



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0115922
(43) 공개일자 2018년10월24일

(51) 국제특허분류(Int. Cl.)
C07K 14/47 (2006.01) C12N 9/12 (2006.01)
G01N 33/574 (2006.01)
(52) CPC특허분류
C07K 14/4716 (2013.01)
C12N 9/12 (2013.01)
(21) 출원번호 10-2017-0048416
(22) 출원일자 2017년04월14일
심사청구일자 2017년04월14일

(71) 출원인
대한민국 (식품의약품안전처장)
충청북도 청주시 흥덕구 오송읍 오송생명2로 187
한림대학교 산학협력단
강원도 춘천시 한림대학길 1, 한림대학교(옥천동)
(72) 발명자
박지원
세종특별자치시 달빛로 135, 605동 1003호(상록어
울림아파트 6단지)
장대영
서울특별시 서초구 신반포로 9, 반포아파트 76동
203호
고영호
경기도 안양시 동안구 관평로170번길 15, 일송빌
딩 801호
(74) 대리인
조정환

전체 청구항 수 : 총 15 항

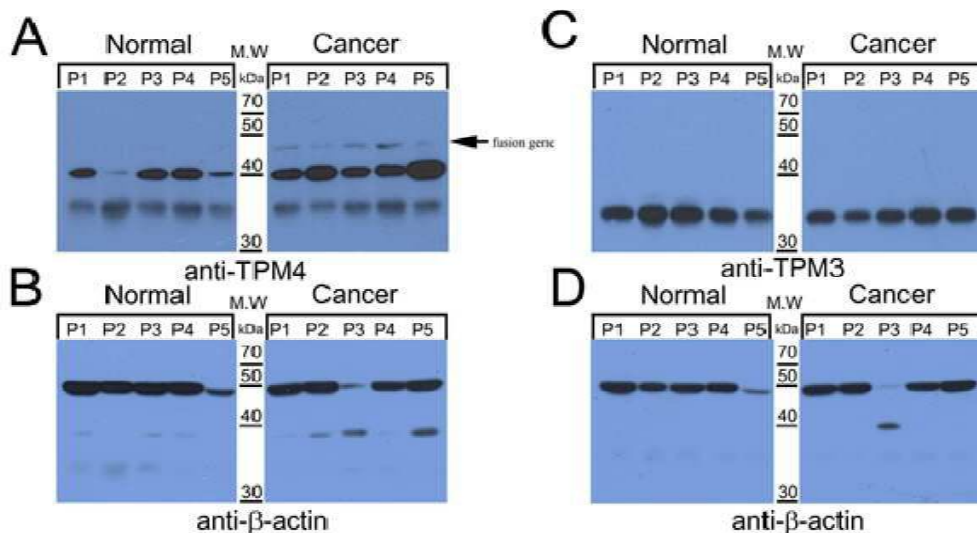
(54) 발명의 명칭 TPM4-ALK 융합단백질 및 이를 포함하는 암 진단용 조성물

(57) 요약

배경: 최근 위암의 유전체 및 전사체 연구는 위암이 환경 위험인자와 조합한 다양한 유전적 결함으로 인해 발생 하는 이중적 질병이라고 제안되었다. 본 발명에서는 정량적 무표지 프로테옴 분석을 수행하여 한국인 위암 조직 에서만 발현되는 융합단백질을 탐지하였다.

결과: 정상 위 조직에서는 발현되지 않고 위암 조직에서만 발현되는 TPM4 (Tropomyosin 4)-ALK (Anaplastic lymphoma receptor tyrosine kinase) 융합단백질을 확인하였다. TPM4 항체로 확인한 결과, 위암 조직에서만 고 분자 TPM4 밴드가 존재하였다.

대표도 - 도4



(52) CPC특허분류

G01N 33/574 (2013.01)

C07K 2319/00 (2013.01)

G01N 2333/4712 (2013.01)

G01N 2333/912 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 15182 한임평 512

부처명 식품의약품안전처

연구관리전문기관 독성평가연구부 임상연구과

연구사업명 주요사업

연구과제명 한국인 특이적 위암 원인유전자 발굴 및 표적치료제 선별 연구

기 여 율 1/2

주관기관 한림대학교 산학협력단

연구기간 2015.06.01 ~ 2017.11.30

이 발명을 지원한 국가연구개발사업

과제고유번호 HURF-2015-38

부처명 한림대학교

연구관리전문기관 의학과

연구사업명 의료원임상연구과제

연구과제명 위장관 선암 연관 융합유전자들의 병인기전과 진단 및 치료연구(융합유전자병인기전연구, 진단방법개발, 융합단백질활성조절물질발굴)

기 여 율 1/2

주관기관 한림대학교 산학협력단

연구기간 2015.05.01 ~ 2016.04.30

명세서

청구범위

청구항 1

N-말단에 TPM4 (Tropomyosin 4) 단백질 또는 그의 단편과 C-말단에 ALK (Anaplastic lymphoma receptor tyrosine kinase) 단백질 또는 그의 단편이 융합된 암 진단용 TPM4-ALK 융합단백질.

청구항 2

제 1 항에 있어서,

상기 TPM4-ALK 융합단백질은 서열번호 1의 아미노산 서열을 갖는 것을 특징으로 하는 암 진단용 TPM4-ALK 융합단백질.

청구항 3

제 1 항 또는 제 2 항의 융합단백질을 암호화하는 암 진단용 TPM4-ALK 융합유전자.

청구항 4

제 3 항에 있어서, 상기 융합유전자는 서열번호 2의 뉴클레오타이드 서열을 갖는 것임을 특징으로 하는 암 진단용 TPM4-ALK 융합유전자.

청구항 5

제 1 항에 있어서,

상기 암은 폐암, 간암, 대장암, 췌장암, 위암, 유방암, 난소암, 신장암, 갑성선암, 식도암, 전립선암 및 뇌암 중에서 선택된 어느 하나인 것을 특징으로 하는 암 진단용 TPM4-ALK 융합단백질.

청구항 6

시험대상 환자의 암을 진단하기 위한 조성물로서, 상기 조성물은 제 1 항 또는 제 2항의 융합단백질에 특이적으로 결합하는 분자 및 상기 융합단백질을 암호화하는 융합유전자와 혼성화 가능한 폴리뉴클레오타이드로 이루어진 군에서 선택된 분자를 1종 이상 포함하는 암 진단용 조성물.

청구항 7

제 6 항에 있어서,

상기 융합단백질에 특이적으로 결합하는 분자는 항체 및 압타머로 이루어진 군에서 선택된 하나 이상인 것임을 특징으로 하는 암 진단용 조성물.

청구항 8

제 6 항에 있어서,

상기 암은 폐암, 간암, 대장암, 췌장암, 위암, 유방암, 난소암, 신장암, 갑성선암, 식도암, 전립선암 및 뇌암 중에서 선택된 어느 하나인 것을 특징으로 하는 암 진단용 조성물.

청구항 9

시험대상 환자의 암 진단정보를 제공하는 방법에 관한 것으로서, 제 1 항 또는 제 2 항의 융합단백질, 상기 융합단백질을 암호화하는 융합유전자 또는 상기 융합유전자에 상응하는 mRNA를 검출하는 단계를 포함하고, 상기 융합단백질, 융합유전자 또는 mRNA가 검출되면 상기 환자를 암 환자로 결정하는 것을 특징으로 하는 암 진단 정보를 제공하는 방법.

청구항 10

제 9 항에 있어서,

상기 암은 폐암, 간암, 대장암, 췌장암, 위암, 유방암, 난소암, 신장암, 갑상선암, 식도암, 전립선암 및 뇌암 중에서 선택된 어느 하나인 것을 특징으로 하는 암 진단 정보를 제공하는 방법.

청구항 11

제 9 항 또는 제 10 항에 있어서, 상기 검출하는 단계는 면역크로마토그래피, 면역조직화학염색, ELISA, 방사선 면역측정법, 효소 면역분석, 형광면역분석, 발광면역분석, 웨스턴블롯팅 및 FACS 중에서 선택된 어느 하나의 방법을 적용하는 것을 특징으로 하는 암 진단 정보를 제공하는 방법.

청구항 12

제 1 항 또는 제 2 항의 융합단백질의 억제제 및 상기 융합단백질을 암호화하는 폴리뉴클레오타이드 분자 억제제로 이루어진 군에서 선택된 1종 이상을 유효성분으로 포함하는 암 예방 또는 치료용 조성물.

청구항 13

제 12 항에 있어서,

상기 융합 단백질의 억제제는 항체, 압타머, 키나제 저해제 및 신호 전달 저해제로 이루어진 군에서 선택된 물질을 1종 이상 포함하는 것을 특징으로 하는 암 예방 또는 치료용 조성물.

청구항 14

제 12 항에 있어서,

상기 폴리뉴클레오타이드 분자 억제제는 siRNA, shRNA 및 압타머로 이루어진 군에서 선택된 물질을 1종 이상 포함하는 것을 특징으로 하는 암 예방 또는 치료용 조성물.

청구항 15

제 12 항에 있어서,

상기 암은 폐암, 간암, 대장암, 췌장암, 위암, 유방암, 난소암, 신장암, 갑상선암, 식도암, 전립선암 및 뇌암 중에서 선택된 어느 하나인 것을 특징으로 하는 암 예방 또는 치료용 조성물.

발명의 설명

기술 분야

[0001] 본 발명은 TPM4-ALK 융합단백질 및 이를 포함하는 암 진단용 조성물에 관한 것으로서, 좀 더 자세히는 정상 위 조직에서는 발현되지 않고, 위암 조직에서만 발현되는 TPM4-ALK 융합단백질 및 이를 이용한 암 진단용 조성물에 관한 것이다.

배경 기술

[0002] 위암은 세계적으로 종양 관련 사망 중 세 번째 원인이다. 또한, 위암은 2013년도 기준으로 갑상선암 다음으로 흔한 종양이었다. 위암의 유전적 기초는 최근 광범위하게 연구되고 있으며, 위암은 동질성 질병(homogenous disease)이 아니라 수많은 유전적 돌연변이와 환경 위험인자의 조합에 의해 일어나는 질병의 복합체인 것으로 드러나고 있다. 위암에 관한 몇몇 유전자 연구는 위암 조직에서 확인된 돌연변이들이 위암 유형, 발병 연령, 인종 및 위암 부위에 따라 다양함을 제시하였다. 생식세포 돌연변이 및 체세포 돌연변이 모두를 포함하는 수많은 돌연변이는 개별 위암조직에서 다르게 나타난다. 유전자 연구와 대조적으로, 위암에 관한 프로테옴 연구는 신호 전달경로가 세포와 조직에서 발현되는 단백질에 의해 대부분 조절됨에도 불구하고 현재까지 광범위하게 수행되고 있지 못한 것이 현실이다.

[0003] 종양 프로테옴 연구의 주요 목표 중 하나는 종양 조직에서 다르게 발현되는 단백질들을 식별하여 종양 환자를 위한 진단 바이오마커를 확인하려는 것이다. 다른 종양과 비교하여 위암 진단 및 예후 바이오마커는 오늘날까지

광범위하게 정해지지 않았다. 위암에 대하여 임상적용 가능한 진단 마커는 HER2 (human epidermal growth factor receptor 2)인데, 이것은 처음에는 유방암에서 확인되었고 재발성 또는 전이성 종양환자의 약 20%에서 관찰된다.

[0004] 종양의 원인이 되는 돌연변이 중에서 염색체 전이 및 재배열에서 유래하는 융합유전자들은 수많은 유형의 종양에서 중요한 역할을 하는 유전자 변이 (driver mutations)로 잘 알려져 있다. 처음으로 연구되고 가장 잘 연구된 융합유전자는 만성 골수성 백혈병 (chronic myelogenous leukemia; CML)에서 확인되었고, BCR (breakpoint cluster region) 유전자와 ABL1 (Abelson murine leukemia viral oncogene homolog 1) 유전자의 융합을 통해 발생했다. 만성 골수성 백혈병에서 BCR-ABL1 융합유전자 확인은 타이로신 카이네이즈 저해제인 표적화 항암제 이마티니브 (imatinib) 개발을 유도했다. 융합유전자는 고형암에서도 발견되었다. 예를 들면, EML4 (echinoderm microtubule-associated protein-like 4) 유전자와 ALK (anaplastic lymphoma receptor tyrosine kinase) 유전자의 융합은 4%의 비소세포성폐암 환자에서 일어나는 것으로 보고되었다. EML4-ALK-양성 환자들은 ALK 저해제인 크리조티니브 (crizotinib) (PF02341066)를 처방했을 때 현저한 임상적 효용을 나타내었다.

[0005] 암 처방 치료의 효과에 대한 최근 연구들은 암 환자 간의 유전적 차이가 치료 효과를 결정하는 가장 중요한 인자가 될 수 있음을 제안했다. 그리하여, 암환자 치료의 최근 경향은 암환자의 유전적 결함을 확인한 후 이 돌연변이에 특이적인 표적 약물을 처방하는 것이다. 표적화된 위암 약물의 한 가지 예는 HER2에 특이적으로 결합하여 활성을 저해하는 단일클론 항체인 헤르세핀이다. 헤르세핀은 약 20%의 HER2-양성 위암환자에게 화학요법과 조합하여 처방하였을 때 훌륭한 임상 효과를 거둘 수 있다. 개발 중인 표적화 위암 치료의 다른 표적들로는 혈관 내피 성장인자 수용체 2 (vascular endothelial growth factor receptor 2), 인간 상피세포 성장인자 수용체 3 (Human Epidermal Growth Factor Receptor 3), 포스파티딜이노시톨 4, 5-이인산 3-카이네이즈 촉매 서브유닛- α , MET 및 섬유아세포 성장인자 수용체 등이다.

[0006] 완전한 맞춤의학을 위해서는 위암에 존재하는 다른 유전적 결함이 밝혀져야 한다. 본 발명에서는 정상 조직에는 존재하지 않지만 위암 조직에만 존재하는 발현이 다른 단백질들 (DEPs)과 융합단백질을 확인하기 위하여 정량적 및 정성적 무표지 프로테오믹스 분석을 이용하였다. 밝혀진 DEPs와 융합유전자들은 암의 진단과 예후 마커 및 정밀의학을 개발하는 도구의 표적으로 이용할 수 있다. 따라서 위암의 새로운 신약 타겟(druggable targets) 및 표적 치료를 개발하기 위하여 광범위한 유전적 연구가 요구된다.

[0007] 선행기술문헌

[0008] -비특허문헌

[0009] 1. McLen MH, El-Omar EM: **Genetics of gastric cancer**. *NATURE REVIEWS GASTROENTEROLOGY & HEPATOLOGY* 2014, **11**(11):664-674.

[0010] 2. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: **GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide**. In: *IARC CancerBase No 11 [Internet]*. vol. 1.0: International Agency for Research on Cancer; 2015.

[0011] 3. TCGA: **Comprehensive molecular characterization of gastric adenocarcinoma**. *Nature* 2014, **513**(7517):202-209.

[0012] 본 출원에서 임의의 참조 문헌의 언급은, 참조 문헌이 본 출원에 대한 관련 선행 기술이라는 것을 허용하는 것은 아니다.

발명의 내용

해결하려는 과제

[0013] 본 발명은 좀 더 정확하고 간편한 위암 진단과 예후 마커를 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0014] 위암에 대한 새로운 치료방법의 개발은 세계적으로 사망률을 현저히 감소시키는데 기여하고 있다. 그럼에도 불구하고, 위암은 여전히 한국과 아시아에서 가장 흔한 암 중 하나이다. 위암이 이종 유전자 결함과 관련이 있음을 고려하면, 맞춤의학의 효능을 확장하기 위하여 각 환자의 단백질 수준에서 특이한 유전적 결함을 밝혀야 한다.

- [0015] 본 발명에서는 한국인 위암 조직에서 발현이 변화하는 단백질들 (DEPs)과 융합단백질을 연구하기 위하여 무표지 프로테옴 분석을 수행하였다. 본 발명에서 밝혀진 DEPs는 여섯 개의 밀접한 관련이 있는 클러스터와 다섯 개의 상대적으로 연관된 그룹들로 나눌 수 있다. 이 다섯 개의 그룹은 보고되거나 또는 예측된 기능을 기초로 하여 수동으로 정리되었는데, 이는 이들 간의 직접적인 상호작용이 아직 명확해지지 않았기 때문이다 (도 2).
- [0016] 본 발명에서 확인된 여섯 개의 클러스터는 발암과 전이에 중요한 역할을 하는 것으로 알려진 단백질뿐만 아니라 최근 처방된 암 치료제 또는 그 유사체에 표적을 두었다. 예컨대, 다양한 β -튜블린은 종양세포에서 염색체 유사분열과 혈관신생을 저해하는 파클리탁셀 (Paclitaxel) 및 빈카 알칼로이드 (Vinca alkaloids)의 알려진 표적이며, 이 표적들은 클러스터를 형성하였다 (도 2 II). 또한, 대부분의 암에서 증가하는 열충격 분자 샤페론은 위암에서도 현저히 증가하였으며 클러스터를 형성하였다 (도 2 V). 뿐만 아니라, 소포체 및 골지체에 존재하며 단백질 접합, 변형 및 교차를 조절하는 요소들 또한 현저히 증가하였고 클러스터를 형성하였다. 이 단백질들 중 유방암과 비소세포폐암에서 항암제 저항성과 관련 있는 N-올리고당 전이효소 복합체의 한 구성요소인 RPN2는 현저히 증가하여 위암에서 약물 저항성 및 전이 기작의 보존 가능성을 제시한다. 또한, 위암에서 HSPA5, CALR 및 CANX의 현저한 증가는 위암에서 저산소증 및 낮은 영양공급을 극복하기 위하여 UPR (unfolded protein response)이 활성화될 것으로 보이는데, 이는 헬리코박터-유도 위암에서 관찰된 UPR 활성화와 유사하고, 저산소증 및 낮은 영양공급은 종양 세포에서 관찰되는 두 초기 스트레스 조건이며, 흑색종 위 모델에서 공격성을 유도한다. 또한, 본 발명자들은 PKM, PGK1, ENO1, ENO2 및 GPI와 같은 현저히 상향조절된 당 분해효소들이 속한 클러스터를 확인하였다 (도 2 IV). 종양 대사에 관한 최근의 연구들은 당 분해효소들의 상향조절은 발암 및 약제 저항성과 관련이 있고 예측 바이오마커로 이용될 수 있음을 제시한다. 이와 같은 발견을 종합하면, 본 발명에서 확인된 대부분의 DEPs는 이미 발암, 전이 및 약제 저항성과 관련이 있는 것으로 알려졌다.
- [0017] 본 발명에서 도 3과 도 5a~5n은 위암에서 확인된 TPM4-ALK 융합단백질을 나타낸다. 표적화된 항암제 사용은 몇 가지 장점이 있는데, 부작용이 적고 반응률이 높다는 점이다. 그러나, 단지 1종의 표적화된 위암 항암제인 헤르세핀만이 현재 위암 치료에 이용되고 있으며, 헤르세핀은 HER2-양성 위암 환자의 약 20%에만 효과가 있다. 따라서, 위암에서 정밀의학을 실현하기 위하여 새로운 표적 유전자 또는 단백질이 밝혀져야 한다. ALK 융합유전자가 몇몇 종양에서 중요한 역할을 담당하는 변이 (driver mutation)이며 표적화 항암제 개발의 표적임을 고려하면 본 발명에서 확인된 TPM4-ALK 융합단백질은 매우 흥미롭다. 융합단백질들의 3차원 상동성 모델링은 그들이 정상 단백질 또는 융합단백질과 이종이량체 또는 동종이량체를 형성할 수 있음을 제시한다. 이량체화 결과를 보면, 융합단백질은 정상 단백질이 위치하지 않는 곳에서 발견되었으며, 결과적으로 ALK의 비정상적 활성을 일으킨다. 이러한 TPM4-ALK 융합단백질의 비정상적 활성화는 정상 세포를 암세포로 바꿀 수 있다.
- [0018] 본 발명은 N-말단에 TPM4 (Tropomyosin 4) 단백질 또는 그의 단편과 C-말단에 ALK (Anaplastic lymphoma receptor tyrosine kinase) 단백질 또는 그의 단편이 융합된 TPM4-ALK 융합단백질을 제공한다. 또한, 본 발명은 상기 TPM4-ALK 융합단백질이 서열번호 1의 아미노산 서열을 갖는 것을 특징으로 한다. 또한, 본 발명은 상기 융합단백질을 암호화하는 TPM4-ALK 융합유전자를 제공한다. 또한, 본 발명은 상기 융합유전자가 서열번호 2의 뉴클레오타이드 서열을 갖는 것을 특징으로 한다.
- [0019] 또한, 본 발명은 상기 TPM4-ALK 융합단백질에 특이적으로 결합하는 분자 및 상기 융합단백질을 암호화하는 융합유전자에 특이적으로 결합하는 분자로 이루어진 군에서 선택된 1종 이상을 포함하는 암 진단용 조성물을 제공한다. 또한, 본 발명은 상기 융합단백질에 특이적으로 결합하는 분자가 항체 및 압타머로 이루어진 군에서 선택된 것임을 특징으로 한다. 또한, 본 발명은 상기 암이 위암인 것을 특징으로 한다. 또한, 본 발명은 시험대상 환자로부터 분리된 생물학적 시료에서 상기 융합단백질, 상기 융합단백질을 암호화하는 융합유전자 또는 상기 융합유전자에 상응하는 mRNA를 검출하는 단계를 포함하고, 상기 융합단백질, 융합유전자 또는 mRNA가 검출되면 상기 환자를 암 환자로 결정하는 것을 특징으로 하는, 암 진단 정보를 제공하는 방법에 관한 것이다. 또한, 본 발명은 상기 융합단백질의 억제제 및 상기 융합단백질을 암호화하는 폴리뉴클레오타이드 분자 억제제로 이루어진 군에서 선택된 1종 이상을 유효성분으로 포함하는 암 예방 또는 치료용 조성물에 관한 것이다. 또한, 상기 암 진단정보 제공방법에서 상기 암은 위암인 것을 특징으로 한다. 하나의 예시에서, 본 발명에 따른 융합단백질 및/또는 융합유전자는 고품암, 구체적으로 위암 환자에게서 특이적으로 발견되거나 발현되는 것으로 확인되었으므로, 상기 융합단백질 및/또는 이를 암호화하는 융합유전자는 고품암, 구체적으로 위암의 진단 마커로서 유용하다.
- [0020] 본 발명은 상기 융합단백질에 특이적으로 결합하는 분자 및/또는 상기 융합유전자와 혼성화 가능한 폴리뉴클레오타이드를 포함하는 암 진단용 조성물을 제공한다. 상기 융합단백질에 특이적으로 결합하는 분자는 항체, 압타머 등으로 이루어진 군에서 선택된 것일 수 있다. 또한, 상기 융합단백질을 암호화하는 융합유전자와 혼성화 가

능한 폴리뉴클레오타이드는 상기 융합유전자 내의 연속하는 융합 부위를 포함하는 연속하는 50 내지 250개, 구체적으로 100 내지 200개의 염기로 이루어진 DNA 분자를 증폭할 수 있도록 상기 DNA 분자의 양 말단에 인접하는 20 내지 100개, 구체적으로 25 내지 50개의 염기서열 또는 이와 상보적인 서열과 완전히 상보적이거나 80% 이상, 바람직하게는 90% 이상의 상보적인 염기서열을 가지며 상기 DNA 분자와 특이적으로 결합 가능한 폴리뉴클레오타이드를 의미한다.

[0021] 본 발명은 환자로부터 얻은 생물 시료에서 상기 융합단백질의 발현을 측정하는 단계를 포함하는 암 진단정보를 제공하는 방법을 제공한다. 상기 방법에 있어서 생물 시료에서 상기 융합단백질의 발현이 검출되면, 상기 환자는 암 (고형암), 구체적으로 위암 환자로 결정할 수 있다. 상기 융합단백질 발현 검출은 생물 시료에서 융합단백질 존재를 검출하거나, 상기 융합단백질을 암호화하는 융합유전자 또는 이에 상응하는 mRNA의 존재를 검출하여 수행할 수 있다. 예컨대, 상기 융합단백질의 존재는 이 융합단백질에 특이적으로 결합하는 분자 (예컨대, 항체 또는 압타머)를 이용하여 상기 융합단백질과 상기 분자 (예컨대, 항체 또는 압타머)와의 상호작용을 검출하는 통상적인 분석법, 예컨대, 면역크로마토그래피, 면역조직화학염색, ELISA (enzyme linked immunosorbent assay), 방사선 면역측정법 (radioimmunoassay: RIA), 효소 면역분석 (enzyme immunoassay: EIA), 형광면역분석 (Florescence immunoassay: FIA), 발광면역분석 (luminescence immunoassay: LIA), 웨스턴블롯팅 (Western blotting), FACS 등에 의하여 검출할 수 있다.

[0022] 통상적인 단백질 발현 검출 분석법으로 면역분석 방법을 수행할 수 있다. 유용한 면역분석은 동종 면역분석 또는 이종 면역분석일 수 있다. 동종 분석에서 그 면역학적 반응은 종종 융합단백질 특이적 시약, 예를 들어, 융합 폴리펩타이드 특이적 항체, 표지된 분석체 및 분석 대상 생물시료가 관련된다. 표지로부터 일어나는 신호는 표지된 분석체에 항체의 결합에 의하여 직접 또는 간접적으로 변형된다. 사용 가능한 면역화학적 표지들은 자유 라디칼, 방사성 동위원소, 형광염료, 효소, 박테리오파지, 조효소 등으로 이루어진 군에서 선택된 1종 이상일 수 있다.

[0023] 발명의 상세한 설명, 청구범위 등에 기재된 방법의 실행에 유용한 항체들은 진단 분석에 적합한 고체 지지체 (예컨대 라텍스 또는 폴리스티렌과 같은 물질로 제조된 웰, 비드, 플레이트 또는 슬라이드)에 부착될 수 있다. 항체들 또는 다른 융합단백질 결합 시약들은 공지된 기술에 따라 방사성 표지 (예컨대, ³⁵S, ¹²⁵I, ¹³¹I 등), 효소 표지 (예컨대 호스래디쉬 퍼옥시데이즈, 알칼라인 포스파테이즈 등), 및 형광 표지들과 같은 검출될 수 있는 작용기들에 마찬가지로 부착될 수 있다.

[0024] 상기 환자는 인간을 포함하는 인간, 원숭이 등의 영장류, 마우스, 래트 등의 설치류를 포함하는 포유류일 수 있다. 상기 생물 시료는 상기 환자로부터 분리된 세포, 조직, 체액 등일 수 있다.

[0025] 본 발명의 암 환자에서 특이적으로 발현하는 융합단백질은 새로운 암 치료 타겟으로서 사용될 수 있다. 따라서, 본 발명의 또 다른 예는 상기 융합단백질 억제제 및 상기 융합단백질을 암호화하는 폴리뉴클레오타이드 분자 억제제로 이루어진 군에서 선택된 1종 이상을 유효성분으로 포함하는 암 예방 및/또는 치료용 조성물을 제공한다. 상기 융합단백질의 억제제는 상기 융합단백질에 결합하여 그 기능을 상실하게 하거나 저하하는 물질로서, 상기 융합단백질에 대한 항체, 압타머, 또는 통상적인 키나제 저해제, 신호 전달 저해제 등으로 이루어진 군에서 선택된 1종 이상일 수 있다. 상기 융합단백질을 암호화하는 융합 DNA 분자 억제제는 상기 DNA 분자에 결합하여 융합단백질로 발현하지 못하도록 하는 물질로서, 상기 DNA 분자에 특이적으로 결합하는 siRNA, shRNA, 압타머 등으로 이루어진 군에서 선택된 1종 이상의 것일 수 있다.

[0026] 상기 암 진단용 조성물, 암 진단에 정보를 제공하는 방법, 및 암의 예방 및/또는 치료용 조성물의 진단 또는 치료 대상이 되는 암은 모든 종류의 고형암일 수 있다. 예컨대 상기 고형암은 폐암, 간암, 대장암, 췌장암, 위암, 유방암, 난소암, 신장암, 갑상선암, 식도암, 전립선암, 뇌암 등일 수 있으며, 구체적으로는 상기 고형암은 위암 일 수 있다.

발명의 효과

[0027] 본 발명에서는 한국인 위암에서 발현되는 융합단백질들을 확인하였으며, 확인된 TPM4 (Tropomyosin 4)-ALK (Anaplastic lymphoma receptor tyrosine kinase) 융합단백질은 위암 마커로 이용할 수 있다.

도면의 간단한 설명

[0028] 도 1은 유전자 온톨로지: 위암에서 상향조절 또는 하향조절되는 유전자를 분석한 결과이다. 위암에서 확인된 상향조절 단백질 (A) 또는 하향조절 단백질 (B)들은 다양한 분자적 기능 및 생물학적 과정과 관련되어 있다.

도 2는 정상 위 조직과 비교하여 위암 조직에서 확인된 상향조절 (적색) 또는 하향조절 (청색) 단백질을 STRING (Search Tool for Recurring Instances of Neighboring Genes) 분석한 결과이다. 상향조절 또는 하향조절된 유전자들을 그들의 주요 기능적 특성에 따라 분류하였다. 여섯 개의 밀접하게 관련된 클러스터가 확인되었다. 보고되었거나 예측된 생물학적 및 분자적 기능에 따라 부가적으로 다섯 개의 군이 더 분류되었다.

도 3은 TPM4-ALK 융합단백질의 예상 3차원 구조이며, TPM4-ALK 융합단백질의 3차원 상동관계 (homology) 모델링은 융합단백질에서 트로포마이오신 도메인을 통해 가능한 이량체화 (dimerization)를 보여준다.

도 4는 단일클론 TPM4 항체가 위암 조직 유래 단백질 추출물에서 큰 융합단백질을 인식하지만, 정상 위 조직 유래 단백질 추출물에서는 그렇지 않음을 보여준다.

A: 단일클론 항-TPM4 항체를 이용하면 두 개의 TPM4 밴드가 위암 조직에서 탐지되지만, 정상 대조군 위 조직에서는 탐지되지 않았다. 위쪽 밴드는 융합단백질로 예상되었다.

B: 단일클론 β -액틴 항체가 로딩 대조군으로 이용되었다.

C: 단일클론 TMP3 항체는 위암 조직에서 하나의 밴드만을 탐지하였다.

D: 각 레인에 로딩한 단백질의 양은 단일클론 β -액틴 항체로 확인하였다. 단일클론 β -액틴 항체를 로딩 대조군으로 이용하였다.

도 5a-5n은 무표지 프로테오믹스를 통해 확인한 27개의 융합단백질 후보를 나타낸 것이다. 도 5a-5n은 편의상 나누어 나타내었으나, 실제로는 세로로 이어진 하나의 도면이다. 커버리지 맵 (Coverage maps)은 아래 실시예의 <재료와 방법>에 설명한 바와 같은 색으로 표시하였다.

발명을 실시하기 위한 구체적인 내용

[0029] 아래에서는 구체적인 실시예를 들어 본 발명의 구성을 좀 더 자세히 설명한다. 그러나, 본 발명의 범위가 실시예의 기재에만 한정되는 것이 아님은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 자명하다.

[0031] <재료와 방법>

[0032] 임상 조직시료

[0033] 총 9쌍의 위암 조직 및 인접 정상 조직을 2011년에 한림대학교 성심병원에서 위 절제 수술을 받은, 정보를 제공하고 동의한 위암 환자로부터 수집하였다. 자세한 환자의 임상 데이터는 표 1에 요약하였다. 병리학자가 확인한 바에 따라, 괴사성 조직을 피하여 종양의 중심 부위와 인접한 정상 조직을 본 발명의 실시예에 이용하였다. 본 연구는 한림대학교 성심병원 임상시험위원회의 승인을 받았다 (승인번호: 2011-1057).

[0035] 단백질 추출

[0036] 무표지 프로테옴 분석 (label-free proteomic analysis)을 위해 위암 조직 및 인접 정상 조직으로부터 T-PER 조직 단백질 추출 키트 (Tissue protein extraction kit) (Pierce Biotechnology, Rockford, IL, USA)를 이용하여 단백질을 추출하였다. 간단히 설명하면, 20 내지 30 mg의 조직을 유리 비드와 함께 200 μ L의 T-PER 시약 (단백질 분해효소 포함) (Roche Diagnostics, Basel, Switzerland)에 넣고 간 다음, 30초씩 여섯 번 초음파처리하였다. 단백질 추출물은 15,000g로 10분간 원심분리하여 수용성 분획을 얻었다. 모든 단계는 얼음 위에서 수행하였다. 수용성 분획 내의 총 단백질은 BCA 분석을 이용하여 정량하였다. 상층액을 로딩 완충액 (Tris 40 mM pH 7.5, 2% SDS, 10% 글리세롤, 25 mM DTT)과 혼합하고, 혼합물을 95°C에서 5분간 가열한 다음, 12% SDS-PAGE 젤의 각 레인마다 30 μ g의 단백질을 로딩하였다. 그 후 젤을 잘라 각 시료 레인을 분리하고, 각 레인은 다섯 개의 조각으로 나누었다. 최종적으로 각 조각은 트립신 골드로 처리한 다음, 펩타이드를 추출하여 완전히 건조하였다.

[0038] 무표지 정량적 프로테오믹스 (Label-free quantitative proteomics)

[0039] 트립신 처리하여 얻은 펩타이드를 Synapt G1 HDMS mass spectrometer (Waters)와 짝지은 nanoAcquity UPLC

(Waters, Milford, MA)를 이용하여 세 번 분석하였다. 입자 크기 1.7 μm 의 BEH 130 C18 75 μm x 250 mm 컬럼 (Waters)을 이용하여 펩타이드를 분리하고, Symmetry C18 RP (180 μm x 20 mm, 입자 크기 5 μm)로 농축하였다. 각 실험에서 트립신 처리된 펩타이드 2 μL 씩을 이동상 A (0.1% 포름산 함유한 물)와 함께 농축 컬럼에 로딩하였다. 단계별 농도구배는 280 nL/min 속도로 적용하였으며, 55분간 5~45%의 이동상 B (0.1% 포름산 함유 아세토나이트릴)를 포함하며, 이후 10분 동안 90% 이동상 B까지 가파르게 농도를 증가하였다. 용출된 펩타이드는 데이터-독립적인 MS^E 모드를 이용하여 양성 이온화 모드에서 분석하였다. [Glu1]-피브리노펩타이드 (400 fmol/ μL)의 MS/MS 피크는 m/z 50 내지 1990 범위의 TOF (time-of-flight analyzer) 분석기를 계산하기 위하여 채용하였으며, 이중 전하 [Glu1]-피브리노펩타이드 이온 (m/z 785.8426)은 lock mass 보정을 위해 채용하였다. 데이터를 얻는 동안 모세관 전압은 3.2 kV로 맞추었고, 전원 온도는 100 $^{\circ}\text{C}$ 로 맞추었다. 저에너지 MS 모드 (온전한 펩타이드 이온)와 상승 에너지 모드의 붕괴 에너지는 각각 6 eV와 15 내지 40 eV로 맞추었다. 스캔 시간은 1.0 초로 맞추었다.

[0040] LC-MS^E 미가공 데이터 파일은 가공하였고, 단백질 확인 및 상대적 정량 분석은 ProteinLynx Global Server (PLGS 2.5.1, Waters)를 이용하여 수행하였다.

[0041] 처리계수 (processing parameter)는 전구체와 결과물 이온에 대한 자동 내성, 펩타이드당 최소 세 개의 단편 이온 매치, 단백질당 최소 일곱 개의 단편 이온 매치, 최대 4%의 FRP (false positive rate)를 갖는 단백질당 최소 두 개의 펩타이드 매치, 고정된 변형으로서 시스테인의 카바마이드메틸레이션 (+57 Da)과 다양한 변형으로서 메티오닌의 산화 (+16 Da) 및 잘못된 절단을 허용하는 하나 (one allowed missed cleavage)를 포함하였다. 단백질은 PLGS 소프트웨어 이온 계수 알고리즘 (ion accounting algorithm)으로 Unitprot website (<http://www.uniprot.org>) 상에서 호모 사피엔스 데이터베이스 (70,718 entries)를 서치하여 확인하였다.

[0042] 정량적 분석은 낮은 충돌에너지 모드에서 세 번 반복하여 측정, 관찰된 펩타이드 이온 피크 강도 측정에 기초하였으며, PLGS 2.5.1의 일부인 Waters expression을 이용하여 수행하였다. 데이터셋트는 자동 표준화 기능을 이용하여 표준화하였다. 모든 단백질은 >95%의 신뢰도로 확인하였고, 각 시료에 대하여 세 번 반복 실험에서 동일한 펩타이드는 질량 정밀도와 머무름 시간 허용오차 <0.25 min에 기초하여 PLGS 2.5.1에 포함된 클러스터링 소프트웨어를 이용하여 클러스터링하였다. 80 이상의 단백질 가능성 스코어를 가진 최소 2/3의 기술장치 반복에서 확인된 단백질들만이 정량 및 정성 분석을 위해 선택되었다.

[0043] 위암에서 융합단백질을 확인하기 위하여 본 발명자들은 Catalogue of Somatic Mutations in Cancer (COSMICv77: <http://cancer.sanger.ac.uk/cosmic>)로부터 융합단백질 데이터베이스를 이용하였다. 펩타이드와 매치되는 융합단백질의 아미노산 서열 부분은 다음에 따라 색깔로 표시하였다.

[0044] 하나의 펩타이드와 매치됨: 청색,

[0045] 펩타이드 일부와 매치됨: 적색,

[0046] 변형된 펩타이드와 매치됨: 녹색,

[0047] 부분적으로 변형된 펩타이드와 매치됨: 노란색.

[0048] 26개의 융합단백질 후보들을 매뉴얼에 따라 확인하여 펩타이드가 두 단백질의 연결점을 묶는지를 시험하였다 (Additional file 3).

[0050] 융합단백질의 3차원 구조 예측

[0051] TPM4-ALK 융합단백질의 3차원 유사 모델은 MArkovian TRAnsition of Structure evolution: Protein 3-D Structure Comparison (<http://strcomp.protein.osaka-u.ac.jp/matrix/>)을 이용하여 생성되었다. 3차원 융합단백질 모델의 시각화와 변형에는 Swiss PDB viewer 4.0.1 (Swiss Institute of Bioinformatics)을 이용하였다.

[0053] 웨스턴 블롯 분석

[0054] 정상 조직 및 암 조직 유래 단백질 추출물을 12% SDS-PAGE를 통해 분리한 다음 나이트로셀룰로스 막으로 전이하였다. 5% 탈지분유 함유 TBST에 넣어 블로킹한 다음, 마우스 항-TPM3 및 항-TPM4 항체 (Developmental Study Hybridism Bank, University of Iowa, Ames, IO, USA)를 이용하여 다섯 쌍의 위암 시료 및 정상 대조군으로부터

터 단백질을 탐지하였다. 단일클론 항 β -액틴 항체 (Sigma-Aldrich, St. Louis, MO, USA)를 로딩 대조군으로 이용하였다.

[0056] 한국인 위암 환자에서 다르게 발현되는 단백질

[0057] 위암 조직에서 다르게 발현되는 단백질을 확인하기 위하여, 아홉 쌍의 암 및 이에 매치되는 정상 위 조직을 이용하여 무표지 프로테오믹스를 수행하였다. 본 발명자들은 정상 대조군 조직과 비교하여 최소 다섯 개의 위암 조직에서 각각 상향조절되는 72개의 단백질 또는 하향조절되는 29개의 단백질을 발견하였다 (도 1).

[0058] 다르게 발현되는 단백질의 유전자 온톨로지 카테고리를 연구하기 위하여 상향조절 단백질 또는 하향조절 단백질을 Panther database (www.pantherdb.org)에 올려 생물학적 공정에 따라 카테고리화하였다. 위암 조직에서 상향조절된 72개의 단백질 중에서 42, 34, 23 및 21 단백질은 각각 대사과정, 세포과정, 위치 및 세포 구성요소 조직 또는 생합성 카테고리에 할당되었다. 이와 대조적으로, 하향조절된 29개의 단백질 중에서 13, 11, 10 및 9 단백질은 각각 다세포 조직 과정, 대사 과정, 발생 과정 및 세포 과정 카테고리에 할당되었다 (도 1).

[0059] 한국인 위암 발병 기저의 분자적 및 세포적 병인을 좀 더 연구하기 위하여 한국인 위암 조직에서 다르게 발현되는 단백질 (DEP) 간의 관계를 Search Tool for Recurring Instances of Neighboring Genes (STRING) database (www.STRING-db.org)를 이용하여 연구하였다. 도 2는 STRING 분석 결과를 나타낸다. 붉은 점은 암 조직에서 상향조절된 단백질을 가리키고, 청색 점은 하향조절된 단백질을 나타낸다. 점 사이의 색으로 표시된 선은 다양한 형태의 관계를 나타낸다 [14]. 다르게 발현되는 단백질 (DEP)은 대부분의 매우 관련이 높은 DEPs를 포함하는 여섯 개의 클러스터와 그룹 간에 관계가 없지만 알려지거나 예측한 바 유사한 기능을 가지는 다섯 개의 그룹으로 나뉘었다.

[0061] 액틴 세포골격 및 운동 활성을 조절하는 단백질 발현의 변화

[0062] 클러스터 I은 두 개의 주요 경로 즉, 액틴 세포골격 및 운동 활성 조절의 구성요소들을 포함하였다. 흥미롭게도, 액티닌 α -4 (ACTN4), 액티닌 α -1, 액티닌 α -1 골격근, 모에신 (Moesin; MSN), 빈쿨린 (Vinculin) 및 트랜스젤린 (Transgelin)과 같은 액틴-결합 단백질 및 가교제 (cross-linkers)는 위암에서 일관되게 하향조절되었다. 췌장, 난소, 폐 및 침샘암 [15]에서 ACTN4의 발현 수준과 유방암 [16]에서 MSN 발현수준은 변화하였다. 이와 대조적으로, 액틴-관련 단백질 3 유사체 (Actin-related protein 3 Homolog), IQGAP1 (IQ motif-containing GTPase activating protein 1), 아데닐레이트 사이클레이즈-관련 단백질 1 (Adenylate cyclase-associated protein 1), GDP 분리 저해제 2 (GDP dissociation inhibitor 2) 및 Rho GTPase 활성화 단백질 1과 같은 액틴 폴리머화 조절 단백질의 수준은 위암에서 현저히 높아졌다. 이 단백질 중 IQGAP1의 상향조절은 폐, 난소, 위 및 대장암의 발암과 관련이 있는 것으로 알려져 있다 [17].

[0063] MYL9 (Myosin light chain 9, regulatory), 미오신 경쇄 6, 알칼리, 평활근 및 비근육, TPM1 (Tropomyosin 1- α), 트로포마이오신 2- β 및 트로포마이오신 3 (TPM3)과 같은 운동 활성 조절 요소는 위암에서 하향조절되었다. 그러나, MYH9 (Myosin Heavy Chain 9, non-muscle)는 상향조절되었다. 이들 단백질 중 MYL9 [18], TPM1 [19] 및 MYH9 [20]는 다양한 종양의 발암에 관련되어 있는 것으로 알려져 있다. 다른 하향조절 요소들은 비멘틴 (Vimentin)과 데스민 (Desmin) 같은 중간 필라멘트 요소들을 포함하며 (도 1I), 이들은 각각 간세포암 [21] 및 대장암 [22]과 매우 밀접하게 관련되어 있다.

[0065] 현저하게 상향조절된 미세소관의 주요 구성요소

[0066] 인간에게 존재하는 다양한 튜불린 중에서 여섯 개의 α -튜불린 [Tubulin α (TUBA)-1A, TUBA-1B, TUBA-1C, TUBA-3E, TUBA-4A 및 TUBA-8], 일곱 개의 β -튜불린 [Tubulin β (TUBB), TUBB-2A, TUBB-2B, TUBB-3, TUBB-4A, TUBB-4B, TUBB-6 및 TUBB-8], 그리고 두 개의 조절 인자 [Filamin A (FLNA), TCP1 서브유닛 6A를 포함하는 새퍼론]는 위암에서 현저히 상향조절되었다 (도 2 II). β -튜불린은 항암제의 확립된 표적이고 [23], FLNA는 유방암 조직에서 상향조절된다고 보고되었다 [24].

[0068] 위암에서 발현이 변화한 이온결합 단백질들

- [0069] 발현이 변화한 단백질들 (DEPs) 중에서 헤모글로빈 (HB) 서브유닛- β , HB 서브유닛- α_2 , HB 서브유닛- δ , HB 서브유닛- ϵ_1 및 HB 서브유닛- α_1 과 같은 철- 및 산소-결합 단백질들은 위암에서 현저히 하향조절되었다. 뿐만 아니라, Ca^{2+} , Na^+ 및 K^+ 에 대한 알부민의 결합도 위암에서 하향조절되었다. 반면, 다른 철-결합 단백질, 트랜스페린 (Transferrin)과 아연-함유 탄산 탈수소효소 (Carbonic anhydrase)는 위암에서 현저히 상향조절되었다 (도 2 III).
- [0071] 위암에서 상향조절되는 당 분해대사-관련 단백질들
- [0072] 클러스터 IV의 단백질들은 다양한 대사과정에 관련되어 있다. 이 단백질들 중 글리세르알데하이드-3-인산 탈수소효소, 이놀레이즈 1 (Enolase 1; ENO1), 이놀레이즈 2 (ENO2), 글루코스-6-인산 아이소머레이즈 (GPI), PKM (Pyruvate kinase muscle) 및 PGK1 (Phosphoglycerate Kinase 1)은 당분해와 관련성이 높다. 이 유전자들은 상향조절되었으며, 서로 밀접하게 관련된다. 이 클러스터는 또한 하향조절되는 말레이트 탈수소효소 1, ATP 합성효소 H^+ 전이 미토콘드리아 F1 복합체 α -서브유닛과 심근 및 상향조절되는 미토콘드리아에서 발견된 시트르산 합성효소를 포함한다 (도 2 IV).
- [0074] 대부분의 분자 샤페론 관련 단백질은 위암에서 상향조절되었다.
- [0075] 클러스터 V (도 2V)는 두 개의 중심 분자 샤페론을 포함하는데, 열충격 단백질 (HSP) 90 kDa 세포질- α 클래스 A 멤버 1과 HSP90 α -패밀리 클래스 b 멤버 1이며, 이는 HSP 패밀리 A 멤버 1 유사, HSP 패밀리 A 멤버 2, HSP 패밀리 A 멤버 5, HSP 패밀리 A 멤버 6, HSP 패밀리 A 멤버 8, HSP 패밀리 A 멤버 9, HSP 패밀리 D 멤버 1, 저산소증 상향조절 1 (Hypoxia up-regulated 1) 및 HSP B 멤버 1 (HSPB1)과 강하게 반응한다. HSPB1을 제외하고, 모든 열충격 단백질들은 위암에서 상향조절되었다.
- [0077] 단백질 접힘과 교역 활성을 갖는 상향조절된 단백질들
- [0078] 클러스터 VI (도 2VI)은 골지체 (GA) 또는 소포체 (ER)에 위치한 상향조절된 단백질들을 포함하며, 단백질 접힘과 교역을 조절하였다. 발로신 (Valosin)-함유 단백질, 주요 조직적합성 복합체 클래스 1 (Major histocompatibility complex class 1) (HLA)-B, HLA-C, RPN2 (Ribophorin II), PDIA (Protein disulfide isomerase family A) 멤버-3, PDIA 멤버-6, 프롤릴 4-하이드록실레이즈 서브유닛- β (Prolyl 4-hydroxylase subunit- β), CANX (Calnexin), CALR (Calreticulin), 티오레독신 도메인 함유 5 (Thioredoxin domain containing 5) 및 글루코시데이즈 II α -서브유닛 (Glucosidase II α -subunit)은 현저히 상향조절된 단백질들이다.
- [0080] 단백질 합성에 관련된 상향조절된 단백질들
- [0081] 단백질 합성에 관여하는 세 개의 단백질: 진핵생물 번역연장인자 (EEF)-2, EEF-1A1 및 EEF-1A2가 상향조절되었고, 다른 클러스터의 요소들과 상호작용하였다. 이 단백질들 중 EEF1A2는 난소암 발암과정과 관련이 있다 [25].
- [0083] 위암에서 하향조절된 원암 유전자 (proto-oncogene)와 상향조절된 단백질 분해효소 저해제
- [0084] 전방 구배 2 (Anterior gradient 2) [26, 27], 전방 구배 3 [27], Ras 억제제-1 [28], SET 핵 원암 유전자 [29], 트립토판- $tRNA$ 합성효소 [30] 및 유비퀴틴-유사 변형자 활성화 효소 1 [31]과 같은 원암 유전자들은 위암에서 현저히 하향조절되었다 (도 2B). 또한, 위암 [32]과 대장암 [33]에서 종양 진행과 관련된 세르핀 펩티데이즈 저해제 클레이드 (SERPIN)-A 멤버 1 (SERPINA1) 및 초기 단계 간세포종양 [34]의 강력한 바이오마커로 확인된 SERPIN-H 멤버 1은 본 발명에서 위암에서 현저히 증가하는 것으로 밝혀졌다.
- [0086] 에너지 대사 및 세포 구조인자와 관련된 발현 변화 단백질들

- [0087] 크레아틴 카이네이즈 B는 조직에서 에너지 항상성을 조절하고 자궁경부암에서 감소하는데 [35], 위암에서는 하향조절되었다. 이와 대조적으로, 에너지 항상성을 유지하는 ATPase Na^+/K^+ -전이 서브유닛 α -1, 알데하이드를 산화하여 카복실산을 생성하는 미토콘드리아 알데하이드 탈수소효소 2 패밀리를, 그리고 음식의 트리글리세라이드 소화에 관여하는 효소인 위형 라이페이스 F (Gastric type Lipase F)는 위암에서 현저히 상향조절되었다.
- [0088] 세포 구조 조직화에 중요한 여러 개 도메인을 갖는 다른 단백질들 또한 발현 변화를 나타내었다. 예컨대, POTE 앵키린 도메인 패밀리의 멤버 (Ankyrin domain family member) (POTE)-J 및 POTE-I는 위암에서 각각 하향조절 및 상향조절되었다. 콜라겐 원섬유 조직화를 조절하고 전립선암 [36]에 관여하는 류신이 풍부한 작은 프로테오글라이칸 패밀리에 속하는 루미칸 (Lumican)은 위암에서 현저히 하향조절되었다. 주요 볼트 단백질 (Major vault protein)은 약물저항성 암에서 매우 과발현된다 [37]. 카리오페린 서브유닛 β -1, 정선 플라코글로빈 (Junction Plakoglobin), 라민 (Lamin) A/C, 다중 PDZ 도메인 크럼프 세포 극성 복합체 인자 (Multiple PDZ domain crumbs cell polarity complex component), 클래스린 중쇄 (Clathrin heavy chain) 및 류신-풍부 펜타트리고펩타이드 반복-함유 (Leucine-rich pentatricopeptide repeat-containing)는 상향조절되었다.
- [0090] **무표지 프로테옴 분석으로 TPM4-ALK 융합단백질을 확인하였다.**
- [0091] 본 발명자들은 COSMIC 융합유전자 데이터베이스를 이용하여 아홉 쌍의 위암 조직 및 정상 조직으로부터 얻은 MS 데이터를 연구하여 27개의 융합단백질 후보들을 얻었다 (도 5a-5n). 본 발명자들은 질량분석기로 펩타이드가 두 단백질의 연결점을 묶는지를 매뉴얼에 따라 시험하였다. TPM4-ALK 융합단백질만이 두 단백질을 묶는 해당 펩타이드를 가지고 있음을 확인하였다.
- [0092] 발암과정에서 위 TPM4-ALK 융합단백질의 가능한 역할을 좀 더 연구하기 위하여, TPM4-ALK 융합단백질의 3차원 구조를 상동관계 모델링 (homology modeling)을 수행함으로써 예측하였다. TPM4-ALK 융합단백질의 3차원 구조는 이것이 하나의 ALK 카이네이즈 도메인과 하나의 TPM이 우선적으로 하나의 알파 나선을 이루고 정상 TPM4 또는 다른 TPM4-ALK 융합단백질과 병렬 이량체 이중 나선 (parallel dimeric coiled-coils)으로 조립될 수 있음을 보여주었다 (도 3). 융합단백질이 이량체화되는 능력은 융합단백질의 위치가 정상 단백질의 위치와 다를 수 있고, 그리하여 세포 과정 및 분자 과정을 변화시킬 수 있음을 암시한다.
- [0094] **위암 조직에서 크기가 커진 밴드는 단일클론 TPM4 항체로 인식되었다.**
- [0095] TPM4-ALK 융합단백질의 존재를 확인하기 위하여 본 발명자들은 TPM3 또는 TPM4에 특이적인 단일클론항체를 이용하여 웨스턴 블롯 분석을 수행하였다. 항-TPM4 항체는 정상 위 조직에서는 탐지되지 않는 크기가 커진 밴드를 인식하였다. 그러나, 항-TPM3 항체는 부가된 밴드를 탐지하지 못하였다. 위암 조직으로부터 단지 하나의 밴드만이 정상 위 조직에서 관찰된 것과 동일한 이동 패턴을 보여주었다 (도 4).
- [0097] 표 1은 아홉 명의 위암 환자의 임상 및 병리학적 데이터이다.

표 1

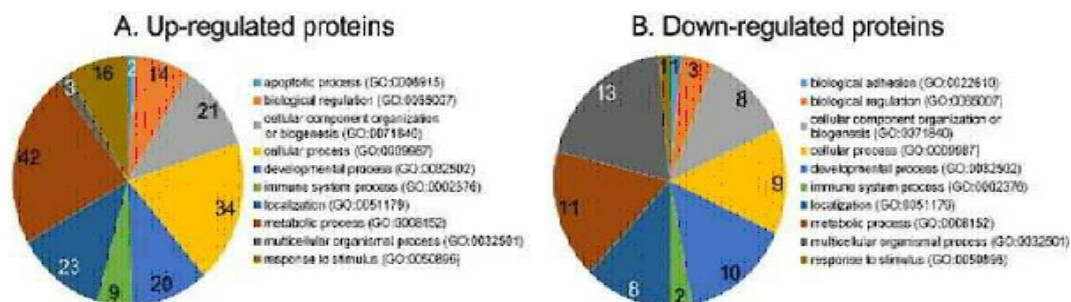
No	Sex	Age	Stage	Tumor location	Differentiation	Lauren classification
1	F	82	IIB	Antrum	Poorly differentiated	Intestinal
2	M	57	IIA	Antrum	Moderately differentiated	Intestinal
3	M	56	IIIB	Antrum	Poorly differentiated	Mixed
4	M	79	IIIB	Antrum	Poorly differentiated	Mixed
5	M	78	IIIB	Antrum	Signet ring cell	Diffuse
6	M	78	IIA	Body	Moderately differentiated	Intestinal
7	F	74	IIIC	Body	Signet ring cell	Diffuse
8	F	78	IIIC	Body	Poorly differentiated	Diffuse
9	M	59	IIIA	Body	Poorly differentiated	Mixed

[0098]

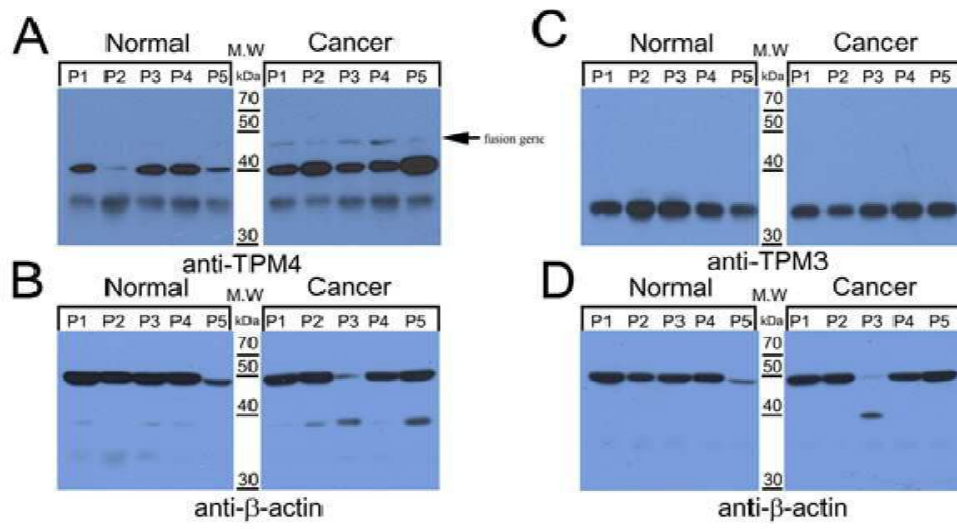
도면

도면1

Gene Ontology: Biological Process



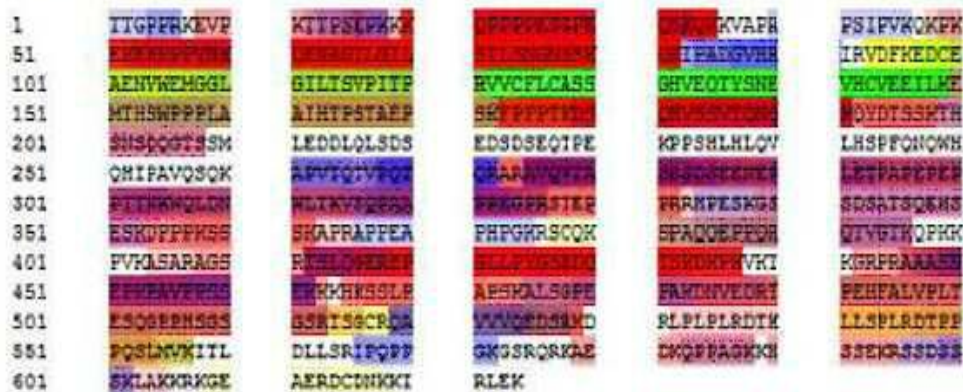
도면4



도면5a

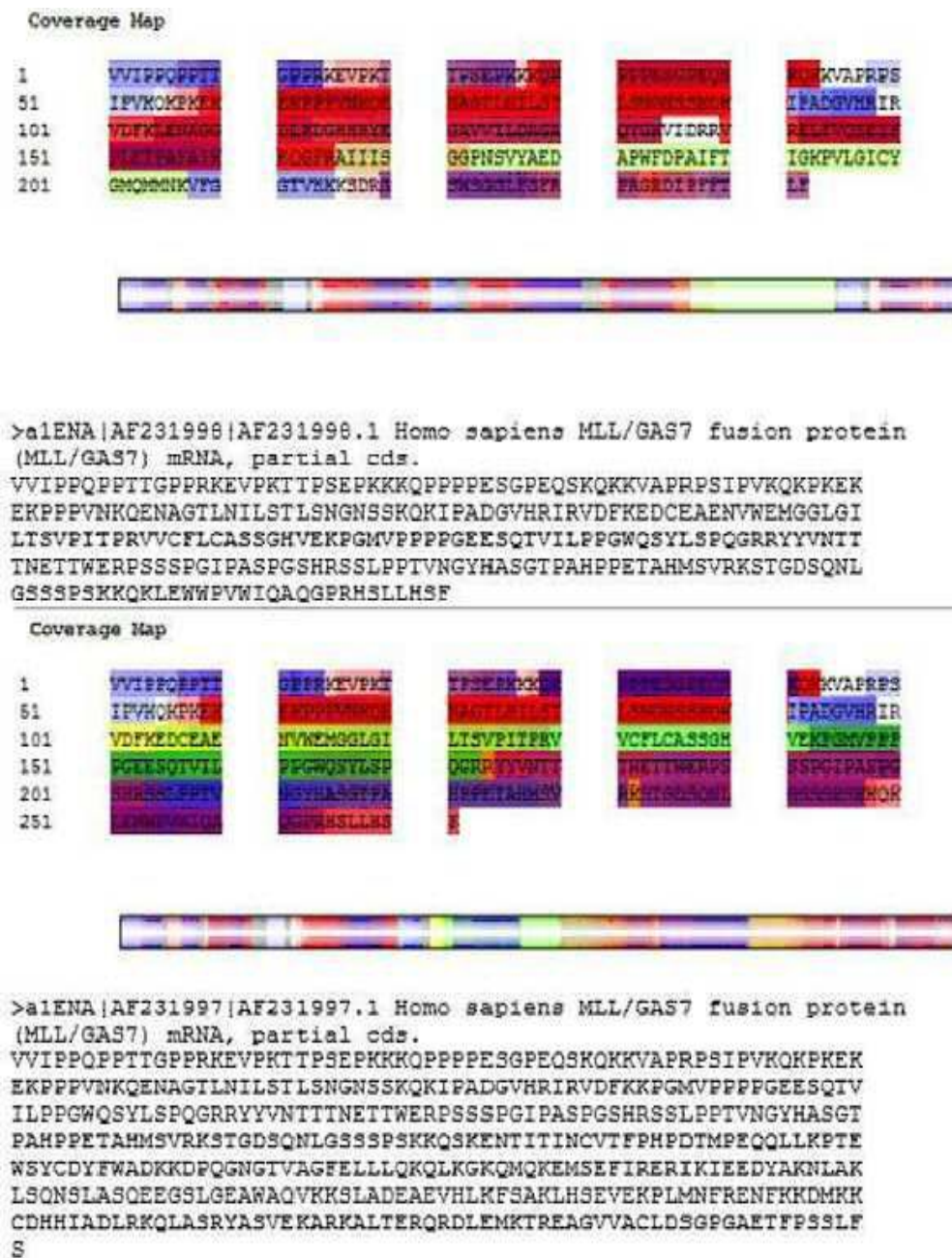
>AF487905|AF487905.1 Homo sapiens MLL/AF4 fusion protein mRNA, partial cds.
 TTGPPRKEVPKTTTPSEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKKEKPPPVNK
 QENAGTLNILSTLSNGNSSKQKIPADGVHRIRVDFKEDCEAENVWEMGGILGILTSVPITP
 RVVCFCLCASSGHVEQTYSNVHCVEEILKEMTHSWPPPLAAIHTPSTAEPKFPFPFKDS
 QHVSSVTQNKQYDTSSKTHSNSQQTSSMLEDDLQLSDSESDSEQTPEKPPSHLHLQV
 LHSPTQNWQHQPAPVQSQKAPVTQTVPQTQRARAVQVTASSSDSENEPLETPAPEPEP
 PTINKWQLDNWLIKVSQPAAPPEGPRSTEPFRRHPEESKGSSDSATISQEHSESKDPPPKSS
 SKAPRAPPEAPHPGKRSCOKSPAQQEPPQRQTIVGKQPKKFKVKSARAGSRTISLQGEREP
 GLLPYGSRDQTSKDKPKVKTGRPRAAASNEPKPAVFPSSSEKKKKHSSLPAPSKALSGPE
 PAKDNVEDRTPEHFALVPLTESQGPPHSGSGSRTSGCRQAVVVQEDSRKDRPLPLPLRDTK
 LLSPLRDTFPPQSLMVKITL DLLSRIPQPPGKGSRQRKAEDKQPPAGKKHSSSEKRSSDSS
 SKLAKKRKGEAERDCDNKKIRLEK

Coverage Map

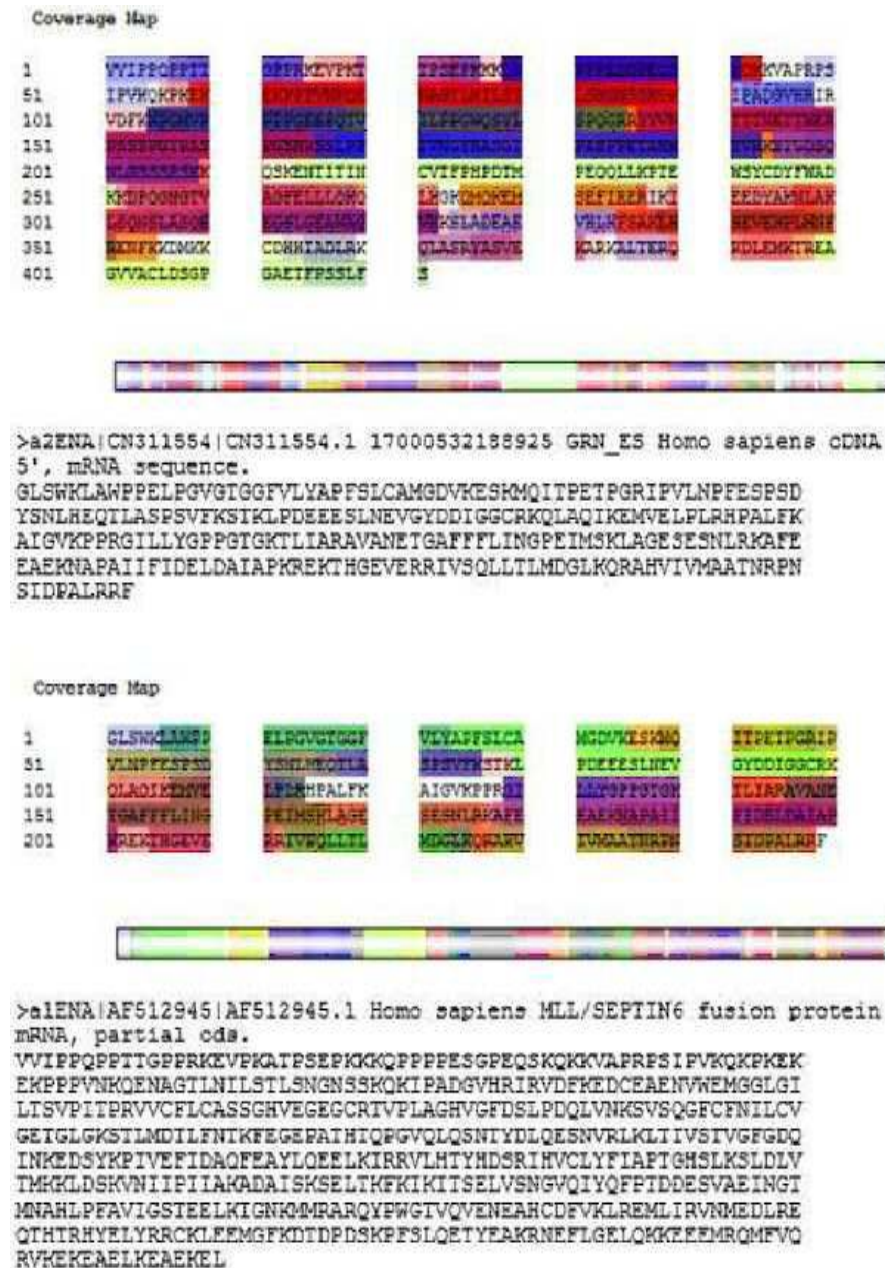


>AF297746|AF297746.1 Homo sapiens clone 1426-11 MLL/GMPS fusion protein mRNA, partial cds.
 VVIPPOFPPTTGPPEKEVPKTTTPSEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKKE
 EKPPPVNKQENAGTLNILSTLSNGNSSKQKIPADGVHRIRVDFKLENAGGDLKDGHHHHE
 GAVVILDAGAQQYGVKIDRRVRELFVQSEIFFLETFAFAIKEQGFRATIIISGGPNSVYAE
 APWFDPATFTTIGKPVLGICYGMQMMNKVFGGTVHHKSDRGSWSGGGLFGFRPRGRDIPFT
 LF

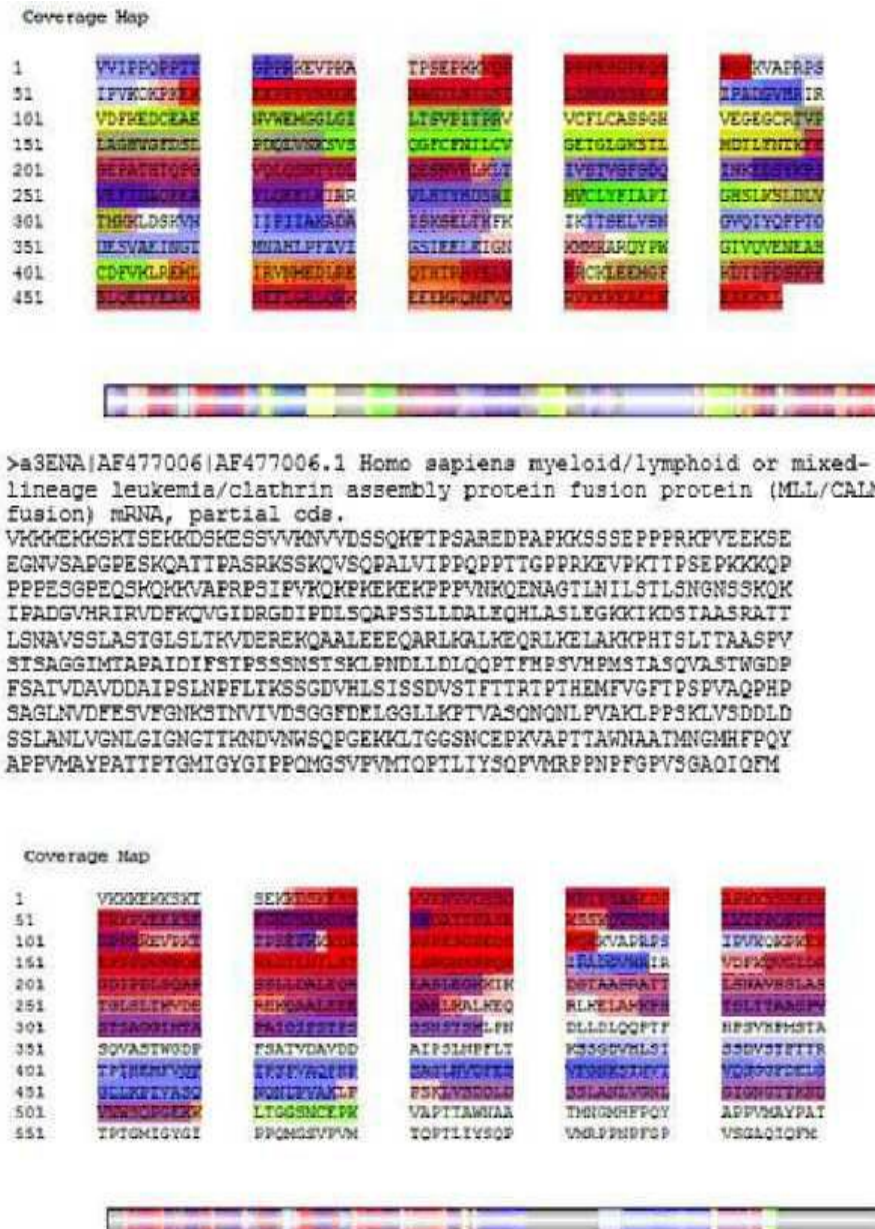
도면5b



도면5c



도면5d

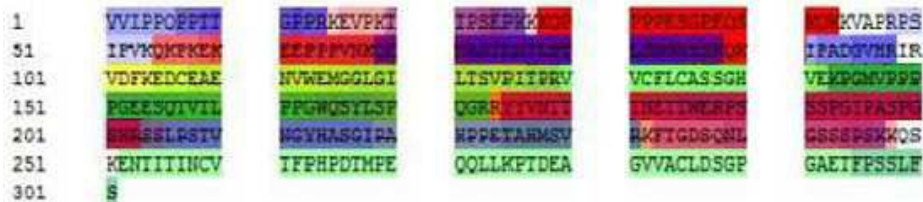


도면5e

>AF231999|AF231999.1 Homo sapiens MLL/GAS7 fusion protein (MLL/GAS7) mRNA, partial cds.

VVIPQPPTTGPPRKEVPKTIPTSEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKEK
EEPPFVNKQENAGTLNILSTLSNGNSSKQKIPADGVHRIRVDFKEDCEAENVWEMGGGLGI
LTSVPITPRVVCFLCASSGHVEKPGMVPPPPGEESSQTVILPPGWQSYLSPQGRYYVNTI
TNETTWERPSSSPGIPASPGSHRSSLPSTVNGYHASGIPAHPPETAHMSVRKFTGDSQNL
GSSSPSKKQSKENTITINCVTFFPHPDTPMEQQLLKPTDEAGVVACLDGSGPGAETFPSSLF
S

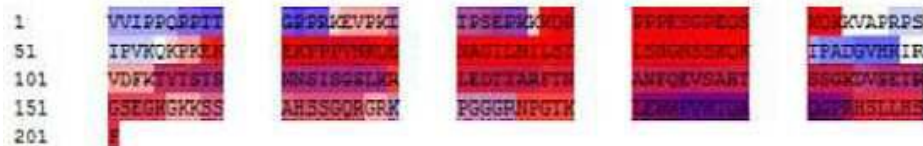
Coverage Map



>AF272375|AF272375.1 Homo sapiens clone 1370-44 MLL/AF-10 fusion protein mRNA, partial cds.

VVIPQPPTTGPPRKEVPKTIPTSEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKEK
EKPPFVNKQENAGTLNILSTLSNGNSSKQKIPADGVHRIRVDFKTYTSTSNNSISGSLKR
LEDITAREFTNANFQEVSAHTSSGKDVSEIRGSEGGKGGKSSAHSSGQRGRKPGGGGRNPGTK
LEWWPVWIAQGGPRHSLHSEF

Coverage Map



도면5f

>AF512943|AF512943.1 Homo sapiens clone 1838-11 MLL/SEPTIN6 fusion protein mRNA, partial cds.
 WVIPPQPTTIGPFRKEVFKATPSEPKKKQPPPPESGPEQSKQKKVAFRPSIPVKQKPKKEK
 EKPPPVNKKQENAGTILNLTSLNNGNSSKQKIPADGVHRIKRVDFKEDCEAENVWEMGGLGI
 LITVPITPRVVCFLCASSGHVEGECRTVFLAGHVGFDSLPDQLVNKSVSQGFCFNILCV
 GETGLGKSTLMDTILFNKFEGETPATHTQPGVQLQSNITYDLQESNVRLKLIIVSTVGFSDQ
 INKEDSYKPIVEFIDAQFEAYLQEEELKIRRVLHTYHDSRIHVCLYFIAPTGHSLKSLDLV
 TMKGLDSKVNIIPIIAKADAISKSELTIKFKIKITSELVSNVQIYQFPIDDESVAEINGT
 MNAHLPPFAVIGSTEELKIGNKMMRARQYFPGTVQVENEAHCDFVKLRREMLIRVNMEDLRE
 QTHTRHYELRRCKLEEMGFKDIDPDSKPFSLQETYEAKRNEFLGELQKK

Coverage Map



>AF297749|AF297749.1 Homo sapiens MLL/GMP5 fusion protein mRNA, partial cds.
 RPSIPVKQKPKKEKPPFVNKKQENAGTILNLTSLNNGNSSKQKIPADGVHRIKRVDFKEDC
 EAENVWEMGGLGILITVPEITPRVVCFLCASSGHVELENAGGDLKDGHVHVEGAVVILDAG
 AQYGVKVIDRRVRELFVQSEIFPLETPAFATKEQGFRA

Coverage Map



도면5g

>AF186110|AF186110.1 Homo sapiens TPM4-ALK fusion oncoprotein type 1 (TPM4-ALK fusion) mRNA, partial cds.

LSARRIEGWNSNLAFFPRFILAEGDVRAALNRRIQLVEEELDRAQERLATALQRLZEAEKA
ADESERGMKVIEWRAMKDEEFMEIQEMQLKARKHIAEEADRKYEEVARKLVILEGELERA
EERAEVSELRCSGLEEDLNVTNNLKSLEAASENYSEKSDKYEESEIKLLSDKLEAEATRA
ETAEATVAKLENTIDDLVYRRRNQELQAMQMELOSPYKLSKLRSTSTIMTDYNNPYCTA
GKTSISISDLREVFARNITLIRGLGHGAFGEVYEGQVSGMFNDPSPQLQVAVRTLP

Coverage Map



>AM491362|AM491362.1 Homo sapiens partial mRNA for bcr-abl1 e6a2 chimera protein

YQPYQSIYVGGMMEGEGKGPLLRSCQSTSECEKRLTWPRRSYSPRSYEDCGGGYTPDCSEN
ENLTSSEEDFSGQSSSRVSPFTTYRMFRDKSRSPSONSOOSFDS3SPFTPOCHRRHRHC
FVVVSEATIVGVKRTGQIWFNDGEGAFHGDADGSGTGTFPGYGCAADRAECCRRMQDGLPY
IDDSFSSSFHLSSFRGRSDALVSGALESTKASELOLEKGLMKRVLSGILASEETVLS
HLEALLPMKPLKAAATTSQFVLTSQQIETIFFEVFELYEIMKLFYDGLFPRVQQW3HQ
RVGDLFQKLASQLGVYRASVDNYGVAMEMMAEMCCQANAQFAEISENLRRASNKDAKDETT
KNSLEKALQRFPVADDFCPQGLSZAARNW3KEDLLAGPSENDPNLFVALYDTVA3GDNTLS
ITKGEKLRVLGYNNHNGEWCEAQTKNGQGWVPSNYITPVNSLERH3WYHGPVSRNAAEYLL
SSGINGSFLVRESESSPGQRSISLRYEGRVYHYRINTASDGKLYV3SESRFNTLAEVLVH
HSTVADGLITTLHYFAPKRNKFTVYGV3PNYDKMEMERTDITMKM

Coverage Map



도면5h

>AM491363|AM491363.1 Homo sapiens partial mRNA for bcr-abl1 e19a2 chimeric protein
 PSMAFRVHSRNGKSYTFLISSDYERAERENIREQOKKCFRSFSLSLTSVELQMPNTNSCVKL
 QTVHSIPLTINKEDDESPGLYGFLNVIHSAIGFKQSSNLYCTLEVDSFGYFVNKAKTRV
 YRDTAEFNWNEEFIELEGSQILRILCYEKCYNKTKIPKEDGESIDRLMGKGQVQLDPQA
 LQDRDWQRTVIAMNGIEVKLSVKFNSREFSLKRMPSRKOTGVLGVKIADVTKRERSKVPPY
 IVRQCVEEIERRGHEEVGIYRVSGVATDIOALKAADFVKALQRPVADDFEPQGLSEAAARW
 NSKENLLAGPSENDPNLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEWCEAQTNGQ
 GWVPSNYITFPVNSLEKHSWYHGFVSRNAEHLSSGINGSFLVRESESSPGQRSISLRYE
 GRVYHYRINTASDGKLYVSSSESRNTLAELVHHHSTVADGLITTLHYPAPKRNKPSVYGV
 SPNYDKWEMERIDIIMKH

Coverage Map



>AJ131466|AJ131466.1 Homo sapiens partial mRNA for bcr-abl1 e14a2 chimeric protein
 ANKGSKATERLKKKLSEQESLLLLMSPSMAFRVHSRNGKSYTFLISSDYERAERENIRE
 QOKKCFRSFSLSLTSVELQMLTNSCVKLQTVHSIPLTINKEDDESPGLYGFLNVIHSAIGF
 KQSSKALQRPVADDFEPQGLSEAAARWNSKENLLAGPSENDPNLFVALYDFVASGDNTLSI
 TKGEKLRVLGYNHNGEWCEAQTNGQGWVPSNYITFPVNSLEKHSWYHGFVSRNAEYLLS
 SGINGSFLVRESESSPGQRSISLRYEGRVYHYRINTASDGKLYVSSSESRNTLAELVHHH
 STVADGLITTLHYPAPKRNKPTVYGVSPNYDK

Coverage Map



도면5i

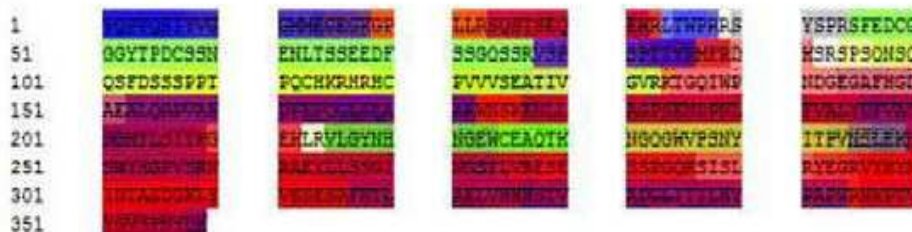
>a1ENA|AJ131467|AJ131467.1 Homo sapiens partial mRNA for bcr-abl1 e13a2 chimeric protein
 ANKGSKATERLKKKLSEQESLLLMSPSMAFRVHSRNGKSYTFLISSDYERAEWRENIRE
 QQKKCFRSFSLTSVELQMLTNSCVKLQTVHSIPLTINKEEALQRPVASFEPQGLSEAR
 WNSKENLLAGPSENDPNLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEWCEAQTKNG
 QGWVPSNYITPVNSLEKHSWYHGFVSRNAAEYLLSSGINGSFLVRESESSPGORSISLRY
 EGRVYHYRINTASDGKLYVSSESFRNTLAEVHHHSTVADGLITTLHYPAFKRNKPTVYG
 VSPNYDK

Coverage Map



>a2ENA|AF113911|AF113911.1 Homo sapiens BCR-ABL1 e1a2 chimeric protein (BCR/ABL fusion) mRNA, partial cds.
 YQPYQSIYVGGMMEGEGKGPLLRSQSTSEQEKRLTWPRRSYSPRSFEDCGGGYTPDCSSN
 ENLTSSEEDFSSGQSSRVSPSPITTYRMFRDKSRSPSQNSQQSFDSSSPPTFQCHKRHRHC
 FVVVSEATIVGVKRTGQIWPNDGEGAFHGDAEALQRPVASFEPQGLSEARWNSKENLL
 AGPSENDPNLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEWCEAQTKNGQGWVPSNY
 ITFVNSLEKHSWYHGFVSRNAAEYLLSSGINGSFLVRESESSPGORSISLRYEGRVYHYR
 INTASDGKLYVSSESFRNTLAEVHHHSTVADGLITTLHYPAFKRNKPTVYGVSPNYDK

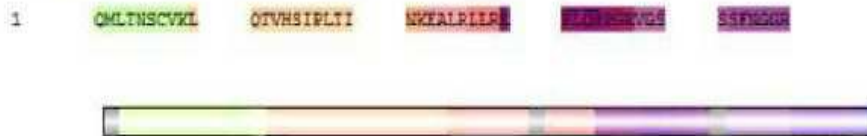
Coverage Map



도면5j

>alENA/DQ898314|DQ898314.1 Homo sapiens isolate el3a4 BCR/ABL fusion protein (BCR/ABL fusion) mRNA, partial cds, alternatively spliced.
QMLTNSCVKLQTVHSIPLTIKALRLIREPLQHPGRVGSSEFNNGGR

Coverage Map



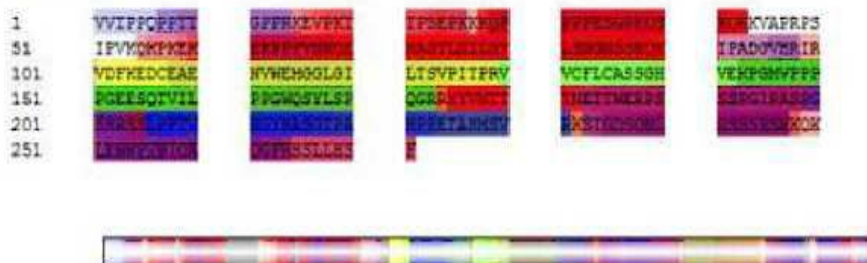
>DQ451148|DQ451148.1 Homo sapiens MLL-AF4 fusion protein (MLL/AF4 fusion) mRNA, partial cds.
TTPSEPKGKQPPFPESGPEQSKQKGVAPRPSIPVKQKPKKDSQHVSSVTQNRQYDTSS
KTHSNSQQGTS

Coverage Map



>alENA/AF231996|AF231996.1 Homo sapiens MLL/GAS7 fusion protein (MLL/GAS7) mRNA, partial cds.
VVIFPQPPTTGPPRKEVEKTIIPSEPKKKQPPFPESGPEQSKQKGVAPRPSIPVKQKPKK
EKPPFPVKNQENAGTLNLTSLNGNSSKQKIPADGVHRIRVDFKKPGMVFPFPGEESQTV
ILPPGWQSYLSPQGRYYVNTTNETTWERPSSSPGIPASPGSHRSSLPPTVNGYHASGT
PAHPPETAHMSVRKSIGYSQNLGSSSPSKKQKLEW*PVNIQAQGRHSLIHSE

Coverage Map



도면5k

>a3ENA|AF031404|AF031404.1 Homo sapiens MLL-AF4 fusion protein mRNA, partial cds.
 KQKKVAPRPSIPVKQKPKKEKPPPVNKQENAGTLNILSTLSNGNSSKQKIPADGVHRIR
 VDFKEDCEAENVWEMGGGLGILTSVPITPRVVCFLCASSGHVEQTYSNVHCVVEEILKEMT
 HSWPPELTAIHTPSTAEPKFPFPFTKDSQHVSSVTQNKQYDTSSKTHSNSQOGTSSMLE
 DDLQLSDSEDSDSEQTPEKPPSSSAPPSAPQSLPEPVASAHSSSSAESESTSDSDSSSDSE
 SESSSS

Coverage Map

1	KQKKVAPRPS	IPVKQKPKK	KPPPVNKQ	AGTLNILST	SSKQKIPAD
51	IPADGVHRIR	VDFKEDCEAE	NWEMGGGLG	LTSVPITPRV	VCFLCASSGH
101	VEQTYSNVH	CVVEEILKEM	HSWPPELTA	HTPSTAEPK	F
151	FPFTKDSQHV	SSVTQNKQY	DTSSKTHSN	SQOGTSSMLE	DDLQLSDSED
201	SDSEQTPEK	PPSSSAPPS	APQSLPEPV	ASAHSSSSAE	SESTSDSDSS
	SDSE	SESSSS			



>a1ENA|AM491360|AM491360.1 Homo sapiens partial mRNA for bcr-abl1
 e14a3 chimeric protein
 PSMAFRVHSRNGKSYTFLISSDYERAEWRENIREQQKKCFRFSFLTSVELQMLTNSCVKL
 QTVHSIPLTIINKEDDESPGLYGFNLNVIHVSATGFKQSSSEKLRVLGYNHNGEWCEAQTKN
 GQGWVPSNYITPVNSLEKHSWYHGFVSRNAEYLLSSGINGSFLVRESESSPGQRSISLR
 YEGRVYHYRINTASDGKLYVSSSESRFNTLAELVHHHSTIVADGLITTLHYPPAPKRNKPTVY
 GVSPNYDKWEMERTDITMKH

Coverage Map

1	PSMAFRVHSR	NGKSYTFLI	SSDYERAEW	RENIREQQK	CFRFSFLTS
51	VELQMLTNS	CVKLQTVH	SIPLTIINK	EDDESPGLY	GFNLNVIH
101	VSATGFKQS	SSSEKLRVL	GYNHNGEW	CEAQTKNG	QGWVPSNY
151	ITPVNSLEK	HSWYHGFV	SRNAEYLL	SSGINGSF	LVRESESS
201	PGQRSISLR	YEGRVYHY	RINTASDG	KLYVSSSE	SRFNTLAEL
251	VHHHSTIV	ADGLITTL	HYPPAPKR	NKPTVYGV	SPNYDKWE
	MERTDITM	KH			



도면51

>alENA|AY187921|AY187921.1 Homo sapiens MLL/ENL fusion mRNA, partial sequence.

EDPAFKKSSSEPPPRKPVEERKSEEGNVSAAPGPESKQATTPASRRSSKQVSOBALVIPPQP
PTTGPPRKEVPKTTTPEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKKEKPPPVN
KOENAGTLNLTSLNGN88KQKIPADGVHRIGVDFKEDCEAENVWEMGGLGILTSVPIT
PRVVCFLCAS8SGHVECTVQVRLELGHRAQLRKKFTTEGFTHDWMVFVRGPEQ

Coverage Map



>alENA|AF422798|AF422798.1 Homo sapiens MLL/LAF4 fusion protein mRNA, partial cds.

EDPAFKKSSSEPPPRKPVEERKSEEGNVSAAPGPESKQATTPASRRSSKQVSOBALVIPPQP
PTTGPPRKEVPKTTTPEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKKEKMTWLPP
LSAIQAPGKVEPTKFFPNKDSQLVSSGHNNPKKGDAEPESPDNG

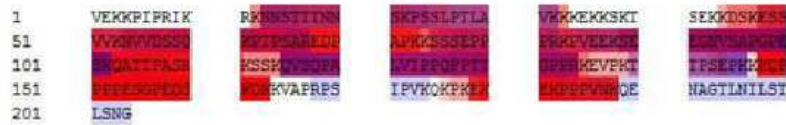
Coverage Map



도면5m

>a1ENA|FM165198|FM165198.1 Homo sapiens partial mRNA for TET1-MLL fusion transcript in acute lymphoblastic leukemia, patient 4765
VEKKPIPRIKRNNSTTTNNSKPSLPTLAVKKKKKSKTSEKKDSKSSVVKNVVDSSQ
KPTPSAREDPAPKSSSEPPPRKPVEEKSEEGNVSAAGPESKQATTPASRKSSKQVSQPA
LVIPPQPPTTGPPRKEVPKTTTPSEPKKKQPPPPESGPEQSKQKKVAPRPSIPVKQKPKK
ERFPPVNRQENAGTLNILSTLSNG

Coverage Map



>a2ENA|AF533682|AF533682.1 Homo sapiens IL1RAPL1/dystrophin fusion protein mRNA, partial cds.
DFLLPTREPEILWYKECKRTKTWRPSIVFKRDTLLIREVREDDIGNVTCELYGGGFVVRRT
TELTFTVERIQNQWDEVQEHQLNRRQQLNEMLKDSTQWLEAKEEAEQVLGQARAKLESWK
EGPYTVDAIQKKI

Coverage Map



도면5n

>DT217282|DT217282.1 KB-EST0002366 BNS6 Homo sapiens cDNA, mRNA sequence. hnrNFA2/B1/FAM96A fusion gene
VAAAPGPFVRRCSQSCFSSSGSSHYARTSPVVRVPRRSLSSRSAAGNRAEATESAMERE
KEQFRKLFIGGLSFETTESLRNYEQWGKLTDCVVMRDPASRRSRGFGFVTFSSMAEVD
AAMAAAPHSIDGRVVEPKRAVAREESGKPGAHVTVKKLFVGGIKEDTEHHLRDYFEEYL
VIIRFTPTVPHCSLTTLTVGNLHF

a2ENA Coverage Map



서열 목록

- <110> KOREA FOOD & DRUG ADMINISTRATION
Industry Academic Cooperation Foundation, Hallym University
- <120> TPM4-ALK fusion protein and composition for diagnosing cancer
- <130> 17-FP-011
- <160> 2
- <170> Kopatent In 2.0
- <210> 1

<211> 784

<212> PRT

<213> Artificial Sequence

<220><223> TPM4-ALK_fusion_protein

<400> 1

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

1 5 10 15

Leu Gln Gln Gln Ala Asp Glu Ala Glu Asp Arg Ala Gln Gly Leu Gln

20 25 30

Arg Glu Leu Asp Gly Glu Arg Glu Arg Arg Glu Lys Ala Glu Gly Asp

35 40 45

Val Ala Ala Leu Asn Arg Arg Ile Gln Leu Val Glu Glu Glu Leu Asp

50 55 60

Arg Ala Gln Glu Arg Leu Ala Thr Ala Leu Gln Lys Leu Glu Glu Ala

65 70 75 80

Glu Lys Ala Ala Asp Glu Ser Glu Arg Gly Met Lys Val Ile Glu Asn

85 90 95

Arg Ala Met Lys Asp Glu Glu Lys Met Glu Ile Gln Glu Met Gln Leu

100 105 110

Lys Glu Ala Lys His Ile Ala Glu Glu Ala Asp Arg Lys Tyr Glu Glu

115 120 125

Val Ala Arg Lys Leu Val Ile Leu Glu Gly Glu Leu Glu Arg Ala Glu

130 135 140

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

145 150 155 160

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

165 170 175

Lys Tyr Ser Glu Lys Glu Asp Lys Tyr Glu Glu Glu Ile Lys Leu Leu

180 185 190

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

195 200 205

Thr Val Ala Lys Leu Glu Lys Thr Ile Asp Asp Leu Glu Val Tyr Arg

210 215 220
 Arg Lys His Gln Glu Leu Gln Ala Met Gln Met Glu Leu Gln Ser Pro
 225 230 235 240

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

 305 310 315 320
 Ser Glu Gln Asp Glu Leu Asp Phe Leu Met Glu Ala Leu Ile Ile Ser
 325 330 335
 Lys Phe Asn His Gln Asn Ile Val Arg Cys Ile Gly Val Ser Leu Gln
 340 345 350
 Ser Leu Pro Arg Phe Ile Leu Leu Glu Leu Met Ala Gly Gly Asp Leu
 355 360 365
 Lys Ser Phe Leu Arg Glu Thr Arg Pro Arg Pro Ser Gln Pro Ser Ser
 370 375 380

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

 450 455 460

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

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

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

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

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

725

730

735

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

740

745

750

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

755

760

765

Glu Asp Thr Ile Leu Lys Ser Lys Asn Ser Met Asn Gln Pro Gly Pro

770

775

780

<210> 2

<211> 2355

<212> DNA

<213> Artificial Sequence

<220><223> TPM4-ALK_fusion_protein_cDNA

<400> 2

atggccggcc tcaactccct ggaggcgggtg aaacgcaaga tccaggccct gcagcagcag 60

gcggacgagg cggaagaccg cgcgcagggc ctgcagcggg agctggacgg cgagcgcgag 120

cggcgcgaga aagctgaagg tgatgtggcc gccctcaacc gacgcatcca gtcgttgag 180

gaggagttag acagggctca ggaacgactg gccacggccc tgcagaagct ggaggaggca 240

gaaaaagctg cagatgagag tgagagagga atgaaggtga tagaaaaccg ggccatgaag 300

gatgaggaga agatggagat tcaggagatg cagctcaaag aggccaagca cattgcggaa 360

gaggctgacc gcaaatagca ggaggtagct cgtaagctgg tcattcctgga gggtgagctg 420

gagagggcag aggagcgtgc ggaggtgtct gaactaaaat gtggtgacct ggaagaagaa 480

ctcaagaatg ttactaaca tctgaaatct ctggaggctg catctgaaaa gtattctgaa 540

aaggaggaca aatatgaaga agaaattaaa ctctgtctg acaaactgaa agaggctgag 600

accctgtctg aatttgaga gagaacggtt gcaaaactgg aaaagacaat tgatgacctg 660

gaagtgtacc gccggaagca ccaggagctg caagccatgc agatggagct gcagagccct 720

gagtacaagc tgagcaagct ccgacctcg accatcatga ccgactacaa ccccaactac 780

tgctttgctg gcaagacctc ctccatcagt gacctgaagg aggtgccgcg gaaaaacatc 840

accctcattc ggggtctggg ccatggcgcc ttggggagg tgatgaagg ccaggtgtcc 900

ggaatgccc acgaccaag cccctgcaa gtggctgtga agacgtgcc tgaagtgtgc 960

tctgaacagg acgaactgga tttctcatg gaagccctga tcatcagcaa attcaaccac 1020

cagaacattg ttcgctgcat tggggtgagc ctgcaatccc tgccccggtt catcctgctg	1080
gagctcatgg cgggggggaga cctcaagtcc ttcttccgag agaccgccc tcgcccgagc	1140
cagccctcct ccttgccat gctggacctt ctgcacgtgg ctggggacat tgcctgtggc	1200
tgtcagtatt tggaggaaaa ccacttcac caccgagaca ttgctgccag aaactgcctc	1260
ttgacctgtc caggccctgg aagagtggcc aagattggag acttcgggat ggcccagac	1320
atctacaggg cgagctacta tagaaaggga ggtgtgcca tgctgccagt taagtggatg	1380
ccccagagg ccttcatgga aggaatattc acttctaaaa cagacacatg gtcctttgga	1440
gtgctgctat gggaaatctt ttctcttga tatatgcat accccagcaa aagcaaccag	1500
gaagtcttgg agttgtcac cagtggaggc cggatggacc cacccaagaa ctgccctggg	1560
cctgtatacc ggataatgac tcagtgtgga caacatcagc ctgaagacag gcccaacttt	1620
gccatcattt tggagaggat tgaatactgc acccaggacc cggatgtaat caacaccgt	1680
ttgccgatag aatatgggcc acttgtggaa gaggaagaga aagtcctgt gaggcccaag	1740
gacctgagg gggttcctcc tctcctggc tctcaacagg caaacggga ggaggagcgc	1800
agcccagctg ccccaaccac tctgcctacc acctcctctg gcaaggctgc aaagaaacc	1860
acagctgcag agatctctgt tcgagtcct agagggccgg ccgtggaagg gggacacgtg	1920
aatatggcat tctctcagtc caacctcct tcggagtgc acaaggcca cggatccaga	1980
aacaagccca ccagcttgtg gaaccaacg tacggctcct gggttacaga gaaaccacc	2040
aaaaagaata atcctatagc aaagaaggag ccacacgaca ggggtaacct ggggctggag	2100
ggaagctgta ctgtccacc taacgttgca actgggagac ttccgggggc ctcaactgtc	2160
ctagagccct ctctgctgac tgccaatatg aaggaggtac ctctgttcag gctacgtcac	2220
ttcccttgtg ggaatgtcaa ttacggctac cagcaacagg gcttgccctt agaagccgt	2280
actgcccctg gagctggta ttacaggat accattctga aaagcaagaa tagcatgaac	2340
cagcctgggc cctga	2355