacatg cggtaaatct catattttct tctggacctg
1381 gagcggcaat tcactaacaa gaaaacaggg aatttttggg aaatatgaaa agccaaaatt
1441 tgtgcagtgt ttagcattct tggggaatgg agatgttctt actggagact caggtggagt
1501 catgcttata tggagcaaaa ctactgtaga gccacacact gggaaaggac ctaaagggtg
1561 atatcaaatc agcaaacaaa tcaaaagctc tgaatggcag gtgttcacac tttgtcagat
1621 gagaaatggg atgttattaa ctggaggagg gaaagacaga aaaataattc tgtgggatca
1681 tgatctgaat cctgaaagag aaatagagat atgctggatg agccctgagt acaagctgag
1741 caagctccgc acctcgacca tcatgaccga ctacaacccc aactactgct ttgctggcaa
1801 gacctcctcc atcagtgaac tgaaggaggt gccgcggaag aacatcaccc tcattcgggg
1861 tctgggcat ggagcctttg gggaggtgta tgaaggccag gtgtccggaa tgcccaacga
1921 cccaagcccc ctgcaagtgg ctgtgaagac gctgcctgaa gtgtgctctg aacaggacga
1981 actggatttc ctcatggaag cctgatcat cagcaaattc aaccaccaga acattgttcg
2041 ctgcatggg gtgagcctgc aatccctgcc cgggttcac cctgctggagc tcatggcggg
2101 gggagacctc aagtccttcc tccgagagac ccgccctcgc ccgagccagc cctcctccct
2161 ggccatgctg gaccttctgc acgtggctcg ggacattgcc tgtggtgtgc agtatttggg
2221 ggaaccacac ttcatccacc gagacattgc tgccagaaac tgcccttga cctgtccagg
2281 cctggaaga gtggccaaga ttggagactt cgggatggcc cgagacatct acagggcgag
2341 ctactataga aaggagggct gtgccatgct gccagttaa tggatgcccc cagaggcctt
2401 catggaagga atattcactt ctaaaacaga cacatggtcc tttggagtgc tgctatggga
2461 aatcttttct cttggatata tgccataccc cagcaaaagc aaccaagaag tcttgaggtt
2521 tgtcaccagt ggaggccgga tggaccacc caagaactgc cctgggcctg tataccggat
2581 aatgactcag tgtggcaac atcagcctga agacaggccc aactttgcca tcatttttga
2641 gaggattgaa tactgcacc aggaccgga tgaatcaac accgctttgc cgatagaata
2701 tgggtccactt gtggaagagg aagagaaagt gctgtgagg cccaaggacc ctgagggggt
2761 tctctctctc ctggtctctc aacaggcaaa acgggaggag gagcgagcc cagctgcccc
2821 accacctctg cctaccacct cctctggcaa ggctgcaaag aaaccacag ctgcagaggt
2881 ctctgttcga gtccttagag ggccggccgt ggaaggggga cactgaata tggcattctc
2941 tcaagccaac cctccttcgg agttgcacag ggtccacgga tccagaaaca agcccaccag
3001 cttgtggaac ccaacgtacg gctcctggtt tacagagaaa cccacaaaaa agaataatcc
3061 tatagcaaaag aaggagccac acgagagggt taacctgggg ctggagggaa gctgtactgt
3121 cccacctaac gttgcaactg ggagacttcc gggggcctca ctgctcctag agccctcttc
3181 gctgactgcc aatatgaagg aggtacctct gttcaggcta cgtcacttcc cttgtgggaa
3241 tgtcaattac ggctaccagc aacagggctt gcccttagaa gccgctactg cccctggagc
3301 tggtcattac gaggatacca ttctgaaaag caagaatagc atgaaccagc ctgggcccctg
3361 agctcggctg cacactcact tctcttctt ggatcccta agacogtgg

```

[0041]

EML4-ALK 변이체 4 단백질 서열 (BAG75147.1; GI:209837704)

1 mdgfagsladd ssaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva
61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
121 kpggqrekke eshsndgspq iraspspqs sqplqihrt pesknaptk sikrpspaek
181 shnswnsdd srnklskips tpklikvktk tadkhkdvii ngegeyikmf mrgrpitmfi
241 psdvdnyddi rtelpppekkl lewaygyrgk dcranyvllp tgeivyfias vvvlfnyeer
301 tqrhylghtd cvkclaihp d kiriatgqia gvdkgdrplq phvrwdsvt lsltqiiglg
361 tfergvgcld fskadsgvhl cviddsnehm ltvwdwqrka kgaeiktne vvlavefhpt
421 dantiitcgk shiffwtwsg nsltrkqgif gkyekpkfvq clafllngdv ltgdsggvml
481 iwskttept pgkpgkgyvq iskqikahdg svftlcqmrn gmltgggkd rkiilwdhdl
541 npereieicw mspeyklsl rtstimtdyn pnycfagkts sisdlkevpr knitlirglg
601 hgafgevyeg qvsgmpndps plqvavktlp evcseqdeld flmealisk fnhqnivrci
661 gvsllqslprf illelmagdp lksflretrp rpsqsslam ldllhvardi acgcqyleen
721 hfihrdiaar nclltcpgpg rvakigdfgm ardiyrasy rkggcamlpv kwmppeafme
781 giftsktdtw sfgvllweif slgypypsk sngevelevt sggrmdppkn cpgpvyrmt
841 qcwqhgpdr pnfaiileri eyctqdpdvi ntalpieygp lveeeekvpv rpkdpegvpp
901 llvsqqakre eerspaapp lpttssgkaa kkptaeevsv rvprgpaveg ghvnmfssqs
961 nppselhrvh gsrnkptslw nptygswfte kptkknnpia kkephergnl glegscvpp
1021 nvatgrlpga slillepsit anmkevplfr lrhfpognvn ygyqqgglpl eaatapgagh
1081 yedtilkskn smnqpgp

EML4-ALK 변이체 5a 핵산 서열 (AB374364.1; GI:209837705)

1 actctgtcgg tccgctgaat gaagtcccg cccctctaag cccggagccc ggcgctttcc
61 ccgcaagatg gacgggttcc cggcagttc ccatgatagt atttctgctg caagtacttc
121 tgatgttcaa gatcgctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat
181 cactgtgcta aagcggtt tggctgatgt tttagagcgt cttgcaatct ctgaagatca
241 tgtggcctca gtgaaaaaat cagctctcag taaagtgtac cgccggaagc accaggagct
301 gcaagccatg cagatggagc tgcagagccc tgagtacaag ctgagcaagc tccgcacctc
361 gaccatcatg accgactaca accccaacta ctgctttgct ggcaagacct cctccatcag
421 tgacctaag gaggtgccc ggaaaaacat caccctcatt cggggtctgg gccatggagc
481 ctttggggag gtgtatgaag gccaggtgtc cggaatgccc aacgacccaa gcccctgca
541 agtggctgtg aagacgtgtc ctgaagtgtg ctctgaacag gacgaactgg atttctcat
601 ggaagccctg atcatcagca aattcaacca ccagaacatt gtctcgtgca ttgggggtgag
661 cctgcaatcc ctgcccgtg tcatcctgct ggagctcatg gcggggggag acctcaagtc
721 ctctctccga gagacccgcc ctgcccagc ccagccctcc tccctggcca tgctggacct
781 tctgcacgtg gctcgggaca ttgctgtgtg ctgtcagtat ttggaggaaa accacttcat
841 caaccagagc attgctgcca gaaactgcct cttgaactgt ccaggccctg gaagagtggc
901 caagattgga gacttcggga tggcccagga catctacagg gcgagctact atagaaaggg
961 aggtgtgtcc atgtgccc ttaagtggat gcccacagag gccttcatgg aaggaatatt
1021 cacttctaaa acagacacat ggtcctttgg agtgcgtgta tgggaaatct tttctcttgg
1081 atatatgcca taccacagca aaagcaacca ggaagttctg gagtttgtca ccagtggagg
1141 ccggatggac ccacccaaga actgccttgg gctgtgtatc cggataatga ctcagtgtgt
1201 gcaacatcag cctgaagaca ggcccaactt tgccatcatt ttggagagga ttgaatactg
1261 caccagagac ccggatgtaa tcaacaccgc ttgcccgaata gaatatggtc cacttgggga
1321 agaggaagag aaagtgcctg tagggcccaa ggaccctgag ggggttccct ctctcctgtg
1381 ctctcaacag gcaaaaaggg aggaggagcg cagcccagct gcccaccac ctctgcctac
1441 caccctctct ggcaaggctg caaagaaacc cacagctgca gaggtctctg ttcgagtcct
1501 tagaggggcg gccgtggaag ggggacacgt gaatatggca ttctctcagt ccaacctcc
1561 ttcggagtgg cacagggtcc acggatccag aaacaagccc accagcttgt ggaacccaac
1621 gtacggctcc tgggtttacag agaaacccac caaaaagaat aatcctatag caaagaagga
1681 gccacacgag aggggttaacc tggggctgga gggagctgt actgtccac ctaacgttgc
1741 aactgggaga ctccggggg cctcactgct cctagagccc tcttcgctga ctgccaatat
1801 gaaggaggta cctctgttca ggctacgtca ctcccttgt gggaatgtca attacggcta
1861 ccagcaacag ggcttgcctt tagaagccgc tactgcccct ggagctggtc attacggaga
1921 taccattctg aaaagcaaga atagcatgaa ccagcctggg ccctgagctc ggtcgacac
1981 tcaattctct tcttgggat cctaagacc gtgg

EML4-ALK 변이체 5a 단백질 서열 (BAG75148.1; GI:209837706)

1 mdgfagsladd ssaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva

61 svkksvsskv yrrkhqelqa mqlmqlspspey klslkrtsti mtdynpnycf agktssisd1
 121 kevprknitl irglghgafg evyegqvsgm pndpsplqva vktlpevcse qdeldflmea
 181 liiskfnhqn ivrcigvslq slprfillel maggdlsfl retrprpsqp sslamlldilh
 241 vardiacgcq yleenhfihr diaarncllt cpypgrvaki gdfgmardiy rasyyrkggc
 301 amlpvkwmp eafmegifts ktdtwsfgvl lweifslgym pypsksnqev lefvtsggrm
 361 dppkncpgpv yrimtqcwqh qpedrpnfai ilerietctq dpdvintalp ieygplveee
 421 ekvpvrpkdp egvppllvsq qakreersp aappplptts sgkaakkpta aevsvrvprg
 481 pavegghvnm afsqsnppse lhrvhgsrnk ptslwnptyg swftekptkk nnpiakkeph
 541 ergnlglegs ctvpnpvatg rlpgasille pssltanmke vplfirlhfp cgnvnygyqg
 601 qglpleaata pgaghyedti lksksnmnp gp

EML4-ALK 변이체 5b 단백질 서열 (AB374365.1; GI:209837707)

1 actctgtcgg tccgctgaat gaagtgcctg cccctctaag ccggagagcc ggcgctttcc
 61 ccgcaagatg gacggtttcc cggcagctct ccatgatagt atttctgtg caagtacttc
 121 tgatgttcaa gatcgctgt cagctcttga gtcacagatt cagcaacaag aagatgaaat
 181 cactgtgcta aaggcggtct tggctgatgt tttgagcgct cttgcaatct ctgaagatca
 241 tgtggcctca gtgaaaaaat cagctctcaag taaaggttca gagctcaggg gaggatattg
 301 agatccaggg aggccttctg taggaagtgg cctgtgtagt gcttcaaggg ccaggctgcc
 361 aggcctatgt gcagctgacc accacacgtg agtgtagcgc cggaagcacc aggagctgca
 421 agccatgcag atggagctgc agagccctga gtacaagctg agcaagctcc gcacctogac
 481 catcatgacc gactacaacc ccaactactg ctttctgtgc aagacctcct ccatcagtga
 541 cctgaaggag gtgcccgcga aaaacatcac cctcattcgg ggtctgggcc atggagcctt
 601 tggggagggtg tatgaaggcc aggtgtccgg aatgcccac gacccaagcc ccctgcaagt
 661 ggctgtgaag acgctgcctg aagtgtgctc tgaacaggac gaactggatt tctcatgga
 721 agccctgatc atcagcaaat tcaaccacca gaacattgtt cgctgcattg gggtgagcct
 781 gcaatccctg ccccggttca tctgtctgga gctcatggcg gggggagacc tcaagtcctt
 841 cctccgagag acccgccctc gcccgagcca gccctcctcc ctggccatgc tggaccttct
 901 gcacgtggct cgggacattg cctgtggctg tcagtatttg gaggaaaacc acttcatcca
 961 ccgagacatt gctgcagaa actgcctctt gacctgtcca ggccctggaa gagtggccaa
 1021 gattggagac ttccggatgg cccgagacat ctacagggcg agctactata gaaagggagg
 1081 ctgtgccatg ttcgcatgta agtggatgcc cccagagggc ttcattggaag gaattatcac
 1141 ttctaaaaa gacacatggt cctttggagt gctgctatgg gaaatctttt ctcttgata
 1201 tatgccatc cccagcaaaa gcaaccagga agttctggag ttgtcacca gtggaggccg
 1261 gatggaccca cccaagaact gccctgggccc tgtataccgg ataagtactc agtgcctgga
 1321 acatcagcct gaagacaggg ccaactttgc catcattttg gagaggattg aatactgcac
 1381 ccaggagccc gatgtaatca acaccgcttt gccgatagaa tatgggtccac ttgtggaaga
 1441 ggaagagaaa gtgctctgga ggcccaagga cctgagggg gtctcctctc tctgggtctc
 1501 tcaacaggca aaacgggagg aggagcgag cccagctgcc ccaccacctc tgcctaccac
 1561 ctctctgtgc aaggctgcaa agaaaaccac agctgcagag gtctctgttc gagtccctag
 1621 agggccggcg gtggaagggg gacacgtgaa tatggcattc tctcagtcga accctccttc
 1681 ggagttgcac aggtgccacg gatccagaaa caagcccacc agcttctgga acccaacgta
 1741 cggtctctgg tttaacagaga aaccaccaa aaagaataat cctatagcaa agaaggagcc
 1801 acacgagagg ggtaacctgg ggctggaggg aagctgtact gtcccaccta acgttgcaac
 1861 tgggagactt ccgggggccc cactgctcct agagccctct tcgctgactg ccaatatgaa
 1921 ggaggtacct ctgttcaggc tacgtcactt ccctgtggg aatgtcaatt acggctacca
 1981 gcaacagggc ttgcccttag aagccgctac tgcccctgga gctggctcatt acgaggatac
 2041 cattctgaaa agcaagaata gcatgaacca gcctgggccc tgagctcggg cgcacactca
 2101 cttctcttcc ttgggatccc taagaccgtg g

EML4-ALK 변이체 5b 단백질 서열 (BAG75149.1; GI:209837708)

1 mdgfgslldd ssaastsdv qdrlsalesr vqqqedeitv lkaaladvlr rlaisedhva
 61 svkksvsskv selrggygdp grlpvsglc sasrarlpgh vaadhppavy rrkhhqelqam
 121 qmelqspeyk lsklrtstim tdynpnycfa gktssisdik evprknitli rglghgafge
 181 vyegqvsgmp ndpsplqvav ktlpevcseq deldflmeal iiskfnhqn vrcigvslqs
 241 lprfillelm aggdlsflr etrprpsqps slamlldlhv ardiacgcqy leenhfihrd
 301 iaarnclltc ppggrvakig dfgmardiy asyrrkggca mlpvkwmppe afmegiftsk
 361 tdtwsfgvll weifslgym ypsksnqevl efvtsggrmd ppkncpgpv rmtqcwqh
 421 pdrpnfai lerietctq pdvintalpi eygplveeee kvprvpkdp gvppllvssq
 481 akreerspa appplptts gkaakkpta evsvrvprgp avegghvnm fsqsnppsel

[0043]

541 hrvhgsrnkp tslwnptygs wftektkkn npiakkephe rgnlglegsc tvppnvatgr
601 lpgaslillep ssltanmkev plfrlrhfpv gnvnygyqqg glpleaatap gaghyedtil
661 ksknsmnqpg p

EML4-ALK 변이체 6 핵산 서열 (AB462411.1; GI:227452648)

1 tactctgtcg gtccgctgaa tgaagtcccc gccctcttaa gcccgagacc cggcgcttcc
61 cccgcaagat ggacgggttc gccggcagtc tcgatgatat tatttctgct gcaagtactt
121 ctgatgttca agatcgcttg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa
181 tcaactgtgt aaaggcggtt ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc
241 atgtggcctc agtgaaaaaa tcagctctca gtaaaaggcca accaagccot cgagcagtta
301 ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg
361 ctgtctcaat tgcaggaaaa gaaactcttt catctgctgc taaaagtggg acagaaaaaa
421 agaaagaaaa accacaaggc cagagagaaa aaaaagagga atctcattct aatgatcaaa
481 gtccacaaat tcgagcatca ccttctcccc agccctcttc acaacctctc caaatacaca
541 gacaaactcc agaaagcaag aatgtacttc ccaccaaaa cataaaacga ccatcaccag
601 ctgaaaagtc acataattct tgggaaaatt cagatgatat ccgtaataaa ttgtcgaaaa
661 taccttcaac acccaaatat ataccaaaag ttaccaaaac tgcagacaag cataaagatg
721 tcatcatcaa ccaagaagga gaatatatta aatgtttat gcgcggtcgg ccaattacca
781 tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaaactg cctcctgaga
841 agctcaaaat ggagtgggca tatggttato gaggaagga ctgtagagct aatgtttacc
901 ttcttccgac cggggaaaata gtttatttca ttgcatcagt agtagtacta tttaattatg
961 aggagagaa ctcagcgacac tacctggggc atacagactg tgtgaaatgc cttgctatcc
1021 atcctgacaa aattagattt gcaactggac agatagctgg cgtggataaa gatggaaggc
1081 ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg
1141 gacttggcac ttttgagcgt ggagtaggat gcctggattt ttcaaaagca gattcagggtg
1201 ttcatctata tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga
1261 ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggct gtggagtctc
1321 acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatatttcc ttctggacct
1381 ggagcggcaa ttactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat
1441 ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag
1501 tcatgcttat atggagcaaa actactgtag agccacacac tgggaaagga cctaaaggaa
1561 gtggcctgtg tagtgcttca agggccaggc tgccaggcca tgttgagact gaccacccac
1621 ctgcagtgta ccgcccgaag caccaggagc tgcaagccat gcagatggag ctgcagagcc
1681 ctgagtacaa gctgagcaag ctccgcacct cgaccatcat gaccgactac aaccccaact
1741 actgctttgc tggcaagacc tcctccatca gtgacctgaa ggaggtgccc cggaaaaaca
1801 tcaccctcat tcggggtctg ggccatggag cctttgggga ggtgtatgaa ggccaggtgt
1861 ccggaatgcc caacgaccca agcccccctg aagtggctgt gaagacgctg cctgaagtgt
1921 gctctgaaca ggaagcaact gatttctcca tggaaagccct gatcatcagc aaattcaacc
1981 accagaacat tgttcgctgc attggggtga gcctgcaatc cctgccccgg ttcactctgc
2041 tggagctcat ggccggggga gacctcaagt ccttctcccg agagacccc cctcgcccga
2101 gccagccctc ctccctggcc atgctggacc ttctgcacgt ggctcgggac attgctgtg
2161 gctgtcagta tttggaggaa aaccacttca tccaccgaga cattgctgcc agaaactgcc
2221 tcttgacctg tcacggccct ggaagagtgg ccaagattgg agacttcggg atggcccag
2281 acatctacag ggcgagctac tatagaaagg gaggtctgtc catgctgcca gtttaagtga
2341 tgcccccaag ggccttcatg gaaggaatat tcacttctaa aacagacaca tggctccttg
2401 gagtgtctgt atgggaaatc ttttctcttg gatatatgcc ataccccagc aaaagcaacc
2461 aggaagtctt ggagtgtgtc accagtggag gccggatgga cccacccaag aactgccttg
2521 ggctgttata ccggataaat actcagtgtc ggcaacatca gcctgaagac aggcccaact
2581 ttgccatcat tttggagagg attgaatact gcacccagga cccggatgta atcaacaccg
2641 ctttggcgat agaatatggt ccacttgttg aagaggaaga gaaagtgcct gtgaggccca
2701 aggaccctga ggggggttct cctctccttg tctctcaaca ggcaaaacgg gagggaggag
2761 gcagcccagc tgccccacca cctctgccta ccacctctc tggcaaggct gcaaaagaaac
2821 ccacagctgc agagggtctc gttcagatcc ctgaggggcc ggccgtggaa gggggacacg
2881 tgaatatggc attctctcag tccaaccctc ctcggaggtt gcacagggtc cacggatcca
2941 gaaacaagcc caccagcttg tggaaaccaa cgtacggctc ctggtttaca gagaaccca
3001 ccaaaaagaa taatctata gcaaaagagg agccacacga gaggggtaac ctggggctgg
3061 agggaaagct tactgtccca cctaactgtg caactgggag acttcggggg gcctcactgc
3121 tcctagagcc ctcttcgctg actgccaata tgaaggaggt acctctgttc aggctacgtc
3181 acttcccttg tgggaatgtc aattacggct accagcaaca gggtctgccc ttagaagccg
3241 ctactgcccc tggagctggt cattacgagg ataccattct gaaaagcaag aatagcatga

[0044]

3301 accagcctgg gccctgagct cggtcgcaca ctcaactctc ttccttgga tccctaagac
3361 cgtgg

EML4-ALK 변이체 6 단백질 서열 (BAH57335.1; GI:227452649)

1 mdgfasldd sisaastdv qdrlesalesr vqqgedeiv lkaaladvlr rlaisedhva
61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
121 kpggqrekke eshsndqspq iraspspqs sqplqihrt pesknaptk sikrpspaek
181 shnswensdd srnklskips tpklipkvtk tadkhkdvii ngegeyikmf mrgrpitmfi
241 psdvdyddi rtelppekkl lewaygyrgk dcranvylp tgeivvyfias vvvlfnyeer
301 tqrhylghtd cvkclaihp kiriatggia gvdkgdrplq phvrwdsvt lstlqiiglg
361 tfergvvgld fskadsgvhl cviddsnehm ltvdwqgrka kgaeikttne vvlavefhpt
421 dantiitcgk shiffwtwsg nsltrkggif gkyekpkfvq clafingdv ltgdsqgvml
481 iwskttept pgkpgksgl csasrarlpq hvaadhppav yrrkhqelqa mqmelqspey
541 klsklrtsti mtdynpnycf agktssisd1 kevprknitl irglghgafg evyegqvsgm
601 pndpsplqva vktlpevcse qdeldflema liiskfnhqn ivrcigvslq slprfillel
661 maggdlsfsl retrprsqps sslamlldllh vardiacgcq yleenhfihv diaarncllt
721 cpqpggrvaki gdfgmardiy rasyyrkqgc amlpvkwmp eafmegifts ktdtwsfgvl
781 lweifslgym pypsksnqev lefvtsggrm dppkncpgpv yrintqcwqh qpdrpnfai
841 ilerleyctg dpdvintalp ieygplveee ekvprpkdp egvppllvsq qakreersp
901 aappplptts sgkaakkpta aevsvrvprg pavegghvnm afsqsnppse lhrvhgsrnk
961 ptslwnptyg swftekptk nnpiakkeph ergnllegs ctvppnvatg rlpgasllle
1021 pssltanmke vplfrlrhfp cgnvnygyq qglpleaata pgaghyedti lksknsnmqp
1081 gp

EML4-ALK 변이체 7 핵산 서열 (AB462412.1; GI:227452650)

1 tactctgtcg gtcgcgtgaa tgaagtgcgc gccctctaa gcccgagcc cggcgcttcc
61 ccgcgaagat ggacggttcc gccgcagtc togatgatag tattctgtc gcaagtactt
121 ctgatgttca agatgcgctg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa
181 tcactgtgct aaaggcggtc ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc
241 atgtggcctc agtgaaaaaa tcagtctcaa gtaaaaggcca accaagccct cgagcagtta
301 ttcccatgtc ctgtataacc aatggaaagt gtgcaaacag aaaaccaagt cataccagtg
361 ctgtctcaat tgcagaaaaa gaaactcttt catctgctgc taaaagtgtg acagaaaaaa
421 agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa
481 gtccacaaat tcgagcatca ctttctcccc agccctcttc acaacctctc caaatacaca
541 gacaaactcc agaaagcaag aatgctactc ccacaaaag cataaaaaga ccatcaccag
601 ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa
661 taccttcaac acccaaatat ataccaaaag ttaccaaaac tgcagacaag cataaagatg
721 tcatcatcaa ccaagaagga gaatatatta aatgtttat gcgcggtcgg ccaattacca
781 tgtttcattcc ttccgatgtt gacaaactat atgacatcag aacggaactg cctcctgaga
841 agctcaaaact ggagtgggca tatgggtatc gaggaagga ctgtagagct aatgtttacc
901 ttcttccgac cggggaata gtttatttca ttgcatcagt agtagtacta ttttaattatg
961 agggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgtctatc
1021 atctgacaaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc
1081 ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg
1141 gacttggcac ttttgagcgt ggagtaggat gcctggattt ttcaaaaaga gattcaggtg
1201 ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga
1261 ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggtc gtggagtctc
1321 acccaacaga tgcataatcc ataattacat gcggtaaatc tcatattttc ttctggacct
1381 ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat
1441 ttgtgcagtg tttagcatc ttgggggaat gagatgttct tactggagac tcagggtggg
1501 tcatgcttat atggagcaaa actactgtag agccacacc tgggaaaaga cctaaaggtg
1561 tatatcaaat cagcaaacaa atcaaagctc atgatggcag tgtgttcaca cttgtgcaga
1621 tgagaaatgg gatgttatta actggaggag ggaagacag aaaaataatt ctgtgggac
1681 atgatctgaa tcttgaaga gaaatagagc accaggagct gcaagccatg cagatggagc
1741 tgcagagccc tgagtacaag ctgagcaagc tccgcacctc gaccatcatg accgactaca
1801 accccaacta ctgctttgct ggcaagacct cctccatcag tgacctgaag gaggtgccgc
1861 ggaaaaacat caccctcatt cggggtcttg gccatggagc ctttggggag gtgtatgaag
1921 gccaggtgtc cggaatgcc aacgaccaa gccccctgca agtggctgtg aagacgctgc
1981 ctgaagtgtg ctctgaacag gacgaactgg atttctcat ggaagccctg atcatcagca

[0045]

2041 aattcaacca ccagaacatt gtctgctgca ttgggggtgag cctgcaatcc ctgccccggt
 2101 tcatcctgct ggagctcatg gcggggggag acctcaagtc cttcctccga gagaccgccc
 2161 ctgccccgag ccagccctcc tccctggcca tgctggacct tctgcacgtg gctcgggaca
 2221 ttgctctgtg ctgtcagtat ttggaggaaa accacttcat ccaccgagac attgtgtcca
 2281 gaaactgcct cttgacctgt ccaggccctg gaagagtggc caagattgga gacttcggga
 2341 tggccccaga catctacagg gcgagctact atagaaaggg aggtgtgtcc atgtgtccag
 2401 ttaagtggat gccccagag gccttcatgg aaggaaatatt cacttctaaa acagacacat
 2461 ggtccttttg agtgctgcta tgggaaatct tttctcttgg atatatgcca taccacagca
 2521 aaagcaacca ggaagtctct gagttttgtc ccagtggagg ccggtatggc ccaccaaga
 2581 actgccctgg gcctgtatac cggataatga ctcaagtgtg gcaacatcag cctgaagaca
 2641 ggcccaactt tgccatcatt ttggagagga ttgaatactg caccacaggc ccggtatgta
 2701 tcaacaccgc ttgtccgata gaatatggtc cacttgtgga agaggaagag aaagtgcctg
 2761 tgaggcccaa ggaccctgag ggggttcctc ctctcctggt ctctcaacag gcaaaacggg
 2821 agggaggagc cagccagctg gcccaccac ctctgcctac cactcctctt ggcaaggctg
 2881 caaagaaaacc cacagctgca gaggtctctg ttcgagtccc tagaggggccg ccgtggaag
 2941 ggggacacgt gaatatggca ttctctcagt ccaaccctcc ttcggagtgt cacaagggtc
 3001 acggtatccg aaacaagccc accagcttgt ggaacccaac gtacggctcc tggtttacag
 3061 agaaacccac caaaaaaat aatcctatag caaagaagga gccacacgac aggggtaacc
 3121 tggggctgga gggaaactgt actgtcccac ctaacgttgc aactgggaga cttccggggg
 3181 cctcactgct cctagagccc tcttcgctga ctgccaatat gaaggaggtg cctctgttca
 3241 ggctacgtca cttcccttgt gggaaatgtc attacggcta ccagcaacag ggcttgccct
 3301 tagaagccgc tactgcccct ggagctggtc attacagga taccattctg aaaagcaaga
 3361 atagcatgaa ccagcctggg ccttagctc ggtcgcacac tcacttctct tccttgggat
 3421 ccctaagacc gtgga

EML4-ALK 변이체 7 단백질 서열 (BAH57336.1; GI:227452651)

1 mdgfagslldd sistaatsdv qdrilsalesr vqqgedeitv lkaaladvlr rlaisedhva
 61 svkksvsskg qpspravipm scitngsgan rkpshtsavs iagketlssa aksgtekkke
 121 kpqgqrekke eshsndqspg iraspspqs sqplqihrt pesknaptk sikrpspaek
 181 shnswnsdd srnklskips tpklipkvtk tadkhkdvii ngegeyikmf mrgrpitmfi
 241 psdvndyddi rtelppekkl lewaygyrgk dcranvylp tgeivyfias vvvlfnyeer
 301 tqrhylghtd cvkclaihp d kiriatggia gvdkgdgrplq phvrwdsvt lstlqiiglg
 361 tfergvglcd fskadsgvhl cviddsnehm ltvwdwqrka kgaektne vvlavefhpt
 421 dantiitcgk shiffwtwsq nsltrkqgif gkyekpkfvq clafngndv lgtgsggvml
 481 iwskttevept pgkpgkvyq iskqikahdg svftlcqmrn gmltgggkd rkiilwdhdl
 541 npereiehqe lqamqmlqs peyklsklrt stimtdynpn ycfagktssi sdlkevprkn
 601 itlirglghg afgevyeggv sgmpndpspl qvavktlpev cseqdeldfl mealiskfn
 661 hqnivrcigv slqslprfil lelmaggdlk sfldretrpr sqpsslamld llhvardiac
 721 gcqyleenhf ihrdiaarn lltcpgpgrv akigdfgmar diyasyyrk ggcamlpvkw
 781 mpeafmegi ftsktdtwsf gvllweifsl gypypsksn qevlefvtsq grmdppkncp
 841 gpvyrimtgc wqhqpdrpn failerley ctqdpdvint alpieygplv eeeekvpvrp
 901 kdpegvplll vsqakreee rspaappplp ttssgkaakk ptaaevsrv prgpaveggh
 961 vnmafsqsnp pselhkhvhs rnkptslwnp tygswftekp tkknnpiaak ephdrnlgf
 1021 egscvtppnv atgrlpgasl llepssltan mkevplfrlr hfpcgnvnyg yqqqglplea
 1081 atapgaghye dtilksknsm nqpgp

KIF5B-ALK 핵산 서열 (AB462413.1; GI:227452652)

1 tgcgagaaaag atggcggacc tggccgagtg caacatcaaa gtgatgtgtc gcttcagacc
 61 tctcaacgag tctgaagtga accgcggcga caagtacatc gccaaagtttc agggagaaga
 121 cacggtcgtg atcgcgctoca agccttatgc atttgatcgg gtgttccagt caagcacatc
 181 tcaagagcaa gtgtataatg actgtgcaaa gaagattgtt aaagatgtac ttgaaggata
 241 taatggaaca atatttgcac atggacaaac atcctctggg aagacacaca caatggaggg
 301 taaacttcat gatccagaag gcatgggaat tattccaaga atagtgaag atatttttaa
 361 ttatatattc tccatggatg aaaatttggg atttcatatt aaggtttcat attttgaaat
 421 atatttggat aagataaggg acctgttaga tgtttcaaag accaaacctt cagttcatga
 481 agacaaaaac cgagttccct atgtaaaggg gtgcacagag cgttttgtat gtagtccaga
 541 tgaagttatg gataccatag atgaaggaaa atccaacaga catgtagcag ttacaaatat
 601 gaatgaacat agctctagga gtcacagtat atttcttatt aatgtcaaac aagagaacac
 661 acaaacggaa caaaagctga gtggaaaact ttatctggtt gatttagctg gtagtgaaaa

[0046]

```

721 ggtagtagtaaa actggagctg aaggtgctgt gctggatgaa gctaaaaaca tcaacaagtc
781 actttctgct cttggaaaatg ttatttctgc tttggctgag ggtagtacat atgttccata
841 tcgagatagt aaaaatgacaa gaatccttca agattcatta ggtggcaact gtagaaccac
901 tattgtaatt tgctgctctc catcatcata caatgagctt gaaacaaaat ctacactctt
961 atttggccaa agggccaaaa caattaagaa cacagtgtgt gtcaatgtgg agttaactgc
1021 agaacagtgg aaaaagaagt atgaaaaaga aaaagaaaaa aataagatcc tgcggaacac
1081 tattcagtggt cttgaaaatg agctcaacag atggcgtaat ggggagacgg tgccatttga
1141 tgaacagttt gacaaaagaga aagccaactt ggaagctttc acagtggata aagatattac
1201 tottaccaat gataaaccag caaccgcaat tggagttata ggaattttta ctgatgctga
1261 aagaagaaaag tgtgaagaag aaattgctaa attatacaaa cagcttgatg acaaggatga
1321 agaaaattaac cagcaaagtc aactggtaga gaaactgaag acgcaaatgt tggatcagga
1381 ggagcttttg gcatctacca gaagggatca agacaatatg caagctgagc tgaatcgctt
1441 tcaagcagaa aatgatgcct ctaaaagaag agtgaaagaa gttttacagg ccctagaaga
1501 acttgctgtc aattatgatc agaagtctca ggaagttaa gacaaaacta aggaatatga
1561 attgcttagt gatgaattga atcagaaatc ggcaacttta gcgagtatag atgctgagct
1621 tcagaaactt aaggaaatga ccaaccacca gaaaaaacga gcagctgaga tgatggcatc
1681 tttactaaaa gaccttgcat aaataggaat tgctgtggga aataatgatg taaagcagcc
1741 tgagggaaact ggcctgatag atgaagagtt cactgttgca agactctaca ttacgaaaaa
1801 gaagtcagaa gtaaaaacca tggtagaaac ttgcaagcag ttagaaagca cacaaactga
1861 gagcaacaaa aaaatggaag aaaatgaaaa ggaagttagca gcattgtcagc ttctgtatctc
1921 tcaacatgaa gccaaaatca agtcattgac tgaatacctt caaaatgtgg aacaaaagaa
1981 aagacagttg gaggaatctg tcgatgcctt cagtgaagaa ctagtccagc ttctgagcaca
2041 agagaaagtc catgaaatgg aaaaggagca cttaaataag gttcagactg caaatgaagt
2101 taagcaagct gttgaacagc agatccagag ccatagagaa actcatcaaa aacagatcag
2161 tagtttgaga gatgaagtga aagcaaaagc aaaacttatt actgatcttc aagacaaaaa
2221 ccagaaaatg atgttagagc aggaacgtct aagagtagaa catgagaagt tgaagccac
2281 agatcaggaa aagagcagaa aactacatga acttacggtt atgcaagata gacgagaaca
2341 agcaagacaa gacttgaagg gtttgaaga gacagtggca aaagaacttc agactttaca
2401 caacctgctc aaactctttg ttcaggacct ggctacaaga gttaaaaaga gtgctgagat
2461 tgattctgat gacaccggag gcagcgctgc tcagaagcaa aaaatctcct ttcttgaaaa
2521 taatcttgaa cagctcacta aagtgacaaa acagtgtgta cgtgataatg cagatctccg
2581 ctgtgaactt cctaatgttg aaaagcgact tcgagctaca gctgagagag tgaagctttt
2641 ggaatcagca ctgaaaagaag ctaaaagaaa tgcattctgt gatcgcaaac gctatcagca
2701 agaagtagat cgcataaagg aagcagtcag gtcaaaagaa atggccagaa gagggcattc
2761 tgacacagatt gtgtaccgcc ggaagcaca ggagctgcaa gccatgcaga tggagctgca
2821 gagccctgag tacaagctga gcaagctccg cactcgacc atcatgaccg actacaaccc
2881 caactactgc tttgctggca agacctctc catcagtgac ctgaaggagg tgcgcgga
2941 aaacatcacc ctcatctggg gctctggcca tggcgctttt ggggaggtgt atgaaggcca
3001 ggtgtccgga atgcccacac acccaagccc cctgcaagtg gctgtgaaga cgtgctctga
3061 agtgtgctct gaacaggagc aactggattt cctcatggaa gccctgatca tcagcaaat
3121 caaccaccag aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat
3181 cctgctggag ctcatggcgg ggggagacct caagtcttct ctccgagaga cccgcctcg
3241 cccgagccag cctctctccc tggccatgct ggacctctg cactgtgctc gggacattgc
3301 ctgtggctgt cagtatttgg aggaaaacca ctcatccac cgagacattg ctgccagaaa
3361 ctgctctctt acctgtccag gccctggaag agtggccaag attggagact tcgggatggc
3421 ccgagacatc tacagggcga gctactatag aaaggagggc tgtgccatgc tggcagttaa
3481 gtggatgccc ccagaggcct tcatggaagg aatattcact tctaaaaacag acacatggtc
3541 ctttggagtg ctgctatggg aaatcttttc tcttgatat atgccatacc ccagcaaaag
3601 caaccaggaa gttctggagt ttgtcaccag tggaggccgg atggaccac ccaagaactg
3661 ccttgggctt gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc
3721 caactttgcc atcatttttg agaggattga atactgcacc caggaccggg atgtaatcaa
3781 caccgctttg ccgatagaat atggtccact tgtggaagag gaagagaaaag tgcctgtgag
3841 gcccaaggac cctgaggggg ttctctctct cctggctctt caacaggcaa aacgggagga
3901 ggagcgcagc ccagctgccc caccacctct gctaccacc tctctggcca aggctgcaaa
3961 gaaacccaca gctgcagagg tctctgttct agtccctaga gggccggcgg tggaaagggg
4021 acacgtgaat atggcattct ctcagtccaa cctctcttct gagttgcaca aggtccacgg
4081 atccagaaaac aagcccacca gcttgtggaa cccaacgtac ggtcctggtt ttacagagaa
4141 acccaacaaa aagaataatc ctatagcaaa gaaggagcca cagcacaggg gtaacctggg
4201 gctggaggga agctgtactg tcccacctaa cgttgcaact gggagacttc cgggggctc
4261 actgctccta gagcctctt cgtgactgc caatatgaag gaggtacctc tgttcaggct
4321 acgtcacttc ccttgtggga atgtcaatta cggctaccag caacagggct tgccttaga

```

[0047]

4381 agccgctact gcccttgag ctggtcatta cgaggatacc attctgaaaa gcaagaatag
4441 catgaaccag cctgggcccct gagctcggtc gcacactca

KIF5B-ALK 단백질 서열 (BAH57337.1; GI:227452653)

1 madlaecnk vmcrfplne sevnrgdkiy akfqqgedtvv iaskpyafdr vfqsstsqeq
61 vyndcakkiv kvlegyngt ifayggtssg kthtmegklh dpegmglipr ivqdifnyiy
121 smdenlefhi kvsyfeiyld kirdlldvsk tnlsvhedkn rvpyvkgcte rfvcspdevm
181 dtidegksnr hvavtnmneh ssrshsifli nvkqentgte qklsqklylv dlagsekvsk
241 tgaegavide akninksisa lgnvisalae gstyvpyrds kmtrilqdsi ggnrcrttvi
301 ccspssyns etkstllfqq raktikntvc vnveltaeqw kkkyekekek nkilrntiqw
361 lenelnrwrn getvpideqf dkekanleaf tvdkditltn dkpataigvi gnfdaerrk
421 ceeeiaklyk qlddkdeein qqsqlevklk tqmldqeell astrrdqdnm gaelnrlgae
481 ndaskeevke vlqaleelav nydqksqeve dktkeyells delnqksatl asidaelqkl
541 kemtnhqkkr aaemmasilk dlaeigiavg nndvkqpegt gmiddeftva rlyiskmkse
601 vktmvkroqk lestqtesnk kmeenekela acqlrisqhe akikslteyl qnveqkkrql
661 eesvdalsee lvqlraqekv hemekehlk vtanevkqa veqqiqshre thqkqisslr
721 deveakakli tdlqdnqkqm mleqerlrve hekldatdqe ksrklheltyv mqrdrreqarq
781 dlkgleetva kelqtlhnlr klfvqdlatr vkksaeidsd dtggsaaqkq kisflennle
841 qltkvhkqlv rdnadlrcel pklekrllrat aervkalesa lkeakenasr drkryqqevd
901 rikeavrskn margrhsaqi vyrkqhqlq amqmelqspe yklsklrtst imtdynpnyc
961 fagktssisd lkevprknit lirlghgaf gevyeqqvsg mpndpsplqv avktlpevcs
1021 eqdeldfllme aliiskfnhq nivrcigvsi qslprfille lmaggdlksf lretrprpsq
1081 psslamldll hvardiacgc qyleenhfih rdiaarncll tcpgpgvrak igdfgmardi
1141 yrasyyrkkg camlpvkmpw peafmegift sktdtwsfgv llweifslgy mpypsksnqe
1201 vlefvtsggr mdpkncpgp vyrmtqcwq hqpdrpnfa iilerieyct qdpdvintal
1261 pieygpvlvee eekvpvrpkl pegvpllvvs qqakreeers paappplptt ssgkaakkpt
1321 aaevsvrvpr gpaveghvnd mafsqsnpps elhkvhgsrn kptslwnpty gswftekptk
1381 knnpiakkep hdnrgnlqleg sctvppnvat grlpqasill epssltanmk evplfrlrhf
1441 pcgnvnygyq qqglpleaat apgaghyedt ilksknsnmq pgp

NPM-ALK 서열 (t(2;5) (p23;q35) 염색체 전위) *

TPM3-ALK 서열 (t(1;2) (p25;p23) 염색체 전위) *

TFGXL-ALK 핵산 서열 (AF390893.1; GI:20269389)

1 atgaacggac agttgatct aagtgggaag ctaatcatca aagctcaact tggggaggat
61 attcggcgaa ttccatttca taatgaagat attacttatg atgaattagt gctaattgatg
121 caacaggttt tcagaggaaa acttctgagt aatgatgaag taacaataaa gtataaagat
181 gaagatggag atcttataac aatttttgat agttctgacc ttccctttgc aattcagtg
241 agtaggatac tgaaactgac attatttgtt aatggccagc caagaccctt tgaatcaagt
301 caggtgaaat atctccgtcg agaactgata gaacttcgaa ataaagtga tctgttattg
361 gatagcttgg aaccacctgg agaaccagga ccttccacca atattcctga aatgatact
421 gtggatggta gggaagaaaa gtctgcttct gattcttctg gaaaacagtc tactcagggt
481 atggcagcaa gtatgtctgc ttttgatcct ttaaaaaacc aagatgaaat caataaaaaat
541 gttatgtcag cgtttgtctt aacagatgat caggtttcag ggccaccacg tgctcctgca
601 gaagatcggt caggaaacac cgacagcatt gcttccctct cctcagcagc tccccacca
661 ggcgttcagc cacagcagcc accatatata ggagctcaga ctcaagcagg tcagattgaa
721 gtgtaccgcc ggaagaccca ggagctgcaa gccatgcaga tggagctgca gagccctgag
781 tacaagctga gcaagctccg cactcagacc atcatgaccg actacaaccc caactactgc
841 tttgctggca agacctctct catcagtgac ctgaaggagg tgcccgcgaa aaacatcacc
901 ctcatctcgg gtctgggcca tggcgccctt ggggaggtgt atgaaggcca ggtgtccgga
961 atgcccacac acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct
1021 gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaat caaccaccag
1081 aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag
1141 ctcatggcgg ggggagacct caagtctctc ctccgagaga cccgccctcg cccgagccag
1201 ccctcctccc tggccatgct ggacctctct cactggtgctc gggacattgc ctgtggtgt
1261 cagttatttg aggaataaca cttcatccac cgagacattg ctgccagaaa ctgcctcttg
1321 aacctgccag gccttggaag agtggccaag attggagact tcgggatggc ccgagacatc

[0048]

1381 tacagggcga gctactatag aaagggagggc tgtgccatgc tggcagttaa gtggatgccc
1441 ccagaggcct tcatgggaag aatattcact tctaaaacag acacatgggc ctttggagtg
1501 ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag caaccaggaa
1561 gttctggagt ttgtcaccag tggaggcccg atggacccac ccaagaactg ccctgggctt
1621 gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc caacttttgc
1681 atcatttttg agaggattga atactgcacc caggaccccg atgtaatcaa caccgctttg
1741 ccgatagaat atgggccact tgtggaagag gaagagaaag tgcctgtgag gcccaaggac
1801 cctgaggggg ttctctctct cctggtctct caacaggcaa aacgggagga ggagcgcagc
1861 ccagctgccc caccacctct gcctaccacc tctctggca aggctgcaaa gaaaccoca
1921 gctgcagagg tctctgttct agtccctaga gggccggccg tggaggggg acacgtgaat
1981 atggcattct ctcagtccaa cctcctctcg gaggttgcaca aggtccacgg atccagaaac
2041 aagcccacca gcttgttgaa cccaacgtac ggctcctggt ttacagagaa acccaccaaa
2101 aagaataatc ctatagcaaa gaaggagcca cagcagagg gtaacctggg gctggaggga
2161 agctgtactg tcccacctaa cgttgcaact gggagacttc cgggggctc actgctccta
2221 gagccctctt cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc
2281 ccttgtggga atgtcaatta cggctaccag caacagggtc tgcctctaga agcgcctact
2341 gccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag
2401 cctgggcccct ga

TFGXL-ALK 단백질 서열 (AAM17922.1; GI:20269390)*

1 mngqldlsgk liikaqlged irripihned itydelvlmm qrvfrgklls ndevtikykd
61 edgdlitifd ssdlsfaiqc srilkltlfv ngqprpless qvkyllrrel elrnkvnrll
121 dsleppgepg pstnipendt vdgreekssas dssgkqstqv maasmsafdp lknqdeinkn
181 vmsafgltd qvsgppsapa edrsgtpdsi asssaaahpp gvqqqppyt gaqtqaggie
241 vyrrkhqelq amqmelgspe yklsklrltst imtdynpnyc fagktssisd lkevprknit
301 lirlghgaf gevyeqvgv mpndpsplqv avktlpevcs eqdeldflme aliiskfnhq
361 nivrcigvsl qslprfille lmaggdlksf lretrprpsq psslamldll hvardiacgc
421 qyleenhfih rdiaarncll tpgpgprvak igdfgmardi yrasyyrkkg camlpvkwmp
481 peafmegift sktdtwsfgv llweifslgy mpypsksnqe vlefvtsggr mdppkncpgp
541 vyrimtqcvw hqpdrpnfa iilerieyct qdpdvintal pieygpivee eekvpvrpkd
601 pegvppllvs qqakreeers paappplptt ssgkaakkpt aaevsrvrpr gpavegghvn
661 mafsqsnpps elhkvghsrn kptslwnpty gswftekptk knnpiakkep hdrnlgleg
721 sctvppnvat grlpqaslll epssltanmk evplfrlrhf pcgnvnvygq qqglpleaat
781 apgaghyedt ilksknsnmn pgp

TFGL-ALK 핵산 서열 (AF143407.1; GI:6739534)

1 cctccgcaag ccgtctttct ctatagttgt atatataaga catcctggag tccaccatga
61 acggacagtt gcatctaagt ggaagctaa tcatcaaaag tcaacttggg gaggatattc
121 ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcata atgatgcaac
181 gagttttcag aggaaaaact ctgagtaatg atgaagtaac aataaagtat aaagatgaag
241 atggagatct tataacaatt ttgatagtt ctgaccttct ctttgcaatt cagtgcagta
301 ggatactgaa actgacatta ttgttaatg gccagccaag accccttgaa tcaagtccag
361 tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata
421 gcttggaaac acctggagaa ccaggacctt ccaccaatat tccctgaaaa gatactgttg
481 atggtaggga agaaaagtct gcttctgatt cttctggaaa acagtctact caggttatgg
541 cagcaagtat gtctgctttt gatcctttta aaaaccaaga tgaaatcaat aaaaatgtta
601 tgtcagcggt tggcttaaca gatgatcagg ttctagtgta ccgcccgaag caccaggagc
661 tgcaagccat gcagatggag ctgcagagcc ctgagtacaa gctgagcaag ctccgcaact
721 cgaccatcat gaccgactac aaccccaact actgctttgc tggcaagacc tctccatca
781 gtgacctgaa ggaggtgccg cggaaaaaca tcacctcat tcggggtctg ggccatggcg
841 cctttgggga ggtgtatgaa ggcaggtgt ccggaatgcc caacgaccca agccccctgc
901 aagtggctgt gaagacgctg cctgaagtgt gctctgaaca ggacgaactg gatttctca
961 tggaaagccct gatcatcagc aaattcaacc accagaacat tgttcgctgc attggggtga
1021 gcctgcaatc cctgccccgg ttcatcctgc tggagctcat ggcgggggga gacctcaagt
1081 ccttccctcg agagccccgc cctcgccccg gccagccctc ctccctggcc atgctggacc
1141 ttctgcacgt ggctcgggac attgctgtgt gctgtcagta tttggaggaa aaccacttca
1201 tccaccgaga cattgctgcc agaaaactgc tcttgacctg tccaggccct ggaagagtgg
1261 ccaagattgg agacttcggg atggcccag acatctacag ggcgagctac tatagaaaag
1321 gaggtgtgc catgctgcca gtttaagtga tgcgccaga ggccttcctg gaaggaatat

[0049]

1381 tcactttctaa aacagacaca tggctccttg gagtgtgtgt atgggaaatc ttttctcttg
 1441 gatatatgcc ataccccgagc aaaagcaacc aggaagtctt ggagtttgtc accagtggag
 1501 gccggatgga cccacccaag aactgccctg gccctgtata ccggataatg actcagtgtc
 1561 ggcaacatca gcttgaagac agggcccaact ttgccatcat tttggagagg attgaatact
 1621 gcacccagga ccgggatgta atcaacaccg ctttgccgat agaataatgt coacttgttg
 1681 aagaggaaga gaaagtgcct gtgaggccca aggaccctga gggggttcct cctctccttg
 1741 tctctcaaca gccaaaacg gagggaggag gcaagccagc tgccccacca cctctgccta
 1801 ccacctctc tggcaaggct gcaaagaaac ccacagctgc agaggtctct gttcgagtcc
 1861 ctgaggggccc ggccgtggaa gggggacacg tgaatatggc attctctcag tccaacccct
 1921 cttcgagggt gcacaaggtc cacggatcca gaaacaagcc caccagcttg tggaaaccaa
 1981 cgtacggctc ctggttttaca gaaaaaccca ccaaaaagaa taatcctata gcaaagaagg
 2041 agccacacga caggggtaac ctggggctgg agggaagctg tactgtccca cctaacgttg
 2101 caactgggag acttccgggg gcctcactgc tcttagagcc ctcttcgctg actgccaata
 2161 tgaaggaggt acctctgttc aggtacgtc acttcccttg tgggaatgtc aattacgggt
 2221 accagcaaca gggcttgcct ttagaagccg ctactgcccc tggagcttgt cattacgagg
 2281 ataccattct gaaaaagcaag aatagcatga accagcctgg gccctgagct cggctgcaca
 2341 ctacatttct ttccttggga tccctaagac cgtggaggag agagaggcaa tggctccttc
 2401 acaaaccaga gaccaaagt cactgtttgt tttgtgcaa cctattttga agtaccacca
 2461 aaaaagctgt attttgaaaa tgcttttaga aggttttgag catgggttca tctatttct
 2521 tcgaaagaag aaaatatcat aaaatgagt gataaatata aggcccatg gtggttgcat
 2581 aaggttttta tgcattgttg ttgtatactt ccttatgctt cttttaaatt gtgtgtgctc
 2641 tgcttcaatg tagtcagaat tagctgcttc tatgtttcat agttgggtgc atagatgttt
 2701 ccttgccctg ttgatgtgga catgagccat ttgaggggag agggaacgga aataaaggag
 2761 ttatttggta tgaactaaa

TFGL-ALK 단백질 서열 (AAF27292.1; GI:6739535)*

1 mngqldlsgk liikaqlged irripihned itydelvlmm qrvfrgklls ndevtikykd
 61 edgdltitfd ssdlsfaiqc srilkltlfv ngqprpless qvkylrreli elrnkvnrll
 121 dsleppgepg pstnipendt vdgreeksas dssgkqstqv maasmsafdp lknqdeinkn
 181 vmsafgltdt qvsvyrrkhq elqamqmelq speyklslklr tstimtdynp nycfagktss
 241 isdlkevprk nitlirglgh gafgevyegq vsgmpndpsp lqvavktlpe vcseqdelfd
 301 lmealiiskf nhqnivrcig vslgslprfi llelmaggdl ksflretrpr psqpslaml
 361 dlhvardia cgcqyleenh fihrdiaarn cltctpgpgr vakigdfgma rdiyrasyr
 421 kggcamlpvk wmppeafmeg iftsktdtws fgvilweifs lgypypysks nqevlefvts
 481 ggrmdppknc pgpvyrmtq cwqhqpdrp nfailerie yctqdpdvin talpieygpl
 541 veeeekvprv pkdpegvppl lvsqqakree erspaapppl pttssgkaak kptaaevsvr
 601 vprgpavegg hvnmfsgsn ppselkhvhg srnkptslwn ptygswftek ptkknpiak
 661 kephdrnlg legscvppn vatgripas lllpepslta nmkevpflrl rhfpcgnvny
 721 gyqqqglple aatapgaghy editlkskns mnqpgp

TFGS-ALK 핵산 서열 (AF125093.1; GI:7229260)

1 cctccgcaag ccgtctttct ctgagttgt atatatagaa catcctggag tccaccatga
 61 acggacagtt ggaatcaagt gggaagctaa tcatcaaaag ccaactggg gaggatattc
 121 ggcgaaattcc tattcataat gaagatatta cttagatga attagtgtca atgatgcaac
 181 gaggttttcag aggaaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag
 241 atggagatct tataacaatt tttgatagtt ctgacctttc ctttgcaatt cagtgcagta
 301 ggatactgaa actgacatta tttgttaatg gccagccaag accccttgaa tcaagttagg
 361 tgaatatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata
 421 gcttggaacc acctggagaa ccaggacctt ccaccaatat tctgaaaaat gtgtaccgcc
 481 ggaagcacca ggagctgcga gccatgcaga tggagctgca gagccctgag tacaagctga
 541 gcaagctccg cacctcgacc atcatgacgg actacaaccc caactactgc tttgctggca
 601 agacctctc catcagtgac ctgaaggagg tgccgcggaa aaacatcacc ctcatcggg
 661 gtctgggcca tggcgctctt ggggaggtgt atgaaggcca ggtgtccgga atgcccacg
 721 acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct gaacaggacg
 781 aactggattt cctcatggaa gccctgatca toagcaaat caaccaccag aacattgttc
 841 gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag ctcatggcgg
 901 ggggagacct caagtctctc ctccgagaga cccgcccctg cccgagccag cctcctccc
 961 tggccatgct ggacctctc cactgtgctc gggacattgc ctgtggctgt cagtatttgg
 1021 aggaanaacca cttcatccac cgagacattg ctgcagaaaa ctgcctcttg acctgtccag

[0050]

1081 gccctggaag agtggccaag attggagact tgggatggc ccgagacatc tacagggcga
1141 gctactatag aaaggaggagc tgtgccatgc tggcagttaa gtggatgccc ccagaggcct
1201 tcatggaagg aatattcact tctaaaacag acacatggtc ctttggatgt ctgctatggg
1261 aaatcttttc tcttgatgat atgccatacc ccagcaaaag caaccaggaa gttctggagt
1321 ttgtcaccag tggaggccgg atggaccac ccaagaactg ccctgggcct gtataccgga
1381 taatgactca gtgctggcaa catcagcctg aagacaggcc caactttgcc atcattttgg
1441 agaggattga atactgcacc caggaccgg atgtaataca caccgctttg ccgatagaat
1501 atggtccact tgtggaagag gaagagaaaag tgcctgtgag gcccaaggac cctgaggggg
1561 ttcctcctct cctgtgtctc caacaggcaa aacgggagga ggagcgcagc ccagctgccc
1621 caccacctct gctaccacc tctctggca aggtgcaca atccagaaac aagccacca
1681 tctctgttgc agtccctaga gggccggccg tggaggggg acacgtgaat atggcattct
1741 ctcagtcaca cctcctctc gagttgcaca aggtccacgg atccagaaac aagccacca
1801 gcttgtggaa cccaacgtac ggtcctggt ttaacagaaa acccaccaaa aagaataatc
1861 ctatagcaaa gaaggagcca cagcagagg gtaacctggg gctggaggga agctgtactg
1921 tccacacctaa cgttgcaact gggagacttc cgggggcctc actgctccta gagcctctt
1981 cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc ccttgtggga
2041 atgtcaatta cggctaccag caacagggtc tgccttaga agcgcctact gccctgggag
2101 ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag cctgggcctc
2161 gagctcggtc gcacactcac tctcttctc tgggatccct aagaccgtgg aggagagaga
2221 ggcaatgctc ccttcacaaa ccagagacca aatgtcacgt tttgtttgt gccaacctat
2281 tttgaagtac caccaaaaaa gctgtatttt gaaaatgctt tagaaagggt ttgagcatgg
2341 gttcatccta tcttctcgaa agaagaaaa atcataaaaa tgagtataa atacaaggcc
2401 cagatgtggt tgcataaggt ttttatgcat gtttgttga tacttcctta tgcttcttt
2461 aaattgtgtg tgctctgctt caatgtagtc agaattagct gcttctatgt ttcagattg
2521 gggtcacata tgtttcctg ccttgttgat gtggacatga gccatttgag gggagaggga
2581 acggaataa aggagtatt tgtaatgact aaaa

TFGS-ALK 단백질 서열 (AAF42734.1; GI:7229261)*

1 mngqldlsgk liikaqlged irriphined itydelvlmm qrvfrgklls ndevtikykd
61 edgdlitifd ssdlsfaiqc srilkltlfv nggprpless qvkyllrreli elrnkvnrll
121 dsleppgepg pstnipenvy rrkhqelqam qmelqspeyk lsklrtstim tdylnpnycf
181 gktssisdik evprknitli rglghgafge vyegqvsgmp ndpsplqvav ktlpevcseq
241 deldflmeal iiskfnhqn vrcigvslqs lprfillelm aggdlsfllr etrprpsqps
301 slamlldllh ardiacgcqy leenhfihrd iaarnclltc pgpgrvakig dfgmardiyr
361 asyyrkqgca mlpvkwmppe afmegiftsk tdtwsfgvll weifslgym ypsksnqevl
421 efvtsggrmd pknpcgppy rimtgcwqhq pedrpnfaii lerieyctqd pdvintalpi
481 eygpiveeee kvprpkdpe gvppllvssq akreeerspa appplptss gkaakkpata
541 evsvrvprgp aveghvnma fsqsnppsel hkvhgsrnkp tslwnptygs wftkptkkn
601 npiakkephd rgnlglegsc tvppnvatgr lpgaslllep ssltanmkev plfirlrhfpc
661 gnvnygyqqq glpleaatap gaghyedtil ksknsmnpgg p

ATIC-ALK 서열 (inv(2)(p23;q35) 염색체 전위)*

CLTC-ALK 서열 (t(2;17)(p23;q23) 염색체 전위)*

MSN-ALK 핵산 서열 (AF295356.1; GI:14625823)

1 aactccgctg cctttgccgc caccatgccc aaaacgatca gtgtgcgtgt gaccaccatg
61 gatgcagagc tggagtttgc catccagccc aacaccaccg ggaagcagct atttgaccag
121 gtggtgaaaa ctattggcct gagggaaagt tggttctttg gtctgcagta ccaggacact
181 aaaggtttct ccacctggct gaaactcaat aagaaggtga ctgcccagga tgtgcggaag
241 gaaagcccc tgctctttaa gttccgtgcc aagttctacc ctgaggatgt gtccgaggaa
301 ttgattcagg acatcactca gcgcctgttc tttctgcaag tgaaagaggg cattctcaat
361 gatgatattt actgcccgc tgagaccgct gtgctgctgg cctcgtatgc tgtccagtct
421 aagtatggcg acttcaataa ggaagtgcac aagtctggct acctggccgg agacaagttg
481 ctcccgcaga gactcctgga acagcacaaa ctcaacaagg accagtggga ggagcggatc
541 caggtgtgag atgaggaaca ccgtggcatg ctcaaggagg atgctgtcct ggaatatctg
601 aagattgctc aagatctgga gatgtatggt gtgaactact tcagcatcaa gaacaagaaa

[0051]

661 ggctcagagc tgtggctggg ggtggatgcc ctgggtctca acatctatga gcagaatgac
721 agactaactc ccaagatagg cttccctgg agtgaaatca ggaacatctc tttcaatgat
781 aagaaatttg tcatcaagcc cattgacaaa aaagcccggt acttctgtctt ctatgctccc
841 cggctgcgga ttaacaagcg gatcttgccc ttgtgcatgg ggaacatga actatacatg
901 cgccgtcgca agcctgatac cattgaggtg cagcagatga aggcacagcg ccgggaggag
961 aagcaccaga agcagatgga cgtgctatg ctggaaaatg agaagaagaa gcgtgaaatg
1021 gcagagaagg agaaaagagaa gattgaacgg gagaaggagg agctgatgga gaggctgaa
1081 cagatcgagg aacagactaa gaaggctcag caagaactgg aagaacagac ccgtagggct
1141 ctggaacttg agcaggaacg gaagcgtgcc cagagcgagg ctgaaaagct ggccaaggag
1201 cgtcaagaag ctgaagaggc caaggaggcc ttgctgcagg cctcccgga ccagaaaaag
1261 actcaggaac agctggcctt ggaatggca gagctgacag ctcgaatctc ccagctggag
1321 atggcccgac agaagaagga gaggtaggct gtggagtggc agcagaagca ggagctgcaa
1381 gccatgcaga tggagctgca gagccctgag tacaagctga gcaagctccg cacctcgacc
1441 atcatgaccg actacaaccc caactactgc ttgtctggca agacctctc catcagtgac
1501 ctgaaggagg tgccgoggaa aaacatcacc ctcatcggg gtctgggcca tggcgccctt
1561 ggggagggtg atgaaggcca ggtgtccgga atgccaacg acccaag

MSN-ALK 단백질 서열 (AAK71522.1; GI:14625824)*

1 mpktisrvrt tmdaelefai qnnttqkqlf dqvvktiglr ewvffglqyg dtkgfstwlk
61 lnkkvtaqdv rkespllkf rakfypedvs eeliqditqr lflqvkqegi lnddiycppe
121 tavllasyav qskygdfnke vhsqylagd kllpqrveqg hklndqwee riqvweehhr
181 gmlredavle ylkiaqdlm ygvnyfsikn kkselwlgv dalglniyeq ndrtpkigf
241 pwseirnif ndkkfvikpi dkkapdfvfy aprlrinkri lalcmgnhel ymrrrkpdti
301 evqgmkaqar eekhgkqmer amlenekkr emaekekeki erekeelmer lkqieeqtkk
361 aqgeleeqtr raleleqerk raqseaecla kerqaeek eallqasrdq kktqeqlale
421 maeltarisq lemarqkes eavewqkqe lqamqmelqs peyklsklrt stimtdynpn
481 ycfagktssi sdlkevprkn itlirglghg afgevyegqv sgmpndp

TPM4-ALK 작은 변이체 핵산 서열 (AF362887.1; GI:14010353)

1 cgagaagttg agggagaaag gcgggcccgg gaacaggctg aggctgaggt ggccctcctg
61 aaccgtagga tccagctggt tgaagaagag ctggaccgtg ctcaggagcg tgcggagggtg
121 tctgaactaa aatgtgtgga cctggaagaa gaactcaaga atgttactaa caatctgaaa
181 tctctggagg ctgcatctga aaagtattct gaaaaggagg acaaatatga agaagaaatt
241 aaactcttgt ctgacaaact gaaagaggct gagaccctg ctgaatttgc agagagaacg
301 gttgcaaaac tggaagaaac aattgatgac ctggaagtgt acctccgga gcccaagag
361 ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa gctccgcacc
421 ctcgac

TPM4-ALK 작은 변이체 단백질 서열 (AAK51964.1; GI:14010354)

1 revegerarr eqaeaevasl nrriqlveee ldraqeraev selkcgdlle elknvtnnk
61 sleaasekys ekedkyeeei klldklkea etraefaert vaklektidd levylrkhqe
121 lqamqmelqs peyklsklrt ld

TPM4-ALK 큰 변이체 핵산 서열 (AF362886.1; GI:14010351)

1 ctggcagagt cccgttgccg agagatggat gagcagatta gactgatgga ccagaacctg
61 aagtgtctga gtgctgctga agaaaagtac tctcaaaaag aagataaata tgaggaagaa
121 atcaagattc ttactgataa actcaaggag gcagagaccc gtgctgaatt tgcagagaga
181 acggttgcaa aactggaaaa gacaattgat gacctggaag tgtaccgccg gaagaccag
241 gagctgcaag ccatgcagat ggagctgcag agccctgagt acaagctgag caagctccgc
301 acctcgac

TPM4-ALK 큰 변이체 단백질 서열 (AAK51963.1; GI:14010352)

1 laesrcremd eqirlmdqnl kclsaaeeky sqkedkyeee ikiltdklke aetraefaer
61 tvaklektid dlevyrkhq elqamqmelq speyklsklr tst

MYH9-ALK 서열 (t(2;22)(p23;q11.2) 염색체 전위)*

RANBP2-ALK 서열 (t(2;2)(p23;q13) 또는 inv(2)(p23;q11-13)

염색체 전위)*

ALO17-ALK 서열 (t(2;17)(p23;q25) 염색체 전위)*

CARS-ALK 서열 (t(2;11;2)(p23;p15;q31) 염색체 전위)*

*** MSN-ALK 및 MYH-9 를 제외하고는, 모든 융합 단백질은 최종 563 개의
아미노산의 ALK 를 함유한다. MSN-ALK 및 MYH9 는 각각 최종 567 개 및
566 개의 아미노산을 함유한다.**

"ALK 돌연변이"는 일반적으로 참조 역형성림프종 키나제 서열에 비한 핵산 및/또는 아미노산 서열에서의 변경을 의미한다. 그러나, 일부 구체예에서, "ALK 돌연변이"는 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료에 대한 반응이 예측되는 특정 역형성림프종 키나제 돌연변이를 의미할 수 있다. 예를 들어, 야생형

ALK 단백질(NP_004295)의 위치 1156(C1156)의 시스테인 아미노산 및/또는 위치 1196(L1196)의 루신 아미노산의 상이한 아미노산으로의 돌연변이는 ALK 억제제에 대한 내성을 부여하는 것으로 본원에 기재된다. 한 구체예에서, C1156 위치는 티로신 아미노산을 포함하고/하거나 L1196 위치는 메티오닌 아미노산을 포함한다. 당업자는 또한 야생형 ALK 단백질의 "C1156" 및 "L1196" 돌연변이에 해당하는 아미노산 위치는 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료에 대한 반응의 예측 값에 영향을 미치지 않으면서 다양한 참조 서열(예를 들어, ALK 동족체, ALK 융합 단백질 등)에 비해 상이한 수를 가질 것임을 인지할 것이다. 당업자는 특정 단백질의 아미노산 서열과 유전부호(하기 제시됨)에 의해 정의되는 바와 같은 상기 단백질을 코딩할 수 있는 뉴클레오티드 서열 사이에 공지되고 명확한 상응이 존재하는 것을 추가로 인지할 것이다. 마찬가지로, 특정 핵산의 뉴클레오티드 서열과 유전부호에 의해 정의되는 상기 핵산에 의해 엔코딩되는 아미노산 서열 사이에 공지되고 명확한 상응이 존재한다.

[0055]	유전부호	
[0056]	알라닌(Ala, A)	GCA, GCC, GCG, GCT
[0057]	아르기닌(Arg, R)	AGA, ACG, CGA, CGC, CGG, CGT
[0058]	아스파라긴(Asn, N)	AAC, AAT
[0059]	아스파르트산(Asp, D)	GAC, GAT
[0060]	시스테인(Cys, C)	TGC, TGT
[0061]	글루탐산(Glu, E)	GAA, GAG
[0062]	글루타민(Gln, Q)	CAA, CAG
[0063]	글리신(Gly, G)	GGA, GGC, GGG, GGT
[0064]	히스티딘(His, H)	CAC, CAT
[0065]	이소루신(Ile, I)	ATA, ATC, ATT
[0066]	루신(Leu, L)	CTA, CTC, CTG, CTT, TTA, TTG
[0067]	리신(Lys, K)	AAA, AAG
[0068]	메티오닌(Met, M)	ATG
[0069]	페닐알라닌(Phe, F)	TTC, TTT
[0070]	프롤린(Pro, P)	CCA, CCC, CCG, CCT
[0071]	세린(Ser, S)	AGC, AGT, TCA, TCC, TCG, TCT
[0072]	트레오닌(Thr, T)	ACA, ACC, ACG, ACT
[0073]	트립토판(Trp, W)	TGG
[0074]	티로신(Tyr, Y)	TAC, TAT
[0075]	발린(Val, V)	GTA, GTC, GTG, GTT
[0076]	종료 신호(종지)	TAA, TAG, TGA
[0077]	유전부호의 중요하고 널리 공지된 특징은 이의 중복성이며, 이에 의해, 단백질을 만드는데 이용되는 아미노산 대부분에 대해 하나 이상의 코딩 뉴클레오티드 트리플렛이 이용될 수 있다(예를 들어, 상기 예시됨). 따라서, 다수의 상이한 뉴클레오티드 서열이 제공된 아미노산 서열을 코딩할 수 있다. 이러한 뉴클레오티드 서열은 기능적으로 동등한 것으로 간주되는데, 이는 이들이 모든 유기체에서 동일한 아미노산 서열을 생성시키기 때문이다(그러나, 특정 유기체는 다른 서열을 번역하는 것보다 일부 서열을 더 효과적으로 번역할 수 있음). 또한, 가끔, 퓨린 또는 피리미딘의 메틸화된 변이체가 제공된 뉴클레오티드 서열에서 발견될 수 있다. 이러한 메틸화는 트리뉴클레오티드 코돈과 상응하는 아미노산 사이의 코딩 관계에 영향을 미치지 않는다. 또한 당업자는 관련된 코돈 도표를 기초로 하여 엔코딩되는 아미노산을 특이적으로 변경시키기 위해 특정 코돈의 뉴클레오티드를	

돌연변이화시키는 방법을 이해할 것이다. 예를 들어, Cys-1156에 대한 코돈은 "TGC"이고, Tyr에 대한 코돈은 "TAT" 또는 "TAC"일 수 있다. 따라서, 코돈의 위치 2에서의 단일 뉴클레오타이드 G에서 A로의 치환은 시스테인이 아닌 티로신을 엔코딩할 것이다. 당업자는 다른 돌연변이를 디자인하기 위해 유사한 조작을 수행할 수 있다.

[0078] "결합 화합물"은 결합 조성물, 예를 들어, 소분자, 항체, 펩티드, 펩티드 또는 비-펩티드 리간드, 단백질, 올리고뉴클레오타이드, 올리고뉴클레오타이드 유사체, 예를 들어, 펩티드 핵산, 핵틴, 또는 표적 단백질 또는 분자에 특이적으로 결합할 수 있는 임의의 다른 분자 존재물 또는 관심 분석물과의 안정적 복합체 형성, 예를 들어, 단백질의 복합체를 의미한다.

[0079] "결합 모이어티"는 분석물에 특이적으로 결합할 수 있는 분자 태그가 직접 또는 간접적으로 부착될 수 있는 임의의 분자를 의미한다. 결합 모이어티는 약 1000 달톤 이하의 분자량을 갖고, 수소, 탄소, 산소, 질소, 황 및 인으로 구성된 군으로부터 선택되는 원자를 함유하는 항체, 항체 결합 조성물, 펩티드, 단백질, 핵산 및 유기 분자를 포함하나, 이에 제한되지는 않는다.

[0080] "바이오마커" 또는 "마커"는 변경될 수 있는 유전자, mRNA, 또는 단백질이며, 상기 변경은 암과 관련된다. 변경은 정상 또는 건강한 조직 또는 세포(예를 들어, 대조군)에서의 양, 구조, 및/또는 활성과 비교하는 경우의 암 조직 또는 암 세포에서의 양, 구조, 및/또는 활성일 수 있고, 이는 질병 상태, 예를 들어, 암과 관련된다. 예를 들어, 암과 관련되거나, 항암 요법에 대한 반응을 예측하는 본 발명의 마커는 정상 건강한 조직 또는 세포에 비해 암 조직 또는 암 세포에서의 변경된 뉴클레오타이드 서열, 아미노산 서열, 염색체 전위, 염색체내 역위, 카피수, 발현 수준, 단백질 수준, 단백질 활성, 또는 메틸화 상태를 가질 수 있다. 또한, "마커"는 질병 상태, 예를 들어, 암과 관련된 조직 또는 세포에 존재하는 경우, 예를 들어, 치환, 결실 또는 삽입에 의해 구조가 변경된, 예를 들어, 돌연변이된(돌연변이를 함유함) 분자, 예를 들어, 뉴클레오타이드 또는 아미노산 수준에서 야생형 서열과 상이한 분자를 포함한다.

[0081] 용어 "암" 또는 "종양"은 암-유발 세포의 전형적인 특징, 예를 들어, 조절되지 않는 증식, 불사, 전이 잠재성, 신속한 성장 및 증식 속도, 및 일정한 특징적인 형태 특징을 갖는 세포의 존재를 의미한다. 암 세포는 종종 종양의 형태로 존재하지만, 이러한 세포는 동물 내에서 단독으로 존재할 수 있거나, 비-종양형성 암 세포, 예를 들어, 백혈병 세포일 수 있다. 본원에서 사용되는 용어 "암"은 전암성 뿐만 아니라 악성 암을 포함한다. 암은 B 세포 암, 예를 들어, 다발골수종, 발덴스트롬 마크로글로블린혈증, 중쇄 질병, 예를 들어, 알파 쇄 질병, 감마 쇄 질병, 무 쇄 질병, 양성 모노클로날 감마글로블린병증, 면역세포 아밀로이드증, 흑색종, 유방암, 폐암(예를 들어, 비소세포폐암종 또는 NSCLC), 기관지암, 결장직장암, 전립선암, 췌장암, 위암, 난소암, 방광암, 뇌 또는 중추신경계 암, 말초신경계 암, 식도암, 자궁경부암, 자궁암 또는 자궁내막암, 구강 또는 인두암, 간암, 신장암, 고환암, 담관암, 소장 또는 충수암, 침샘암, 갑상샘암, 부신암, 골육종, 연골육종, 혈액 조직암, 샘암종, 염증성 근섬유모세포 종양, 위장관기질종양(GIST), 결장암, 다발골수종(MM), 골수형성이상 증후군(MDS), 골수증식질환(MPD), 급성 림프구성 백혈병(ALL), 급성 골수구성 백혈병(AML), 만성 골수구성 백혈병(CML), 만성 림프구성 백혈병(CLL), 진성적혈구증가증, 호지킨 림프종, 비-호지킨 림프종(NHL), 연성-조직 육종, 섬유육종, 점액육종, 지방육종, 골육종, 척삭종, 혈관육종, 내피육종, 림프관육종, 림프관내피육종, 윤활막종, 중피종, 유잉종양, 평활근육종, 횡문근육종, 편평세포 암종, 기저세포 암종, 샘암종, 땀샘 암종, 피지샘 암종, 유두암종, 유두모양샘암종, 낭샘암종, 수질암종, 기관지원성 암종, 신세포 암종, 간암, 담관 암종, 융모막암종, 고환종, 배아암종, 윌름즈 종양, 방광암종, 상피암종, 신경아교종, 별아교세포종, 수모세포종, 두개인두종, 뇌실막종, 송과체종, 혈관모세포종, 청신경종, 희소돌기아교세포종, 수막종, 신경모세포종, 망막모세포종, 소포림프종, 광범위큰B세포림프종, 외투세포 림프종, 간세포 암종, 갑상샘암, 위암, 두경부암, 소세포암, 본태성 혈소판 증가증, 원인불명골수화생, 과다호산구증후군, 전신비만세포증, 가족성 과다호산구증가증, 만성 호산구백혈병, 신경내분비암, 카르시노이드 종양 등을 포함하나, 이에 제한되지는 않는다.

[0082] "화학요법제"는 질환, 예를 들어, 암을 치료하는데 사용되는 화학 물질, 예를 들어, 세포독성제 또는 세포증식 억제제를 의미한다.

[0083] "상보적"은 2개의 핵산 가닥의 영역 사이 또는 동일한 핵산 가닥의 2개의 영역 사이의 서열 상보성의 광범위한 개념을 의미한다. 첫번째 핵산 영역의 아데닌 잔기는 잔기가 티민 또는 우라실인 경우에 첫번째 영역에 역평행한 두번째 핵산 영역의 잔기와 함께 특이적 수소결합("염기쌍 형성(base pairing)")을 형성할 수 있는 것으로 공지되어 있다. 유사하게, 첫번째 핵산 가닥의 사이토신 잔기는 잔기가 구아닌인 경우 첫번째 가닥에 역평행한 두번째 핵산 가닥의 잔기와 염기쌍을 형성할 수 있는 것이 공지되어 있다. 핵산의 첫번째 영역은, 두 영역이 역평행 방식으로 배열되는 경우, 첫번째 영역의 적어도 하나의 뉴클레오타이드 잔기가 두번째 영역의 잔기와 염기

쌍을 형성할 수 있으면, 동일하거나 상이한 핵산의 두번째 영역에 상보적이다. 특정 구체예에서, 첫번째 영역은 첫번째 부분을 포함하고, 두번째 영역은 두번째 부분을 포함하며, 이에 의해, 첫번째 및 두번째 부분이 역평행 방식으로 배열되는 경우, 첫번째 부분의 뉴클레오티드 잔기의 적어도 약 50%, 적어도 약 75%, 적어도 약 90%, 또는 적어도 약 95%는 두번째 부분의 뉴클레오티드 잔기와 염기쌍을 형성할 수 있다. 다른 구체예에서, 첫번째 부분의 모든 뉴클레오티드 잔기는 두번째 부분의 뉴클레오티드 잔기와 염기쌍을 형성할 수 있다.

[0084] "유전자의 카피수" 또는 "마커의 카피수"는 특정 유전자 생성물을 엔코딩하는 세포내의 DNA 서열의 수를 의미한다. 일반적으로, 제공된 유전자에 대해, 포유동물은 각각의 유전자의 2개의 카피를 갖는다. 그러나, 카피수는 유전자 증폭 또는 유전자 복사에 의해 증가되거나, 결실에 의해 감소될 수 있다.

[0085] 마커는, 이러한 마커가 기질과 공유적 또는 비공유적으로 회합하여, 마커의 많은 분획이 기질로부터 해리됨이 없이 기질이 유체(예를 들어, 표준 염수 시트레이트, pH 7.4)로 행구어질 수 있는 경우에 기질에 "고정"된다.

[0086] 본원에서 사용되는 "위험비(Hazard ratio)"는 상대적 위험에 대한 평가를 발생시키는데 사용되는 통계 방법을 의미한다. "위험비"는 한 그룹의 예측 위험 대 또 다른 그룹의 예측 위험 사이의 비이다. 예를 들어, ALK 억제제로 치료되지 않은 환자 집단에 비해 ALK 억제제로 처리된 환자 집단은, ALK 억제제가, 특히 ALK 돌연변이 상태와 관련하여 질병의 재발이 멀어지도록 하는 시간을 증가시키는데 효과적인지 아닌지의 여부에 대해 평가될 수 있다. 예를 들어, 본원에 기재된 바와 같은 양성 조직 내에 ALK 돌연변이를 갖는 환자를 치료하는 것은 양성 조직 내에 상기 ALK 돌연변이를 갖지 않는 피검체에 비해 ALK 억제제로부터의 치료 이익을 증가시킨다.

[0087] 본원에서 사용되는 "ALK 억제 작용제" 또는 "ALK 억제제"는 ALK의 생물학적 활성을 억제할 수 있는 화합물을 의미한다. 생물학적 활성은 또한 본 출원에 기재된 바와 같은 환자 반응을 포함할 수 있다. 예시적 ALK 억제제는 PF-02341066, PDD, 2-메틸-11-(2-메틸프로필)-4-옥소-4,5,6,11,12,13-헥사하이드로-2H-인다졸로[5,4- α]피롤로[3,4-c]카르바졸-8-일[4-(디메틸아미노)벤질]카르바메이트, (1S,2S,3R,4R)-3-({5-클로로-2-[(1-에틸-2,3,4,5-테트라하이드로-6-메톡시-2-옥소-1H-1-벤즈아제핀-7-일)아미노]-4-피리미디닐}아미노)바이시클로[2.2.1]헵트-5-엔-2-카르복사미드, 및 NVP-TAE684를 포함하나, 이에 제한되지는 않는다(예를 들어, 참조로서 본원에 포함되는 문헌[PNAS 104:270-275, 2007; Choi, Y.L. et al. (2008) *Cancer Res.* 68:4971-2976; 및 *Biochemistry* 48:3600-3609, 2009] 참조).

[0088] 본원에서 상호교환적으로 사용되는 용어 "상동성" 또는 "동일성"은 2개의 폴리뉴클레오티드 서열 사이 또는 2개의 폴리펩티드 서열 사이의 서열 유사성을 의미하며, 동일성이 더욱 엄격한 비교이다. 구 "동일성 또는 상동성 퍼센트" 및 "동일성 또는 상동성 %"는 2개 이상의 폴리뉴클레오티드 서열 또는 2개 이상의 폴리펩티드 서열의 비교에서 발견되는 서열 유사성의 백분율을 의미한다. "서열 유사성"은 2개 이상의 폴리뉴클레오티드 서열 사이의 염기쌍 서열에서의 유사성 퍼센트(임의의 적합한 방법에 의해 결정됨)를 의미한다. 2개 이상의 서열은 0 내지 100%의 유사성 사이 어디일 수 있거나, 0 내지 100% 사이의 어떠한 정수 값일 수 있다. 동일성 또는 유사성은 비교 목적을 위해 정렬될 수 있는 각각의 서열 내의 위치를 비교함으로써 결정될 수 있다. 비교된 서열 내의 위치가 동일한 뉴클레오티드 염기 또는 아미노산이 차지하는 경우, 분자는 상기 위치에서 동일하다. 폴리뉴클레오티드 서열 사이의 유사성 또는 동일성의 정도는 폴리뉴클레오티드 서열에 의해 공유되는 위치에서 동일하거나 매치되는 뉴클레오티드의 수의 함수이다. 폴리펩티드 서열의 동일성의 정도는 폴리펩티드 서열에 의해 공유되는 위치에서 동일한 아미노산의 수의 함수이다. 폴리펩티드 서열의 상동성 또는 유사성의 정도는 폴리펩티드 서열에 의해 공유되는 위치에서 아미노산의 수의 함수이다. 본원에서 사용되는 용어 "실질적인 상동성"은 적어도 50%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95% 또는 이를 초과하는 상동성을 의미한다.

[0089] 암은, 암의 적어도 하나의 증상이 경감되거나, 종결되거나, 둔화되거나, 예방되는 경우에 "억제"된다. 본원에서 사용되는 용어 암은 또한, 암의 재발 또는 전이가 감소되거나, 둔화되거나, 지연되거나, 예방되는 경우에 "억제"된다.

[0090] "마커 핵산" 또는 "바이오마커 핵산"은 본 발명의 마커에 의해 엔코딩되거나 이에 상응하는 핵산(예를 들어, DNA, mRNA, cDNA)이다. 예를 들어, 이러한 마커 핵산 분자는 표 1에 나열된 핵산 서열 중 임의의 핵산 서열의 전체 또는 부분적 서열, 또는 이러한 서열의 상보적 또는 하이브리드화 단편을 포함하는 DNA(예를 들어, 유전체 DNA 및 cDNA)를 포함한다. 마커 핵산 분자는 또한 표 1에 나열된 핵산 서열 중 임의의 핵산 서열의 전체 또는 부분적 서열, 또는 이러한 서열의 상보체를 포함하는 RNA를 포함하며, 여기서 모든 티미딘 잔기는 유리딘 잔기로 대체된다. "마커 단백질"은 본 발명의 마커에 의해 엔코딩되거나 이에 상응하는 단백질이다. 마커 단백질은 표 1에 나열된 서열 중 임의의 서열 또는 이의 단편에 의해 엔코딩되는 단백질의 전체 또는 부분적 서열을

포함한다. 용어 "단백질" 및 "폴리펩티드"는 본원에서 상호교환적으로 사용된다.

- [0091] 마커의 "정상" 카피수 또는 마커 발현의 "정상" 수준은 암에 걸리지 않은 피검체, 예를 들어, 인간으로부터의 생물학적 샘플, 예를 들어, 가래, 기관지폐포 세척액(bronchoalveolar lavage), 흉막삼출액(pleural effusion), 조직, 전혈, 혈청, 혈장, 구강 채취물(buccal scrape), 타액, 뇌척수액, 소변, 대변, 및 골수 중의 마커의 발현 수준 또는 카피수이다.
- [0092] ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 "과다발현" 또는 "발현, 카피수, 및/또는 활성의 현저히 더 높은 수준"은 발현 또는 카피수를 평가하는데 사용되는 검정의 표준 오차보다 큰 시험 샘플에서의 발현 수준, 카피수, 및/또는 활성을 의미하며, 이는 대조군 샘플(예를 들어, 암에 걸리지 않은 건강한 피검체로부터의 샘플) 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 발현 수준 또는 카피수, 또는 여러 대조군 샘플 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 평균 발현 수준 또는 카피수의 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배 또는 적어도 10배 또는 이를 초과할 수 있다.
- [0093] 용어 "프로브"는 특별한 소기의 표적 분자, 예를 들어, 본 발명의 마커에 선택적으로 결합할 수 있는 임의의 분자를 의미한다. 프로브는 당업자에 의해 합성되거나, 적절한 생물학적 제조물로부터 유래될 수 있다. 표적 분자의 검출 목적상, 프로브는 본원에 기재된 바와 같이 표지되도록 특이적으로 디자인될 수 있다. 프로브로서 이용될 수 있는 분자의 예는 RNA, DNA, 단백질, 항체, 및 유기 모노머를 포함하나, 이에 제한되지는 않는다.
- [0094] "RECIST"는 "고형 종양의 반응 평가 기준(고형 종양 그룹의 반응 평가 기준)"을 의미하는 약어를 의미하며, 이는 치료 동안 암환자가 개선("반응")되거나, 동일하게 유지("안정적")되거나, 악화("진행")되는 경우를 정의하는 공개된 규칙의 세트이다. RECIST 기준에 의해 정의된 반응은, 예를 들어, 문헌[Journal of the National Cancer Institute, Vol. 92, No. 3, Feb. 2, 2000]에서 공개되었고, RECIST 기준은 다른 유사한 공개된 정의 및 규칙 세트를 포함할 수 있다. 당업자는 본원에서 사용되는 바와 같은 RECIST 기준에 따른 정의, 예를 들어, "PR", "CR", "SD" 및 "PD"를 이해할 것이다.
- [0095] 본원에서 사용되는 바와 같은 치료에 대한 "반응성" 내지 "반응하다", 및 이러한 동사의 다른 형태는 ALK 억제제를 이용한 치료에 대한 피검체의 반응을 의미한다. 예로서, 피검체는, 이러한 피검체에서의 종양의 성장이 약 10%, 약 20%, 약 30%, 약 40%, 약 50%, 약 60%, 약 70%, 약 80%, 약 90% 또는 이 초과로 지연되는 경우에 ALK 억제제를 이용한 치료에 대해 반응한다. 또 다른 예에서, 피검체는, 이러한 피검체에서의 종양이 임의의 적절한 척도, 예를 들어, 질량 또는 부피에 의해 결정시 약 5%, 약 10%, 약 20%, 약 30%, 약 40%, 약 50% 또는 이 초과로 수축하는 경우에 ALK 억제제를 이용한 치료에 대해 반응한다. 또 다른 예에서, 피검체는, 이러한 피검체가 치료가 적용되지 않은 경우에 예측된 기대여명의 약 5%, 약 10%, 약 20%, 약 30%, 약 40%, 약 50% 또는 이 초과로 연장된 기대여명을 경험하는 경우에 ALK 억제제를 이용한 치료에 대해 반응한다. 또 다른 예에서, 피검체는, 이러한 피검체가 증가된 질병에 구애되지 않는 생존, 전반적 생존 또는 증가된 진행 기간을 갖는 경우에 ALK 억제제를 이용한 치료에 대해 반응한다. 상기 나열된 바와 같은 RECIST 기준을 포함하는, 환자가 치료에 반응하는지 결정하기 위한 여러 방법이 이용될 수 있다.
- [0096] "샘플", "조직 샘플", "환자 샘플", "환자 세포 또는 조직 샘플" 또는 "시료"는 각각 피검체 또는 환자의 조직으로부터 수득된 유사한 세포의 집합물을 의미한다. 조직 샘플의 공급원은 신선한, 동결된 및/또는 보존된 기관, 조직 샘플, 생검, 또는 흡인물로부터의 것과 같은 고형 조직; 혈액 또는 임의의 혈액 성분; 체액, 예를 들어, 뇌척수액, 양수, 복수 또는 간질액; 또는 피검체의 임신 또는 발단에서의 임의의 시점으로부터 세포일 수 있다. 조직 샘플은 자연 조직과 자연적으로 혼합되지 않는 화합물, 예를 들어, 보존제, 항응고제, 완충제, 고정제, 영양소, 항생제 등을 함유할 수 있다.
- [0097] 마커의 양, 예를 들어, 피검체에서의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 발현 또는 카피수는, 이러한 마커의 양이 양을 평가하기 위해 이용되는 검정의 표준 오차보다 큰 양에 의해 정상 수준보다 각각 크거나 작거나, 상기 양의 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배, 적어도 10배 또는 이를 초과하여 크거나 작은 경우에, 마커의 정상 양보다 "현저히" 크거나 작다. 대안적으로, 피검체에서의 마커의 양은, 이러한 양이 마커의 정상적인 양보다 각각 적어도 2배, 적어도 약 3배, 적어도 약 10배, 또는 적어도 약 5배 크거나 작은 경우에 정상 양에 비해 "현저히" 높거나 낮은 것으로 간주될 수 있다.
- [0098] 본원에서 사용되는 용어 "유의한 사건"은 당업자에 의해 결정시 중요한 환자의 질병에서의 사건을 의미한다. 유의한 사건의 예는, 일차 진단, 사망, 재발, 환자의 질병이 전이성인 것의 결정, 환자의 질병의 재발 또는 상

기 언급된 단계 중 어느 하나로부터 다른 단계까지의 환자의 질병의 진행을 포함하나, 이에 제한되지는 않는다. 유의한 사건은 당업자에 의해 결정시 OS, TTP를 평가하는데 사용되고/되거나 RECIST 또는 다른 반응 기준을 이용한 임의의 중요한 사건일 수 있다.

[0099] 본원에서 사용되는 용어 "피검체" 및 "환자"는 상호교환적으로 사용된다. 본원에서 사용되는 용어 "피검체" 및 "피검체들"은 동물, 예를 들어, 포유동물, 예를 들어, 비-영장류(예를 들어, 소, 돼지, 말, 당나귀, 염소, 낙타, 고양이, 개, 기니아 피그, 래트, 마우스 양) 및 영장류(예를 들어, 원숭이, 예를 들어, 시노몰구스(cynomolgous) 원숭이, 고릴라, 침팬지 및 인간)를 의미한다.

[0100] 본원에서 사용되는 용어 "시간 경과"는 최초 사건과 이후의 사건 사이의 시간의 양을 의미한다. 예를 들어, 환자의 암과 관련하여, 시간 경과는 환자의 질병과 관련될 수 있고, 이는 질병의 경과에서 유의한 사건을 평가함으로써 측정될 수 있으며, 여기서, 예를 들어, 첫번째 사건은 진단일 수 있고, 이후의 사건은 전이일 수 있다.

[0101] "질병 진행 기간(Time to progression)" 또는 "TTP"는 치료 시작으로부터 진행 또는 암 또는 감시(censor)까지 측정된 시간을 의미한다. 감시는 연구 종료 또는 치료에서의 변화로부터 발생할 수 있다. 질병 진행 기간은 또한, 예를 들어, 질병 진행 기간이 특정 시간에 걸쳐 진행이 없을 가능성을 나타낼 수 있으며, 상기 시간이 치료 시작과 진행 또는 감시 사이의 시간인 카플란-마이어(Kaplan-Meier) 플롯에서와 같이 가능성으로 표시될 수 있다.

[0102] "전사된 폴리뉴클레오티드"는 본 발명의 마커의 전사 및 존재시 전사체의 일반적인 전사후 처리(예를 들어, 스플라이싱), 및 전사체의 역전사에 의해 제조된 성숙 RNA의 전부 또는 일부에 상보적이거나 상동성인 폴리뉴클레오티드(예를 들어, RNA, cDNA, 또는 RNA 또는 cDNA 중 하나의 유사체)이다.

[0103] "치료하다", "치료" 및 이러한 용어의 다른 형태는 암의 성장을 방해하거나, 암의 중량 또는 부피를 수축시키거나, 피검체의 예상 생존 기간 및/또는 종양의 진행 기간을 연장시키는 것 등을 위한 ALK 억제제의 투여를 의미한다.

[0104] ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 "저발현(underexpression)" 또는 "발현, 카피수, 및/또는 활성의 현저하게 낮은 수준"은, 예를 들어, 대조군 샘플(예를 들어, 암에 걸리지 않은 건강한 피검체로부터의 샘플) 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 발현 수준, 카피수, 및/또는 활성, 또는 여러 대조군 샘플 내의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)의 평균 발현 수준, 카피수, 및/또는 활성보다 적어도 2배, 적어도 3배, 적어도 4배, 적어도 5배, 또는 적어도 10배 또는 이 초과로 적은, 발현 또는 카피수를 평가하는데 이용되는 검정의 표준 오차보다 큰 시험 샘플 내의 발현 수준 또는 카피수를 의미한다.

[0105] II. 본 발명의 예시적 방법

[0106] 본 발명은, 적어도 부분적으로, 암의 치료에서 ALK 억제제의 효능 예측과 관련된, 예를 들어, ALK 돌연변이를 포함하는 유전체의 특정 영역의 확인을 기초로 한다. ALK 유전자 발현 서열의 분석은 폴리펩티드가 ALK 억제제를 이용한 치료에 대해 적어도 부분적으로 내성이 되도록 할 수 있는 ALK 폴리펩티드(예를 들어, EML4-ALK 폴리펩티드를 포함하는 표 1에 나열된 바이오마커)에 대한 신규한 돌연변이의 확인을 야기시켰다. 따라서, 본원에 기재된 다양한 방법에서 상기 바이오마커 중 하나 이상의 존재 및/또는 부재는 본 발명의 범위 내에 속한다.

[0107] 일부 구체예에서, 본 발명의 방법은 피검체에서의 암의 진행을 모니터링하는데 사용될 수 있으며, 여기서 피검체 내의 샘플에 암의 진행 동안, 예를 들어, 첫번째 시점 및 이후의 시점에서 본원에서 확인된 하나 이상의 ALK 돌연변이(예를 들어, EML4-ALK 돌연변이)가 존재하는 경우, 암은 ALK 억제제-매개 치료에 반응할 가능성이 적거나, 그 반대이다. 또 다른 구체예에서, 첫번째 시점과 이후의 시점 사이에서, 피검체는 치료, 예를 들어, 화학요법, 방사선 요법, 수술, 또는 암을 억제하는데 유용한 임의의 다른 치료 방법을 받거나, 치료 완료되거나, 회복된다.

[0108] 본원에 추가로 기재된 바와 같이, 본 발명의 하나 이상의 바이오마커(예를 들어, EML4-ALK 돌연변이를 포함하는 ALK 돌연변이)는 참조 서열, 예를 들어, SEQ ID NO:1과 비교하는 경우 유전체(예를 들어, 점라인 및/또는 체세포) 내의 존재에 의해 특별히 확인될 수 있다. 예를 들어, 본원에 기재된 방법은 표 1에 나열된 하나 이상의 돌연변이를 포함하는 DNA의 스트레치(stretch)의 표적 핵산 증폭 반응을 수행함으로써 본 발명의 바이오마커를 검출하는 단계, 및 하나 이상의 돌연변이의 존재에 대해 증폭된 표적 핵산을 분석하는 단계를 수반할 수 있다.

[0109] 핵산을 증폭시키는 다양한 방법은 당 분야에 공지되어 있고, 다음과 같다: 미국 특허 번호 4,683, 195호(참조로

서 포함됨), 미국 특허 번호 4,683,202호(참조로서 포함됨) 및 미국 특허 번호 4,800,159호(참조로서 포함됨)에 기재된, PCR(중합효소 연쇄반응), 및 특허 EP-B-0.569.272호에 개시된 바와 같은, 특허 1-단계(one-step) 포맷의 PCR의 RT-PCR 대체(역전사 PCR), 특허 출원 EP-A-0.201.184호에, 예를 들어, 기재된 바와 같은 LCR(리가제 연쇄반응), 국제 출원 WO-A-90/01069호(참조로서 포함됨)에, 예를 들어, 기재된 바와 같은 RCR(복구 연쇄반응), 특허 출원 WO-A-90/06995호(참조로서 포함됨)에, 예를 들어, 기재된 바와 같은 3SR(자기부양 염기서열 복제반응), 주형으로 이중 가닥 DNA를 이용하는 EP-B-0.397.269호 및 미국 특허 번호 5,466,586호(참조로서 포함됨)에, 예를 들어, 기재된 바와 같은 NASBA(핵산 서열-기반 증폭반응), 및 미국 특허 번호 5,399,491호(참조로서 포함됨)에 기재된 바와 같은 TMA(전사 매개 증폭반응).

[0110] 증폭된 생성물 내의 하나 이상의 돌연변이의 존재의 검출은 당 분야에 널리 공지된 다양한 방식, 예를 들어, DNA 시퀀싱 방법, 예를 들어, 생거 시퀀싱(Sanger sequencing) 및 딥 시퀀싱(deep sequencing), 제한 효소의 사용, 대립유전자 특이적 증폭, 펩티드 핵산(PNA)-매개 PCR, 구조 차이의 검출, 예를 들어, 단일가닥 구조 다형성(SSCP) 및 표지된 올리고뉴클레오타이드 프로브를 이용한 막 상에서의 검출 단계(도트 블롯(dot blot))를 갖는 변성 구배 겔 전기영동(DGGE) 검정, 미세역가 플레이트에서의 검출 단계를 갖는 검정, 예를 들어, 역 하이브리드화, 올리고뉴클레오타이드 라이게이션 검정(OLA, MLPA), 첫번째 뉴클레오타이드 변화(First Nucleotide Change, FNC) 기술, 가교 기술, 신속 주기 PCR 및 동시 형광 분석(예를 들어, 5' 뉴클레아제/Taqman), 및 PCR 후 질량분광법 또는 모세관 전기영동을 이용한 미니-시퀀싱으로 수행될 수 있다.

[0111] III. 예시적인 분리된 핵산 분자

[0112] 본 발명의 한 양태는 본 발명의 마커에 상응하는 폴리펩티드 또는 이러한 폴리펩티드의 일부를 엔코딩하는 핵산을 포함하는, 본 발명의 바이오마커에 상응하는 분리된 핵산 분자에 속한다. 본 발명의 핵산 분자는 본원에서 확인된 ALK 또는 ALK-관련 유전체(예를 들어, 점라인 및/또는 체세포) 영역에 존재하고/하거나 ALK 또는 ALK-관련(예를 들어, EML4-ALK) 폴리펩티드를 엔코딩하는 핵산 분자를 포함한다. 일부 구체예에서, 본 발명의 핵산 분자는 표 1에 제시된 핵산 서열 또는 이의 단편을 포함하거나, 이로 본질적으로 구성되거나, 이로 구성된다. 본 발명의 분리된 핵산 분자는 또한 본 발명의 마커에 상응하는 폴리펩티드를 엔코딩하는 핵산 분자, 및 이러한 핵산 분자의 단편을 포함하는 본 발명의 마커에 상응하는 핵산 분자를 확인하기 위한 하이브리드화 프로브로서 사용하기에 충분한 핵산 분자, 예를 들어, 핵산 분자의 증폭 또는 돌연변이에 대한 PCR 프라이머로 사용하기에 적합한 것을 포함한다. 본원에서 사용되는 용어 "핵산 분자"는 DNA 분자(예를 들어, cDNA 또는 유전체 DNA) 및 RNA 분자(예를 들어, mRNA) 및 뉴클레오타이드 유사체를 이용하여 생성된 DNA 또는 RNA의 유사체를 포함한다. 핵산 분자는 단일-가닥 또는 이중-가닥일 수 있고, 특정 구체예에서, 핵산 분자는 이중-가닥 DNA이다.

[0113] "분리된" 핵산 분자는 핵산 분자의 자연원에 존재하는 다른 핵산 분자로부터 분리된 것이다. 특정 구체예에서, "분리된" 핵산 분자는 핵산이 유래되는 유기체의 유전체 DNA 내의 핵산에 자연적으로 플랭킹(flancking)된 서열(예를 들어, 단백질-엔코딩 서열)(즉, 핵산의 5' 및 3' 말단에 위치한 서열)이 존재하지 않는다. 예를 들어, 다양한 구체예에서, 분리된 핵산 분자는 핵산이 유래되는 세포의 유전체 DNA 내의 핵산 분자에 자연적으로 플랭킹된 약 5 kB 미만, 약 4 kB 미만, 약 3 kB 미만, 약 2 kB 미만, 약 1 kB 미만, 약 0.5 kB 미만 또는 약 0.1 kB 미만의 뉴클레오타이드 서열을 함유할 수 있다. 또한, "분리된" 핵산 분자, 예를 들어, cDNA 분자는 재조합 기술에 의해 생성되는 경우 다른 세포 물질 또는 배지를 실질적으로 함유하지 않을 수 있거나, 화학적으로 합성되는 경우 화학 전구체 또는 다른 화합물을 실질적으로 함유하지 않을 수 있다.

[0114] 용어 "다른 세포 물질 또는 배지를 실질적으로 함유하지 않음"은 핵산 분자의 제조물을 포함하며, 여기서 상기 분자는 이러한 분자가 분리되는 세포의 세포 구성성분으로부터 분리되거나 재조합적으로 생성된다. 따라서, 세포 물질을 실질적으로 함유하지 않는 핵산 분자는 약 30% 미만, 약 20% 미만, 약 10% 미만, 또는 약 5% 미만(건조 중량 기준)의 다른 세포 물질 또는 배지를 갖는 핵산 분자의 제조물을 포함한다.

[0115] 본 발명의 핵산 분자, 예를 들어, 표 1에 나열된 ALK 유전자 돌연변이는 표준 분자생물학 기술 및 본원에 기재된 데이터베이스 기록의 서열 정보를 이용하여 분리될 수 있다. 상기 핵산 서열의 전부 또는 일부를 이용하여, 본 발명의 핵산 분자는 표준 하이브리드화 및 클로닝 기술을 이용하여 분리될 수 있다(예를 들어, 문헌 [Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989]에 기재됨).

[0116] 본 발명의 핵산 분자는 주형으로서 cDNA, mRNA, 또는 유전체 DNA(예를 들어, 점라인 및/또는 체세포 유전체 DNA) 및 표준 PCR 증폭 기술에 따른 적절한 올리고뉴클레오타이드 프라이머를 이용하여 증폭될 수 있다. 상기와 같이 증폭된 핵산 분자는 적절한 벡터로 클로닝될 수 있고, DNA 서열 분석에 의해 특정규명될 수 있다. 또한,

본 발명의 핵산 분자의 전부 또는 일부에 상응하는 올리고뉴클레오티드는 표준 합성 기술, 예를 들어, 자동화된 DNA 합성기를 이용하여 제조될 수 있다.

[0117] 또 다른 구체예에서, 본 발명의 분리된 핵산 분자는 본 발명의 마커에 상응하는 핵산의 뉴클레오티드 서열 또는 본 발명의 마커에 상응하는 단백질을 엔코딩하는 핵산의 뉴클레오티드 서열에 상보적인 뉴클레오티드 서열을 갖는 핵산 분자를 포함한다. 제공된 뉴클레오티드 서열에 상보적인 핵산 분자는 제공된 뉴클레오티드 서열에 하이브리드화됨으로써 안정적인 듀플렉스(duplex)를 형성할 수 있는 제공된 뉴클레오티드 서열에 충분히 상보적인 것이다.

[0118] 또한, 본 발명의 핵산 분자는 전장 핵산 서열이 본 발명의 마커를 포함하는 핵산 서열의 일부만을 포함하거나, 본 발명의 마커에 마커에 상응하는 폴리펩티드를 엔코딩하는 핵산 서열의 일부만을 포함할 수 있다. 이러한 핵산 분자는, 예를 들어, 프로브 또는 프라이머로 사용될 수 있다. 프로브/프라이머는 통상적으로 하나 이상의 실질적으로 정제된 올리고뉴클레오티드로 사용된다. 올리고뉴클레오티드는 통상적으로 엄격한 조건하에서 본 발명의 핵산의 적어도 약 7개, 적어도 약 8개, 적어도 약 9개, 적어도 약 10개, 적어도 약 11개, 적어도 약 12개, 적어도 약 13개, 적어도 약 14개, 적어도 약 15개, 적어도 약 16개, 적어도 약 17개, 적어도 약 18개, 적어도 약 19개, 적어도 약 20개, 적어도 약 21개, 적어도 약 22개, 적어도 약 23개, 적어도 약 24개, 적어도 약 25개, 적어도 약 26개, 적어도 약 27개, 적어도 약 28개, 적어도 약 29개, 적어도 약 30개, 적어도 약 31개, 적어도 약 32개, 적어도 약 33개, 적어도 약 34개, 적어도 약 35개, 적어도 약 36개, 적어도 약 37개, 적어도 약 38개, 적어도 약 39개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적어도 약 85개, 적어도 약 90개, 적어도 약 95개, 적어도 약 100개 또는 이 초과인 연속적 뉴클레오티드에 하이브리드화되는 뉴클레오티드 서열의 영역을 포함한다.

[0119] 본 발명의 핵산 분자의 서열을 기초로 한 프로브는 본 발명의 하나 이상의 마커에 상응하는 전사체 또는 유전체 서열을 검출하는데 사용될 수 있다. 프로브는 이에 부착되는 표지 기, 예를 들어, 방사성 동위원소, 형광 화합물, 효소, 또는 효소 보조인자를 포함한다. 이러한 프로브는, 예를 들어, mRNA 수준을 검출하는 것과 같이 피검체로부터의 세포의 샘플 내의 단백질을 엔코딩하는 핵산 분자의 수준을 측정하거나, 단백질을 엔코딩하는 유전자가 돌연변이되었거나 결실되었는지의 여부를 결정하는 것과 같이, 단백질을 잘못 발현하는 세포 또는 조직을 확인하기 위한 진단 시험 키트의 일부로 사용될 수 있다.

[0120] 본 발명은 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)과 실질적으로 상동성인 핵산 분자를 추가로 포함하며, 이들은 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5% 또는 이 초과로 상동성이다. 다른 구체예에서, 본 발명은 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)과 실질적으로 상동성인 핵산 분자를 추가로 포함하며, 이들은 단지 또는 단지 또는 적어도 1개, 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 11개, 적어도 12개, 적어도 13개, 적어도 14개, 적어도 15개, 적어도 16개, 적어도 17개, 적어도 18개, 적어도 19개, 적어도 20개, 적어도 21개, 적어도 22개, 적어도 23개, 적어도 24개, 적어도 25개, 적어도 26개, 적어도 27개, 적어도 28개, 적어도 29개, 적어도 30개, 적어도 31개, 적어도 32개, 적어도 33개, 적어도 34개, 적어도 35개, 적어도 36개, 적어도 37개, 적어도 38개, 적어도 39개, 적어도 40개, 적어도 45개, 적어도 50개, 적어도 55개, 적어도 60개, 적어도 65개, 적어도 70개, 적어도 75개, 적어도 80개, 적어도 85개, 적어도 90개, 적어도 95개, 적어도 100개의 뉴클레오티드 또는 이 사이의 임의의 범위로 상이하다.

[0121] 용어 "단일 뉴클레오티드 다형성"(SNP)은 대립유전자 서열 사이의 변화 부위에 존재하는 단일 뉴클레오티드에 의해 차지되는 다형성 부위를 의미한다. 상기 부위는 보통 대립유전자의 고도로 보존된 서열(예를 들어, 집단의 1/100 또는 1/1000 일원에서 상이한 서열)에 선행하거나 후속된다. SNP는 보통 다형성 부위에서의 하나의 뉴클레오티드의 또 다른 뉴클레오티드로의 치환으로 인해 발생한다. SNP는 또한 참조 대립유전자에 비한 뉴클레오티드의 결실 또는 뉴클레오티드의 삽입으로부터 발생할 수 있다. 통상적으로, 다형성 부위는 참조 염기가 아닌 염기에 의해 차지된다. 예를 들어, 참조 대립유전자가 다형성 부위에서 염기 "T"(티미딘)를 함유하는 경우, 변경된 대립유전자는 다형성 부위에서 "C"(시티딘), "G"(구아닌), 또는 "A"(아데닌)를 함유할 수 있다. SNP는 단백질-코딩 핵산 서열에서 발생할 수 있고, 이러한 경우 이들은 결함이 있거나 그렇지 않으면 변이체 단백질, 또는 유전병을 발생시킬 수 있다. 이러한 SNP는 유전자의 코딩 서열을 변경시킬 수 있고, 이에 따라 또 다른 아미노산("미스센스" SNP)을 특정할 수 있거나, SNP는 정지 코돈("논센스" SNP)를 도입시킬 수 있다. SNP

가 단백질의 아미노산 서열을 변경시키지 않는 경우, SNP는 "침묵(silent)"으로 언급된다. SNP는 또한 뉴클레오티드 서열의 비코딩 영역에서 발생할 수 있다. 이는, 예를 들어, 대안적 스플라이싱의 결과로서 결합이 있는 단백질 발현을 발생시킬 수 있거나, 이는 단백질의 기능에 영향을 미치지 않을 수 있다.

[0122] 또 다른 구체예에서, 본 발명의 분리된 핵산 분자는 적어도 7, 적어도 15, 적어도 20, 적어도 25, 적어도 30, 적어도 35, 적어도 40, 적어도 45, 적어도 50, 적어도 55, 적어도 60, 적어도 65, 적어도 70, 적어도 75, 적어도 80, 적어도 85, 적어도 90, 적어도 95, 적어도 100, 적어도 125, 적어도 150, 적어도 175, 적어도 200, 적어도 250, 적어도 300, 적어도 350, 적어도 400, 적어도 450, 적어도 550, 적어도 650, 적어도 700, 적어도 800, 적어도 900, 적어도 1000, 적어도 1200, 적어도 1400, 적어도 1600, 적어도 1800, 적어도 2000, 적어도 2200, 적어도 2400, 적어도 2600, 적어도 2800, 적어도 3000, 적어도 3500, 적어도 4000, 적어도 4500, 또는 이 초과 뉴클레오티드 길이이고, 이는 엄격한 조건하에서 본 발명의 마커에 상응하는 핵산 분자 또는 본 발명의 마커에 상응하는 단백질을 엔코딩하는 핵산 분자에 하이브리드화된다. 본원에서 사용되는 용어 "엄격한 조건하에서 하이브리드화되는"은 하이브리드화 및 세척을 위한 조건을 기재하기 위한 것으로, 이러한 조건하에서 서로 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 85% 동일한 뉴클레오티드 서열은 통상적으로 서로 하이브리드화된 채로 유지된다. 이러한 엄격한 조건은 당업자에게 공지되어 있고, 이는 문헌[*Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989)]의 섹션 6.3.1-6.3.6에서 발견될 수 있다. 엄격한 하이브리드화 조건의 또 다른 비제한적인 예는 45°C에서의 6X 염화나트륨/시트르산 나트륨(SSC) 중에서의 하이브리드화 후, 50-65°C에서의 0.2X SSC, 0.1% SDS에서의 1회 이상의 세척이다.

[0123] 본 발명은 또한 본 발명의 핵산 분자에 상보적인 적어도 하나의 영역을 갖는 분자 비콘(beacon) 핵산 분자를 포함하며, 상기 분자 비콘은 샘플 내의 본 발명의 핵산 분자의 존재를 측정하는데 유용하다. "분자 비콘" 핵산은 상보적 영역의 쌍을 포함하고, 형광단 및 이에 결합된 형광 켄처(quencher)를 갖는 핵산 분자이다. 형광단 및 켄처는 상기 배향에서 핵산의 상이한 부분과 결합되고, 상보적 영역이 서로 어닐링되는 경우, 형광단의 형광은 켄처에 의해 켄칭된다. 핵산 분자의 상보적 영역이 서로 어닐링되지 않는 경우, 형광단의 형광은 보다 덜한 정도로 켄칭된다. 분자 비콘 핵산 분자는, 예를 들어, 미국 특허 5,876,930호(참조로서 포함됨)에 기재되어 있다.

[0124] IV. 예시적인 분리된 단백질 및 항체

[0125] 본 발명의 한 양태는 본 발명의 개별적 마커에 상응하는 분리된 단백질, 및 이의 생물학적으로 활성인 부분에 속한다. 한 구체예에서, 마커에 상응하는 자연 폴리펩티드는 표준 단백질 정제 기술을 이용한 적절한 정제 계획에 의해 세포 또는 조직원으로부터 분리될 수 있다. 또 다른 구체예에서, 본 발명의 마커에 상응하는 폴리펩티드는 재조합 DNA 기술에 의해 생성된다. 재조합 발현에 대한 대안으로, 본 발명의 마커에 상응하는 폴리펩티드는 표준 펩티드 합성 기술을 이용하여 화학적으로 합성될 수 있다.

[0126] "분리된" 또는 "정제된" 단백질 또는 이의 생물학적 활성 부분은 단백질이 유래되는 세포 또는 조직원으로부터의 세포 물질 또는 다른 오염 단백질을 실질적으로 함유하지 않거나, 화학적으로 합성되는 경우 화학 전구체 또는 다른 화합물을 실질적으로 함유하지 않는다. 용어 "세포 물질을 실질적으로 함유하지 않는"은, 분리되거나 재조합적으로 생성되는 세포의 세포 구성요소로부터 단백질이 분리된 상기 단백질의 제조물을 포함한다. 따라서, 세포 물질을 실질적으로 함유하지 않는 단백질은 약 30% 미만, 약 20% 미만, 약 10% 미만, 또는 약 5% 미만(건조 중량)의 이중성 단백질(본원에서 "오염 단백질"로도 언급됨)을 갖는 단백질의 제조물을 포함한다. 단백질 또는 이의 생물학적 활성 단백질이 재조합적으로 생성되는 경우, 이는 배지를 실질적으로 함유하지 않을 수 있고, 즉, 배지는 단백질 제조물 부피의 약 20% 미만, 약 10% 미만, 또는 약 5% 미만이 존재한다. 단백질이 화학적 합성에 의해 생성되는 경우, 이는 화학 전구체 또는 다른 화합물을 실질적으로 함유하지 않을 수 있고, 즉, 이는 단백질의 합성과 관련된 화학 전구체 또는 다른 화합물로부터 분리된다. 따라서, 이러한 단백질의 제조물은 약 30% 미만, 약 20% 미만, 약 10% 미만, 약 5% 미만(건조 중량)의 관심 폴리펩티드가 아닌 화학 전구체 또는 화합물을 갖는다.

[0127] 본 발명의 마커에 상응하는 폴리펩티드의 생물학적 활성 부분은 전장 단백질보다 적은 아미노산을 포함하는 본 발명의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)에 상응하는 단백질의 아미노산 서열과 충분히 동일하거나 이로부터 유래된 아미노산 서열을 포함하는 폴리펩티드를 포함하고, 이는 상응하는 전장 단백질의 적어도 하나의 활성을 나타낸다. 통상적으로, 생물학적 활성 부분은 상응하는 단백질의 적어도 하나의 활성을 갖는 도메인 또는 모티프를 포함한다. 본 발명의 단백질의 생물학적 활성 부분은, 예를 들어, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65,

70, 75, 80, 85, 90, 95, 100 또는 이 초과인 아미노산 길이인 폴리펩티드일 수 있다. 또한, 단백질의 다른 영역이 결실된 다른 생물학적 활성 부분은 제조법 기술에 의해 제조될 수 있고, 본 발명의 폴리펩티드의 자연 형태의 기능적 활성 중 하나 이상에 대해 평가될 수 있다.

[0128] 특정 구체예에서, 폴리펩티드는 표 1에 나열된 핵산 분자에 의해 인코딩되는 단백질의 아미노산 서열을 갖는다. 다른 유용한 단백질은 상기 서열 중 하나와 실질적으로 동일(예를 들어, 적어도 60, 적어도 65, 적어도 70, 적어도 75, 적어도 80, 적어도 85, 적어도 86, 적어도 87, 적어도 88, 적어도 89, 적어도 90, 적어도 91, 적어도 92, 적어도 93, 적어도 94, 적어도 95, 적어도 96, 적어도 97, 적어도 98, 적어도 99, 적어도 99.5% 또는 이 초과)하고, 아미노산 서열에서 아직 상이한 상응하는 전장 단백질의 단백질의 기능적 활성(예를 들어, ALK 억제제에 대한 내성 또는 민감성 부여)을 보유한다.

[0129] 2개의 아미노산 서열 또는 2개의 핵산의 동일성 퍼센트를 결정하기 위해, 서열은 최적 비교 목적에 대해 정렬된다(예를 들어, 두번째 아미노 또는 핵산 서열과의 최적 정렬을 위해 첫번째 아미노산 또는 핵산 서열의 서열에 갭이 도입될 수 있음). 이후, 상응하는 아미노산 위치 또는 뉴클레오타이드 위치에서의 아미노산 잔기 또는 뉴클레오타이드가 비교된다. 첫번째 서열 내의 위치가 두번째 서열 내의 상응하는 위치와 동일한 아미노산 잔기 또는 뉴클레오타이드에 의해 차지되는 경우, 분자는 상기 위치에서 동일하다. 2개의 서열 사이의 동일성 퍼센트는 서열에 의해 공유되는 동일한 위치의 수의 함수이다(즉, 동일성 % = 동일한 위치의 # / 위치의 전체 # (예를 들어, 중첩 위치) x 100). 일 구체예에서, 2개의 서열은 동일한 길이이다.

[0130] 2개의 서열 사이의 동일성 퍼센트의 결정은 수학적 알고리즘을 이용하여 달성될 수 있다. 2개의 서열의 비교에 사용되는 수학적 알고리즘의 또 다른 비제한적인 예는 문헌[Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877]에서와 같이 변형된 문헌[Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268]의 알고리즘이다. 이러한 알고리즘은 문헌[Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410]의 NBLAST 및 XBLAST 프로그램에 통합된다. BLAST 뉴클레오타이드 검색은 본 발명의 핵산 분자에 상동성인 뉴클레오타이드 서열을 수득하기 위해 NBLAST 프로그램, 스코어 = 100, 워드길이(wordlength) = 12로 수행될 수 있다. BLAST 단백질 검색은 본 발명의 단백질 분자에 상동성인 아미노산 서열을 수득하기 위해 XBLAST 프로그램, 스코어 = 50, 워드길이 = 3으로 수행될 수 있다. 비교 목적상 갭이 도입된 정렬을 수득하기 위해, 문헌[Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402]에 기재된 바와 같이 갭이 도입된 BLAST(Gapped BLAST)가 이용될 수 있다. 대안적으로, 분자 사이의 먼 관계를 검출하는 반복 검색을 수행하는데 PSI-Blast가 이용될 수 있다. BLAST, 갭이 도입된 BLAST, 및 PSI-Blast 프로그램을 이용하는 경우, 각각의 프래그먼트(예를 들어, XBLAST 및 NBLAST)의 디폴트 파라미터가 이용될 수 있다(예를 들어, ncbi.nlm.nih.gov의 월드 와이드 웹의 NCBI 웹사이트 참조). 서열의 비교에 사용되는 수학적 알고리즘의 또 다른 비제한적인 예는 문헌[Myers and Miller, (1988) *Comput Appl Biosci*, 4: 11-7]의 알고리즘이다. 이러한 알고리즘은 GCG 서열 정렬 소프트웨어 패키지의 부분인 ALIGN 프로그램(버전 2.0)에 통합된다. 아미노산 서열을 비교하기 위해 ALIGN 프로그램을 이용하는 경우, PAM120 중량 잔기 표, 12의 갭 길이 페널티, 및 4의 갭 페널티가 이용될 수 있다. 국소 서열 유사성 및 정렬의 영역을 확인하기 위한 또 다른 유용한 알고리즘은 문헌[Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448]에 기재된 바와 같은 FASTA 알고리즘이다. 뉴클레오타이드 또는 아미노산 서열을 비교하기 위해 FASTA 알고리즘을 이용하는 경우, 예를 들어, PAM120 중량 잔기 표가 2의 k -터플(k -tuple) 값과 함께 이용될 수 있다.

[0131] 2개의 서열 사이의 동일성 퍼센트는 갭을 허용하거나 갭을 허용함이 없이 상기 기재된 것과 유사한 기술을 이용하여 결정될 수 있다. 동일성 퍼센트 계산에서, 정확한 매치만이 계수된다.

[0132] 본 발명의 마커에 상응하는 분리된 폴리펩티드 또는 이의 단편은 폴리클로날 및 모노클로날 항체 제조물에 대한 표준 기술을 이용하여 항체를 생성시키기 위한 면역원으로 사용될 수 있다. 전장 폴리펩티드 또는 단백질이 사용될 수 있거나, 대안적으로 본 발명은 면역원으로 사용하기 위한 항원성 펩티드 단편을 제공한다. 본 발명의 단백질의 항원성 펩티드는 본 발명의 폴리펩티드 중 하나의 아미노산 서열의 적어도 8개(또는 적어도 10개, 적어도 15개, 적어도 20개, 또는 적어도 30개 또는 이 초과)의 아미노산 잔기를 포함하고, 이는 단백질의 에피토프를 포함하여, 펩티드에 대해 발생된 항체가 단백질에 상응하는 본 발명의 마커와 특이적 면역 복합체를 형성한다. 항원성 펩티드에 의해 포함되는 예시적 에피토프는 단백질의 표면에 위치되는 영역, 예를 들어, 친수성 영역이다. 친수성 영역을 확인하기 위해 소수성 서열 분석, 친수성 서열 분석, 또는 유사한 분석이 이용될 수 있다.

[0133] 면역원은 통상적으로 적합한(즉, 면역적격) 피검체, 예를 들어, 토끼, 염소, 마우스, 또는 다른 포유동물 또는

척추동물을 면역화시킴으로써 항체를 제조하는데 사용된다. 적절한 면역원 제조물은, 예를 들어, 재조합적으로 발현되거나 화학적으로 합성된 폴리펩티드를 함유할 수 있다. 제조물은 애쥬번트, 예를 들어, 프로인트(Freund) 완전 또는 불완전 애쥬번트, 또는 유사한 면역자극제를 추가로 포함할 수 있다.

[0134] 따라서, 본 발명의 또 다른 양태는 본 발명의 폴리펩티드에 특이적인 항체에 속한다. 본원에서 상호교환적으로 사용되는 용어 "항체" 및 "항체 물질"은 면역글로불린 분자 및 면역글로불린 분자의 면역학적 활성 부분, 즉, 본 발명의 폴리펩티드와 같은 항원에 특이적으로 결합하는 항원 결합 부위를 함유하는 분자를 의미한다. 본 발명의 제공된 폴리펩티드에 특이적으로 결합하는 분자는 폴리펩티드에 결합하나, 샘플, 예를 들어, 폴리펩티드를 자연적으로 함유하는 생물학적 샘플 내의 다른 분자에는 실질적으로 결합하지 않는 분자이다. 면역글로불린 분자의 면역학적 활성 부분의 예는 항체에 펩신과 같은 효소를 처리함으로써 생성될 수 있는 F(ab) 및 F(ab')₂를 포함한다. 본 발명은 폴리클로날 및 모노클로날 항체를 제공한다. 본원에서 사용되는 용어 "모노클로날 항체" 또는 "모노클로날 항체 조성물"은 특정 에피토프와 면역반응할 수 있는 항원 결합 부위의 하나의 종만을 함유하는 항체 분자의 집단을 의미한다.

[0135] 폴리클로날 항체는 면역원으로서 본 발명의 폴리펩티드로 적합한 피검체를 면역화시킴으로써 상기 기재된 바와 같이 제조될 수 있다. 면역화된 피검체에서의 항체 역가는 표준 기술, 예를 들어, 고정된 폴리펩티드를 이용하는 효소면역측정법(ELISA)을 이용하여 시간 경과에 걸쳐 모니터링될 수 있다. 요망시, 항체 분자는 피검체로부터 수거되거나 분리(예를 들어, 피검체의 혈액 또는 혈청으로부터 수거 또는 분리)될 수 있고, IgG 분획을 수득하기 위해 널리 공지된 기술, 예를 들어, 단백질 A 크로마토그래피에 의해 추가로 정제될 수 있다. 면역화 후 적절한 시점에서, 예를 들어, 특정 항체 역가가 가장 높은 경우, 항체 생성 세포가 피검체로부터 수득될 수 있고, 표준 기술, 예를 들어, 문헌[Kohler and Milstein (1975) *Nature* 256:495-497]에 본래 기재된 하이브리도마 기술, 인간 B 세포 하이브리도마 기술(예를 들어, Kozbor *et al.*, 1983, *Immunol. Today* 4:72 참조), EBV-하이브리도마 기술(예를 들어, Cole *et al.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985 참조) 또는 트리오마(trioma) 기술에 의해 모노클로날 항체를 제조하는데 사용될 수 있다. 하이브리도마를 생성시키기 위한 기술은 널리 공지되어 있다(예를 들어, 일반적으로 *Current Protocols in Immunology*, Coligan *et al.* ed., John Wiley & Sons, New York, 1994 참조). 본 발명의 모노클로날 항체를 생성하는 하이브리도마 세포는, 예를 들어, 표준 ELISA 검정을 이용하여 관심 폴리펩티드에 결합하는 항체에 대해 하이브리도마 배양 상층액을 스크리닝함으로써 검출된다.

[0136] 모노클로날 항체-분비 하이브리도마를 제조하는 것에 대한 대안으로, 본 발명의 폴리펩티드에 특이적인 모노클로날 항체는 관심 폴리펩티드로 재조합의 조합 면역글로불린 라이브러리(예를 들어, 항체 파지 디스플레이 라이브러리)를 스크리닝함으로써 확인되고 분리될 수 있다. 파지 디스플레이 라이브러리를 생성시키기 스크리닝하기 위한 키트가 시판된다(예를 들어, Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; 및 Stratagene SurfZAP Phage Display Kit, Catalog No. 240612). 추가로, 항체 디스플레이 라이브러리를 생성시키고 스크리닝하는데 특히 사용할 수 있는 방법 및 시약의 예는, 예를 들어, 미국 특허 번호 5,223,409호(참조로서 포함됨); PCT 공개 번호 WO 92/18619호(참조로서 포함됨); PCT 공개 번호 WO 91/17271호(참조로서 포함됨); PCT 공개 번호 WO 92/20791호(참조로서 포함됨); PCT 공개 번호 WO 92/15679호(참조로서 포함됨); PCT 공개 번호 WO 93/01288호(참조로서 포함됨); PCT 공개 번호 WO 92/01047호(참조로서 포함됨); PCT 공개 번호 WO 92/09690호(참조로서 포함됨); PCT 공개 번호 WO 90/02809호(참조로서 포함됨); 문헌[Fuchs *et al.* (1991) *Bio/Technology* 9: 1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246: 1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734]에서 발견될 수 있다.

[0137] 추가로, 표준 재조합 DNA 기술을 이용하여 제조될 수 있는 인간 및 비-인간 부분을 포함하는 키메라 및 인간화된 모노클로날 항체와 같은 재조합 항체가 본 발명의 범위에 속한다. 이러한 키메라 및 인간화된 모노클로날 항체는 당 분야에 공지된 재조합 DNAs 기술, 예를 들어, PCT 공개 번호 WO 87/02671호(참조로서 포함됨); 유럽 특허 출원 184,187호; 유럽 특허 출원 171,496호; 유럽 특허 출원 173,494호; PCT 공개 번호 WO 86/01533호(참조로서 포함됨); 미국 특허 번호 4,816,567호(참조로서 포함됨); 유럽 특허 출원 125,023호; 문헌[Better *et al.* (1988) *Science* 240: 1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; 및 Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80: 1553-1559; Morrison (1985) *Science* 229: 1202-1207; Oi *et al.* (1986) *Bio/Techniques* 4:214]; 미국 특허 5,225,539호(참조로서 포함됨); 문헌[Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239: 1534; 및 Beidler *et al.* (1988) *J. Immunol.* 141:4053-

4060]에 기재된 방법을 이용하여 생성될 수 있다.

[0138] 완전한 인간 항체가 인간 피검체의 치료적 치료에 특히 요망된다. 이러한 항체는 내인성 면역글로불린 중쇄 및 경쇄 유전자를 발현할 수 없으나, 인간 중쇄 및 경쇄 유전자를 발현할 수 있는 트랜스제닉 마우스를 이용하여 생성될 수 있다. 트랜스제닉 마우스는 선택 항원, 예를 들어, 본 발명의 마커에 상응하는 폴리펩티드의 전부 또는 일부로 일반적인 방식으로 면역화된다. 항원에 특이적인 모노클로날 항체는 통상적인 하이브리도마 기술을 이용하여 수득될 수 있다. 트랜스제닉 마우스에 의해 포함된 인간 면역글로불린 트랜스젠(transgene)은 B 세포 분화 동안 재배열되고, 이후에 종류 변환(class switching) 및 체세포 돌연변이를 겪는다. 따라서, 상기 기술을 이용하여, 치료적학적으로 유용한 IgG, IgA 및 IgE 항체를 생성시키는 것이 가능하다. 인간 항체를 생성시키기 위한 상기 기술의 개관을 위해, 문헌[Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93]을 참조하라. 인간 항체 및 인간 모노클로날 항체를 생성시키기 위한 상기 기술 및 상기 항체를 생성시키기 위한 프로토콜의 상세한 논의를 위해, 예를 들어, 미국 특허 5,625,126호(참조로서 포함됨); 미국 특허 5,633,425호(참조로서 포함됨); 미국 특허 5,569,825 (참조로서 포함됨); 미국 특허 5,661,016호(참조로서 포함됨); 및 미국 특허 5,545,806호(참조로서 포함됨)를 참조하라. 또한, Abgenix, Inc.(Freemont, CA)와 같은 회사가 상기 기재된 것과 유사한 기술을 이용하여 선택된 항원에 특이적인 인간 항체를 제공하는 것을 책임질 수 있다.

[0139] 선택 에피토프를 인지하는 완전한 인간 항체는 "유도 선택(guided selection)"으로 언급되는 기술을 이용하여 생성될 수 있다. 이러한 방법에서, 선택 비-인간 모노클로날 항체, 예를 들어, 무린 항체는 동일 에피토프를 인지하는 완전한 인간 항체의 선택을 유도하기 위해 사용된다(Jespers *et al.*, 1994, *Bio/technology* 12:899-903).

[0140] 본 발명의 마커에 상응하는 폴리펩티드에 특이적인 항체(예를 들어, 모노클로날 항체)는 표준 기술, 예를 들어, 친화성 크로마토그래피 또는 면역침전법에 의해 폴리펩티드를 분리하는데 이용될 수 있다. 또한, 상기 항체는 마커 발현의 수준 및 패턴을 평가하기 위해 마커(예를 들어, 세포 용해질 또는 세포 상층액 중 마커)를 검출하는데 이용될 수 있다. 항체는 또한 임상 시험 절차의 일부로서, 예를 들어, 제공된 치료 요법의 효능을 결정하기 위해 조직 또는 체액(예를 들어, 종양 세포 함유 체액) 중의 단백질 수준을 진단적으로 모니터링하는데 이용될 수 있다. 검출은 검출가능한 물질에 항체를 커플링시킴으로써 촉진될 수 있다. 검출가능한 물질의 예는 다양한 효소, 보결분자단(prosthetic group), 형광 물질, 발광 물질, 생물발광 물질, 및 방사성 물질을 포함하나, 이에 제한되지는 않는다. 적합한 효소의 예는 호스라디쉬 퍼옥시다제(horseradish peroxidase), 알칼리성 포스파타제, β -갈락토시다제, 또는 아세틸콜린에스테라제를 포함하나, 이에 제한되지 않고; 적합한 보결분자단 복합체의 예는 스트렙타비딘/비오틴 및 아비딘/비오틴을 포함하나, 이에 제한되지는 않고; 적합한 형광 물질의 예는 움벨리페론, 플루오레세인, 플루오레세인 이소티오시아네이트, 로다민, 디클로로트리아지닐아민 플루오레세인, 단실 클로라이드 또는 피코에리트린을 포함하나, 이에 제한되지는 않고; 발광 물질의 예는 루미놀을 포함하나, 이에 제한되지는 않고; 생물발광 물질의 예는 루시페라제, 루시페린, 및 애큐오린을 포함하나, 이에 제한되지는 않고; 적합한 방사성 물질의 예는 ^{125}I , ^{131}I , ^{35}S 또는 ^3H 를 포함하나, 이에 제한되지는 않는다.

[0141] V. 예시적 재조합 발현 벡터 및 숙주 세포

[0142] 본 발명의 또 다른 양태는 본 발명의 마커에 상응하는 폴리펩티드(또는 이러한 폴리펩티드의 부분)를 엔코딩하는 핵산을 함유하는 벡터, 예를 들어, 발현 벡터에 속한다. 본원에서 사용되는 용어 "벡터"는 연결된 또 다른 핵산을 운반할 수 있는 핵산 분자를 의미한다. 벡터의 한 유형은 추가 DNA 세그먼트가 연결될 수 있는 원형의 이중 가닥 DNA 루프를 의미하는 "플라스미드"이다. 벡터의 또 다른 유형은 추가 DNA 세그먼트가 바이러스 유전체에 라이게이션될 수 있는 바이러스 벡터이다. 특정 벡터는 이들이 도입되는 숙주 세포에서 자율 복제할 수 있다(예를 들어, 박테리아 복제 기점 및 에피솜 포유동물 벡터를 갖는 박테리아 벡터). 다른 벡터(예를 들어, 비-에피솜 포유동물 벡터)는 숙주 세포로 도입시 숙주 세포의 유전체로 통합됨으로써, 숙주 유전체와 함께 복제된다. 또한, 특정 벡터, 즉 발현 벡터는 이들이 작동가능하게 연결되는 유전자의 발현을 유도할 수 있다. 일반적으로, 재조합 DNA 기술에 유용한 발현 벡터는 종종 플라스미드(벡터)의 형태로 존재한다. 그러나, 본 발명은 동등한 기능을 제공하는 발현 벡터의 다른 형태, 예를 들어, 바이러스 벡터(예를 들어, 복제 결함 레트로바이러스, 아데노바이러스 및 아데노-관련 바이러스)를 포함한다.

[0143] 본 발명의 재조합 발현 벡터는 숙주 세포에서의 핵산의 발현에 적합한 형태의 본 발명의 핵산을 포함한다. 이는 재조합 발현 벡터가 발현되는 핵산 서열에 작동가능하게 연결된, 발현에 사용되는 숙주 세포를 기초로 하여 선택된 하나 이상의 조절 서열을 포함하는 것을 의미한다. 재조합 발현 벡터 내에서, "작동가능하게 연결된"은 관심 뉴클레오티드 서열이 뉴클레오티드 서열의 발현(예를 들어, 벡터가 숙주 세포에 도입되는 경우 시험관내

전사/번역 시스템 또는 숙주 세포 내에서 발현)을 가능케 하는 방식으로 조절 서열(들)에 연결되는 것을 의미한다. 용어 "조절 서열"은 프로모터, 인핸서 및 다른 발현 조절 요소(예를 들어, 아데닐중합체형성 신호)를 포함한다. 이러한 조절 서열은, 예를 들어, 문헌[Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991)]에 기재되어 있다. 조절 서열은 많은 유형의 숙주 세포에서 누클레오타이드 서열의 항시성(constitutive) 발현을 유도하는 것 및 특정 숙주 세포에서만 누클레오타이드 서열의 발현을 유도하는 것(예를 들어, 조직-특이적 조절 서열)을 포함한다. 발현 벡터의 디자인은 형질전환되는 숙주 세포의 선택, 요망되는 단백질의 발현 수준 등과 같은 요인에 좌우될 수 있음이 당업자에 의해 인식될 것이다. 본 발명의 발현 벡터는 숙주 세포에 도입됨으로써, 본원에 기재된 바와 같은 핵산에 의해 엔코딩되는 융합 단백질 또는 펩티드를 포함하는 단백질 또는 펩티드를 생성할 수 있다.

[0144] 본 발명의 재조합 발현 벡터는 원핵생물(예를 들어, 이. 콜라이(E. coli)) 또는 진핵생물 세포(예를 들어, 곤충 세포(배큘로바이러스 발현 벡터를 이용함), 효모 세포 또는 포유동물 세포)에서의 본 발명의 마커에 상응하는 폴리펩티드의 발현을 위해 디자인될 수 있다. 적합한 숙주 세포는 문헌[Goeddel, 상기]에 추가로 논의되어 있다. 대안적으로, 재조합 발현 벡터는, 예를 들어, T7 프로모터 조절 서열 및 T7 중합효소를 이용하여 시험관내에서 전사되고 번역될 수 있다.

[0145] 원핵생물에서의 단백질의 발현은 가장 흔하게는 융합 또는 비-융합 단백질의 발현을 유도하는 항시성 또는 유도성 프로모터를 함유하는 벡터를 이용하여 이. 콜라이에서 수행된다. 융합 벡터는 이에 의해 엔코딩되는 단백질, 보통 재조합 단백질의 아미노 말단에 다수의 아미노산을 첨가한다. 이러한 융합 벡터는 통상적으로, 1) 재조합 단백질의 발현을 증가시키고; 2) 재조합 단백질의 용해도를 증가시키고; 3) 친화성 정제에서 리간드로서 작용함으로써 재조합 단백질의 정제를 돕는 3개의 목적에 도움이 된다. 종종, 융합 발현 벡터에서, 단백질분해성 절단 부위가 융합 단백질의 정제 후에 융합 모이어티로부터 재조합 단백질의 분리를 가능케 하기 위해 융합 모이어티 및 재조합 단백질의 접합부에 도입된다. 이러한 효소 및 이의 인지체 인지 서열은 인자 Xa, 트롬빈 및 엔테로키나제를 포함한다. 통상적 융합 벡터는 표적 재조합 단백질에 글루타티온 S-트랜스페라제(GST), 말토오스 E 결합 단백질, 또는 단백질 A를 각각 융합시키는 pGEX(Pharmacia Biotech Inc; Smith and Johnson, 1988, *Gene* 67:31-40), pMAL(New England Biolabs, Beverly, MA) 및 pRIT5(Pharmacia, Piscataway, NJ)를 포함한다.

[0146] 적합한 유도성 비-융합 이. 콜라이 발현 벡터의 예는 pTrc(Amann *et al.*, 1988, *Gene* 69:301-315) 및 pET 11d(Studier *et al.*, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991)를 포함한다. pTrc 벡터로부터의 표적 유전자 발현은 하이브리드 trp-lac 융합 프로모터로부터의 숙주 RNA 중합효소 전사에 좌우된다. pET 11d 벡터로부터의 표적 유전자 발현은 공동 발현되는 바이러스 RNA 중합효소(T7 gn1)에 의해 매개되는 T7 gn10-lac 융합 프로모터로부터의 전사에 좌우된다. 이러한 바이러스 중합효소는 lacUV 5 프로모터의 전사 조절하에서 T7 gn1 유전자를 갖는 상재균 프로파지로부터의 숙주 균주 BL21(DE3) 또는 HMS174(DE3)에 의해 공급된다.

[0147] 이. 콜라이에서의 재조합 단백질 발현을 최대화시키기 위한 한 방법은 재조합 단백질을 단백질분해적으로 절단시키기 위해 손상된 능력을 갖는 숙주 박테리아에서 단백질을 발현시키는 것이다(Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990). 또 다른 방법은 발현 벡터에 삽입되는 핵산의 핵산 서열을 변경시켜, 각각의 아미노산에 대한 개별적 코돈이 이. 콜라이에서 사용되는 것 중 하나일 수 있도록 하는 것이다(Wada *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118). 이러한 본 발명의 핵산 서열의 변경은 표준 DNA 합성 기술에 의해 수행될 수 있다.

[0148] 또 다른 구체예에서, 발현 벡터는 효모 발현 벡터이다. 효모 에스. 세레비지애(S. cerevisiae)에서의 발현을 위한 벡터의 예는 pYepSec1(Baldari *et al.*, 1987, *EMBO J.* 6:229-234), pMFa(Kurjan and Herskowitz, 1982, *Cell* 30:933-943), pJRY88(Schultz *et al.*, 1987, *Gene* 54:113-123), pYES2(Invitrogen Corporation, San Diego, CA), 및 pPicZ(Invitrogen Corp, San Diego, CA)를 포함한다.

[0149] 대안적으로, 발현 벡터는 배큘로바이러스 발현 벡터이다. 배양된 곤충 세포(예를 들어, Sf9 세포)에서의 단백질의 발현에 이용가능한 배큘로바이러스 벡터는 pAc 시리즈(Smith *et al.*, 1983, *Mol. Cell Biol.* 3:2156-2165) 및 pVL 시리즈(Lucklow and Summers, 1989, *Virology* 170:31-39)를 포함한다.

[0150] 또 다른 구체예에서, 본 발명의 핵산은 포유동물 발현 벡터를 이용하여 포유동물 세포에서 발현된다. 포유동물 발현 벡터의 예는 pCDM8(Seed, 1987, *Nature* 329:840) 및 pMT2PC(Kaufman *et al.*, 1987, *EMBO J.* 6: 187-195)를 포함한다. 포유동물 세포에서 사용되는 경우, 발현 벡터의 조절 기능은 종종 바이러스 조절 요소에 의

해 제공된다. 예를 들어, 일반적으로 사용되는 프로모터는 폴리오마, 아데노바이러스 2, 사이토메갈로바이러스 및 원숭이 바이러스 40으로부터 유래된다. 원핵생물 및 진핵생물 세포 둘 모두에 대한 다른 적합한 발현 시스템에 대해, 문헌[Sambrook *et al.*, 상기]의 챕터 16 및 17을 참조하라.

[0151] 또 다른 구체예에서, 재조합 포유동물 발현 벡터는 특정 세포 유형에서 핵산의 발현을 유도할 수 있다(예를 들어, 조직-특이적 조절 요소가 핵산을 발현하는데 사용된다). 조직-특이적 조절 요소는 당 분야에 공지되어 있다. 적합한 조직-특이적 프로모터의 비제한적 예는 알부민 프로모터(간-특이적; Pinkert *et al.*, 1987, *Genes Dev.* 1:268-277), 림프-특이적 프로모터(Calame and Eaton, 1988, *Adv. Immunol.* 43:235-275), 예를 들어, T 세포 수용체의 프로모터(Winoto and Baltimore, 1989, *EMBO J.* 8:729-733) 및 면역글로불린(Banerji *et al.*, 1983, *Cell* 33:729-740; Queen and Baltimore, 1983, *Cell* 33:741-748), 뉴런-특이적 프로모터(예를 들어, 신경미세섬유 프로모터; Byrne and Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), 췌장-특이적 프로모터(Edlund *et al.*, 1985, *Science* 230:912-916), 및 유선-특이적 프로모터(예를 들어, 유청단백질 프로모터; 미국 특허 번호 4,873,316호(참조로서 포함됨) 및 유럽 출원 공개 번호 264,166호)를 포함한다. 발달적으로 조절되는 프로모터, 예를 들어, 유린 호스(hox) 프로모터(Kessel and Gruss, 1990, *Science* 249:374-379) 및 α -태아단백질 프로모터(Camper and Tilghman, 1989, *Genes Dev.* 3:537-546)가 또한 포함된다.

[0152] 본 발명은 안티센스 배향으로 발현 벡터로 클로닝되는 본 발명의 DNA 분자를 포함하는 재조합 발현 벡터를 추가로 제공한다. 즉, DNA 분자는 본 발명의 폴리펩티드를 엔코딩하는 mRNA에 대해 안티센스인 RNA 분자의 발현(DNA 분자의 전사에 의함)을 가능케 하는 방식으로 조절 서열에 작동가능하게 연결된다. 다양한 세포 유형에서 안티센스 RNA 분자의 연속적 발현을 유도하는 안티센스 배향으로 클로닝되는 핵산에 작동가능하게 연결되는 조절 서열이 선택될 수 있고, 예를 들어, 안티센스 RNA의 항시성의 조직-특이적 또는 세포 유형 특이적 발현을 유도하는 바이러스 프로모터 및/또는 인핸서, 또는 조절 서열이 선택될 수 있다. 안티센스 발현 벡터는, 안티센스 핵산이 고효율 조절 영역의 조절하에서 생성되는 재조합 플라스미드, 파지미드, 또는 약독화 바이러스의 형태로 존재할 수 있고, 이의 활성은 벡터가 도입되는 세포 유형에 의해 결정될 수 있다. 안티센스 유전자를 이용한 유전자 발현의 조절의 논의에 대해, 문헌[Weintraub *et al.*, 1986, *Trends in Genetics*, Vol. 1(1)]을 참조하라.

[0153] 본 발명의 또 다른 양태는 본 발명의 재조합 발현 벡터가 도입되는 숙주 세포에 속한다. 용어 "숙주 세포" 및 "재조합 숙주 세포"는 본원에서 상호교환적으로 사용된다. 이러한 용어가 특정 피검체 세포 뿐만 아니라 이러한 세포의 프로제니(progeny) 또는 잠재적 프로제니를 의미하는 것이 이해된다. 특정 변형이 돌연변이 또는 환경적 영향으로 인해 후속 세대에서 발생할 수 있으므로, 상기 프로제니는 사실상 모(parent) 세포와 동일하지 않을 수 있으나, 본원에서 사용되는 용어의 범위 내에 여전히 포함된다.

[0154] 숙주 세포는 임의의 원핵생물(예를 들어, 이. 콜라이) 또는 진핵생물 세포(예를 들어, 곤충 세포, 효모 또는 포유동물 세포)일 수 있다.

[0155] 벡터 DNA는 통상적인 형질전환 또는 트랜스펙션 기술을 통해 원핵생물 또는 진핵생물 세포에 도입될 수 있다. 본원에서 사용되는 용어 "형질전환" 및 "트랜스펙션"은 인산칼슘 또는 염화칼슘 공동-침전, DEAE-텍스트란-매개 트랜스펙션, 리포펙션, 또는 전기천공을 포함하는 숙주 세포로 외래 핵산을 도입시키기 위한 다양한 당 분야에 인지되는 기술을 의미한다. 숙주 세포를 형질전환시키거나 트랜스펙션시키기 위한 적합한 방법은 문헌[Sambrook, *et al.* (상기)], 및 다른 실험실 메뉴얼에서 발견될 수 있다.

[0156] 포유동물 세포의 안정적 트랜스펙션을 위해, 사용되는 발현 벡터 및 트랜스펙션 기술에 따라, 세포의 단지 작은 분획이 외래 DNA를 이에 유전체에 통합시킬 수 있는 것이 공지되어 있다. 상기 통합체(integrand)를 확인하고 선택하기 위해, 선택가능한 마커(예를 들어, 항생제에 대한 내성)를 엔코딩하는 유전자는 일반적으로 관심 유전자와 함께 숙주 세포에 도입된다. 예시적 선택가능한 마커는 약물, 예를 들어, G418, 하이그로마이신 및 메토트렉세이트에 대한 내성을 부여하는 것을 포함한다. 도입된 핵산으로 안정적으로 트랜스펙션된 세포는 약물 선택에 의해 확인될 수 있다(예를 들어, 선택가능한 마커 유전자가 통합된 세포는 생존하는 반면, 다른 세포는 사멸할 것이다).

[0157] 배양되는 본 발명의 숙주 세포, 예를 들어, 원핵생물 또는 진핵생물 숙주 세포는 본 발명의 마커에 상응하는 폴리펩티드를 생성시키는데 사용될 수 있다. 따라서, 본 발명은 본 발명의 숙주 세포를 이용한 본 발명의 마커에 상응하는 폴리펩티드를 생성시키는 방법을 추가로 제공한다. 한 구체예에서, 상기 방법은 적합한 배지에서 본 발명의 숙주 세포(본 발명의 폴리펩티드를 엔코딩하는 재조합 발현 벡터가 도입된 숙주 세포)를 배양하여 마커를 생성시키는 것을 포함한다. 또 다른 구체예에서, 상기 방법은 배지 또는 숙주 세포로부터 마커 폴리펩티드

를 분리시키는 것을 추가로 포함한다.

[0158] 본 발명의 숙주 세포는 또한 비인간 트랜스제닉 동물을 생성시키는데 사용될 수 있다. 예를 들어, 한 구체예에서, 본 발명의 숙주 세포는 본 발명의 마커에 상응하는 폴리펩티드를 엔코딩하는 서열이 도입된 수정된 난모세포 또는 배아 줄기세포이다. 이러한 숙주 세포는 이후에 본 발명의 마커 단백질을 엔코딩하는 외인성 서열이 숙주 세포의 유전체에 도입된 비-인간 트랜스제닉 동물, 또는 본 발명의 서열의 마커에 상응하는 폴리펩티드를 엔코딩하는 내인성 유전자(들)이 변경된 상동성 재조합 동물을 생성시키는데 이용될 수 있다. 이러한 동물은 마커에 상응하는 폴리펩티드의 기능 및/또는 활성을 연구하고, 폴리펩티드 활성의 조절인자를 확인하고/하거나 평가하는데 유용할 뿐만 아니라, 마커 발견 또는 평가, 예를 들어, 치료 및 진단 마커 발견 또는 평가를 위한 치료 또는 진단 분자의 전임상 시험에서 유용하거나, 약물 효능 및 특이성의 대용약으로 유용하다.

[0159] 본원에서 사용되는 용어 "트랜스제닉 동물"은 비-인간 동물, 예를 들어, 포유동물, 예를 들어, 설치류, 예를 들어, 래트 또는 마우스이며, 상기 동물의 세포의 하나 이상은 트랜스진을 포함한다. 트랜스제닉 동물의 다른 예는 비-인간 영장류, 양, 개, 소, 염소, 닭, 양서류 등을 포함한다. 트랜스진은 트랜스제닉 동물이 발달하는 세포의 유전체로 통합되고, 성숙 동물의 유전체에 남아있음으로써, 트랜스제닉 동물의 하나 이상의 세포 유형 또는 조직에서 엔코딩되는 유전자 생성물의 발현을 유도하는 외인성 DNA이다. 본원에서 사용되는 용어 "상동성 재조합 동물"은 비-인간 동물, 예를 들어, 포유동물, 예를 들어, 마우스이며, 여기서, 내인성 유전자는 동물의 발달 전에 내인성 유전자와 동물의 세포, 예를 들어, 동물의 배아 세포로 도입되는 외인성 DNA 분자 사이의 상동성 재조합에 의해 변경된다. 트랜스제닉 동물은 또한, 예를 들어, 문헌[Chan I.T., *et al.* (2004) *J Clin Invest.* 113(4):528-38 and Chin L. *et al* (1999) *Nature* 400(6743):468-72]에 기재된 것과 같은 유도성 트랜스제닉 동물을 포함한다.

[0160] 본 발명의 트랜스제닉 동물은, 예를 들어, 미세주사, 레트로바이러스 감염에 의해 수정된 난모세포의 수컷 전핵(pronucleus)으로 본 발명의 마커에 상응하는 폴리펩티드를 엔코딩하는 핵산을 도입시키고, 난모세포가 가임신 암컷 양육 동물에서 발달하도록 함으로써 생성될 수 있다. 인트론 서열 및 아데닐중합체형성 신호는 또한 트랜스진 발현의 효율을 증가시키기 위해 트랜스진에 포함될 수 있다. 조직-특이적 조절 서열(들)은 특정 세포로의 본 발명의 폴리펩티드의 발현을 유도시키기 위해 트랜스진에 작동가능하게 연결될 수 있다. 배아 조작 및 미세주사를 통한 트랜스제닉 동물, 특히 마우스와 같은 동물을 생성시키는 방법은 당 분야에서 통상적이며, 이는, 예를 들어, 미국 특허 번호 4,736,866호(참조로서 포함됨) 및 4,870,009호(참조로서 포함됨), 미국 특허 번호 4,873,191호(참조로서 포함됨) 및 문헌[Hogan, *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986]에 기재되어 있다. 유사한 방법이 다른 트랜스제닉 동물의 생성에 사용된다. 트랜스제닉 시조 동물은 이의 유전체 내의 트랜스진의 존재 및/또는 동물의 조직 또는 세포에서의 트랜스진을 엔코딩하는 mRNA의 발현을 기초로 하여 확인될 수 있다. 이후, 트랜스제닉 시조 동물은 트랜스진을 갖는 추가 동물을 개량하는데 사용될 수 있다. 또한, 트랜스진을 갖는 트랜스제닉 동물은 다른 트랜스진을 갖는 다른 트랜스제닉 동물로 추가로 개량될 수 있다.

[0161] 상동성 재조합 동물을 생성시키기 위해, 결실, 첨가 또는 치환이 도입됨으로써 유전자를 변경, 예를 들어, 유전자를 기능적으로 붕괴시키는, 본 발명의 마커에 상응하는 폴리펩티드를 엔코딩하는 유전자의 적어도 일부를 함유하는 벡터가 제조된다. 또 다른 구체예에서, 벡터는 상동성 재조합시 내인성 유전자가 기능적으로 붕괴(즉, 기능성 단백질을 더 이상 엔코딩하지 않음; "녹아웃(knock out)" 벡터로도 언급됨)되도록 디자인된다. 대안적으로, 벡터는 상동성 재조합시 내인성 유전자가 돌연변이되거나 달리 변경되지만, 여전히 기능성 단백질을 엔코딩하도록 디자인될 수 있다(예를 들어, 업스트림 조절 영역은 내인성 단백질의 발현을 변경하도록 변경될 수 있음). 상동성 재조합 벡터에서, 유전자의 변경된 부분은 벡터가 갖는 외인성 유전자와 배아 줄기세포 내의 내인성 유전자 사이에서 상동성 재조합이 발생하도록 하기 위해 유전자의 추가 핵산 옆에 5' 및 3' 말단에 플랭킹된다. 추가 플랭킹 핵산 서열은 내인성 유전자를 이용한 성공적인 상동성 재조합을 위해 충분한 길이이다. 통상적으로, 수 킬로베이스의 플랭킹 DNA(5' 및 3' 말단 둘 모두)가 벡터에 포함된다(예를 들어, 상동성 재조합 벡터에 대해 문헌[Thomas and Capecchi, 1987, *Cell* 51:503] 참조). 벡터는 배아 줄기세포주(예를 들어, 전기천공에 의해)로 도입되고, 도입된 유전자가 내인성 유전자와 재조합적으로 조합된 세포가 선택된다(예를 들어, 문헌 Li *et al.*, 1992, *Cell* 69:915] 참조). 이후, 선택된 세포는 동물(예를 들어, 마우스)의 주머니배에 주입되어 응집 키메라가 형성된다(예를 들어, 문헌[Bradley, *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, Ed., IRL, Oxford, 1987, pp. 113-152] 참조). 이후, 키메라 배아는 적합한 가임신 암컷 양육 동물에 이식될 수 있고, 배아가 분만된다. 생식세포 내에 상동성 재조합 DNA를 갖는 프로젝트가, 동물의 모든 세포가 트랜스진의 생식선 전달에 의해 상동성 재조합 DNA를 함유하는 동물을 개량시키는데

사용될 수 있다. 상동성 재조합 벡터 및 상동성 재조합 동물을 작제하기 위한 방법은 문헌[Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829] 및 PCT 공개 번호 WO 90/11354호(참조로서 포함됨), WO 91/01140호(참조로서 포함됨), WO 92/0968호(참조로서 포함됨), 및 WO 93/04169호(참조로서 포함됨)에 추가로 기재되어 있다.

[0162] 또 다른 구체예에서, 트랜스진의 조절된 발현을 가능케 하는 선택된 시스템을 함유하는 트랜스제닉 비-인간 동물이 생성될 수 있다. 이러한 시스템의 한 예는 박테리오파지 P1의 *cre/loxP* 재조합효소 시스템이다. *cre/loxP* 재조합효소 시스템의 기재에 대해, 예를 들어, 문헌[Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236]를 참조하라. 재조합효소 시스템의 또 다른 예는 사카로마이세스 세레비지애(*Saccharomyces cerevisiae*)의 FLP 재조합효소 시스템이다(O'Gorman *et al.*, 1991, *Science* 251:1351-1355). *cre/loxP* 재조합효소 시스템이 트랜스진의 발현을 조절하기 위해 사용되는 경우, *Cre* 재조합효소 및 선택된 단백질 둘 모두를 엔코딩하는 트랜스진을 함유하는 동물이 필요하다. 이러한 동물은, 예를 들어, 선택된 단백질을 엔코딩하는 트랜스진을 함유하는 동물 및 재조합효소를 엔코딩하는 트랜스진을 함유하는 동물의 2개의 트랜스제닉 동물을 교배시킴에 의한 "이중" 트랜스제닉 동물의 작제를 통해 제공될 수 있다.

[0163] 본원에 기재된 비-인간 트랜스제닉 동물의 클론은 또한 문헌[Wilmut *et al.* (1997) *Nature* 385:810-813] 및 PCT 공개 번호 WO 97/07668호(참조로서 포함됨) 및 WO 97/07669호(참조로서 포함됨)에 기재된 방법에 따라 생성될 수 있다.

[0164] V. 예시적 키트

[0165] 키트는, 본 발명의 방법을 수행하기 위한 유닛으로 판매 촉진되거나, 배포되거나, 판매되는 제품인, 적어도 하나의 시약, 예를 들어, 본 발명의 마커를 특이적으로 검출하기 위한 프로브를 포함하는 임의의 제품(예를 들어, 패키지 또는 용기)이다. 본 발명의 조성물, 키트, 및 방법이 본 발명의 방법을 수행하기 위해 사용되는 경우, 본 발명의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)은 양성 결과가 상응하는 단계, 등급, 조직학적 유형, 또는 양성/전암성/악성 특성의 암에 걸린 피검체의 적어도 약 20%, 적어도 약 40%, 적어도 약 60%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 99% 또는 100%에서 수득되도록 선택될 수 있다. 특정 구체예에서, 본 발명의 마커 또는 마커의 패널은 약 10%를 초과하는 PPV(양성예측도)가 전반적 집단에서 수득(예를 들어, 99.5%를 초과하는 검정 특이성과 커플링됨)되도록 선택될 수 있다.

[0166] 본 발명의 다수의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)이 본 발명의 조성물, 키트, 및 방법에서 사용되는 경우, 각각의 마커의 양, 구조, 및/또는 활성 또는 발현의 수준 또는 카피수는 동일 유형의 비암성 샘플, 단일 반응 혼합물(즉, 각각의 마커에 대해 상이한 형광 프로브와 같은 시약을 이용함) 또는 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커) 중 하나 이상에 상응하는 개별적 반응 혼합물에서의 다수의 마커 각각의 정상적인 양, 구조, 및/또는 활성 또는 발현 수준 또는 카피수와 비교될 수 있다. 다수의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)이 사용되는 경우, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 이 초과 개개의 마커가 사용되거나 확인될 수 있다.

[0167] 본 발명은 샘플(예를 들어, 집적된 조직(archived tissue) 샘플 또는 피검체로부터 수득된 샘플) 내의 암 세포를 검정하기 위한 조성물, 키트, 및 방법을 포함한다. 이러한 조성물, 키트, 및 방법은 필요시 조성물, 키트, 및 방법이 특정 유형의 샘플과 함께 사용하기 위해 적합화되는 것을 제외하고는 상기 기재된 것과 실질적으로 동일하다.

[0168] 따라서, 본 발명은 ALK 억제제(예를 들어, 피검체 샘플과 같은 샘플 내의 ALK 억제제)에 대한 감소된 반응을 갖거나 가질 수 있는 암 세포의 존재를 평가하기 위한 키트를 포함한다. 키트는 본 발명의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)을 확인할 수 있는, 예를 들어, 본 발명의 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)에 상응하는 핵산 또는 폴리펩티드와 특이적으로 결합할 수 있는 하나 이상의 시약을 포함할 수 있다. 본 발명의 마커에 상응하는 폴리펩티드와 결합하는 적합한 시약은 항체, 항체 유도체, 항체 단편 등을 포함한다. 핵산(예를 들어, 유전체 DNA, mRNA, 스플라이싱된 mRNA, cDNA 등)과 결합하는 적합한 시약은 상보적 핵산을 포함한다. 예를 들어, 핵산 시약은 기관에 고정된 올리고뉴클레오티드(표지되거나 표지되지 않음), 기관과 결합하지 않은 표지된 올리고뉴클레오티드, PCR 프라이머의 쌍, 분자 비콘 프로브 등을 포함할 수 있다. 일부 구체예에서, 키트는 표 1에 나열된 적어도 하나의 돌연변이를 포함하는 핵산의 적어도 하나의 스트레치 주위의 핵산의 스트레치를 인지하고 하이브리드화되는 적어도 한 쌍의 프라이머 및 상기 돌연변이의 존재에 대해 증폭된 표적 핵산을 검출하기 위한 수단을 포함하는 것과 같은 본원에 기재된 방법을 수행하는데 유용한 시약을 포함할 수 있다.

- [0169] 본 발명의 키트는 임의로 본 발명의 방법을 수행하는데 유용한 추가 성분을 포함할 수 있다. 예를 들어, 키트는 상보적 핵산을 어닐링시키거나 항체와 이러한 항체가 특이적으로 결합하는 단백질을 결합시키기에 적합한 유체(예를 들어, SSC 완충액), 하나 이상의 샘플 구획, 본 발명의 방법의 수행을 기재하는 설명서 자료, 정상 세포의 샘플, 암 세포의 샘플 등을 포함할 수 있다.
- [0170] 본 발명의 키트는 마커의 단백질 수준 또는 단백질 활성을 결정하는데 유용한 시약을 포함할 수 있다. 또 다른 구체예에서, 본 발명의 키트는 마커의 메틸화 상태를 결정하기 위한 시약을 포함할 수 있거나, 마커의 구조의 변경, 예를 들어, 돌연변이의 존재를 결정하기 위한 시약을 포함할 수 있다.
- [0171] **VI. 예측 약물**
- [0172] 본 발명은 또한 진단 검정, 약리유전학, 및 임상 시험 모니터링이 개체를 예방적으로 치료하기 위한 예측 목적에 유용한 예측 약물 분야에 속한다. 따라서, 본 발명의 한 양태는 암을 갖거나 암이 발병할 위험이 있는 개체가 ALK 억제제-매개 요법에 더욱 반응할 수 있을지의 여부를 결정하기 위한 본 발명의 하나 이상의 마커에 상응하는 폴리펩티드 또는 핵산의 양, 구조, 및/또는 활성을 결정하기 위한 검정에 관한 것이다.
- [0173] 따라서, 한 양태에서, 본 발명은 암을 가진 피검체가 ALK 억제제를 이용한 치료에 반응할 수 있는지의 여부를 결정하기 위한 방법에 관한 것이다. 또 다른 양태에서, 본 발명은 질병의 시간 경과를 예측하는 방법에 관한 것이다. 또 다른 양태에서, 상기 방법은 질병의 시간 경과에서 유의한 사건의 가능성을 예측하기 위한 방법에 관한 것이다. 특정 구체예에서, 상기 방법은 본원에 기재된 바와 같은 ALK 억제제(예를 들어, ALK 돌연변이)를 이용한 치료에 대한 반응과 관련된 바이오마커 또는 바이오마커의 조합을 검출하고, 피검체가 ALK 억제제를 이용한 치료에 반응할 수 있는지의 여부를 결정하는 것을 포함한다.
- [0174] 일부 구체예에서, 상기 방법은 ALK 돌연변이(예를 들어, 표 1에 나열된 것)를 검출하기 위해 암(예를 들어, 암의 증상으로 제공됨)을 갖는 것으로 진단되거나 암을 갖는 것으로 여겨지는 환자로부터의 생물학적 조직 샘플의 세포유전 스크리닝을 포함한다.
- [0175] 스크리닝 방법의 결과 및 이의 해석은 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료에 대한 환자 반응을 예측한다. 본 발명에 따르면, ALK 돌연변이의 존재는 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료가 ALK 돌연변이를 갖지 않는 환자의 치료 이익에 비해 암 세포에 대해 향상된 치료 이익을 제공할 것을 예측한다.
- [0176] 한 구체예에서, 본 발명의 방법은 환자로부터 분리된 세포로부터 수득된 DNA 샘플, 예를 들어, 점라인 및/또는 체세포 DNA를 함유하는 샘플, 예를 들어, 염색체 샘플을 환자의 세포 내의 하나 이상의 이상(예를 들어, 돌연변이)의 존재 또는 부재를 결정하기 위해 세포유전 이상(예를 들어, 본원에 기재된 ALK 돌연변이)과 관련된 염색체 영역 내의 유전체 DNA에 특이적이고 엄격한 조건하에서 이에 하이브리드화되는 폴리뉴클레오티드 프로브에 접촉시키는 것을 포함한다. 분석 결과는 치료제, 특히 ALK를 억제하는 작용제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료에 대해 환자가 반응할 가능성을 예측한다.
- [0177] 또 다른 구체예에서, 시간 경과를 환자의 질병의 경과에서 유의한 사건 사이의 시간을 결정함으로써 측정되며, 상기 측정은 환자가 긴 시간 경과를 갖는지의 여부를 예측한다. 또 다른 구체예에서, 유의한 사건은 일차 진단으로부터 사망까지의 진행이다. 또 다른 구체예에서, 유의한 사건은 일차 진단으로부터 전이성 질병까지의 진행이다. 또 다른 구체예에서, 유의한 사건은 일차 진단으로부터 재발까지의 진행이다. 또 다른 구체예에서, 유의한 사건은 전이성 질병으로부터 사망까지의 진행이다. 또 다른 구체예에서, 유의한 사건은 재발로부터 사망까지의 진행이다. 특정 구체예에서, 시간 경과는 전체 생존률, 진행까지의 시간 및/또는 RECIST의 이용 또는 다른 반응 기술과 관련하여 측정된다.
- [0178] 특정 구체예에서, 소정의 측정은 환자 샘플을 적어도 2개의 환자 서브그룹으로 나눔으로써 생성된다. 특정 구체예에서, 서브그룹의 수는 2개이며, 환자 샘플은 ALK 돌연변이(들)을 갖는 환자의 서브그룹 및 ALK 돌연변이(들)을 갖지 않는 서브그룹으로 나누어진다. 특정 구체예에서, 피검체에서의 ALK 돌연변이 상태는 ALK 돌연변이(들)을 갖거나 갖지 않는 서브그룹에 비교되고; 환자가 ALK 돌연변이(들)을 갖는 경우, 환자는 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)에 반응하지 않을 수 있고/있거나 환자는 긴 시간 경과를 가질 수 있다. 특정 구체예에서, 서브그룹의 수는 특정 ALK 돌연변이와 관련된 것과 같은 예측된 ALK 억제제 효능의 층화(stratification)에 따라 2개를 초과하며, 이는 3개의 서브그룹, 4개의 서브그룹, 5개의 서브그룹 및 6개의 서브그룹을 포함하나 이에 제한되지는 않는다. 특정 구체예에서, 반응 가능성은 전체 생존률, 진행까지의 시간

및/또는 RECIST 기준 이용과 관련하여 측정된다. 특정 구체예에서, ALK 억제제는 PF-02341066 및/또는 PDD이다.

[0179] 또 다른 양태에서, 본 발명은 ALK 돌연변이 양성 암을 갖는 피검체가 ALK 억제제(예를 들어, PF-02341066 및/또는 PDD)를 이용한 치료에 반응할 수 있는지의 여부 및/또는 질병의 시간 경과가 긴지의 여부를 결정하는 방법에 관한 것이다. 또 다른 양태에서, 본 발명은 ALK 돌연변이 양성 암을 갖는 피검체에서 질병의 시간 경과를 예측하는 방법에 관한 것이다. 또 다른 양태에서, 본 발명은 ALK 돌연변이 양성 암을 갖는 피검체에서 유의한 사건의 가능성을 예측하는 방법에 관한 것이다.

1. ALK 돌연변이를 검출하기 위한 방법

[0181] 하이브리드화-기반 검정을 포함하는 ALK 유전자 돌연변이 및/또는 유전자 생성물(예를 들어, 표 1에 나열된 마커)을 평가하는 방법은 당업자에게 널리 공지되어 있다. 예를 들어, 샘플 내의 엔코딩 핵산의 카피수를 평가하는 한 방법은 서던 블롯(Southern Blot)을 포함한다. 서던 블롯에서, 유전체 DNA(통상적으로, 전기영동 겔에서 단편화 및 분리됨)는 표적 영역에 특이적인 프로브에 하이브리드화된다. 표적 영역에 대한 프로브로부터의 하이브리드화 신호와 정상 유전체 DNA의 분석으로부터의 대조군 프로브 신호의 강도의 비교(예를 들어, 동일하거나 관련된 세포, 조직, 기관 등의 증폭되지 않은 부분)는 표적 핵산의 존재/부재 및 상대 카피수의 평가를 제공한다. 대안적으로, 노던 블롯은 샘플 내의 엔코딩 핵산의 카피수를 평가하기 위해 이용될 수 있다. 노던 블롯에서, mRNA는 표적 영역에 특이적인 프로브에 하이브리드화된다. 표적 영역에 대한 프로브로부터의 하이브리드화 신호와 정상 mRNA의 분석으로부터의 대조군 프로브 신호의 강도의 비교(예를 들어, 동일하거나 관련된 세포, 조직, 기관 등의 증폭되지 않은 부분)는 표적 핵산의 존재/부재 및 상대 카피수의 평가를 제공한다.

[0182] 카피수를 결정하기 위한 대안적 수단은 인 시츄(*in situ*) 하이브리드화이다(예를 들어, Angerer (1987) *Meth. Enzymol* 152: 649). 일반적으로, 인 시츄 하이브리드화는 하기 단계를 포함한다: (1) 분석되는 조직 또는 생물학적 구조의 고정; (2) 표적 DNA의 접근성을 증가시키고, 비특이적 결합을 감소시키기 위한 생물학적 구조의 예비하이브리드화 처리; (3) 생물학적 구조 또는 조직 내의 핵산에 대한 핵산의 혼합물의 하이브리드화; (4) 하이브리드화에서 발견되지 않는 핵산 단편을 제거하기 위한 하이브리드화후 세척; 및 (5) 하이브리드화된 핵산 단편의 검출. 사용하기 위한 상기 단계 및 조건 각각에서 사용되는 시약은 특정 적용에 따라 다양하다.

[0183] 예시적 하이브리드화-기반 검정은 전통적인 "직접 프로브" 방법, 예를 들어, 서던 블롯 또는 인 시츄 하이브리드화(예를 들어, FISH 및 FISH + SKY), 및 "비교 프로브" 방법, 예를 들어, 비교 유전체 하이브리드화(CGH), 예를 들어, cDNA-기반 또는 올리고뉴클레오타이드-기반 CGH를 포함하나, 이에 제한되지는 않는다. 상기 방법은 기관(예를 들어, 막 또는 유리) 결합 방법 또는 어레이-기반 방법을 포함하나 이에 제한되지는 않는 매우 다양한 포맷으로 사용될 수 있다.

[0184] 한 양태에서, FISH 분석이 사용된다. 세포 샘플은 당 분야에 공지된 적절한 세포유전 시험 방법, 예를 들어, FISH 방법에 의해 시험되기 위해 당 분야에 널리 공지된 방법에 따라 환자로부터 취득된다. 한 구체예에서, FISH는 Vysis™ 시스템(Abbott Molecular)에 따라 수행될 수 있고, 이의 제조업체의 프로토콜은 참조로서 본원에 포함된다.

[0185] 염색체의 다양한 부분에 존재하는 DNA 염기 서열에 본질적으로 상보적인 DNA 세그먼트를 함유하는 프로브가 사용된다. 본 발명에 따른 유용한 프로브, 표지화 및 샘플에 대한 프로브의 하이브리드화의 예는 Vysis, Inc.의 2개의 미국 특허(Bittner, et al.)인 미국 특허 번호 5,491,224호(참조로서 포함됨) 및 6,277,569호(참조로서 포함됨)에 기재되어 있다.

[0186] 염색체 프로브는 통상적으로 약 50 내지 약 10^5 개의 뉴클레오타이드 길이이다. 보다 긴 프로브는 통상적으로 약 100 내지 약 500개의 뉴클레오타이드 길이의 보다 작은 단편을 포함한다. 중심질 DNA 및 유전자좌-특이적 DNA와 하이브리드화되는 프로브는, 예를 들어, Vysis, Inc.(Downers Grove, Ill.), Molecular Probes, Inc.(Eugene, Oreg.) 또는 Cytocell(Oxfordshire, UK)로부터 시판된다. 대안적으로, 프로브는 표준 기술을 통해 염색체 또는 유전체 DNA로부터 비-상업적으로 제조될 수 있다. 예를 들어, 사용될 수 있는 DNA의 공급원은 숙주의 정상적인 염색체 상보체(complement)와 함께 하나의 염색체(예를 들어, 인간 염색체) 또는 하나의 염색체의 일부, 및 유세포측정 또는 미세해부에 의해 정제된 염색체를 함유하는 유전체 DNA, 클로닝된 DNA 서열, 체세포 하이브리드를 포함한다. 관심 영역은 클로닝을 통해 분리될 수 있거나, 중합효소 연쇄반응(PCR)을 통해 부위-특이적 증폭에 의해 분리될 수 있다. 예를 들어, 문헌[Nath and Johnson, *Biotechnic Histochem.*, 1998, 73(1):6-22, Wheelless *et al.*, *Cytometry* 1994, 17:319-326], 및 미국 특허 번호 5,491,224호(참조로서 포함됨)를

참조하라.

- [0187] 사용되는 프로브는 세포유전 이상이 상기 영역에 존재하는지의 여부를 결정하기 위해 염색체의 특정 영역에 하이브리드화된다. 세포유전 이상의 한 유형은 결실이다. 결실은 하나 이상의 전체 염색체에 존재할 수 있으나, 결실은 보통 하나 이상의 염색체의 일부의 상실을 포함한다. 프로브에 함유되는 염색체의 전체 영역이 세포로부터 결실되는 경우, 세포로부터의 DNA에 대한 상기 프로브의 하이브리드화는 보통 발생하지 않을 것이며, 상기 염색체에는 신호가 존재하지 않을 것이다. 프로브 내에 부분적으로 함유된 염색체의 영역이 세포로부터 결실되는 경우, 세포로부터의 DNA에 대한 프로브의 하이브리드화가 여전히 발생할 수 있으나, 보다 덜한 신호가 존재할 수 있다. 예를 들어, 신호의 상실은 프로브가 검출하려는 유전 이상을 함유하지 않는 대조군 세포로부터의 DNA에 대한 프로브 하이브리드화와 비교된다. 일부 구체예에서, 적어도 1개, 5개, 10개, 20개, 30개, 40개, 50개, 60개, 70개, 80개, 90개, 100개, 110개, 120개, 130개, 140개, 150개, 160개, 170개, 180개, 190개, 200개 또는 이 초과인 세포가 세포유전 이상의 존재에 대해 계수된다.
- [0188] 검출되는 세포유전 이상은 비-상호 전위(non-reciprocal translocation), 염색체내 역전, 점 돌연변이, 결실, 유전자 카피수 변화, 유전자 발현 수준 변화, 및 점라인 돌연변이를 포함할 수 있으나, 이에 제한되지는 않는다. 특히, 세포유전 이상의 한 유형은 중복이다. 중복은 전체 염색체에 존재할 수 있거나, 전체 염색체보다 작은 영역에 존재할 수 있다. 프로브에 함유된 염색체의 영역이 세포에서 중복되는 경우, 세포로부터의 DNA에 대한 상기 프로브의 하이브리드화는 보통 프로브 내에 함유된 염색체 영역의 이상을 갖지 않는 대조군 세포에 존재하는 신호의 수에 비해 적어도 하나의 추가 신호를 발생시킬 것이다. 인간 염색체 2p23 또는 이의 오솔로그(ortholog) 또는 2p23 또는 이의 오솔로그의 ALK 유전자와의 전위를 포함하는 임의의 염색체 영역을 검출하는 임의의 프로브가 이용될 수 있으나, 적합한 프로브는 당 분야에 널리 공지되어 있다(예를 들어, Vysis, Inc. (Downers Grove, IL)로부터 이용가능함).
- [0189] 염색체 프로브는 하이브리드화되는 염색체 영역이 검출될 수 있도록 표지화된다. 프로브는 통상적으로 보다 낮은 파장/보다 높은 에너지의 광을 흡수한 후에 형광을 발하는 유기 분자인 형광단으로 직접 표지된다. 형광단은 이차 검출 분자 없이 프로브가 시각화되도록 한다. 뉴클레오티드에 형광단을 공유적으로 부착시킨 후, 뉴클레오티드는 표준 기술, 예를 들어, 닉(nick) 번역, 무작위 프라이밍, 및 PCR 라벨링을 이용하여 프로브로 직접 통합될 수 있다. 대안적으로, 프로브 내의 테옥시시티딘 뉴클레오티드는 링커와 함께 아미노전이될 수 있다. 이후, 형광단은 아미노전이된 테옥시시티딘 뉴클레오티드에 공유적으로 부착된다. 예를 들어, 미국 특허 번호 5,491,224호(참조로서 포함됨)를 참조하라.
- [0190] 미국 특허 번호 5,491,224호에는 공유적으로 결합된 형광 표지를 갖는 다수의 시토신 잔기로서 표지되는 프로브가 기재되어 있다. 형광 표지된 시토신 염기의 수는 상기과 같이 DNA 세그먼트가 표지된 개체가 검출되는 염색체 또는 염색체 영역과 관련하여 이의 특이적 상보성 결합(하이브리드화) 특성을 본질적으로 유지하면서 검출가능한 형광 신호를 발생시키기에 충분하다. 이러한 프로브는 표지되지 않은 DNA 프로브 세그먼트를 취하고, 이러한 세그먼트 내의 다수의 테옥시시티딘 뉴클레오티드를 연결기와 함께 아미노전이지키고, 아미노전이된 테옥시시티딘 염기의 적어도 일부에 형광 표지를 공유적으로 결합시킴으로써 제조된다.
- [0191] 프로브는 또한 닉 번역, 무작위 프라이머 표지화 또는 PCR 표지화에 의해 표지될 수 있다. 표지화는 형광 (직접)-또는 합텐(haptene) (간접)-표지된 뉴클레오티드를 이용하여 수행된다. 표지의 대표적인 비제한적 예는 AMCA-6-dUTP, 캐스케이드블루(CascadeBlue)-4-dUTP, 플루오레세인(Fluorescein)-12-dUTP, 로다민(Rhodamine)-6-dUTP, 텍사스레드(TexasRed)-6-dUTP, Cy3-6-dUTP, Cy5-dUTP, 비오틴(Biotin)(BIO)-11-dUTP, 디곡시게닌(Digoxigenin)(DIG)-11-dUTP 또는 디니트로페닐 (DNP)-11-dUTP를 포함한다.
- [0192] 프로브는 또한 비오틴 또는 디곡시게닌으로 간접적으로 표지되거나, ^{32}P 및 ^3H 와 같은 방사성 동위원소로 표지될 수 있으나, 프로브를 시각화시키기 위해 이차 검출 분자 또는 추가 처리가 요구된다. 예를 들어, 비오틴으로 표지된 프로브는 검출가능한 마커에 컨쥬게이션된 아비딘에 의해 검출될 수 있다. 예를 들어, 아비딘은 효소 마커, 예를 들어, 알칼리성 포스파타제 또는 호스라디쉬 퍼옥시다제에 컨쥬게이션될 수 있다. 효소 마커는 효소에 대한 기질 및/또는 촉매를 이용한 표준 비색 반응으로 검출될 수 있다. 알칼리성 포스파타제에 대한 촉매는 5-브로모-4-클로로-3-인돌릴포스파타제 및 니트로 블루 테트라졸륨을 포함한다. 호스라디쉬 퍼옥시다제에 대한 촉매로서 디아미노벤조에이트가 사용될 수 있다.
- [0193] 프로브는 또한 형광 또는 다른 표지가 하이브리드화 전 또는 하이브리드화 동안 DNA의 일부가 아니고, 염색체에 하이브리드화된 프로브를 검출하기 위해 하이브리드화 후에 첨가되도록 제조될 수 있다. 예를 들어, DNA로 통합된 항원 분자를 갖는 프로브가 이용될 수 있다. 하이브리드화 후, 이러한 항원 분자는 항원 분자와 반응하는

특이적 항체를 이용하여 검출된다. 이러한 항체는 자신이 플루오로크롬을 포함할 수 있거나, 결합된 플루오로크롬을 갖는 이차 항체를 이용하여 검출될 수 있다.

- [0194] 그러나, 처리되거나 변형되는 프로브 DNA는 하이브리드화에서 사용하기 전에 반응되지 않은 잔여 생성물(예를 들어, DNA에 통합되지 않은 플루오로크롬 분자)을 제거하기 위해 보통 정제된다.
- [0195] 하이브리드화 전, 염색체 프로브는 당 분야에 널리 공지된 방법에 따라 변성된다. 일반적으로, 하이브리드화 단계는 과량의 블로킹 DNA를 표지된 프로브 조성물에 첨가하는 단계, 하이브리드화 조건하에서 블로킹된 프로브 조성물을, 예를 들어, DNA가 변성된 슬라이드 상의 검출되는 염색체 영역과 접촉시키는 단계, 하이브리드화되지 않은 프로브를 세척하는 단계, 및 염색체 또는 염색체 영역에 대한 프로브 조성물의 결합을 검출하는 단계를 포함한다.
- [0196] 프로브는 하이브리드화 조건하에서 염색체 DNA에 하이브리드화되거나 어닐링된다. "하이브리드화 조건"은 프로브와 표적 염색체 DNA 사이의 어닐링을 촉진시키는 조건이다. 다양한 프로브의 어닐링은 프로브 길이, 염기 농도 등에 따라 다양할 것이므로, 어닐링은 프로브 농도, 하이브리드화 온도, 염 농도 및 당 분야에 널리 공지된 다른 요인을 다양화시킴으로써 촉진된다.
- [0197] 하이브리드화 조건은 프로브의 농도, 염기 조성, 복잡도 및 길이 뿐만 아니라 염 농도, 온도, 및 인큐베이션 길이를 다양화시킴으로써 촉진된다. 예를 들어, 인 시츄 하이브리드화는 통상적으로 1-2x SSC, 50-65% 포름아미드 및 비-특이적 하이브리드화를 억제하기 위한 블로킹 DNA를 함유하는 하이브리드화 완충액에서 수행된다. 일반적으로, 상기 기재된 바와 같은 하이브리드화 조건은 약 25℃ 내지 약 55℃의 온도, 및 약 0.5시간 내지 약 96시간의 인큐베이션 길이를 포함한다.
- [0198] 표적 영역 외부의 DNA에 대한 염색체 프로브의 비특이적 결합은 일련의 세척에 의해 제거될 수 있다. 각 세척에서의 온도 및 염의 농도는 세척의 엄격성을 조절하기 위해 다양화된다. 예를 들어, 높은 엄격성 조건에 대해, 세척은 0.2x 내지 약 2x SSC, 및 약 0.1% 내지 약 1%의 비-이온성 세제, 예를 들어, Nonidet P-40(NP40)을 이용하여 약 65℃ 내지 약 80℃에서 수행될 수 있다. 엄격성은 세척의 온도를 감소시키거나, 세척에서의 염의 농도를 증가시킴으로써 저하될 수 있다. 일부 적용에서, 반복 서열의 하이브리드화 능력을 블로킹시키는 것이 필요하다. 따라서, 일부 구체예에서, tRNA, 인간 유전체 DNA, 또는 Cot-I DNA가 비-특이적 하이브리드화를 블로킹시키기 위해 사용된다.
- [0199] 세척 후, 슬라이드는 배수 및 공기 건조된 후, 배지를 마운팅시키고, DAPI와 같은 대조염색을 수행하고, 커버슬립이 슬라이드에 적용된다. 슬라이드는 즉시 관찰되거나, 시험 전에 -20℃에서 저장될 수 있다.
- [0200] 형광 인 시츄 하이브리드화(FISH) 기술에서 사용되는 형광 프로브에 대해, 형광은 각각의 형광단에 대해 적절한 필터가 장비된 형광 현미경으로 관찰될 수 있거나, 다수의 형광단을 관찰하기 위해 이중 또는 삼중 밴드-패스(band-pass) 필터 세트를 이용하여 관찰될 수 있다. 예를 들어, 미국 특허 번호 5,776,688호(참조로서 포함됨)를 참조하라. 대안적으로, 염색체 프로브의 하이브리드화 패턴을 시험하기 위해 유세포측정과 같은 기술이 이용될 수 있다. 염색체 카피수 또는 염색체 영역의 재배열을 검출하기 위해 FISH가 이용될 수 있다. 이러한 프로브는 상보적 DNA에 하이브리드화되거나 결합하며, 이들은 형광 태그로 표지되므로, 연구자들이 형광 현미경을 이용하여 DNA의 서열의 위치를 확인하는 것이 가능해진다. 세포가 활발히 분열되는 것을 필요로 하는 염색체를 연구하는데 이용되는 대부분의 다른 기술과는 달리, FISH는 또한 분열하지 않는 세포에서 수행될 수 있어, 매우 용도가 넓은 절차로 만든다. 따라서, FISH는 간기 세포, 또는 세포 분열 주기 중 중기의 세포를 이용하여 수행될 수 있다. FISH 분석과 관련된 기술 중 많은 기술은 미국 특허 번호 5,447,841호(참조로서 포함됨)(Gray and Pinkel)에 기재되어 있다.
- [0201] FISH 결과는 프로브가 검출하도록 디자인된 특정 세포유전 이상을 함유하지 않는 것으로 공지된 대조군 세포를 참조로 하여 해석될 수 있다. 대조군 세포로부터의 DNA에 대한 프로브의 FISH 하이브리드화 패턴은 특이적 세포유전 이상에 대해 시험되거나 검정되는 세포로부터의 DNA에 대한 동일 프로브의 하이브리드화와 비교된다. 프로브가 염색체 또는 염색체 영역의 결실을 검출하기 위해 디자인되는 경우, 대조군 세포로부터의 DNA보다 시험되는 세포로부터의 DNA에 대한 프로브의 보다 덜한 하이브리드화가 보통 존재한다. 보통, 시험되는 세포에서 프로브 신호가 부재하며, 이는 프로브가 보통 하이브리드화되는 염색체의 영역의 상실을 나타낸다. 프로브가 염색체 중복 또는 첨가를 검출하기 위해 디자인되는 경우, 대조군 세포로부터의 DNA보다 시험되는 세포로부터의 DNA에 대한 프로브의 보다 많은 하이브리드화가 보통 존재한다. 보통, 시험되는 세포에서의 프로브 신호의 추가가 존재하며, 이는 프로브가 보통 하이브리드화되는 추가의 염색체 영역의 존재를 나타낸다.

- [0202] CGH 방법에서, 핵산의 첫번째 수집물(예를 들어, 샘플, 예를 들어, 가능한 종양으로부터의 수집물)은 첫번째 표지로 표시되고, 핵산의 두번째 수집물(예를 들어, 대조군, 예를 들어, 건강한 세포/조직으로부터의 수집물)은 두번째 표지로 표시된다. 핵산의 하이브리드화의 비는 어레이 내의 각각의 섬유에 결합하는 2개(첫번째 및 두번째) 표지의 비에 의해 결정된다. 염색체 결실 또는 증가가 존재하는 경우, 2개의 표지로부터의 신호의 비에서 차이가 검출될 것이고, 상기 비는 카피수의 측정을 제공할 것이다. 어레이-기반 CGH는 또한 단색 표지화로 수행될 수 있다(2개의 상이한 염료를 이용하여 대조군 및 가능한 종양 샘플을 표지화시키고, 이들을 하이브리드화 전에 혼합시켜, 어레이에서의 프로브의 경쟁적 하이브리드화로 인해 비를 발생시키는 것과 반대됨). 단색 CGH에서, 대조군은 표지되고, 하나의 어레이에 하이브리드화되고, 절대 신호가 판독되며, 가능한 종양 샘플은 표지되고, 이차 어레이(동일한 내용물을 가짐)에 하이브리드화되고, 절대 신호가 판독된다. 카피수 차이는 2개의 어레이로부터의 절대 신호를 기초로 하여 계산된다. 본 발명의 방법과 함께 사용하기에 적합한 하이브리드화 프로토콜은, 예를 들어, 문헌[Albertson (1984) *EMBO J.* 3:1227-1234; Pinkel (1988) *Proc. Natl. Acad. Sci. USA* 85: 9138-9142; EPO Pub. No. 430,402; *Methods in Molecular Biology*, Vol. 33: *In situ Hybridization Protocols*, Choo, ed., Humana Press, Totowa, N.J. (1994) 등]에 기재되어 있다. 한 구체예에서, 문헌[Pinkel, *et al.* (1998) *Nature Genetics* 20:207-211] 또는 문헌[Kallioniemi (1992) *Proc. Natl. Acad. Sci. USA* 89:5321-5325 (1992)]의 하이브리드화 프로토콜이 사용된다. 어레이-기반 CGH는 각각의 전체내용이 참조로서 본원에 포함되는 미국 특허 번호 6,455,258호에 기재되어 있다.
- [0203] 또 다른 구체예에서, 증폭-기반 검정은 존재/부재 및 카피수를 측정하는데 이용될 수 있다. 이러한 증폭-기반 검정에서, 핵산 서열은 증폭 반응(예를 들어, 중합효소 연쇄반응(PCR)에서 주형으로 작용한다. 정량적 증폭에서, 증폭 생성물의 양은 본래 샘플 내의 주형의 양에 비례될 것이다. 적절한 대조군, 예를 들어, 건강한 조직에 대한 비교는 카피수의 측정을 제공한다.
- [0204] "정량적" 증폭 방법은 당업자에게 널리 공지되어 있다. 예를 들어, 정량적 PCR은 동일 프라이머를 이용하여 공지된 양의 대조군 서열을 동시에 공동 증폭시키는 것을 포함한다. 이는 PCR 반응을 교정하는데 사용될 수 있는 내부 표준을 제공한다. 정량적 PCR에 대한 상세한 프로토콜은 문헌[Innis, *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.]에 제공된다. 정량적 PCR 분석을 이용한 마이크로새틀라이트(microsatellite) 유전자좌에서의 DNA 카피수의 측정은 문헌[Ginzinger, *et al.* (2000) *Cancer Research* 60:5405-5409]에 기재되어 있다. 유전자에 대한 공지된 핵산 서열은 당업자가 유전자의 임의의 부분을 증폭시키기 위해 프라이머를 통상적으로 선택하는 것을 가능케 하기에 충분하다. 본 발명의 방법에서 플루오로게닉(Fluorogenic) 정량적 PCR이 또한 이용될 수 있다. 플루오로게닉 정량적 PCR에서, 정량은 형광 신호, 예를 들어, TaqMan 및 sybr 그린(green)의 양을 기초로 한다.
- [0205] 다른 적합한 증폭 방법은 리가제 연쇄반응(LCR)(Wu and Wallace (1989) *Genomics* 4:560, Landegren, *et al.* (1988) *Science* 241:1077, 및 Barringer *et al.* (1990) *Gene* 89:117 참조), 전사 증폭(Kwoh, *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173), 자기부양 염기서열 복제반응(Guatelli, *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874), 도트 PCR, 및 링커 어댑터 PCR 등을 포함하나, 이에 제한되지는 않는다.
- [0206] 이형접합 상실(LOH) 맵핑(Wang, Z.C., *et al.* (2004) *Cancer Res* 64(1):64-71; Seymour, A. B., *et al.* (1994) *Cancer Res* 54, 2761-4; Hahn, S. A., *et al.* (1995) *Cancer Res* 55, 4670-5; Kimura, M., *et al.* (1996) *Genes Chromosomes Cancer* 17, 88-93)이 또한 증폭 또는 결실의 영역을 확인하는데 사용될 수 있다.
- [0207] **2. 유전자 발현을 평가하기 위한 방법**
- [0208] 마커 발현 수준이 또한 검정될 수 있다. 본 발명의 마커의 발현은 전사된 분자 또는 단백질의 발현을 검출하는 매우 다양한 널리 공지된 방법 중 임의의 방법에 의해 평가될 수 있다. 이러한 방법의 비제한적인 예는 분비된 단백질, 세포-표면 단백질, 세포질 단백질, 또는 핵 단백질의 검출을 위한 면역학적 방법, 단백질 정제 방법, 단백질 기능 또는 활성 검정, 핵산 하이브리드화 방법, 핵산 역전사 방법, 및 핵산 증폭 방법을 포함한다.
- [0209] 특정 구체예에서, 특정 유전자의 활성은 유전 전사체(예를 들어, mRNA)의 측정, 번역된 단백질 양의 측정, 또는 유전자 생성물 활성의 측정에 의해 특정구명된다. 마커 발현은 mRNA 수준, 단백질 수준, 또는 단백질 활성을 검출하는 것을 포함하는 다양한 방식으로 모니터링될 수 있으며, 이 중 임의의 것은 표준 기술을 이용하여 측정될 수 있다. 검출은 유전자 발현(예를 들어, 유전체 DNA, cDNA, mRNA, 단백질, 또는 효소 활성)의 수준의 정량을 포함할 수 있거나, 대안적으로, 특히 대조군 수준과의 비교에서의 유전자 발현의 수준의 정성적 평가일 수 있다. 검출되는 수준의 유형은 상황으로부터 명백할 것이다.

- [0210] 핵산 하이브리드화 기술을 이용하여 유전자 전사체(mRNA 또는 이로부터 제조된 cDNA)를 검출하고/하거나 정량하는 방법은 당업자에게 공지되어 있다(문헌[Sambrook *et al.* 상기] 참조). 예를 들어, cDNA의 존재, 부재, 또는 양을 평가하기 위한 한 방법은 상기 기재된 바와 같은 서던(Southern) 전달을 포함한다. 간단히, mRNA는 분리(예를 들어, 문헌[Sambrook *et al.* 상기]의 산 구아니디늄-페놀-클로로포름 추출 방법을 이용함)되고, 역전사되어 cDNA가 생성된다. cDNA는 이후 임의로 절단되고, 완충액 중에서 겔 상에서 영동되고, 막으로 옮겨진다. 이후, 하이브리드화는 표적 cDNA에 특이적인 핵산 프로브를 이용하여 수행된다.
- [0211] 상기 진단 및 예후 검정의 일반적 원리는 마커 및 프로브가 상호작용하고 결합함으로써 반응 혼합물에서 분리되고/되거나 검출될 수 있는 복합체를 형성시키기에 적절한 조건 하 및 충분한 시간 동안, 마커 및 프로브를 함유할 수 있는 샘플 또는 반응 혼합물을 제조하는 것을 포함한다. 이러한 검정은 다양한 방식으로 수행될 수 있다.
- [0212] 예를 들어, 상기 검정을 수행하는 한 방법은 마커 또는 프로브를 기관으로도 언급되는 고체상 지지체에 부착시키고, 반응 종료시 고체상에 부착된 표적 마커/프로브 복합체를 검출하는 것을 포함할 것이다. 이러한 방법의 한 구체예에서, 마커의 존재 및/또는 농도에 대해 검정되는 피검체로부터의 샘플이 담체 또는 고체상 지지체에 부착될 수 있다. 또 다른 구체예에서, 프로브가 고체상에 부착될 수 있고, 피검체로부터의 샘플이 검정의 부착되지 않은 성분으로서 반응될 수 있는 반대 상황이 가능하다.
- [0213] 검정 성분을 고체상에 부착시키는 많은 확립된 방법이 존재한다. 이는 비오틴 및 스트렙타비딘의 컨쥬게이션을 통해 고정되는 마커 또는 프로브 분자를 포함하나, 이에 제한되지는 않는다. 이러한 비오틴화된 검정 성분은 당 분야에 공지된 기술(예를 들어, 비오틴화 키트, Pierce Chemicals, Rockford, IL)을 이용하여 비오틴-NHS(N-히드록시-숙신이미드)로부터 제조될 수 있고, 이는 스트렙타비딘-코팅된 96 웰 플레이트(Pierce Chemical)의 웰에 고정된다. 특정 구체예에서, 고정된 검정 성분을 갖는 표면은 미리 제조되거나 저장될 수 있다.
- [0214] 상기 검정에 대한 다른 적합한 담체 또는 고체상 지지체는 마커 또는 프로브가 속하는 분자 부류에 결합할 수 있는 임의의 물질을 포함한다. 널리 공지된 지지체 또는 담체는 유리, 폴리스티렌, 나일론, 폴리프로필렌, 폴리에틸렌, 텍스트란, 아밀라제, 자연 및 변형 셀룰로오스, 폴리아크릴아미드, 반려암, 및 자철광을 포함하나, 이에 제한되지는 않는다.
- [0215] 상기 언급된 방법으로 검정을 수행하기 위해, 이차 성분이 부착된 후에 고정되지 않은 성분이 고체상에 첨가된다. 반응이 완료된 후, 복합체화되지 않은 성분은, 형성된 임의의 복합체가 고체상에 고정된 채로 유지되는 조건하에서 제거(예를 들어, 세척에 의함)될 수 있다. 고체상에 부착된 마커/프로브 복합체의 검출은 본원에 개시된 다수의 방법으로 달성될 수 있다.
- [0216] 또 다른 구체예에서, 부착되지 않은 검정 성분인 경우 프로브는 본원에 논의된 검출가능한 표지를 직접적 또는 간접적으로 가지며 당업자에게 널리 공지된 검정의 검출 및 판독 목적으로 위해 표지될 수 있다.
- [0217] 예를 들어, 형광 에너지 전달 기술을 이용함으로써 성분(마커 또는 프로브)의 추가 조작 또는 표지화 없이 마커/프로브 복합체 형성을 직접적으로 검출하는 것이 또한 가능하다(예를 들어, Lakowicz *et al.*, 미국 특허 번호 5,631,169호(참조로서 포함됨); Stavrianopoulos, *et al.*, 미국 특허 번호 4,868,103호(참조로서 포함됨) 참조). 첫번째 '공여자' 분자 상의 형광단 표지는 적절한 파장의 입사광을 이용한 여기 시에 이의 방출된 형광 에너지가 두번째 '수용자' 분자 상의 형광 표지에 의해 흡수되고, 차례로 흡수된 에너지로 인해 형광을 발할 수 있도록 선택된다. 대안적으로, '공여자' 단백질 분자는 단순히 트립토판 잔기의 자연 형광 에너지를 이용할 수 있다. 광의 다양한 파장을 방출하는 표지는, '수용자' 분자 표지가 '공여자'의 표지와 구별될 수 있도록 선택된다. 표지 사이의 에너지 전달의 효율은 분자 사이에 떨어진 거리와 관련이 있으므로, 분자 사이의 공간적 관계가 평가될 수 있다. 분자 사이에 결합이 발생하는 상황에서, 검정에서 '수용자' 분자 표지의 형광 방출은 최대가 되어야 한다. FET 결합 사건은 편리하게는 당 분야에 널리 공지된 표준 형광측정 검출 수단(예를 들어, 형광계 이용)을 통해 측정될 수 있다.
- [0218] 또 다른 구체예에서, 마커를 인지하는 프로브 능력의 결정은 실시간 생체분자 상호작용 분석(Biomolecular Interaction Analysis, BIA)과 같은 기술을 이용함으로써 검정 성분(프로브 또는 마커)을 이용한 표지화 없이 달성될 수 있다(예를 들어, Sjolander, S. and Urbaniczky, C, 1991, *Anal. Chem.* 63:2338-2345 및 Szabo *et al.*, 1995, *Curr. Opin. Struct. Biol.* 5:699-705 참조). 본원에서 사용되는 용어 "BIA" 또는 "표면 플라즈몬 공명"은 상호작용체(예를 들어, BIAcore) 중 임의의 것을 표지화시키지 않고 실시간으로 생체특이적 상호작용을

연구하는 기술이다. 결합 표면에서의 질량의 변화(결합 사건을 나타냄)는 표면 근처에서의 광의 굴절률을 변경(표면 플라즈몬 공명(SPR)의 광학 현상)시키고, 이는 생물학적 분자 사이의 실시간 반응의 표시로 이용될 수 있는 검출가능한 신호를 발생시킨다.

[0219] 대안적으로, 또 다른 구체예에서, 유사한 진단 및 예후 검정이 액체상의 용질로서의 마커 및 프로브로 수행될 수 있다. 이러한 검정에서, 복합체화된 마커 및 프로브는 차별적 원심분리, 크로마토그래피, 전기영동 및 면역침전법을 포함하나 이에 제한되지는 않는 다수의 표준 기술 중 임의의 기술에 의해 복합체화되지 않은 성분으로부터 분리된다. 차별적 원심분리에서, 마커/프로브 복합체는 복합체의 상이한 크기 및 밀도를 기초로 하여 복합체의 상이한 침강 평형으로 인해 일련의 원심분리 단계를 통해 복합체화되지 않은 검정 성분으로부터 분리될 수 있다(예를 들어, Rivas, G., and Minton, A.P., 1993, *Trends Biochem Sci.* 18(8):284-7). 복합체화되지 않은 것으로부터 복합체화된 분자를 분리시키기 위해 표준 크로마토그래피 기술이 또한 이용될 수 있다. 예를 들어, 겔 여과 크로마토그래피는 크기를 기초로 하고, 컬럼 형식의 적절한 겔 여과 수지의 이용을 통해 분자를 분리시키고, 예를 들어, 상대적으로 더 큰 복합체는 비교적 더 작은 복합체화되지 않은 성분으로부터 분리될 수 있다. 유사하게, 복합체화되지 않은 성분에 비해 마커/프로브 복합체의 상대적으로 상이한 전하 특성은, 예를 들어, 이온-교환 크로마토그래피 수지의 이용을 통해 복합체화되지 않은 성분으로부터 복합체를 구별하는데 이용될 수 있다. 이러한 수치 및 크로마토그래피 기술은 당업자에게 널리 공지되어 있다(예를 들어, Heegaard, N.H., 1998, *J. Mol. Recognit.* Winter 11(1-6): 141-8; Hage, D.S., 및 Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 Oct 10;699(1-2):499-525 참조). 겔 전기영동은 또한 결합되지 않은 성분으로부터 복합체화된 검정 성분을 분리시키는데 이용될 수 있다(예를 들어, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987-1999 참조). 이러한 기술에서, 단백질 또는 핵산 복합체는, 예를 들어, 크기 또는 전하를 기초로 하여 분리된다. 전기영동 과정 동안 결합 상호작용을 유지시키기 위해, 완원제의 부재하에서 비-변성 겔 매트릭스 물질 및 조건이 통상적이다. 특정 검정에 대한 적절한 조건 및 이의 성분은 당업자에게 널리 공지될 것이다.

[0220] 한 특정 구체예에서, 마커에 상응하는 mRNA의 수준은 당 분야에 공지된 방법을 이용하여 생물학적 샘플 내에서 인 시추 및 시험관내 형식 둘 모두에 의해 결정될 수 있다. 용어 "생물학적 샘플"은 피검체로부터 분리된 조직, 세포, 생물학적 유체 및 이의 분리물 뿐만 아니라 피검체 내에 존재하는 조직, 세포 및 유체를 포함한다. 많은 발현 검출 방법은 분리된 RNA를 이용한다. 생체내 방법을 위해, mRNA의 분리에 대해 선택하지 않는 임의의 RNA 분리 기술이 세포로부터 RNA의 정제에 이용될 수 있다(예를 들어, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York 1987-1999 참조). 추가로, 많은 수의 조직 샘플이 당업자에게 널리 공지된 기술, 예를 들어, 문헌[Chomczynski (1989, U.S. Patent No. 4,843,155 (참조로서 포함됨))]의 1-단계 RNA 분리 방법을 이용하여 용이하게 처리될 수 있다.

[0221] 분리된 핵산은 서던 또는 노던 분석, 중합효소 연쇄반응 분석 및 프로브 어레이를 포함하나 이에 제한되지는 않는 하이브리드화 또는 증폭 검정에 사용될 수 있다. mRNA 수준의 검출을 위한 한 진단 방법은 분리된 mRNA와 검출되는 유전자에 의해 엔코딩된 mRNA에 하이브리드화될 수 있는 핵산 분자(프로브)를 접촉시키는 것을 포함한다. 핵산 프로브는, 예를 들어, 전장 cDNA 또는 이의 일부, 예를 들어, 적어도 7개, 15개, 30개, 50개, 100개, 250개 또는 500개 뉴클레오타이드 길이이며, 엄격한 조건하에서 본 발명의 마커를 엔코딩하는 mRNA 또는 유전체 DNA에 특이적으로 하이브리드화되기에 충분한 올리고뉴클레오타이드일 수 있다. 본 발명의 진단 검정에 사용하기 위한 다른 적합한 프로브가 본원에 기재된다. mRNA와 프로브의 하이브리드화는 당해 마커가 발현되는 것을 나타낸다.

[0222] 한 형식에서, 예를 들어, 아가로오스 겔 상에서 분리된 mRNA를 영동시키고, 겔로부터의 mRNA를 막, 예를 들어, 니트로셀룰로오스에 옮김으로써 mRNA는 고체 표면 상에 고정되고, 프로브와 접촉된다. 한 대안적 형식에서, 프로브(들)는 고체 표면 상에 고정되고, mRNA가, 예를 들어, Affymetrix 유전자 칩 어레이 중의 프로브(들)와 접촉된다. 당업자는 본 발명의 마커에 의해 엔코딩되는 mRNA의 수준을 검출하는데 사용하기 위한 공지된 mRNA 검출 방법을 용이하게 적합화시킬 수 있다.

[0223] 프로브는 단백질을 엔코딩하는 핵산 서열의 전장이거나 전장 미만일 수 있다. 보다 짧은 프로브는 특이성에 대해 경험적으로 시험된다. 예시적 핵산 프로브는 20개 또는 이를 초과하는 염기 길이이다(예를 들어, 핵산 하이브리드화에서 사용하기 위한 핵산 프로브 서열을 선택하는 방법에 대해 문헌[Sambrook *et al.*] 참조). 하이브리드화된 부분의 시각화는 cDNA의 존재 또는 부재의 정성적 결정을 가능케 한다.

[0224] 샘플 내의 본 발명의 마커에 상응하는 전사체의 수준을 결정하기 위한 대안적 방법은, 예를 들어, rtPCR(문헌

[Mullis, 1987], 미국 특허 번호 4,683,202호(참조로서 포함됨)에 나열된 실험 구체예), 리가제 사슬 반응(Barany, 1991, *Proc. Natl. Acad. Sci. USA*, 88:189-193), 자기부양 염기서열 복제반응(Guatelli *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), 전사 증폭 시스템(Kwoh *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-베타 레플리카제(Lizardi *et al.*, 1988, *Bio/Technology* 6: 1197), 회전 환 복제(rolling circle replication)(Lizardi *et al.*, 미국 특허 번호 5,854,033호(참조로서 포함됨)) 또는 임의의 다른 핵산 증폭 방법에 의한 핵산 증폭 과정 후, 당업자에게 널리 공지된 기술을 이용한 증폭된 분자의 검출을 포함한다. 플루오로게닉 rtPCR이 또한 본 발명의 방법에 사용될 수 있다. 플루오로게닉 rtPCR에서, 정량은 TaqMan 및 sybr 그린과 같은 형광 신호의 양을 기초로 한다. 이러한 검출 계획은 핵산 분자가 매우 적은 수로 존재하는 경우에 상기 핵산 분자의 검출에 특히 유용하다. 본원에서 사용되는 바와 같은 증폭 프라이머는 유전자의 5' 또는 3' 영역(플러스 및 마이너스 가닥 각각, 또는 반대)으로 어닐링될 수 있고, 이들 사이의 짧은 영역을 함유할 수 있는 핵산 분자의 쌍으로 정의된다. 일반적으로, 증폭 프라이머는 약 10 내지 30개의 뉴클레오타이드 길이이고, 약 50 내지 200개의 뉴클레오타이드 길이의 영역에 플랭킹된다. 적절한 조건하에서 및 적절한 시약을 이용하여, 상기 프라이머는 상기 프라이머에 의해 플랭킹된 뉴클레오타이드 서열을 포함하는 핵산 분자의 증폭을 가능케 한다.

[0225] 인 시추 방법에 대해, mRNA는 검출 전에 세포로부터 분리될 필요가 없다. 이러한 방법에서, 세포 또는 조직 샘플은 공지된 조직학적 방법을 이용하여 제조되고/처리된다. 이후, 샘플은 지지체, 통상적으로 유리 슬라이드 상에 고정된 후, 마커를 엔코딩하는 mRNA에 하이브리드화될 수 있는 프로브와 접촉된다.

[0226] 마커의 절대 발현 수준을 기초로 하여 결정을 하는 것에 대한 대안으로서, 마커의 표준화된 발현 수준을 기초로 하여 결정이 이루어질 수 있다. 발현 수준은 마커의 발현을 마커가 아닌 유전자, 예를 들어, 항시적으로 발현되는 하우스키핑(housekeeping) 유전자의 발현과 비교함으로써 마커의 절대 발현 수준을 교정함으로써 표준화된 다. 표준화에 적합한 유전자는 하우스키핑 유전자, 예를 들어, 액틴 유전자, 또는 상피세포-특이적 유전자를 포함한다. 이러한 표준화는 하나의 샘플, 예를 들어, 피검체 샘플에서의 또 다른 샘플, 예를 들어, 비-암성 샘플에 대한 발현 수준의 비교, 또는 상이한 공급원으로부터의 샘플 사이의 발현 수준의 비교를 가능케 한다.

[0227] 대안적으로, 발현 수준은 상대 발현 수준으로 제공될 수 있다. 마커의 상대 발현 수준을 결정하기 위해, 당해 샘플에 대한 발현 수준의 결정 전에 정상 분리물 대 암 세포 분리물의 10개 이상의 샘플, 또는 심지어 50개 이상의 샘플에 대해 마커의 발현 수준이 결정된다. 보다 많은 수의 샘플에서 검정된 유전자 각각의 평균 발현 수준이 결정되고, 이는 마커의 기준선 발현 수준으로 사용된다. 시험 샘플에 대해 결정된 마커의 발현 수준(절대 발현 수준)은 이후 상기 마커에 대해 수득된 평균 발현 값에 의해 나누어진다. 이는 상대 발현 수준을 제공한다.

[0228] 특정 구체예에서, 기준선 결정에 사용되는 샘플은 동일 조직 유형의 암세포 또는 정상 세포로부터 유래될 것이다. 세포 공급원의 선택은 상대 발현 수준의 이용에 좌우된다. 평균 발현 스코어로서 정상 조직에서 발견되는 발현을 이용하는 것은 검정되는 마커가 세포가 유래되는 조직(정상 세포에 비함)에 특이적인지의 여부를 확인하는 것을 돕는다. 또한, 더욱 많은 데이터가 축적됨에 따라, 평균 발현 값은 교정되어, 축적된 데이터를 기초로 한 개선된 상대 발현 값을 제공할 수 있다. 정상 세포로부터의 발현 데이터는 암 상태의 중증도를 등급화시키기 위한 수단을 제공한다.

[0229] 또 다른 구체예에서, 마커의 발현은 피검체 샘플 내의 세포로부터 유전체 DNA 또는 mRNA/cDNA(즉, 전사된 폴리뉴클레오타이드)를 제조하고, 유전체 DNA 또는 mRNA/cDNA와 마커 및 이의 단편을 포함하는 뉴클레오타이드의 상보체인 참조 폴리뉴클레오타이드를 하이브리드화시킴으로서 평가된다. cDNA는 임의로 참조 폴리뉴클레오타이드와의 하이브리드화 전에 다양한 중합효소 연쇄반응 중 임의의 중합효소 연쇄반응을 이용하여 증폭될 수 있다. 하나 이상의 마커의 발현은 마찬가지로 마커(들)의 발현 수준을 평가하기 위해 정량적 PCR(QPCR)을 이용하여 검출될 수 있다. 대안적으로, 본 발명의 마커의 돌연변이 또는 변이체(예를 들어, 단일 뉴클레오타이드 다형성, 결실)를 검출하기 위한 많은 공지된 방법 중 임의의 방법이 피검체에서 돌연변이된 마커의 발생을 검출하는데 사용될 수 있다.

[0230] 한 관련 구체예에서, 샘플로부터 수득된 전사된 폴리뉴클레오타이드의 혼합물은 기관에 고정된 본 발명의 마커의 적어도 일부(예를 들어, 적어도 7개, 적어도 10개, 적어도 15개, 적어도 20개, 적어도 25개, 적어도 30개, 적어도 40개, 적어도 50개, 적어도 100개, 적어도 500개 또는 이 초과 뉴클레오타이드 잔기)에 상보적이거나 이에 상동성인 폴리뉴클레오타이드를 갖는 기관과 접촉된다. 본 발명의 마커에 상보적이거나 이에 상동성인 폴리뉴클레오타이드가 기관 상에서 차별적으로 검출가능(예를 들어, 상이한 발색단 또는 형광단을 이용하거나, 상이한 선

택 위치에 고정되어 검출가능함)한 경우, 다수의 마커의 발현 수준은 단일 기관(예를 들어, 선택 위치에서 고정된 폴리뉴클레오티드의 "유전자 칩" 마이크로어레이)을 이용하여 동시에 평가될 수 있다. 하나이 핵산과 또 다른 핵산의 하이브리드화를 포함하는 마커 발현을 평가하는 방법이 사용되는 경우, 하이브리드화는 엄격한 하이브리드화 조건하에서 수행될 수 있다.

[0231] 또 다른 구체예에서, 마커의 발현을 평가하기 위한 방법의 조합이 이용된다.

[0232] 본 발명의 조성물, 키트, 및 방법은 본 발명의 하나 이상의 마커의 발현 수준 또는 카피수에서의 차이의 검출에 의존하므로, 특정 구체예에서, 마커의 발현 수준 또는 카피수는 정상 세포 및 암성 세포 중 적어도 하나에서의 발현 또는 카피수를 평가하기 위해 사용되는 방법의 최소 검출 한도보다 현저하게 크다.

[0233] 3. 발현된 단백질을 평가하기 위한 방법

[0234] 마커 단백질의 활성 또는 수준은 또한 발현된 폴리펩티드를 검출하거나 정량함으로써 검출되고/되거나 정량될 수 있다. 폴리펩티드는 당업자에게 널리 공지된 다수의 수단 중 임의의 수단에 의해 검출되고 정량될 수 있다. 이들은 분석 생화학 방법, 예를 들어, 전기영동, 모세관 전기영동, 고성능 액체 크로마토그래피(HPLC), 박층 크로마토그래피(TLC), 과확산(hyperdiffusion) 크로마토그래피 등, 또는 다양한 면역학적 방법, 예를 들어, 액체 또는 겔 침전소 반응, 면역확산(단일 또는 이중), 면역전기영동, 방사면역측정법(RIA), 효소면역측정법(ELISA), 면역형광 검정, 웨스턴 블로팅, 면역조직화학 등을 포함할 수 있다. 당업자는 세포가 본 발명의 마커를 발현하는지의 여부를 결정하는데 사용하기 위한 공지된 단백질/항체 검출 반응을 용이하게 적합화시킬 수 있다.

[0235] 본 발명의 폴리펩티드를 검출하기 위한 또 다른 작용제는 본 발명의 마커에 상응하는 폴리펩티드에 결합할 수 있는 항체, 예를 들어, 검출가능한 표지를 갖는 항체이다. 항체는 폴리클로날 또는 모노클로날 항체일 수 있다. 온전한 항체, 또는 이의 단편(예를 들어, Fab 또는 F(ab')₂)이 이용될 수 있다. 프로브 또는 항체와 관련된 용어 "표지된"은 프로브 또는 항체에 대한 검출가능한 물질의 커플링(즉, 물리적 연결)에 의한 프로브 또는 항체의 직접적 표지화, 뿐만 아니라 직접적으로 표지되는 또 다른 시약과의 반응성에 의한 프로브 또는 항체의 간접적 표지화를 포함한다. 간접적 표지화의 예는 형광 표지된 이차 항체를 이용한 일차 항체의 검출 및 형광 표지된 스트렙타비딘을 이용하여 검출될 수 있는 비오틴을 이용한 DNA 프로브의 말단-표지화를 포함한다.

[0236] 또 다른 구체예에서, 항체는 표지되고, 예를 들어, 방사성-표지되거나, 발색단-표지되거나, 형광단-표지되거나, 효소-표지되는 항체이다. 또 다른 구체예에서, 마커에 상응하는 단백질과 특이적으로 결합하는 항체 유도체(예를 들어, 기질 또는 단백질-리간드쌍(예를 들어, 비오틴-스트렙타비딘)의 단백질 또는 리간드와 컨쥬게이션된 항체) 또는 항체 단편(예를 들어, 단일-사슬 항체, 분리된 항체 과가변 도메인 등), 예를 들어, 마커에 상응하는 열린해독물에 의해 엔코딩되는 단백질 또는 정상 번역후 변형의 전부 또는 일부를 겪는 상기 단백질이 이용된다.

[0237] 면역조직화학 또는 IHC는 생물학적 조직에서 항원에 특이적으로 결합하는 항체의 원리를 이용하는 조직 섹션의 세포 내에 항원(예를 들어, 단백질)을 국소화시키는 방법을 의미한다. 면역조직화학 염색은 암성 종양에서 발견된 것과 같은 이상 세포의 진단에서 널리 사용된다. 특이적 분자 마커는 증식 또는 세포 사멸(아포토시스)와 같은 특정 세포 사건의 특징이다. IHC는 또한 바이오마커의 분포 및 국소화 및 생물학적 조직의 상이한 부분에서의 차별적으로 발현되는 단백질을 이해하기 위한 연구에서 널리 사용된다. 항체-항원 상호작용을 시각화시키는 것은 다수의 방식으로 달성될 수 있다. 대부분의 일반적인 예에서, 항체는 색-발생 반응을 촉매할 수 있는 효소, 예를 들어, 퍼옥시다제에 컨쥬게이션된다. 대안적으로, 항체는 또한 형광단, 예를 들어, 플루오레세인, 로다민, DyLight Fluor 또는 Alexa Fluor에 태깅될 수 있다.

[0238] 세포로부터의 단백질은 당업자에게 널리 공지된 기술을 이용하여 분리될 수 있다. 단백질 분리 방법은, 예를 들어, 문헌[Harlow and Lane (Harlow and Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York)]에 기재된 것일 수 있다.

[0239] 한 형식에서, 항체 또는 항체 단편은 발현된 단백질을 검출하기 위한 웨스턴 블롯 또는 면역형광 기술과 같은 방법에 이용될 수 있다. 상기 용도에서, 고체 지지체에 항체 또는 단백질을 고정시킬 수 있다. 적합한 고체상 지지체 또는 담체는 항원 또는 항체가 결합할 수 있는 임의의 지지체를 포함한다. 널리 공지된 지지체 또는 담체는 유리, 폴리스티렌, 폴리프로필렌, 폴리에틸렌, 텍스트란, 나일론, 아밀라아제, 자연 및 변형 셀룰로오스, 폴리아크릴아미드, 반려암, 및 자철광을 포함한다.

[0240] 당업자는 항체 또는 항원을 결합시키기 위한 많은 다른 적합한 담체를 인지할 것이며, 본 발명과 함께 사용하기

위한 상기 지지체를 적합화시킬 수 있을 것이다. 예를 들어, 세포로부터 분리된 단백질은 폴리아크릴아미드 겔 전기영동에서 영동될 수 있고, 고상 지지체, 예를 들어, 니트로셀룰로오스에 고정될 수 있다. 이후, 지지체는 적합한 완충액으로 세척된 후, 검출가능하게 표지된 항체로 처리될 수 있다. 고상 지지체는 이후 결합되지 않은 항체를 제거하기 위해 두번째 시간 동안 완충액으로 세척될 수 있다. 고체 지지체 상의 결합되지 않은 표지의 양은 이후 통상적인 수단에 의해 검출될 수 있다. 전기영동 기술을 이용하여 단백질을 검출하는 수단은 당업자에게 널리 공지되어 있다(일반적으로, R. Scopes (1982) *Protein Purification*, Springer-Verlag, N.Y.; Deutscher, (1990) *Methods in Enzymology* Vol. 182: *Guide to Protein Purification*, Academic Press, Inc., N.Y.).

[0241] 또 다른 구체예에서, 웨스턴 블롯(면역블롯) 분석은 샘플 내의 폴리펩티드의 존재를 검출하고 정량하는데 사용된다. 이러한 기술은 일반적으로 분자량을 기초로 하여 겔 전기영동에 의해 샘플 단백질을 분리시키고, 분리된 단백질을 적합한 고체 지지체(예를 들어, 니트로셀룰로오스 필터, 나일론 필터, 또는 유도체화된 나일론 필터)로 옮기고, 샘플을 폴리펩티드에 특이적으로 결합하는 항체와 함께 인큐베이션시키는 것을 포함한다. 항-폴리펩티드 항체는 고체 지지체 상의 폴리펩티드에 특이적으로 결합한다. 이러한 항체는 직접 표지될 수 있거나, 대안적으로 항-폴리펩티드에 특이적으로 결합하는 표지된 항체(예를 들어, 표지된 항-인간 항체)를 이용하여 이후에 검출될 수 있다.

[0242] 또 다른 구체예에서, 폴리펩티드는 면역검정을 이용하여 검출된다. 본원에서 사용되는 바와 같은 면역검정은 분석물에 특이적으로 결합하는 항체를 이용하는 검정이다. 따라서, 면역검정은 분석물을 분리시키고, 표적화시키고, 정량하기 위한 다른 물리적 또는 화학적 특성의 사용과 반대되는 항-항체에 대한 폴리펩티드의 특이적 결합의 검출을 특징으로 한다.

[0243] 폴리펩티드는 다수의 널리 인지된 면역학적 결합 검정 중 임의의 검정을 이용하여 검출되고/되거나 정량된다(예를 들어, 미국 특허 번호 4,366,241호(참조로서 포함됨); 4,376,110호(참조로서 포함됨); 4,517,288호(참조로서 포함됨); 및 4,837,168호(참조로서 포함됨) 참조). 일반적 면역검정의 개관을 위해, 문헌[Asai (1993) *Methods in Cell Biology* Volume 37: *Antibodies in Cell Biology*, Academic Press, Inc. New York; Stites & Terr (1991) *Basic and Clinical Immunology* 7th Edition]을 또한 참조하라.

[0244] 면역학적 결합 검정(또는 면역검정)은 통상적으로 분석물(폴리펩티드 또는 서브서열(subsequence))에 특이적으로 결합하고, 이를 종종 고정시키는 "포획 작용제"를 이용한다. 포획 작용제는 분석물에 특이적으로 결합하는 모이어티이다. 또 다른 구체예에서, 포획 작용제는 폴리펩티드에 특이적으로 결합하는 항체이다. 항체(항-펩티드)는 당업자에게 널리 공지된 다수의 수단 중 임의의 수단에 의해 생성될 수 있다.

[0245] 면역검정은 또한 종종 포획 작용제 및 분석물에 의해 형성된 결합 복합체에 특이적으로 결합하고 이를 표지시키는 표지화 작용제를 이용한다. 표지화 작용제 자체는 항체/분석물 복합체를 포함하는 모이어티 중 하나일 수 있다. 따라서, 표지화 작용제는 표지된 폴리펩티드 또는 표지된 항-항체일 수 있다. 대안적으로, 표지화 작용제는 세번째 모이어티, 예를 들어, 항체/폴리펩티드 복합체에 특이적으로 결합하는 또 다른 항체일 수 있다.

[0246] 한 구체예에서, 표지화 작용제는 표지를 갖는 두번째 인간 항체이다. 대안적으로, 두번째 항체는 표지가 결핍되어 있을 수 있으나, 이는 다음에 두번째 항체가 유래되는 종의 항체에 특이적인 표지된 세번째 항체에 의해 결합될 수 있다. 두번째 항체는 검출가능한 모이어티, 예를 들어, 세번째 표지된 분자가 특이적으로 결합할 수 있는 비오틴, 예를 들어, 효소-표지된 스트렙타비딘으로 변형될 수 있다.

[0247] 면역글로불린 불변 영역에 특이적으로 결합할 수 있는 다른 단백질, 예를 들어, 단백질 A 또는 단백질 G가 또한 표지 작용제로 사용될 수 있다. 이러한 단백질은 스트렙토코쿠스 박테리아의 세포벽의 일반적인 구성성분이다. 이들은 다양한 종으로부터의 면역글로불린 불변 영역과의 강한 비-면역원성 반응성을 나타낸다(일반적으로, Kronval, et al. (1973) *J. Immunol.*, 111:1401-1406, 및 Akerstrom (1985) *J. Immunol*, 135: 2589-2542 참조).

[0248] 상기 표시한 바와 같이, 폴리펩티드의 검출 및/또는 정량을 위한 면역검정은 당업자에게 널리 공지된 매우 다양한 형식을 취할 수 있다.

[0249] 폴리펩티드를 검출하기 위한 예시적 면역검정은 경쟁적 또는 비경쟁적 면역검정일 수 있다. 비경쟁적 면역검정은 포획된 분석물의 양이 직접적으로 측정되는 검정이다. 한 "샌드위치" 검정에서, 예를 들어, 포획 작용제(항-펩티드 항체)는 이들이 고정되는 고체 기관에 직접적으로 결합될 수 있다. 이러한 고정된 항체는 이후 시험 샘플에 존재하는 폴리펩티드를 포획한다. 따라서, 고정된 폴리펩티드는 이후 표지를 갖는 두번째 인간 항체와

같은 표지화 작용제에 의해 결합된다.

- [0250] 경쟁적 검정에서, 샘플에 존재하는 분석물(폴리펩티드)의 양은 샘플 내에 존재하는 분석물에 의해 포획 작용제(항-펩티드 항체)로부터 대체되는(또는 경쟁하는) 첨가된(외인성) 분석물(폴리펩티드)의 양을 측정함으로써 간접적으로 측정된다. 한 경쟁적 검정에서, 상기 경우에, 공지된 양의 폴리펩티드가 샘플에 첨가되고, 샘플은 이후 포획 작용제와 접촉된다. 항체에 결합된 폴리펩티드의 양은 샘플에 존재하는 폴리펩티드의 농도에 역비례한다.
- [0251] 또 다른 구체예에서, 항체는 고체 기관 상에 고정된다. 항체에 결합된 폴리펩티드의 양은 폴리펩티드/항체 복합체에 존재하는 폴리펩티드의 양을 측정하거나, 대안적으로 복합체화되지 않은 채로 남아있는 폴리펩티드의 양을 측정함으로써 결정될 수 있다. 폴리펩티드의 양은 표지된 폴리펩티드를 제공함으로써 검출될 수 있다.
- [0252] 본원에 기재된 검정은 당업자에게 널리 공지된 표준 방법에 따라 스코어링(양성 또는 음성 또는 폴리펩티드의 양)된다. 스코어링의 특정 방법은 검정 형식 및 표지의 선택에 좌우될 것이다. 예를 들어, 웨스턴 블롯 검정은 효소 표지에 의해 생성된 착색된 생성물을 시각화시킴으로써 스코어링될 수 있다. 정확한 분자량에서 명백히 보이는 착색된 밴드 또는 스팟이 양성 결과로 스코어링되는 반면, 명백히 보이는 스팟 또는 밴드의 부재는 음성으로 스코어링된다. 밴드 또는 스팟의 강도는 폴리펩티드의 정량적 측정을 제공할 수 있다.
- [0253] 본원에 기재된 다양한 면역검정에 사용하기 위한 항체는 본원에 기재된 바와 같이 생성될 수 있다.
- [0254] 또 다른 구체예에서, 수준(활성)은 유전자 생성물의 효소 활성을 측정함으로써 검정된다. 효소의 활성을 검정하는 방법은 당업자에게 널리 공지되어 있다.
- [0255] 마커 단백질의 검출을 위한 생체내 기술은 단백질에 대해 특이적인 표지된 항체를 피검체에 도입시키는 것을 포함한다. 예를 들어, 항체는 방사성 마커로 표지될 수 있고, 피검체에서의 상기 방사성 마커의 존재 및 위치는 표준 영상화 기술에 의해 검출될 수 있다.
- [0256] 본 발명의 방법에 의해 확인되는 특정 마커는 분비된 단백질일 수 있다. 당업자가 임의의 특정 마커 단백질이 분비된 단백질인지의 여부를 결정하는 것은 간단한 문제이다. 상기 결정을 하기 위해, 마커 단백질이, 예를 들어, 포유동물 세포, 예를 들어, 인간 세포주에서 발현되고, 세포의 유체가 수거되고, 세포의 유체 내의 단백질의 존재 또는 부재가 평가(예를 들어, 단백질에 특이적으로 결합하는 표지된 항체를 이용함)된다.
- [0257] 하기는 단백질의 분비를 검출하는데 사용될 수 있는 방법의 한 예이다. 약 8×10^5 개의 293T 세포가 5% (v/v) CO₂, 95% 공기 하에서 성장 배지(10% 우태아 혈청이 보충된 돌베코(Dulbecco) 변형 이글 배지(Dulbecco's modified Eagle's medium, DMEM))를 함유하는 웰 중에서 37°C에서 약 60 내지 70% 컨플루언스(confluence)까지 인큐베이션된다. 이후, 세포는 웰 당 단백질을 엔코딩하는 발현 벡터를 포함하는 2 마이크로그램의 DNA 및 10 마이크로리터의 LipofectAMINE™(GIBCO/BRL Catalog no. 18342-012)을 포함하는 표준 트랜스펙션 혼합물을 이용하여 트랜스펙션된다. 트랜스펙션 혼합물은 약 5시간 동안 유지된 후, 새로운 성장 배지로 대체되고, 대기 중에서 유지된다. 각각의 웰은 메티오닌 또는 시스테인을 함유하지 않는 DMEM(DMEM-MC; ICN Catalog no. 16-424-54)으로 2회 가볍게 행구어진다. 약 1 밀리리터의 DMEM-MC 및 약 50 마이크로리터의 Trans-³⁵S™ 시약(ICN Catalog no. 51006)이 각각의 웰에 첨가된다. 웰은 상기 기재된 5% CO₂ 대기하에서 유지되고, 선택된 기간 동안 37°C에서 인큐베이션된다. 인큐베이션 후, 150 마이크로리터의 적응용 배지(conditioned medium)가 분리되고, 원심분리되어 부유 세포 및 조직파편이 제거된다. 상층액에서의 단백질의 존재는 단백질이 분비된 것을 나타낸다.
- [0258] 피검체 샘플, 예를 들어, 가래, 기관지폐포 세척액, 흉막삼출액, 조직, 전혈, 혈청, 혈장, 구강 채취물, 타액, 뇌척수액, 소변, 대변, 및 골수가 이 안에 세포를 함유할 수 있으며, 특히 세포가 암성 세포이고, 더욱 특히 암이 전이되는 암인 경우, 상기 피검체 샘플이 본 발명의 방법에서 사용될 수 있음이 인지될 것이다. 세포 샘플은 물론 샘플 내의 마커의 발현 수준을 평가하기 전에 다양한 널리 공지된 수거후 예비 및 저장 기술(예를 들어, 핵산 및/또는 단백질 추출, 고정, 저장, 동결, 초여과, 농축, 증발, 원심분리 등)에 적용될 수 있다. 따라서, 본 발명의 조성물, 키트, 및 방법은 단백질을 발현하는 세포의 표면에 전시되는 적어도 하나의 부분을 갖는 단백질에 상응하는 마커의 발현을 검출하는데 사용될 수 있다. 당업자가 임의의 특정 마커에 상응하는 단백질이 세포-표면 단백질을 포함하는지의 여부를 결정하는 것은 간단한 문제이다. 예를 들어, 전체 세포 상에서 상기 단백질을 검출하기 위해 면역학적 방법이 이용될 수 있거나, 널리 공지된 컴퓨터-기반 서열 분석 방법(예를 들어, SIGNALP 프로그램; Nielsen *et al.*, 1997, *Protein Engineering* 10:1-6)이 적어도 하나의 세포의 도

메인(즉, 분비된 단백질 및 적어도 하나의 세포-표면 도메인을 갖는 단백질 둘 모두를 포함함)의 존재를 예측하는데 이용될 수 있다. 단백질을 발현하는 세포의 표면 상에 전시되는 적어도 하나의 부분을 갖는 단백질에 상응하는 마커의 발현은 반드시 세포를 용해시킬 필요 없이 검출될 수 있다(예를 들어, 단백질의 세포-표면 도메인에 특이적으로 결합하는 표지된 항체를 이용함).

[0259] 본 발명은 또한 생물학적 샘플, 예를 들어, 가래, 기관지세포 세척액, 흉막삼출액, 조직, 전혈, 혈청, 혈장, 구강 채취물, 타액, 뇌척수액, 소변, 대변, 골수를 함유하는 샘플 내의 본 발명의 마커에 상응하는 폴리펩티드 또는 핵산의 존재를 검출하기 위한 키트를 포함한다. 이러한 키트는 피검체가 암에 걸렸는지 또는 암이 발병할 위험이 증가되었는지 결정하기 위해 사용될 수 있다. 예를 들어, 키트는 생물학적 샘플 내의 본 발명의 마커에 상응하는 폴리펩티드 또는 폴리펩티드를 엔코딩하는 mRNA를 검출할 수 있는 표지된 화합물 또는 작용제 및 샘플 내의 폴리펩티드 또는 mRNA의 양을 결정하기 위한 수단(예를 들어, 폴리펩티드에 결합하는 항체 또는 DNA 또는 폴리펩티드를 엔코딩하는 mRNA에 결합하는 올리고뉴클레오티드 프로브)을 포함할 수 있다. 키트는 또한 키트를 이용하여 수득된 결과를 해석하기 위한 설명서를 포함할 수 있다.

[0260] 항체-기반 키트에 대해, 키트는, 예를 들어, (1) 본 발명의 마커에 상응하는 폴리펩티드에 결합하는 첫번째 항체(예를 들어, 고체 지지체에 부착됨); 및 (2) 폴리펩티드 또는 첫번째 항체에 결합하고, 검출가능한 표지에 컨쥬게이션된, 두번째의 상이한 항체를 포함할 수 있다.

[0261] 올리고뉴클레오티드-기반 키트에 대해, 키트는, 예를 들어, (1) 본 발명의 마커에 상응하는 폴리펩티드를 엔코딩하는 핵산 서열에 하이브리드화되는 올리고뉴클레오티드, 예를 들어, 검출가능하게 표지된 올리고뉴클레오티드 또는 (2) 본 발명의 마커에 상응하는 핵산 분자를 증폭시키는데 유용한 프라이머 쌍을 포함할 수 있다. 키트는 또한, 예를 들어, 완충제, 보존제, 또는 단백질 안정화제를 포함할 수 있다. 키트는 검출가능한 표지(예를 들어, 효소 또는 기질)를 검출하는데 필요한 성분을 추가로 포함할 수 있다. 키트는 또한 검정되고 시험 샘플과 비교될 수 있는 대조군 샘플 또는 일련의 대조군 샘플을 함유할 수 있다. 키트의 각각의 성분은 개별적 용기 내에 동봉될 수 있고, 다양한 용기 모두는 키트를 이용하여 수행된 검정의 결과를 해석하기 위한 설명서와 함께 단일한 패키지 내에 존재할 수 있다.

[0262] 4. 구조적 변화를 평가하기 위한 방법

[0263] 본 발명은 또한 구조적 변화, 예를 들어, 돌연변이의 존재를 평가하기 위한 방법을 제공한다.

[0264] 또 다른 검출 방법은 다형성 부위에 중첩되고, 다형성 영역 주위의 약 5개, 약 10개, 약 20개, 약 25개, 또는 약 30개의 뉴클레오티드를 갖는 프로브를 이용한 대립유전자 특이적 하이브리드화이다. 본 발명의 또 다른 구체예에서, 돌연변이에 특이적으로 하이브리드화될 수 있는 여러 프로브가 고상 지지체, 예를 들어, "칩"에 부착된다. 올리고뉴클레오티드는 리소그래피(lithography)를 포함하는 다양한 방법에 의해 고체 지지체에 결합될 수 있다. 예를 들어, 칩은 250,000개의 올리고뉴클레오티드를 가질 수 있다(GeneChip, Affymetrix™). "DNA 프로브 어레이"로도 언급되는 올리고뉴클레오티드를 포함하는 상기 칩을 이용한 돌연변이 검출 분석은, 예를 들어, 문헌[Cronin *et al.* (1996) *Human Mutation* 7:244]에 기재되어 있다. 한 구체예에서, 칩은 유전자의 적어도 하나의 다형성 영역의 돌연변이 모두를 포함한다. 고상 지지체는 이후 시험 핵산과 접촉되고, 특정 프로브에 대한 하이브리드화가 검출된다. 따라서, 하나 이상의 유전자의 다수의 돌연변이의 확인은 간단한 하이브리드화 실험으로 확인될 수 있다. 예를 들어, 5' 업스트림 조절 요소에서의 뉴클레오티드 다형성의 돌연변이의 확인은 단일한 하이브리드화 실험으로 결정될 수 있다.

[0265] 다른 검출 방법에서, 돌연변이를 확인하기 전에 마커의 적어도 일부를 먼저 증폭시키는 것이 필요하다. 증폭은 당 분야에 공지된 방법에 따라, 예를 들어, PCR 및/또는 LCR(Wu and Wallace (1989) *Genomics* 4:560 참조)에 의해 수행될 수 있다. 한 구체예에서, 세포의 유전체 DNA는 2개의 PCR 프라이머 및 증폭된 DNA의 필요량을 생성시키기에 충분한 다수의 주기의 증폭에 노출된다. 특정 구체예에서, 프라이머는 150 내지 350개의 염기쌍만큼 떨어져 위치된다.

[0266] 대안적 증폭 방법은 다음을 포함한다: 자기부양 염기서열 복제반응(Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), 전사 증폭 시스템(Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-메타 레플리카제(Lizardi, P.M. *et al.*, (1988) *Bio/Technology* 6: 1197), 및 자기부양 염기서열 복제반응(Guatelli *et al.*, (1989) *Proc. Nat. Acad. Sci.* 87:1874), 및 핵산 기반 서열 증폭(NABSA), 또는 임의의 다른 핵산 증폭 방법 후, 당업자에게 널리 공지된 기술을 이용한 증폭된 분자의 검출. 이러한 검출 계획은 핵산 분자가 매우 적은 양으로 존재하는 경우에 상기 핵산 분자의 검출에 특히 유용하다.

- [0267] 한 구체예에서, 당 분야에 공지된 다양한 시퀀싱 반응 중 임의의 시퀀싱 반응이 마커의 적어도 일부를 직접 시퀀싱하고, 상응하는 참조(대조군) 서열을 갖는 샘플 서열의 서열과 비교함으로써 돌연변이를 검출하는데 이용될 수 있다. 예시적 시퀀싱 반응은 문헌[Maxam and Gilbert (*Proc. Natl Acad Sci USA* (1977) 74:560) 또는 Sanger (Sanger *et al.* (1977) *Proc. Nat. Acad. Sci* 74:5463)]에 의해 개발된 기술을 기초로 한 것을 포함한다. 질량분광법에 의한 시퀀싱(예를 들어, H. Koster에 의한 DNA Sequencing by Mass Spectrometry를 제목으로 하는 미국 특허 번호 5,547,835호(참조로서 포함됨) 및 국제 특허 출원 공개 번호 WO 94/16101호(참조로서 포함됨); H. Koster에 의한 DNA Sequencing by Mass Spectrometry Via Exonuclease Degradation을 제목으로 하는 미국 특허 번호 5,547,835호(참조로서 포함됨) 및 국제 특허 출원 공개 번호 WO 94/21822호(참조로서 포함됨), 및 H. Koster에 의한 DNA Diagnostics Based on Mass Spectrometry를 제목으로 하는 미국 특허 번호 5,605,798호(참조로서 포함됨) 및 국제 특허 출원 번호 PCT/US96/03651호(참조로서 포함됨); 문헌[Cohen *et al.* (1996) *Adv Chromatogr* 36: 127-162; 및 Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38: 147-159] 참조)을 포함하는 다양한 자동화된 시퀀싱 절차 중 임의의 절차가 피검체 검정을 수행하는 경우에 이용될 수 있는 것이 또한 고려된다(*Biotechniques* (1995) 19:448). 특정 구체예에서, 단지 1개, 2개 또는 3개의 핵산 염기의 존재가 시퀀싱 반응에서 결정되는 것이 필요함이 당업자에게 명백할 것이다. 예를 들어, 단지 1개의 누클레오티드가 검출되는 경우 A-트랙(A-track) 등이 수행될 수 있다.
- [0268] 또 다른 시퀀싱 방법은, 예를 들어, "Method of DNA sequencing employing a mixed DNA-polymer chain probe"를 제목으로 하는 미국 특허 번호 5,580,732호(참조로서 포함됨) 및 "Method for mismatch-directed in vitro DNA sequencing"을 제목으로 하는 미국 특허 번호 5,571,676호(참조로서 포함됨)에 개시되어 있다.
- [0269] 일부 경우에, 피검체로부터의 DNA에 마커의 특정 대립유전자의 존재가 제한 효소 분석에 의해 제시될 수 있다. 예를 들어, 특이적 누클레오티드 다형성은 또 다른 돌연변이의 누클레오티드 서열에 부재하는 제한 부위를 포함하는 누클레오티드 서열을 발생시킬 수 있다.
- [0270] 한 추가 구체예에서, 절단 작용제(예를 들어, 누클레아제, 히드록실아민 또는 오스뮴 테트록시드 및 피페리딘)로부터의 보호는 RNA/RNA, DNA/DNA, 또는 RNA/DNA 헤테로듀플렉스(heteroduplex)에서 미스매치된 염기를 검출하는데 이용될 수 있다(Myers, *et al.* (1985) *Science* 230: 1242). 일반적으로, "미스매치 절단" 기술은 마커 돌연변이의 누클레오티드 서열을 포함하는 임의로 표시되는 대조군 핵산, 예를 들어, RNA 또는 DNA와 조직 샘플로부터 수득된 샘플 핵산, 예를 들어, RNA 또는 DNA를 하이브리드화시킴으로써 형성된 헤테로듀플렉스를 제공함으로써 개시된다. 이중-가닥의 듀플렉스는 대조군 가닥과 샘플 가닥 사이의 염기쌍 미스매치를 기초로 하여 형성된 듀플렉스와 같은 듀플렉스의 단일-가닥 영역을 절단하는 작용제로 처리된다. 예를 들어, RNA/DNA 듀플렉스는 RNase로 처리될 수 있고, DNA/DNA 하이브리드는 미스매치된 영역을 효소적으로 절단하는 S1 누클레아제로 처리될 수 있다. 다른 구체예에서, DNA/DNA 또는 RNA/DNA 듀플렉스는 미스매치된 영역을 절단하기 위해 히드록실아민 또는 오스뮴 테트록시드 및 피페리딘으로 처리될 수 있다. 미스매치된 영역의 절단 후, 생성된 물질은 이후 대조군 및 샘플 핵산이 동일한 누클레오티드 서열을 갖는지 또는 대조군 및 샘플 핵산의 누클레오티드가 상이한지의 여부를 결정하기 위해 변성 폴리아크릴아미드 겔 상에서의 크기에 의해 분리된다. 예를 들어, 문헌[Cotton *et al* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol.* 217:286-295]을 참조하라. 또 다른 구체예에서, 대조군 또는 샘플 핵산은 검출을 위해 표시된다.
- [0271] 또 다른 구체예에서, 돌연변이는 변성 고-성능 액체 크로마토그래피(DHPLC)에 의해 확인될 수 있다(Oefner and Underhill, (1995) *Am. J. Human Gen.* 57:Suppl. A266). DHPLC는 단편 내의 특정 누클레오티드 유전자좌에서 이형접합인 개체로부터의 PCR 단편의 증폭 동안 생성되는 헤테로듀플렉스를 검출하기 위해 역상 이온-쌍형성 크로마토그래피(reverse-phase ion-pairing chromatography)를 이용한다(Oefner and Underhill (1995) *Am. J. Human Gen.* 57:Suppl. A266). 일반적으로, PCR 생성물은 관심 DNA에 플랭킹된 PCR 프라이머를 이용하여 생성된다. DHPLC 분석이 수행되고, 생성된 크로마토그램은 특정 크로마토그래피 프로파일을 기초로 하여 염기쌍 변화 또는 결실을 확인하기 위해 분석된다(O' Donovan *et al.* (1998) *Genomics* 52:44-49 참조).
- [0272] 다른 구체예에서, 마커 돌연변이의 유형을 확인하기 위해 전기영동 이동에서의 변화가 이용된다. 예를 들어, 단일가닥 구조 다형성(SSCP)이 돌연변이와 야생형 핵산 사이의 전기영동 이동성에서의 차이를 검출하기 위해 사용될 수 있다(Orita *et al.* (1989) *Proc Natl. Acad. Sci USA* 86:2766, 또한 Cotton (1993) *Mutat Res* 285:125-144; 및 Hayashi (1992) *Genet Anal Tech Appl* 9:73-79 참조). 샘플 및 대조군 핵산의 단일-가닥 DNA 단편이 변성되고, 복원된다. 단일-가닥 핵산의 이차 구조는 서열에 따라 다양하고, 전기영동 이동성에서의 발생된 변화는 심지어 단일 염기 변화의 검출을 가능케 한다. DNA 단편은 표시되거나, 표시된 프로브를 이용하여 검출될 수 있다. 검정의 민감성은 이차 구조가 서열에서의 변화에 대해 더욱 민감한 RNA(DNA가 아님)를 이용함

으로써 향상될 수 있다. 또 다른 구체예에서, 본 발명의 방법은 전기영동 이동성에서의 변화를 기초로 하여 이중 가닥 헤테로듀플렉스 분자를 분리하는 헤테로듀플렉스 분석을 이용한다(Keen *et al.* (1991) *Trends Genet* 7:5).

[0273] 또 다른 구체예에서, 다형성 영역의 돌연변이의 확인은 변성의 구배를 함유하는 폴리아크릴아미드 겔 내에서의 다형성 영역을 포함하는 핵산의 이동을 분석함으로써 수득되며, 이는 변성 구배 겔 전기영동(DGGE)을 이용하여 검정된다(Myers *et al.* (1985) *Nature* 313:495). 분석 방법으로서 DGGE가 사용되는 경우, DNA는, 예를 들어, PCR에 의해 약 40 bp의 고-용해 GC-풍부 DNA의 GC 클램프(clamp)를 첨가함으로써 완전히 변성되지 않는 것을 보장하기 위해 변형될 것이다. 한 추가 구체예에서, 대조군 및 샘플 DNA의 이동성에서의 차이를 확인하기 위해 변성 작용제 구배 대신에 온도 구배가 이용된다(Rosenbaum and Reissner (1987) *Biophys Chem* 265:1275).

[0274] 2개의 핵산 사이의 적어도 하나의 뉴클레오타이드의 차이를 검출하기 위한 기술의 예는 선택적 올리고뉴클레오타이드 하이브리드화, 선택적 증폭, 또는 선택적 프라이머 신장(extension)을 포함하나, 이에 제한되지는 않는다. 예를 들어, 공지된 다형성 뉴클레오타이드가 중심에 위치(대립유전자-특이적 프로브)된 후, 완전한 매치가 발견되는 경우에만 하이브리드화를 허용하는 조건하에서 표적 DNA에 하이브리드화되는 올리고뉴클레오타이드 프로브가 제조될 수 있다(Saiki *et al.* (1986) *Nature* 324: 163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230; 및 Wallace *et al.* (1979) *Nucl. Acids Res.* 6:3543). 이러한 대립유전자 특이적 올리고뉴클레오타이드 하이브리드화 기술은 마커의 다양한 다형성 영역에서의 여러 뉴클레오타이드 변화의 동시 검출에 이용될 수 있다. 예를 들어, 특이적 돌연변이의 뉴클레오타이드 서열을 갖는 올리고뉴클레오타이드는 하이브리드화 막에 부착되고, 이러한 막은 이후에 표지된 샘플 핵산과 하이브리드화된다. 이후, 하이브리드화 신호의 분석은 샘플 핵산의 뉴클레오타이드의 정체를 나타낼 것이다.

[0275] 대안적으로, 선택적 PCR 증폭에 의존하는 대립유전자 특이적 증폭 기술은 본 발명과 함께 이용될 수 있다. 특정 증폭을 위한 프라이머로서 사용되는 올리고뉴클레오타이드는 분자의 중심(증폭이 차별적 하이브리드화에 의존함)(Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448), 또는 적절한 조건하에서 미스매치를 방지할 수 있거나, 증합효소 신장을 감소시킬 수 있는 프라이머의 극도의 3' 말단(Prossner (1993) *Tibtech* 11:238; Newton *et al.* (1989) *Nucl. Acids Res.* 17:2503)에 관심 돌연변이를 가질 수 있다. 이러한 기술은 프로브 올리고 염기 신장(Probe Oligo Base Extension)을 위한 "PROBE"로도 언급된다. 또한, 절단-기반 검출을 발생시키기 위해 돌연변이의 영역 내에 신규한 제한 부위를 도입시키는 것이 요망될 수 있다(Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1).

[0276] 또 다른 구체예에서, 돌연변이의 확인은, 예를 들어, 미국 특허 번호 4,998,617호 및 문헌[Landegren, U. *et al.*, (1988) *Science* 241:1077-1080]에 기재된 바와 같은 올리고뉴클레오타이드 라이게이션 검정(OLA)을 이용하여 수행된다. OLA 프로토콜은 표적의 단일 가닥의 인접한 서열에 하이브리드화될 수 있도록 디자인된 2개의 올리고뉴클레오타이드를 이용한다. 올리고뉴클레오타이드 중 하나는 분리 마커, 예를 들어, 비오틴화된 마커에 연결되고, 나머지는 검출가능하게 표지된다. 정확한 상보적 서열이 표적 분자에서 발견되는 경우, 올리고뉴클레오타이드는 이의 말단이 인접하고, 라이게이션 기질을 생성시키도록 하이브리드화될 것이다. 이후, 라이게이션은 표지된 올리고뉴클레오타이드가 아비딘 또는 또 다른 비오틴 리간드를 이용하여 회수되도록 한다. 문헌[Nickerson, D. A. *et al.*]에는 PCR 및 OLA의 특성을 조합시키는 핵산 검출 검정이 기재되어 있다(Nickerson, D.A. *et al.*, (1990) *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8923-8927). 이러한 방법에서, PCR은 표적 DNA의 지수적 증폭을 달성하기 위해 이용되고, 상기 증폭은 이후에 OLA를 이용하여 검출된다.

[0277] 본 발명은 추가로 마커에서 단일 뉴클레오타이드 다형성을 검출하는 방법을 추가로 제공한다. 단일 뉴클레오타이드 다형성은 불변 서열의 영역에 의해 플랭킹된 변화 부위를 구성하므로, 이의 분석은 변화 부위에 존재하는 단일 뉴클레오타이드의 정체의 결정 이상을 필요로 하지 않으며, 각각의 피검체에 대해 완전한 유전 서열을 결정하는 것이 불필요하다. 상기 단일 뉴클레오타이드 다형성의 분석을 촉진하기 위해 여러 방법이 개발되었다.

[0278] 한 구체예에서, 단일 염기 다형성은, 예를 들어, 문헌[Mundy, C.R. (미국 특허 번호 4,656,127호(참조로서 포함됨))]에 개시된 바와 같은 특화된 엑소뉴클레아제-내성 뉴클레오타이드를 이용하여 검출될 수 있다. 상기 방법에 따르면, 다형성 부위 3'에 가까운 대립유전자 서열에 상보적인 프라이머가 특정 동물 또는 인간으로부터 수득된 표적 분자에 하이브리드화되는 것이 가능해진다. 표적 분자 상의 다형성 부위가 존재하는 특정 엑소뉴클레아제-내성 뉴클레오타이드 유도체에 상보적인 뉴클레오타이드를 함유하는 경우, 상기 유도체는 하이브리드화된 프라이머의 말단에 통합될 것이다. 이러한 통합은 프라이머가 엑소뉴클레아제에 대해 내성이 되도록 함으로써, 이의 검출이 가능해진다. 샘플의 엑소뉴클레아제-내성 유도체의 정체가 공지되어 있으므로, 프라이머가 엑소뉴클레아

제에 대해 내성이 되었다는 발견은 표적 분자의 다형성 부위에 존재하는 뉴클레오티드가 상기 반응에서 사용되는 뉴클레오티드 유도체의 뉴클레오티드에 상보적인 것을 나타낸다. 이러한 방법은 많은 양의 관계 없는 서열 데이터의 결정을 필요로 하지 않는다는 장점을 갖는다.

[0279] 본 발명의 또 다른 구체예에서, 용액-기반 방법이 다형성 부위의 뉴클레오티드의 정체를 결정하는데 이용된다 (Cohen, D. *et al.* French Patent 2,650,840; PCT 출원 번호 W091/02087호(참조로서 포함됨)). 미국 특허 번호 4,656,127호(참조로서 포함됨)의 먼디(Mundy) 방법에서와 같이, 다형성 부위 3'에 가까운 대립유전자 서열에 상보적인 프라이머가 이용된다. 상기 방법은 다형성 부위의 뉴클레오티드에 상보적인 경우 프라이머의 말단에 통합되는 표지된 디데옥시뉴클레오티드 유도체를 이용하여 상기 부위의 뉴클레오티드의 정체를 결정한다.

[0280] 유전 비트 분석(Genetic Bit Analysis) 또는 GBA로 공지된 대안적 방법은 문헌[Goelet, P. *et al.* (PCT 출원 번호 92/15712호(참조로서 포함됨))]에 기재되어 있다. 상기 [Goelet, P. *et al.*]의 방법은 표지된 종료자 및 다형성 부위 3' 서열에 상보적인 프라이머의 혼합물을 이용한다. 따라서, 통합되는 표지된 종료자는 평가되는 표적 분자의 다형성 부위에 존재하는 뉴클레오티드에 의해 결정되고, 이에 상보적이다. 문헌[Cohen *et al.* (프랑스 특허 2,650,840호; PCT 출원 번호 W091/02087호(참조로서 포함됨))]과 대조적으로, 문헌[Goelet, P. *et al.*]의 방법은 이중성 상 검정이며, 여기서 프라이머 또는 표적 분자는 고상에 고정된다.

[0281] DNA 내의 다형성 부위를 검정하기 위한 여러 프라이머-유도 뉴클레오티드 통합 절차가 기재되어 있다(Komher, J. S. *et al.*, (1989) *Nucl. Acids. Res.* 17:7779-7784; Sokolov, B. P., (1990) *Nucl. Acids Res.* 18:3671; Syvanen, A. -C., *et al.*, (1990) *Genomics* 8:684-692; Kuppuswamy, M. N. *et al.*, (1991) *Proc. Natl. Acad. Sci. (U.S.A.)* 88: 1143-1147; Prezant, T. R. *et al.*, (1992) *Hum. Mutat.* 1:159-164; Ugozzoli, L. *et al.*, (1992) *GATA* 9:107-112; Nyren, P. (1993) *et al.*, *Anal. Biochem.* 208: 171-175). 이러한 방법은 이들이 모두 다형성 부위의 염기 사이를 구별하기 위한 표지된 테옥시뉴클레오티드의 통합에 의존한다는 점에서 GBA와 상이하다. 이러한 형식에서, 신호는 통합된 테옥시뉴클레오티드의 수에 비례하므로, 동일 뉴클레오티드의 수행에서 발생하는 다형성은 수행 길이에 비례하는 신호를 발생시킬 수 있다(Syvanen, A.C., *et al.*, (1993) *Amer. J. Hum. Genet.* 52:46-59).

[0282] 마커의 코딩 영역에 위치한 다형성 영역의 돌연변이의 정체를 결정하기 위해, 상기 기재된 것과 다른 방법이 이용될 수 있다. 예를 들어, 돌연변이된 마커를 엔코딩하는 돌연변이의 확인은, 예를 들어, 번역조직화학 또는 면역침전에서 돌연변이 단백질을 특이적으로 인지하는 항체를 이용함으로써 수행될 수 있다. 야생형 마커 또는 마커의 돌연변이된 형태에 대한 항체는 당 분야에 공지된 방법에 따라 제조될 수 있다.

[0283] 대안적으로, 마커의 활성, 예를 들어, 마커 리간드에 대한 결합을 또한 측정할 수 있다. 결합 검정은 당 분야에 공지되어 있고, 이는, 예를 들어, 피검체로부터 세포를 수득하고, 단백질의 야생형에 대한 결합과 상이한 단백질의 돌연변이된 형태에 결합하는지의 여부를 결정하기 위해 표지된 리간드를 이용한 결합 실험을 수행하는 것을 포함한다.

[0284] VI. ALK-억제를 기초로 한 예시적 스크리닝 방법

[0285] 본 발명은 암 세포 증식, 성장, 분화, 아포토시스, 및/또는 전이를 억제하기 위해 ALK 폴리펩티드(예를 들어, EML4-ALK 폴리펩티드)를 억제하는 물질을 확인하기 위한 방법을 추가로 제공한다. 상기 방법은 시험 화합물과 ALK 폴리펩티드(예를 들어, 표 1에 나열된 폴리펩티드)를 접촉시키는 것을 포함한다. 일부 구체예에서, ALK 폴리펩티드는 하나 이상의 ALK 억제제에 의한 억제에 대한 부분적 반응 또는 무반응의 위험을 증가시키는 변이체(예를 들어, 표 1에 나열된 폴리펩티드)를 포함한다. 종양 전이의 억제제인 화합물은 ALK 폴리펩티드 변이체(예를 들어, 리간드 결합, 예를 들어, ATP 결합 및/또는 티로신 키나제 활성을 포함함)의 활성에 대한 시험 화합물의 효과를 결정함으로써 확인될 수 있다. 한 특정 예에서, 시험 화합물의 부재하에서의 활성에 비한 티로신 키나제 활성을 억제하는 시험 화합물은 종양 전이의 억제제로서 시험 화합물을 확인한다. 화합물이 ALK 변이체의 활성을 억제하는 경우, 이는 종양 성장 또는 전이를 억제하는 능력에 대해 추가로 평가될 수 있다.

[0286] 특히, 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)를 포함하는 티로신 키나제 돌연변이를 활성화시키는 것은, 예를 들어, 암 세포 증식, 성장, 분화, 아포토시스, 및/또는 전이를 억제하거나 예방함으로써 신생물을 치료하거나, 개선시키거나, 예방하는데 사용될 수 있는 화합물을 확인하는데 유용하다. 조절, 예를 들어, 억제하거나, 길항시키거나, 효능화시키거나, 모방하는 분자에 대한 화학적 라이브러리를 스크리닝하는 것은 당 분야에 공지되어 있다. 화학적 라이브러리는, 예를 들어, 펩티드 라이브러리, 펩티도미메틱(peptidomimetic) 라이브러리, 화학적으로 합성된 라이브러리, 재조합, 예를 들어, 파지 디스플레이

라이브러리, 및 시험관내 번역-기반 라이브러리, 다른 비-펩티드 합성 유기 라이브러리일 수 있다.

[0287] 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)의 적절한 고친화성 억제제의 스크리닝 또는 생성, 확인 및 선택은 다양한 방법에 의해 달성될 수 있다. 대략적으로 말하면, 이들은 2개의 일반적 방법을 포함할 수 있으나, 이에 제한되지는 않는다. 한 방법은 정확히 상호작용하는 후보 분자를 디자인하기 위해 표적 단백질에 대한 구조적 지식을 사용하는 것이다. 예로는, 특히 도 6으로 본원에 개시된 신규한 구조-기능 정보를 기초로 한 컴퓨터 보조 분자 디자인일 것이다. 두번째 방법은 분자의 조합 또는 다른 라이브러리를 사용함으로써 분자의 많은 라이브러리가 표적 효소와 관련한 친화성, 또는 표적 효소의 활성을 억제하는 능력에 대해 스크리닝된다. 한 추가 예에서, 항체의 패널은 표적 효소를 억제하는 능력에 대해 스크리닝될 수 있다.

[0288] 본원에 제공된 일부 구체예는 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)를 억제하는 제공된 화합물의 능력을 결정하는 것을 포함한다. 시험 화합물은 신생물을 치료하는데 유용한 것으로 공지된 화합물에 대한 활성을 직접적 또는 간접적으로 비교함으로써 신생물 병소를 치료하는 예상 능력에 대해 평가될 수 있다. 예를 들어, 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)에 대한 ATP 결합 및/또는 티로신 키나제 활성과 같은 리간드 결합을 억제하는 시험 화합물의 능력은 PF-02341066 및/또는 PDD와 같은 공지된 ALK 억제제의 능력과 비교될 수 있다. 한 구체예에서, 이러한 시험 화합물은 동일 검정 조건하에서 공지된 ALK 억제제의 억제에 비해 적어도 100%, 적어도 99.9%, 적어도 99.8%, 적어도 99.7%, 적어도 99.6%, 적어도 99.5%, 적어도 99.4%, 적어도 99.3%, 적어도 99.2%, 적어도 99.1%, 적어도 99%, 적어도 98.5%, 적어도 98%, 적어도 97.5%, 적어도 97%, 적어도 96.5%, 적어도 96%, 적어도 95.5%, 적어도 94%, 적어도 93.5%, 적어도 93%, 적어도 92.5%, 적어도 92%, 적어도 91.5%, 적어도 91%, 적어도 90.5%, 적어도 90%, 적어도 89.5%, 적어도 89%, 적어도 88.5%, 적어도 88%, 적어도 87.5%, 적어도 87%, 적어도 86.5%, 적어도 86%, 적어도 85.5%, 적어도 85%, 적어도 84.5%, 적어도 84%, 적어도 83.5%, 적어도 83%, 적어도 82.5%, 적어도 82%, 적어도 81.5%, 적어도 81 %, 적어도 80.5%, 적어도 80%, 적어도 79%, 적어도 78%, 적어도 77%, 적어도 76%, 적어도 75%, 적어도 74%, 적어도 73%, 적어도 72%, 적어도 71%, 적어도 70%, 적어도 69%, 적어도 68%, 적어도 67%, 적어도 66%, 적어도 65%, 적어도 64%, 적어도 63%, 적어도 62%, 적어도 61%, 적어도 60%, 적어도 59%, 적어도 58%, 적어도 57%, 적어도 56%, 적어도 55%, 적어도 54%, 적어도 53%, 적어도 52%, 적어도 51%, 적어도 50%, 또는 이들 사이의 임의의 범위의 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)의 억제를 갖는다. 특정 구체예에서, 세포는 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)를 엔코딩하는 작제물로 트랜스펙션되고, 검출가능한 마커로 태깅되거나 표지되는 시험 화합물과 접촉되고, 결합된 시험 화합물의 존재에 대해 분석될 수 있다. 특정 구체예에서, 트랜스펙션된 세포는 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)로 트랜스펙션되지 않은 세포에 비해 시험 화합물에 결합하는 것으로 관찰되며, 이는 상기 세포에 의해 발현되는 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)에 결합하는 것을 나타낸다. 화합물의 결합은 통상적으로 ELISA, RIA, 및/또는 BIAcore 검정과 같은 당 분야에 공지된 매우 다양한 검정 중 어느 하나에 의해 결정된다.

[0289] 화합물은 효소의 발현된 재조합 형태, 또는 또 다른 종으로부터 분리된 유사체 또는 오쏘로그를 이용하여 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)의 활성에 대해 억제 또는 다른 효과에 대해 스크리닝될 수 있다. 대안적으로, 상기 신규한 바이오마커 폴리펩티드 중 하나를 발현하는 세포는 시험 화합물로 처리될 수 있고, 특정 표적의 인산화에 대한 시험 화합물의 효과는, 예를 들어, 본원에 기재된 기술 중 하나를 이용하여 결정될 수 있다. 한 예에서, 티로신 키나제 활성이 결정된다. 활성(예를 들어, 억제)에 영향을 주는 티로신 키나제 인산화를 결정하기 위한 방법은 당업자에게 널리 공지되어 있다. 일부 예에서, 티로신 키나제 활성은 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)(예를 들어, 단백질 또는 펩티드 단편, 특히 다운스트림 신호전달 성분의 단백질 또는 펩티드 단편)에 의해 인산화될 수 있는 기질의 표지된 인산염(예를 들어, ³²P-표지된 인산염)의 통합을 평가함으로써 결정될 수 있다. 다른 구체예에서, 티로신 키나제 활성은 유니버설(universal) 티로신 키나제 활성 키트(예를 들어, Universal Tyrosine Kinase Assay Kit (Takara Bio, Inc., Madison, Wis.); Tyrosine Kinase Assay Kit (Millipore, Billerica, Mass.))를 이용하여 측정될 수 있다.

[0290] 또 다른 구체예에서, 화합물이 종양 세포, 예를 들어, 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)와 같은 활성화된 티로신 키나제 돌연변이를 발현하는 것으로 공지된 종양 세포의 성장을 감소시키는지의 여부를 결정하는 것을 추가로 포함하는 스크리닝 방법이 제공된다. 당업자에게 널리 공지된 시험되는 조직을 기초로 하여 선택될 수 있는 다양한 세포주(예를 들어, BA/F3 세포)가 사용될 수 있다. 예를 들어, 많은 세포주가 널리 특성규명되어 있고, 이는 신규한 항암 약물을 위한 스크리닝 프로그램으로 미국 국립 암 협회

(United States National Cancer Institute, NCI)에 의해 사용된다.

- [0291] 100 μ M, 90 μ M, 80 μ M, 70 μ M, 60 μ M, 50 μ M, 40 μ M, 30 μ M, 20 μ M, 10 μ M, 9 μ M, 8 μ M, 7 μ M, 6 μ M, 5 μ M, 4.5 μ M, 4 μ M, 3.5 μ M, 3 μ M, 2.5 μ M, 2 μ M, 1.5 μ M, 1 μ M, 900 nM, 850 nM, 800 nM, 750 nM, 700 nM, 650 nM, 600 nM, 550 nM, 500 nM, 450 nM, 400 nM, 350 nM, 300 nM, 250 nM, 200 nM, 150 nM, 100 nM, 95 nM, 90 nM, 85 nM, 80 nM, 75 nM, 70 nM, 65 nM, 60 nM, 55 nM, 50 nM, 45 nM, 40 nM, 35 nM, 30 nM, 25 nM, 20 nM, 15 nM, 10 nM, 5 nM, 4 nM, 3 nM, 2 nM, 1 nM 또는 이 미만의 용량에서 약 50%를 초과하여 발생하는 것과 같은 유의한 및 통계적으로 유의한 종양 세포 성장 억제는 상기 화합물이 신생물 병소를 치료하는데 유용한 것을 추가로 나타낸다. IC₅₀ 값이 결정될 수 있고, 비교 목적을 위해 사용될 수 있다. 이러한 값은 대조군에 비해 50%까지 종양 세포 성장을 억제하는데 필요한 약물의 농도이다.
- [0292] 이러한 값은 다른 기준에 추가로 적용될 수 있다. 예를 들어, 다른 구체예에서, 본원에 제공된 스크리닝 방법은 시험 화합물이 종양 세포의 배양에서 아포토시스를 유도하는지의 여부를 결정하는 것을 추가로 포함한다. 괴사 및 아포토시스의 세포 사멸의 2개의 상이한 형태가 형태 및 생화학 기준에 의해 기재될 수 있다. 괴사는 형질막의 투과성 증가가 동반됨으로써 세포가 팽창하고, 형질막이 수분 내에 파열된다. 아포토시스는 막 기포 형성(blebbing), 세포질의 응축, 및 내인성 엔도뉴클레아제의 활성화를 특징으로 한다.
- [0293] 아포토시스는 정상 조직 전환 및 기관 및 수족의 배아 발달 동안 자연적으로 발생한다. 아포토시스는 또한 이온화 방사선 및 특정 화학요법 약물에 의해 세포독성 T-림프구 및 자연살세포를 포함하는 다양한 자극에 의해 유도될 수 있다. 아포토시스의 부적절한 조절이 암, AIDS, 또는 알츠하이머병 등을 포함하는 많은 병리 질환에서 중요한 역할을 하는 것으로 생각된다.
- [0294] 시험 화합물은 상기 기재된 바와 같은 조건하에서 유지된 종양 세포의 배양물을 이용하여 아포토시스의 유도에 대해 스크리닝될 수 있다. 상기 스크리닝 방법의 일부 예에서, 시험 화합물을 이용한 세포의 처리는 컨플루언스 전 및 컨플루언스 후 배양 및 1 내지 7일 동안의 시험 화합물의 다양한 농도에서의 처리를 포함한다. 아포토시스 세포는 배양물의 부착된 부분 및 "부유" 부분 둘 모두에서 측정될 수 있다. 둘 모두는 상층액을 분리시키고, 부착된 세포에 트립신 처리하고, 원심분리 세척 단계(10분, 2000 rpm) 후 둘 모두의 제조물을 조합시킴으로써 수거된다. 시험 화합물을 이용한 처리 후, 배양물을, 예를 들어, 아크리딘 오렌지 및 에티디움 브로마이드를 이용한 표지화 후에 형광 현미경에 의해 아포토시스 및 괴사에 대해 검정될 수 있다. 아포토시스 세포를 측정하기 위한 많은 방법은 당업자에게 공지되어 있으며, 예를 들어, 아포토시스 세포수를 측정하기 위한 방법은 문헌[Duke & Cohen (Curr. Prot. Immuno., Coligan *et al.*, eds., 3.17.1-3.17.1, 1992)에 기재되어 있다. 예를 들어, 부유 및 부착 세포는 트립신처리에 의해 수거되고, PBS 중에서 3회 세척된다. 이후, 세포의 분취량이 원심분리된다. 펠렛은 배지 및 PBS 중에 제조된 아크리딘 오렌지 및 에티디움 브로마이드를 함유하는 염료 혼합물에 재현탁되고, 가볍게 혼합된다. 이후, 혼합물은 현미경 슬라이드 상에 배치될 수 있고, 아포토시스의 형태 특징에 대해 시험될 수 있다. 아포토시스는 또한 시험 화합물로 처리된 세포에서의 DNA 단편화에서의 증가를 측정함으로써 정량될 수 있다. 세포질 히스톤-결합-DNA-단편(모노-뉴클레오솜 및 올리고-뉴클레오솜)의 정량적 시험관내 결정에 대해 시판되는 광도계 효소 면역검정(EIA)이 이용가능하다(예를 들어, Cell Death Detection ELISA, Boehringer Mannheim).
- [0295] 추가 구체예에서, 본원에 제공된 스크리닝 방법은 시험 화합물이, 예를 들어, 전이의 동물 모델에서 종양 전이를 감소시키는지의 여부를 결정하는 것을 추가로 포함한다. 종양 전이를 평가하기 위한 방법은 당업자에게 공지되어 있다(예를 들어, Khanna and Hunter, Carcinogenesis 26:513-523, 2005 참조). 전이의 한 모델은 인간 암세포주 또는 조직이 면역약화 마우스(예를 들어, SCID 마우스 또는 누드 마우스)에 이식된 인간-마우스 이종 이식을 포함한다. 유사한 방법에서, 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)를 발현하도록 조작된 세포주가 면역약화 마우스에 이식될 수 있다. 한 예에서, 종양 세포 또는 세포주는 전신 순환에 직접 주사된다. 주사 부위는 주로 상기 실험 시스템에서 전이가 발달하는 부위를 규정한다. 실험 전이 모델에 대해 이용되는 종양 세포 주사의 대부분의 공통 부위는 마우스의 측면 꼬리 정맥이며, 이는 주로 폐 전이를 발생시킨다. 대조적으로, 종양 세포의 비장내 또는 문맥 주사가 간에서의 전이를 발달시키는데 이용되는 가장 흔한 부위이며, 세포의 심장내 주사는 뼈를 포함하는 여러 부위로의 전이를 발생시킬 수 있다. 순환으로의 종양 세포 또는 다른 세포주 주사 후, 관심 부위(예를 들어, 폐)에서의 전이의 발달이 수일 또는 수주의 기간에 걸쳐 모니터링된다.
- [0296] 종양 전이를 평가하기 위한 또 다른 모델은 같은자리 이식을 이용하며, 여기서 암세포는 해부학적 위치 또는 종양이 유래되는 조직으로 이식된다(예를 들어, 종양 단편의 직접적 주사 또는 외과적 이식에 의함). 같은자리

종양으로부터 발생하는 자연 전이는 수일 또는 수주의 기간에 걸쳐 평가될 수 있다. 종양 전이를 감소시키거나 예방하는 시험 화합물의 능력은 종양 세포의 피하, 근내 또는 순환으로의 주사 후 동물로 시험 화합물을 투여하거나, 같은자리 이식에 의해 평가될 수 있다. 전이 발달의 수, 크기 또는 시간이 평가될 수 있다. 종양 전이를 억제하는 화합물은 대조군 샘플에 비해 전이의 수를, 예를 들어, 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 심지어 100%까지 감소시킬 수 있다. 종양 전이를 억제하는 화합물은 또한 대조군 샘플에 비해 전이의 크기를 감소시킬 수 있다. 유사하게, 종양 전이를 억제하는 화합물은 전이의 발달의 발생을, 예를 들어, 적어도 1주, 2주, 1개월, 6개월, 1년, 또는 심지어 무기한까지 지연시킬 수 있다.

[0297] **VII. 예시적 ALK 억제제**

[0298] 본원에 개시된 방법은 종양 세포 사멸을 유도하거나, 종양 성장을 감소시키거나, 종양 전이의 위험을 감소시키기 위해 표 1에 나열된 본 발명의 신규한 바이오마커(예를 들어, ALK 돌연변이)의 억제제를 이용한 치료에 대한 후보자로서 피검체를 확인하는 것을 포함한다. ALK 폴리펩티드의 억제제는 당업자에게 공지되어 있다. 예를 들어, PF-02341066, PDD, 2-메틸-11-(2-메틸프로필)-4-옥소-4,5,6,11,12,13-헥사하이드로-2H-인다졸로[5,4-a]피롤로[3,4-c]카르바졸-8-일[4-(디메틸아미노)벤질]카르바메이트, (1S,2S,3R,4R)-3-({5-클로로-2-[(1-에틸-2,3,4,5-테트라하이드로-6-메톡시-2-옥소-1H-1-벤즈아제핀-7-일)아미노]-4-피리미디닐}아미노)바이시클로[2.2.1]헵트-5-엔-2-카르복사미드, 및 NVP-TAE684가 있으며, 이에 대해서는, 예를 들어, 참조로서 본원에 포함되는 문헌[PNAS 104:270-275, 2007; Choi, Y.L. et al. (2008) *Cancer Res.* 68:4971-2976; 및 *Biochemistry* 48:3600-3609, 2009]을 참조하라.

[0299] **참조로서의 포함**

[0300] 본원에 언급된 모든 간행물, 특허, 및 특허 출원은, 각각의 개별적 간행물, 특허 또는 특허 출원이 참조로서 포함되는 것으로 특별히 및 개별적으로 지시된 것과 같이, 전체내용이 참조로서 본원에 포함된다. 불일치의 경우, 본원의 임의의 정의를 포함하는 본 출원이 우선할 것이다.

[0301] 또한, tigr.org의 월드 와이드 웹 주소의 미국 유전체 연구소(The Institute for Genomic Research, TIGR) 및/또는.ncbi.nlm.nih.gov의 월드 와이드 웹 주소의 미국 국립생물공학정보센터(미국 국립생물공학정보센터, NCB I)에 의해 유지되는 것과 같은 공적인 데이터베이스 내의 기재에 따른 등록 번호를 참조로 하는 임의의 폴리뉴클레오티드 및 폴리펩티드 서열은 이의 전체내용이 참조로서 포함된다.

[0302] **예시**

[0303] 본 발명은 하기 실시예에 의해 추가로 예시되나, 이로 제한되는 것으로 간주되어선 안된다. 본 출원을 통해 인용된 모든 참고문헌, 도면, 서열 목록, 특허 및 공개된 특허 출원의 내용은 참조로서 본원에 포함된다.

[0304] **실시예 1 - 실시예 2-4에 대한 재료 및 방법**

[0305] a. DNA 시퀀싱

[0306] EZ1 시스템(Qiagen, Valencia, CA)을 사용하여 추출된 시료 RNA로부터 올리고(dT)-프라이밍된 cDNA를 생성시키고, PrimeSTAR HS DNA 중합효소(Takara Bio Inc., Shiga, Japan) 및 프라이머 ALK-TK-F(5'-TACAACCCCACTACTGCTTGTCT-3') 및 ALK-TK-R1(5'-AGGCACCTTCTCTCTCTCCAC-3')를 이용하여 30주기(10초 동안 98°C 및 1분 동안 68°C로 구성됨)의 중합효소 연쇄반응(PCR)에 적용시켰다. 이후, ALK의 키나제 도메인에 상응하는 PCR 생성물을 단편화시키고, 페어드-엔드 시퀀싱 시스템(paired-end sequencing system)(Illumina, San Diego, CA)에 의해 양 말단으로부터 76개의 염기에 대해 Illumina Genome Analyzer II(GAII)로 시퀀싱하였다. 가공되지 않은 판독 데이터를 PCR 프라이머 서열의 존재 및 모든 염기에 대해 20 이상의 Q 값을 기초로 하여 특성 여과(quality-filter)시켰다. 특성 여과된 판독값을 이후 Bowtie 알고리즘(bowtie-bio.sourceforge.net/index.shtml의 월드 와이드 웹에서 이용가능함)을 사용하여 ALK cDNA 서열에 정렬시켰다.

[0307] 3130xl Genetic Analyzer(Applied Biosystems, Foster City, CA)를 이용한 모세관 시퀀싱(capillary sequencing)을 위해, 동일 프라이머 세트 또는 EA-F-g-S(5'-CCACACCTGGGAAAGGACCTAAAG-3') 및 ALK-TK-R2(5'-CCTCCAAATACTGACAGCCACAGG-3') 프라이머의 조합을 이용하여 cDNA로부터 PCR 생성물을 제조하였다.

[0308] b. 돌연변이 EML4-ALK

[0309] FLAG 에피토프-태깅된 EML4-ALK 변이체 1을 엔코딩하는 cDNA(Soda, M. et al. (2007) *Nature* 448:561-566)를

FLAG-태깅된 EML4-ALK 및 마우스 CD8의 동시 발현을 위해 pMX-iresCD8 레트로바이러스 벡터(Yamashita Y. et al. (2001) J. Biol. Chem. 276:39012-39020)에 삽입하였다. ALK의 C1156Y 및 L1196M 돌연변이에 상응하는 누클레오타이드 변화를 EML4-ALK(C1156Y), EML4-ALK(L1196M), 또는 EML4-ALK(C1156Y/L1196M)의 발현을 위해 개별적으로 또는 조합하여 도입시켰다. 상기 플라스미드를 기초로 한 재조합 레트로바이러스를 패키징 세포주, BOSC23(Pear, W.S. et al. (1993) Proc. Natl. Acad. Sci. USA 90:8392-8396)을 사용하여 생성시키고, 마우스 인터루킨-3-의존성 세포주 BA/F3(Palacios, R. et al. (1985) Cell 41:727-734)를 감염시키는데 사용하였다. 생성된 CD8-양성 세포를 miniMACS 세포 분리 컬럼 및 CD8에 대한 항체에 컨주게이션된 자기 비드(둘 모두 Miltenyi Biotec, Gladbach, Germany)를 사용하여 정제시켰다. PF-02341066을 Selleck으로부터 수득하였다.

[0310] EML4-ALK의 티로신 인산화의 시험을 위해, 융합 단백질을 발현하는 BA/F3 세포를 15시간 동안 ALK 억제제에 노출시킨 후, FLAG에 대한 항체(Sigma-Aldrich, St. Louis, MO)를 이용하여 세포 용해질로부터 EML4-ALK를 면역침전시키고, Tyr¹⁶⁰⁴-인산화된 ALK에 대한 항체(Cell Signaling Technology, Danvers, MA)를 이용하여 면역블롯 분석하였다. 시험관내 키나제 검정을 합성 YFF 펩티드(Operon Biotechnologies, Huntsville, AL)를 이용하여 이전에 기재(Donella-Deana, A. et al. (2005) Biochemistry 44:8533-8542)된 바와 같이 30분 동안 실온에서 수행하였다.

[0311] **실시예 2 - ALK 티로신 키나제 억제제에 대한 내성과 관련된 신규한 ALK 돌연변이**

[0312] 환자는 흡연 경력이 없는 28세의 남성이었고, 2008년 4월에 T4N3M1의 임상 단계에서 폐 샘암종을 갖는 것으로 진단되었다. 종양이 임의의 EGFR 돌연변이를 포함하지 않는 경우, 환자를 통상적인 화학요법으로 치료하였고, 이는 뇌 및 뼈에서 다수의 전이가 형성되어 질병 진행을 발생시켰다. 2008년 11월에, 종양에서 EML4-ALK 변이체 1에 대한 mRNA의 존재를 가래의 역전사-PCR 분석 뿐만 아니라 생검 시료의 인 시츄 하이브리드화에서의 형광에 의해 확인하였다. 따라서, 환자는 PF-02341066의 시험에 등록되었고, 환자의 수행 상태에서 현저한 개선을 경험하였다(수준 4로부터 2로의 감소). 환자는 치료에 대한 "부분적 반응"을 나타내었으나, 환자의 흉막삼출액은 전체적으로 근치되지 않았다. 그러나, 치료 5개월 후, 종양은 돌연히 다시 성장하기 시작하여, 흉막삼출액의 증가 및 양쪽 폐에서의 다수의 암 결절 형성을 발생시켰다. 환자는 2009년 5월에 상기 시험으로부터 중단되었고, 이후 분자 분석을 위해 흉막삼출액이 수득되었다.

[0313] 종양이 ALK 억제제의 지속된 투여에도 불구하고 다시 성장하기 시작한 경우, 종양이 약물에 대한 내성을 부여하는 이차 유전적 변화를 획득하였는지의 여부를 결정하였다. 또한, TKI에 대한 내성이 종종 표적 키나제 내에서 획득된 돌연변이로부터 발생하는 경우, EML4-ALK 자체가 아미노산 변화를 겪을 가능성을 시험하였다.

[0314] 가래(ID J-#1) 및 흉막삼출액(ID J-#113) 시료는 치료 전 및 치료 후 환자 종양의 분자 분석에 각각 이용가능하였다. 2개의 시료에서의 종양 세포의 비율이 상이한 경우, 상기 시료로부터 유래된 EML4-ALK cDNA의 딥 시퀀싱을 수행하기 위해 차세대 서열분석기를 사용하였다. 따라서, ALK의 티로신 키나제 도메인에 상응하는 cDNA를 둘 모두의 시료로부터 증폭시키고(도 1A), 단편화시키고, GAII 시스템을 이용한 분자 시퀀싱에 적용시켰다. 비교를 위해, 융합 단백질에 대해 또한 양성인 EML4-ALK-양성 NSCLC 세포주, H2228, 및 3개의 다른 임상 시료를 유사하게 분석하였다. 공지된 단일 누클레오타이드 다형성, rs3795850을 4개의 시료로부터의 cDNA에서 검출하였다(도 1B). 또한, 인간 야생형 ALK cDNA(GenBank accession number, NM_004304)의 누클레오타이드 4230에 상응하는 위치에서의 T→C 변화를 J-#1 cDNA에서 낮은 빈도(8.9%)로 검출하였다. 또한, 야생형 ALK cDNA의 누클레오타이드 4374 및 4493에 상응하는 위치에서의 G→A 및 C→A 변화를 J-#113 cDNA에서 각각 41.8 및 14.0%의 빈도로 검출하였다. 상기 시료 중 임의의 시료로부터 유래된 키나제-도메인 cDNA에서 다른 회귀성 변화(관독에서 5% 이상으로 존재)는 없었다.

[0315] 이러한 누클레오타이드 변화를 생거 서열분석기를 이용하여 확인하였다. EML4-ALK에서가 아닌 내인성 야생형 ALK에서 돌연변이가 발생할 가능성을 배제시키기 위해, EML4 cDNA를 표적으로 하는 정방향 프라이머로 PCR을 또한 수행하여, 융합 cDNA만 증폭시켰다(도 1A). T4230C 변화가 J-#1으로부터 유래된 수백개의 융합 cDNA에서 검출되지 않았고, 이는 최초 PCR 또는 GAII 시퀀싱 단계에서 발생한 인공물이었음을 나타낸다.

[0316] 그러나, G4374A 및 C4493A 변화 둘 모두는 생거 시퀀싱에 의해 용이하게 확인되었다. J-#113에 대해 시퀀싱된 73개의 융합 cDNA 클론에서, 34개의 클론(46.6%)이 G4374A에 대해 양성이었고, 11개의 클론(15.1%)이 C4493A에 대해 양성이었고, 나머지(38.4%)는 야생형이었다(도 1C). PCR 분석은 동일 생성물 내의 둘 모두의 누클레오타이드 위치를 포함하나, 생성물 어느 것도 둘 모두의 돌연변이를 함유하지 않았으며, 이는 각각의 돌연변이가 독립적으로 발생한 것을 나타낸다. G4374 또는 C4493 위치를 포함하는 유전체 단편을 또한 PCR에 의해 증폭시키고,

뉴클레오티드 시퀀싱에 적용시켜, 종양 유전체 내에서의 각각의 변화를 확인하였다(도 2).

[0317] G4374A 및 C4493A 치환은 야생형 인간 ALK의 아미노산 1156 및 1196에 상응하는 위치에서 각각 Cys→Tyr 및 Leu→Met 변화를 발생시켰다.

[0318] **실시예 3 - 신규한 ALK 돌연변이는 ALK 티로신 키나제 억제제에 대한 내성을 부여한다.**

[0319] 상기 아미노산 변화가 ALK 억제제에 대한 EML4-ALK의 민감성에 영향을 미치는지의 여부를 다음으로 시험하였다. 야생형 EML4-ALK, 단일 돌연변이 EML4-ALK(C1156Y) 및 EML4-ALK(L1196M), 및 이중 돌연변이 EML4-ALK(C1156Y/L1196M)를 BA/F3 세포에서 개별적으로 발현시킨 후, 세포를 ALK 억제제에 노출시켰다. PF-02341066은 야생형 EML4-ALK를 발현하는 BA/F3 세포의 성장을 농도-의존 방식으로 억제하였다(도 4A). 대조적으로, C1156Y 또는 L1196M 돌연변이를 발현하는 세포는 상기 약물에 대해 현저하게 감소된 민감성을 나타내었고, 반복된 실험은 EML4-ALK(L1196M)를 발현하는 BA/F3 세포가 EML4-ALK(C1156Y)를 발현하는 BA/F3 세포보다 PF-02341066에 대해 더욱 내성인 것을 나타내었다(도 3). 둘 모두의 돌연변이의 존재는 PF-02341066에 대한 세포의 내성에 추가 영향을 발생시키지 않았다. 따라서, 이러한 데이터는 C1156Y 및 L1196M 돌연변이가 각각 상기 약물에 대해 내성을 부여하는 것을 나타내었다.

[0320] EML4-ALK의 티로신 인산화를 Tyr1604에서 인산화된 ALK에 특이적인 항체를 이용한 면역블롯 분석으로 시험하였다. PF-02341066에 대한 BA/F3 세포의 노출은 야생형 EML4-ALK의 티로신 인산화를 현저히 억제하였으나, EML4-ALK(C1156Y) 또는 EML4-ALK(L1196M)의 티로신 인산화에 대해서는 실질적 효과가 없었다(도 4B). 상기 발견과 일치하게, 시험관내 키나제 검정은 EML4-ALK의 C1156Y 및 L1196M 돌연변이가 야생형 단백질보다 PF-02341066에 의한 효소 활성의 억제에 덜 민감한 것을 나타내었다(도 4C). 세포 성장의 억제에 대한 경우와 같이(도 4A), L1196M 돌연변이는 C1156Y 돌연변이보다 PF-02341066에 의한 키나제 활성의 억제에 대해 더욱 무반응성이었다(도 4C).

[0321] **실시예 4 - 신규한 ALK 돌연변이 사이의 구조-기능 관계 및 ALK 티로신 키나제 억제제에 대한 내성**

[0322] 도 5는 인슐린 수용체인 관련 키나제의 결정 구조를 기초로 한 ALK의 키나제 도메인의 3차원 구조 모델에서의 Cys1156 및 Leu1196의 위치를 도시한다. Cys1156은 예측 헬릭스 αC의 아미노-말단에 인접할 뿐만 아니라 ATP-결합 포켓의 상부 리드(lid)에 가깝게 위치된다. 다른 티로신 키나제에서 상기 위치에서의 활성화 돌연변이는 보고되지 않았다. ALK의 Leu1196은 ABL1의 Thr315 및 EGFR의 Thr790에 상응하며, 이들 각각은 상기 키나제에서 TKI에 대한 내성을 부여하는 가장 빈번히 획득되는 돌연변이의 부위이다(Deininger, M. *et al.* (2005) *Blood* 105:2640-2653; Linardou, H. *et al.* (2009) *Nat. Rev. Clin. Oncol.* 6:352-366). 이러한 "게이트키퍼(gatekeeper)" 부위는 ATP-결합 포켓의 하부면에 위치되며(도 5), 상기 위치에서 부피가 큰 측쇄를 갖는 아미노산의 존재는 많은 TKI의 결합을 간섭하는 것으로 공지되어 있다(Shah, N.P. *et al.* (2002) *Cancer Cell* 2:117-125; Tsao, M.S. *et al.* (2005) *N. Engl. J. Med.* 353: 133-144).

[0323] 따라서, 다수의 ALK 억제제에 대한 내성을 부여하는 EML4-ALK의 키나제 도메인 내의 2개의 새로운 돌연변이가 확인되었다. 둘 모두의 돌연변이를 포함하는 EML4-ALK cDNA가 관찰되지 않는 경우, 각각의 돌연변이가 종양의 별개의 서브클론에서 독립적으로 발달한 것으로 생각된다.

[0324] 이론으로 제한하고자 하는 바는 아니지만, 치료 전의 환자의 가래로부터 제조된 cDNA가 C1156Y 또는 L1196M 돌연변이에 상응하는 뉴클레오티드 변화를 함유하지 않는 경우, 종양 서브클론이 PF-02341066을 이용한 치료 동안 새로이 돌연변이를 획득했을 수 있다. 그러나, 흉막삼출액은 처리 전에 시험될 수 없으므로, C1156Y 또는 L1196M 돌연변이를 포함하는 종양 세포가 환자의 최초 입원시 흉막삼출액에 이미 존재했을 가능성이 완전히 배제될 수 없었다. 이러한 경우, 종양은 이의 이후의 신속한 성장을 가능케 하는 PF-02341066을 이용한 5개월의 치료 동안 다른 아직 공지되지 않은 돌연변이를 획득했을 수 있다. 그러나, C1156Y 또는 L1196M 돌연변이를 갖는 종양 세포의 서브클론은 최초 치료에 대해 난치성일 것이며, 치료 과정 동안 확장될 것이다. 이와는 반대로, 적어도 5개월 동안 환자에서 종양 확장의 징후가 없었고, 이는 C1156Y 및 L1196M 돌연변이가 PF-02341066을 이용한 치료 동안 발달한 것을 나타낸다. 이러한 생각은 제피티니브 및 에를로티니브에 대한 내성을 부여하는 EGFR의 T790M 돌연변이가 TKI로 이전에 치료된 환자에서 빈번히 검출되나, 치료되지 않은 개체에서는 드물게 발견된다는 사슬에 의해 추가로 뒷받침된다(Pao, W. *et al.* (2005) *PLoS Med.* 2:e73).

[0325] 여러 티로신 키나제의 게이트키퍼 위치에서의 아미노산 치환이 TKI로 치료된 종양에서 검출되었다(Kobayashi, S. *et al.* (2005) *N. Engl. J. Med.* 352:786-792; Pao, W. *et al.* (2005) *PLoS Med.* 2:e73; Shah, N.P. *et al.* (2002) *Cancer Cell* 2:117-125; Cools, J. *et al.* (2003) *N. Engl. J. Med.* 348:1201-1214; Tamborini,

E. *et al.* (2004) *Gastroenterology* 127:294-299). 상기 위치에서의 돌연변이가 EML4-ALK 또는 ALK에 대해 보고되지 않은 반면, ALK에 대한 또 다른 융합형 온코키나제(oncokinase)인 NPM-ALK의 게이트키퍼 위치에서의 다양한 인공적 아미노산 치환의 효과가 최근에 시험되었다(Lu, L. *et al.* (2009) *Biochemistry* 48:3600-3609). 생체내에서의 종양 세포의 본 발명의 분석과 일치하게, 상기 위치에서의 Met의 도입은 NPM-ALK가 다수의 ALK 억제제에 대해 가장 내성이 되도록 하는 것을 발견하였다.

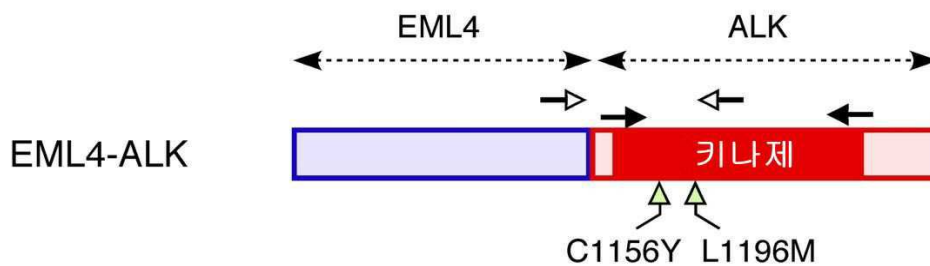
[0326] 게이트키퍼 치환과 대조적으로, αC 헬릭스(ALK 중 Cys1156)에 대해 아미노-말단에 가까운 위치에서의 활성화 돌연변이는 다른 티로신 키나제에 대해 보고되지 않았다. EGFR의 상응하는 위치에서의 Thr→Ile 변화가 하나의 NSCLC 경우에서 기재된 반면, 이의 약물 민감성에 대한 관련성은 시험되지 않았다(Tsao, M.S. *et al.* (2005) *N. Engl. J. Med.* 353:133-144). 효소 활성의 알로스테리 조절에 대한 헬릭스 αC의 중요성이 세린-트레오닌 키나제에 대해 입증되었다(Hindie, V. *et al.* (2009) *Nat. Chem. Biol.* 5:758-764). 따라서, ALK의 Cys1156에서의 변화는 TKI 결합을 알로스테릭하게 간섭하거나, Cys1156이 키나제 도메인과 TKI 사이의 물리적 상호작용과 직접적으로 관련될 것이다.

[0327] 동등부

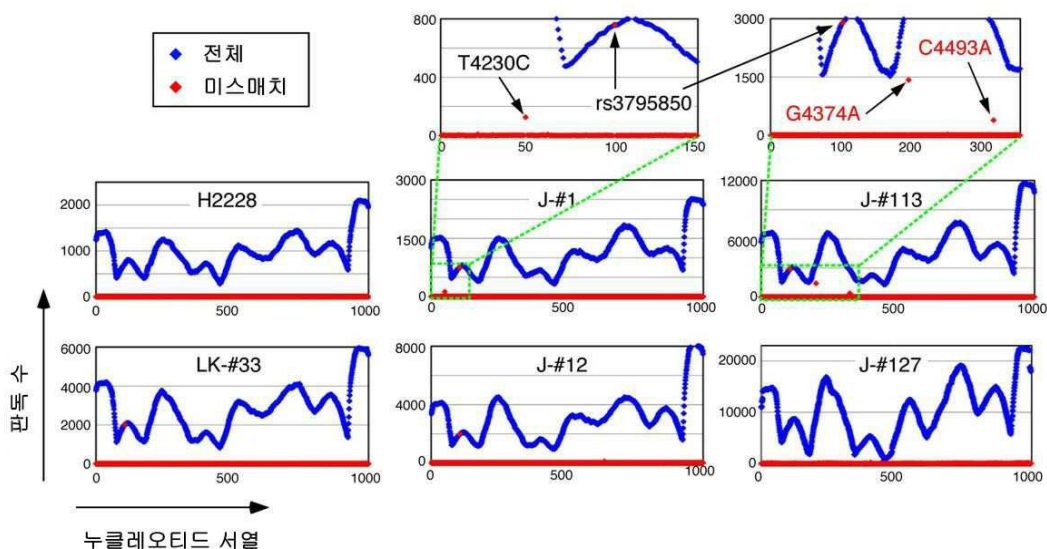
[0328] 당업자는 본원에 기재된 본 발명의 특정 구체예에 대한 많은 동등부를 인지하거나, 통상적인 것을 넘어서지 않는 실험을 이용하여 상기 동등부를 확인할 수 있을 것이다. 이러한 동등부는 하기 첨부되는 청구항에 포함된다.

도면

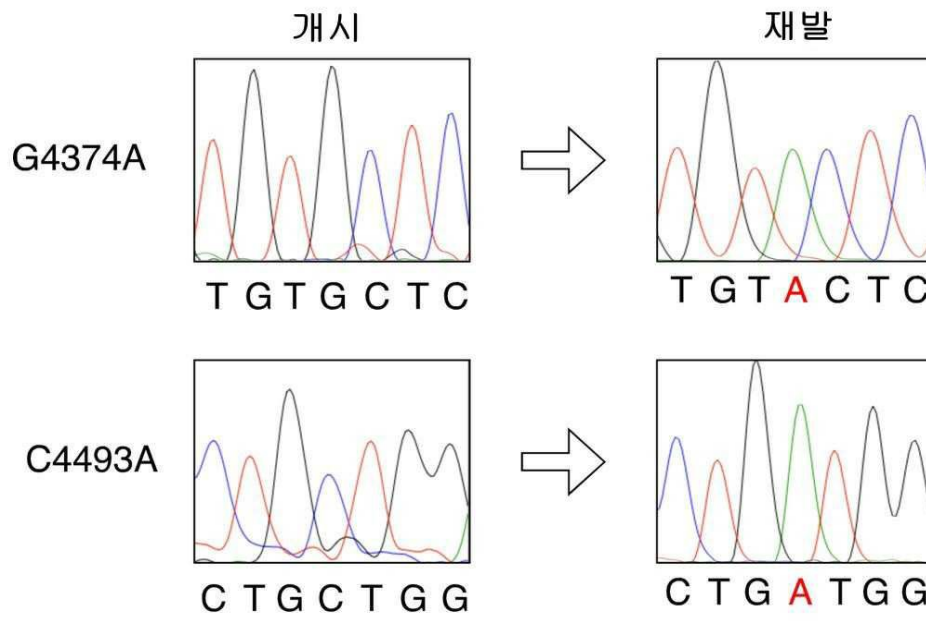
도면1a



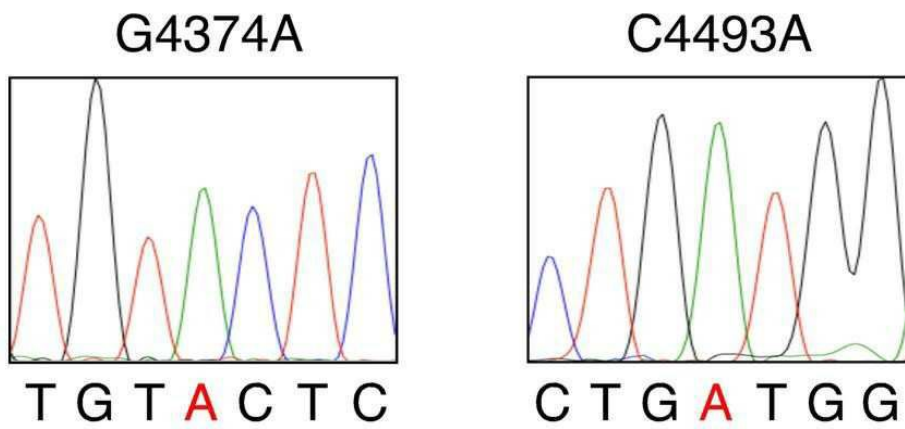
도면1b



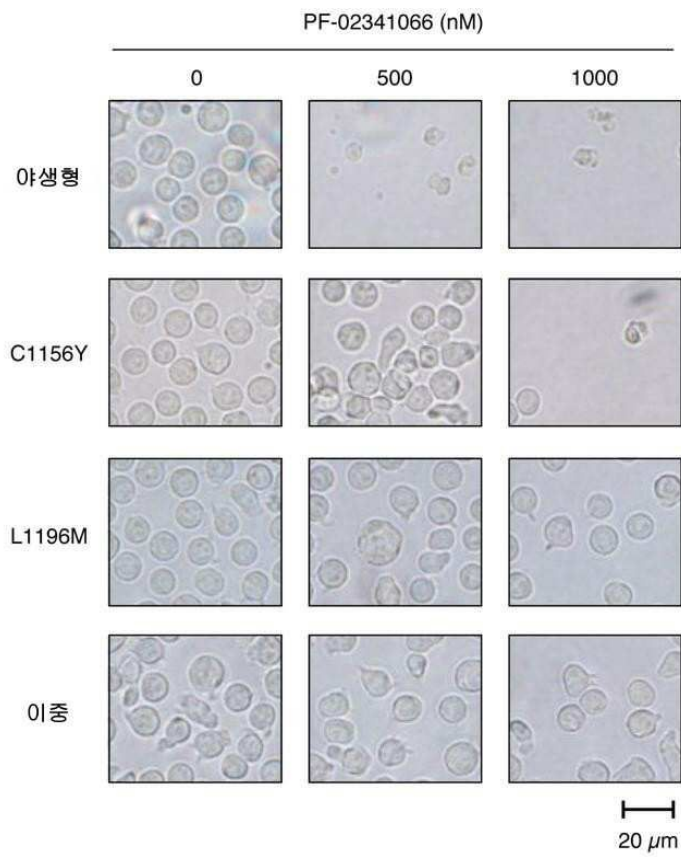
도면1c



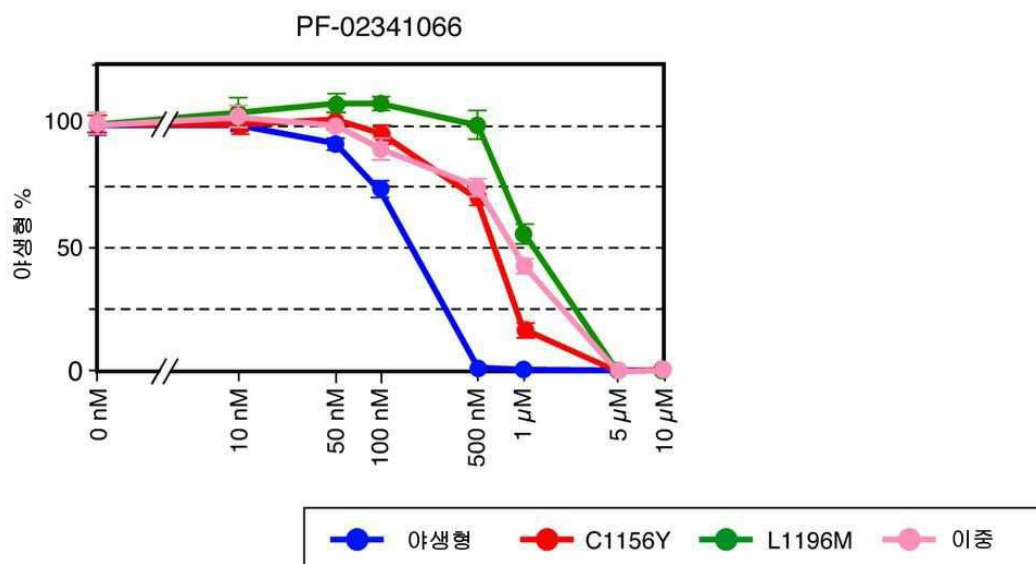
도면2



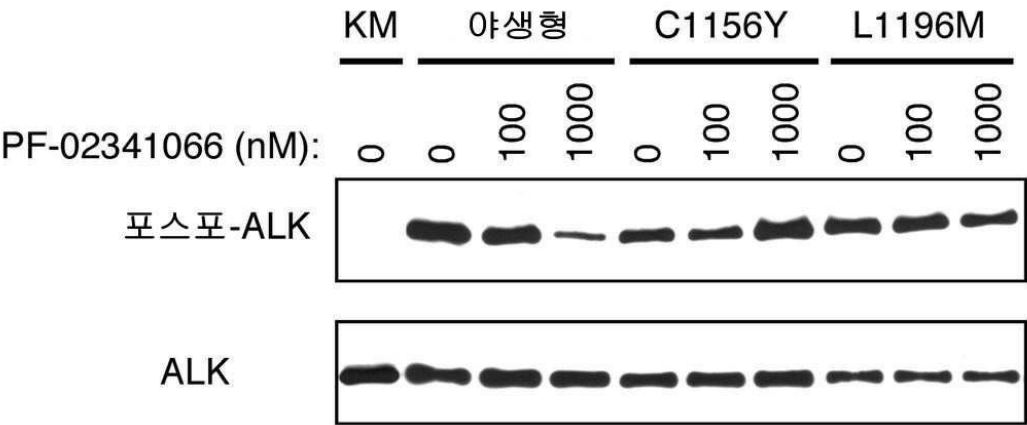
도면3



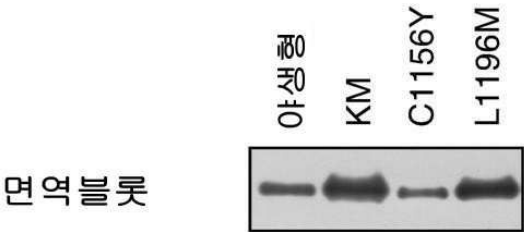
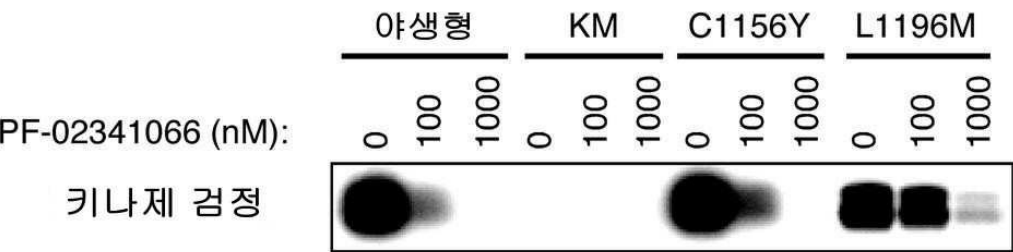
도면4a



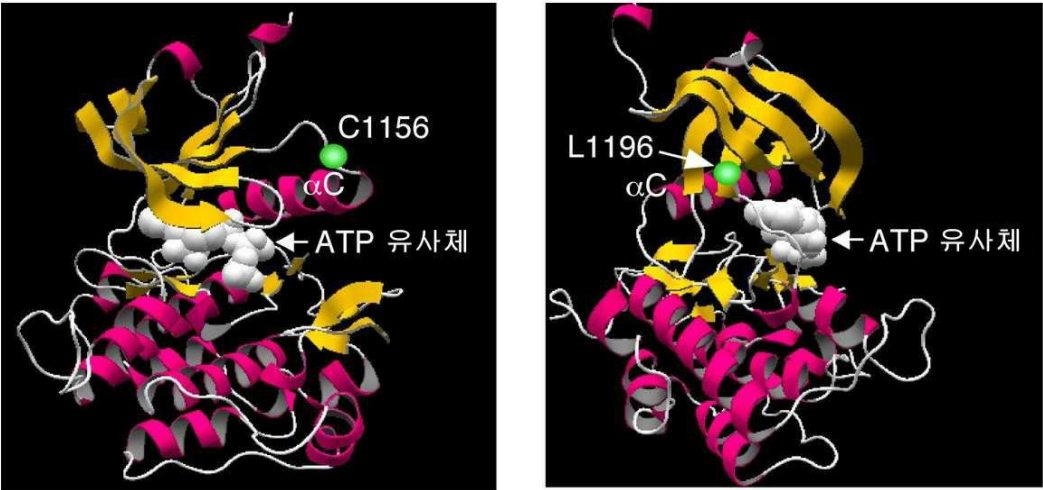
도면4b



도면4c



도면5



서열목록

SEQUENCE LISTING

<110> JICHI MEDICAL UNIVERSITY

<120> IDENTIFICATION, ASSESSMENT, AND THERAPY OF CANCERS WITH
INNATE OR ACQUIRED RESISTANCE TO ALK INHIBITORS

<130> CGX-003.25

<140> PCT/IB2011/000382

<141> 2011-02-04

<150> 61/337,465

<151> 2010-02-04

<160> 40

<170> PatentIn version 3.5

<210> 1

<211> 6222

<212> DNA

<213> Homo sapiens

<400> 1

gggggcggca gcggtggttag cagctggtac ctcccgcgc ctctgttcgg agggtcgcgg	60
ggcaccgagg tgctttccgg ccgccctctg gtcggccacc caaagccgcg ggcgctgatg	120
atgggtgagg agggggcggc aagatttcgg gcgccctgc cctgaacgcc ctgagctgct	180
gccgccgggg ccgctccagt gcctgcgaac tctgaggagc cgaggcgccg gtgagagcaa	240
ggacgttgca aacttgcgca gcgcgggggc tgggattcac gcccagaagt tcagcaggca	300
gacagtccga agccttcccg cagcggagag atagcttgag ggtgcgcaag acggcagcct	360
ccgccctcgg ttcccgcga gaccgggcag aagagcttgg aggagccaaa aggaacgcaa	420
aaggcggcca ggacagcgtg cagcagctgg gagccgccgt tctcagcctt aaaagttgca	480
gagattggag gctgccccga gaggggacag accccagctc cgactgcggg gggcaggaga	540
ggacggtacc caactgccac ctcccttcaa ccatagtagt tcctctgtac cgagcgcagc	600
gagctacaga cgggggcgcg gcactcggcg cggagagcgg gaggtcaag gtcccagcca	660
gtgagcccag tgtgcttgag tgtctctgga ctgcgccctg agcttccagg tctgtttcat	720
ttagactcct gctcgcctcc gtgcagttgg gggaaagcaa gagacttgcg cgacgcaca	780
gtcctctgga gatcaggtgg aaggagccgc tgggtaccaa ggactgttca gacctcttc	840
ccatctcggg gagagcgaag ggtgaggctg ggcccgga gacgtgtaaa cggcctcctc	900

cgggcgggatg ggagccatcg ggctcctgtg gtccttgccg ctgctgcttt ccacggcagc 960

tgtgggctcc ggcatgggga ccggccagcg cgcgggctcc ccagctgcgg ggccgccgct 1020

gcagccccgg gagccactca gctactcgcg cctgcagagg aagagtctgg cagttgactt 1080

cgtggtgccc tcgtctctcc gtgtctacgc ccgggaccta ctgctgccac catcttctc 1140

ggagctgaag gctggcagcg ccgaggcccg cggctcgcta gctctggact gcgccccgct 1200

gctcaggttg ctggggccgg cgcgggggt ctcctggacc gccggttcac cagccccggc 1260

agaggcccg acgtgtcca ggggtctgaa gggcggctcc gtgcgcaagc tccggcgtgc 1320

caagcagttg gtgctggagc tgggcgagga ggcgatcttg gagggttgcg tcgggcccc 1380

cggggaggcg gcgtggggc tgctccagtt caatctcagc gagctgttca gttgggtgat 1440

tcgccaagcg gaagggcgac tgaggatccg cctgatgccc gagaagaagg cgtcggaagt 1500

gggcagagag ggaagctgt ccgcgccaat tcgcgctcc cagccccgcc ttctcttcca 1560

gatcttcggg actggtcata gctccttga atcaccaaca aacatgcctt ctcttctcc 1620

tgattatattt acatggaatc tcacctggat aatgaaagac tccttcctt tcctgtctca 1680

tcgcagccga tatggtctgg agtgcagctt tgacttcccc tgtgagctgg agtattcccc 1740

tcactgcat gacctagga accagagctg gtcttgccgc cgcattccct ccgaggaggc 1800

ctcccagatg gacttgctgg atgggcctgg ggcagagcgt tctaaggaga tgcccagagg 1860

ctctttctc ctctcaaca cctcagctga ctccaagcac accatcctga gtccgtggat 1920

gaggagcagc agtgagcact gcacactggc cgtctcggtg cacaggcacc tgcagccctc 1980

tggaaggtag attgcccagc tgctgcccc caacgaggct gcaagagaga tcctctgat 2040

gcccactcca gggaagcatg gttggacagt gctccaggga agaactgggc gtccagacaa 2100

ccatttcga gtggccctgg aatacatctc cagtggaaac cgcagcttgt ctgcagtgga 2160

cttctttgcc ctgaagaact gcagtgaagg aacatcccca ggctccaaga tggccctgca 2220

gagctccttc acttggttga atgggacagt cctccagctt gggcaggcct gtgacttcca 2280

ccaggactgt gccaggagg aagatgagag ccagatgtgc cggaactgc ctgtgggttt 2340

ttactgcaac ttigaagatg gcttctgtgg ctggacccaa ggcacactgt caccacacac 2400

tcctcaatgg caggtcagga ccctaaagga tgcccgttc caggaccacc aagaccatgc 2460

tctattgctc agtaccactg atgtccccgc ttctgaaagt gctacagtga ccagtctac 2520

gtttctgca ccgatcaaga gctctccatg tgagctccga atgtcctggc tcattcgtgg 2580

agtcttgagg ggaacgtgt ccttggtgct agtggagaac aaaaccggga aggagcaagg 2640

caggatggtc tggcatgtcg ccgcctatga aggcttgagc ctgtggcagt ggatgggtgtt 2700
 gcctctcttc gatgtgtctg acaggtttctg gctgcagatg gtcgcatggg ggggacaagg 2760
 atccagagcc atcgtggctt ttgacaatat ctccatcagc ctggactgct acctcaccat 2820
 tagcggagag gacaagatcc tgcagaatac agcacccaaa tcaagaaacc tgtttgagag 2880
 aaacccaaac aaggagctga aaccggggga aaattcacca agacagacc ccattttga 2940
 ccctacagtt cattggctgt tcaccacatg tggggccagc gggcccatg gcccaccca 3000
 ggcacagtgc aacaacgcct accagaactc caacctgagc gtggaggtgg ggagcgaggg 3060

 cccctgaaa ggcatccaga tctggaaggt gccagccacc gacacctaca gcatctcggg 3120
 ctacggagct gctggcggga aaggcgggaa gaacacatg atgcgggtccc acggcgtgtc 3180
 tgtgctgggc atcttcaacc tggagaagga tgacatgctg tacatcctgg ttgggcagca 3240
 gggagaggac gcctgcccc gtacaaacca gttaatccag aaagtctgca ttggagagaa 3300
 caatgtgata gaagaagaaa tccgtgtgaa cagaagcgtg catgagtggg caggaggcgg 3360
 aggaggaggg ggtggagcca cctacgtatt taagatgaag gatggagtgc cgggtccct 3420
 gatcatgca gccggaggtg gtggcagggc ctacggggcc aagacagaca cgttccacc 3480

 agagagactg gagaataact cctcggttct agggctaaac ggcaattccg gagccgcagg 3540
 tgggtggaggt ggcgtggaatg ataacacttc ctgtctctgg gccggaaaat ctttgcagga 3600
 ggggtgccacc ggaggacatt cctgccccca ggccatgaag aagtgggggt gggagacaag 3660
 aggggggttc ggagggggtg gaggggggtg ctctcaggt ggaggaggcg gaggatatat 3720
 aggcggcaat gcagcctcaa acaatgacct cgaaatggat ggggaagatg gggtttcctt 3780
 catcagtcca ctgggcatcc tgtacacccc agctttaaaa gtgatggaag gccacgggga 3840
 agtgaatatt aagcattatc taaactgcag tcaactgtgag gtacacgaat gtcacatgga 3900

 ccctgaaagc cacaaggtca tctgtttctg tgaccacggg acggtgctgg ctgaggatgg 3960
 cgtctctctg attgtgtcac ccacccgga gccacacctg ccactctcgc tgatcctctc 4020
 tgtggtgacc tctgccctg tggccgcctt ggtcctggct ttctccggca tcatgattgt 4080
 gtaccgccgg aagcaccagg agctgcaagc catgcagatg gagctgcaga gccctgagta 4140
 caagctgagc aagctccgca cctcgacct catgaccgac tacaaccca actactgctt 4200
 tgctggaag acctctcca tcagtacct gaaggagtg ccgcgaaaa acataccct 4260
 cattcgggt ctggccatg gcgcctttgg ggaggtgtat gaaggccagg tgtccggaat 4320

 gcccacgac ccaagcccc tgcaagtggc tgtgaagacg ctgcctgaag tgtgctctga 4380
 acaggacgaa ctggatttcc tcatggaagc cctgatcatc agcaaattca accaccagaa 4440
 cattgttcgc tgcatgggg tgagcctgca atccctgcc cggttcatcc tgctggagct 4500

catggcgggg ggagacctca agtccttctc ccgagagacc cgccttcgcc cgagccagcc 4560
ctcctccctg gccatgctgg accttctgca cgtggctcgg gacattgcct gtggctgtca 4620
gtatttggag gaaaaccact tcatccaccg agacattgct gccagaaact gcctcttgac 4680
ctgtccaggc cciggaagag tggccaagat tggagacttc gggatggccc gagacatcta 4740

cagggcgagc tactatagaa agggaggctg tgccatgctg ccagttaagt ggatgcccc 4800
agaggccttc atggaaggaa tattcacttc taaaacagac acatggctct ttggagtgt 4860
gctatgggaa atcttttctc ttggatata gccatacccc agcaaaagca accaggaagt 4920
tctggagttt gtcaccagtg gaggccggat ggaccacccc aagaactgcc ctgggcctgt 4980
ataccggata atgactcagt gctggcaaca tcagcctgaa gacaggccca actttgccat 5040
cattttggag aggattgaat actgcacca ggaccggat gtaatcaaca ccgctttgcc 5100
gatagaatat ggtccacttg tggaagagga agagaaagtg cctgtgaggc ccaaggaccc 5160

tgagggggtt cctcctctcc tggctctca acaggcaaaa cgggaggagg agcgcagccc 5220
agctgcccc ccacctctgc ctaccacctc ctctggcaag gctgcaaaga aaccacagc 5280
tgagagatc tctgttcgag tccctagagg gccggccgtg gaagggggac acgtgaatat 5340
ggcattctct cagtcacacc ctccttcgga gttgcacaag gtccacggat ccagaaacaa 5400
gcccaccagc ttgtggaacc caacgtacgg ctcttggttt acagagaaac ccacaaaaa 5460
gaataatcct atagcaaaga aggagccaca cgacaggggt aacctggggc tggagggag 5520
ctgtactgtc ccacctaacg ttgcaactgg gagacttcg ggggcctcac tgctcctaga 5580

gccctcttcg ctgactgcca atatgaagga ggtacctctg ttcaggctac gtcacttccc 5640
ttgtgggaat gtcaattacg gctaccagca acagggttg cccttagaag ccgctactgc 5700
ccctggagct ggtcattacg aggataccat tctgaaaagc aagaatagca tgaaccagcc 5760
tgggccctga gtcggctgc acactcactt ctcttccttg ggatccctaa gaccgtggag 5820
gagagagagg caatggctcc ttcacaaacc agagaccaa tgtcacgttt tgttttgtc 5880
caacattttt tgaagtacca ccaaaaaagc tgtattttga aaatgcttta gaaagtttt 5940
gagcatgggt tcatcctatt ctttcgaaag aagaaaatat cataaaaatg agtgataaat 6000

acaaggccca gatgtggttg cataagggtt ttatgcatgt ttgttgtata cttccttatg 6060
cttctttcaa attgtgtgtg ctctgcttca atgtagtcag aattagctgc ttctatgttt 6120
catagttggg gtcatagatg tttccttgcc ttgttgatgt ggacatgagc catttgaggg 6180
gagagggaac ggaaataaag gagttatttg taatgactaa aa 6222

<210> 2

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 2

ggtaagaagt ggctcactct tgag

24

<210> 3

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 3

cacaacaact gcagcaaaga ctgg

24

<210> 4

<211> 1620

<212> PRT

<213> Homo sapiens

<400> 4

Met Gly Ala Ile Gly Leu Leu Trp Leu Leu Pro Leu Leu Ser Thr

1 5 10 15

Ala Ala Val Gly Ser Gly Met Gly Thr Gly Gln Arg Ala Gly Ser Pro

20 25 30

Ala Ala Gly Pro Pro Leu Gln Pro Arg Glu Pro Leu Ser Tyr Ser Arg

35 40 45

Leu Gln Arg Lys Ser Leu Ala Val Asp Phe Val Val Pro Ser Leu Phe

50 55 60

Arg Val Tyr Ala Arg Asp Leu Leu Leu Pro Pro Ser Ser Ser Glu Leu

65 70 75 80

Lys Ala Gly Arg Pro Glu Ala Arg Gly Ser Leu Ala Leu Asp Cys Ala

85 90 95

Pro Leu Leu Arg Leu Leu Gly Pro Ala Pro Gly Val Ser Trp Thr Ala

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

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

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

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

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

595 600 605
 Leu Gln Met Val Ala Trp Trp Gly Gln Gly Ser Arg Ala Ile Val Ala
 610 615 620
 Phe Asp Asn Ile Ser Ile Ser Leu Asp Cys Tyr Leu Thr Ile Ser Gly
 625 630 635 640
 Glu Asp Lys Ile Leu Gln Asn Thr Ala Pro Lys Ser Arg Asn Leu Phe
 645 650 655
 Glu Arg Asn Pro Asn Lys Glu Leu Lys Pro Gly Glu Asn Ser Pro Arg
 660 665 670

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

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

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

Tyr Gly Ala Lys Thr Asp Thr Phe His Pro Glu Arg Leu Glu Asn Asn
850 855 860

Ser Ser Val Leu Gly Leu Asn Gly Asn Ser Gly Ala Ala Gly Gly Gly
865 870 875 880

Gly Gly Trp Asn Asp Asn Thr Ser Leu Leu Trp Ala Gly Lys Ser Leu
885 890 895

Gln Glu Gly Ala Thr Gly Gly His Ser Cys Pro Gln Ala Met Lys Lys
900 905 910

Trp Gly Trp Glu Thr Arg Gly Gly Phe Gly Gly Gly Gly Gly Gly Cys
915 920 925

Ser Ser Gly Gly Gly Gly Gly Gly Tyr Ile Gly Gly Asn Ala Ala Ser
930 935 940

Asn Asn Asp Pro Glu Met Asp Gly Glu Asp Gly Val Ser Phe Ile Ser
945 950 955 960

Pro Leu Gly Ile Leu Tyr Thr Pro Ala Leu Lys Val Met Glu Gly His
965 970 975

Gly Glu Val Asn Ile Lys His Tyr Leu Asn Cys Ser His Cys Glu Val
980 985 990

Asp Glu Cys His Met Asp Pro Glu Ser His Lys Val Ile Cys Phe Cys
995 1000 1005

Asp His Gly Thr Val Leu Ala Glu Asp Gly Val Ser Cys Ile Val
1010 1015 1020

Ser Pro Thr Pro Glu Pro His Leu Pro Leu Ser Leu Ile Leu Ser
1025 1030 1035

Val Val Thr Ser Ala Leu Val Ala Ala Leu Val Leu Ala Phe Ser
1040 1045 1050

Gly Ile Met Ile Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala
1055 1060 1065

Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu
1070 1075 1080

Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe

1085	1090	1095
Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg		
1100	1105	1110
Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly		
1115	1120	1125
Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser		
1130	1135	1140
Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu		
1145	1150	1155
Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys		
1160	1165	1170
Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln		
1175	1180	1185
Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp		
1190	1195	1200
Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro		
1205	1210	1215
Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile		
1220	1225	1230
Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg		
1235	1240	1245
Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly		
1250	1255	1260
Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr		
1265	1270	1275
Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val		
1280	1285	1290
Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser		
1295	1300	1305
Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe		
1310	1315	1320

Ser Leu Gly Tyr Met Pro Tyr	Pro Ser Lys Ser Asn	Gln Glu Val
1325	1330	1335
Leu Glu Phe Val Thr Ser Gly	Gly Arg Met Asp Pro	Pro Lys Asn
1340	1345	1350
Cys Pro Gly Pro Val Tyr Arg	Ile Met Thr Gln Cys	Trp Gln His
1355	1360	1365
Gln Pro Glu Asp Arg Pro Asn	Phe Ala Ile Ile Leu	Glu Arg Ile
1370	1375	1380
Glu Tyr Cys Thr Gln Asp Pro	Asp Val Ile Asn Thr	Ala Leu Pro
1385	1390	1395
Ile Glu Tyr Gly Pro Leu Val	Glu Glu Glu Glu Lys	Val Pro Val
1400	1405	1410
Arg Pro Lys Asp Pro Glu Gly	Val Pro Pro Leu Leu	Val Ser Gln
1415	1420	1425
Gln Ala Lys Arg Glu Glu Glu	Arg Ser Pro Ala Ala	Pro Pro Pro
1430	1435	1440
Leu Pro Thr Thr Ser Ser Gly	Lys Ala Ala Lys Lys	Pro Thr Ala
1445	1450	1455
Ala Glu Ile Ser Val Arg Val	Pro Arg Gly Pro Ala	Val Glu Gly
1460	1465	1470
Gly His Val Asn Met Ala Phe	Ser Gln Ser Asn Pro	Pro Ser Glu
1475	1480	1485
Leu His Lys Val His Gly Ser	Arg Asn Lys Pro Thr	Ser Leu Trp
1490	1495	1500
Asn Pro Thr Tyr Gly Ser Trp	Phe Thr Glu Lys Pro	Thr Lys Lys
1505	1510	1515
Asn Asn Pro Ile Ala Lys Lys	Glu Pro His Asp Arg	Gly Asn Leu
1520	1525	1530
Gly Leu Glu Gly Ser Cys Thr	Val Pro Pro Asn Val	Ala Thr Gly
1535	1540	1545
Arg Leu Pro Gly Ala Ser Leu	Leu Leu Glu Pro Ser	Ser Leu Thr

1550	1555	1560	
Ala Asn Met Lys Glu Val Pro	Leu Phe Arg Leu Arg	His Phe Pro	
1565	1570	1575	
Cys Gly Asn Val Asn Tyr Gly	Tyr Gln Gln Gln Gly	Leu Pro Leu	
1580	1585	1590	
Glu Ala Ala Thr Ala Pro Gly	Ala Gly His Tyr Glu	Asp Thr Ile	
1595	1600	1605	
Leu Lys Ser Lys Asn Ser Met	Asn Gln Pro Gly Pro		
1610	1615	1620	
<210> 5			
<211> 3926			
<212> DNA			
<213> Homo sapiens			
<400> 5			
ggcggcgcg	cgcgcgctc	gcggctgctg	cctgggaggg aggccgggca ggcggctgag 60
cgcgcgcggt	ctcaacgtga	cggggaagtg	gttcgggcgg ccgcggctta ctacccagg 120
gcgaacggac	ggacgacgga	ggcgggagcc	ggtagccgag ccgggcgacc tagagaacga 180
gcgggtcagg	ctcagcgtcg	gccactctgt	cggtccgctg aatgaagtgc ccgccctct 240
gagcccgagg	cccggcgctt	tccccgaag	atggacggtt tcgccggcag tctcgatgat 300
agtatttctg	ctgcaagtac	ttctgatgtt	caagatcgcc tgtcagctct tgagtcacga 360
gttcagcaac	aagaagatga	aatcactgtg	ctaaaggcgg ctttggctga tgttttgagg 420
cgtcttgcaa	tctctgaaga	tcatgtggcc	tcagtgaaaa aatcagtctc aagtaaaggc 480
caaccaagcc	ctcgagcagt	tattcccatg	tctgtataa ccaatggaag tgggtgcaaac 540
agaaaaccaa	gtcataccag	tgctgtctca	attgcaggaa aagaaactct ttcattctgct 600
gctaaaagtg	gtacagaaaa	aaagaaagaa	aaaccacaag gacagagaga aaaaaaagag 660
gaatctcatt	ctaatgatca	aagtccacaa	attcgagcat caccttctcc ccagccctct 720
tcacaacctc	tccaaatata	cagacaaact	ccagaaagca agaatgtac tcccacaaaa 780
agcataaaac	gaccatcacc	agctgaaaag	tcacataatt cttgggaaaa ttcagatgat 840
agccgtaata	aattgtcgaa	aataccttca	acacccaaat taataccaaa agttacaaaa 900
actgcagaca	agcataaaga	tgatcatc	aaccaagaag gagaatatat taaaatgttt 960
atgcgcggtc	ggccaattac	catgttcatt	ccttccgatg ttgacaacta tgatgacatc 1020

agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaag 1080

gactgtagag ctaatgttta cttcttccg accgggaaa tagttatatt cattgcatca 1140

gtagtagtac tattaatta tgaggagaga actcagcgac actacctggg ccatacagac 1200

tgtgtgaaat gccttgctat acatctgac aaaattagga ttgcaactgg acagatagct 1260

ggcgtggata aagatggaag gcctctacaa cccacgtca gagtgtggga ttctgttact 1320

ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat 1380

ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactcaa tgagcatatg 1440

cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa 1500

gttgttttgg ctgtggagtt tcaccaaca gatgcaaata ccataattac atgcggtaaa 1560

tctcatattt tcttctggac ctggagcggc aattcactaa caagaaaaca gggaattttt 1620

gggaaatatg aaaagccaaa atttgtgcag tgtttagcat tcttggggaa tggagatgtt 1680

cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagcccaca 1740

cctgggaaag gacctaaagt gtaccgccgg aagcaccagg agctgcaagc catgcagatg 1800

gagctgcaga gccctgagta caagctgagc aagctccgca cctcgaccat catgaccgac 1860

tacaacccca actactgctt tgctggcaag acctcctcca tcagtacact gaaggaggtg 1920

ccgcggaaaa acatcacctt cattcggggt ctgggccatg gagcctttgg ggaggtgtat 1980

gaaggccagg tgtccggaat gcccaacgac ccaagccccc tgcaagtggc tgtgaagacg 2040

ctgcctgaag tigtctctga acaggacgaa ctggatttcc tcatggaagc cctgatcatc 2100

agcaaattca accaccagaa cattgttcgc tgcattgggg tgagcctgca atccctgccc 2160

cggttcatcc tgctggagct catggcgggg ggagacctca agtcttctc cagagagacc 2220

cgcctcgcgc cgagccagcc ctctcctctg gccatgctgg accttctgca cgtggctcgg 2280

gacattgcct gtggctgtca gtatttgag gaaaaccact tcatccacg agacattgct 2340

gccagaaact gcctcttgac ctgtccaggc cctggaagag tggccaagat tggagacttc 2400

gggatggccc gagacatcta caggcgagc tactatagaa agggaggctg tgccatgctg 2460

ccagttaagt ggatgcccc agaggccttc atggaaggaa tattcacttc taaaacagac 2520

acatggctct ttggagtgt gctatgggaa atcttttctc ttggatatat gccatacccc 2580

agcaaaagca accaggaagt tctggagttt gtcaccagtg gaggccgat ggaccaccc 2640

aagaactgcc ctgggcctgt ataccggata atgactcagt gctggcaaca tcagcctgaa 2700

gacaggccca actttgcat cattttggag aggattgaat actgcacca ggaccggat 2760

gtaatcaaca ccgtttgcc gatagaatat ggtccacttg tggaagagga agagaaagt 2820

cctgtgaggc ccaaggaccc tgagggggtt cctcctctcc tggctctca acaggcaaaa 2880

cgggaggagg agcgagccc agctgcccc aacacctctgc ctaccacctc ctctggcaag 2940
 gctgcaaaga aaccacagc tgcagaggtc tctgttcgag tccctagagg gccggccgtg 3000
 gaagggggac acgtgaatat ggcatctctc cagtccaacc ctcttcgga gttgcacagg 3060
 gtccacggat ccagaaacaa gccaccagc ttgtggaacc caacgtacgg ctctggttt 3120

 acagagaaac ccacaaaaa gaataatcct atagcaaaga aggagccaca cgagagggt 3180
 aacctggggc tggagggaag ctgtactgtc ccacctaacg ttgcaactgg gagacttccg 3240
 ggggcctcac tgctcctaga gccctcttcg ctgactgcca atatgaagga ggtacctctg 3300
 ttcaggtac gtcattccc ttgtgggaat gtcaattacg gctaccagca acagggttg 3360
 cccttagaag ccgtactgc ccctggagct ggctattacg aggataccat tctgaaaagc 3420
 aagaatagca tgaaccagc tgggcctga gtcggtcac acactcactt ctcttccttg 3480
 ggatccctaa gaccgtggag gagagagagg caatcaatgg ctcttcaca aaccagagac 3540

 caaatgtcac gttttgttt gtgccaacct attttgaagt accacaaaa aagctgtatt 3600
 ttgaaaatgc tttagaaagg ttttgagcat gggttcatcc tattctttcg aaagaagaaa 3660
 atatcataaa aatgagtgat aaatacaagg ccagatgtg gttgcataag gtttttatgc 3720
 atgtttgttg tatacttctc tatgttctt ttaaattgtg tgtgctctgc ttcaatgtag 3780
 tcagaattag ctgcttctat gtttcatagt tggggtcata gatgtttcct tgccttgttg 3840
 atgtggacat gagccatttg aggggagagg gaacggaaat aaaggagtta ttgtaatga 3900
 aaaaaaaaa aaaaaaaaa aaaaaa 3926

<210> 6

<211> 1059

<212> PRT

<213> Homo sapiens

<400> 6

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val

35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

50

55

60

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

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

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

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

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

Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly

565 570 575

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

580 585 590

Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu

595 600 605

Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val

610 615 620

Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly

625 630 635 640

Gly Asp Leu Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln

645 650 655

Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile

660 665 670

Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp

675 680 685

Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val

690 695 700

Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser

705 710 715 720

Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro

725 730 735

Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp

740 745 750

Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro

755 760 765

Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly

770 775 780

Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile

785 790 795 800

Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala

805	810	815	
Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile			
820	825	830	
Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu			
835	840	845	
Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu			
850	855	860	
Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro			
865	870	875	880
Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr			
885	890	895	
Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly			
900	905	910	
Gly His Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu			
915	920	925	
His Arg Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro			
930	935	940	
Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro			
945	950	955	960
Ile Ala Lys Lys Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly			
965	970	975	
Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala			
980	985	990	
Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val			
995	1000	1005	
Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr			
1010	1015	1020	
Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro			
1025	1030	1035	
Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser			
1040	1045	1050	

Met Asn Gln Pro Gly Pro

1055

<210> 7

<211> 4679

<212> DNA

<213> Homo sapiens

<400> 7

ggcggcgcgcg cgcggcgctc gcggtgctg cctgggaggg aggccgggca ggcggctgag 60

cggcgcggtt ctcaacgtga cggggaagtg gttcgggagg cgcggctta ctacccagg 120

gcgaacggac ggacgacgga ggcgggagcc ggtagccgag ccgggagacc tagagaacga 180

gcgggtcagg ctacgctcg gccactctgt cggccgctg aatgaagtgc ccgccctct 240

gagcccgag cccggcgctt tcccgcgaag atggacggtt tcgccggcag tctcgatgat 300

agtatttctg ctgcaagtac ttctgatgtt caagatcgcc tgtcagctct tgagtcacga 360

gttcagcaac aagaagatga aatcactgtg ctaaaggcgg ctttggtga tgttttgagg 420

cgtcttgcaa tctctgaaga tcatgtggcc tcagtgaata aatcagtctc aagtaaaggc 480

caaccaagcc ctcgagcagt tattcccatg tctgtataa ccaatggaag tggcgcaaac 540

agaaaaccaa gtcataccag tgctgtctca attgcaggaa aagaaactct ttcactctgt 600

gctaaaagtg gtacagaaaa aaagaagaa aaaccacaag gacagagaga aaaaaagag 660

gaatctcatt ctaatgatca aagtccacaa attcgagcat caccttctcc ccagccctct 720

tcacaacctc tccaaatata cagacaaact ccagaaagca agaattgtac tcccacaaa 780

agcataaac gaccatcacc agctgaaaag tcacataatt cttgggaaaa ttcagatgat 840

agccgtaata aattgtcgaa aataccttca acacccaaat taatacaaaa agttacaaa 900

actgcagaca agcataaaga tgtcatcatc aaccaagaag gagaatatat taaaatgttt 960

atgcgcggtc ggccaattac catgttcatt ccttcgatg ttgacaacta tgatgacatc 1020

agaacggaac tgcctcctga gaagctcaaa ctggagtggg catatggtta tcgaggaaag 1080

gactgtagag ctaatgttta ccttcttccg accggggaaa tagtttattt cattgcatca 1140

gtagtagtac tatttaatta tgaggagaga actcagcgac actacctggg ccatacagac 1200

tgtgtgaaat gccttgctat acatcctgac aaaattagga ttgcaactgg acagatagct 1260

ggcgtggata aagatggaag gcctctacaa cccacgtca gagtgtggga ttctgttact 1320

ctatccacac tgcagattat tggacttggc acttttgagc gtggagtagg atgcctggat 1380

ttttcaaaag cagattcagg tgttcattta tgtgttattg atgactcaa tgagcatatg 1440

cttactgtat gggactggca gaagaaagca aaaggagcag aaataaagac aacaaatgaa	1500
gttgttttgg ctgtggagtt tcaccaaca gatgcaaata ccataattac atgcggtaaa	1560
tctcatatth tcttctggac ctggagcggc aattcactaa caagaaaaca gggaatthtt	1620
gggaaatatg aaaagccaaa atthgtgcag tgtttagcat tcttggggaa tggagatgtt	1680
cttactggag actcaggtgg agtcatgctt atatggagca aaactactgt agagcccaca	1740
cctgggaaag gacctaaagg tgtatatcaa atcagcaaac aaatcaaagc tcatgatggc	1800
agtgtgttca cactthtga gatgagaaat gggatgttat taactggagg agggaaagac	1860
agaaaaataa tctgtggga tcatgatctg aatcctgaaa gagaaataga ggttctctgat	1920
cagtatggca caatcagagc tgtagcagaa ggaaaggcag atcaatthtt agtaggcaca	1980
tcacgaaact ttatthtacg aggaacatth aatgatggct tccaaataga agtacagggt	2040
catacagatg agctthgggg tcttgccaca catcccttca aagatthtct cttagacatg	2100
gctcaggaca ggcaggtgtg cctgtggaac tcaatggaac acaggctgga atggaccagg	2160
ctggtagatg aaccaggaca ctgtgcagat thtcatccaa gtggcacagt ggtggccata	2220
ggaacgcact caggcaggtg gthtgttctg gatgcagaaa ccagagatct agthtctatc	2280
cacacagacg ggaatgaaca gctctctgtg atgcgctact caatagatgg taccttctctg	2340
gctgtaggat ctcatgacaa cthtatttac ctctatgtag tctctgaaaa tggagaaaa	2400
tatagcagat atggaagggt cactggacat tccagctaca tcacacacct tgactggtec	2460
ccagacaaca agtatataat gtctaactcg ggagactatg aaatattgta cttgtaccgc	2520
cggagcacc aggagctgca agccatgcag atggagctgc agagccctga gtacaagctg	2580
agcaagctcc gcacctgac catcatgacc gactacaacc ccaactactg cthtgtctggc	2640
aagacctct ccatcagta cctgaaggag gtgccgcgga aaaacatcac cctattcgg	2700
ggtctgggcc atggagctth tggggagggt tatgaaggcc aggtgtccgg aatgcccaac	2760
gaccaagcc cctgcaagt ggctgtgaag acgtgcctg aagtgtgctc tgaacaggac	2820
gaactggatt tctcatgga agccctgatc atcagcaaata tcaaccacca gaacattgtt	2880
cgctgcattg ggtgagcct gcaatccctg ccccggttca tctgtctgga gctcatggcg	2940
gggggagacc tcaagctctt cctccgagag acccgccctc gcccgagcca gcctctctc	3000
ctggccatgc tggaccttct gcacgtggct cgggacattg cctgtggctg tcagtatthg	3060
gaggaaaacc acttcatcca ccgagacatt gctgccagaa actgcctctt gacctgtcca	3120
ggccctggaa gagtggccaa gattggagac ttcgggatgg cccgagacat ctacagggcg	3180
agctactata gaaagggagg ctgtgccatg ctgccagtta agtggatgcc cccagaggcc	3240
ttcatggaag gaatattcac ttctaaaaca gacacatggc cthtggagt gctgctatgg	3300

gaaatctttt ctcttggata tatgccatac cccagcaaaa gcaaccagga agttctggag 3360
 ttgtcacca gtggaggccg gatggacca cccaagaact gccctgggcc tgtataccgg 3420

ataatgactc agtgctggca acatcagcct gaagacaggc ccaactttgc catcattttg 3480
 gagaggattg aatactgcac ccaggaccgc gatgtaatca acaccgcttt gccgatagaa 3540
 tatggccac ttgtggaaga ggaagagaaa gtgcctgtga ggcccaagga ccctgagggg 3600
 gtctctctc tcttggctc tcaacaggca aaacgggagg aggagcgcag cccagctgcc 3660
 ccaccacctc tgctaccac ctctcttggc aaggctgcaa agaaaccac agctgcagag 3720
 gtctctgttc gattccctag agggccggcc gtggaagggg gacacgtgaa tatggcattc 3780
 tctcagtcca accctccttc ggagttgcac aggggtccacg gatccagaaa caagcccacc 3840

agcttgtgga acccaacgta cggctcctgg ttacagaga aaccaccaa aaagaataat 3900
 cctatagcaa agaaggagcc acacgagagg ggtaacctgg ggctggaggg aagctgtact 3960
 gtccaccta acgttgcaac tgggagactt cgggggcct cactgctcct agagccctct 4020
 tcgtgactg ccaatatgaa ggaggtacct ctgttcaggc tacgtcactt cccttgtggg 4080
 aatgtcaatt acggctacca gcaacagggc ttgcccttag aagccgtac tgcccttga 4140
 gctggtcatt acgaggatc cattctgaaa agcaagaata gcatgaacca gcctgggccc 4200
 tgagctcggc cacacactca cttctcttc ttgggatccc taagaccgtg gaggagagag 4260

aggcaatcaa tggctcctc acaaaccaga gaccaaagt cacgtttgt tttgtgcaa 4320
 cctattttga agtaccacca aaaaagctgt atttgaaaa tgctttagaa aggttttgag 4380
 catgggttca tctattctt tcgaaagaag aaaatatcat aaaaatgagt gataaataca 4440
 aggcccagat gtggttgcat aaggttttta tgcattgttg ttgtatactt ccttatgctt 4500
 cttttaaat gtgtgtgctc tgcttcaatg tagtcagaat tagctgcttc tatgtttcat 4560
 agttggggtc atagatgttt ctttgccttg ttgatgtgga catgagccat ttgaggggag 4620
 agggaacgga aataaaggag ttatttghta tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4679

<210> 8

<211> 1310

<212> PRT

<213> Homo sapiens

<400> 8

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30
 Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val
 35 40 45
 Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

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

 115 120 125
 Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
 130 135 140
 Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
 145 150 155 160
 Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
 165 170 175
 Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg

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

 245 250 255
 Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
 260 265 270

Arg Ala Asn Val Tyr Leu Leu Pro Thr Gly Glu Ile Val Tyr Phe Ile
 275 280 285
 Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
 290 295 300
 Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp

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

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

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

 500 505 510
 Phe Thr Leu Cys Gln Met Arg Asn Gly Met Leu Leu Thr Gly Gly Gly

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

 565 570 575
 Arg Gly Thr Phe Asn Asp Gly Phe Gln Ile Glu Val Gln Gly His Thr
 580 585 590
 Asp Glu Leu Trp Gly Leu Ala Thr His Pro Phe Lys Asp Leu Leu Leu
 595 600 605
 Thr Cys Ala Gln Asp Arg Gln Val Cys Leu Trp Asn Ser Met Glu His
 610 615 620
 Arg Leu Glu Trp Thr Arg Leu Val Asp Glu Pro Gly His Cys Ala Asp

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

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

 755 760 765

Lys Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro
 770 775 780
 Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu
 785 790 795 800
 Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala
 805 810 815
 Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro
 820 825 830
 Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu
 835 840 845
 Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe
 850 855 860
 Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu
 865 870 875 880
 Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser
 885 890 895
 Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala
 900 905 910
 Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln
 915 920 925
 Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn
 930 935 940
 Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp
 945 950 955 960
 Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly
 965 970 975
 Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met
 980 985 990
 Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu
 995 1000 1005
 Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys

1010	1015	1020
Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met		
1025	1030	1035
Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr		
1040	1045	1050
Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile		
1055	1060	1065
Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile		

1070	1075	1080
Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu		
1085	1090	1095
Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro		
1100	1105	1110
Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro		
1115	1120	1125
Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala		

1130	1135	1140
Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly		
1145	1150	1155
Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser		
1160	1165	1170
Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys		
1175	1180	1185
Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu		

1190	1195	1200
Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His		
1205	1210	1215
Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro		
1220	1225	1230
Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu		
1235	1240	1245

Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg

1250 1255 1260
 Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln
 1265 1270 1275
 Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His
 1280 1285 1290
 Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro
 1295 1300 1305
 Gly Pro

1310

<210> 9

<211> 2473

<212> DNA

<213> Homo sapiens

<400> 9

actctgtcgg tccgtgaat gaagtgcccg cccctctaag cccggagccc ggcgctttcc	60
ccgcaagatg gacggtttcg ccggcagttc cgatgatagt atttctgctg caagtacttc	120
tgatgttcaa gatcgctgt cagctcttga gtcacgagtt cagcaacaag aagatgaaat	180
caactgtgcta aaggcggctt tggctgatgt tttaggcgt cttgcaatct ctgaagatca	240
tgtggcctca gtgaaaaaat cagttcctca taaaggccaa ccaagccctc gagcagttat	300
tccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc	360
tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa	420
gaaagaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag	480
tccacaaatt cgagcatcac cttctccca gccctcttca caacctctcc aaatacacag	540
acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc	600
tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat	660
accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt	720
catcatcaac caagtgtacc gccggaagca ccaggagctg caagccatgc agatggagct	780
gcagagccct gagtacaagc tgagcaagct ccgcacctcg accatcatga ccgactacaa	840
ccccaaactac tgccttctgt gcaagacctc ctccatcagt gacctgaagg aggtgccgcg	900
gaaaaacatc accctcattc ggggtctggg ccatggagcc tttggggagg tgtatgaagg	960

ccagggtgtcc ggaatgccca acgacccaag cccctgcaa gtggctgtga agacgtgcc 1020
tgaagtgtgc tctgaacagg acgaactgga tttcctcatg gaagccctga tcatcagcaa 1080
attcaaccac cagaacattg ttcgtctcat tggggtgagc ctgcaatccc tgccccggtt 1140
catcctgtctg gagctcatgg cgggggggaga cctcaagtcc ttcctccgag agaccgccc 1200
tcgcccgagc cagccctcct ccctggccat gctggacctt ctgcacgtgg ctcgggacat 1260

tgctgtggc tgtcagtatt tggaggaaaa ccacttcac caccgagaca ttgctgccag 1320
aaactgcctc ttgacctgtc caggccctgg aagagtggcc aagattggag acttcgggat 1380
ggcccagac atctacagg cgagctacta tagaaaggga ggctgtgcca tgctgccagt 1440
taagtggatg cccccagagg ctttcatgga aggaatattc acttctaaaa cagacacatg 1500
gtcctttgga gtgctgctat gggaaatctt tttctttgga tataatgcat acccagcaa 1560
aagcaaccag gaagtcttgg agtttctcac cagtggaggc cggatggacc cacccaagaa 1620
ctgccctggg cctgtatacc ggataatgac tcagtgtctg caacatcagc ctgaagacag 1680

gcccacttt gccatcattt tggagaggat tgaatactgc acccaggacc cggatgtaat 1740
caacaccgct ttgccgatag aatatgggtcc acttgtgga gaggaagaga aagtgcctgt 1800
gaggcccaag gacctgagg gggttcctcc tctcctggc tctcaacagg caaacggga 1860
ggaggagcgc agcccagctg cccaccacc tctgctacc acctcctctg gcaaggctgc 1920
aaagaaacc acagctgcag aggtctctgt tcgagtcct agagggccg ccgtggaagg 1980
gggacacgtg aatatggcat tctctcagtc caacctcct tcggagtgtc acagggtcca 2040
cggatccaga aacaagccca ccagcttgtg gaaccaacg tacggctcct ggtttacaga 2100

gaaaccacc aaaaagaata atcctatagc aaagaaggag ccacacgaga ggggtaacct 2160
ggggctggag ggaagctgta ctgtcccacc taacgttgca actgggagac ttccgggggc 2220
ctcactgtc ctagagccct cttcgtgac tgccaatatg aaggaggtag ctctgttcag 2280
gtacgtcac ttcccttgtg ggaatgtcaa ttacggctac cagcaacagg gcttgccctt 2340
agaagccgt actgccctg gagctggta ttacaggat accattctga aaagcaagaa 2400
tagcatgaac cagcctgggc cctgagctcg gtgcacact cacttctctt ccttgggac 2460
cctaagaccg tgg 2473

<210> 10

<211> 785

<212> PRT

<213> Homo sapiens

<400> 10

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

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

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

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

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

Ser Lys Ser Asn Gln Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg

500 505 510

Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr

515 520 525

Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile

530 535 540

Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr

545 550 555 560

Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val

565 570 575

Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser

580 585 590

Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro

595 600 605

Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala

610 615 620

Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His

625 630 635 640

Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg

645 650 655

Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr

660 665 670

Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala

675 680 685

Lys Lys Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys

690 695 700

Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu

705 710 715 720

Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu

725 730 735

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

740	745	750	
Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His			
755	760	765	
Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly			
770	775	780	
Pro			
785			
<210> 11			
<211> 2506			
<212> DNA			
<213> Homo sapiens			
<400> 11			
actctgtcgg tccgctgaat gaagtgcccg ccctcttaag cccggagccc ggcgctttcc			60
ccgcaagatg gacggtttcg cggcagctct cgatgatagt atttctgctg caagtacttc			120
tgatgttcaa gatcgctgtg cagctcttga gtcacgagtt cagcaacaag aagatgaaat			180
cactgtgcta aaggcggctt tggctgatgt ttgaggcgt cttgcaatct ctgaagatca			240
tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat			300
tcccatgtcc tgtataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc			360
tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa			420
gaaagaaaaa ccacaaggac agagagaaaa aaaagaggaa tctcattcta atgatcaaag			480
tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag			540
acaaactcca gaaagcaaga atgctactcc caccaaaaagc ataaaacgac catcaccagc			600
tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat			660
accttcaaca cccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt			720
catcatcaac caagcaaaaa tgtcaactcg cgaaaaaaac agccaagtgt accgccggaa			780
gcaccaggag ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa			840
gctccgacc tcgacctca tgaccgacta caacccaac tactgctttg ctggcaagac			900
ctctccatc agtgacctga aggaggtgcc gcggaaaaac atcacctca ttcggggtct			960
gggccatgga gcctttgggg aggtgtatga aggccaggtg tccggaatgc ccaacgaccc			1020
aagccccctg caagtggctg tgaagacgct gcctgaagtg tgctctgaac aggacgaact			1080
ggatttcctc atggaagccc tgatcatcag caaattcaac caccagaaca ttgttcgctg			1140

cattgggggtg agcctgcaat ccctgccccg gttcatcctg ctggagctca tggcgggggg 1200
 agacctcaag tccttctccc gagagaccgc ccctcgcccc agccagccct cctccctggc 1260
 catgctggac ctctgcacg tggctcgga cattgcctgt ggctgtcagt atttgagga 1320
 aaaccacttc atccaccgag acattgctgc cagaaactgc ctcttgacct gtccaggccc 1380
 tggaagagtg gccaagattg gagacttcgg gatggcccga gacatctaca gggcgagcta 1440
 ctatagaaag ggaggctgtg ccatgctgcc agttaagtgg atgccccag aggccttcat 1500

ggaaggaata ttcacttcta aaacagacac atggtccttt ggagtgctgc tatgggaaat 1560
 cttttctctt ggatatatgc cataccccag caaaagcaac caggaagtgc tggagtttgt 1620
 caccagtggg ggccggatgg acccacccaa gaactgcctt gggcctgtat accggataat 1680
 gactcagtgc tggcaacatc agcctgaaga caggcccaac tttgcatca ttttgagag 1740
 gattgaatac tgcaccagg acccgatgt aatcaacacc gctttgccga tagaatatgg 1800
 tccacttgtg gaagaggaag agaaagtgc tgtgaggccc aaggaccctg aggggggttcc 1860
 tcctctctg gtctctcaac aggcacaaac ggaggaggag cgcagcccag ctgccccacc 1920

acctctgcct accacctcct ctggcaaggc tgcaaagaaa cccacagctg cagaggtctc 1980
 tgttcgagtc cctagagggc cggccgtgga agggggacac gtgaatatgg cattctctca 2040
 gtccaacctt ccttcggagt tgcacagggt ccacggatcc agaaacaagc ccaccagctt 2100
 gtggaaccca acgtacggct cctggtttac agagaaacce accaaaaaga ataatcctat 2160
 agcaaagaag gagccacacg agaggggtaa cctggggctg gagggaagct gtactgtccc 2220
 acctaacgtt gcaactggga gacttcggg ggcctcactg ctcttagagc cctcttcgct 2280
 gactgccaat atgaaggagg tacctctgtt caggctacgt cacttccctt gtgggaatgt 2340

caattacggc taccagcaac agggcttgcc cttagaagcc gctactgccc ctggagctgg 2400
 tcattacgag gataccattc tgaaaagcaa gaatagcatg aaccagcctg ggccctgagc 2460
 tcggtcgac actcacttct cttccttggg atccctaaga ccgtgg 2506

<210> 12

<211> 796

<212> PRT

<213> Homo sapiens

<400> 12

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

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

Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro
 275 280 285

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

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

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

 Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp
 485 490 495
 Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln
 500 505 510
 Glu Val Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys

515	520	525	
Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His			
530	535	540	
Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu			
545	550	555	560
Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu			
565	570	575	
Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys			
580	585	590	
Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg			
595	600	605	
Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser			
610	615	620	
Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg			
625	630	635	640
Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe			
645	650	655	
Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg			
660	665	670	
Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr			
675	680	685	
Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His			
690	695	700	
Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn			
705	710	715	720
Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser			
725	730	735	
Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His			
740	745	750	
Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro			
755	760	765	

Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr Ile

770

775

780

Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro

785

790

795

<210> 13

<211

> 3409

<212> DNA

<213> Homo sapiens

<400> 13

actctgtcgg tccgctgaat gaagtgcccg ccctcttaag cccggagccc ggcgctttcc	60
ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc	120
tgatgttcaa gatcgctgtg cagctcttga gtcacgagtt cagcaacaag aagatgaaat	180
cactgtgcta aaggcggctt tggctgatgt tttaggcgt cttgcaatct ctgaagatca	240
tgtggcctca gtgaaaaaat cagtctcaag taaaggccaa ccaagccctc gagcagttat	300
tcccatgtcc tgiataacca atggaagtgg tgcaaacaga aaaccaagtc ataccagtgc	360
tgtctcaatt gcaggaaaag aaactctttc atctgctgct aaaagtggta cagaaaaaaa	420
gaaagaaaaa ccacaaggag agagagaaaa aaaagaggaa tctcattcta atgatcaaag	480
tccacaaatt cgagcatcac cttctcccca gccctcttca caacctctcc aaatacacag	540
acaaactcca gaaagcaaga atgctactcc caccaaaagc ataaaacgac catcaccagc	600
tgaaaagtca cataattctt gggaaaattc agatgatagc cgtaataaat tgtcgaaaat	660
accttcaaca ccaaattaa taccaaaagt taccaaaact gcagacaagc ataaagatgt	720
catcatcaac caagaaggag aatatattaa aatgtttatg cgcggtcggc caattacat	780
gttcattcct tccgatgttg acaactatga tgacatcaga acggaactgc ctctgagaa	840
gctcaaaactg gagtgggcat atggttatcg aggaaaggac tgtagagcta atgtttacct	900
tcttccgacc ggggaaatag ttattttcat tgcattcagta gtagtactat ttaattatga	960
ggagagaact cagcgacact acctgggcca tacagactgt gtgaaatgcc ttgctataca	1020
tcctgacaaa attaggattg caactggaca gatagctggc gtggataaag atggaaggcc	1080
tctacaacce cagtcagag tgtgggattc tgttactcta tccacactgc agattattgg	1140
acttggcact tttagcgtg gagtaggatg cctggatttt tcaaaagcag attcaggtgt	1200
tcatttatgt gttattgatg actccaatga gcatatgctt actgtatggg actggcagag	1260
gaaagcaaaa ggagcagaaa taaagacaac aatgaagtt gttttggctg tggagtttca	1320

cccaacagat gcaaatacca taattacatg cggtaaactt catatcttct tctggacctg 1380
 gagcggaat tctaataca gaaacaggg aatttttggg aaatatgaa agccaaaatt 1440
 tgtgcagtgt ttagcattct tggggaatgg agatgttctt actggagact caggtggagt 1500
 catgcttata tggagcaaaa ctactgtaga gccacacct gggaaaggac ctaaagggtg 1560
 atatcaaatc agcaacaaa tcaaagctca tgatggcagt gtgttcacac tttgtcagat 1620

 gagaaatggg atgttattaa ctggaggagg gaaagacaga aaaataattc tgtgggatca 1680
 tgatctgaat ccigaaagag aaatagagat atgctggatg agccctgagt acaagctgag 1740
 caagctccgc acctcgacca tcatgaccga ctacaacccc aactactgct ttgctggcaa 1800
 gacctcctcc atcagtgacc tgaaggaggt gccgcggaaa aacatcaccc tcattcgggg 1860
 tctgggcat ggagcctttg gggagggtga tgaaggccag gtgtccgaa tgcccaacga 1920
 cccaagcccc ctgcaagtgg ctgtgaagac gctgcctgaa gtgtgctctg aacaggacga 1980
 actggatttc ctcatggaag ccctgatcat cagcaaattc aaccaccaga acattgttcg 2040

 ctgcattggg gtgagcctgc aatccctgcc cgggttcac ctgctggagc tcatggcggg 2100
 gggagacctc aagtccttcc tccgagagac ccgcccctgc ccgagccagc cctcctcct 2160
 ggccatgctg gaccttctgc acgtggctcg ggacattgcc tgtggctgtc agtatttga 2220
 ggaaaaccac ttcaccacc gagacattgc tgccagaaac tgcctcttga cctgtccagg 2280
 ccctggaaga gtggccaaga ttgggagactt cgggatggcc cgagacatct acaggcgag 2340
 ctactataga aaggagggt gtgccatgct gccagttaag tggatgcccc cagaggcctt 2400
 catggaagga atattcatt ctaaaacaga cacatggtcc tttggagtgc tgctatggga 2460

 aatcttttct ctiggaata tgccataccc cagcaaaagc aaccaagaag ttctggagtt 2520
 tgtcaccagt ggaggccga tggaccacc caagaactgc cctgggcctg tataccggat 2580
 aatgactcag tctggcaac atcagcctga agacaggccc aactttgcca tcattttga 2640
 gaggattgaa tactgcacc aggaccgga tgtaatcaac accgctttgc cgatagaata 2700
 tggctcactt gtggaagagg aagagaaagt gcctgtgagg ccaaggacc ctgagggggt 2760
 tctctctctc ctggtctctc aacaggcaaa acgggaggag gagcgagcc cagctgcccc 2820
 accacctctg cctaccact cctctggcaa ggctgcaaag aaaccacag ctgcagaggt 2880

 ctctgttga gtccttagag ggccggcgt ggaaggggga cacgtgaata tggcattctc 2940
 tcagtccaac cctccttcgg agttgcacag ggtccacga tccagaaaca agcccaccag 3000
 cttgtggaac ccaacgtacg gctcctggtt tacagagaaa cccacaaaa agaataatcc 3060
 tatagcaag aaggagccac acgagagggg taacctgggg ctggagggaa gctgtactgt 3120
 cccacctaac gttgcaactg ggagacttcc gggggcctca ctgctcctag agccctcttc 3180

gtgactgcc aatatgaagg aggtacctct gttcaggcta cgtcacttcc ctgtgggaa 3240

tgtcaattac ggctaccagc aacagggcct gcccttagaa gccgtactg cccctggagc 3300

tggtcattac gaggatacca ttctgaaaag caagaatagc atgaaccagc ctgggcctg 3360

agctcggtcg cacactcact tctcttcctt gggatcccta agaccgtgg 3409

<210> 14

<211> 1097

<212> PRT

<213> Homo sapiens

<400> 14

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val

35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met

65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr

85 90 95

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

100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys

115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser

130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr

145 150 155 160

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

165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile

225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

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

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

290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly
325 330 335

Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
340 345 350

Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys

355 360 365

Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp
370 375 380

Asp Ser Asn Glu His Met Leu Thr Val Trp Asp Trp Gln Arg Lys Ala
385 390 395 400

Lys Gly Ala Glu Ile Lys Thr Thr Asn Glu Val Val Leu Ala Val Glu
405 410 415

Phe His Pro Thr Asp Ala Asn Thr Ile Ile Thr Cys Gly Lys Ser His

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

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

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

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

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

675 680 685

Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu Leu

690 695 700

His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn

705 710 715 720

His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys

725 730 735

Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg

740 745 750

Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu

755 760 765

Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr

770 775 780

Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe

785 790 795 800

Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu

805 810 815

Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro

820 825 830

Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro Glu

835 840 845

Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr

850 855 860

Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro

865 870 875 880

Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu

885 890 895

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

900 905 910

Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys

915 920 925
Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg

930 935 940
Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser

945 950 955 960
Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys Pro

965 970 975
Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro

980 985 990
Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg Gly

995 1000 1005
Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala

1010 1015 1020
Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser

1025 1030 1035
Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His

1040 1045 1050
Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu

1055 1060 1065
Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp

1070 1075 1080
Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro

1085 1090 1095

<210> 15

<211> 2014

<212> DNA

<213> Homo sapiens

<400> 15

actctgtcgg tccgctgaat gaagtgcccg ccctcttaag cccggagccc ggcgctttcc 60

ccgcaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc 120

tgatgttcaa gatcgctgtg cagctcttga gtcacgagtt cagcaacaag aagatgaaat 180

cactgtgcta aaggcggtt tggctgatgt tttaggcgt cttgcaatct ctgaagatca	240
tgtggcctca gtgaaaaat cagtctcaag taaagtgtac cgccggaagc accaggagct	300
gcaagccatg cagatggagc tgcagagccc tgagtacaag ctgagcaagc tccgcacctc	360
gaccatcatg accgactaca accccaacta ctgctttgct ggcaagacct cctccatcag	420
tgacctgaag gaggtgccgc ggaaaaacat caccctcatt cgggggtctgg gccatggagc	480
ctttggggag gtgtatgaag gccagggtgc cggaatgccc aacgacccaa gccccctgca	540
agtggctgtg aagacgtgc ctgaagtgtg ctctgaacag gacgaactgg atttcctcat	600
ggaagccctg atcatcagca aattcaacca ccagaacatt gttcgctgca ttggggtgag	660
cctgcaatcc ctgccccgtt tcatcctgct ggagctcatg gcggggggag acctcaagtc	720
cttcctccga gagaccgcc ctgccccgag ccagccctcc tccctggcca tgctggacct	780
tctgcacgtg gctcgggaca ttgctgtggt ctgtcagtat ttggaggaaa accacttcat	840
ccaccgagac attgctgcca gaaactgcct cttgacctgt ccaggccctg gaagagtggc	900
caagattgga gacttcggga tggcccgaga catctacagg gcgagctact atagaaaggg	960
aggctgtgcc atgtgccag ttaagtggat gccccagag gccttcatgg aaggaatatt	1020
cacttctaaa acagacacat ggtccttttg agtgcctgta tgggaaatct tttctcttg	1080
atatatgcca taccacgca aaagcaacca ggaagtcttg gatttgtca ccagtggagg	1140
ccgatggac ccaccaaga actgccctgg gctgtatac cggataatga ctcagtctg	1200
gcaacatcag ctgaagaca ggccaactt tgccatcatt ttggagagga ttgaatactg	1260
caccaggac ccgatgtaa tcaacaccgc ttgcccata gaatatggtc cacttgtgga	1320
agaggaagag aaagtgcctg tgaggcccaa ggaccctgag ggggttcctc ctctcctggt	1380
ctctcaacag gcaaacggg aggaggagcg cagcccagct gcccaccac ctctgcctac	1440
cacctcctct ggcaagctg caaagaaacc cacagctgca gaggtctctg ttcgagtccc	1500
tagagggccg gccgtggaag ggggacacgt gaatatggca ttctctcagt ccaacctcc	1560
ttcggagtgt cacagggtcc acggatccag aaacaagccc accagcttgt ggaaccaac	1620
gtacggctcc tggtttacag agaaaccac caaaaagaat aatcctatag caaagaagga	1680
gccacacgag aggggtaacc tggggtgga gggaagctgt actgtccac ctaacgttgc	1740
aactgggaga cttccggggg cctcactgct cctagagccc tcttcgtga ctgccaatat	1800
gaaggaggta ccctgttca ggctacgtca cttcccttgt gggaatgtca attacggcta	1860
ccagcaacag ggcttgcct tagaagccgc tactgcccct ggagctggtc attacgagga	1920
taccattctg aaaagcaaga atagcatgaa ccagcctggg cctgagctc ggtcgcacac	1980
tcacttctct tccttgggat ccctaagacc gtgg	2014

<210> 16

<211> 632

<212> PRT

<213> Homo sapiens

<400> 16

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val

35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

50 55 60

Ser Val Ser Ser Lys Val Tyr Arg Arg Lys His Gln Glu Leu Gln Ala

65 70 75 80

Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg

85 90 95

Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys Phe Ala Gly

100 105 110

Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg Lys Asn Ile

115 120 125

Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu Val Tyr Glu

130 135 140

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

145 150 155 160

Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe

165 170 175

Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val

180 185 190

Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu

195 200 205

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

210

215

220

Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu Leu His

225

230

235

240

Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu Asn His

245

250

255

Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr Cys Pro

260

265

270

Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala Arg Asp

275

280

285

Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met Leu Pro

290

295

300

Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe Thr Ser

305

310

315

320

Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser

325

330

335

Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val Leu Glu

340

345

350

Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys Pro Gly

355

360

365

Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro Glu Asp

370

375

380

Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys Thr Gln

385

390

395

400

Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu

405

410

415

Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly

420

425

430

Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg

435

440

445

Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala

450 455 460
Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly

465 470 475 480
Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln Ser Asn

485 490 495
Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys Pro Thr

500 505 510
Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys Pro Thr

515 520 525
Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg Gly Asn

530 535 540
Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala Thr Gly

545 550 555 560
Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala

565 570 575
Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly

580 585 590
Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala

595 600 605
Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys

610 615 620
Asn Ser Met Asn Gln Pro Gly Pro

625 630

<210> 17

<211> 2131

<212> DNA

<213> Homo sapiens

<400> 17

actctgtcgg tccgtgaat gaagtgcccg ccctctaag cccggagccc ggcgctttcc	60
ccgaagatg gacggtttcg ccggcagtct cgatgatagt atttctgctg caagtacttc	120
tgatgttcaa gatcgctgtg cagctcttga gtcacgagtt cagcaacaag aagatgaaat	180

cactgtgcta aaggcggtt tggctgatgt tttaggcgt cttgcaatct ctgaagatca 240
 tgtggcctca gtgaaaaat cagtctcaag taaaggttca gagctcaggg gaggatatgg 300
 agatccaggg aggcttcctg taggaagtgg cctgtgtagt gcttcaaggg ccaggctgcc 360
 aggccatgtt gcagctgacc acccacctgc agtgtaccgc cggaagcacc aggagctgca 420
 agccatgcag atggagctgc agagccctga gtacaagctg agcaagctcc gcacctcgac 480
 catcatgacc gactacaacc ccaactactg ctttgctggc aagacctcct ccatcagtga 540
 cctgaaggag gtgccgcgga aaaacatcac cctcattcgg ggtctgggcc atggagcctt 600

 tggggaggtg tatgaaggcc aggtgtccgg aatgcccaac gaccaagcc cctgcaagt 660
 ggctgtgaag acgtgcctg aagtgtgctc tgaacaggac gaactggatt tcctcatgga 720
 agccctgac atcagcaat tcaaccacca gaacattgtt cgctgcattg gggtagcct 780
 gcaatccctg ccccggttca tctgctgga gctcatggcg gggggagacc tcaagtcctt 840
 cctccgagag acccgccctc gcccagacca gccctcctcc ctggccatgc tggaccttct 900
 gcacgtggct cgggacattg cctgtggctg tcagtatttg gaggaaaacc acttcatcca 960
 ccgagacatt gctgccagaa actgcctctt gacctgtcca ggccctggaa gagtggccaa 1020

 gattggagac ttcgggatgg cccgagacat ctacagggcg agctactata gaaaggagg 1080
 ctgtgccatg ctgccagta agtggatgcc cccagaggcc ttcattggaag gaatatcac 1140
 ttctaaaaca gacacatggt cctttggagt gctgctatgg gaaatctttt ctcttgata 1200
 tatgccatac cccagcaaaa gcaaccagga agttctggag tttgtacca gtggaggccg 1260
 gatggacca cccaagaact gccctgggcc tgtataccgg ataatgactc agtgcaggca 1320
 acatcagcct gaagacaggc ccaactttgc catcattttg gagaggattg aatactgcac 1380
 ccaggaccg gatgtaatca acaccgcttt gccgatagaa tatgtccac ttgtggaaga 1440

 ggaagagaaa gtgcctgtga ggccaagga cctgagggg gttcctctc tctgtgtctc 1500
 tcaacaggca aaacgggagg aggagcgag cccagctgcc ccaccacctc tgcctaccac 1560
 ctctctggc aaggctgcaa agaaaccac agctgcagag gtctctgttc gagtccctag 1620
 agggccggcc gtggaagggg gacacgtgaa tatggcattc tctcagtcca accctccttc 1680
 ggagttgcac aggttcacg gatccagaaa caagcccacc agcttgtgga acccaacgta 1740
 cggctcctgg ttacagaga aaccaccaa aaagaataat cctatagcaa agaaggagcc 1800
 acacgagagg ggtaacctgg ggctggaggg aagctgtact gtccacctc acgttgcaac 1860

 tgggagactt ccgggggcct cactgtcct agagccctct tcgtgactg ccaatatgaa 1920
 ggaggtacct ctgttcaggc tacgtcactt ccttgtggg aatgtcaatt acggctacca 1980
 gcaacagggc ttgcccttag aagccgtac tgcccctgga gctggtcatt acgaggatac 2040

cattctgaaa agcaagaata gcatgaacca gcctgggccc tgagctcggt cgcacactca 2100
 cttctcttcc ttgggatccc taagaccgtg g 2131
 <210> 18
 <211> 671
 <212> PRT
 <213> Homo sapiens
 <400> 18
 Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

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

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

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

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

195 200 205
Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys

210 215 220
Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser
225 230 235 240

Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys
245 250 255
Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu

260 265 270
Ala Met Leu Asp Leu Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys
275 280 285

Gln Tyr Leu Glu Glu Asn His Phe Ile His Arg Asp Ile Ala Ala Arg
290 295 300

Asn Cys Leu Leu Thr Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly
305 310 315 320
Asp Phe Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys

325 330 335
Gly Gly Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe
340 345 350

Met Glu Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val
355 360 365

Leu Leu Trp Glu Ile Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys
370 375 380

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

385 390 395 400
Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys
405 410 415

Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu
420 425 430

Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu

435 440 445
Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val

450 455 460
Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln

465 470 475 480
Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro

485 490 495
Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val

500 505 510
Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn

515 520 525
Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Arg Val His

530 535 540
Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser

545 550 555 560
Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys

565 570 575
Glu Pro His Glu Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val

580 585 590
Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu

595 600 605
Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg

610 615 620
Leu Arg His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln

625 630 635 640
Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu

645 650 655
Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro

660 665 670

<210> 19

<211> 3365

<212> DNA

<213> Homo sapiens

<400> 19

tactctgtcg gtccgctgaa tgaagtgcc gccctctaa gcccgagcc cggcgctttc	60
cccgcaagat ggacggtttc gccggcagtc tcgatgatag ttttctgtc gcaagtactt	120
ctgatgttca agatcgctg tcagctcttg agtcacgagt tcagcaaca gaagatgaaa	180
tcactgtgct aaaggcggtt ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc	240
atgtggcctc agtgaaaaa tcagctctca gtaaaggcca accaagcct cgagcagtta	300
ttcccatgtc ctgtataacc aatggaagt gtgcaaacag aaaaccaagt cataccagt	360
ctgtctcaat tgcaggaaaa gaaactctt catctgtgc taaaagtgtt acagaaaaaa	420
agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa	480
gtccacaaat tcgagcatca ctttctccc agccctcttc acaacctctc caaatacaca	540
gacaaactcc agaaagcaag aatgtactc ccacaaaag cataaacga ccatcaccag	600
ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa	660
tacctcaac acccaaatta ataccaaaag ttacaaaac tgcagacaag cataaagatg	720
tcacatcaa ccaagaagga gaatatatta aaatgtttat gcgcggtcgg ccaattacca	780
tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga	840
agctcaaact ggagtgggca tatggttatt gaggaagga ctgtagagct aatgtttacc	900
ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg	960
aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc cttgctatac	1020
atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc	1080
ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg	1140
gacttggcac ttttgagcgt ggagttagat gcctggattt ttcaaaagca gattcaggtg	1200
ttcatttatg tgttattgat gactccaatg agcatatgct tactgtatgg gactggcaga	1260
ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttggt gtggagtttc	1320
acccaacaga tgcaaatacc ataattacat gcggtaaata tcatattttc ttctggacct	1380
ggagcggcaa ttactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat	1440
ttgtgcagtg ttiagcattc ttggggaatg gagatgttct tactggagac tcaggtggag	1500
tcatgcttat atggagcaaa actactgtag agcccacacc tgggaaagga cctaaaggaa	1560
gtggcctgtg tagtgcttca agggccaggc tgccaggcca tgttgcagct gaccaccac	1620
ctgcagtgtg ccgccggaag caccaggagc tgcaagccat gcagatggag ctgcagagcc	1680

ctgagtacaa gctgagcaag ctccgcacct cgaccatcat gaccgactac aaccccaact 1740
 actgctttgc tggcaagacc tctccatca gtgacctgaa ggaggtgccg cggaaaaaca 1800
 tcacctcat tcggggtctg ggccatggag cctttgggga ggtgtatgaa ggccaggtgt 1860
 ccggaatgcc caacgaccca agccccctgc aagtggctgt gaagacgctg cctgaagtgt 1920

gctctgaaca ggacgaactg gatttctca tggaagccct gatcatcagc aaattcaacc 1980
 accagaacat tgttcgctgc attggggtga gcctgcaatc cctgccccgg ttcattcctgc 2040
 tggagctcat ggcgggggga gacctcaagt ccttctccg agagaccgc cctcgccga 2100
 gccagccctc ctccctggcc atgctggacc ttctgcacgt ggctcgggac attgcctgtg 2160
 gctgtcagta ttggaggaa aaccattca tccaccgaga cattgctgcc agaaactgcc 2220
 tcttgacctg tccaggccct ggaagagtgg ccaagattgg agacttcggg atggcccag 2280
 acatctacag ggcgagctac tatagaaagg gaggtgtgc catgctgcca gttaagtga 2340

tgccccaga ggccttcattg gaaggaatat tcacttctaa aacagacaca tggctccttg 2400
 gagtgtgct atgggaaatc ttttctcttg gatatatgcc atacccagc aaaagcaacc 2460
 aggaagtct ggagtttgc accagtggag gccggatgga cccaccaag aactgcctg 2520
 gccctgtata ccgataatg actcagtgt ggcaacatca gcctgaagac aggcccaact 2580
 ttgccatcat ttggagagg attgaatact gcaccagga cccggatgta atcaacaccg 2640
 ctttgccgat agaatatggt ccacttgtgg aagaggaaga gaaagtgcct gtgaggccca 2700
 aggaccctga gggggttct cctctcctgg tctctcaaca ggcaaacgg gaggaggagc 2760

gcagcccagc tgccccacca cctctgccta ccacctctc tggcaaggct gcaaagaaac 2820
 ccacagctgc agaggtctct gtctgagtc ctgagggcc ggccgtggaa gggggacacg 2880
 tgaatatggc attctctcag tccaacctc ctctggagtt gcacagggtc cacggatcca 2940
 gaaacaagcc caccagcttg tggaacccaa cgtacggctc ctggtttaca gagaaacca 3000
 ccaaaaagaa taatctata gcaaagaagg agccacacga gaggggtaac ctggggctgg 3060
 agggaagctg tactgtccca cctaacttg caactgggag acttccgggg gcctcactgc 3120
 tcttagagcc ctcttcgctg actgccaata tgaaggaggt acctctgttc aggctacgtc 3180

acttccttg tgggaatgc aattacggct accagcaaca ggccttgccc ttagaagccg 3240
 ctactcccc tggagctggt cattacagg ataccattt gaaaagcaag aatagcatga 3300
 accagcctgg gccctgagct cggtcgcaca ctcaattctc ttccttggga tccctaagac 3360
 cgtgg 3365

<210> 20

<211> 1082

<212> PRT

<213> Homo sapiens

<400> 20

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val

35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met

65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr

85 90 95

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

100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys

115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser

130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr

145 150 155 160

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

165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg

180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val

195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly

210 215 220

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

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

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

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

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

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

 595 600 605
 Val Ala Val Lys Thr Leu Pro Glu Val Cys Ser Glu Gln Asp Glu Leu
 610 615 620
 Asp Phe Leu Met Glu Ala Leu Ile Ile Ser Lys Phe Asn His Gln Asn
 625 630 635 640
 Ile Val Arg Cys Ile Gly Val Ser Leu Gln Ser Leu Pro Arg Phe Ile
 645 650 655
 Leu Leu Glu Leu Met Ala Gly Gly Asp Leu Lys Ser Phe Leu Arg Glu

 660 665 670
 Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser Leu Ala Met Leu Asp Leu
 675 680 685
 Leu His Val Ala Arg Asp Ile Ala Cys Gly Cys Gln Tyr Leu Glu Glu
 690 695 700
 Asn His Phe Ile His Arg Asp Ile Ala Ala Arg Asn Cys Leu Leu Thr
 705 710 715 720

Cys Pro Gly Pro Gly Arg Val Ala Lys Ile Gly Asp Phe Gly Met Ala

725 730 735

Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly Cys Ala Met

740 745 750

Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu Gly Ile Phe

755 760 765

Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu Trp Glu Ile

770 775 780

Phe Ser Leu Gly Tyr Met Pro Tyr Pro Ser Lys Ser Asn Gln Glu Val

785 790 795 800

Leu Glu Phe Val Thr Ser Gly Gly Arg Met Asp Pro Pro Lys Asn Cys

805 810 815

Pro Gly Pro Val Tyr Arg Ile Met Thr Gln Cys Trp Gln His Gln Pro

820 825 830

Glu Asp Arg Pro Asn Phe Ala Ile Ile Leu Glu Arg Ile Glu Tyr Cys

835 840 845

Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile Glu Tyr Gly

850 855 860

Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro Lys Asp Pro

865 870 875 880

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

885 890 895

Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly

900 905 910

Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val Arg Val Pro

915 920 925

Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala Phe Ser Gln

930 935 940

Ser Asn Pro Pro Ser Glu Leu His Arg Val His Gly Ser Arg Asn Lys

945 950 955 960

Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe Thr Glu Lys

965 970 975
 Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro His Glu Arg

 980 985 990
 Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro Asn Val Ala
 995 1000 1005
 Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro Ser Ser
 1010 1015 1020
 Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg His
 1025 1030 1035
 Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu

1040 1045 1050
 Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp
 1055 1060 1065
 Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro
 1070 1075 1080

<210> 21

<211> 3435

<212> DNA

<213> Homo sapiens

<400> 21

tactctgtcg gtccgctgaa tgaagtgcc gccctcttaa gcccgagacc cggcgctttc 60
 cccgcaagat ggacggtttc gccggcagtc tcgatgatag tatttctgct gcaagtactt 120

ctgatgttca agatcgcttg tcagctcttg agtcacgagt tcagcaacaa gaagatgaaa 180
 tcaactgtgct aaaggcggtt ttggctgatg ttttgaggcg tcttgcaatc tctgaagatc 240
 atgtggcctc agtgaacaaa tcagctctca gtaaaggcca accaagccct cgagcagtta 300
 ttcccatgtc ctgtataacc aatggaagtg gtgcaaacag aaaaccaagt cataccagtg 360
 ctgtctcaat tgcaggaaaa gaaactcttt catctgtctg taaaagtggc acagaaaaaa 420
 agaaagaaaa accacaagga cagagagaaa aaaaagagga atctcattct aatgatcaaa 480
 gtccacaaat tcgagcatca ctttctcccc agccctcttc acaacctctc caaatacaca 540

gacaaactcc agaaagcaag aatgtacttc ccacaaaag cataaaacga ccatcaccag 600
 ctgaaaagtc acataattct tgggaaaatt cagatgatag ccgtaataaa ttgtcgaaaa 660

taccttcaac acccaaatta ataccaaaaag ttaccaaaaac tgcagacaag cataaagatg	720
tcatcatcaa ccaagaagga gaatatatta aatgtttat gcgcggtcgg ccaattacca	780
tgttcattcc ttccgatgtt gacaactatg atgacatcag aacggaactg cctcctgaga	840
agctcaaact ggagtgggca tatggttatac gaggaagga ctgtagagct aatgtttacc	900
ttcttccgac cggggaaata gtttatttca ttgcatcagt agtagtacta tttaattatg	960
aggagagaac tcagcgacac tacctgggcc atacagactg tgtgaaatgc ctgtctatac	1020
atcctgacaa aattaggatt gcaactggac agatagctgg cgtggataaa gatggaaggc	1080
ctctacaacc ccacgtcaga gtgtgggatt ctgttactct atccacactg cagattattg	1140
gacttggcac ttttgagcgt ggagttaggat gcctggattt ttcaaaagca gattcaggtg	1200
ttcatttatg tgitattgat gactccaatg agcatatgct tactgtatgg gactggcaga	1260
ggaaagcaaa aggagcagaa ataaagacaa caaatgaagt tgttttgct gtggagttc	1320
acccaacaga tgcaaatacc ataattacat gcggtaaatc tcatattttc ttctggacct	1380
ggagcggcaa ttcactaaca agaaaacagg gaatttttgg gaaatatgaa aagccaaaat	1440
ttgtgcagtg tttagcattc ttggggaatg gagatgttct tactggagac tcaggtggag	1500
tcatgcttat atggagcaaa actactgtag agccacacc tgggaaagga cctaaaggtg	1560
tatatcaaat cagcaacaa atcaaagctc atgatggcag tgtgttcaca ctttgtcaga	1620
tgagaaatgg gatgttatta actggaggag ggaaagacag aaaaataatt ctgtgggatc	1680
atgatctgaa tctgaaaga gaaatagac accaggagct gcaagccatg cagatggagc	1740
tgcagagccc ttagtacaag ctgagcaagc tccgcacctc gaccatcatg accgactaca	1800
accccaacta ctgctttgct ggcaagacct cctccatcag tgacctgaag gaggtgccgc	1860
ggaaaaacat caccctcatt cggggtctgg gccatggagc ctttggggag gtgtatgaag	1920
gccaggtgtc cggaatgccc aacgacccaa gccccctgca agtggctgtg aagacgtgc	1980
ctgaagtgtg ctctgaacag gacgaactgg atttcctcat ggaagccctg atcatcagca	2040
aattcaacca ccagaacatt gtctgctgca ttgggtgtag cctgcaatcc ctgccccgt	2100
tcatctgct ggagctcatg gcggggggag acctcaagtc cttctccga gagaccgcc	2160
ctgccccgag ccagccctcc tccctggcca tgctggacct tctgcacgtg gctcgggaca	2220
ttgcctgtgg ctgtcagtat ttggaggaaa accacttcat ccaccgagac attgctgcca	2280
gaaactgcct ctigacctgt ccaggccctg gaagagtggc caagattgga gacttcggga	2340
tggcccgaga catctacagg gcgagctact atagaaaggg aggcctgtgcc atgctgccag	2400
ttaagtggat gccccagag gccttcatgg aaggaatatt cacttctaaa acagacacat	2460
ggtcctttgg agtctgcta tgggaaatct tttctcttgg atatatgcca taccacagca	2520

aaagcaacca ggaagtcttg gagtttgtca ccagtggagg ccgcatggac ccaccaaga 2580
 actgccttgg gcctgtatag cggataatga ctcatgtctg gcaacatcag cctgaagaca 2640

ggcccaactt tgccatcatt ttggagagga ttgaatactg caccaggac ccgcatgtaa 2700
 tcaacaccgc ttigccgata gaatatggtc cacttgttga agaggaagag aaagtgcctg 2760
 tgaggcccaa ggacctgag ggggttctc ctctcttggg ctctcaacag gcaaacggg 2820
 aggaggagcg cagcccagct gccccaccac ctctgcctac cacctcctct ggcaaggctg 2880
 caaagaaacc cacagctgca gaggtctctg ttcgagtccc tagagggccg gccgtggaag 2940
 ggggacacgt gaatatggca ttctctcagt ccaaccctcc ttcgagttg cacaaggctc 3000
 acgcatccag aaacaagccc accagcttgt ggaaccaac gtacggctcc tggtttacag 3060

agaaaccac caaaaagaat aatcctatag caaagaagga gccacacgac aggggtaacc 3120
 tgggctgga gggaagctgt actgtccac ctaacgttg aactgggaga ctccggggg 3180
 cctcactgct cctagagccc tcttcgctga ctgccaatat gaaggagta cctctgttca 3240
 ggctacgtca ctcccttgt gggaatgtca attacgcta ccagcaacag ggcttgcct 3300
 tagaagccg tactgccct ggagctggc attacgagga taccattctg aaaagcaaga 3360
 atagcatgaa ccagcctggg ccctgagctc ggtcgcacac tcaattctct tccttgggat 3420
 ccctaagacc gtgga 3435

<210> 22

<211> 1105

<212> PRT

<213> Homo sapiens

<400> 22

Met Asp Gly Phe Ala Gly Ser Leu Asp Asp Ser Ile Ser Ala Ala Ser

1 5 10 15

Thr Ser Asp Val Gln Asp Arg Leu Ser Ala Leu Glu Ser Arg Val Gln

20 25 30

Gln Gln Glu Asp Glu Ile Thr Val Leu Lys Ala Ala Leu Ala Asp Val

35 40 45

Leu Arg Arg Leu Ala Ile Ser Glu Asp His Val Ala Ser Val Lys Lys

50 55 60

Ser Val Ser Ser Lys Gly Gln Pro Ser Pro Arg Ala Val Ile Pro Met

65 70 75 80

Ser Cys Ile Thr Asn Gly Ser Gly Ala Asn Arg Lys Pro Ser His Thr
85 90 95

Ser Ala Val Ser Ile Ala Gly Lys Glu Thr Leu Ser Ser Ala Ala Lys
100 105 110

Ser Gly Thr Glu Lys Lys Lys Glu Lys Pro Gln Gly Gln Arg Glu Lys
115 120 125

Lys Glu Glu Ser His Ser Asn Asp Gln Ser Pro Gln Ile Arg Ala Ser
130 135 140

Pro Ser Pro Gln Pro Ser Ser Gln Pro Leu Gln Ile His Arg Gln Thr
145 150 155 160

Pro Glu Ser Lys Asn Ala Thr Pro Thr Lys Ser Ile Lys Arg Pro Ser
165 170 175

Pro Ala Glu Lys Ser His Asn Ser Trp Glu Asn Ser Asp Asp Ser Arg
180 185 190

Asn Lys Leu Ser Lys Ile Pro Ser Thr Pro Lys Leu Ile Pro Lys Val
195 200 205

Thr Lys Thr Ala Asp Lys His Lys Asp Val Ile Ile Asn Gln Glu Gly
210 215 220

Glu Tyr Ile Lys Met Phe Met Arg Gly Arg Pro Ile Thr Met Phe Ile
225 230 235 240

Pro Ser Asp Val Asp Asn Tyr Asp Asp Ile Arg Thr Glu Leu Pro Pro
245 250 255

Glu Lys Leu Lys Leu Glu Trp Ala Tyr Gly Tyr Arg Gly Lys Asp Cys
260 265 270

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

Ala Ser Val Val Val Leu Phe Asn Tyr Glu Glu Arg Thr Gln Arg His
290 295 300

Tyr Leu Gly His Thr Asp Cys Val Lys Cys Leu Ala Ile His Pro Asp
305 310 315 320

Lys Ile Arg Ile Ala Thr Gly Gln Ile Ala Gly Val Asp Lys Asp Gly

325 330 335
 Arg Pro Leu Gln Pro His Val Arg Val Trp Asp Ser Val Thr Leu Ser
 340 345 350
 Thr Leu Gln Ile Ile Gly Leu Gly Thr Phe Glu Arg Gly Val Gly Cys
 355 360 365
 Leu Asp Phe Ser Lys Ala Asp Ser Gly Val His Leu Cys Val Ile Asp

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

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

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

 565 570 575

Tyr Asn Pro Asn Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp
 580 585 590
 Leu Lys Glu Val Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly
 595 600 605
 His Gly Ala Phe Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro
 610 615 620
 Asn Asp Pro Ser Pro Leu Gln Val Ala Val Lys Thr Leu Pro Glu Val

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

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

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

820	825	830	
Met Asp Pro Pro Lys Asn Cys Pro Gly Pro Val Tyr Arg Ile Met Thr			
835	840	845	
Gln Cys Trp Gln His Gln Pro Glu Asp Arg Pro Asn Phe Ala Ile Ile			
850	855	860	
Leu Glu Arg Ile Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr			
865	870	875	880
Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val			
885	890	895	
Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser			
900	905	910	
Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro			
915	920	925	
Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala			
930	935	940	
Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His			
945	950	955	960
Val Asn Met Ala Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Lys			
965	970	975	
Val His Gly Ser Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr			
980	985	990	
Gly Ser Trp Phe Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala			
995	1000	1005	
Lys Lys Glu Pro His Asp Arg Gly Asn Leu Gly Leu Glu Gly Ser			
1010	1015	1020	
Cys Thr Val Pro Pro Asn Val Ala Thr Gly Arg Leu Pro Gly Ala			
1025	1030	1035	
Ser Leu Leu Leu Glu Pro Ser Ser Leu Thr Ala Asn Met Lys Glu			
1040	1045	1050	
Val Pro Leu Phe Arg Leu Arg His Phe Pro Cys Gly Asn Val Asn			
1055	1060	1065	

Tyr Gly Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala

1070	1075	1080
Pro Gly Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn		
1085	1090	1095
Ser Met Asn Gln Pro Gly Pro		
1100	1105	
<210> 23		
<211> 4479		
<212> DNA		
<213> Homo sapiens		
<400> 23		
tgcgagaaag atggcggacc tggccgagtg caacatcaaa gtgatgtgtc gcttcagacc	60	
tctcaacgag tctgaagtga accgcggcga caagtacatc gccaagtttc agggagaaga	120	
cacggtcgtg atcgctcca agccttatgc atttgatcgg gtgttccagt caagcacatc	180	
tcaagagcaa gtgtataatg actgtgcaaa gaagattgtt aaagatgtac ttgaaggata	240	
taatggaaca atatttgcac atggacaaac atcctctggg aagacacaca caatggaggg	300	
taaaattcat gatccagaag gcatgggaat tattccaaga atagtgaag atatttttaa	360	
ttatatattac tccatggatg aaaatttggg atttcatatt aaggtttcat attttgaaat	420	
atatttggat aagataaggg acctgttaga tgtttcaaag accaaccttt cagttcatga	480	
agacaaaaac cgagttccct atgtaaaggg gtgcacagag cgttttgtat gtagtccaga	540	
tgaagttatg gataccatag atgaaggaaa atccaacaga catglagcag ttacaaatat	600	
gaatgaacat agctctagga gtcacagtat atttcttatt aatgtcaaac aagagaacac	660	
acaaacggaa caaaagctga gtggaaaact ttatctgggt gatttagctg gtagtgaaaa	720	
ggttagtaaa actggagctg aagggtgctgt gctggatgaa gctaaaaaca tcaacaagtc	780	
actttctgct cttggaaatg ttatttctgc ttggctgag ggtagtacat atgttcata	840	
tcgagatagt aaaatgacaa gaatccttca agattcatta ggtggcaact gtagaaccac	900	
tattgtaatt tgctgctctc catcatcata caatgagtct gaaacaaaat ctacactctt	960	
atttgccaa agggccaaaa caattaagaa cacagtttgt gtcaatgtgg agttaactgc	1020	
agaacagtgg aaaaagaagt atgaaaaaga aaaagaaaaa aataagatcc tgcggaacac	1080	
tattcagtgg cttgaaaatg agctcaacag atggcgtaat ggggagacgg tgcctattga	1140	
tgaacagttt gacaaagaga aagccaactt ggaagctttc acagtggata aagatattac	1200	

tcttaccaat gataaaccag caaccgcaat tggagttata ggaaatttta ctgatgctga	1260
aagaagaag tgtaagaag aaattgctaa attatacaaa cagcttgatg acaaggatga	1320
agaaattaac cagcaaagtc aactggtaga gaaactgaag acgcaaatgt tggatcagga	1380
ggagcttttg gcatctacca gaagggatca agacaatatg caagctgagc tgaatcgctt	1440
tcaagcagaa aatgatgcct ctaaagaaga agtgaaagaa gttttacagg ccctagaaga	1500
acttgctgtc aattatgatc agaagtctca ggaagttgaa gacaaaacta aggaatatga	1560
attgcttagt gatgaattga atcagaaatc ggcaacttta gcgagtatag atgctgagct	1620
tcagaaactt aaggaaatga ccaaccacca gaaaaaacga gcagctgaga tgatggcatc	1680
tttactaaaa gaccttgtag aaataggaat tgctgtggga aataatgatg taaagcagcc	1740
tgagggaact ggcatgatag atgaagagtt cactgttgca agactctaca ttagcaaat	1800
gaagtcagaa gtaaaaacca tggtagaacg ttgcaagcag ttagaaagca cacaaactga	1860
gagcaacaaa aaaatggaag aaaatgaaaa ggagtttagca gcatgtcagc ttcgtatctc	1920
tcaacatgaa gccaaaatca agtcattgac tgaatacctt caaaatgtgg aacaaaagaa	1980
aagacagttg gaggaatctg tcgatgcctt cagtgaagaa ctagtccagc ttcgagcaca	2040
agagaaagtc catgaaatgg aaaaggagca cttaataaag gttcagactg caaatgaagt	2100
taagcaagct gttgaacagc agatccagag ccatagagaa actcatcaaa aacagatcag	2160
tagtttgaga gatgaagtag aagcaaaagc aaaacttatt actgatcttc aagacaaaaa	2220
ccagaaaatg atgttagagc aggaacgtct aagagtagaa catgagaagt tgaaagccac	2280
agatcaggaa aagagcagaa aactacatga acttacggtt atgcaagata gacgagaaca	2340
agcaagacaa gacttgaagg gtttgaaga gacagtggca aaagaacttc agactttaca	2400
caacctgcgc aaactctttg ttcaggacct ggctacaaga gttaaaaaga gtgctgagat	2460
tgattctgat gacaccggag gcagcgctgc tcagaagcaa aaaatctcct ttccttgaana	2520
taatcttgaa cagctcacta aagtgcacaa acagttggta cgtgataatg cagatctccg	2580
ctgtgaactt cctaagttgg aaaagcgact tcgagctaca gctgagagag tgaaagcttt	2640
ggaatcagca ctgaaagaag ctaaagaaaa tgcattctgt gatcgcaaac gctatcagca	2700
agaagtagat cgcataaagg aagcagtcag gtcaaagaat atggccagaa gagggcattc	2760
tgacagatt gtgtaccgc ggaagcacca ggagctgcaa gccatgcaga tggagctgca	2820
gagccctgag tacaagctga gcaagctccg cacctcgacc atcatgaccg actacaaccc	2880
caactactgc tttgctggca agacctctc catcagtgac ctgaaggagg tgcccgga	2940
aaacatcacc ctcatcggg gtctgggcca tggcgctttt ggggaggtgt atgaaggcca	3000
ggtgtccgga atgcccacg acccaagccc cctgcaagtg gctgtgaaga cgctgcctga	3060

agtggtgctct gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaatt 3120

caaccaccag aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat 3180

cctgtctggag ctcatggcgg ggggagacct caagtccttc ctccgagaga cccgccctcg 3240

cccagaccag ccctcctccc tggccatgct ggaccttctg cacgtggctc gggacattgc 3300

ctgtggtgtg cagtatttgg aggaaaacca ctcatccac cgagacattg ctgccagaaa 3360

ctgcctcttg acctgtccag gccctggaag agtggccaag attggagact tcgggatggc 3420

ccgagacatc tacagggcga gctactatag aaaggaggc tgtgcatgc tgccagttaa 3480

gtggatgccc ccagaggcct tcatggaagg aatattcact tctaaaacag acacatggtc 3540

ctttggagtg ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag 3600

caaccaggaa gttctggagt ttgtcaccag tggaggccgg atggaccac ccaagaactg 3660

ccctgggctt gtataccga taatgactca gtgctggcaa catcagcctg aagacaggcc 3720

caactttgcc atcattttgg agaggattga atactgcacc caggacccgg atgtaatcaa 3780

caccgctttg ccgatagaat atggtccact tgtggaagag gaagagaaag tgcctgtgag 3840

gccaaggac cctgaggggg ttctcctctt cctggtctct caacaggcaa aacgggagga 3900

ggagcgcagc ccagctgccc caccacctct gcctaccacc tcctctggca aggctgcaaa 3960

gaaaccacaca gctgcagagg tctctgttcg agtcctaga gggccggccg tggaaagggg 4020

acacgtgaat atggcattct ctcatgcaa ccctccttcg gagttgcaca aggtccacgg 4080

atccagaaac aagcccacca gcttgtggaa cccaacgtac ggctcctggt ttacagagaa 4140

accacacaaa aagaataatc ctatagcaaa gaaggagcca cacgacagg gtaacctggg 4200

gctggaggga agctgtactg tccaccta cgttgcaact gggagacttc cgggggcctc 4260

actgctccta gagccctctt cgctgactgc caatatgaag gaggtacctc tgttcaggct 4320

acgtcacttc ccttgtggga atgtcaatta cggctaccag caacagggt tgccttaga 4380

agccgtact gccctggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag 4440

catgaaccag cctgggccct gagctcggtc gcacactca 4479

<210> 24

<211> 1483

<212> PRT

<213> Homo sapiens

<400> 24

Met Ala Asp Leu Ala Glu Cys Asn Ile Lys Val Met Cys Arg Phe Arg

1 5 10 15

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

260 265 270
 Thr Tyr Val Pro Tyr Arg Asp Ser Lys Met Thr Arg Ile Leu Gln Asp
 275 280 285
 Ser Leu Gly Gly Asn Cys Arg Thr Thr Ile Val Ile Cys Cys Ser Pro

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

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

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

 485 490 495
 Glu Leu Ala Val Asn Tyr Asp Gln Lys Ser Gln Glu Val Glu Asp Lys
 500 505 510

Thr Lys Glu Tyr Glu Leu Leu Ser Asp Glu Leu Asn Gln Lys Ser Ala
 515 520 525
 Thr Leu Ala Ser Ile Asp Ala Glu Leu Gln Lys Leu Lys Glu Met Thr
 530 535 540
 Asn His Gln Lys Lys Arg Ala Ala Glu Met Met Ala Ser Leu Leu Lys
 545 550 555 560
 Asp Leu Ala Glu Ile Gly Ile Ala Val Gly Asn Asn Asp Val Lys Gln
 565 570 575
 Pro Glu Gly Thr Gly Met Ile Asp Glu Glu Phe Thr Val Ala Arg Leu
 580 585 590
 Tyr Ile Ser Lys Met Lys Ser Glu Val Lys Thr Met Val Lys Arg Cys
 595 600 605
 Lys Gln Leu Glu Ser Thr Gln Thr Glu Ser Asn Lys Lys Met Glu Glu
 610 615 620
 Asn Glu Lys Glu Leu Ala Ala Cys Gln Leu Arg Ile Ser Gln His Glu
 625 630 635 640
 Ala Lys Ile Lys Ser Leu Thr Glu Tyr Leu Gln Asn Val Glu Gln Lys
 645 650 655
 Lys Arg Gln Leu Glu Glu Ser Val Asp Ala Leu Ser Glu Glu Leu Val
 660 665 670
 Gln Leu Arg Ala Gln Glu Lys Val His Glu Met Glu Lys Glu His Leu
 675 680 685
 Asn Lys Val Gln Thr Ala Asn Glu Val Lys Gln Ala Val Glu Gln Gln
 690 695 700
 Ile Gln Ser His Arg Glu Thr His Gln Lys Gln Ile Ser Ser Leu Arg
 705 710 715 720
 Asp Glu Val Glu Ala Lys Ala Lys Leu Ile Thr Asp Leu Gln Asp Gln
 725 730 735
 Asn Gln Lys Met Met Leu Glu Gln Glu Arg Leu Arg Val Glu His Glu
 740 745 750
 Lys Leu Lys Ala Thr Asp Gln Glu Lys Ser Arg Lys Leu His Glu Leu

755	760	765	
Thr Val Met Gln Asp Arg Arg Glu Gln Ala Arg Gln Asp Leu Lys Gly			
770	775	780	
Leu Glu Glu Thr Val Ala Lys Glu Leu Gln Thr Leu His Asn Leu Arg			
785	790	795	800
Lys Leu Phe Val Gln Asp Leu Ala Thr Arg Val Lys Lys Ser Ala Glu			
	805	810	815
Ile Asp Ser Asp Asp Thr Gly Gly Ser Ala Ala Gln Lys Gln Lys Ile			
820	825	830	
Ser Phe Leu Glu Asn Asn Leu Glu Gln Leu Thr Lys Val His Lys Gln			
835	840	845	
Leu Val Arg Asp Asn Ala Asp Leu Arg Cys Glu Leu Pro Lys Leu Glu			
850	855	860	
Lys Arg Leu Arg Ala Thr Ala Glu Arg Val Lys Ala Leu Glu Ser Ala			
865	870	875	880
Leu Lys Glu Ala Lys Glu Asn Ala Ser Arg Asp Arg Lys Arg Tyr Gln			
885	890	895	
Gln Glu Val Asp Arg Ile Lys Glu Ala Val Arg Ser Lys Asn Met Ala			
900	905	910	
Arg Arg Gly His Ser Ala Gln Ile Val Tyr Arg Arg Lys His Gln Glu			
915	920	925	
Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu Ser			
930	935	940	
Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn Tyr Cys			
945	950	955	960
Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val Pro Arg			
965	970	975	
Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe Gly Glu			
980	985	990	
Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro Ser Pro Leu			
995	1000	1005	

Gln Val	Ala Val Lys Thr	Leu Pro Glu Val	Cys Ser	Glu Gln Asp
1010		1015		1020
Glu Leu	Asp Phe Leu Met	Glu Ala Leu Ile	Ile Ser	Lys Phe Asn
1025		1030		1035
His Gln	Asn Ile Val Arg	Cys Ile Gly Val	Ser Leu	Gln Ser Leu
1040		1045		1050
Pro Arg	Phe Ile Leu Leu	Glu Leu Met Ala	Gly Gly	Asp Leu Lys
1055		1060		1065
Ser Phe	Leu Arg Glu Thr	Arg Pro Arg Pro	Ser Gln	Pro Ser Ser
1070		1075		1080
Leu Ala	Met Leu Asp Leu	Leu His Val Ala	Arg Asp	Ile Ala Cys
1085		1090		1095
Gly Cys	Gln Tyr Leu Glu	Glu Asn His Phe	Ile His	Arg Asp Ile
1100		1105		1110
Ala Ala	Arg Asn Cys Leu	Leu Thr Cys Pro	Gly Pro	Gly Arg Val
1115		1120		1125
Ala Lys	Ile Gly Asp Phe	Gly Met Ala Arg	Asp Ile	Tyr Arg Ala
1130		1135		1140
Ser Tyr	Tyr Arg Lys Gly	Gly Cys Ala Met	Leu Pro	Val Lys Trp
1145		1150		1155
Met Pro	Pro Glu Ala Phe	Met Glu Gly Ile	Phe Thr	Ser Lys Thr
1160		1165		1170
Asp Thr	Trp Ser Phe Gly	Val Leu Leu Trp	Glu Ile	Phe Ser Leu
1175		1180		1185
Gly Tyr	Met Pro Tyr Pro	Ser Lys Ser Asn	Gln Glu	Val Leu Glu
1190		1195		1200
Phe Val	Thr Ser Gly Gly	Arg Met Asp Pro	Pro Lys	Asn Cys Pro
1205		1210		1215
Gly Pro	Val Tyr Arg Ile	Met Thr Gln Cys	Trp Gln	His Gln Pro
1220		1225		1230
Glu Asp	Arg Pro Asn Phe	Ala Ile Ile Leu	Glu Arg	Ile Glu Tyr

1235	1240	1245
Cys Thr Gln Asp Pro Asp Val	Ile Asn Thr Ala Leu	Pro Ile Glu
1250	1255	1260
Tyr Gly Pro Leu Val Glu Glu	Glu Glu Lys Val Pro	Val Arg Pro
1265	1270	1275
Lys Asp Pro Glu Gly Val Pro	Pro Leu Leu Val Ser	Gln Gln Ala
1280	1285	1290
Lys Arg Glu Glu Glu Arg Ser	Pro Ala Ala Pro Pro	Pro Leu Pro

1295	1300	1305
Thr Thr Ser Ser Gly Lys Ala	Ala Lys Lys Pro Thr	Ala Ala Glu
1310	1315	1320
Val Ser Val Arg Val Pro Arg	Gly Pro Ala Val Glu	Gly Gly His
1325	1330	1335
Val Asn Met Ala Phe Ser Gln	Ser Asn Pro Pro Ser	Glu Leu His
1340	1345	1350
Lys Val His Gly Ser Arg Asn	Lys Pro Thr Ser Leu	Trp Asn Pro

1355	1360	1365
Thr Tyr Gly Ser Trp Phe Thr	Glu Lys Pro Thr Lys	Lys Asn Asn
1370	1375	1380
Pro Ile Ala Lys Lys Glu Pro	His Asp Arg Gly Asn	Leu Gly Leu
1385	1390	1395
Glu Gly Ser Cys Thr Val Pro	Pro Asn Val Ala Thr	Gly Arg Leu
1400	1405	1410
Pro Gly Ala Ser Leu Leu Leu	Glu Pro Ser Ser Leu	Thr Ala Asn

1415	1420	1425
Met Lys Glu Val Pro Leu Phe	Arg Leu Arg His Phe	Pro Cys Gly
1430	1435	1440
Asn Val Asn Tyr Gly Tyr Gln	Gln Gln Gly Leu Pro	Leu Glu Ala
1445	1450	1455
Ala Thr Ala Pro Gly Ala Gly	His Tyr Glu Asp Thr	Ile Leu Lys
1460	1465	1470
Ser Lys Asn Ser Met Asn Gln	Pro Gly Pro	

1475

1480

<210> 25

<211> 2412

<212> DNA

<213> Homo sapiens

<400> 25

atgaacggac agttaggatct aagtgggaag ctaatcatca aagctcaact tggggaggat	60
attcgcgcaa ttcctattca taatgaagat attacttatg atgaattagt gctaatgatg	120
caacgagttt tcagaggaaa acttctgagt aatgatgaag taacaataaa gtataaagat	180
gaagatggag atcttataac aatttttgat agttctgacc tttcctttgc aattcagtg	240
agtaggatac tgaactgac attatttgtt aatggccagc caagaccct tgaatcaagt	300
caggtgaaat atctccgtcg agaactgata gaacttcgaa ataaagtga tcgtttattg	360
gatagcttgg aaccacctgg agaaccagga ccttcacca atattcctga aaatgatact	420
gtggatggta gggaagaaaa gtctgcttct gattcttctg gaaaacagtc tactcaggtt	480
atggcagcaa gtatgtctgc ttttgatcct ttaaaaaacc aagatgaaat caataaaat	540
gttatgtcag cgtttggctt aacagatgat caggtttcag ggccaccag tgctcctgca	600
gaagatcgtt caggaacacc cgacagcatt gcttctcct cctcagcagc tcaccacca	660
ggcgttcagc cacagcagcc accatataca ggagctcaga ctcaagcagg tcagattgaa	720
gtgtaccgcc ggaagcacca ggagctgcaa gccatgcaga tggagctgca gagccctgag	780
tacaagctga gcaagctccg cacctcgacc atcatgaccg actacaacc caactactgc	840
tttgctggca agacctctc catcagtgac ctgaaggagg tgccgcggaa aaacatcacc	900
ctcattcggg gtctgggcca tggcgctttt ggggaggtgt atgaaggcca ggtgtccgga	960
atgccaacg acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgctct	1020
gaacaggacg aactggattt cctcatggaa gccctgatca tcagcaaatt caaccaccag	1080
aacattgttc gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgctggag	1140
ctcatggcgg ggggagacct caagtccttc ctccgagaga cccgccctcg cccgagccag	1200
ccctctctcc tggccatgct ggaccttctg cacgtggctc gggacattgc ctgtggtgt	1260
cagtatttgg aggaaaacca cttcatccac cgagacattg ctgccagaaa ctgcctcttg	1320
acctgtccag gccctggaag agtggccaag attggagact tcgggatggc ccgagacatc	1380
tacagggcga gctactatag aaaggaggc tgtgccatgc tgccagttaa gtggatgccc	1440
ccagaggcct tcatggaagg aatattcact tctaaaacag acacatggtc ctttggagt	1500

ctgctatggg aaatcttttc tcttggatat atgccatacc ccagcaaaag caaccaggaa 1560
gttctggagt ttgtcaccag tggaggccgg atggacccac ccaagaactg cctgggcct 1620

gtataccgga taatgactca gtgctggcaa catcagcctg aagacaggcc caactttgcc 1680
atcatttttg agaggattga atactgcacc caggaccggg atgtaatcaa caccgctttg 1740
ccgatagaat atggtccact tgtggaagag gaagagaaag tgcctgtgag gccaaggac 1800
cctgaggggg ttctctctct cctggtctct caacaggcaa aacgggagga ggagcgcagc 1860
ccagctgccc caccacctct gcctaccacc tcctctggca aggctgcaaa gaaaccaca 1920
gctgcagagg tctctgttcg agtcctaga gggccggccg tggaaggggg acacgtgaat 1980
atggcattct ctcagtccaa ccctccttcg gagttgcaca aggtccacgg atccagaaac 2040

aagcccacca gcttgtggaa cccaacgtac ggctcctggt ttacagagaa acccaccaaa 2100
aagaataatc ctatagcaaa gaaggagcca cagcacaggg gtaacctggg gctggaggga 2160
agctgtactg tccaccta cgttgcaact gggagacttc cgggggcctc actgctccta 2220
gagccctctt cgtgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc 2280
ccttgtggga atgtcaatta cggctaccag caacagggtc tgccttaga agccgtact 2340
gccccgggag ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag 2400
cctgggcct ga 2412

<210> 26

<211> 803

<212> PRT

<213> Homo sapiens

<400> 26

Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln

1 5 10 15

Leu Gly Glu Asp Ile Arg Arg Ile Pro Ile His Asn Glu Asp Ile Thr

20 25 30

Tyr Asp Glu Leu Val Leu Met Met Gln Arg Val Phe Arg Gly Lys Leu

35 40 45

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

50 55 60

Leu Ile Thr Ile Phe Asp Ser Ser Asp Leu Ser Phe Ala Ile Gln Cys

65 70 75 80

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

325 330 335
 Glu Val Cys Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu
 340 345 350
 Ile Ile Ser Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val
 355 360 365
 Ser Leu Gln Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly

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

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

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

 565 570 575

Asn Thr Ala Leu Pro Ile Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu
 580 585 590
 Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu Leu
 595 600 605
 Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro
 610 615 620
 Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr

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

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

 755 760 765
 Tyr Gln Gln Gln Gly Leu Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala
 770 775 780
 Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln
 785 790 795 800
 Pro Gly Pro

<210> 27

<211> 2779

<212> DNA

<213> Homo sapiens

<400> 27

cctccgaag ccgtctttct ctagagtgtg atatatagaa catcctggag tccacatga	60
acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc	120
ggcgaattcc tattcataat gaagatatta cttatgatga attagtgcata atgatgcaac	180
gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag	240
atggagatct tataacaatt ttgatagtt ctgacctttc ctttgcaatt cagtgcagta	300
ggatactgaa actgacatta ttgtttaatg gccagccaag accccttgaa tcaagtcagg	360
tgaaatatct ccgtcgagaa ctgatagaac ttcgaaataa agtgaatcgt ttattggata	420
gcttggaaac acctggagaa ccaggacctt ccaccaatat tctgaaaat gatactgtgg	480
atggtaggga agaaaagtct gcttctgatt cttctggaaa acagtcctact caggttatgg	540
cagcaagtat gtctgctttt gatcctttta aaaaccaaga tgaaatcaat aaaaatgtta	600
tgtcagcgtt tggcttaaca gatgatcagg ttccagtgtg ccgccggaag caccaggagc	660
tgcaagccat gcagatggag ctgcagagcc ctgagtacaa gctgagcaag ctccgcacct	720
cgaccatcat gaccgactac aaccccaact actgctttgc tggcaagacc tctccatca	780
gtgacctgaa ggaggtgccg cggaaaaaca tcacctcat tcggggtctg ggccatggcg	840
cctttgggga ggtgtatgaa ggccaggtgt ccggaatgcc caacgacca agccccctgc	900
aagtggctgt gaagacgtg cctgaagtgt gctctgaaca ggacgaactg gatttcctca	960
tggaagccct gatcatcagc aaattcaacc accagaacat tgttcgctgc attggggtga	1020
gcctgcaatc cctgccccgg ttcatcctgc tggagctcat ggccggggga gacctcaagt	1080
ccttctctcg agagaccgc cctgccccga gccagccctc ctcctggcc atgctggacc	1140
ttctgcacgt ggctcgggac attgctgtg gctgtcagta ttggaggaa aaccacttca	1200
tccaccgaga cattgtgcc agaaactgcc tcttgacctg tccaggccct ggaagagtgg	1260
ccaagattgg agacttcggg atggcccgag acatctacag ggcgagctac tatagaaagg	1320
gaggctgtgc catgctgcca gttaagtga tgccccaga ggccttcag gaaggaatat	1380
tcaattctaa aacagacaca tggctctttg gactgtgtgt atgggaaatc ttttctcttg	1440
gatatatgcc ataccaccgc aaaagcaacc aggaagtctt ggagtttgc accagtggag	1500
gccggatgga cccaccaag aactgccctg ggctgtata ccggataatg actcagtgt	1560
ggcaacatca gctgaagac aggcccaact ttgcatcat ttggagagg attgaatact	1620
gcaccaggga ccggatgta atcaacaccg ctttgccgat agaataatgt cacttgttg	1680

aagaggaaga gaaagtgcct gtgaggccca aggacctga gggggttcct cctctcctgg 1740
tctctcaaca ggcaaacgg gaggaggagc gcagcccagc tgccccacca cctctgccta 1800

ccacctcttc tggcaaggct gcaaagaaac ccacagctgc agaggtctct gttcgagtec 1860
ctagagggcc ggccgtggaa gggggacacg tgaatatggc attctctcag tccaaccctc 1920
cttcggagtt gcacaaggct cacggtacca gaaacaagcc caccagcttg tggaacccaa 1980
cgtacggctc ctggtttaca gagaaaccca ccaaaaagaa taatcctata gcaaagaagg 2040
agccacacga caggggtaac ctggggctgg agggaagctg tactgtccca cctaacttg 2100
caactgggag acttcggggg gcctcactgc tcctagagcc ctcttcgctg actgccaata 2160
tgaaggaggt acctctgttc aggctacgtc acttccttg tgggaatgtc aattacggct 2220

accagcaaca gggttgccc ttagaagccg ctactgccc tggagctggt cattacgagg 2280
ataccattct gaaaagcaag aatagcatga accagcctgg gcctgagct cggtcgcaca 2340
ctcacttctc ttcttggga tcctaagac cgtggaggag agagaggcaa tggtccttc 2400
acaaaccaga gaccaaagt cagttttgt tttgtgcaa cctattttga agtaccacca 2460
aaaaagctgt atttgaaaa tgctttagaa aggttttgag catgggttca tcctattctt 2520
tcgaaagaag aaaatatcat aaaaatgagt gataaataca aggcccagat gtggttgcatt 2580
aaggttttta tgcattttg ttgtatactt ccttatgctt cttttaaat gtgtgtgctc 2640

tgcttcaatg tagtcagaat tagctgcttc tatgtttcat agttggggtc atagatgttt 2700
ccttgcttg ttgatgtgga catgagccat ttgaggggag agggaacgga aataaaggag 2760
ttatttgtaa tgactaaaa 2779

<210> 28
<211> 756
<212> PRT
<213> Homo sapiens
<400> 28

Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln
1 5 10 15
Leu Gly Glu Asp Ile Arg Arg Ile Pro Ile His Asn Glu Asp Ile Thr
20 25 30

Tyr Asp Glu Leu Val Leu Met Met Gln Arg Val Phe Arg Gly Lys Leu
35 40 45
Leu Ser Asn Asp Glu Val Thr Ile Lys Tyr Lys Asp Glu Asp Gly Asp

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

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

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

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

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

 Glu Lys Val Pro Val Arg Pro Lys Asp Pro Glu Gly Val Pro Pro Leu

545 550 555 560
 Leu Val Ser Gln Gln Ala Lys Arg Glu Glu Glu Arg Ser Pro Ala Ala
 565 570 575
 Pro Pro Pro Leu Pro Thr Thr Ser Ser Gly Lys Ala Ala Lys Lys Pro
 580 585 590
 Thr Ala Ala Glu Val Ser Val Arg Val Pro Arg Gly Pro Ala Val Glu
 595 600 605

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

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

Ala Gly His Tyr Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn
 740 745 750
 Gln Pro Gly Pro
 755

<210> 29

<211> 2614

<212> DNA

<213> Homo sapiens

<400> 29

cctccgcaag ccgtctttct cttagattgt atatatagaa catcctggag tccacatga 60
 acggacagtt ggatctaagt gggaagctaa tcatcaaagc tcaacttggg gaggatattc 120
 ggcaattcc tattcataat gaagatatta cttagatga attagtcta atgatgcaac 180
 gagttttcag aggaaaactt ctgagtaatg atgaagtaac aataaagtat aaagatgaag 240

 atggagatct tataacaatt tttagattgt ctgaccttct ctttgaatt cagtgcagta 300
 ggatactgaa actgacatta ttgttaatg gccagccaag accccttgaa tcaagtcagg 360
 tgaaatatct ccgtcgagaa ctgatagaac ttgaaataa agtgaatcgt ttattggata 420
 gcttggaaac acctggagaa ccaggacctt ccaccaatat tctgaaaat gtgtaccgcc 480
 ggaagcacca ggagctgcaa gccatgcaga tggagctgca gagccctgag tacaagctga 540
 gcaagctccg cacctcgacc atcatgaccg actacaacc ccaactactgc ttgtctggca 600
 agacctctc catcagtac ctgaaggagg tgcgcggaa aaacatcacc ctcatcggg 660

 gtctgggcca tggcgctttt ggggaggtgt atgaaggcca ggtgtccgga atgccaacg 720
 acccaagccc cctgcaagtg gctgtgaaga cgctgcctga agtgtgtct gaacaggacg 780
 aactggattt cctcatggaa gccctgatca tcagcaaatt caaccaccag aacattgttc 840
 gctgcattgg ggtgagcctg caatccctgc cccggttcat cctgtggag ctcatggcgg 900
 ggggagacct caagtcttc ctccgagaga cccgcctcg cccgagccag cctctctccc 960
 tggccatgct ggacctctg cacgtggctc gggacattgc ctgtggctgt cagtatttgg 1020
 aggaaaacca ctcatccac cgagacattg ctgccagaaa ctgcctcttg acctgtccag 1080

 gccctggaag agtggccaag attggagact tcgggatggc ccgagacatc tacaggcgga 1140
 gctactatag aaaggaggcg tgtgcatgc tgccagttaa gtggatgccc ccagaggcct 1200
 tcatggaagg aatattcact tctaaaacag acacatggtc ctttggagt ctgctatggg 1260
 aaatcttttc tcttgatatt atgccatacc ccagcaaaag caaccaggaa gttctggagt 1320
 ttgtcaccag tggaggccgg atggaccac ccaagaactg cctgggcct gtataccgga 1380
 taatgactca gtctggcaa catcagcctg aagacaggcc caactttgcc atcattttg 1440
 agaggattga atactgcacc caggaccggg atgtaatcaa caccgctttg ccgatagaat 1500

 atgttccact tgtggaagag gaagagaaag tgctgtgag gcccaaggac cctgaggggg 1560
 ttctctctct cctggtctct caacaggcaa aacgggagga ggagcgcagc ccagctgccc 1620
 caccacctct gcctaccacc tctctggca aggttgcaaa gaaaccaca gctgcagagg 1680
 tctctgttcg agtccctaga gggccggccg tggaaggggg acacgtgaat atggcattct 1740
 ctcatgcaa cctctctcg gagttgcaca aggtccacgg atccagaaac aagcccacca 1800
 gcttgtggaa cccaacgtac ggctcctggt ttacagagaa acccaccaa aagaataatc 1860

ctatagcaaa gaaggagcca cacgacaggg gtaacctggg gctggaggga agctgtactg 1920

tcccacctaa cgttgcaact gggagacttc cgggggcctc actgctccta gagccctctt 1980

cgctgactgc caatatgaag gaggtacctc tgttcaggct acgtcacttc ccttgtggga 2040

atgtcaatta cggctaccag caacagggct tgcccttaga agccgctact gcccctggag 2100

ctggtcatta cgaggatacc attctgaaaa gcaagaatag catgaaccag cctgggcctt 2160

gagctcggtc gcacactcac ttctcttctt tgggatccct aagaccgtgg aggagagaga 2220

ggcaatggct ccttcacaaa ccagagacca aatgtcacgt tttgttttgt gccaacctat 2280

tttgaagtac caccaaaaaa gctgtatttt gaaaatgctt tagaaagggtt ttgagcatgg 2340

gttcatccta ttctttcgaa agaagaaaat atcataaaaa tgagtataa atacaaggcc 2400

cagatgtggt tgcataaggt ttttatgcat gtttgttgta tacttctta tgettctttt 2460

aaattgtgtg tgctctgctt caatgtagtc agaattagct gcttctatgt ttcatagttg 2520

gggtcataga tgtttccttg ccttgttgat gtggacatga gccatttgag gggagaggga 2580

acggaataa aggagttatt tgtaatgact aaaa 2614

<210> 30

<211> 701

<212> PRT

<213> Homo sapiens

<400> 30

Met Asn Gly Gln Leu Asp Leu Ser Gly Lys Leu Ile Ile Lys Ala Gln

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

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

Gly Met Ala Arg Asp Ile Tyr Arg Ala Ser Tyr Tyr Arg Lys Gly Gly
 355 360 365
 Cys Ala Met Leu Pro Val Lys Trp Met Pro Pro Glu Ala Phe Met Glu
 370 375 380
 Gly Ile Phe Thr Ser Lys Thr Asp Thr Trp Ser Phe Gly Val Leu Leu

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

 450 455 460
 Glu Tyr Cys Thr Gln Asp Pro Asp Val Ile Asn Thr Ala Leu Pro Ile
 465 470 475 480
 Glu Tyr Gly Pro Leu Val Glu Glu Glu Glu Lys Val Pro Val Arg Pro
 485 490 495
 Lys Asp Pro Glu Gly Val Pro Pro Leu Leu Val Ser Gln Gln Ala Lys
 500 505 510
 Arg Glu Glu Glu Arg Ser Pro Ala Ala Pro Pro Pro Leu Pro Thr Thr

 515 520 525
 Ser Ser Gly Lys Ala Ala Lys Lys Pro Thr Ala Ala Glu Val Ser Val
 530 535 540
 Arg Val Pro Arg Gly Pro Ala Val Glu Gly Gly His Val Asn Met Ala
 545 550 555 560
 Phe Ser Gln Ser Asn Pro Pro Ser Glu Leu His Lys Val His Gly Ser
 565 570 575
 Arg Asn Lys Pro Thr Ser Leu Trp Asn Pro Thr Tyr Gly Ser Trp Phe

 580 585 590
 Thr Glu Lys Pro Thr Lys Lys Asn Asn Pro Ile Ala Lys Lys Glu Pro

595	600	605	
His Asp Arg Gly Asn Leu Gly Leu Glu Gly Ser Cys Thr Val Pro Pro			
610	615	620	
Asn Val Ala Thr Gly Arg Leu Pro Gly Ala Ser Leu Leu Leu Glu Pro			
625	630	635	640
Ser Ser Leu Thr Ala Asn Met Lys Glu Val Pro Leu Phe Arg Leu Arg			
	645	650	655
His Phe Pro Cys Gly Asn Val Asn Tyr Gly Tyr Gln Gln Gln Gly Leu			
	660	665	670
Pro Leu Glu Ala Ala Thr Ala Pro Gly Ala Gly His Tyr Glu Asp Thr			
	675	680	685
Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro			
	690	695	700
<210> 31			
<211> 1607			
<212> DNA			
<213> Homo sapiens			
<400> 31			
aactccgctg cctttgccgc caccatgccc aaaacgatca gtgtgcgtgt gaccaccatg	60		
gatgcagagc tggagtttgc catccagccc aacaccaccg ggaagcagct atttgaccag	120		
gtggtgaaaa ctattggctt gaggaagt tggttctttg gtctgcagta ccaggacact	180		
aaaggtttct ccacctggct gaaactcaat aagaaggtga ctgcccagga tgtgcggaag	240		
gaaagccccc tgctctttaa gttccgtgcc aagttctacc ctgaggatgt gtccgaggaa	300		
ttgattcagg acatcactca gcgcctgttc tttctgcaag tgaaagaggg cattctcaat	360		
gatgatattt actgcccgcc tgagaccgct gtgctgctgg cctcgtatgc tgtccagtct	420		
aagtatggcg acttcaataa ggaagtgc at aagtctggct acctggccgg agacaagttg	480		
ctccgcaga gagtcttga acagcacaaa ctcaacaagg accagtggga ggagcggatc	540		
caggtgtggc atgaggaaca ccgtggcatg ctgaggagg atgtgtctt ggaatatctg	600		
aagattgctc aagatctgga gatgtatggt gtgaactact tcagcatcaa gaacaagaaa	660		
ggctcagagc tgtggctggg ggtggatgcc ctgggtctca acatctatga gcagaatgac	720		
agactaactc ccaagatagg ctccccctgg agtgaaatca ggaacatctc ttccaatgat	780		
aagaaatttg tcatcaagcc cattgacaaa aaagcccg acttcgtctt ctatgtccc	840		

cggctgcgga ttaacaagcg gatcttggcc ttgtgcatgg ggaacatga actatacatg 900

cgccgtcgca agcctgatac cattgaggtg cagcagatga aggcacaggc ccgggaggag 960

aagcaccaga agcagatgga gcgtgctatg ctggaaaatg agaagaagaa gcgtgaaatg 1020

gcagagaagg agaaagagaa gattgaacgg gagaaggagg agctgatgga gaggctgaag 1080

cagatcgagg aacagactaa gaaggctcag caagaactgg aagaacagac ccgtagggct 1140

ctggaacttg agcaggaacg gaagcgtgcc cagagcgagg ctgaaaagct ggccaaggag 1200

cgtcaagaag ctgaagaggc caaggaggcc ttgctgcagg cctcccgga ccagaaaaag 1260

actcaggaac agctggcctt ggaaatggca gagctgacag ctggaatctc ccagctggag 1320

atggcccgac agaagaagga gagtgaggct gtggagtggc agcagaagca ggagctgcaa 1380

gccatgcaga tggagctgca gagccctgag tacaagctga gcaagctccg cacctcgacc 1440

atcatgaccg actacaacc caactactgc tttgtggca agacctctc catcagtac 1500

ctgaaggagg tgccgcggaa aaacatcacc ctattcggg gtctgggcca tggcgcttt 1560

ggggaggtgt atgaaggcca ggtgtccgga atgccaacg acccaag 1607

<210> 32

<211> 527

<212> PRT

<213> Homo sapiens

<400> 32

Met Pro Lys Thr Ile Ser Val Arg Val Thr Thr Met Asp Ala Glu Leu

1 5 10 15

Glu Phe Ala Ile Gln Pro Asn Thr Thr Gly Lys Gln Leu Phe Asp Gln

20 25 30

Val Val Lys Thr Ile Gly Leu Arg Glu Val Trp Phe Phe Gly Leu Gln

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

Ile Thr Gln Arg Leu Phe Phe Leu Gln Val Lys Glu Gly Ile Leu Asn

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

Gln Ile Glu Glu Gln Thr Lys Lys Ala Gln Gln Glu Leu Glu Glu Gln
355 360 365
Thr Arg Arg Ala Leu Glu Leu Glu Gln Glu Arg Lys Arg Ala Gln Ser
370 375 380
Glu Ala Glu Lys Leu Ala Lys Glu Arg Gln Glu Ala Glu Glu Ala Lys
385 390 395 400
Glu Ala Leu Leu Gln Ala Ser Arg Asp Gln Lys Lys Thr Gln Glu Gln
405 410 415
Leu Ala Leu Glu Met Ala Glu Leu Thr Ala Arg Ile Ser Gln Leu Glu
420 425 430
Met Ala Arg Gln Lys Lys Glu Ser Glu Ala Val Glu Trp Gln Gln Lys
435 440 445
Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys
450 455 460
Leu Ser Lys Leu Arg Thr Ser Thr Ile Met Thr Asp Tyr Asn Pro Asn
465 470 475 480
Tyr Cys Phe Ala Gly Lys Thr Ser Ser Ile Ser Asp Leu Lys Glu Val
485 490 495
Pro Arg Lys Asn Ile Thr Leu Ile Arg Gly Leu Gly His Gly Ala Phe
500 505 510
Gly Glu Val Tyr Glu Gly Gln Val Ser Gly Met Pro Asn Asp Pro

515 520 525

<210> 33

<211> 426

<212> DNA

<213> Homo sapiens

<400> 33

cgagaagttg agggagaaag gcgggcccg gaacaggctg aggctgaggt ggcctccttg	60
aaccgtagga tccagctggt tgaagaagag ctggaccgtg ctcaggagcg tgcggaggtg	120
tctgaactaa aatgtggtga cctggaagaa gaactcaaga atgttactaa caatctgaaa	180
tctctggagg ctgcatctga aaagtattct gaaaaggagg acaaatatga agaagaaatt	240

aaacttctgt ctgacaaact gaaagaggct gagaccctg ctgaatttgc agagagaacg 300

gttgcaaac tggaagac aattgatgac ctggaagtgt acctccgaa gcaccaagag 360

ctgcaagcca tgcagatgga gctgcagagc cctgagtaca agctgagcaa gctccgcacc 420

ctcgac 426

<210> 34

<211> 142

<212> PRT

<213> Homo sapiens

<400> 34

Arg Glu Val Glu Gly Glu Arg Arg Ala Arg Glu Gln Ala Glu Ala Glu

1 5 10 15

Val Ala Ser Leu Asn Arg Arg Ile Gln Leu Val Glu Glu Glu Leu Asp

20 25 30

Arg Ala Gln Glu Arg Ala Glu Val Ser Glu Leu Lys Cys Gly Asp Leu

35 40 45

Glu Glu Glu Leu Lys Asn Val Thr Asn Asn Leu Lys Ser Leu Glu Ala

50 55 60

Ala Ser Glu Lys Tyr Ser Glu Lys Glu Asp Lys Tyr Glu Glu Glu Ile

65 70 75 80

Lys Leu Leu Ser Asp Lys Leu Lys Glu Ala Glu Thr Arg Ala Glu Phe

85 90 95

Ala Glu Arg Thr Val Ala Lys Leu Glu Lys Thr Ile Asp Asp Leu Glu

100 105 110

Val Tyr Leu Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu

115 120 125

Gln Ser Pro Glu Tyr Lys Leu Ser Lys Leu Arg Thr Leu Asp

130 135 140

<210> 35

<211> 308

<212> DNA

<213> Homo sapiens

<400> 35

ctggcagagt cccgttgccg agagatggat gacgagatta gactgatgga ccagaacctg 60

aagtgtctga gtgctgctga agaaaagtac tctcaaaaag aagataaata tgaggaagaa 120

atcaagattc ttactgataa actcaaggag gcagagaccc gtgctgaatt tgcagagaga 180

acggttgcaa aactggaaaa gacaattgat gacctggaag tgtaccgccg gaagcaccag 240

gagctgcaag ccatgcagat ggagctgcag agccctgagt acaagctgag caagctccgc 300

acctcgac 308

<210> 36

<211> 103

<212> PRT

<213> Homo sapiens

<400> 36

Leu Ala Glu Ser Arg Cys Arg Glu Met Asp Glu Gln Ile Arg Leu Met

1 5 10 15

Asp Gln Asn Leu Lys Cys Leu Ser Ala Ala Glu Glu Lys Tyr Ser Gln

20 25 30

Lys Glu Asp Lys Tyr Glu Glu Glu Ile Lys Ile Leu Thr Asp Lys Leu

35 40 45

Lys Glu Ala Glu Thr Arg Ala Glu Phe Ala Glu Arg Thr Val Ala Lys

50 55 60

Leu Glu Lys Thr Ile Asp Asp Leu Glu Val Tyr Arg Arg Lys His Gln

65 70 75 80

Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro Glu Tyr Lys Leu

85 90 95

Ser Lys Leu Arg Thr Ser Thr

100

<210> 37

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 37

tacaacccca actactgctt tgct 24

<210> 38

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 38

aggcactttc tcttctctt ccac 24

<210> 39

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 39

ccacacctgg gaaaggacct aaag 24

<210> 40

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 40

cctccaaata ctgacagcca cagg 24