



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년08월13일
(11) 등록번호 10-2010655
(24) 등록일자 2019년08월07일

- (51) 국제특허분류(Int. Cl.)
G01N 33/68 (2006.01) C12Q 1/68 (2018.01)
G01N 33/53 (2006.01)
- (52) CPC특허분류
G01N 33/6896 (2013.01)
C12Q 1/6883 (2018.05)
- (21) 출원번호 10-2017-0118125
- (22) 출원일자 2017년09월15일
심사청구일자 2017년09월15일
- (65) 공개번호 10-2019-0030813
- (43) 공개일자 2019년03월25일
- (56) 선행기술조사문헌
US8257922 B2
WO2012075043 A1
US20150198618 A1

- (73) 특허권자
재단법인대구경북과학기술원
대구 달성군 현풍면 테크노중앙대로 333,
- (72) 발명자
최영식
대구광역시 동구 첨단로 61
윤종혁
대구광역시 동구 첨단로 61
(뒷면에 계속)
- (74) 대리인
이중승

전체 청구항 수 : 총 15 항

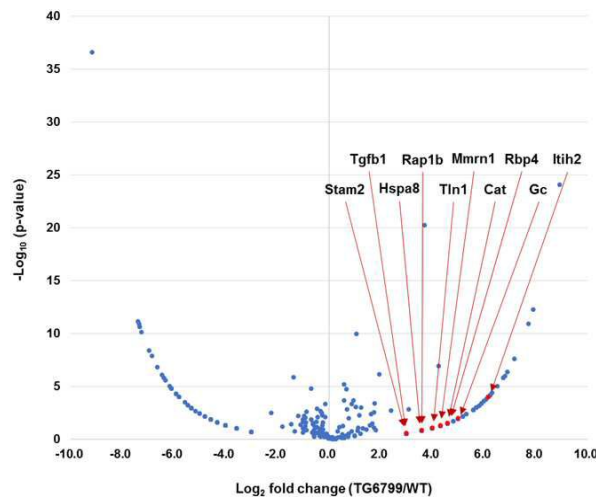
심사관 : 차명훈

(54) 발명의 명칭 알츠하이머병 진단을 위한 신규한 단백질 마커 및 이의 용도

(57) 요약

본 발명에 따르면, 혈액 EV로부터 알츠하이머병 진단에 이용될 수 있는 단백질 마커가 성공적으로 동정되었다. 먼저, 비표지정량분석법을 사용하여 야생형 대비 특이적으로 증가된 혈액 EV 단백질을 TG6799 마우스로부터 동정한 후, 바이오인포메틱스 및 생물화학 실험을 통해 11개 신규한 AD 바이오마커 후보를 선별하였다. 다음, 사람 알츠하이머병 환자의 시료를 사용하여 상기 선별된 바이오마커 후보를 평가하였다. 이와 같이, 본 발명은 알츠하이머병의 진단을 위한 신규한 단백질마커 및 효과적인 알츠하이머병 진단방법을 제공한다.

대표도 - 도3



(52) CPC특허분류

G01N 33/53 (2018.05)
C12Q 2600/158 (2013.01)
G01N 2500/04 (2013.01)

(72) 발명자

정현진

대구광역시 동구 첨단로 61

이미숙

대구광역시 동구 첨단로 61

김희정

대구광역시 동구 첨단로 61

장재명

대구광역시 동구 첨단로 61

김다솜

대구광역시 동구 첨단로 61

최유라

대구광역시 동구 첨단로 61

이 발명을 지원한 국가연구개발사업

과제고유번호 17-BR-03

부처명 미래창조과학부

연구관리전문기관 한국뇌연구원, 한국연구재단

연구사업명 한국뇌연구원연구운영비지원

연구과제명 뇌발달 및 뇌질환의 원인규명, 진단 및 제어법 개발

기여율 1/1

주관기관 한국뇌연구원

연구기간 2015.01.01 ~ 2017.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

a) 정상인 및 알츠하이머병의 진단이 필요한 피험자로부터 분리된 혈액 샘플로부터 HSPA8 (heat shock cognate 71 kDa protein), ITIH2 (inter-alpha trypsin inhibitor, heavy chain 2), DBP (vitamin D-binding protein), TLN1 (talin-1), CAT (catalase), STAM2 (signal transducing adaptor molecule 2), PARVB (beta-parvin), GNAI2 (guanine nucleotide-binding protein G(i) subunit alpha-2), ALAD (delta-aminolevulinic acid dehydratase), LDHA (L-lactate dehydrogenase), CCT4 (T-complex protein 1 subunit delta), 및 PRDX2 (peroxiredoxin 2)로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질 (exosomal proteins)을 검출하고, 및

b) 상기 (a)에서 검출된 단백질의 양이 정상인과 비교하여 피험자에서 증가되거나 감소된 경우, 피험자를 알츠하이머병 환자로 판정하는 것을 포함하는,

알츠하이머병 진단을 위한 정보를 제공하는 방법.

청구항 2

제 1항에 있어서, 정상인과 비교하여 피험자에서 증가된 양을 갖는 단백질이 HSPA8, ITIH2, DBP, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA 및 CCT4로 구성된 군으로부터 선택되는 어느 하나의 단백질인 것인, 알츠하이머병 진단을 위한 정보를 제공하는 방법.

청구항 3

제 1항에 있어서, 정상인과 비교하여 피험자에서 감소된 양을 갖는 단백질이 PRDX2인 것인, 알츠하이머병 진단을 위한 정보를 제공하는 방법.

청구항 4

제 1항 내지 제 3항 중 어느 한 항에 있어서, 상기 단백질의 검출이 상기 단백질 또는 이것의 면역원성 단편과 이에 대한 항체에 의해 형성된 항원-항체 복합체를 검출 하는 것인, 알츠하이머병 진단을 위한 정보를 제공하는 방법.

청구항 5

제 4항에 있어서, 상기 항원-항체 복합체의 검출이 웨스턴블로팅 (western blot), ELISA (enzyme linked immunosorbent assay), 면역침전분석법 (immunoprecipitation Assay), 보체고정분석법 (complement fixation assay), 유세포분석 (fluorescence activated cell sorter, FACS), 또는 단백질칩 (protein chip)에 의한 검출인 것인, 알츠하이머병 진단을 위한 정보를 제공하는 방법.

청구항 6

정상인의 혈액과 비교하여 알츠하이머병 환자의 혈액에서 상대적으로 발현이 증가하거나 감소하는 엑소솜 단백질로서, HSPA8, ITIH2, DBP, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA 및 CCT4로 구성된 군으로부터 선택된 하나 이상의 단백질 또는 이것의 면역원성 단편에 특이적으로 결합하는 항체를 유효성분으로 포함하는, 알츠하이머병 진단을 보조하기 위한 조성물.

청구항 7

제 6항에 있어서, 상기 항체는 모노클로날 항체, 다가항체, 또는 재조합 항체인 것인, 알츠하이머병 진단을 보조하기 위한 조성물.

청구항 8

정상인의 혈액과 비교하여 알츠하이머병 환자의 혈액에서 상대적으로 발현이 증가하거나 감소하는 엑소솜 단백질로서, HSPA8, ITIH2, DBP, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 단백질의 발현 수준을 측정하는 체제를 포함하는, 알츠하이머병 진단키트.

청구항 9

제 8항에 있어서, 상기 단백질 발현 수준을 측정하는 체제가 상기 단백질 또는 이의 면역원성단편에 특이적으로 결합하는 항체인 것인, 알츠하이머병 진단키트.

청구항 10

- a) 피험자에서 분리된 혈액 샘플로부터 혈액 세포외소낭을 제조하는 단계;
- b) 서열번호 13 내지 서열번호 23 및 서열번호 24로 구성된 군으로부터 선택된 염기서열에 의해 구성되는 *Hspa8*, *Iyih2*, *Gc*, *Tln1*, *Cat*, *Stam2*, *Parvb*, *Gnai2*, *Alad*, *Ldha*, *Cct4* 및 *Prdx2*의 유전자 중 하나 이상의 발현 수준을 측정하는 단계; 및
- c) 상기 b)의 유전자 중 하나 이상이 정상인 대조군과 비교하여 증가되거나 감소된 것을 확인하는 단계를 포함하여, 알츠하이머병 모니터링 또는 진단을 위한 정보를 제공하는 방법.

청구항 11

제 10항에 있어서, 상기 b)의 발현 수준은 RT-PCR, ELISA, 웨스턴블로팅 및 면역조직화학법으로 구성된 군으로부터 선택된 어느 하나를 측정하는 것인, 알츠하이머병 모니터링 또는 진단을 위한 정보를 제공하는 방법.

청구항 12

- a) 피험자로부터 분리된 혈액 시료로부터 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질의 발현 수준을 측정하는 단계; 및
- b) 상기 단계 a)의 단백질의 발현 수준이 정상인 대조군과 비교하여 증가하거나 감소된 것을 확인하는 단계를 포함하여, 알츠하이머병 진단의 정보를 제공하기 위한 엑소솜 단백질 발현 수준의 측정방법.

청구항 13

삭제

청구항 14

제 12항에 있어서, 상기 단계 a)의 단백질 발현 수준 측정은 상기 단백질 각각에 대해 특이적으로 결합하는 항체를 이용하는 것인, 엑소솜 단백질 발현 수준의 측정방법.

청구항 15

- a) 알츠하이머병 환자로부터 분리한 혈액 세포외소낭에 피검물질을 처리하는 단계;
 - b) 상기 단계 a)의 세포외소낭에서 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질의 발현 수준을 측정하는 단계; 및
 - c) 상기 단계 b)의 단백질 발현 수준을 무처리 대조군에 비해 증가시키거나 감소시키는 피검물질을 선별하는 단계를 포함하여,
- 알츠하이머병 치료제 후보물질을 스크리닝 하는 방법.

청구항 16

제 15항에 있어서, 상기 단계 b)에서 단백질 발현 수준 측정은 면역형광법, 질량분석법, 단백질 칩, 웨스턴블로팅 및 ELISA로 구성된 군으로부터 선택되는 것인, 알츠하이머병 치료제 후보물질을 스크리닝 하는 방법.

발명의 설명

기술 분야

[0001] 본 발명은 알츠하이머병 (Alzheimers disease, AD) 진단을 위한 바이오마커에 관한 것이다. 보다 구체적으로 본 발명은 알츠하이머병의 조기 진단에 이용될 수 있는 단백질 마커에 관한 것이다.

배경 기술

[0002] 알츠하이머병은 치매를 유발하는 가장 흔한 질환의 하나이며, 사고력과 기억력을 포함하는 인지능력 손상을 초래한다. 알츠하이머병의 치료를 위해, 아세틸콜린에스테라아제 저해제 (Acetylcholinesterase (AChE) inhibitors) 또는 NMDA (N-Methyl-D-aspartate) 수용체 길항제 등의 약물이 사용되고 있으나, 이들 약물의 대부분은 초기단계 알츠하이머병 환자에 대해 효과가 있는 것으로 알려져 있다. 알츠하이머병 진단을 위해, 임상 조사 (clinical exam) 및 뇌이미징 (brain imaging) 등과 같은 방법이 이용되고 있으나, 이와 같은 진단방법으로 초기단계의 알츠하이머병을 진단하는 것은 매우 어려운 일이다. 생물학적 유체 (bio fluid)를 포함하는 다양한 검체를 사용하여 알츠하이머병을 진단하고자 하는 연구가 계속되어 왔으며, 그 결과 신뢰도 있는 알츠하이머병 진단방법이 개발되었다. 그러나, 이들 진단방법 역시 후기단계 알츠하이머병의 진단에만 유용하다는 한계가 있다. 알츠하이머병의 효과적인 치료를 위해서는 병의 조기 진단이 필수적으로 요구됨에도 불구하고, 현재 초기단계 알츠하이머병을 진단할 수 있는 방법은 거의 없는 실정이다.

[0003] 질량분석 (mass spectrometry) 기술의 발전에 따라, 서브프로테오 (subproteome) 연구가 바이오마커 개발을 위한 차세대 패러다임이 되었다. 예를 들어, 세포외소낭 (엑소솜, extracellular vesicle, EV)은 이것의 연구 초기에는 단순한 분비체 (secretome)에 불과한 것으로 여겨졌으나, 이들 세포외소낭이 그 내부에 세포 단백질, 핵산, 및 지질 등을 포함하고, 측분비 (paracrine) 또는 내분비 (endocrine) 작용을 통해 다양한 시그널 매개에 관여함에 의해 셔틀 비이클 (shuttling vehicle)과 같은 특성을 나타내는 것이 밝혀짐에 따라, 최근에는 새로운 시그널 매개체로서 인정되고 있다. 세포외소낭은 시그널링 단백질, 핵산 및 지질을 포함하는 이유로 바이오마커 개발을 위한 흥미로운 타겟이 되었음에도 불구하고, 혈액 EV의 경우 이의 샘플 제조에 어려움이 있어 이를 대상으로 한 연구가 거의 없었다.

[0004] 본 발명자들은 혈액으로부터 알츠하이머병의 진단을 위한 바이오마커를 발견하고자 부단히 노력하였으며, 그 결과 혈액에서 분리한 엑소솜 단백질에서 알츠하이머병 진단에 이용될 수 있는 신규한 단백질 바이오마커를 발견하고 그 특성을 확인함으로써 본 발명을 완성하였다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국특허공개 제 10-2017-0005940호

비특허문헌

[0006] (비특허문헌 0001) Choi, D. S. & Gho, Y. S. Isolation of extracellular vesicles for proteomic profiling. Methods in molecular biology 1295, 167-177, doi:10.1007/978-1-4939-2550-6_14 (2015)

발명의 내용

해결하려는 과제

- [0007] 본 발명은 알츠하이머병 진단을 위한 단백질 마커를 제공하는 것을 목적으로 한다.
- [0008] 또한, 본 발명은 알츠하이머병 진단을 위한 단백질 마커로서 엑소솜 단백질을 제공하는 것을 목적으로 한다.
- [0009] 또한, 본 발명은 알츠하이머병 진단을 위한 조성물을 제공하는 것을 목적으로 한다.
- [0010] 또한, 본 발명은 알츠하이머병 진단에 필요한 정보의 제공에 이용되는 단백질 마커를 제공하는 것을 목적으로 한다.

한다.

[0011] 또한, 본 발명은 단백질 마커를 이용하여 알츠하이머병 치료를 위한 약물 후보물질을 스크리닝하는 방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0012] 상기 목적을 달성하기 위해, 본 발명은 혈액 세포외소낭으로부터 분리한 알츠하이머병 진단을 위한 단백질 바이오마커를 제공한다. 본 발명의 일 실시예에서, 본 발명자들은 TG6799 마우스의 혈액 EV (extracellular vesicles)에 대한 프로테옴 분석을 수행하고, 신규한 AD 바이오마커 단백질을 동정하였다. 구체적으로, 혈액 EV 내에 존재하는 바이오마커 후보를 확인하기 위해, 메탄올-클로로포름 침전 및 비이온세제-용합 겔-내 분해를 포함하는 2가지의 다른 EV 단백질 제조방법을 병행하였다. 다음, 바이오인포메틱스 분석을 통해, AD 혈액에서 특이적으로 증가된 18개의 바이오마커 후보 단백질을 스크리닝 하였다. 이어서, 웨스턴블로팅으로 이들 단백질의 증가 수준을 평가하고, 사람 AD 환자로부터 분리한 혈액 EV를 사용하여 이들에 대한 임상적 신뢰도를 검증하였다.

[0013] 또한, 본 발명은, a) 정상인 및 알츠하이머병의 진단이 필요한 피험자로부터 분리된 혈액샘플로부터, HSPA8 (heat shock cognate 71 kDa protein), ITIH2 (inter-alpha trypsin inhibitor, heavy chain 2), DBP (vitamin D-binding protein), TLN1 (talin-1), CAT (catalase), STAM2 (signal transducing adaptor molecule 2), PARVB (beta-parvin), GNAI2 (guanine nucleotide-binding protein G(i) subunit alpha-2), ALAD (delta-aminolevulinic acid dehydratase), LDHA (L-lactate dehydrogenase), CCT4 (T-complex protein 1 subunit delta), 및 PRDX2 (peroxiredoxin 2)로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질 (exosomal proteins)을 검출하고, 및 b) 상기 a)에서 검출된 단백질의 양이 정상인과 비교하여 피험자에서 증가되거나 감소된 경우, 피험자를 알츠하이머병 환자로 판단하는 것인, 알츠하이머병을 진단하는 방법을 제공한다.

[0014] 본 발명의 일 실시예에서, 정상인과 비교하여 피험자에서 증가된 양을 갖는 단백질은 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA 및 CCT4로 구성된 군으로부터 선택되는 단백질이고, 정상인과 비교하여 피험자에서 감소된 양을 갖는 단백질은 PRDX2이다. 본 발명의 다른 실시예에서, 단백질 검출은 상기 단백질 또는 이것의 면역원성 단편과 이에 대한 항체에 의해 형성된 항원-항체 복합체를 검출하며, 이러한 검출은 웨스턴 블로팅 (western blot), ELISA (enzyme linked immunosorbent assay), 면역침전분석법 (immunoprecipitation Assay), 보체고정분석법 (complement fixation assay), 유세포분석 (fluorescence activated cell sorter, FACS), 또는 단백질칩 (protein chip)에 의해 수행될 수 있다.

[0015] 또한, 본 발명은 정상인의 혈액에서와 비교하여 알츠하이머병 환자의 혈액에서 상대적으로 발현이 증가하거나 감소하는 엑소솜 단백질 또는 이것의 면역원성 단편에 특이적으로 결합하는 항체를 유효성분으로 포함하는, 알츠하이머병 진단을 위한 조성물을 제공한다.

[0016] 본 발명의 일 실시예에서, 상기 항체는 모노클로날 항체, 다클론항체, 또는 재조합 항체이다. 본 발명의 다른 일 실시예에서, 상기 알츠하이머병 진단을 위한 조성물은, HSPA8 (Heat shock cognate 71 kDa protein; NCBI Genbank 등록번호:15481), ITIH2 (inter-alpha trypsin inhibitor, heavy chain 2; NCBI Genbank 등록번호:16425), Gc (vitamin D-binding protein; NCBI Genbank 등록번호:14473), TLN1 (talin-1; NCBI Genbank 등록번호:21894), CAT (catalase; NCBI Genbank 등록번호:12359), STAM2 (signal transducing adaptor molecule 2; NCBI Genbank 등록번호:56324), PARVB (beta-parvin; NCBI Genbank 등록번호:170736), GNAI2 (guanine nucleotide-binding protein G(i) subunit alpha-2; NCBI Genbank 등록번호:14678), ALAD (delta-aminolevulinic acid dehydratase; NCBI Genbank 등록번호:17025), LDHA (L-lactate dehydrogenase; NCBI Genbank 등록번호:16828) 및 CCT4 (T-complex protein 1 subunit delta; NCBI Genbank 등록번호:12464)로 이루어진 군에서 선택되는 하나 이상의 단백질의 발현 수준을 측정하는 체제를 추가로 포함할 수 있다. 상기 단백질의 발현 수준을 측정하는 체제는 상기 단백질 또는 이것의 면역원성 단편에 특이적인 항체일 수 있으며, 여기서 항체는 모노클로날항체, 다클론항체 및 재조합 항체를 모두 포함한다. 본 발명에서 모노클로날항체는 당해 분야에 널리 공지된 하이브리도마 방법 (Kohler 및 Milstein(1976)European journal of Immunology 6:511-519), 또는 파지 항체 라이브러리 (Clarkson et al, Nature, 352:624-628, 1991; Marks et al, J. Mol. Biol.,222:58, 1-597, 1991) 기술을 이용하여 제조될 수 있다. 다클론항체는 상기 단백질 항원을 동물에 주입하고 동물로부터 항체를 포함하는 혈청을 수득하는 것을 포함하여 종래 방법에 의해 제조될 수 있다. 이러한 다클론항체는 개, 염소, 양, 토끼, 원숭이, 말, 돼지, 및 소 등을 포함하는 임의의 동물로부터 제조하는 것이

가능하다. 또한, 본 발명의 항체에는 키메라 항체, 인간화 항체, 인간항체 등이 포함된다. 또한, 본 발명에 사용되는 항체는 2개의 전체 길이의 경쇄 및 2개의 전체 길이의 중쇄를 가지는 완전한 형태의 항체 및 이것의 기능적 단편을 포함한다. 상기 항체의 기능적 단편은 항원 결합기능을 보유하는 단편을 의미하며 Fab, F(ab'), F(ab')₂ 및 Fv 등을 포함한다. 또한, 본 발명은 정상인의 혈액과 비교하여 알츠하이머병 환자의 혈액에서 상대적으로 발현량이 증가하거나 감소하는 엑소솜 단백질로서, HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 단백질의 발현 수준을 측정하는 체제를 포함하는, 알츠하이머병 진단키트를 제공한다.

[0017] 본 발명의 다른 일 실시예에서, 상기 키트는 피험자로부터 분리된 생체시료 내 상기 단백질 마커의 수준을 측정하는 체제 뿐만 아니라, 단백질 수준의 분석에 적합한 하나 이상의 조성물, 용액 또는 장치를 포함할 수 있다. 예를 들어, 상기 키트는 항체의 면역학적 검출을 위하여 기질, 완충용액, 검출 라벨로 표지된 2차 항체, 및 발색 기질 등을 포함할 수 있다. 본 발명의 또 다른 일 실시예에서, 상기 알츠하이머병 진단을 위한 키트는 ELISA 키트, 샌드위치 ELISA 등의 ELISA 방법을 수행하기 위해, 필요한 성분을 포함하는 키트를 제공한다. 즉, 상기 ELISA 키트는 상기 단백질 마커에 특이적인 항체를 포함하고, 결합된 항체를 검출할 수 있는 시약, 예를 들면, 표지된 2차 항체, 발색단 (chromophores), 효소 및 그의 기질 또는 항체와 결합할 수 있는 다른 물질 등을 포함할 수 있다.

[0018] 본 발명의 다른 일 실시예에서, 상기 알츠하이머병 진단을 위한 키트는 웨스턴 블로팅, 면역침전분석법, 보체고정분석법, 유세포분석, 또는 단백질칩 등에 대한 키트일 수 있다.

[0019] 본 발명의 또 다른 일 실시예에서, 본 발명은 상기 알츠하이머병의 진단용 조성물을 포함하는 알츠하이머병 진단을 위한 단백질 칩에 관한 것이다. 상기 단백질칩은 상기 본 발명의 단백질 마커에 대한 하나 이상의 항체가 칩 위에 고밀도로 배치된 것으로 구성된다. 본 발명에서 상기 검출 라벨은 효소, 형광체, 리간드, 발광체, 마이크로입자, 레독스 분자 및 방사선 동위원소로 이루어진 그룹 중에서 선택될 수 있으나, 이에 제한되는 것은 아니다.

[0020] 또한, 본 발명은 a) 피험자에서 분리된 혈액 샘플로부터 혈액 세포외소낭을 제조하는 단계; b) 각각 서열번호 13 내지 서열번호 23 및 서열번호 24의 염기서열로 구성되는 *Hspa8*, *Iyih2*, *Gc*, *Tln1*, *Cat*, *Stam2*, *Parvb*, *Gnai2*, *Alad*, *Ldha*, *Cct4* 및 *Prdx2*로 구성된 유전자 중 하나 이상의 발현 수준을 측정하는 단계; 및 c) 상기 b)의 유전자 중 하나 이상이 정상인 대조군과 비교하여 증가되거나 감소된 경우, 알츠하이머병에 걸리거나 위험도가 증가한 것으로 판정하는 단계를 포함하여, 알츠하이머병 모니터링 또는 진단을 위한 정보를 제공하는 방법을 제공한다. 본 발명의 일 실시예에서, 상기 b)의 발현 수준은 RT-PCR, ELISA, 웨스턴블로팅 및 면역조직화학법으로 구성된 군으로부터 선택된 어느 하나의 방법에 의해 측정될 수 있다.

[0021] 또한, 본 발명은 a) 피험자로부터 분리된 생물학적 시료로부터 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질의 발현 수준을 측정하는 단계; 및 b) 상기 단계 a)의 단백질의 발현 수준이 정상인 대조군과 비교하여 증가하거나 감소된 경우, 상기 피험자가 알츠하이머병에 걸릴 위험도가 높거나 알츠하이머병에 걸린 객체인 것으로 판정하는 단계를 포함하여, 알츠하이머병 진단을 위한 정보를 제공하기 위한 엑소솜 단백질 발현수준의 측정방법을 제공한다. 여기서, 상기 단계 a)의 생물학적 시료는 골수, 혈청, 혈장 및 혈액으로 구성된 군으로부터 선택될 수 있으며, 상기 단계 a)에서 단백질 발현수준 측정을 위해 상기 단백질 각각에 대해 특이적으로 결합하는 항체를 이용할 수 있다.

[0022] 또한, 본 발명은 a) 알츠하이머병 환자로부터 분리한 혈액 세포외소낭에 피검물질을 처리하는 단계; b) 상기 단계 a)의 세포외소낭에서 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 구성된 군으로부터 선택된 하나 이상의 엑소솜 단백질의 발현 수준을 측정하는 단계; 및 c) 상기 단계 b)의 단백질 발현 수준을 무처리 대조군에 비해 증가시키거나 감소시키는 피검물질을 선별하는 단계를 포함하여, 알츠하이머병 치료제 후보물질을 스크리닝 하는 방법을 제공한다. 본 발명의 일 실시예에서, 상기 단계 b)에서의 단백질 발현 수준 측정은 면역형광법, 질량분석법, 단백질칩, 웨스턴블로팅 및 ELISA로 구성된 군으로부터 선택된다.

[0023] 본 발명의 일 실시예에서, TG6799 마우스 혈액 EV (extracellular vesicles) 샘플로부터 발현 수준의 차이가 특징적으로 나타나는 12개의 단백질을 선별하였으며 (HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2), 이들 중 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA 및 CCT4 단백질의 발현수준이, 정상인과 비교하여 AD 환자에서 높게 나타난 것을 확인하였다 (도 3).

[0024] 본 발명에서 용어, "진단"은 질환 또는 질병의 존재 또는 특성을 확인하는 것을 의미한다. 본 발명의 목적상, 진단은 알츠하이머병의 발병 여부를 확인하거나, 질환의 진행 또는 심화 여부를 확인하는 것을 의미할 수 있다.

[0025] 본 발명에서 용어, "진단을 위한 마커 또는 진단 진단 마커"는 알츠하이머병을 정상인 상태와 구분 할 수 있는 물질로서, 정상인 개체로부터 분리한 생체시료 와 비교하여 알츠하이머병을 가진 개체의 시료에서 증가 또는 감소의 변화된 양상을 보이는 단백질, 펩티드, 폴리펩타이드 또는 핵산, 지질, 당지질, 또는 당 (단당류, 이당류, 올리고당류 등) 등과 같은 유기 생체분자를 포함한다. 본 발명의 목적상, 본 발명의 진단을 위한 마커는 알츠하이머병 환자의 시료에서 특이적으로 증가된 발현을 나타내는 HSPA8 (Heat shock cognate 71 kDa protein), ITIH2 (inter-alpha trypsin inhibitor, heavy chain 2), DBP (vitamin D-binding protein), TLN1 (talin-1), CAT (catalase), STAM2 (signal transducing adaptor molecule 2), PARVB (beta-parvin), GNAI2 (guanine nucleotide-binding protein G(i) subunit alpha-2), ALAD (delta-aminolevulinic acid dehydratase), LDHA (L-lactate dehydrogenase) 및 CCT4 (T-complex protein 1 subunit delta), 그리고 알츠하이머병 환자의 시료에서 특이적으로 감소된 발현을 나타내는 페록시레독신2 (PRDX2, Peroxiredoxin 2) 단백질을 말한다. 본 발명의 상기 마커 단백질의 아미노산 서열 및 염기서열은 공지된 유전자 데이터베이스에서 그 서열 정보를 확인할 수 있다. 예를 들어, 인간 HSPA8 단백질의 아미노산 서열은 NCBI Genbank 등록번호 15481에서 확인할 수 있고, 이를 서열번호 1에 기재하였으며, 이것의 염기서열은 서열번호 13에 기재하였다. 마찬가지로, ITIH2 단백질의 아미노산 서열은 NCBI Genbank 등록번호 16425 에서 확인할 수 있고, 이를 서열번호 2에 기재하였으며, 이것의 염기서열은 서열번호 14에 기재하였다. DBP 단백질의 아미노산 서열은 NCBI Genbank 등록번호 14473에서 확인할 수 있고, 이를 서열번호 3에 기재하였으며, 이것의 염기서열은 서열번호 15에 기재하였다. TLN1 (talin-1) 단백질의 아미노산 서열은 NCBI Genbank 등록번호 21894에서 확인할 수 있고, 이를 서열번호 4에 기재하였으며, 이것의 염기서열은 서열번호 16에 기재하였다. CAT (catalase) 단백질의 아미노산 서열은 NCBI Genbank 등록번호 12359에서 확인할 수 있고, 이를 서열번호 5에 기재하였으며, 이것의 염기서열은 서열번호 17에 기재하였다. STAM2 단백질의 아미노산 서열은 NCBI Genbank 등록번호 56324에서 확인할 수 있고, 이를 서열번호 6에 기재하였으며, 이것의 염기서열은 서열번호 18에 기재하였다. PARVB 단백질의 아미노산 서열은 NCBI Genbank 등록번호 170736에서 확인할 수 있고, 이를 서열번호 7에 기재하였으며, 이것의 염기서열은 서열번호 19에 기재하였다. GNAI2 단백질의 아미노산 서열은 NCBI Genbank 등록번호 14678에서 확인할 수 있고, 이를 서열번호 8에 기재하였으며, 이것의 염기서열은 서열번호 20에 기재하였다. ALAD 단백질의 아미노산 서열은 NCBI Genbank 등록번호 17025에서 확인할 수 있고, 이를 서열번호 9에 기재하였으며, 이것의 염기서열은 서열번호 21에 기재하였다. LDHA 단백질의 아미노산 서열은 NCBI Genbank 등록번호 16828에서 확인할 수 있고, 이를 서열번호 10에 기재하였으며, 이것의 염기서열은 서열번호 22에 기재하였다. CCT4 단백질의 아미노산 서열은 NCBI Genbank 등록번호 12464에서 확인할 수 있고, 이를 서열번호 11에 기재하였으며, 이것의 염기서열은 서열번호 23에 기재하였다. PRDX2 단백질의 아미노산 서열은 NCBI Genbank 등록번호 21672에서 확인할 수 있고, 이를 서열번호 12에 기재하였으며, 이것의 염기서열은 서열번호 24에 기재하였다.

[0026] 본 발명에서 용어, "검체 시료"는 본 발명의 단백질 마커의 발현 수준이 차이 나는 조직, 세포, 전혈, 혈장, 혈청, 혈액, 타액, 객담, 림프액, 뇌척수액, 세포간액 또는 소변과 같은 시료를 포함하며, 이에 제한되는 것은 아니다. 바람직하게는, 혈액의 세포외소낭이다. 본 발명의 일 실시예에 따르면, 혈액 시료에서 단백질을 분리하는 과정은 일반적으로 알려진 공정을 이용하여 수행할 수 있으며, 단백질 수준은 상기와 같은 다양한 항원-항체 결합체 측정방법으로 측정할 수 있다.

발명의 효과

[0027] 본 발명은 HSPA8, ITIH2, GC, TLN1, CAT, STAM2, PARVB, GNAI2, ALAD, LDHA, CCT4 및 PRDX2로 이루어진 군에서 선택되는 하나 이상의 단백질의 발현 수준을 측정하는 제제를 이용하여, 알츠하이머병을 조기에 진단할 수 있는 방법을 제공한다.

도면의 간단한 설명

[0028] 도 1a는 본 발명에 따라 TG6799 마우스 혈액 EV에서 알츠하이머병 진단을 위한 단백질 마커를 검출하는 방법을 개략적으로 나타낸 것이다.

도 1b는 아밀로이드베타의 단량체 및 사량체가 TG6799 마우스 뇌에서만 존재하는 것을 확인한 결과이다.

도 1c는 입자분석기를 사용하여 세포외소낭의 크기를 분석한 것이다.

도 1d는 본 발명에서 제조한 혈액 EV 중에 본 발명의 단백질 마커가 많은 양 존재하는 것을 나타낸다.

도 1e는 본 발명에서 제조한 혈액 EV 중에 아밀로이드베타가 존재하는 것을 나타낸 것이다.

도 2a 및 2b는 본 발명에 따른 혈액 EV 내에 발현 수준이 증가된 단백질 (71개) 및 감소된 단백질 (60개)을 나타낸다.

도 3은 본 발명에서 선별한 10개의 단백질이 이들 단백질 선별기준의 신뢰도를 나타내는 볼캐노 형태의 선상에 있음을 제시한다.

도 4는 비이온 세제-융합 SDS-PAGE에 의해 수득한 TG6799 마우스 혈액 EV를 사용한 분석으로서, 혈액 EV 막에 존재하는 단백질을 추출하고 쿠마시블루 염색한 결과 특이적으로 나타나는 3개 밴드를 표시한 나타낸 것이다.

도 5a 및 5b는 TG6799 마우스 혈액 EV를 사용하여 스크리닝한 바이오마커 후보 단백질에 대한 평가결과로서, 이들 단백질이 AD 혈액 EV에서 뚜렷하게 증가한 것을 나타낸다.

도 6a 및 6b는 사람 유래 시료를 사용하여 TG6799 마우스 혈액 EV에서 분리한 AD 관련 단백질의 바이오마커로서의 가능성을 평가한 것으로서, 상기 단백질의 수준이 AD 혈액 EV에서 뚜렷하게 증가한 것을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0029]

재료 및 방법

[0030]

(1) 실험동물 관리

[0031]

모든 동물실험은 한국 뇌연구원의 실험동물 취급 및 보호 위원회의 승인 하에 수행되었다. TG6799 마우스 (입수처: The Jackson Laboratory)를 12시간 명/암 주기 조건에서 자유롭게 물을 섭취할 수 있도록 하였다.

[0032]

(2) 혈액 세포외소낭의 제조

[0033]

EDTA (ethylenediaminetetraacetic acid)를 처리하고 원심분리를 수행하여 혈장을 제조한 후, 상기 혈장을 HBSS 완충액 (cat#14175095, Invitrogen)에 10배 희석하였다. 다음, 4°C에서 60분간 인큐베이션 한 후, 상기 희석용액을 테이블탑 원심분리기를 사용하여 4°C, 20분간 12,000 rpm으로 원심분리 하였다. 생성된 펠릿을 HBS 1 ml에 용해시키고, 4°C, 90분간 120,000 x g에서 원심분리 하였다. 침전된 펠릿을 HBS 200 μl에 다시 현탁시켰다. 엑소좀 단백질 농축물을 BCA (bicinchoninic acid assay) 어세이로 측정하였다 (Cat# 23225, ThermoFisher scientific).

[0034]

(3) 질량분석을 위한 혈액 세포외소낭의 펩타이드 제조

[0035]

메탄올-클로로포름 (methanol-chloroform) 방법을 수행하였다. 요약하면, 혈액 세포외소낭 100 μl를 메탄올 400 μl에 혼합하고 충분히 보텍싱 (vortexing) 하였다. 상기 혼합물을 9,000 x g에서 5분간 원심분리 하였다. 다음, 클로로포름 100 μl를 가하고 충분히 보텍싱 하였다. 9,000 x g에서 5분간 원심분리 한 후, 300 L 클로로포름을 가하고 충분히 보텍싱 하였다. 상기 혼합물을 16,000 x g에서 5분 원심분리 한 후, 상층부를 조심스럽게 제거하였다. 메탄올 300 μl을 상기 상층부에 가한 후 가볍게 두드려서 혼합하였다. 16,000 X g에서 5분 원심분리한 후, 상층액을 제거하였다. 다음, 5분 동안 건조 한 후, 펠릿을 수거하였다. 용액 내 분해 (in-solution digestion)를 위해, 상기 펠릿을 40 mM NH₄HCO₃ 중의 6 M 우레아 100μl에 재현탁하였다. 5 mM 디티오 트라이톨 (dithiothreitol, DTT)과 함께 실온에서 1시간 인큐베이션 한 후, 암실에서 이오도아세트아미드 (iodoacetamide, IAA)를 실온으로 30분간 처리하였다. 단백질 대비 1:20 비율의 트립신 (Promega, Madison, WI, USA)을 37°C에서 16시간 처리하였다. 동결건조기를 사용하여 건조한 후, 탈염 컬럼 (#89873, Thermo Fisher scientific, San Jose, CA, USA)을 제조사의 지시에 따라 사용하였다. 비이온성 세제 융합 SDS-PAGE 방법을 수행하기 위해, 혈액 EV를 1% DDM (n-dodecyl-β-D-maltoside)으로 재현탁하고 37°C에서 10분간 인큐베이션 하였다. 5% 베타-메르kapto에탄올 (beta-mercaptoethanol)과 SDS (Sodium Dodecyl Sulfate) 샘플 완충액을 상기 재현탁 혼합물과 혼합하고 10분 동안 끓인 다음, 상기 샘플을 SDS-PAGE로 분석하였다. 겔 플러그를 30% ACN (acetonitrile)로 세척하고, 10 mM DTT로 56°C에서 30분 처리하였다. 겔 플러그를 100% ACN로 세척한 후, 100 mM IAA를 암실에서 실온으로 30분 처리하였다. 겔 플러그를 100% ACN로 세척한 후, 겔 플러그를 동결건조기로 건조한 다음, 트립신 (1:50 비율)을 37°C에서 16시간 처리하였다.

[0036]

(4) 질량 분석 및 DB 분석

[0037] 트립신 분해물을 나노엘렉트로스프레이 이온소스 (nano-electrospray ion source)가 결합된 Q-exactive plus hybrid quadrupole orbit-trap 질량분석기 (Thermo Fisher scientific, San Jose, CA, USA)로 분석하였다. 펩타이드의 크로마토그래피 분리를 로딩 컬럼으로서 Acclaim PepMap™ 100 (75 μm x 2 cm, Thermo Fisher scientific, San Jose, CA, USA) 및 Acclaim PepMap™ RSLC (75 μm x 15 cm, 100 resin, Thermo Fisher scientific, San Jose, CA, USA)이 장착된 Ultimate 3000 RSLCnano system (Thermo Fisher scientific, San Jose, CA, USA) 상에서 수행하였다. 펩타이드를 RS auto-sampler로부터 로딩하고, 0.1% 포름산 (formic acid)을 함유하는 ACN/water의 유속 300 nl/min의 선형구배로 분리하였다. LC 용리제를 분석컬럼으로부터 전자분무하고, 2.0 kV의 전압이 나노스프레이 소스의 액체 접촉부를 통해 인가되도록 하였다. 펩타이드 혼합물을 40 분 동안 5%에서 40% 농도구배를 갖도록 한 ACN을 사용하여 분리하였다. 상기 분석법은 350 ~ 2000 m/z 범위의 전체 MS 스캔 (full MS scan), 및 상기 전체 MS 스캔으로부터 10개의 가장 강한 이온 상에 데이터-의존적인 MS/MS (MS2)로 구성하였다. 상기 질량분석기는 데이터 의존적 모드로 포착하도록 프로그래밍 하였다. 질량분석기 교정 (calibration)은 제조자의 지시에 따라 제안된 교정 솔루션으로 수행되었다. 본 발명에서 프로테오믹스 분석은, 각 조건 당 3개의 생물학적 리플리케이트 (biological replicates) 샘플을 사용하고, 각 샘플에 대해 2회 반복 (technical replicates) 하여 수행하였다. 데이터베이스 검색을 수행하기 위해, 탠덤질량스펙트럼 (tandem mass spectra)을 Thermo Fisher Scientific Proteome Discoverer software version 2.1.1.21로 프로세싱 하였다. 상기 스펙트럼 데이터를 마우스 Uniprot database (release version 2016_06)에 대해 검색하였다. 이때 사용된 분석 플로우에는 4개의 노드 (node) 즉, Spectrum Files (data input), Spectrum Selector (spectrum and feature retrieval), Sequest HT (sequence database search), 및 Percolator (peptide spectral match or PSM Validation and FDR analysis)를 포함하였다. 모든 동정된 단백질은 펩타이드 수준에서 계산하여 ≤ 1%의 FDR (false discovery rate)를 가졌다. 평가는 q-value에 기초하였다. 검색 파라미터는 가변수식화 (dynamic modification)로서 메티오닌 산화 및 고정수식화 (fixed modification)로서 시스테인의 메틸티오-수식화 함께 미절단 (missed cleavage) 개수 최대값 2까지의 트립신분해 특이도 (tryptic specificity)를 허용하였다. +1, +2, 및 +3 이온에 대한 질량검색 파라미터는, 전구이온 (precursor ions)에 대해 20 ppm 및 조각이온 (fragment ions)에 대해 0.6 Da의 질량에러 허용 오차를 포함하였다.

[0038] (5) 웨스턴블로팅

[0039] 혈액 EVs를 SDS 샘플완충액 (Bio-Rad) 및 5% 베타-메르카пто에탄올 (beta-mercaptoethanol)과 혼합한 다음, 10 분 동안 끓였다. 마우스로부터 뇌 조직을 적출하고, 얼음-냉각 PBS로 세척한 다음, 1x Halt 프로테아제 및 프로테아제 저해제 각테일을 함유하는 RIPA 완충액 (radioimmunoprecipitation assay buffer) (Thermo fisher)으로 균질화 하였다. 생성된 균질물을 소니케이션하고, 세포파편을 4℃, 15분 동안 13,000 x g로 원심분리 하여 제거하였다. 단백질 농도를 BCA 어세이 (ThermoFisher scientific)로 측정된 후, 용해물을 SDS 샘플완충액과 혼합하였다. SDS-PAGE로 단백질을 분리한 후, Bio-Rad wet transfer system을 사용하여 단백질을 PVDF (polyvinylidene fluoride) 멤브레인으로 옮기고 5% 스킴밀크를 함유하는 PBS-T로 30분간 블로킹한 다음, 항체와 함께 4℃에서 16시간 동안 인큐베이션하였다. 상기 멤브레인을 PBS-T로 3회 세척한 후, 블롯을 HRP (horseradish peroxidase)-컨쥬게이션된 안티-마우스 또는 안티-래빗 항체와 함께 실온에서 1시간 인큐베이션 하였다. 다음, 상기 멤브레인을 PBS-T (phosphate buffered saline with Tween 20, Thermo Fisher Scientific)로 세척하고 ECL 웨스턴블로팅시스템 (GE Healthcare)으로 현상하였다. Inter-alpha trypsin inhibitor heavy chain 2 (ab118257, Abcam), vitamin D-binding protein (ab153922, Abcam), catalase (ab52477, Abcam), talin-1 (4021S, Cell Signaling Technology), heat shock cognate 71 kDa protein (ab154415, Abcam), signal transducing adaptor molecule (SH3 domain and ITAM motif) 2 (226050, United States Biological), Peroxiredoxin 2 (ab71533, Abcam), Beta-parvin (sc-375581, Santa Cruz Biotechnology), Guanine nucleotide-binding protein G(i) subunit alpha-2 (sc-13534, Santa Cruz Biotechnology), porphobilinogen synthase (sc-271585, Santa Cruz Biotechnology), L-lactate dehydrogenase A chain (PA5-23036, Thermo Fisher Scientific), T-complex protein 1 subunit delta (ab129072, Abcam), CD9 (#553758, BD Biosciences), CD63 (sc-15363, Santa Cruz Biotechnology), CD81 (ab33697, Abcam), 및 Alix (2171S, Cell Signaling Technology)에 대한 웨스턴블로팅을 수행하였다.

[0040] (6) 통계분석

[0041] 대조군 및 AD군에서 각각 동정된 단백질의 상대적 양을 각 조건으로부터의 PSMs (peptide spectrum matches)와 대비하여 비교하였다. 각 단백질에 대해, 2개 조건 사이의 양적 차이의 p-값은 카이제곱검정 (Chi-square

test)으로 평가하였다. 카이제곱검정 값은 다음 수식 1과 같이 세팅된 자유도와 함께 G test (test of goodness of fit; 하기 수식 1 참조)로 측정하였다:

[0042] (수식 1) $G=4.60517*[f1*\log(f1)+ f2*\log(f2)- f1*\log(f^1)- f2*\log(f^2)]$,

[0043] 상기 식에서, f1 및 f2는 각각 대조군 및 AD군의 혈액 EVs의 LC-MS/MS 결과로부터의 PSMs를 나타내며, f^1과 f^2는 예측된 PSMs를 나타낸다. 상기 예측된 PSMs는 PSMs의 합계 간의 비율로 정상화하였다. 예측된 PSMs를 2개 조건 모두에서의 PSMs의 합으로 평가하였으며, 각 조건의 PSMs의 합과 전체 PSMs의 비율을 곱하였다. 배타적으로 동정된 단백질에 대해서는, p-값의 평가를 위해, 비검출된 조건에 0.25 PSM 값을 주었다.

[0044] **실시예**

[0045] 실시예 1: TG6799 마우스로부터 혈액 세포외소낭의 제조

[0046] TG6799 마우스로부터 혈액 세포외소낭의 제조하는 과정을 도 1a에 개략적으로 도시하였다. TG6799 마우스에서만 관찰되는 아밀로이드베타 발현을 확인하기 위해, 본 발명자들은 TG6799 마우스 (6개월령, 입수처: The Jackson Laboratory)의 전체 뇌 용해물을 사용하여 웨스턴블로팅을 수행하였다. 본 발명자들은 아밀로이드베타의 단량체 및 사량체 (tetramers)가 TG6799 뇌에서만 존재하는 것을 확인하였다 (도 1b). 혈액 EV에서 AD의 존적으로 변화된 단백질을 찾기 위해, 본 발명자들은 먼저 야생형 및 TG6799 마우스 혈액으로부터 혈액 EV를 분리하였다. 혈액의 혈장을 제조한 다음, 상기 혈장을 농도구배 초원심분리에 적용하여 혈액 EV를 수득하였다. 혈액 EV의 품질 평가를 위해, 입자분석기 (Nanosight, 제조사)를 사용하여 소낭 (vesicle) 크기를 분석하였다. 그 결과, 100 nm 크기의 소낭이 대부분인 것으로 나타났다 (도 1c). 다음, 본 발명자들은 상기 제조한 혈액 EV로부터 소낭 바이오마커 단백질을 확인하였다. 고품질의 소낭 바이오마커 단백질 예컨대, CD9, CD63, CD81 및 Alix가 상기 제조된 혈액 EV 중에 매우 많은 양으로 존재하였다 (도 1d). 몇몇 소낭 내에 아밀로이드베타가 존재한다는 것이 보고된 바 있으므로, 본 발명자들은 상기 제조된 혈액 EV 내에서 아밀로이드베타를 유무를 확인하였다. 그 결과, 상기 제조된 혈액 EV 중에 아밀로이드베타가 존재 하는 것이 확인되었다 (도 1e). 이와 같이, 본 발명자들은 TG6799 마우스 및 야생형 마우스로부터 양질의 혈액 EVs가 분리되었음을 확인하였다.

[0047] 실시예 2: 메탄올-클로로포름 침전을 사용한 혈액 EV의 프로테옴 분석

[0048] 제1 EV 프로테옴 제조방법으로서, 상기 수득한 혈액 EV를 메탄올-클로로포름으로 처리하여 EV 단백질을 침전시켰다. 다음, 상기 생성된 침전물을 용액내 분해 (in-solution digestion)에 적용하여 트립신 분해 펩타이드를 제조하였다. 상기 혈액 EVs로부터의 트립신 분해 펩타이드를 LC-electrospray (ESI)-MS/MS로 분석하였다. Sequest search engine-applied Proteome discoverer 2.1로 Uniprot database를 검색하여, 야생형 및 TG6799 마우스로부터 317개의 단백질을 동정하였다. 비표지정량분석법으로 log₂>3의 변화값 (fold change) 또는 높은 통계적 유의성 (P<0.05)의 선별기준을 만족하는 것으로서, 71개의 증가된 단백질 및 60개의 감소된 단백질을 선별하였다 (도 2a 및 도 2b). 상기 프로테옴 결과로부터, 본 발명자들은 AD 바이오마커 후보로서 10개 단백질을 최종 선별하였다 (표 1).

표 1

[0049]

Uniprot	Gene ID	Protein name	% Coverage (aver)	Fold change (log2)
G3X977	Itih2	Inter-alpha trypsin inhibitor, heavy chain 2	8.6	5.91
P21614	Gc	Vitamin D-binding protein	10.4	5.00
H7BWY6	Rbp4	Retinol-binding protein 4	9.0	4.58
P24270	Cat	Catalase	5.4	4.58
B2RPV6	Mmrn1	Multimerin-1	2.1	4.32
Q99JI6	Rap1b	Ras-related protein Rap-1b	14.7	4.00
Q80TM2	Tln1	Talin-1	1.2	4.00
P63017	Hspa8	Heat shock cognate 71 kDa protein	2.0	3.58
Q3TGH8	Stam2	Signal transducing adaptor molecule 2	3.6	3.00
Q3UNK5	Tgfb1	Transforming growth factor beta-1	5.5	3.00

[0051] 상기 선별된 10개 단백질은 inter-alpha 트립신 저해제, 중쇄2 (Itih2), 비타민 D-결합단백질 (vitamin D-binding protein, DBP), Rbp4 (retinol-binding protein 4), 카탈라아제 (Catalase, Cat), 멀티머린-1

(multimerin-1, Mmrn1), ras-관련 단백질 (ras-related protein) Rap-1b (Rap1b), 탈린-1 (talin-1, Tln1), Hspa8 (Heat shock cognate 71 kDa protein), Stam2 (signal transducing adaptor molecule 2) 및 Tgfb1 (transforming growth factor beta-1)를 포함하였다. 상기 선별한 10개 단백질에 대한 볼캐노 플로팅 (volcano plot)은 이들 단백질이 선별기준의 신뢰도를 나타내는 볼캐노 형태의 선상에 있음을 제시하였다 (도 3).

[0052] 실시예 3: 비이온 세제-융합 SDS-PAGE를 이용한 혈액 EV의 프로테옴 분석

[0053] 비이온 세제가 막단백질의 추출에 적합하다는 몇몇 보고가 있다. 본 발명자들은 비이온세제 DDM (n-dodecyl D-maltoside)을 이용하여 EV 막에 존재하는 단백질을 추출하였다. 농도구배 초원심분리로 제조된 혈액 EV를 1% DDM에 가하고, 막단백질을 포함하는 소수성 단백질을 추출하였다. 다음, 상기 추출물을 SDS-PAGE로 분석하였다. 쿠마시브루 (Coomassie Brilliant Blue) 염색에서 특이적으로 나타나는 TG6799 마우스 혈액 EV의 3개 밴드 영역을 발견하였다 (도 4). 상기 3개 밴드영역에 대해, 트립신 기재 겔-내 분해를 수행하였다. 트립신 분해 펩타이드를 사용하는 LC-MS/MS 분석으로 총 160개 단백질을 동정하였다 (결과 제시하지 않음). 이들 전체 단백질로부터, 프로테옴 파라미터에서 높은 신뢰도를 갖는 바이오마커 후보로서, 다음 8개 단백질을 선정하였다: β -parvin (PARVB), delta-aminolevulinic acid dehydratase (ALAD), guanine nucleotide-binding protein G(i) subunit alpha-2 (GNAI2), latent-transforming growth factor beta-binding protein 1 (LTBP1), L-lactate dehydrogenase (LDHA), malate dehydrogenase (MDH2), T-complex protein 1 subunit beta (CCT2) and T-complex protein 1 subunit delta (CCT4) (표 2).

표 2

[0054]

Uniprot	Gene ID	Description	Coverage	# PSMs	# Unique peptides
Q3UGT9	Parvb	Beta-parvin	9.9	3	3
P10518	Alad	Delta-aminolevulinic acid dehydratase (porphobilinogen synthase)	9.1	3	3
P08752	Gnai2	Guanine nucleotide-binding protein G(i) subunit alpha-2	7.3	2	2
Q8CG19	Ltbp1	Latent-transforming growth factor beta-binding protein 1	2.6	4	3
A0A1BOGSX0	Ldha	L-lactate dehydrogenase	15.2	5	5
P08249	Mdh2	Malate dehydrogenase, mitochondrial	15.7	4	4
Q542X7	Cct2	T-complex protein 1 subunit beta	7.7	2	2
P80315	Cct4	T-complex protein 1 subunit delta	6.1	3	3

[0055] 실시예 4: 마우스 혈액 EV로부터 스크리닝된 바이오마커 후보 단백질의 평가

[0056] AD 바이오마커로서의 가능성을 평가하기 위해, 생화학 (biochemical) 실험을 수행하였다. 상기 프로테옴 분석에서와 동일한 방법으로 혈액 EV를 제조한 후, 상기 바이오마커 후보에 대해 웨스턴블로팅을 수행하였다. Heat shock cognate 71 kDa protein, inter-alpha trypsin inhibitor, heavy chain 2, vitamin D-binding protein, talin-1, catalase 및 signal transducing adaptor molecule 2을 포함하는 6개 단백질이 AD 혈액 EV에서 뚜렷하게 증가한 것으로 확인되었다 (도 5a). 프로테옴 분석에서 감소된 것으로 확인된 페록시레독신2 (Peroxiredoxin 2, PRDX2)는 TG6799 혈액 EVs에서 확연히 감소하였다 (도 5a). 또한, 비이온 세제 융합 SDS-PAGE 방법에서 확인한 바이오마커 후보에 대해 웨스턴블로팅을 수행하였다. β -parvin, guanine nucleotide-binding protein G(i) subunit alpha-2, delta-aminolevulinic acid dehydratase, L-lactate dehydrogenase 및 T-complex protein 1 subunit delta를 포함하는 5개 단백질이 AD 혈액 EV에서 현저히 증가하였다 (도 5b). 이러한 웨스턴블로팅 결과는 상기 11개 단백질이 AD 바이오마커로서 높은 가능성이 있음을 제시한다.

[0057] 실시예 5: 단계별로 구분된 사람 AD 환자의 혈액 EV를 사용한 바이오마커 후보 단백질의 평가

[0058] 사람에서 분리한 시료를 사용하여 AD 바이오마커로서의 가능성을 평가하였다. 먼저, 정상, AD 초기단계 및 AD 후기단계의 3종류의 혈액 시료를 제조하였다. 모든 그룹은 60세 이상이었다. 상기 프로테옴 분석에서와 동일한 방법으로 혈액 EV를 제조한 후, 각 바이오마커에 대한 웨스턴블로팅을 수행하였다. 메탄올-클로로포름 방법에서 확인된 inter-alpha trypsin inhibitor, heavy chain 2, L-lactate dehydrogenase A chain, ras-related protein Rap-1b, retinol-binding protein 4, porphobilinogen synthase, Guanine nucleotide-binding

protein G(i) subunit alpha-2 를 포함하는 6개의 단백질의 수준이 AD 혈액 EV에서 뚜렷하게 증가하였다 (도 6a 및 도 6b). 이러한 결과는 상기 6개 단백질이 AD 진단에 직접 적용될 수 있는 바이오마커로서 높은 가능성이 있음을 제시한다.

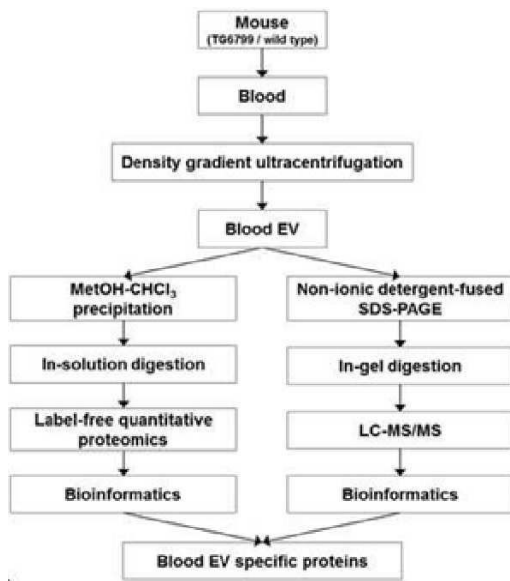
산업상 이용가능성

[0059]

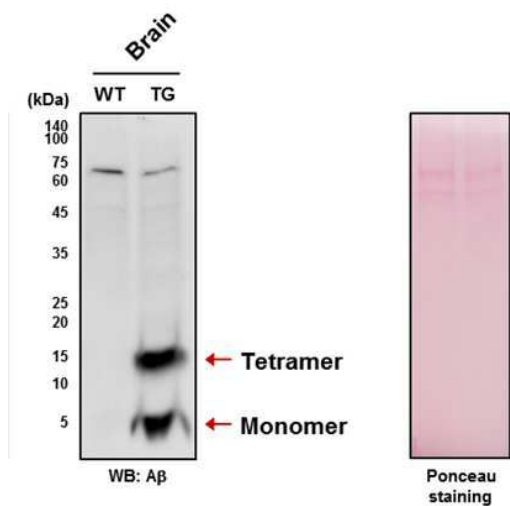
본 발명은 혈액 EV로부터 알츠하이머병 진단에 이용될 수 있는 단백질 마커를 분리한 것으로서, 알츠하이머병의 진단에 이용될 수 있다. 특히, 본 발명의 단백질 마커는 초기단계의 알츠하이머병을 진단할 수 있는 가능성을 제공한다. 따라서, 본 발명은 임상에서 알츠하이머병의 진단 및 치료에 매우 유용하게 이용될 수 있을 것이다.

도면

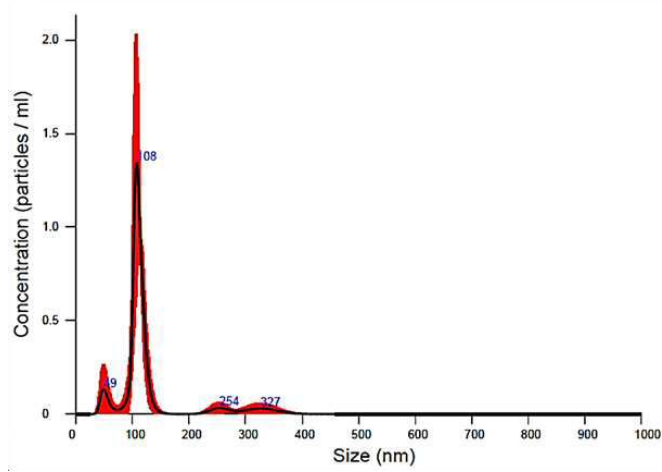
도면1a



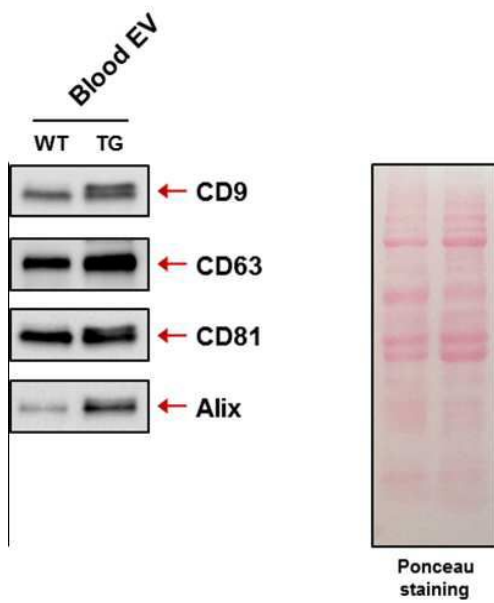
도면1b



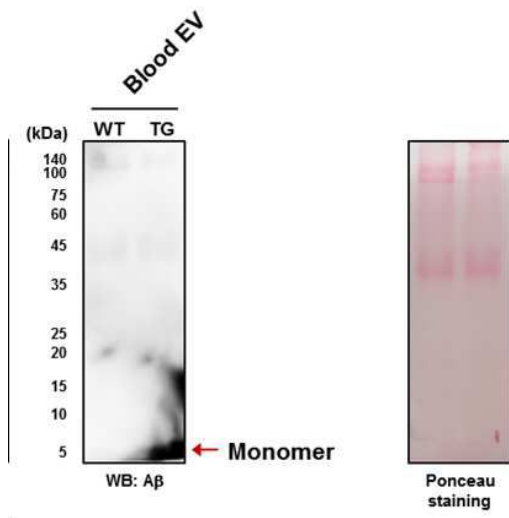
도면1c



도면1d



도면1e



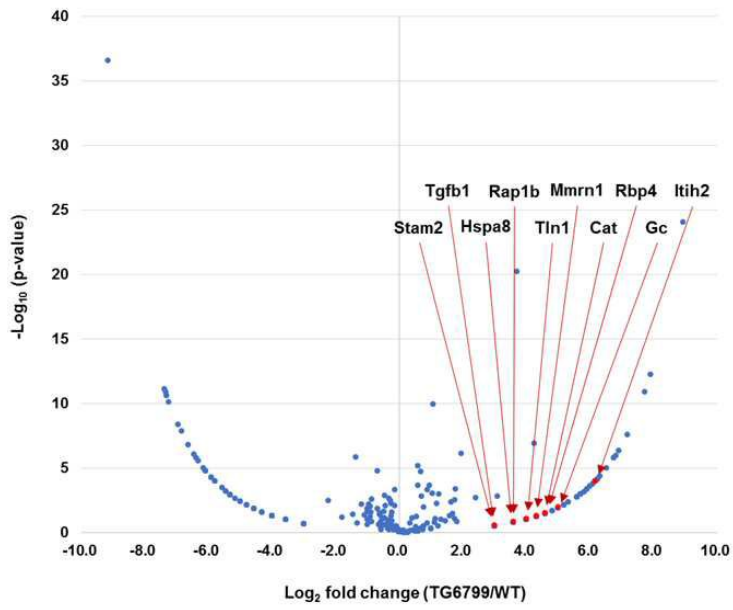
도면2a

Accession	Description	Fold change (log2)	P-value<0.05
A0A0F7R5U8	MAb 110 light chain	8.92	O
Q8R0I8	BC026782 protein	7.91	O
Q8R3H6	Ighg protein	7.73	O
A0A0F7R1P3	MAb 106 heavy chain	7.17	O
Q8BH35	Complement component C8 beta chain	6.91	O
P60710	Actin, cytoplasmic 1	6.81	O
Q92111	Serotransferrin	6.52	O
E9Q8B5	Protein Gm4788	6.32	O
A0A0G2JEU7	Protein Ighv1-82 (Fragment)	6.25	O
I6L998	Uncharacterized protein	6.25	O
F8WJ05	Inter-alpha-trypsin inhibitor heavy chain H1	6.17	O
A0A0R4J0I1	MCG1051009	6.09	O
G3X977	Inter-alpha trypsin inhibitor, heavy chain 2	5.91	O
D3YXF5	Oxidation resistance protein 1	5.91	O
P22599	Alpha-1-antitrypsin 1-2	5.81	O
Q00897	Alpha-1-antitrypsin 1-4	5.70	O
X5J4Q3	IgM heavy chain VDJ region (Fragment)	5.70	O
A0A0A0MQA3	Alpha-1-antitrypsin 1-1	5.58	O
A2NVX0	Light chain variable region (Fragment)	5.58	O
A0A0A6YXT2	Protein Ighv1-36 (Fragment)	5.32	O
Q924Q0	V165-D-J-C mu protein (Fragment)	5.32	O
A0A0R4J038	Kininogen-1	5.17	O
Q924P9	V303-D-J-C mu protein (Fragment)	5.17	O
A2NMA2	Anti-carcinoma embryonic antigen heavy chain variable region (Fragment)	5.00	O
Q566I6	Complement component 1, r subcomponent	5.00	O
U5JE43	Immunoglobulin heavy chain variable region (Fragment)	5.00	O
P21614	Vitamin D-binding protein	5.00	O
Q3TVN6	C-X-C motif chemokine	4.81	O
A0A075B696	Protein Ighv2-3 (Fragment)	4.81	O
P24270	Catalase	4.58	O
P26262	Plasma kallikrein	4.58	O
Q3TJ94	Prothrombin	4.58	O
H7BWY6	Retinol-binding protein 4	4.58	O
E9PZD8	Ceruloplasmin	4.32	O
B2RPV6	Multimerin-1	4.32	O
Q3TUI9	Proteasome subunit alpha type	4.32	O
O09061	Proteasome subunit beta type-1	4.32	O
Q8K182	Complement component C8 alpha chain	4.25	O
Q61129	Complement factor I	4.00	X
A2KTJ9	Immunoglobulin heavy chain variable region (Fragment)	4.00	X
Q7M084	Kininogen, LMW (Fragment)	4.00	X
Q07456	Protein AMBP	4.00	X
A0A0A6YWN3	Protein Ighv13-2 (Fragment)	4.00	X
A0A075B5Q4	Protein Ighv5-12 (Fragment)	4.00	X
Q99J16	Ras-related protein Rap-1b	4.00	X
Q80TM2	Talin-1	4.00	X
P28665	Murinoglobulin-1	3.70	O
Q8R0Z6	Angiotensin-related protein 6	3.58	X
Q08879	Fibulin-1	3.58	X
P63017	Heat shock cognate 71 kDa protein	3.58	X
P01680	Ig kappa chain V-IV region S107B	3.58	X
A0A075B5T1	Protein Ighv12-3 (Fragment)	3.58	X
Q8CIR9	Serine protease MASP3 (Fragment)	3.58	X
Q00898	Alpha-1-antitrypsin 1-5	3.09	O
Q8VIJ1	Anti-DNA heavy chain (Fragment)	3.00	X
E9QP56	Apolipoprotein C-III	3.00	X
O88783	Coagulation factor V	3.00	X
D3YZ61	Complement C1q tumor necrosis factor-related protein 3	3.00	X
Q8K1B8	Fermitin family homolog 3	3.00	X
A0A0A0MQF6	Glyceraldehyde-3-phosphate dehydrogenase	3.00	X
P01627	Ig kappa chain V-II region VKappa167	3.00	X
P98064	Mannan-binding lectin serine protease 1	3.00	X
Q6A0F0	MKIAA0003 protein (Fragment)	3.00	X
E9Q6C2	Protein C1s1	3.00	X
A0A140T8P3	Protein Ighv15-103 (Fragment)	3.00	X
A0A0G2JDU0	Protein Ighv8-34 (Fragment)	3.00	X
Q3U254	Putative uncharacterized protein	3.00	X
Q3TGH8	Signal transducing adaptor molecule (SH3 domain and ITAM motif) 2	3.00	X
P08032	Spectrin alpha chain, erythrocytic 1	3.00	X
Q3UNK5	Transforming growth factor beta-1	3.00	X
Q53VQ0	VH region (Fragment)	3.00	X

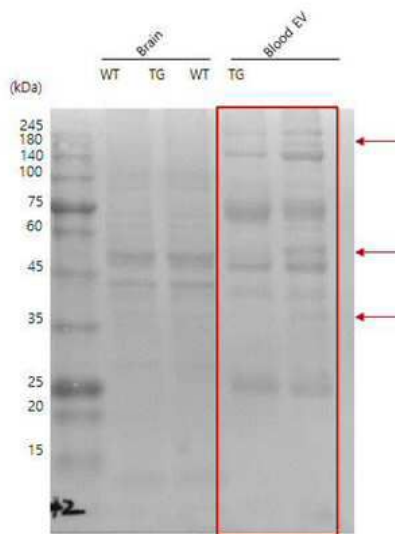
도면2b

Accession	Description	Fold change (log2)	P-value<0.05
Q66J57	Igk protein	-9.17	O
A0A0F7QZE4	MAb 110 heavy chain	-7.36	O
A2MY50	Anti-acid phosphatase variable light chain 11 (Fragment)	-7.25	O
Q19LI2	Alpha-1B-glycoprotein	-6.95	O
A0A075B5M1	Protein Igkv4-63	-6.86	O
Q65Z11	Anti-HIV-1 reverse transcriptase single-chain variable	-6.64	O
A2NN81	B9-scFv	-6.39	O
A0A075B697	Protein Ighv2-9-1 (Fragment)	-6.32	O
A0A0A6YX66	Protein Ighv1-20 (Fragment)	-6.17	O
P01667	Ig kappa chain V-III region PC 6308	-6.09	O
P01646	Ig kappa chain V-V region HP 123E6	-6.09	O
A0A075B5Y2	Protein Ighv1-75 (Fragment)	-6.09	O
A0A075B5T5	Protein Ighv1-5	-5.91	O
A0A075B5W6	Protein Ighv1-55 (Fragment)	-5.91	O
A0A0A6YVS4	Protein Ighv5-9-1 (Fragment)	-5.91	O
A0A075B5N7	Protein Igkv6-13	-5.91	O
A0A075B5W1	Protein Ighv1-50 (Fragment)	-5.81	O
A0A075B5N4	Protein Igkv8-27	-5.81	O
A0A075B5P8	Protein Ighv2-2 (Fragment)	-5.58	O
Q9EQI5	Chemokine (C-X-C motif) ligand 7, isoform CRA_b	-5.46	O
A0A0B4J1I4	Protein Igkv4-72 (Fragment)	-5.46	O
A0A0R4J0X5	Alpha-1-antitrypsin 1-3	-5.32	O
A0A0M5KS39	B2 18-5 anti-human butyrylcholinesterase (BChE) light chain variable r egion (Fragment)	-5.32	O
P01749	Ig heavy chain V region 3	-5.32	O
A0A0G2JGQ8	Ig lambda-3 chain C region (Fragment)	-5.32	O
A0A075B5Y3	Protein Ighv1-80 (Fragment)	-5.32	O
A0A140T8M4	Protein Igkv8-19	-5.32	O
A0A075B677	Protein Igkv4-53	-5.17	O
A0A0B4J1J4	Protein Ighv9-2 (Fragment)	-5.00	O
A0A075B5M3	Protein Igkv4-58	-5.00	O
X5J5T5	IgA heavy chain VDJ region (Fragment)	-4.81	O
A0A0A6YY69	Protein Ighv2-5 (Fragment)	-4.81	O
A0A140T8N1	Protein Igkv1-88 (Fragment)	-4.81	O
E9Q035	Protein Gm20425	-4.58	O
A0A075B5X2	Protein Ighv1-63 (Fragment)	-4.58	O
A0A0B4J1P3	Protein Ighv9-1 (Fragment)	-4.58	O
A0A075B5L7	Protein Igkv4-80 (Fragment)	-4.58	O
Q61408	Complement factor H-related protein	-4.32	O
K7TH85	IgA heavy chain variable region (Fragment)	-4.32	O
Q61750	JH-Cdelta locus (Fragment)	-4.32	O
Q6GQT1	Alpha-2-macroglobulin-P	-4.00	O
Q9Z1C6	Anti-porcine VCAM mAb 2A2 heavy chain variable region (Fragment)	-4.00	O
A0A0R4J0M6	Collectin-11	-4.00	O
Q99LC4	Igh protein	-4.00	O
K7TRE7	IgM heavy chain variable region (Fragment)	-4.00	O
Q5M9N9	Prdx2 protein	-4.00	O
A0A140T8M9	Protein Igkv1-122 (Fragment)	-4.00	O
A0A0G2JDG9	Protein Igkv8-16 (Fragment)	-4.00	O
X5J4T6	IgG1 heavy chain VDJ region (Fragment)	-3.58	X
P28666	Murinoglobulin-2	-3.58	X
A0A0G2JFN9	Protein Ighv1-67 (Fragment)	-3.58	X
C5H0E8	Rap1A-retro1	-3.58	X
Q61852	Smooth muscle gamma-actin	-3.58	X
A2AAE1	Uncharacterized protein KIAA1109	-3.58	X
Q8VDC9	Anti-MOG Z12 variable gamma 2a (Fragment)	-3.00	X
B2RXX8	Cdc42bpa protein	-3.00	X
E9Q0B5	Protein Fcgbp	-3.00	X
A0A0B4J1J0	Protein Igkv4-50 (Fragment)	-3.00	X
E9QQ25	Striated muscle-specific serine/threonine-protein kinase	-3.00	X
B9EKU4	URB2 ribosome biogenesis 2 homolog (S. cerevisiae)	-3.00	X

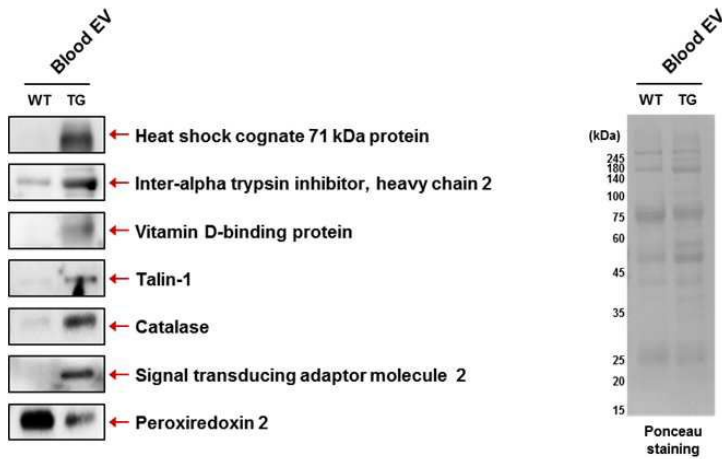
도면3



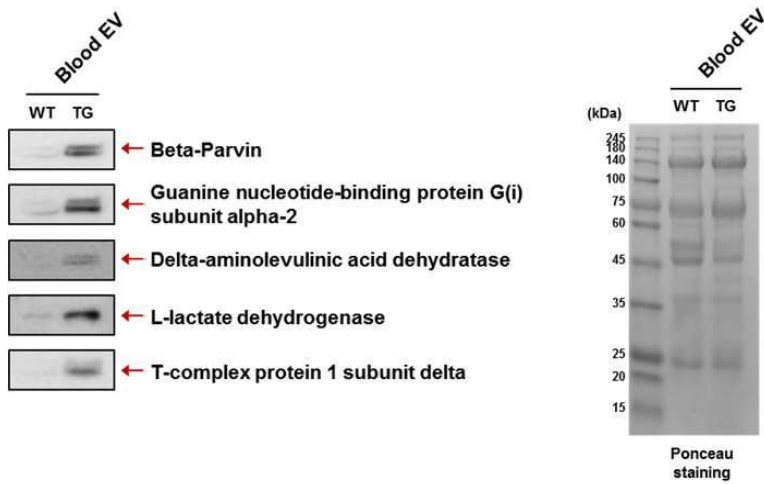
도면4



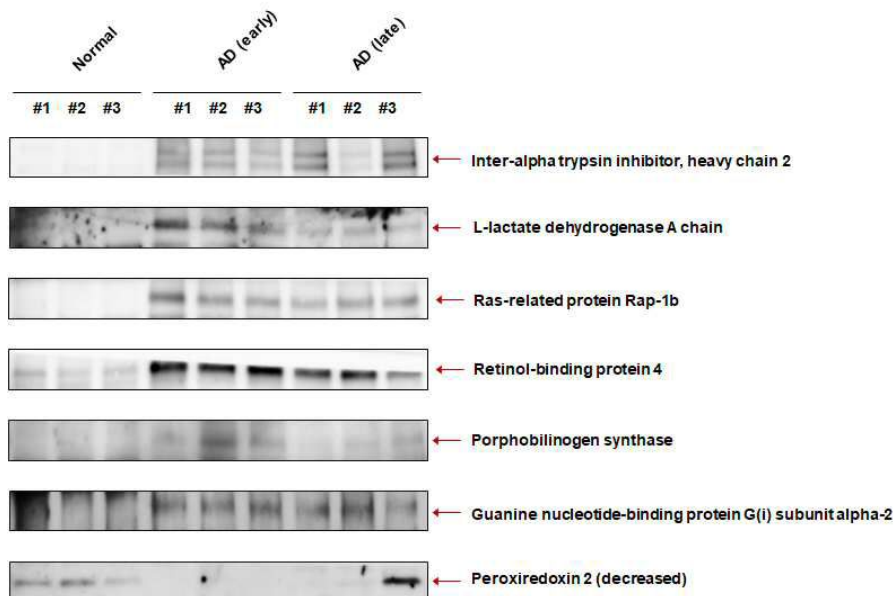
도면5a



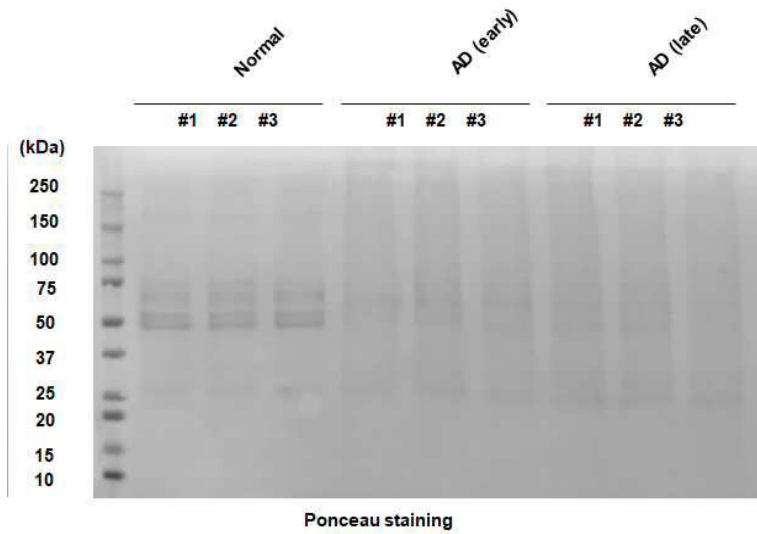
도면5b



도면6a



도면6b



서열 목록

<110> Daegu Gyeongbuk Institute of Science and Technology
 <120> A novel protein marker for diagnosing Alzheimer's disease and the use thereof
 <130> IKR0117P03KR
 <160> 24
 <170> KoPatentIn 3.0
 <210> 1
 <211> 646
 <212> PRT
 <213> Artificial Sequence
 <220><223> Heat shock cognate 71 kDa protein
 <400> 1
 Met Ser Lys Gly Pro Ala Val Gly Ile Asp Leu Gly Thr Thr Tyr Ser
 1 5 10 15
 Cys Val Gly Val Phe Gln His Gly Lys Val Glu Ile Ile Ala Asn Asp
 20 25 30
 Gln Gly Asn Arg Thr Thr Pro Ser Tyr Val Ala Phe Thr Asp Thr Glu
 35 40 45
 Arg Leu Ile Gly Asp Ala Ala Lys Asn Gln Val Ala Met Asn Pro Thr
 50 55 60
 Asn Thr Val Phe Asp Ala Lys Arg Leu Ile Gly Arg Arg Phe Asp Asp

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

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

 450 455 460
 Pro Pro Ala Pro Arg Gly Val Pro Gln Ile Glu Val Thr Phe Asp Ile
 465 470 475 480
 Asp Ala Asn Gly Ile Leu Asn Val Ser Ala Val Asp Lys Ser Thr Gly
 485 490 495
 Lys Glu Asn Lys Ile Thr Ile Thr Asn Asp Lys Gly Arg Leu Ser Lys
 500 505 510
 Glu Asp Ile Glu Arg Met Val Gln Glu Ala Glu Lys Tyr Lys Ala Glu
 515 520 525

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

Ile Ile Ser Trp Leu Asp Lys Asn Gln Thr Ala Glu Lys Glu Glu Phe
 580 585 590

Glu His Gln Gln Lys Glu Leu Glu Lys Val Cys Asn Pro Ile Ile Thr
 595 600 605

Lys Leu Tyr Gln Ser Ala Gly Gly Met Pro Gly Gly Met Pro Gly Gly
 610 615 620

Phe Pro Gly Gly Gly Ala Pro Pro Ser Gly Gly Ala Ser Ser Gly Pro
 625 630 635 640

Thr Ile Glu Glu Val Asp
 645

<210> 2
 <211> 950
 <212> PRT
 <213> Artificial Sequence
 <220><223> Inter-alpha trypsin inhibitor, heavy chain 2
 <400> 2

Met Ser Ser Lys Met Gln Arg Pro Val Cys Leu Leu Ile Trp Leu Phe
 1 5 10 15

Leu Leu Glu Ala Gln Ala Phe Glu Ile Pro Ile Asn Gly Asn Ser Glu
 20 25 30

Phe Ala Glu Tyr Ser Asp Leu Val Glu Leu Ala Pro Asp Lys Leu Pro
 35 40 45

Phe Val Gln Glu Asn Gly Arg His Gln Arg Ser Leu Pro Glu Glu Ser
 50 55 60

Gly Glu Glu Thr Asp Thr Val Asp Pro Val Thr Leu Tyr Ser Tyr Lys
 65 70 75 80

Val Gln Ser Thr Ile Thr Ser Arg Val Ala Thr Thr Thr Ile Gln Ser
 85 90 95

Lys Leu Val Asn Asn Ser Pro Leu Pro Gln Ser Val Val Phe Asp Val
 100 105 110

Gln Ile Pro Lys Gly Ala Phe Ile Ser Asn Phe Thr Met Thr Val Asn
 115 120 125

Gly Met Thr Phe Thr Ser Ser Ile Lys Glu Lys Thr Val Gly Arg Ala
 130 135 140
 Leu Tyr Ser Gln Ala Arg Ala Lys Gly Lys Thr Ala Gly Trp Val Arg

 145 150 155 160
 Ser Arg Thr Leu Asp Met Glu Asn Phe Asn Thr Glu Val Asn Ile Pro
 165 170 175
 Pro Gly Ala Lys Val Gln Phe Glu Leu His Tyr Gln Glu Val Lys Trp
 180 185 190
 Arg Lys Leu Gly Ser Tyr Glu His Lys Ile His Leu Gln Pro Gly Lys
 195 200 205
 Leu Ala Lys His Leu Glu Val Asn Val Trp Ile Ile Glu Pro Gln Gly
 210 215 220

 Met Arg Phe Leu His Val Pro Asp Thr Phe Glu Gly His Phe Gln Gly
 225 230 235 240
 Val Pro Val Ile Ser Lys Gly Gln Gln Lys Ala His Val Ser Phe Lys
 245 250 255
 Pro Thr Val Ala Gln Gln Arg Lys Cys Pro Asn Cys Thr Glu Thr Ala
 260 265 270
 Val Asn Gly Glu Leu Val Val Met Tyr Asp Val Asn Arg Glu Glu Lys
 275 280 285
 Ala Gly Glu Leu Glu Val Phe Asn Gly Tyr Phe Val His Phe Phe Ala

 290 295 300
 Pro Glu Asn Leu Asp Pro Ile Pro Lys Asn Ile Leu Phe Val Ile Asp
 305 310 315 320
 Val Ser Gly Ser Met Trp Gly Ile Lys Met Lys Gln Thr Val Glu Ala
 325 330 335
 Met Lys Thr Ile Leu Asp Asp Leu Arg Thr Asp Asp Gln Phe Ser Val
 340 345 350
 Val Asp Phe Asn His Asn Val Arg Thr Trp Arg Asn Asp Leu Val Ser
 355 360 365

 Ala Thr Lys Thr Gln Ile Ala Asp Ala Lys Arg Tyr Ile Glu Lys Ile

Gly Lys Glu Pro Gly Lys Pro Glu Ala Ser Met Glu Val Lys Gly His
 885 890 895
 Lys Leu Thr Val Thr Arg Gly Leu Gln Lys Asp Tyr Arg Thr Asp Ile
 900 905 910
 Val Phe Gly Thr Asp Val Pro Cys Trp Phe Val His Asn Ser Gly Lys
 915 920 925
 Gly Phe Ile Asp Gly His Tyr Lys Asp Tyr Phe Val Pro Gln Leu Tyr
 930 935 940

Ser Phe Leu Lys Arg Pro
 945 950

<210> 3

<211> 476

<212> PRT

<213> Artificial Sequence

<220><223> Vitamin D-binding protein

<400> 3

Met Lys Arg Val Leu Val Leu Leu Leu Ala Leu Ala Phe Gly His Ala
 1 5 10 15
 Leu Glu Arg Gly Arg Asp Tyr Glu Lys Asp Lys Val Cys Asn Glu Leu
 20 25 30
 Ala Met Leu Gly Lys Glu Asp Phe Arg Ser Leu Ser Leu Ile Leu Tyr
 35 40 45

Ser Arg Lys Phe Ser Ser Ser Thr Phe Glu Gln Val Asn Gln Leu Val
 50 55 60

Lys Glu Val Val Ser Leu Thr Glu Glu Cys Cys Ala Glu Gly Ala Asp
 65 70 75 80

Pro Thr Cys Tyr Asp Thr Arg Thr Ser Glu Leu Ser Val Lys Ser Cys
 85 90 95

Glu Ser Asp Ala Pro Phe Pro Val His Pro Gly Thr Pro Glu Cys Cys
 100 105 110

Thr Lys Glu Gly Leu Glu Arg Lys Leu Cys Met Ala Ala Leu Ser His
 115 120 125

Gln Pro Gln Glu Phe Pro Thr Tyr Val Glu Pro Thr Asn Asp Glu Ile
 130 135 140

Cys Glu Ala Phe Arg Arg Asp Pro Lys Gly Phe Ala Asp Gln Phe Leu
 145 150 155 160

Tyr Glu Tyr Ser Ser Asn Tyr Gly Gln Ala Pro Leu Pro Leu Leu Val
 165 170 175

Ala Tyr Thr Lys Asn Tyr Leu Ser Met Val Gly Ser Cys Cys Thr Ser
 180 185 190

Ala Asn Pro Thr Val Cys Phe Val Lys Glu Arg Leu Gln Met Lys His
 195 200 205

Leu Ser Leu Leu Thr Thr Met Ser Asn Arg Val Cys Ser Gln Tyr Ala
 210 215 220

Ala Tyr Gly Lys Glu Lys Ser Arg Leu Ser His Leu Ile Lys Leu Ala
 225 230 235 240

Gln Lys Val Pro Thr Ala Asn Leu Glu Asn Val Leu Pro Leu Ala Glu
 245 250 255

Asp Phe Thr Glu Ile Leu Ser Arg Cys Cys Glu Ser Thr Ser Glu Asp
 260 265 270

Cys Met Ala Ser Glu Leu Pro Glu His Thr Ile Lys Ile Cys Gln Asn
 275 280 285

Leu Ser Lys Lys Asn Ser Lys Phe Glu Glu Cys Cys Gln Glu Asn Thr
 290 295 300

Pro Met Asn Ile Phe Met Cys Thr Tyr Phe Met Pro Ala Ala Glu Pro
 305 310 315 320

Leu Gln Leu Pro Ala Ile Lys Leu Pro Thr Gly Lys Asp Leu Cys Gly
 325 330 335

Gln Ser Thr Thr Gln Ala Met Asp Gln Tyr Thr Phe Glu Leu Ser Arg
 340 345 350

Arg Thr Gln Val Pro Glu Val Phe Leu Ser Lys Val Leu Glu Pro Thr
 355 360 365

Leu Lys Thr Leu Arg Glu Cys Cys Asp Thr Gln Asp Ser Val Ala Cys
 370 375 380

Phe Ser Thr Gln Ser Pro Leu Leu Lys Arg Gln Leu Thr Ser Phe Ile
 385 390 395 400
 Glu Lys Gly Gln Glu Met Cys Ala Asp Tyr Ser Glu Asn Thr Phe Thr
 405 410 415
 Glu Tyr Lys Lys Lys Leu Ala Glu Arg Leu Arg Thr Lys Thr Pro Asn
 420 425 430
 Thr Ser Pro Ala Glu Leu Lys Asp Met Val Glu Lys His Ser Asp Phe
 435 440 445
 Ala Ser Lys Cys Cys Ser Ile Asn Ser Pro Pro Leu Tyr Cys Ser Ser
 450 455 460
 Gln Ile Asp Ala Glu Met Ile Asp Thr Leu Gln Ser
 465 470 475

<210> 4
 <211> 2541
 <212> PRT
 <213> Artificial Sequence
 <220><223> Talin-1
 <400> 4

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

Met Thr Ile Cys Ala Arg Ile Gly Ile Thr Asn His Asp Glu Tyr Ser
 115 120 125

Leu Val Arg Glu Leu Met Glu Glu Lys Lys Asp Glu Gly Thr Gly Thr
 130 135 140

Leu Arg Lys Asp Lys Thr Leu Leu Arg Asp Glu Lys Lys Met Glu Lys
 145 150 155 160

Leu Lys Gln Lys Leu His Thr Asp Asp Glu Leu Asn Trp Leu Asp His
 165 170 175

Gly Arg Thr Leu Arg Glu Gln Gly Val Glu Glu His Glu Thr Leu Leu
 180 185 190

Leu Arg Arg Lys Phe Phe Tyr Ser Asp Gln Asn Val Asp Ser Arg Asp
 195 200 205

Pro Val Gln Leu Asn Leu Leu Tyr Val Gln Ala Arg Asp Asp Ile Leu
 210 215 220

Asn Gly Ser His Pro Val Ser Phe Asp Lys Ala Cys Glu Phe Ala Gly
 225 230 235 240

Phe Gln Cys Gln Ile Gln Phe Gly Pro His Asn Glu Gln Lys His Lys
 245 250 255

Ala Gly Phe Leu Asp Leu Lys Asp Phe Leu Pro Lys Glu Tyr Val Lys
 260 265 270

Gln Lys Gly Glu Arg Lys Ile Phe Gln Ala His Lys Asn Cys Gly Gln
 275 280 285

Met Ser Glu Ile Glu Ala Lys Val Arg Tyr Val Lys Leu Ala Arg Ser
 290 295 300

Leu Lys Thr Tyr Gly Val Ser Phe Phe Leu Val Lys Glu Lys Met Lys
 305 310 315 320

Gly Lys Asn Lys Leu Val Pro Arg Leu Leu Gly Ile Thr Lys Glu Cys
 325 330 335

Val Met Arg Val Asp Glu Lys Thr Lys Glu Val Ile Gln Glu Trp Ser
 340 345 350

Leu Thr Asn Ile Lys Arg Trp Ala Ala Ser Pro Lys Ser Phe Thr Leu

Asp Ala Thr Ala Lys Met Val Glu Ala Ala Lys Gly Ala Ala Ala His
 865 870 875 880
 Pro Asp Ser Glu Glu Gln Gln Gln Arg Leu Arg Glu Ala Ala Glu Gly
 885 890 895
 Leu Arg Met Ala Thr Asn Ala Ala Ala Gln Asn Ala Ile Lys Lys Lys
 900 905 910
 Leu Val Gln Arg Leu Glu His Ala Ala Lys Gln Ala Ala Ala Ser Ala
 915 920 925
 Thr Gln Thr Ile Ala Ala Ala Gln His Ala Ala Ser Ala Pro Lys Ala
 930 935 940
 Ser Ala Gly Pro Gln Pro Leu Leu Val Gln Ser Cys Lys Ala Val Ala
 945 950 955 960
 Glu Gln Ile Pro Leu Leu Val Gln Gly Val Arg Gly Ser Gln Ala Gln
 965 970 975
 Pro Asp Ser Pro Ser Ala Gln Leu Ala Leu Ile Ala Ala Ser Gln Ser
 980 985 990

 Phe Leu Gln Pro Gly Gly Lys Met Val Ala Ala Ala Lys Ala Ser Val
 995 1000 1005
 Pro Thr Ile Gln Asp Gln Ala Ser Ala Met Gln Leu Ser Gln Cys Ala
 1010 1015 1020
 Lys Asn Leu Gly Thr Ala Leu Ala Glu Leu Arg Thr Ala Ala Gln Lys
 1025 1030 1035 1040
 Ala Gln Glu Ala Cys Gly Pro Leu Glu Met Asp Ser Ala Leu Ser Val
 1045 1050 1055
 Val Gln Asn Leu Glu Lys Asp Leu Gln Glu Ile Lys Ala Ala Ala Arg
 1060 1065 1070
 Asp Gly Lys Leu Lys Pro Leu Pro Gly Glu Thr Met Glu Lys Cys Thr
 1075 1080 1085
 Gln Asp Leu Gly Asn Ser Thr Lys Ala Val Ser Ser Ala Ile Ala Lys
 1090 1095 1100
 Leu Leu Gly Glu Ile Ala Gln Gly Asn Glu Asn Tyr Ala Gly Ile Ala
 1105 1110 1115 1120

Ala Arg Asp Val Ala Gly Gly Leu Arg Ser Leu Ala Gln Ala Ala Arg
 1125 1130 1135

Gly Val Ala Ala Leu Thr Ser Asp Pro Ala Val Gln Ala Ile Val Leu
 1140 1145 1150

Asp Thr Ala Ser Asp Val Leu Asp Lys Ala Ser Ser Leu Ile Glu Glu
 1155 1160 1165

Ala Lys Lys Ala Ser Gly His Pro Gly Asp Pro Glu Ser Gln Gln Arg
 1170 1175 1180

Leu Ala Gln Val Ala Lys Ala Val Thr Gln Ala Leu Asn Arg Cys Val
 1185 1190 1195 1200

Ser Cys Leu Pro Gly Gln Arg Asp Val Asp Asn Ala Leu Arg Ala Val
 1205 1210 1215

Gly Asp Ala Ser Lys Arg Leu Leu Ser Asp Ser Leu Pro Pro Ser Thr
 1220 1225 1230

Gly Thr Phe Gln Glu Ala Gln Ser Arg Leu Asn Glu Ala Ala Ala Gly
 1235 1240 1245

Leu Asn Gln Ala Ala Thr Glu Leu Val Gln Ala Ser Arg Gly Thr Pro
 1250 1255 1260

Gln Asp Leu Ala Arg Ala Ser Gly Arg Phe Gly Gln Asp Phe Ser Thr
 1265 1270 1275 1280

Phe Leu Glu Ala Gly Val Glu Met Ala Gly Gln Ala Pro Ser Gln Glu
 1285 1290 1295

Asp Arg Ala Gln Val Val Ser Asn Leu Lys Gly Ile Ser Met Ser Ser
 1300 1305 1310

Ser Lys Leu Leu Leu Ala Ala Lys Ala Leu Ser Thr Asp Pro Ala Ser
 1315 1320 1325

Pro Asn Leu Lys Ser Gln Leu Ala Ala Ala Ala Arg Ala Val Thr Asp
 1330 1335 1340

Ser Ile Asn Gln Leu Ile Thr Met Cys Thr Gln Gln Ala Pro Gly Gln
 1345 1350 1355 1360

Lys Glu Cys Asp Asn Ala Leu Arg Gln Leu Glu Thr Val Arg Glu Leu

	1365	1370	1375
Leu Glu Asn Pro Val Gln Pro Ile Asn Asp Met Ser Tyr Phe Gly Cys			
	1380	1385	1390
Leu Asp Ser Val Met Glu Asn Ser Lys Val Leu Gly Glu Ala Met Thr			
	1395	1400	1405
Gly Ile Ser Gln Asn Ala Lys Asn Gly Asn Leu Pro Glu Phe Gly Asp			
	1410	1415	1420
Ala Ile Ala Thr Ala Ser Lys Ala Leu Cys Gly Phe Thr Glu Ala Ala			
1425	1430	1435	1440
Ala Gln Ala Ala Tyr Leu Val Gly Val Ser Asp Pro Asn Ser Gln Ala			
	1445	1450	1455
Gly Gln Gln Gly Leu Val Glu Pro Thr Gln Phe Ala Arg Ala Asn Gln			
	1460	1465	1470
Ala Ile Gln Met Ala Cys Gln Ser Leu Gly Glu Pro Gly Cys Thr Gln			
	1475	1480	1485
Ala Gln Val Leu Ser Ala Ala Thr Ile Val Ala Lys His Thr Ser Ala			
	1490	1495	1500
Leu Cys Asn Ser Cys Arg Leu Ala Ser Ala Arg Thr Ala Asn Pro Thr			
1505	1510	1515	1520
Ala Lys Arg Gln Phe Val Gln Ser Ala Lys Glu Val Ala Asn Ser Thr			
	1525	1530	1535
Ala Asn Leu Val Lys Thr Ile Lys Ala Leu Asp Gly Asp Phe Thr Glu			
	1540	1545	1550
Glu Asn Arg Ala Gln Cys Arg Ala Ala Thr Ala Pro Leu Leu Glu Ala			
	1555	1560	1565
Val Asp Asn Leu Ser Ala Phe Ala Ser Asn Pro Glu Phe Ser Ser Val			
	1570	1575	1580
Pro Ala Gln Ile Ser Pro Glu Gly Arg Ala Ala Met Glu Pro Ile Val			
1585	1590	1595	1600
Ile Ser Ala Lys Thr Met Leu Glu Ser Ala Gly Gly Leu Ile Gln Thr			
	1605	1610	1615
Ala Arg Ala Leu Ala Val Asn Pro Arg Asp Pro Pro Arg Trp Ser Val			

1620 1625 1630
 Leu Ala Gly His Ser Arg Thr Val Ser Asp Ser Ile Lys Lys Leu Ile

1635 1640 1645
 Thr Ser Met Arg Asp Lys Ala Pro Gly Gln Leu Glu Cys Glu Thr Ala

1650 1655 1660
 Ile Ala Ala Leu Asn Ser Cys Leu Arg Asp Leu Asp Gln Ala Ser Leu

1665 1670 1675 1680
 Ala Ala Val Ser Gln Gln Leu Ala Pro Arg Glu Gly Ile Ser Gln Glu

1685 1690 1695
 Ala Leu His Thr Gln Met Leu Thr Ala Val Gln Glu Ile Ser His Leu

1700 1705 1710

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

1715 1720 1725

His Lys Val Ser Gln Met Ala Gln Tyr Phe Glu Pro Leu Thr Leu Ala

1730 1735 1740

Ala Val Gly Ala Ala Ser Lys Thr Leu Ser His Pro Gln Gln Met Ala

1745 1750 1755 1760

Leu Leu Asp Gln Thr Lys Thr Leu Ala Glu Ser Ala Leu Gln Leu Leu

1765 1770 1775

Tyr Thr Ala Lys Glu Ala Gly Gly Asn Pro Lys Gln Ala Ala His Thr

1780 1785 1790

Gln Glu Ala Leu Glu Glu Ala Val Gln Met Met Thr Glu Ala Val Glu

1795 1800 1805

Asp Leu Thr Thr Thr Leu Asn Glu Ala Ala Ser Ala Ala Gly Val Val

1810 1815 1820

Gly Gly Met Val Asp Ser Ile Thr Gln Ala Ile Asn Gln Leu Asp Glu

1825 1830 1835 1840

Gly Pro Met Gly Asp Pro Glu Gly Ser Phe Val Asp Tyr Gln Thr Thr

1845 1850 1855

Met Val Arg Thr Ala Lys Ala Ile Ala Val Thr Val Gln Glu Met Val

1860 1865 1870

Thr Lys Ser Asn Thr Ser Pro Glu Glu Leu Gly Pro Leu Ala Asn Gln
 1875 1880 1885
 Leu Thr Ser Asp Tyr Gly Arg Leu Ala Ser Gln Ala Lys Pro Ala Ala
 1890 1895 1900
 Val Ala Ala Glu Asn Glu Glu Ile Gly Ala His Ile Lys His Arg Val
 1905 1910 1915 1920
 Gln Glu Leu Gly His Gly Cys Ser Ala Leu Val Thr Lys Ala Gly Ala
 1925 1930 1935
 Leu Gln Cys Ser Pro Ser Asp Val Tyr Thr Lys Lys Glu Leu Ile Glu
 1940 1945 1950
 Cys Ala Arg Arg Val Ser Glu Lys Val Ser His Val Leu Ala Ala Leu
 1955 1960 1965
 Gln Ala Gly Asn Arg Gly Thr Gln Ala Cys Ile Thr Ala Ala Ser Ala
 1970 1975 1980
 Val Ser Gly Ile Ile Ala Asp Leu Asp Thr Thr Ile Met Phe Ala Thr
 1985 1990 1995 2000
 Ala Gly Thr Leu Asn Arg Glu Gly Ala Glu Thr Phe Ala Asp His Arg
 2005 2010 2015
 Glu Gly Ile Leu Lys Thr Ala Lys Val Leu Val Glu Asp Thr Lys Val
 2020 2025 2030
 Leu Val Gln Asn Ala Ala Gly Ser Gln Glu Lys Leu Ala Gln Ala Ala
 2035 2040 2045
 Gln Ser Ser Val Ala Thr Ile Thr Arg Leu Ala Asp Val Val Lys Leu
 2050 2055 2060
 Gly Ala Ala Ser Leu Gly Ala Glu Asp Pro Glu Thr Gln Val Val Leu
 2065 2070 2075 2080
 Ile Asn Ala Val Lys Asp Val Ala Lys Ala Leu Gly Asp Leu Ile Ser
 2085 2090 2095
 Ala Thr Lys Ala Ala Ala Gly Lys Val Gly Asp Asp Pro Ala Val Trp
 2100 2105 2110
 Gln Leu Lys Asn Ser Ala Lys Val Met Val Thr Asn Val Thr Ser Leu
 2115 2120 2125

2370 2375 2380
 Asp Asp Gly Gln Trp Ser Gln Gly Leu Ile Ser Ala Ala Arg Met Val
 2385 2390 2395 2400
 Ala Ala Ala Thr Asn Asn Leu Cys Glu Ala Ala Asn Ala Ala Val Gln
 2405 2410 2415
 Gly His Ala Ser Gln Glu Lys Leu Ile Ser Ser Ala Lys Gln Val Ala
 2420 2425 2430

 Ala Ser Thr Ala Gln Leu Leu Val Ala Cys Lys Val Lys Ala Asp Gln
 2435 2440 2445
 Asp Ser Glu Ala Met Lys Arg Leu Gln Ala Ala Gly Asn Ala Val Lys
 2450 2455 2460
 Arg Ala Ser Asp Asn Leu Val Lys Ala Ala Gln Lys Ala Ala Ala Phe
 2465 2470 2475 2480
 Glu Asp Gln Glu Asn Glu Thr Val Val Val Lys Glu Lys Met Val Gly
 2485 2490 2495
 Gly Ile Ala Gln Ile Ile Ala Ala Gln Glu Glu Met Leu Arg Lys Glu

 2500 2505 2510
 Arg Glu Leu Glu Glu Ala Arg Lys Lys Leu Ala Gln Ile Arg Gln Gln
 2515 2520 2525
 Gln Tyr Lys Phe Leu Pro Ser Glu Leu Arg Asp Glu His
 2530 2535 2540

 <210> 5
 <211> 527
 <212> PRT
 <213> Artificial Sequence
 <220><223> Catalase
 <400> 5
 Met Ser Asp Ser Arg Asp Pro Ala Ser Asp Gln Met Lys Gln Trp Lys
 1 5 10 15

 Glu Gln Arg Ala Ser Gln Arg Pro Asp Val Leu Thr Thr Gly Gly Gly
 20 25 30
 Asn Pro Ile Gly Asp Lys Leu Asn Ile Met Thr Ala Gly Ser Arg Gly

<212> PRT

<213> Artificial Sequence

<220><223> Signal transducing adaptor molecule 2

<400> 6

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

His Gly Glu Leu Ile Thr Val Leu Asp Asp Ser Asp Ala Asn Trp Trp
 225 230 235 240
 Gln Gly Glu Asn His Arg Gly Thr Gly Leu Phe Pro Ser Asn Phe Val
 245 250 255

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

 325 330 335
 Leu Glu Glu Ile Asp Arg Lys His Ser Glu Leu Ser Glu Leu Asn Val
 340 345 350
 Lys Val Leu Glu Ala Leu Asp Leu Tyr Asn Lys Leu Val Asn Glu Ala
 355 360 365
 Pro Val Tyr Ser Val Tyr Ser Lys Leu His Pro Ala His Tyr Pro Pro
 370 375 380
 Ala Ala Ala Gly Val Pro Val Gln Thr Tyr Pro Val Gln Ser His Gly
 385 390 395 400

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

465 470 475 480
 Met Ser Thr Asp Val Ser Ser Phe Gln Asn Thr Ala Ser Gly Leu Pro
 485 490 495
 Gln Leu Ala Gly Phe Pro Val Ala Val Pro Ala Pro Val Ala Ala Gln
 500 505 510
 Pro Gln Ala Ser Tyr His Gln Gln Pro Leu Leu
 515 520
 <210> 7
 <211> 365
 <212> PRT
 <213> Artificial Sequence
 <220><223> Beta-parvin
 <
 400> 7
 Met Ser Ser Ala Pro Pro Arg Ser Pro Thr Pro Arg Ala Pro Lys Met
 1 5 10 15
 Lys Lys Asp Glu Ser Phe Leu Gly Lys Leu Gly Gly Thr Leu Ala Arg
 20 25 30
 Lys Lys Lys Thr Arg Glu Val Thr Asp Leu Gln Glu Glu Gly Lys Ser
 35 40 45
 Ala Ile Asn Ser Pro Met Ala Pro Ala Leu Val Asp Ile His Pro Glu
 50 55 60
 Asp Thr Gln Leu Glu Glu Asn Glu Glu Arg Thr Met Ile Asp Pro Thr

 65 70 75 80
 Ser Arg Glu Asp Pro Lys Phe Lys Glu Leu Val Lys Val Leu Leu Asp
 85 90 95
 Trp Ile Asn Asp Val Leu Ala Glu Glu Arg Ile Ile Val Lys Gln Leu
 100 105 110
 Glu Glu Asp Leu Tyr Asp Gly Gln Val Leu Gln Lys Leu Leu Glu Lys
 115 120 125
 Leu Ala His Cys Lys Leu Asn Val Ala Glu Val Thr Gln Ser Glu Ile
 130 135 140

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

145 150 155 160
 Leu Arg Pro His Gly Trp Pro Leu Arg Trp Asn Val Asp Ser Ile His
 165 170 175
 Gly Lys Asn Leu Val Ala Ile Leu His Leu Leu Val Ser Leu Ala Met
 180 185 190
 His Phe Arg Ala Pro Ile His Leu Pro Glu His Val Thr Val Gln Val
 195 200 205
 Val Val Val Arg Lys Arg Glu Gly Leu Leu His Ser Ser His Ile Ser

 210 215 220
 Glu Glu Leu Thr Thr Thr Thr Glu Ile Met Met Gly Arg Phe Glu Arg
 225 230 235 240
 Asp Ala Phe Asp Thr Leu Phe Asp His Ala Pro Asp Lys Leu Asn Leu
 245 250 255
 Val Lys Lys Ser Leu Ile Thr Phe Val Asn Lys His Leu Asn Lys Leu
 260 265 270
 Asn Leu Glu Val Thr Asp Leu Glu Thr Gln Phe Ala Asp Gly Val Tyr
 275 280 285

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

 355 360 365

- <210> 8
- <211> 355
- <212> PRT
- <213> Artificial Sequence
- <220><223> Guanine nucleotide-binding protein G(i) subunit alpha-2

<400> 8

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

 Ala Ala Glu Glu Gln Gly Met Leu Pro Glu Asp Leu Ser Gly Val Ile
 115 120 125
 Arg Arg Leu Trp Ala Asp His Gly Val Gln Ala Cys Phe Gly Arg Ser
 130 135 140
 Arg Glu Tyr Gln Leu Asn Asp Ser Ala Ala Tyr Tyr Leu Asn Asp Leu
 145 150 155 160
 Glu Arg Ile Ala Gln Ser Asp Tyr Ile Pro Thr Gln Gln Asp Val Leu
 165 170 175
 Arg Thr Arg Val Lys Thr Thr Gly Ile Val Glu Thr His Phe Thr Phe
 180 185 190
 Lys Asp Leu His Phe Lys Met Phe Asp Val Gly Gly Gln Arg Ser Glu
 195 200 205
 Arg Lys Lys Trp Ile His Cys Phe Glu Gly Val Thr Ala Ile Ile Phe
 210 215 220
 Cys Val Ala Leu Ser Ala Tyr Asp Leu Val Leu Ala Glu Asp Glu Glu
 225 230 235 240
 Met Asn Arg Met His Glu Ser Met Lys Leu Phe Asp Ser Ile Cys Asn

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

<210> 10
 <211> 361
 <212> PRT
 <213> Artificial Sequence
 <220><223> L-lactate dehydrogenase
 <400> 10

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

Val His Ala Leu Ser Cys His Gly Trp Val Leu Gly Glu His Gly Asp
 210 215 220

Ser Ser Val Pro Val Trp Ser Gly Val Asn Val Ala Gly Val Ser Leu
 225 230 235 240

Lys Ser Leu Asn Pro Glu Leu Gly Thr Asp Ala Asp Lys Glu Gln Trp
 245 250 255

Lys Glu Val His Lys Gln Val Val Asp Ser Ala Tyr Glu Val Ile Lys
 260 265 270

Leu Lys Gly Tyr Thr Ser Trp Ala Ile Gly Leu Ser Val Ala Asp Leu
 275 280 285

Ala Glu Ser Ile Met Lys Asn Leu Arg Arg Val His Pro Ile Ser Thr
 290 295 300

Met Ile Lys Gly Leu Tyr Gly Ile Asn Glu Asp Val Phe Leu Ser Val
 305 310 315 320

Pro Cys Ile Leu Gly Gln Asn Gly Ile Ser Asp Val Val Lys Val Thr
 325 330 335

Leu Thr Pro Glu Glu Glu Ala Arg Leu Lys Lys Ser Ala Asp Thr Leu
 340 345 350

Trp Gly Ile Gln Lys Glu Leu Gln Phe
 355 360

<210> 11
 <211> 539
 <212> PRT
 <213> Artificial Sequence
 <220><223> T-complex protein 1 subunit delta
 <400> 11

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

Arg Thr Ser Leu Gly Pro Lys Gly Met Asp Lys Met Ile Gln Asp Gly
 50 55 60
 Lys Gly Asp Val Thr Ile Thr Asn Asp Gly Ala Thr Ile Leu Lys Gln
 65 70 75 80

 Met Gln Val Leu His Pro Ala Ala Arg Met Leu Val Glu Leu Ser Lys
 85 90 95
 Ala Gln Asp Ile Glu Ala Gly Asp Gly Thr Thr Ser Val Val Ile Ile
 100 105 110
 Ala Gly Ser Leu Leu Asp Ser Cys Thr Lys Leu Leu Gln Lys Gly Ile
 115 120 125
 His Pro Thr Ile Ile Ser Glu Ser Phe Gln Lys Ala Leu Glu Lys Gly
 130 135 140
 Leu Glu Ile Leu Thr Asp Met Ser Arg Pro Val Gln Leu Ser Asp Arg

 145 150 155 160
 Glu Thr Leu Leu Asn Ser Ala Thr Thr Ser Leu Asn Ser Lys Val Val
 165 170 175
 Ser Gln Tyr Ser Ser Leu Leu Ser Pro Met Ser Val Asn Ala Val Met
 180 185 190
 Lys Val Ile Asp Pro Ala Thr Ala Thr Ser Val Asp Leu Arg Asp Ile
 195 200 205
 Lys Ile Val Lys Lys Leu Gly Gly Thr Ile Asp Asp Cys Glu Leu Val
 210 215 220

 Glu Gly Leu Val Leu Thr Gln Lys Val Ala Asn Ser Gly Ile Thr Arg
 225 230 235 240
 Val Glu Lys Ala Lys Ile Gly Leu Ile Gln Phe Cys Leu Ser Ala Pro
 245 250 255
 Lys Thr Asp Met Asp Asn Gln Ile Val Val Ser Asp Tyr Ala Gln Met
 260 265 270
 Asp Arg Val Leu Arg Glu Glu Arg Ala Tyr Ile Leu Asn Leu Val Lys
 275 280 285
 Gln Ile Lys Lys Thr Gly Cys Asn Val Leu Leu Ile Gln Lys Ser Ile

<211> 198

<212> PRT

<213> Artificial Sequence

<220><223> peroxiredoxin 2

<400> 12

Met Ala Ser Gly Asn Ala Gln Ile Gly Lys Ser Ala Pro Asp Phe Thr

1 5 10 15

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

20 25 30

Tyr Arg Gly Lys Tyr Val Val Leu Phe Phe Tyr Pro Leu Asp Phe Thr

35 40 45

Phe Val Cys Pro Thr Glu Ile Ile Ala Phe Ser Asp His Ala Glu Asp

50 55 60

Phe Arg Lys Leu Gly Cys Glu Val Leu Gly Val Ser Val Asp Ser Gln

65 70 75 80

Phe Thr His Leu Ala Trp Ile Asn Thr Pro Arg Lys Glu Gly Gly Leu

85 90 95

Gly Pro Leu Asn Ile Pro Leu Leu Ala Asp Val Thr Lys Ser Leu Ser

100 105 110

Gln Asn Tyr Gly Val Leu Lys Asn Asp Glu Gly Ile Ala Tyr Arg Gly

115 120 125

Leu Phe Ile Ile Asp Ala Lys Gly Val Leu Arg Gln Ile Thr Val Asn

130 135 140

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

145 150 155 160

Ala Phe Gln Tyr Thr Asp Glu His Gly Glu Val Cys Pro Ala Gly Trp

165 170 175

Lys Pro Gly Ser Asp Thr Ile Lys Pro Asn Val Asp Asp Ser Lys Glu

180 185 190

Tyr Phe Ser Lys His Asn

195

<210> 13

<211> 2104
 <212> DNA
 <213> Artificial Sequence
 <220><223> Heat shock cognate 71 kDa protein
 <400> 13

ggtctcattg aacgcggagg cagctgcctg gcatttgtgt ggtctcgtcg tcagcgcagc 60
 tgggcctaca cacaagcaac catgtctaag ggacctgcag ttggcattga tctcggcacc 120

 acctactcct gtgtgggtgt cttccagcat ggaaaggtgg aaattattgc caatgaccag 180
 ggtaaccgca ccacgccaag ctatgttgct ttcacggaca cagagagatt aattggggat 240
 gcggccaaga atcaggttgc aatgaacccc accaacacag tttttgatgc caaacgtctg 300
 atcgggcgta ggtttgatga tgctgttggt cagttctgata tgaagcactg gcccttcatg 360
 gtggtgaatg atgcaggcag gcccaaggtc caagtggaat acaaagggga gacaaaaagt 420
 ttctaccag aggaagtgtc ctccatggtt ctgacaaaga tgaaggaaat tgcagaagcg 480
 tacctcgga agaccgttac caacgctgtg gtcacagtgc ccgttactt caatgactct 540

 cagcgacagg caacaaaaga tgctggaact attgctggcc tcaatgtact tcgaatcatc 600
 aatgaaccaa ctgctgctgc tattgcttac ggcttagata agaaggtcgg agctgaaagg 660
 aatgtgctca tttttgactt gggaggtggc acttttgatg tgtcaatcct cactattgag 720
 gatggaattt ttgaggtcaa atcaacagct ggagacaccc acttaggtgg agaagatfff 780
 gacaaccgaa tggtaaatca tttcattgct gagtccaagc gaaagcaca gaaagacatc 840
 agtgagaaca agagagctgt ccgccgtctc cgcacggcct gcgagcgggc caagcgcacc 900
 ctctctcca gcaccaggc cagtattgag attgattctc tctatgaggg aattgacttc 960

 tatactcca ttaccggggc tcgatttgag gagttagaat ctgacctgtt ccgtggcaca 1020
 ctggaccctg tagagaaggc ctttcgagat gccaaagtgg acaagtcaca gatccatgat 1080
 attgtcttgg tgggtggttc taccagaatc cccaagattc agaaacttct gcaagacttc 1140
 ttcaatggaa aagagctgaa caagagcatt aaccccgatg aagctgttgc ctatggtgca 1200
 gctgtccagg cagccattct atctggagac aagtctgaga acgttcagga tttgctgctc 1260
 ttggatgta ctctctttc ctttggtatt gaaactctg gcggagtcat gactgtcctc 1320
 atcaagcga ataccacat ccccaccaag cagacacaga ctttcaccac ctactctgac 1380

 aaccagcctg gtgtactcat tcaggtgtat gaaggtgaaa gggccatgac caaggacaac 1440
 aacctgcttg gaaagttcga gctcacagge atccctccag caccctgtgg ggtccctcag 1500
 attgaggtta cttttgacat cgatgccaat ggcatcctca atgtttctgc tntagataag 1560

agcacaggaa aggagaacaa gatcaccatc accaatgaca agggccgctt gagtaaggaa 1620
 gatattgagc gcatggtcca agaagctgag aagtacaagg ctgaggatga gaagcagaga 1680
 gataaggttt cctccaagaa ctcaactggag tcctatgcct tcaacatgaa agcaactgtg 1740
 gaagatgaga aacttcaagg caagatcaat gatgaggaca aacagaagat tcttgacaag 1800

 tgcaatgaaa tcatcagctg gctggataag aaccagactg cagagaagga agaatttgag 1860
 catcagcaga aagaactgga gaaagtctgc aaccctatca ttaccaagct gtaccagagt 1920
 gcaggtggca tgcttggggg aatgcctggt ggcttcccag gtggaggagc tccccatct 1980
 ggtggtgctt cttcaggccc caccattgaa gaggtggatt aagtcagtcc aagaaggtgt 2040
 agctttgttc cacagggacc caaaacaagt aacatggaat aataaaacta tttaaattgg 2100
 cacc 2104

 <210> 14
 <211> 3139
 <212> DNA

 <213> Artificial Sequence
 <220><223> Inter-alpha trypsin inhibitor, heavy chain 2
 <400> 14

 ttctttttaa aacgatctgc tctcatcaac agacaagttc ctgctaacct ggttccgtag 60
 gtggggagtt ctccccagac catctgctcg ggggagcttg gcacagatgt ccagcaaat 120
 gcagcggcct gtgtgccttc tcatctggct gtttctctg gaagcacaag ctttcgaat 180
 ccccataaac ggaaactcag aatttgcaga atacagtgat cttgtggaac tggccccaga 240
 taaactccca tttgtgcaag agaatggaag acatcagaga agccttcctg aggaatccgg 300
 agaggagacg gacactgttg atccagtjac tctttatagc tacaagtcc agtccactat 360

 tacttctcgg gtggccacaa ctaccatcca gagcaactg gtgaacaatt cccattgcc 420
 tcagagtgtc gtgttcgatg ttcagattcc caaaggagcc tttatctcca acttcacat 480
 gactgtgaac ggcatgacat ttacaagctc cattaaagag aaaactgtgg gccgagctct 540
 ttattcacag gcaagagcaa agggcaagac ggctggatgg gtaaggagca gaactctgga 600
 tatggagaac ttcaacacag aggtgaacat cccgcctgga gccaaagtgc agtttgaact 660
 tcattaccag gaggtgaagt ggaggaagct gggatcctat gagcacaaga ttcatttgca 720
 gcccgggaag ctggccaac acttggaggt gaatgtgtgg attatcgaac cccaaggaat 780

 gagatttctt catgttctcg acacatttga aggccatttc caaggcgtcc ctgtcatatc 840
 taaaggacaa cagaaggcac atgtatcctt caagcccaca gtagcacaac aaagaaaatg 900

ccccaactgc acagagactg cagtgaacgg ggagctggtg gtgatgtatg atgtcaacag 960
 agaagagaag gccggggaac ttgaggtgtt taatggatat tttgttact tctttgctcc 1020
 tgagaacctg gacccaatcc ccaagaacat cctttttgtt attgacgtca gtggctccat 1080
 gtgggggatt aagatgaagc agactgtgga ggcaatgaag acgatattgg atgacctaa 1140
 aacagacgac caattctctg tggttgattt caaccataat gttcgaacct ggagaaatga 1200

 tttagtttca gctactaaaa cacaaattgc agatgccaag agatacattg agaaaattca 1260
 gcctagtgga ggcacgaata tcaacgaggc actgcttcca gcaattttca ttttgaatga 1320
 agccagtaac atgggattgc tgaaccccga ctcagtctct ctgatcattt tggtttctga 1380
 tggagatcca accgtagggg aactaaaatt gtccaaaatt cagaaaaatg tgaagcagag 1440
 catccaagat aacatctccc tgtttagttt gggaatcgga tttgatgtcg actatgattt 1500
 tttgaagagg ctgtccaatg aaaatcgtgg tattgctcaa cggatctacg ggaaccagga 1560
 cacatcctct cagctcaaga aattttacaa ccaggtctct actccactgc tcagaaatgt 1620

 tcagtttaac taccgcagg catcagtaac agatgtcacc caaaacaatt tccacaacta 1680
 cttcggaggt tctgagatcg tgggtggcagg gaaatttgat cccagcaaac tgacagaagt 1740
 ccagagcatc atcacggcaa cttcggctaa cacagaattg gtcttgaaa cctgagcca 1800
 gatggatgac ttggaggagt ttctgtcaaa ggacaagcat gcagaccctg acttcaccaa 1860
 gaagctgtgg gectatctca ccatcaacca actactagca gagagaagtc tggctctac 1920
 tgctgccatc aaaagaaaa tcaccaaacc aatcctgcag atgtctctag accatcacat 1980
 cgtgactccg ctactgcca tggatgata gaatgatgct ggagatgaac gcatgctggc 2040

 tgactccccg ccacaggacc actcctgctg ctgaggcgca ctgtattatg gcaccaaggt 2100
 tgctcgggt cccatccat catgggcaa cccatctccg acaccaatgt cggccatgct 2160
 tgcagtcgga gcaagccac ttgagtctac tccacctact catttgaatc aagtggaaaa 2220
 tgacccccac ttcacattt acctgcccc aagcaaaagg aatatttgtt tcaatattga 2280
 ctcagaacct gggaaaatcc taagcctggt gtctgacca gaatcaggga ttgtagtcaa 2340
 cggtcagctt atcggcgcca agagggcaga gaatggaaag ctaagcacct actttggaaa 2400
 actggggttt tatttccaaa aggaaggcat gaagatagaa atcagcaccg agaccatcac 2460

 cctgagctct ggtctctcta catctgact gtctgtgtec gacacagctc atcttgaaa 2520
 ctcgagggtg ctcatctctg tgaagaaaga aaaatctgtg actctcacc tgaataaaga 2580
 gctgttcttt tccgttctgc tgcacgtgt gtggaggaag caccagtta acgtggactt 2640
 cctggggate tacgtcccc ccatagacaa gttctcact agagtgcag gactgttagg 2700
 ccagttcatg caggagccag ccatccacat cttcaatgag agaccggaa aggagccggg 2760

aaaacccgag gcaagcatgg aggtgaaagg acataagctg actgtcacca gaggcctgca 2820
 gaaggactac aggacggata tcgtgtttgg aacagacgtt ccctgctggg ttgtacacaa 2880

 cagtgggaaa ggtttcatcg atggacatta taaggattat tttgtgcctc agctctacag 2940
 ctttctcaag cggccttaat ggcttatagt tttgaaact gtatgtgtat ccttttctcc 3000
 cttgatactt tttgcagtta ctccccgtt tcagcaactc aaaataaacg cagatattat 3060
 agtggcctaa aaggcctgct aaccaccgg agaaaaataa atatttgcaa aaaagctca 3120
 ggaaaaaaaa aaaaaaaaaa 3139

 <210> 15
 <211> 1812
 <212> DNA
 <213> Artificial Sequence
 <220><223> Vitamin D-binding protein

 <400> 15
 ttgacctat gatgatatta ctatctattt ttgggaacaa ggctgattgt cccttgcctc 60
 tgggcagaga ttaataattg atgagtttct agttaggatc gtttgcaaat cctgtgctcc 120
 tgccitttgc gtggtggccg gcaggtgaga ggaggtgctg caagactctc tggtcgcaga 180
 atgaagaggg ttctggttct cctgctggcc ttagcctttg ggcacgctct agagagaggg 240
 cgagactatg agaaggataa agtttgcaat gaactcgcca tgctggggaa agaggacttc 300
 agatctttat cactaatcct atatagcagg aagttttcca gtagcacatt tgagcaggtc 360
 aaccagcttg tgaaggaagt tgtctccttg actgaggagt gctgtgctga gggagctgac 420

 cccacctgct acgacaccag gacctcagag ctgtctgtta agtctctgta aagtgatgct 480
 ccctttccgg ttcacctgg aactcctgag tgttgaccca aggaggggct ggagcggaaa 540
 ctctgcatgg ctgccctcag tcaccagccg caggaatttc ccacctacgt ggaaccaaca 600
 aatgatgaga tctgtgagcc cttcaggagg gacccaaagg gatttgctga ccagtttctg 660
 tatgagtatt ccagcaatta tggacaagcg cctctgcccac ttttagttgc ttacaccaag 720
 aattatctct ctatggttgg atcctgctgt acttctgcaa acccaactgt atgctttgtg 780
 aaggagagac tccagatgaa gcatttgtca ctctcacca ccatgtcaaa cagagtctgc 840

 tcacaatatg ctgcatatgg aaaggaaaaa tcaaggctga gccatctcat aaaactagcc 900
 caaaaagtgc caactgctaa cctggagaac gttctgccac tagctgaaga ctttactgaa 960
 atcctgtcca gatgtttgta gtctacctca gaggattgca tggccagtga getgcctgag 1020
 cacacaataa aaatctgtca aaacttatcc aaaaagaatt ctaagtttga agagtgtgt 1080

caagaaaaca cacccatgaa catttttatg tgcacctact tcatgccagc tgctgaacca 1140
 cttcaattgc cagctatcaa gttgccaact ggcaaagacc tctgtgtgtca gagtaccaca 1200
 caagccatgg accagtatac atttgaacta agcagaagga ctcaagttcc agaagtgttc 1260

ctcagcaaag ttctggagcc aaccctgaaa acccttaggg agtgcgtgtga cactcaggat 1320
 tctgttgctt gtttcagcac tcagagtccc ctgctgaaga ggcaactaac ttctttcatc 1380
 gaaaaaggtc aagaaatgtg tgcagattat tctgagaaca catttactga gtacaagaaa 1440
 aaattggcag aacggctaag gacaaaaaca cccaacacct ctccggcaga gctgaaagac 1500
 atggtggaga aacactcgga ctttgctctt aagtgtctgt ctataaactc acctcctctc 1560
 tactgcagct cacagattga tgcagaaatg atagacaccc tgcagtctctg atcagggcc 1620
 gtgcactagc ttggatcttg aactgaacca ctctggaaaa ttgccactgg taacaaaatc 1680

aagcaccaca gagatggcct tccaagaaga tcaccaagat aattccattt ttcttagcta 1740
 caatgttttt ggagaattat aaaaaataaa taataaatt gaaatattgt gcaaaactcta 1800
 attatgatgc ag 1812

<210> 16
 <211> 8560
 <212> DNA
 <213> Artificial Sequence
 <220><223> Talin-1
 <400> 16

aactcgaag tggctcctgg gctgcgccac gtcccggggg ctatgcaaat tatggggacg 60
 tcctttcaag cttcccacgc cccagggcgg ggattctgag gatccttgtc cgccttcgcc 120
 tcgagctatt aaaaaaaaaa acttaagcgt ttaaaaggga actccagctt agttcgggcg 180

ttctctccag agctctgcac gctcgttcc tgctggggag gggcggccag gcttcgcggg 240
 cgcccagca tcgagaacta cggccagagc agcttctctgc agtctccgct accatagagc 300
 gcgggcccag ggcgcgccg gcgggtgggg gacgtttcca ggacggaagt ggcccagagt 360
 gtgtcgaagg gagggcgagg ccggagcccc aaaagcgacc gggagaagga gcgggtagcg 420
 ggcccagggc ggggcccaga gccgggcagc gcaggtatcg ccaggccgaa gacgagaagc 480
 tgccaccatg gttgcgcttt cgctgaagat tagcattggg aatgtggtga agacgatgca 540
 atttagcca tctaccatgg tgtacgacgc ctgccgatg attcgtgagc ggatcccaga 600

ggccctggct ggccctccca acgactttgg gctctttctg tcagatgatg accccaaaaa 660
 aggcactctg ctagaggctg ggaaagcttt ggactactac atgctccgaa atggggacac 720

catggagtac agaaaaaagc agagaccct gaagatccgg atgttagacg gaacggtgaa 780
 gactatcatg gtggatgact ccaagaccgt cacagacatg ctcatgacca ttgtgcccc 840
 aattggtatc accaacatg atgagtattc actggttcga gagctgatgg aagaaaagaa 900
 agatgagggg acagggacc tgaagaaaga caagacccta ctgcgagatg aaaagaagat 960
 ggagaagcta aagcagaagt tgcacacaga cgatgagttg aactggctgg accatgggcg 1020

 gacactgagg gaacaggag tggagagca tgagacgttg ctgctgcgga gaaagtctt 1080
 ctactcagac cagaatgtgg attcccaga cctgtacag ctgaacctc tctatgtgca 1140
 ggcacagatg gacatctga atggctcca tctgtctcc ttgacaagg cctgtgaatt 1200
 tgcaggcttc cagtgccaga tccagtttg acctacaat gaacagaagc acaaggctgg 1260
 ctcccttgac ctgaaggact tctgcccga ggagtatgtg aagcagaagg gagagcgtaa 1320
 gatctttcag gcacacaaga attgtggca gatgagtga attgaggcca aggtacgcta 1380
 cgtgaagctg gcccttccc tcaagactta cgggtgtctc ttcttctag tcaaggaaaa 1440

 gatgaagggg aagaataaac tgggtcccag gctgttggc atcactaagg agtgtgtgat 1500
 gcgtgtggat gagaagacca aggagtgat tcaggaatgg agcctacca acatcaaagc 1560
 ctgggctgcc tctccaaga gctttactc ggactttgga gactatcagg atggctacta 1620
 ctcagtacag acgacagaag gcgagcagat cgcacagctc attgctggct acatagatat 1680
 catccttaag aagaaaaaa gcaaggacca ttttggctg gagggagatg aagagtctac 1740
 tatgttggag gactcagtg cacccaaaa gtcaacagtc cttcagcagc agtacaaccg 1800
 agtggggaaa gtggagcacg gctctgtggc tctgccagcc atcatgcgct ctggagcctc 1860

 tggctctgag aatttcagg tggggagcat gccacctgcc cagcagcaga tcaccagtgg 1920
 ccagatgcac cgaggacaca tgccacctc gacttcagcc cagcagcgc tcaactggaac 1980
 cattaactcc agcatgcagg ctgtgcaggc tgcccaggcc actctggatg actttgagac 2040
 tctaccctc ctggccagg atgtgcctc caaagcttg cgtaagaaca agatggatga 2100
 atcaaagcat gagatccact cccagtaga tgccatcaca gctggcactg cctctgtggt 2160
 aaacctgaca gcaggagatc ctgcagagac agactataca gcagtggct gtgcagtcac 2220
 caccatctc tccaacctga cggagatgtc ccgtgggtg aagctattgg ctgccctgct 2280

 ggaggatgaa ggcgcaatg gccggccct tctgcaagca gcaaaggcc ttgcaggggc 2340
 tgtgtcgaa ctgcttcga gtgctcagcc cgccagcgt gagccccgc agaacctgct 2400
 gcaggcagcc ggaacgtgg gccaggccag tggggagctg ttgcagcaa ttgggaaag 2460
 tgacactgac cccacttcc aggatgttct aatgcagcta gcaaaggcag tggcaagtgc 2520
 tgcagccgcc ctggtcctca aggccaagag tgtggcccag cgcacagagg attccggct 2580

tcagacccaa gtgattgctg cagctacgca atgtgctctg tccacttccc agctggtggc	2640
ctgtaccaag gtggtagcac ctacaatcag ctcacctgtg tgccaagaac agctagtcga	2700
ggctggggcga ctggtggcca aagctgtgga gggctgtgtg tctgcctcgc aagcagctac	2760
agaagatgga cagctgcttc gaggggtagg agcagcagcc acggctgtca cccaggccct	2820
caatgagctg ctgcagcagc tgaaggccca cgccaccgga gctgggcctg ccggtcggtg	2880
tgaccaagcc actgatacca tccctaccgt cactgaaaac atcttcagct ccatgggtga	2940
tgctggggag atggtgcgcc agggccgcat cctggcccaa gccacctcag acctggtcaa	3000
tgctatcaag gccgatgctg agggggagag tgatctggag aactctagaa agctcctgag	3060
tgctgccaag atcctcgtg atgccaccgc caagatggtg gaggcggcca agggagcagc	3120
cgcccacct gacagtgagg aacagcagca gcgactgcgt gaagcagctg aggggcttcg	3180
catggccacc aatgcagctg cgcagaacgc catcaagaag aagtgggtgc agcgcctgga	3240
gcatgcagcc aagcaagctg cagcctctgc aacacagacc attgctgcag cccaacatgc	3300
agcctctgcc cccaaggcct ctgcggggccc ccagccccta ctggtgcaga gctgtaagge	3360
cgtggcagag cagattccgt tgctgggtgca ggggtgccga ggaagtcaag ctcagcctga	3420
cagccctagt gctcagctgc ccctcattgc tgccagccag agcttctcgc agccaggtgg	3480
gaagatggtg gcagcggcaa aagcctcagt gccaacaatt caggaccagg cttcagcaat	3540
gcagctgagt cagtgcgcaa agaacctagg cactgccctg gccgagctgc gactgctgc	3600
tcagaaggct caggaggcat gtggaccttt ggagatggat tctgactga gtgttgtaca	3660
aaatctagag aaagacctac aggaaataaa agcagcagct cgagatggca agcttaagcc	3720
cttaccggg gaaacaatgg agaagtgtac ccaagatctt ggcaacagca ccaaagcgg	3780
gagctctgcc atcgccaagc tgctggggga gattgcccag ggcaatgaga actatgcagg	3840
tattgcagct cgggatgtag ctggtggact aaggtcacta gcccagctg cgcggggtgt	3900
ggctgcgctg acatcagatc ctgcagtgca ggccattgtg cttgacacag ccagcgacgt	3960
tctggacaag gccagcagcc tcattgaaga ggcaaaaaa gcatctggac acccagggga	4020
cccagaaagc cagcaaaggc ttgctcaggt agctaaagca gtgaccaag ccctgaaccg	4080
ctgtgtcagc tgctgcctg gccaaagaga cftggataat gccctacgag ccgttgaga	4140
tgccagcaag cgactcctga gtgactcgt tctccaagc acggggacat ttcaagaagc	4200
acagagccga ttgaatgaag ctgcagctgg gttaaatcag gccgccacag aactggtgca	4260
agcctcccga ggaactcctc aggacctggc tgggcctcc ggccgatttg gacaggactt	4320

cagcaccttc ctggaagctg gtgtggagat ggccggacag gccccgagcc aggaggaccg 4380

 agcccagggtg gtgtccaacc tgaaggcat atccatgtct tcaagcaaac ttctccttgc 4440
 cgctaaggcc ttgtccacag accctgcttc tccaacctc aagagtcagc tggctcggc 4500
 tgccccggca gtgacggaca gcatcaacca gctcatcacc atgtgcacc agcaggcacc 4560
 tggccagaag gagtgtgaca atgcaacttc gcagctagag acagtcggag aactcctgga 4620
 gaatccagtc cagcccatca acgacatgtc ctacttcggt tgcttgaca gtgtcatgga 4680
 gaactccaag gtcttaggtg aggccatgac tggcatctcc caaaatgcca agaatggaaa 4740
 tctgccggag ttggagatg ccattgccac agcctccaag gctctctgtg gcttcacgga 4800

 ggcagccgca caggcagcat atctagttag gtctctgac cccaacagcc aagctggaca 4860
 gcaaggactg gtggaacca cacagtttgc cgtgcaaac caggcaattc agatggcctg 4920
 tcagagtctg ggggagcctg gctgtacca ggcccagggt ttatctgcag ccactatgt 4980
 agccaaacac acatctgcat tgtgtaacag ctgtcgctg gcttccgcta ggactgcaa 5040
 tcccactgcc aagcggcagt ttgtacagtc agccaaggag gtggccaaca gtacagcaa 5100
 tcttgtcaag accatcaagg cactagatgg ggacttcaca gaagagaacc gtgccagtg 5160
 ccgagcagcc acagcccctc tgctggaagc tgtggataac ctgagtgcct ttgcctcaa 5220

 cctgagttc tccagctcc ctgcccagat cagccctgag ggcagagcag ccatggagcc 5280
 cattgtaatc tctgctaaga caatgttga gagtgtgga gggctcatcc agacagctcg 5340
 ggcgttagca gtcaatcccc gagaccccc acgttggtct gtgctagctg gccactctcg 5400
 gactgtctca gattccatca agaaacttat tacaagcatg agggacaaag ccccagggea 5460
 gctggagtgt gagacagcca ttgcggctct gaacagctgc ttgcgggacc tagatcaggc 5520
 ttcgcttgcc gctgtcagcc agcagcttgc tccccgtgaa ggaatctctc aagagcttt 5580
 gcacaccag atgctcactg cagtgcaaga aatttctcat ctcatgagc cgctggccag 5640

 cgctgtcgg gctgaagcct cccagctggg acacaagtg tcccaaatgg cccagtactt 5700
 tgaaccactc accctggctg cagtgggtgc tgcgtctaa accctgagcc acccacagca 5760
 gatggcactt ctggaccaga ctaaacgtt ggcagagtct gccttgagc tgctatatac 5820
 tgccaaggag gctggtgta accccaagca agcagcacac acccaggaag ccctggagga 5880
 ggctgtgcag atgatgacag aggcctaga agacctgacg acgacctca acgaagcagc 5940
 cagtgtcga ggagtcgtt gcggcatggt ggactctatc actcaggcca tcaaccagct 6000
 agatgaagga cctatgggtg acccagaagg ctcatctgta gattaccaga caaccatggt 6060

gaggacagcc aaggccattg ctgtcactgt tcaggagatg gtaaccaagt caaacaccag 6120
 ccctgaagag ctaggcctc ttgccaacca gctgaccagt gactatggcc gactggcctc 6180
 acaagccaag cctgcagctg tggctgctga aatgaagag ataggcgctc acatcaaca 6240
 ccgagtacag gagctgggcc atggctgctc tgccttggtc accaaggcag gtgccttgca 6300
 gtgtagcccc agtgatgct acaccaagaa ggagctcata gagtgtgcc gcagagtgtc 6360
 agagaaggtc tcccatgttc tggctgcact ccaggctggg aatcgtggtt cccaggcctg 6420
 cattacagca gccagtgtg tgtctggtat cattgtgac ctcgacacca ccatcatgtt 6480

 cgctactgct ggcacactta accgtgaggg tgctgaaact tttgctgacc accgggaggg 6540
 tatcttaaag acagcaaagg ttcttgtgga ggacaccaag gtcttagtgc agaatgcagc 6600
 tgggagccag gagaagtgg cacaagccgc ccagtcctcc gtggccacca ttaccgcct 6660
 cgctgatgtg gtcaagctg gtgcagccag cctaggagcc gaagaccctg aaactcaggt 6720
 ggtgctgata aatgcagtaa aggacgtagc caaggccctg ggtgacctca tcagcgtac 6780
 gaaggctgca gcgggcaaag ttggggatga cctgcagtg tggcagctca agaactctgc 6840
 caagtgatg gtgaccaatg tgacatcatt gctcaagaca gtgaaggctg tggaaatga 6900

 ggccacaaa ggcacacggg ccctagaggc aaccacagag cacatacgtc aggaactggc 6960
 ggtcttctgt tccccagagc cacctgcaa gacctctacc cctgaagact tcatccgaat 7020
 gaccaaggtt attactatgg caacagccaa agccgttctc gctggcaatt cctgtcgaca 7080
 ggaagatgtc attgccacag ccaatctgag ccgacgggct attgctggaca tgcttcgggc 7140
 ttgcaaggaa gcagcttcc acccagaagt ggcgcctgat gtacggctcc gagccctgca 7200
 ctatggccgg gagtgtgcca atggttacct ggaactgctg gaccacgtgc tgctgacct 7260
 tcagaagcca aaccagacc tgaagcagca gctaactgga cactcaaagc gggttgctgg 7320

 ctctgtgact gagctcatcc aggctgctga agccatgaaa ggaacagagt ggggtgaccc 7380
 agaggacct acigtattg ctgagaatga actcttggga gccgcagccg ccatcgaggc 7440
 tcagccaag aagctggagc agctgaagcc caggccaaa cccaaggagg cggatgagtc 7500
 ctgaaacttt gaggaacaaa tccatagaagc tgccaagtcc atcgctgcag ccaccagtgc 7560
 actggtaaag gctgcgtcag cagcccagag ggagctggtg gctcaaggaa aggtgggctc 7620
 cattccagc aatgcactgg atgatgggca gtggtcgcag ggcctcattt ccgctgccc 7680
 tatggtggct gcggccacca acaatctgtg tgaagcagct aacgcagctg tccaggccca 7740

 tgctagccag gagaaactca tttctcagc caagcaagta gctgcctcca cagctcagct 7800
 cctggtagct tgcaaggtca aagcagatca ggactctgag gcaatgaaac ggctccaggc 7860
 tgctggcaat gcagtgaaga gggcctcaga taacctggtt aaggcggccc agaaagctgc 7920

agccttcgaa gaccaggaga atgagacggt ggtggtgaag gagaaaatgg ttgggggcat 7980
 tgcccagatt atcgagcac aggaagagat gcttcggaag gaacgagagc tggaagaggc 8040
 tcggaaaaag ctgcccaga tccggcagca gcagtacaag ttcttgctt cagagcttcg 8100
 agacgagcac taaaaaagcc ctgtgtattt aatgcagacc cagcccagag actgtgcctg 8160

ccactaccaa agccttctgg gctgctcagg acccaagctg cccaaccca gcactcccc 8220
 aaagtgcctg ccaaacctg ggccctggccc tgcccagtcc cactgcaagc cgtgtcctct 8280
 ccctgactcc caagtgcctt tgcacacctag ggcccctaag tgcctgcccc cctccagagt 8340
 attaacgctc caagagtatt attaatgctg ccgtacctca gtttgaacct gccagggccc 8400
 cagctgctcc agcctgccag cagcttccag ccagtcccca cggccacgtc agctcaactc 8460
 atcccttttt gatactatat ccctacca gctacctatg gggcttgagg gttgtaaac 8520
 caaacaggtc agactccaat aaaggtgatt ctacagctgc 8560

- <210> 17
- <211> 2551
- <212> DNA
- <213> Artificial Sequence
- <220><223> Catalase

<400> 17
 gaagtcacca ctccagcggg cctggccaac aagattgctt tctccgggtg gagaccgctg 60
 cgtccgtccc tctgtctca cgttccgag ctctgcagct ccgcaatcct acaccatgct 120
 ggacagtctg gaccagcca gcgaccagat gaagcagtgg aaggagcagc gggcctcgca 180
 gagacctgat gtctgacca ccggagggcg gaaccaata ggagataaac ttaatatcat 240
 gaccgctggg tcccagggc cctctctctg tcaggatgtg gttttactg acgagatggc 300
 acactttgac agagagcggg ttctgagag agtgggtacac gcaaaaggag caggtgcttt 360

tggatacttt gaggtcacc acgatacac cagatactcc aaggcaaagg tgtttgagca 420
 tattggaaag aggacccta ttgcccctc attctccaca gtcactggag agtcaggctc 480
 agctgacaca gttcgtgacc ctggggggtt tgcagtgaaa ttttactg aagatggtaa 540
 ctgggatctt gtgggaaaca acaccctat tttcttctc agggatgcca tattgtttcc 600
 atcctttatc catagccaga agagaaacc acagactcac ctgaaggatc ctgacatggt 660
 ctgggacttc tggagtcttc gtcccagtc tctccatcag gtttcttct tgttcagtga 720
 ccgagggatt cccgatggc accggccat gaatggctat ggatcacaca cttcaagtt 780

ggttaatgca gatggagagg cagtctattg caagttccat tacaagaccg accagggcat 840

caaaaacttg cctgttgag aggcaggaag gcttgctcag gaagatccgg attatggcct 900
 ccgagatctt ttcaatgcca tgcceaagg caattaccg tcctggacgt ttacatcca 960
 ggtcatgact ttaaggagg cagaaaacttt cccatttaac ccatttgatc tgaccaaggt 1020
 ttggcctcac aaggactacc ctcttatacc agttggcaaa ctggttttaa acaaaaaatcc 1080
 agttaattac tttgctgaag ttgaacagat ggcttttgac ccaagcaata tgccccctgg 1140
 catcgagccc agccctgaca aaatgcttca gggccgcctt tttgcctacc cggacactca 1200

ccgccaccgc ctgggaccca actatctgca gataacctgtg aactgtccct accgcgctcg 1260
 agtggccaac taccagcgtg atggcccat gtgcatgcat gacaaccagg gtggtgcccc 1320
 caactattac cccaacagct tcagcgcacc agagcagcag cgctcagccc tggagcacag 1380
 cgtccagtgc gctgtagatg tgaaacgctt caacagtct aatgaagaca atgtcactca 1440
 ggtgcggaca ttctacaca aggtgttgaa cgaggaggag aggaaacgcc tgtgtgagaa 1500
 cattgccggc cacctgaagg acgctcagct tttcattcag aagaaagcgg tcaagaattt 1560
 cactgacgtc caccctgact atggggcccc catccaggct cttctggaca agtacaacgc 1620

tgagaagcct aagaacgcaa ttcacaccta cagcaggcc ggctctcaca tggctcggaa 1680
 gggaaaagct aacctgtaac tccggtgctc agcctccgct gaggagacct ctctgaagc 1740
 cgagcctgag gatcacctgt aatcaacgct ggatggattc tcccactccg gagcgcagac 1800
 tcacgtgat gactttaaaa cgataatecg ggcttctaga gtgaatgata accatgcttt 1860
 tgatgccgtt tcctgaaggg aaatgaaagg ttagggctta gcaatcattt aacagaaaca 1920
 tggatctaata aggacttctg tttggattat tcatttaaat gactacattt aaaatgatta 1980
 caagaaaggt gtcttagcca gaaacatgac ttgattagac aagataaaaa tcttggcgag 2040

aatagtgtat tctcctatta cctcatggtc tggtatatat acaatacaac acacatacca 2100
 cacacacaca cacatgcaat acacacacta cacacacata cacacactca cacacactea 2160
 tacacacaca tgaagagatg ataaagatgg cccactcaga atttttttt ttatttttct 2220
 aaggtcctta taagcaaac catacttgca tcatgtcttc caaaagtaac tttagcactg 2280
 ttgaaactta atgtttatc ctgtgctgtg cggtgctgtg ctgtgctgtg ctgtgcagct 2340
 aatcagattc ttgtttttc ccacttggat tatgttgatg ttaatacgca gtgattcac 2400
 ataggatgat ttgtacttgc ttacattttt acaataaaat gatctacatg gaaggaccgt 2460

gtttggttgc tttcagctct gtataatgig gaatgtgaag tagagattac cagctctctc 2520
 tgcaagtaaca ataaaagcgc cagcggccag a 2551

<210> 18

<211> 4124

<212> DNA

<213> Artificial Sequence

<220><223> Signal transducing adaptor molecule 2

<400> 18

```

agtcggagtc ccgccagagc gtcgggattc agctcggggt gccggcgtga tgctctgtt      60
cactgccaac ccgttcgagc aagatgtgga aaaagccaca aatgagtaca acaccacaga      120
agactggagt ctgatcatgg acatctgtga cagggttggga agcactccca gtggagcaaa      180

agattgccta aaagccataa tgaaaagggt aaaccataag gtccttcacg ttgctctgca      240
ggcattgacc ctcttggag cttgtgtggc aaactgtgga aagatatttc atttagaagt      300
atgttcccgt gattttgcaa cggaagtacg tctctgtgata aaaaataagg ctcatcctaa      360
agtatgtgaa aaactaaat ctctaattgt ggaatgtgca gaagaattcc agaaggacc      420
ccaatttagt ctgatatctg caactattaa gtctatgaaa gaagaagggg ttacttttcc      480
ttccgcaggc tcccagactg tcgcggctgc tgccaaaaac gggacatcat tgaacaaaaa      540
caaagaagat gaggacatag ctaaagctat tgagttatcg ttgcaagagc agaagcagca      600

gtacacggag acaaaggctt tgtaccacc cgagagaggt cagctcaata acaaggctgc      660
acggagagtc agagctttat acgactttga agctgttgag gacaacgagc tcaccttaa      720
acatggtgag ctcattactg ttttggatga cagtgacgcc aactggtggc aaggagaaaa      780
tcacagagga acaggactct tcccctcaa ctttgtaacg actgacttaa gcacagaggt      840
tgagacagca acggtggaca aattgaatgt aattgatgat gatgtggagg aaattaagaa      900
atcagaacct gagcctgttt atatagatga gggtaaaatg gatagagccc tgcagattct      960
ccagagcata gatccaaaag agtcaaaacc cgactcccaa gacctcttgg acttgaaga      1020

tgtttgcaa cagatgggtc cgatgataga tgaaaaactt gaagagattg ataggaagca      1080
ttcagaactg tctgagttga acgtaaaggt gctggaagcc ctggacctgt acaataagtt      1140
ggtgaatgaa gcgcctgtgt actcggctca ttc aaagctg catcccgca attaccacc      1200
cgcagcagct ggggttccag tgcagacata tccagtcag tcgcatggtg gaaactacct      1260
gggccacggc attaccaag tatctgttgc ccagaactat aacctaggac ctgatcctat      1320
gggctcgctg agatctctgc ctccaatat gaactcgta acagcacaca ccgtccaacc      1380
tccatactta agcactggac aggacactgt ctccaacct tcttacctga accagagctc      1440

tcgtcttcag gcagctgctg gtacagctgc ttacacacag cctgtgggga tgtctacaga      1500
tgtgtcttct ttccagaaca cagcatccgg ttgcctcag ctggtctggt ttccagtggc      1560
agttcctgca cctgtcgctg cacagccaca agcaagctac caccagcagc ctctcctgta      1620

```

gagacgcacc aggacctgct gaacggcttt gtgggtgtgc tgtttaatct aggggattct 1680
 aatctgaaat attaaaagtt tccccctca gtcaaaaaga accatgaata aataaagcac 1740
 aaaacccac cttaccctga aggccataaa aggttttttt ttccagtcca gttggcttga 1800
 agttgacctt ttggtcaaag gcagcttggga tgttctaagt tagtcgggct gaactttctg 1860

 ggcccttact gactagcctt acgacactac atgatagag gtaggagtgt ggaattttct 1920
 cttaaagaa cccgttgitt cctcttaaga tgtgattttt gaggagcact gaagaccttg 1980
 aaacattttt tgcctgcact tccccaccc cttgaatttc ccacacctg ttattttaa 2040
 aggcgagcat cttttagtgc aagccttttag cttgccagca tatttctgt tggtcctta 2100
 aattacagtt gtgtttgcag tactgtcagg tttgctctca gtgtaatgct gagtagtaat 2160
 tagcagatac ttgtctacag aagcaagaac aatttgaag gtacaaaaat ggtttctgtc 2220
 attaccagtg aaaacatgt aatgcaaat attgtggagc agttagccag tcctctgaca 2280

 gacgagggcc agggactgcc ccaggagatg ttagccttcc ccagaaagg catcctca 2340
 tactcgtgtc cctggtatth agaattgtht tctagtgtct tttggtggat gtctaagaa 2400
 ggtgctgtcc tttgatttc ctgctgcta catatgcaga cactgcactt cettcttagg 2460
 gattctcagt gaatctgagc caggagacac tccccataac ctgatgtgat aggaaccatc 2520
 atgggtatca gtgtgaggtg ggcgttctt tcttacagaa agctgtcatc gtgtgccatt 2580
 gggaagggat tgtgtgtccc ctagaagtag ttgtgacct ttgcaggagc agcaggcagc 2640
 ggtgtggtgg ccacaggctg acaggacagc agtcagctgc ttgtctgaag ctctgtctt 2700

 acgttaggtc ttagtacagt tccacatact gtaagagttg gtggcaaacc caaggaatt 2760
 ttctcttaca gctttgcaca ctactttatt atatcagctt ttacacaaac ataaactaga 2820
 aaacatagat gcacaggact cggatccatg atatttacac tgggaaatgt tctaaatggg 2880
 atttttttaa atttaaaatt tttgcatgaa tgcacttctg tatatttttt taagccctta 2940
 aactttatta aaaatccatt gtgagacaat ttggaaatct tatcgtttat ctgtttacga 3000
 tacttggtta taagtatgth ttttttcaa agaaactgta ttctagaacc tactactgca 3060
 taaaatagcc agcacagaag agagegtggg ccgtccctgc aagcctgat cttgtcagtg 3120

 tctggttgtc agatgcagga ccctggaagg cttttcagtc cttagccctc accttctgtg 3180
 tttcacacgg ctttctccct tcagctggac tcattgagtc cttacaagga tggagtgcgt 3240
 gcaaaatggt accttactgt ttctagaaca gccttctgct cttcaaagat taagtaata 3300
 aatglaaaag tccagagctg ccttgataa cgtttgtgta gaggaatgcg tcttccctgc 3360
 ctttccaaac tgtgtttaat aagctcatgc tagtgtttgt gcatttcatg ttcagttatt 3420
 gtctaatta taaaagcttt atttagtcag tgtaaatact gcgttctcat taagtgggga 3480

ccagccattg aagcttagcc gttttggaat actaaaatat gccacatgac ctgtatataa 3540

atagtgtata ttagtgtctg tgaagtttaa tatatgtaga aatttcttcc tcttacattt 3600

tatttttttc cattcataaa tttgtgttgc agagttattg ttagcctaaa atattaaatc 3660

tgtattgac tgatttcagt attaatacga tatgtaaaga tatatatttt cccatgatcc 3720

agttgaactg tacttatttt aagatgtttg tagctattta ttttcagcca aaggaatagt 3780

ttaccttcag atttctacat tgggttcata tttaatgttt tctaacaaaa atttttgaat 3840

ttcatttttc atataaatct tgatatatta gaaaattctt aaaaatatat agaattttat 3900

agtataaaac atttccttgg gtgttaagtg gagtatgttt gaaaagtcaa cagtcagcca 3960

atgttgtctt ttiatttggg aatttgctta caaaattacc tacacattat ctttttgctt 4020

agcatttaaa tatgtatttc cattatgcta tttcaagtac ttgatttttg actaatgatt 4080

gtactgtata ctgaagatga taataaagtt tcatccacat tttg 4124

<210> 19

<211> 3705

<212> DNA

<213> Artificial Sequence

<220><223> Beta-parvin

<400> 19

ggccggcggg ggggtggcgcg cactcgtta tgcctccgc gccaccacgg tcgccacc 60

cgccggcccc caagatgaag aaggacgagt ctttcttggg caagtgggc ggcacactgg 120

caaggaagaa gaagaccagg gaggtgactg acttacagga agaaggcaag agcgccatca 180

actctccaat ggccccgctt ttgggtggaca ttcacctga agacaccag ctagaggaga 240

atgaggagcg taccatgac gateccacct ctagagaaga cccaagttc aaggagtgg 300

tcaaggtact tciggactgg atcaacgatg tctggcaga ggagcgcac attgtgaagc 360

agctggagga ggacctgtac gatggccagg tctccagaa acttctggaa aagctggcgc 420

actgcaagct gaatgtggca gaggtgacgc agtccgagat cgggcagaag cagaagtgc 480

agacggttct ggaagccgtg caagacctgc ttcggcccca tggctggccg ctgcgttga 540

atgtggactc tatccatggg aagaacctag tggccattct gcctctctc gtctccctgg 600

ccatgcactt cagggcccc attcaccttc ctgagcatgt caccgtgcag gtgggtgctc 660

ttcggaaagc agaaggttg ctgcaactca gccacatctc agaggagctg accacgacca 720

cagagatcat gatgggccc tttgagcggg acgccttcca tactcttttc gaccacgccc 780

cggacaagct caaccttgtg aagaagtctc tcatcacgtt tgtgaacaaa cacctgaata 840

agctaaactt ggagggtgact gacctggaga cccagtttgc agatgggtg tacctagttc 900
 tacttctggg ccttcttgaa gactactttg tccccctca caacttctac ctgactcctg 960

 acagctttga ccagaagggtg cacaacgtgg cctttgcectt tgaactgatg ctggatggag 1020
 gcctcaagaa gcccaaggca cgcctgaag atgtggtgaa cttggacctc aagtccactc 1080
 tgcgggtcct ttacacgtg ttaccaagt acaaggatgt ggagtgacgg agcaactgat 1140
 cctccaggga caagtgctag gccagaaaaa tgggtgtatc catccacaca gtccttgctt 1200
 tcccagaaaa cacctccctg gggaggcctc aggcttcctt ggctctgct cctgtcccc 1260
 tgcttggttt ggttttaaca cccccacc ccaccctt cgcctgtggt gttccctagc 1320
 tggttaggtc attgagaact tgggtcttca aataccctct cccgtttaag ttaaggcct 1380

 cacttcttct ttttctctgc ctccccctc cctcccccc cccccccc cccggagaaa 1440
 gatgaaaaag agcaaaggcc cagcctagcc ctcggtgga gagggctgtc cacctgtgct 1500
 ccagcccaca gagcaattac tgcaacagcc ctggctccag ggactccgcc ccatttgctc 1560
 tcaggaatcc cctgggctgg cctctccttg gaagcaagggt tctcagtaaa aaaagatcct 1620
 ctcagttttt cactcctgag tctcagaaga ttagactcc tgatgtgtgt ctccacacac 1680
 cccatccca tccgtcttgc ggtttgctgc ttttaacgga cttttacat gaactctcag 1740
 gaccagaat gccttggggc cccgggtag tgtgtacggt gcatgtctgg gtctcttctg 1800

 ggctctgaat gtgcagcttc agagccggct tggagtagga acctcctgtc tcggaggccc 1860
 tgttagggaa tggacacatt ttagagact ttacaagaag gtagggatcc attacaagg 1920
 ctgcctgacc acaggctacc agcgagcagc cagaattaca gaggccgtgt gctgactgc 1980
 tctggaggct gagggtagaa ggtggagggt ctggcaggct gcctttctct gccttcacac 2040
 cattedctct cctcctctc ctctcctcc tctctatct atctctctt cattccacag 2100
 tggaaaccate cctgcatga cctcgtgac cccatgatgt cagcaagcca cccctcctc 2160
 cccccccc ccagtcgcca aatgcagcct cctccgacc agccaggaat tggtttgg 2220

 gctgtgtctc agctggcagt gggacctaa agatgtccac atccctccta gacctgctg 2280
 gtatcccta cagcatgacc tataagagaa aatcctgctg ctcctcaag cccttcacca 2340
 cccacttccc ttgactctgg aggtgggtgcc atggccggct ggctccttgt ctgaggaatc 2400
 ctgcttctgt ggttaccacc caccaccacc ccttctcac gttcctctg cctctcact 2460
 gccttacagc actgttcatg gcagcttaca ggaaccttc ctttctgat tcccactta 2520
 ccacaagacc cagggtctgt gggtaggtg ttgctaccga gctgaacgcc agcaatgatg 2580

ttccagaaaa cattttaata tcttcccttg gttccactgc tgctaagctg gggacgggtc 2640

ttggaatagc cgctccggtg gaggaggctt ccagcaggg gagagagata attaaaatgg 2700

cattaccgtg tctccctgtg ggatgcggtg acattaaaga gccacactga caaaataccc 2760

gggactggaa ggttctgtgc tgccttctc gcagacacag aaccacagca gtatctggga 2820

gctgtctggga ccgctttgct ctgctcacag gcggtctggg gcggggatcc tagatgcaaa 2880

gacctaccgt gctgaaggga gggaaagaat tggctctggga cgggcggtgg cttcctgggg 2940

ttccctatct ggagggcaca agtctctctt ggaatgtaag caggctcctt tgctgaggc 3000

agaacttggg aagccagggt tctatgaagg tgtgacctcc cccattgact gacgctggag 3060

gtgtggttct cacacaccag tgaagggtc gccttaccgt cccctcactg cggagctgcc 3120

attcatacca tagcctggac aaagcacca ataactaaga tgtctaggcc aacttcaggg 3180

ccaccgcgt aaggcaagg tttctctgta gaaaaaaaaag agagaacccc cacccagatg 3240

acctcaggtg accagtgcc acccccaaac tttccacag ttctctgggt ggggggctgg 3300

gatacaata ctctctttg agctgttaat ttgaggcat tctggaacct tctactgctt 3360

ggatcttcat tctctctga tattgtttag aggtctggcc actaaaactc ttagaaatgt 3420

catcatctgt cacttccatg ggtctttag gaaggggcag gcaggtctgc acccctgcc 3480

cccaggtctg caggccagga ctgggttct ccttctgcc ctgcctgat tggcacatcc 3540

cctccagctc tccagccttc tgtggacctg caggcagctc ttgtgccacc ttcacactca 3600

gttgtcaag tctgagctc agggccctg catccttct cattggctga ccaaggggaa 3660

tcttatatcc aagcagccga atccgttacg tcccctgtgc ggccg 3705

<210> 20

<211> 2177

<212> DNA

<213> Artificial Sequence

<220><223> Guanine nucleotide-binding protein G(i) subunit alpha-2

<400> 20

gtcgctcgga actgccgacc tgagagcttc ccgacaggg ccggcgggtg gagcggagtg 60

ggtcgggagg ggccgagccg ggccgtgggc cgtgtggggg ccaggccggg ccggcggacg 120

gcaggatggg ctgcaccgtg agcgcgagg acaaggcggc agccgagcgc tctaaaatga 180

tcgacaagaa cctgcccggag gacggcgaga agcggcgcg ggaggtgaag ttgcttctgt 240

taggtgctgg agagtccagg aagagcacca tcgtcaagca gatgaagatc atccatgaag 300

atggctactc agaagaggag tgccggcagt accgtgccgt ggtctacagc aacaccatcc 360

agtctatcat ggccatcgtg aaggccatgg gcaacctgca gatcgacttt gctgatcccc 420
 agcgtgcgga tgatgccagg cagctgttcg cctgtcctg tgctgcagag gaacaagga 480

 tgcttcctga agacctgtcc ggtgtcatcc ggaggctctg ggctgaccac ggtgtgcaag 540
 cctgctttgg ccgtcacga gaataccagc tcaatgactc agccgcttac tacctgaatg 600
 atctggagcg cattgcacag agtgactaca tcctacaca gcaggatgtg ctgcggacce 660
 gtgtgaagac cacgggcacg gtggaaacac acttcacctt caaggactta cacttcaaga 720
 tgtttgatgt ggggtgtcag cggctctgagc gcaagaagtg gatccactgc tttgaggcg 780
 tcacggccat catcttctgt gtcgecttga gcgcttatga cttggtgctg gctgaggatg 840
 aggagatgaa ccgcatgcat gagagcatga agctgtttga cagcatctgc aacaacaagt 900

 ggttcacaga cacctccatc atcctcttcc tcaacaagaa ggacctgttt gaagaaaaga 960
 tcacacagag ctccctgacc atctgtttcc ctgagtacac gggggccaac aagtacgacg 1020
 aggcagccag ctacatccag agcaagtttg aggatctaaa taagcgaac gacaccaagg 1080
 agatctacac gcacttcacg tgcgccaccg acaccaagaa cgtgcagttt gtgttcgatg 1140
 ccgtcactga cgtcatcacc aagaacaacc tgaaggactg tggcctcttc tgagggcgag 1200
 tggacctggc aggatgggcc accgctgact gtgctccca ctaccctga ggaagatggg 1260
 ggcaagaaga ccatgttccc tgectgttcc cccagctgct tctccatct tttctctctg 1320

 ttctcagctc cctgtcccc tccctcggct ctgacttgg gggaggggtt gccacaggcc 1380
 tcccagtcta aaaccacct ttgtctgagg tgcgggaggt agccatgta ccccttctct 1440
 gggcatccgt tcgggttttc taaccgttgt cttgttctgg ggtgagcggg gagcgcagtc 1500
 agagatccca aggctatgt ctggagggtta ccaactctc cagcctagac ccctggcttt 1560
 gtccaacacc agccctgacc caagtccaaa tgtttacagg gagcctctg cctacccac 1620
 tctctgccgc tcggaggccc caaaggaaaa agcacaagaa gcgtgagaga taccgccatt 1680
 cctggagaca aagcccact gctcattctc gtactttta agaaaaagaa aaaggaaaaa 1740

 taaggaaaac tgcaaatcta gaaaactttt tagagaaacc tatttaaac tgtcaggtcc 1800
 tgaccagcac cccacccca gccagcccc agccccacc cttccagcg attccgtgcc 1860
 ttgagtgtgt ctgcgtgttt acaccatcc ctctttgggc ggccccctgc tctgcctcc 1920
 acggaattgg attccaaggg ctgttccaga caactgcaa cgtcactgag ggccctgctc 1980
 tgaagccctg ggccctggct cattaacct aaatgtagct ccttagcgt aatctaggaa 2040
 ccgccgtgc ctgctgaggg ccagccccct cacgccctcg ccccaggccc gggtctctca 2100

gcgttgaaca cttccttgct tttttcacat gttttatgga attgttcaca tgatttgaaa 2160

taataaaatg tagaaag 2177

<210> 21

<211> 4447

<212> DNA

<213> Artificial Sequence

<220><223> Delta-aminolevulinic acid dehydratase (porphobilinogen synthase)

<400> 21

gagcgcggtg tccagagccc ggctcggagc ggccggcagc agcgtccttg gtacgtgagc 60

aggcgcgagc ctctgggaag tggcgggttg aagggggacg ggaccccaga acaagagctg 120

aggctgcagc agacaccagc gagccgccac tccatccagc agacttctct gtgttccact 180

gccccgacca tgcaccacca gtctgttctg cacagcggct actttcacc actgcttcgg 240

agctggcaga ctgctgcctc caccgtcagt gcctccaacc tcacttatcc catctttgtc 300

acggatgttc ctgatgatgt ccagcctatc gccagcctcc caggagtggc caggtatggc 360

gtaaaccagc tagaagagat gctgagacct ctggtggaag ctggcctgcg ctgtgtcctg 420

atctttggcg tccccagcag agttcccaag gatgaacagg gctctgcagc tgactctgag 480

gactcccaa ctattgagc tgtecgctg ctgaggaaga ccttccttc cctcctagtg 540

gcctgtgacg tctgcttctg ccctacacc tcccatggcc actgtggcct cctgagttaa 600

aatggagcat tctagcaga ggagagccga cagcggttgg cagaggtggc actggcctat 660

gccaaggcag gctgtcaggt tgtagctccg tcagacatga tggacggagc agttgaggcc 720

atcaaggctg cctgctaaa acatggactt ggcaacaggg tctctgtgat gagctatagt 780

gccaatttg cctcctgttt ctacggctct tccgggatg cagctcagtc aagcccagct 840

tttgagacc gacgtgtta tcagctgcct cctggtgcc gtggcctggc cctccgagca 900

gtggcccagc acattcaaga aggagctgac atgctcatgg tgaagccggg attgccctac 960

ctggacatgg tgcgagaggt gaaggacaag caccccagc tccccctgc agtataccag 1020

gtgtctggag agtttgccat gttgtggcac ggagcccagg ccggggcctt tgatctcagg 1080

actgctgtac tggagacat gacggccttc cgcagagccg gtgccgacat catcatcacc 1140

tactttgcac cgcagctgtt gaagtggctg aaggaagaat gaagaaaaag tgtcggactt 1200

gatcttgacc aagctcctg ggccttacag aaggggaaa gtaaatgcgc tgttagaact 1260

gtgcctgtg cctcttctt ccccacgtg ctggcagggt gcagcgcct ggggtggtttg 1320

ccagcatgct aaactccac tcacagctgc agcctaccag gccctcaag tttcccactc 1380

tggtcagcc cagagacct cctccctgct ctgctgaggt gtggccttag cttggcagaa	1440
cccaagagcc tcaggcaccc gcttggagct ttgggacaag aggagagcag ttggagcttg	1500
aaacctaaag agttattcat taatgaggat caaacccag aagcagcagg gattcctggc	1560
cacaggagat gccagatca gaaagtctgt ctgcacaagt cagttcctga gaacctctgc	1620
ccagggcaga tggacacagc aggatgcaca gcagcatacg tctctgtagg atgtatgcag	1680
agaaagggga gcgtaacaaa gcaggtgaat cgatacttgc taagaagggg aagtccaaga	1740
cagctgtccc tgagttgaaa ccttaaaggg acaaagagga ccagcagacg tgctcagaga	1800
cccctttggt aggaggagct tggcagggct agagggcagg cgacaagag gaacgtggat	1860
aacagagggg aggttggat agttcgttcc ccgacttagc agggacctaa agttttgttt	1920
tgcatctgtg ggacagccat cttaaagatt tgggcgtttg gtttagtttt ctccactaag	1980
tatccctggc ctccccacc taggcaagca ctctgccact gaactgtacc tatagccttc	2040
aaggcacgtt ctggagacca gactagatgt gttgcgggag gtagataagg gaacaagaga	2100
gagtgaggtt gtctttgtgg gtttctgaag ttgccagct ggttttagtt cattcatctg	2160
acatgacacc cgtactcggc ctaatgaagg ttagggccag tttgcaggca cacctgactt	2220
tagaggtgga ggcctactc tagctctgac tctaacgca tgtcattctt aaggtgcaga	2280
attagttaga gttataact ggcaactcct gcaatcagac tggccaata cttggtggct	2340
ttcttttaa gctgtgaata actcaagac tgatcttctt acttctcagc cttctggaag	2400
gtagccacca aaatcaaac agccctaac tgtctcttcc ttggatcctg agcatagatt	2460
taccatctag aaccagatc actaatftaa aacctgctat catcaggcca ggggtatacc	2520
agactagact tgiactatct tggtgttat acagcctctg agctctgtgg ttaaaatggt	2580
tcgtgtgta ggaagccat caccaagctg ttaagagac caggagacat gctaccctgc	2640
atgtgccaag gcctaagt gtccaggecc aatggcaaat gggaaggtca ccaggtcccg	2700
acctttgaag gatcagaaaa cagtcccctg aaggactgat gtttggatga acagctaaag	2760
ctgagacata attgaccaag agccttccag ggaagaacc attgtgcaa gttcctggag	2820
ccgaagtgga gctggatgag cacggggctg gtgtggccc cacaccagga agaatacagg	2880
aacaggtggc ctggacatct cagtctaagc cgtacagggg cttcgatgcc aactcctcc	2940
acctttatcc caagatcaaa gggcacctca aggatgccag caggggaatg accatgttgt	3000
gggttttcac ttagtaacac catggcatta tgcagatgag atgtttgaat tccatggtga	3060
ccatagggat gcaaaaaagg agattattta gattatagtt tagacctgct tctctgaatg	3120

gcaagtggcc tgaccattta atggagaaaag cagactcaag gtagaacctt aagtttagct 3180

tgagtccttg gcaatgagaa aggcaccccc cactaaggct actcacattt tcggcagaac 3240

tgggttaaag gttaaaagga aaaccagatg taacattctg acccagctcc aaagccccga 3300

cccttaacca ctcatgacac ctcccagtaa gggacagtga gaaagatggc tggcctctgc 3360

tgtgtgacta actggatggg gaggaggacc ttggagccca ttttccctg gagctgggtc 3420

ttgtcacctg aagtgtccac ttggttggcag ccacctacat ctttgcaaac aaagcaggca 3480

cagacagaca tggcttacac ctaagggtct cctgcacaga agggcttctt ggacttcaag 3540

actaacctaa gctgcagaat aagatccttc tttaaaatcg ggaagagact tgagagtcat 3600

ctgaaagggt tgiaactgac ctgtgccaca tccatagaga actagatttt cagtctggca 3660

aggatattat cagcctccca gatactggc agtgggtgtc cagggacttg atgagacttg 3720

gttttctcat tcccctctca aatgtgccag tttaggtgtg agtttcaccc tttggcaggc 3780

atcataatca gccattaac tccttaagag tggagaacac ccagtttcca gctagccatc 3840

ccatatagac tgccagtcta gcacaggaag ggcctagtga tttgtggcag tcttgagtct 3900

gggtactctt ggttgcaggg actgaaactc aggtcaatag aaaaagaaac tgaactggtt 3960

cgcatgccta aatcttggtg atgtcaatta tgttgttccc caagtcaaag gctttgatct 4020

tttggccttg gggagacttt attgttggat aagccttccc agccaacaac ttgatccctc 4080

actgaaataa ccctaagat ttgtaatccc agagaaagac ccattctgtc ctgtccctgc 4140

cataacaact ctggagatct gtgattgttc tgcgtaggtc acatgtctgc caattggctc 4200

actcagtatg aactggacag ggatgggtga catggttgac ctgcccccc attacataaa 4260

aggctgacaa ggccttggct ctggtaacaa atggtcctct gcctggaatg ccttctctct 4320

aaatccatgt gtctcctttt atgttcggac ctgagctgac ctgtcacttc ctcaaactg 4380

atccagcaat ctctctaac tcacctctg tattcggatt agtgaataca acaagtcaaa 4440

ataacag 4447

<210> 22

<211> 1854

<212> DNA

<213> Artificial Sequence

<220><223> L-lactate dehydrogenase

<400> 22

gggttcttgc ggggggtgggg gggttaggaa ggaagcttgc gcgtgcgcag gettaagcac 60

gttgctatgc ctgggggtcg caccttgtgg ccgttattgg cgcctctgc tcttgatctt 120

tggacttcc tggagcaact tggcgtctca cttgctgtag ggctctgggt gatgggagaa 180
 gagcgggagg gcagctttct aaccatataa gaggagatac catccccttt tggggttcat 240

 caagatgagt aagtccctcag gcggetacac gtacacggag acctcggtat tatttttcca 300
 tttcaaggtc tcaaaagatt caaagtccaa gatggcaacc ctcaaggacc agctgattgt 360
 gaatcttctt aaggaagagc aggcctccca gaacaagatt acagttgttg gggttgggtc 420
 tgttggcatg gcttgtgcca tcagtatctt aatgaaggac ttggcggatg agcttgccct 480
 tgttgacgtc atggaagaca aactcaaggg cgagatgatg gatctccagc atggcagcct 540
 ctctctaaa acaccaaaaa ttgtctccag caaagactac tgtgtaactg cgaactccaa 600
 gctggtcatt atcaccgagg gggcccgtca gcaagagggg gagagccggc tcaacctggt 660

 ccagcgaaac gtgaacatct tcaagttcat cattccaac attgtcaagt acagtccaca 720
 ctgcaagctg ctgategtct ccaatccagt ggatattctg acctacgtgg ctgggaaat 780
 cagtggcttt ccaaaaaacc gagtaattgg aagtgggtgc aatctggatt cagcgcggtt 840
 ccgttacctg atgggagaga ggctgggggt tcacgcgtg agctgtcacg gctgggtcct 900
 gggagaacat ggcgactcca gtgtgcctgt gtggagtggg gtgaatgttg ccggcgtctc 960
 cctgaagtct cttaaccag aactgggcac tgacgcagac aaggagcagt ggaaggaggt 1020
 tcacaagcag gtggtggaca gtgcctacga ggtgatcaag ctgaaaggtt acacatcctg 1080

 ggccattggc ctctctgtgg cagacttggc tgagagcata atgaagaacc ttaggcgggt 1140
 gcatccatt tccacatga ttaaggtct ctatggaatc aatgaggatg tcttctcag 1200
 tgtccatgt atcctgggac aaaatggaat ctcgatggt gtgaaggatga cactgactec 1260
 tgaggaagag gccgcctga agaagagcgc agacaccctc tggggaatcc agaaggagct 1320
 gcagttctaa agtcttcccc gtgtcttagc acttactgt ccaggctgca gcagggttc 1380
 taggcagacc acaccctct cgtctgagct gtggttagta cagtgggtgt gagatggtgt 1440
 ggggaaacat ctactcccc acagctctgc cctgctgcca agtggactt gtgtagtgt 1500

 gacctggtta gtgtgacagt cccactgtct ctgagacaca ctgccaactg caggcttcca 1560
 ttaccctgt gagcctgctg cattgctgcc ctgcacaaa catgcctagg ccgacgagtt 1620
 ccagttaag tcgtataacc tggctccagt gtgtacgtcc atgatgata tcttgtgcat 1680
 aatgttga caggatattt tataatatt atgtgtctgt agtgtgcat gcaatattat 1740
 gtgagatgta agatctgcat atggatgatg gaaccaacca cccaagtgtc atgccaata 1800
 aaacctgaa cagtgaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaa 1854

<210> 23

<211> 1966

<212> DNA

<213> Artificial Sequence

<220><223> T-complex protein 1 subunit delta

<400> 23

ggctctccag caggcccctt gctcggcttc cgccttctct cccggcgtcg ctttctggaa 60
 ggttcgtgga ggaggcggtc agggagaccg ttactccaca gcaagccgga atccgtgtcc 120
 atccgtcttc ctgaaccgcg gcagacgcca ccaaggtcgc catgccggag aacgtagctt 180
 cccgaagcgg ggcgcccacc gccgggcccg gcagcccgcg gaaaagcgcc taccaggacc 240
 gcgacaagcc agcccagatc cgcttcagca atatttccgc ggccaaagcg gttgctgatg 300
 ctattagaac aagccttggc cctaaggaa tggacaaaat gattcaagat ggaaaaggcg 360

atgtgacat tacaatgat ggtgccacca ttctgaaaca aatgcaggta ttgcatccag 420
 cagccagaat gctggtgaa ttgtctaaag ctcaagacat agaagcagga gatggcacca 480
 cgtcggttgt catcattgct ggctctcttt tagactctctg taccaaactt ctgcagaaag 540
 gtatacatcc aaccatcatt tccgagtcac tccagaaagc tttggaaaag ggtcttgaaa 600
 tccttactga catgtctcga cctgtgcaac tgagcgacag agaaactttg ttaaatagcg 660
 caactacttc attgaattca aaggttgtct ctcaagtatt aagtctactc tctccaatga 720
 gtgtcaatgc ggtgatgaaa gtgattgacc cagccacagc taccagtgtg gatcttcgag 780

atattaaaat agttaagaag cttggtggga caatagatga ctgtgagctg gtggaaggcc 840
 tcgttctcac acagaaagta gcaaatctg gcataacaag agttgaaaag gctaagattg 900
 ggcttattca gttttgctta tctgtcctta aaacagatat ggataatcaa atagtagtat 960
 ctgactatgc ccagatggat cgagtcttc gagaggagag agcctatatt ttaaatttgg 1020
 tgaagcaaat taagaaaaca ggatgtaatg tccttctcat acagaagtct atcctgagag 1080
 atgcccttag tgatcttgca ttacattttc tgaataagat gaagattatg gtggtaagg 1140
 acgttgaaag agaagacatt gaattcatct gtaagacaat tggaaacaaa ccagttgctc 1200

acattgacca gttcactgct gacatgctgg gttctgtgta gttagcagag gaagtcagtt 1260
 taaatggttc tggaaaacta ttcaagatta caggttgtag aagcccaggg aaaacagtta 1320
 caattgtcgt acgtggttct aacaaactgg tgattgaaga agctgagcgc tccattcatg 1380
 atgctctctg tgcacatcca tgcttagtaa agaaaagagc tcttattgca ggagggggtg 1440
 ctccagaaat agagctggcc ctcaactgca cagagtactc ccgaacactg agtgggatgg 1500
 agtctactg tgttcgtgct ttccgggatg ctatggaagt cattccatct acactagctg 1560

aaaatgctgg ccigaatccc atttctacag taacagagct aagaaaccgc catgccaag 1620

 gagaaaaaac tacaggcatt aatgtccgaa aggggtgggat ctccaacatt ttggaggaaa 1680
 tggttgttca gcctctgttg gtgtcagtca gtgctttgac cttagcaact gaaactgtgc 1740
 ggagcattct gaaaatcgat gatgtggtaa atactcgata atctggataa aaggatggtt 1800
 gactgcatca ttatggacag aagtactgtg gctggaatga aggacaacca ccttgttctt 1860
 tgtctggaag cttcagagtt tttggacatt gtcttccagt tggcatttgt ttgaaattgt 1920
 attgaaacaa ttaaatgaaa acattaaata cttggtttta aactcc 1966

 <210> 24
 <211> 1446
 <212> DNA

 <213> Artificial Sequence
 <220><223> peroxiredoxin 2
 <400> 24

 gcccacctcc ccgcgactg cctggtgtgg tttgggccac gcataaaagg ttctccggcc 60
 tagggctctc tcggttttga gatctctttc ctgtctctac ccgtgtctgg acgtccgtgt 120
 gtccggctct tgctcacgca gacttagcac caaaagtgac cttggaaaagg caagacaaag 180
 gatcttctaa ggctctaaag accaagcttg tgtgggtctc tgatttcagt catggcctcc 240
 ggcaacgcgc aaatcggaaa gtcggctcct gacttcacgg ccacagcggg ggtggatggt 300
 gccttcaagg aaatcaagct ttcggactac agaggggaagt acgtggtcct ctttttctac 360

 ccaactggact tcacttttgt ttgccccacg gagatcatcg cttttagcga ccatgctgag 420
 gacttccgaa agctaggctg cgagggtgctg ggagtgtctg tggactctca gttcaccac 480
 ctggcgtgga tcaatacccc acggaaggag ggaggcttgg gccccctgaa tatccctctg 540
 cttgctgacg tgactaaaag cttgtcccag aattacggcg tgttgaaaaa tgatgagggc 600
 attgcttaca ggggtctctt tateatcgat gccaaagggtg tccttcgcca gatcacagtc 660
 aatgacctac ctgtgggacg ctctgtagac gaggctctcc gcctagtcca ggcctttcag 720
 tatacagacg agcatgggga agtctgccct gctggctgga agcccggcag tgacaccatc 780

 aagcccaatg tggatgacag caaggaatac ttctccaaac acaactgaga tgggtaaaaa 840
 tgggtgagcc tgaagcttgg atttcacctg tgcccccaacc tggatgtcct gtgctggccc 900
 agaaaatgct agattttcct ccactctctg aaggggctgg agtctaggct gaggttttct 960
 cattaccac ctggaatctg gtgaatagtg atcctgcctt gagcacacct agctgggccc 1020
 aggtctatag gaaaccaata aagtattagg gacagtgtaa gtctgtgttg gtgtgtgttt 1080

agtagtttta gtagtttaca ctctggcttg tgcaactgac tctgtagtta tgcctttca	1140
ccacaaggtg gaggcattac ctcattcacg ccatggtctg agcatggtgg gacacagtga	1200
gaccaaataa ttactaggcc tggagggatt cttaagagc tctttatggt tctcctgaag	1260
atccaggttc cattcctgga gcttacatag ttgcttaca cctctttta cgccagatcc	1320
aacaccttcg tctgatcctc cagcaccagg catgcacgtg aacatgggca gaacgtgcat	1380
agacatgaaa taaatacttc tgtaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa	1440
aaaaaa	1446