(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0034709
(43) 공개일자 2022년03월18일

(51) 국제특허분류(Int. Cl.)
*G01N 33/68* (2006.01) *G01N 30/72* (2006.01)
*G01N 30/88* (2006.01)
(52) CPC특허분류
*G01N 33/6848* (2013.01)
*G01N 30/7233* (2013.01)
(21) 출원번호 10-2021-0121321
(22) 출원일자 2021년09월10일
심사청구일자 2021년09월10일
(30) 우선권주장
1020200117021 2020년09월11일 대한민국(KR)

(71) 출원인
**서울대학교산학협력단**
서울특별시 관악구 관악로 1 (신림동)
**재단법인 아산사회복지재단**
서울특별시 송파구 올림픽로43길 88 (풍납동)
**울산대학교 산학협력단**
울산광역시 남구 대학로 93(무거동)
(72) 발명자
**김은영**
서울특별시 마포구 독막로 266, 110동 1701호(대흥동, 대흥동태영아파트)
**안용민**
서울특별시 종로구 대학로 71(연건동)
*(뒷면에 계속)*
(74) 대리인
**특허법인이룸리온**

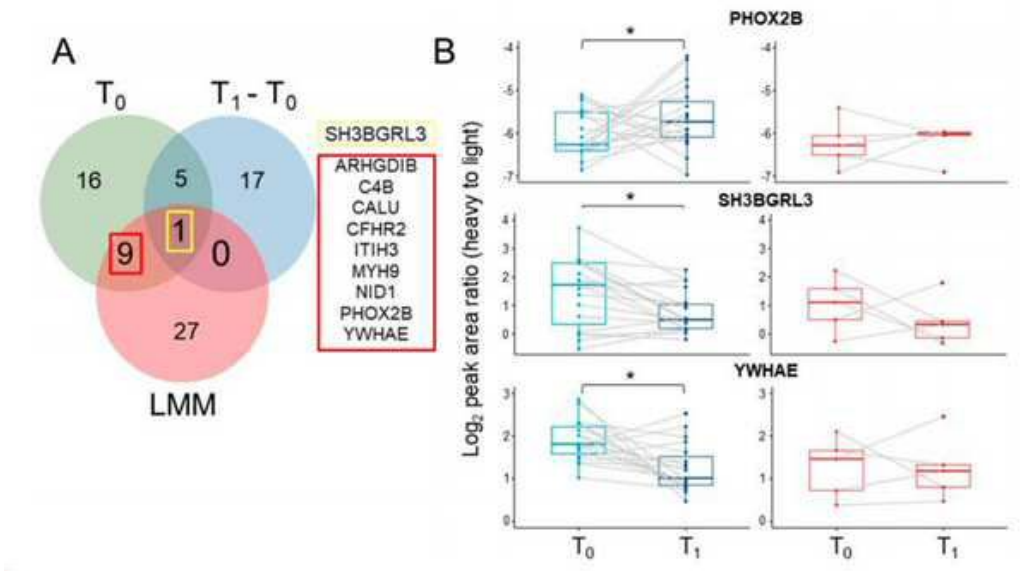
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **주요 우울 장애 치료 효과 조기 예측용 바이오마커**

## (57) 요 약

본 발명은 주요 우울 장애 치료 효과 조기 예측용 바이오마커에 관한 것으로, 더 상세하게는 EPHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 선택적 세로토닌 재흡수 억제제 치료 효과 조기 예측용 바이오마커에 관한 것이다. 본 발명에 따른 바이오마커를 이용하는 경우, 약물 치료의 효능이 좋지 않은 환자를 조기에 발견할 수 있어, 우울증 환자를 치료하는데 필요한 시간과 비용을 줄이고, 우울증 환자의 불필요한 약물에 대한 노출을 최소화할 수 있는 장점이 있다.

*대 표 도* - 도3



(52) CPC특허분류
***G01N 33/6893*** (2013.01)
*G01N 2030/8831* (2013.01)
*G01N 2800/52* (2021.08)

(72) 발명자
**김경곤**
서울특별시송파구 올림픽로 435, 108동 2302호(신천동, 파크리오)

**안희성**
서울특별시 영등포구 당산로 95, 104동 202호(당산동2가, 현대아파트)

이 발명을 지원한 국가연구개발사업

| | |
|---|---|
| 과제고유번호 | 1345313438 |
| 과제번호 | 2017R1D1A1B03028787 |
| 부처명 | 교육부 |
| 과제관리(전문)기관명 | 한국연구재단 |
| 연구사업명 | 이공분야기초연구사업-이공학개인기초연구지원사업-기본연구지원사업 |
| 연구과제명 | 혈중 단백체 및 임상정보의 기계학습 분석을 통한 항우울제 치료반응 조기 예측 알고리듬 개발 및 검증 |
| 기 여 율 | 1/1 |
| 과제수행기관명 | 서울대학교 |
| 연구기간 | 2017.06.01 ~ 2020.05.31 |

# 명 세 서

## 청구범위

### 청구항 1

PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 바이오마커.

### 청구항 2

제1항에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 에스시탈로프람인 것을 특징으로 하는, 바이오마커.

### 청구항 3

제2항에 있어서, 에스시탈로프람 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 하는, 바이오마커.

### 청구항 4

PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 바이오마커 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 제제를 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 조성물.

### 청구항 5

제4항에 있어서, 상기 조성물은 선택적 세로토닌 재흡수 억제제 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 하는, 조성물.

### 청구항 6

제4항에 있어서, 상기 제제는 상기 바이오마커 단백질 또는 이의 단편에 특이적으로 결합하는 항체 또는 이의 항원 결합단편, 또는 압타머인 것을 특징으로 하는, 조성물.

### 청구항 7

제4항에 있어서, 상기 제제는 상기 바이오마커 단백질을 암호화하는 폴리뉴클레오티드에 특이적으로 결합하는 프라이머, 프로브 또는 안티센스 뉴클레오티드인 것을 특징으로 하는, 조성물.

### 청구항 8

제4항 내지 제7항 중 어느 한 항의 조성물을 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 키트.

### 청구항 9

다음 단계를 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측을 위한 정보 제공 방법:

(a) 선택적 세로토닌 재흡수 억제제 치료제가 투여된 대상체에서 분리된 샘플에서 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 바이오마커 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 단계; 및

(b) 상기 측정된 발현 수준을 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하는 단계.

**청구항 10**

제9항에 있어서, 상기 샘플은 혈액, 혈장, 혈청, 소변, 점액, 타액, 눈물, 객담, 척수액, 흉수, 유두 흡인물, 림프액, 기도액, 장액, 비뇨생식관액, 모유, 림프계 체액, 정액, 뇌척수액, 기관계내 체액, 복수, 낭성 종양 체액, 양수액 또는 이들의 조합인 것을 특징으로 하는, 방법.

**청구항 11**

제9항에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 에스시탈로프람인 것을 특징으로 하는, 방법.

**청구항 12**

제9항에 있어서, 상기 샘플은 선택적 세로토닌 재흡수 억제제 투여 후 3 내지 14일에 대상체에서 분리되는 것을 특징으로 하는, 방법.

**청구항 13**

제9항에 있어서, 상기 단백질 또는 그의 단편의 발현 수준을 측정하는 단계는 전기영동, 면역블로팅, 효소 결합 면역흡착 분석법(Enzyme-Linked Immunosorbent Assay: ELISA), 면역 조직 화학 염색, 단백질 칩, 면역침강, 질량분석, 또는 이들의 조합으로 수행; 및/또는

상기 유전자의 발현 수준을 측정하는 단계는 RT-PCR, 경쟁적 RT-PCR(competitive RT-PCR), 실시간 RT-PCR(Real-time RT-PCR), RNase 보호 분석법 (RPA: RNase protection assay), 노던 블랏팅(Northern blotting), DNA 칩, 또는 이들의 조합으로 수행되는 것을 특징으로 하는, 방법.

**청구항 14**

제12항에 있어서, 상기 분리된 샘플에서 측정된 바이오마커 발현 수준이 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여

(i) PHOX2B의 발현이 증가;

(ii) SH3BGRL3 발현이 감소; 및/또는

(iii) YWHAE 발현이 감소; 하는 경우,

상기 선택적 세로토닌 재흡수 억제제는 치료 효과가 있는 것으로 예측하는 것을 특징으로 하는, 방법.

## *발명의 설명*

### *기 술 분 야*

[0001] 본 발명은 주요 우울 장애 치료 효과 조기 예측용 바이오마커에 관한 것으로, 더 상세하게는 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 선택적 세로토닌 재흡수 억제제 치료 효과 조기 예측용 바이오 마커에 관한 것이다.

## 배 경 기 술

[0003] 주요 우울 장애(Major depressive disorder, MDD)는 지속적인 우울감과 활동력 저하를 특징으로 하는 우울증 상태가 지속적 또는 반복적으로 나타나는 정신장애로, 모든 연령과 인종에서 나타나는 질환이다. 증상으로는 우울감과 절망감, 흥미나 쾌락의 현저한 저하, 저하되거나 증가된 식욕과 체중, 수면양의 감소나 증가, 신체적 초조 또는 활동 속도의 지체, 성욕의 상실이나 피로감, 부적절한 죄책감과 책임감, 무가치감, 집중력의 저하 또는 우유부단함, 죽음이나 자살에 대한 생각 등이 있다. 이와 같은 증상이 사회생활은 물론 일상생활에 문제를 일으키는 심각한 수준으로 2주 이상 지속되는 경우, 주요 우울장애로 진단될 수 있다. 주요 우울 장애는 사회적·직업적 생활에 여러 장애를 일으키며, 전체 회복율이 20%에 불과하고 회복된 환자의 80%가 일생 동안 적어도 한 번 이상의 재발을 경험하는 재발률이 매우 높은 질환이다.

[0004] 주요 우울 장애의 치료를 위해서는 항우울제를 처방받게 되는데, 항우울제 중 선택적 세로토닌 재흡수 억제제(selective serotonin reuptake inhibitor, SSRI)는 가장 대표적인 항우울제이다.

[0005] 그러나, 임상에서는 이러한 항우울제를 선택하여 처방하는 과정에서 많은 시행착오를 겪게 되는데, 4-6주 동안 항우울제를 처방받았으나 초기 약물 치료의 효능이 좋지 않은 환자가 상당 비율 발생하고 있고, 이 경우 동일 클래스 내 또는 다른 클래스의 항우울제로 변경하여 치료를 진행하게 된다.

[0006] 이 뿐만 아니라, 항우울제를 이용한 치료 효과가 그다지 우수하지는 않은 것으로 보고되고 있는데, STAR*D(Sequenced Treatment Alternative to Relieve Depression) 연구에 등록한 환자의 약 50 %에서 표준 SSRI 치료에 반응하지 않는 것으로 나타났으며, 약 30% 만이 처음 사용된 항우울제에 반응하여 차도를 나타낸다고 보고되었다. 이 경우, 표준 SSRI 치료에 반응하지 않는 환자에 투여할 두 번째 약물의 선택이 매우 중요하고, 여기서 항우울제의 사용 초기단계에서 항우울제 치료에 대한 반응을 미리 예측할 수 있다면, 우울증 환자를 치료하는데 필요한 시간과 비용을 줄이고 불필요한 약물에 대한 노출을 최소화 할 수 있을 것이다.

[0007] 한편, 프로테오믹스(proteomics)란 샘플에서 발현되는 모든 단백질을 정량 분석하는 기술로, 새로운 분자 바이오마커를 식별하기 위한 강력한 도구 중 하나이며, 주요 우울 장애 환자의 치료에 대한 반응에 있어서도 여러 생물학적 경로를 반영하는 분자 검출이 가능하다(Martins-de-Souza, D. et al., *Eur. Arch. Psychiatry Clin. Neurosci.* 2010, 260, 499-506). 혈장 및 혈청과 같은 말초 체액의 단백질체 분석은 임상 치료에 대한 반응을 예측하고 특히 임상 초기 단계에서 약물 활성을 모니터링하는 데 도움이 된다. 그러나 현재까지 항우울제 치료에 대한 반응을 예측할 수 있는 말초혈액 검체의 단백질체 분석은 거의 없는 실정이며, 액체 크로마토그래피 탠덤 질량 분석(LC-MS/MS) 에서는 여러 혈장 단백질이 6주 치료 기간 동안 항우울제 반응을 예측하는 잠재적 바이오마커가 될 수 있음을 확인하였고, 면역, 내분비 및 대사 기전과 관련된 최대 258개의 혈액 기반 마커에 대한 다중 면역 분석 테스트에서 9개의 마커가 항우울제 치료 반응과 관련된 잠재적인 치료 전 바이오마커가 될 수 있음을 확인하였다 (Turck, C.W. et al., Front. Mol. Neurosci. 2017, 10, 272; Chan, M.K. et al., J. Psychiatr. Res. 2016, 83, 249-259).

[0008] 종단 데이터(longitudinal data)는 일반적으로 생물의학 연구에 사용되는데, 통계 분석에는 혼합 효과 모델(mixed-effect models, MEMs)과 일반화 추정 방정식(generalized estimating equations, GEEs)이 널리 적용된다. 이 중 MEM은 전체 가능성 방법(full-likelihood method)을 기반으로 하는 대상체 수준 접근 방식(subject-level approach)으로, 동일한 통계 단위에서 반복측정이 이루어지거나 관련 통계 단위의 클러스터에서 측정이 이루어지는 설정에서 특히 유용하다. 누락된 값을 처리하는 데 있어 장점이 있기 때문에 혼합 효과 모델은 분산에 대한 반복적인 측정 해석과 같은 보다 전통적인 접근 방식보다 선호된다. 반대로 GEE는 부분 우도 함수(partial-likelihood function)에 의존하는 모집단 수준 모델(population-level model)이다. 프로테오믹스 연구에서는 MEM 방법이 GEE 방법보다 널리 사용되는데, 이는 원하는 효과를 고정한 후 실제 MS로 기술적 또는 생물학적 반복 측정의 무작위 효과를 측정하여 추정할 수 있기 때문이다. 또한 2회 이상의 MS 후 동일한 샘플의 반복 측정에 대한 임의의 영향의 변화(variance of any effect)를 반영하는 것이 기술적으로 더 쉽다. 대조적으로, GEE는 준우도(quasi likelihood)를 이용한 상관 구조의 설정 오류가 강건하여 표본에 대한 수정된 분산 추정 방법이 많이 개발된 바 있다.

[0009] 본 발명자들은 우울증 환자를 치료하는데 필요한 시간과 비용을 줄이고 불필요한 약물에 대한 노출을 최소화고자, 항우울제의 사용 초기단계에서 항우울제 치료에 대한 반응을 미리 예측할 수 있는 바이오마커를 동정하고자 예의 노력하였으며, 그 결과, 에스시탈로프람을 복용하고 있는 우울증 환자 중 에스시탈로프람이 치료 효과를

나타내는 우울증 환자(이하, '반응자'로 지칭됨)과 에스시탈로프람이 치료 효과를 나타내지 않는 우울증 환자(이하, '무반응자'로 지칭됨)에서 발현 양상이 유의하게 차이가 나는 단백질 바이오마커를 새로이 동정함으로써 본 발명을 완성하였다.

## 선행기술문헌

### 비특허문헌

[0011] (비특허문헌 0001) Martins-de-Souza, D. et al., Eur. Arch. Psychiatry Clin. Neurosci. 2010, 260, 499-506

(비특허문헌 0002) Turck, C.W. et al., Front. Mol. Neurosci. 2017, 10, 272

(비특허문헌 0003) Chan, M.K. et al., J. Psychiatr. Res. 2016, 83, 249-259

## 발명의 내용

### 해결하려는 과제

[0012] 본 발명은 선택적 세로토닌 재흡수 억제제의 치료 효과를 조기에 예측할 수 있는 신규한 바이오마커와 이의 용도에 관한 것이다.

### 과제의 해결 수단

[0013] 상기 목적을 달성하기 위하여, 본 발명은 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 바이오마커를 제공한다.

[0014] 본 발명에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 에스시탈로프람인 것을 특징으로 한다.

[0015] 본 발명에 있어서, 상기 바이오마커는 에스시탈로프람 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 한다.

[0016] 본 발명은 또한, PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 제제를 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 조성물.을 제공한다.

[0017] 본 발명에 있어서, 상기 조성물은 선택적 세로토닌 재흡수 억제제 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 한다.

[0018] 본 발명에 있어서, 상기 제제는 상기 바이오마커 단백질 또는 이의 단편에 특이적으로 결합하는 항체 또는 이의 항원 결합단편, 또는 압타머인 것을 특징으로 한다.

[0019] 본 발명에 있어서, 상기 제제는 상기 바이오마커 단백질을 암호화하는 폴리뉴클레오티드에 특이적으로 결합하는 프라이머, 프로브 또는 안티센스 뉴클레오티드인 것을 특징으로 한다.

[0020] 본 발명은 또한, 상기 조성물을 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 키트를 제공한다.

[0021] 본 발명은 또한, 다음 단계를 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측을 위한 정보 제공 방법을 제공한다:

[0022] (a) 선택적 세로토닌 재흡수 억제제 치료제가 투여된 대상체에서 분리된 샘플에서 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 바이오마커 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 단계; 및

[0023] (b) 상기 측정된 발현 수준을 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하는 단계.

[0024] 본 발명에 있어서, 상기 샘플은 혈액, 혈장, 혈청, 소변, 점액, 타액, 눈물, 객담, 척수액, 흉수, 유두 흡인물, 림프액, 기도액, 장액, 비뇨생식관액, 모유, 림프계 체액, 정액, 뇌척수액, 기관계내 체액, 복수, 낭성 종양 체

액, 양수액 또는 이들의 조합인 것을 특징으로 한다.

[0025] 본 발명에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 에스시탈로프람인 것을 특징으로 한다.

[0026] 본 발명에 있어서, 상기 샘플은 선택적 세로토닌 재흡수 억제제 투여 후 3 내지 14일에 대상체에서 분리되는 것을 특징으로 한다.

[0027] 본 발명에 있어서, 상기 단백질 또는 그의 단편의 발현 수준을 측정하는 단계는 전기영동, 면역블로팅, 효소 결합 면역흡착 분석법(Enzyme-Linked Immunosorbent Assay: ELISA), 면역 조직 화학 염색, 단백질 칩, 면역침강, 질량분석, 또는 이들의 조합으로 수행; 및/또는 상기 유전자의 발현 수준을 측정하는 단계는 RT-PCR, 경쟁적 RT-PCR(competitive RT-PCR), 실시간 RT-PCR(Real-time RT-PCR), RNase 보호 분석법 (RPA: RNase protection assay), 노던 블랏팅(Northern blotting), DNA 칩, 또는 이들의 조합으로 수행되는 것을 특징으로 하는, 방법.한다.

[0028] 본 발명에 있어서, 상기 분리된 샘플에서 측정된 바이오마커 발현 수준이 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여

[0029] (i) PHOX2B의 발현이 증가;

[0030] (ii) SH3BGRL3 발현이 감소; 및/또는

[0031] (iii) YWHAE 발현이 감소; 하는 경우,

[0032] 상기 선택적 세로토닌 재흡수 억제제는 치료 효과가 있는 것으로 예측하는 것을 특징으로 하는 한다.

## *발명의 효과*

[0033] 본 발명에 따른 바이오마커를 이용하는 경우, 약물 치료의 효능이 좋지 않은 환자를 조기에 발견할 수 있어, 우울증 환자를 치료하는데 필요한 시간과 비용을 줄이고, 우울증 환자의 불필요한 약물에 대한 노출을 최소화할 수 있는 장점이 있다.

## *도면의 간단한 설명*

[0035] 도 1은 약물 투여에 따라 차등 발현되는 단백질의 변화를 식별하기 위한 혈장 단백질체 분석 결과 및 이들 단백질 기능을 도시한 것으로, 도 1A는 약물 반응자(Responders)와 약물 무반응자(Non-responders)에서 약물 투여의 시작점과 비교하여 차등적으로 발현 양상이 변화하는 단백질(diffrentially abundant proteins)을 약물 투여 1주, 4주, 10주에 확인한 결과이고, 도 1B는 $T_0$와 비교하여 $T_1$, $T_4$, $T_{10}$에서 차등 발현되는 단백질을 벤 다이어그램으로 표시한 결과이며, 도 1C는 $T_0$와 비교하여 $T_1$, $T_4$, $T_{10}$에서 반응자와 무반응자 간 차등적으로 발현 양상이 변화하는 단백질들의 유전자 온톨로지(gene ontology)를 나타낸다.

도 2는 선형 혼합 모델(LMM)의 반응/시간 상호 작용에서 크게 상이한 것으로 확인된 37개 단백질의 친화성 전파 클러스터링, 프로파일 분석 및 데이터베이스 검색 결과를 나타낸 것이다. 도 2A는 t-SNE 기반 친화성 전파 클러스터링 (t-SNE-based affinity propagation clustering)으로 7개의 단백질 클러스터를 식별한 결과이고, 도 2B는 시간 경과에 따른 반응자와 무반응자 사이의 단백질 양의 변화를 확인한 결과이며, 도 2C는 PsyGeNet에서 확인된 정신 질환과 연관된 14개 단백질을 나타낸 결과이고, 도 2D는 Enrichr의 DrugMatrix 범주에서 발견된 10개의 단백질이 선택적 세로토닌 재흡수 억제제(SSRI)에 대한 쥐 조직 및 세포의 반응과의 연관성이 있음을 보여주는 결과이다.

도 3은 약물 투여 후 조기에 반응하는 10가지 후보 예측 바이오마커에 대한 액체 크로마토그래피-다중 반응 모니터링/질량 분석(LC-MRM/MS) 검증 결과이다. 도 3A는 3가지 통계 분석(두 그룹 간의 $T_0$ 대 $T_1$ 및 $T_1$-$T_0$, LMM의 반응/시간 항)에 의한 후보 바이오마커 검출을 위한 벤 다이어그램을 나타낸다. 도 3B는 반응자(파란색) 및 무반응자(빨간색)에서 LC-MRM/MS에 의해 결정된 $T_1$ 대 $T_0$에서 PHOX2B, SH3BGRL3 및 YWHAE의 풍부도(abundance)를 나타내는 박스이다. * 별표는 0.05 수준에서 유의한 조정된 *p*-value를 나타낸다.

## *발명을 실시하기 위한 구체적인 내용*

[0036] 다른 식으로 정의되지 않는 한, 본 명세서에서 사용된 모든 기술적 및 과학적 용어들은 본 발명이 속하는 기술분야에서 숙련된 전문가에 의해서 통상적으로 이해되는 것과 동일한 의미를 갖는다. 일반적으로, 본 명세서에서 사용된 명명법 및 이하에 기술하는 실험 방법은 본 기술분야에서 잘 알려져 있고 통상적으로 사용되는 것이다.

[0038] 본 발명에서는 선택적 세로토닌 재흡수 억제제 중 하나인 에스시탈로프람을 주요 우울 장애 환자에 투여하고, 약물 효능이 발휘되는 반응자와 약물 효능이 발휘되지 않는 무반응자에서 기준일(baseline) 및 치료 후 후속 방문을 통해 시간에 따른 대리 혈장 단백질(surrogate plasma protein)의 변화를 측정하고, 선형 혼합 모델(linear mixed model, LMM) 분석 결과를 중심으로 단백질 변화의 생물학적 의미를 파악하였으며, 동일한 데이터를 GEE로 분석한 후 결과를 비교하였다.

[0039] 이후, 본 발명에서는 항우울제 에스시탈로프람 투여 전후에 초기(0-1주), 중간(1-4주) 또는 후기(4-10주) 반응을 모니터링할 수 있는 후보 혈액 기반 단백질 바이오마커를 식별하기 위하여 LC-MS/MS 프로파일링이 수행하였다. 또한, 본 발명에서는 항우울제 치료 후 혈장 단백질 농도의 변화가 우울 증상의 중증도 변화와 관련이 있는지 여부를 평가하였다.

[0040] 에스시탈로프람을 복용하고 있는 우울증 환자 10명(반응자 5명, 무반응자 5명)을 대상으로 10주간 치료하는 동안 4개의 시점에서 혈액 샘플을 수집하였으며, 두 그룹 간의 단백질 풍부도(protein abundance)에 차이가 나는 혈장 단백질을 프로파일링한 결과, 약물 투여 1주 후, PHOX2B 단백질 수준은 반응자에서 유의하게 증가되는 반면 SH3BGRL3 및 YWHAE 단백질 수준은 반응자에서 유의하게 감소되는 것으로 나타났고, 이들 3개 단백이 무반응자에서는 유의한 변화를 유도하지 못하는 것으로 확인되었다. 이와 같은 결과는 19명의 반응자 및 5명의 무반응자의 24명의 환자를 대상으로 추가 검증되었다.

[0042] 따라서, 본 발명은 일 관점에서, PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 바이오마커에 관한 것이다.

[0043] 본 발명에서, PHOX2B는 paired like homeobox 2B로 CCHS, NBLST2, NBPhox, PMX2B로 표시될 수 있으며, 인간에서 Gene ID: 8929의 뉴클레오티드 서열을 포함하는 암호화된 단백질일 수 있다. 상기 PHOX2B 단백질은 예를 들어, Uniprot Accession No. Q99453, GeneBank Accession No. NM_003924.4, NP_003915.2 등의 뉴클레오티드 또는 아미노산 서열을 포함한 단백질일 수 있다. 상기 PHOX2B 단백질은 동형 단백질(isoform) 또는 이의 전구체를 포함하나, 본 발명에서는 이를 대표하여 서열번호 1로 표시되는 아미노산 서열을 기재하였다.

[0044] 본 발명에서, SH3BGRL3는 SH3 domain binding glutamate rich protein like 3로 HEL-S-297, SH3BP-1, TIP-B1로 표시될 수 있으며, 인간에서 Gene ID: 83442의 뉴클레오티드 서열을 포함하는 암호화된 단백질일 수 있다. 상기 SH3BGRL3 단백질은 예를 들어, Uniprot Accession No. Q9H299, GeneBank Accession No. NM_031286.4, NP_112576.1 등의 뉴클레오티드 또는 아미노산 서열을 포함한 단백질일 수 있다. 상기 SH3BGRL3 단백질은 동형 단백질(isoform) 또는 이의 전구체를 포함하나, 본 발명에서는 이를 대표하여 서열번호 2로 표시되는 아미노산 서열을 기재하였다.

[0045] 본 발명에서, YWHAE는 tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon으로 14-3-3E, HEL2, KCIP-1, MDCR, MDS로 표시될 수 있으며, 인간에서 Gene ID: 7531의 뉴클레오티드 서열을 포함하는 암호화된 단백질일 수 있다. 상기 YWHAE 단백질은 예를 들어, Uniprot Accession No. P62258, GeneBank Accession No. NM_006761.5, NP_006752.1 등의 뉴클레오티드 또는 아미노산 서열을 포함한 단백질일 수 있다. 상기 YWHAE 단백질은 동형 단백질(isoform) 또는 이의 전구체를 포함하나, 본 발명에서는 이를 대표하여 서열번호 3으로 표시되는 아미노산 서열을 기재하였다.

[0047] [PHX2B 서열정보]

[0048] <서열번호 1> UniProtKB Sequence: Q99453 (314 aa)

[0049] MYKMEYSYLNSSAYESCMAGMDTSSLASAYADFSSCSQASGFQYNPIRTTFGATSGCPSLTPGSCSLGTLRDHQSSPYAAVPYKLFTDHGGLNEKRKQRRIR
TTFTSAQLKELERVFAETHYPDIYTREELALKIDLTEARVQVWFQNRRAKFRKQERAAAAAAAAAKNGSSGKKSDSSRDDESKEAKSTDPDSTGGPGPNPNP
TPSCGANGGGGGGPSPAGAPGAAGPGGPGGEPGKGGAAAAAAAAAAAAAAAAAAAAGGLAAAGGPGQGWAPGPGPITSIPDSLGGPFASVLSSLQRPNGAKA

ALVKSSMF

[0051] [SH3BGRL3 서열정보]

[0052] <서열번호 2> UniProtKB Sequence: Q9H299 (93 aa)

[0053] MSGLRVYSTSVTGSREIKSQQSEVTRILDGKRIQYQLVDISQDNALRDEMRALAGNPKATPPQIVNGDQYCGDYELFVEAVEQNTLQEFLKLA

[0055] [YWHAE 서열정보]

[0056] <서열번호 3> UniProtKB Sequence: P62258 (255 aa)

[0057] MDDREDLVYQAKLAEQAERYDEMVESMKKVAGMDVELTVEERNLLSVAYKNVIGARRASWRIISSIEQKEENKGGEDKLKMIREYRQMVETELKLICCDILD
VLDKHLIPAANTGESKVFYYKMKGDYHRYLAEFATGNDRKEAAENSLVAYKAASDIAMTELPPTHPIRLGLALNFSVFYYEILNSPDRACRLAKAAFDDAIA
ELDTLSEESYKDSTLIMQLLRDNLTLWTSDMQGDGEEQNKEALQDVEDENQ

[0059] 본 명세서에서 용어 "바이오마커"는 항우울제 투여에 따른 반응성을 구분할 수 있는 물질로, 항우울제 투여에 따른 반응군과 무반응군에서 발현 양상의 차이를 보이기 때문에, 예컨대 mRNA 또는 단백질 수준에서의 발현 수준이 반응군과 무반응군에서 서로 상이하기 때문에, 그 발현 수준을 평가함으로써 항우울제 투여에 대한 반응성을 판단하거나 평가할 수 있는 물질을 의미한다.

[0060] 본 발명에서 예측이란, 항우울제 투여에 의해 치료 효과가 발휘되는지 여부를 판단하거나 평가하는 것을 의미하며, 예후를 판단하거나 치료 효과를 진단한다는 의미를 모두 포함한다.

[0061] 바람직하게는 본 발명에 따른 바이오마커는 선택적 세로토닌 재흡수 억제제의 치료 효과를 미리 또는 조기에, 예컨대 치료 개시일(기준일)로부터 14일 이내, 예컨대 7일 이내에 예측할 수 있다.

[0062] 본 발명에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 바람직하게는 에스시탈로프람일 수 있다.

[0063] 본 발명에서는 상기 바이오마커로부터 에스시탈로프람 투여 후 4주 이내, 3주 이내, 2주 이내 또는 1주 이내 샘플을 채취하여, 향후 치료 효과 여부의 예측이 가능한 것을 특징으로 할 수 있으나, 샘플 채취 시기가 이에 한정되지는 않는다. 예컨대, 본 발명에서는 에스시탈로프람 투여 후 1일 이내, 2일 이내, 3일 이내, 4일 이내, 5일 이내, 6일 이내, 7일 이내, 8일 이내, 9일 이내, 10일 이내, 11일 이내, 12일 이내, 13일 이내 14일 이내, 15일 이내, 16일 이내, 17일 이내, 18일 이내, 19일 이내, 20일 이내, 21일 이내, 22일 이내, 23일 이내, 24일 이내, 25일 이내, 26일 이내, 27일 이내 또는 28일 이내 샘플을 채취하여 향후 치료 효과 여부의 예측이 가능한 것을 특징으로 할 수 있다.

[0064] 일 양태로서, 본 발명에서는 상기 바이오마커로부터 에스시탈로프람 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 할 수 있다.

[0066] 다른 관점에서, 본 발명은 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 바이오마커 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 제제를 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 조성물에 관한 것이다.

[0067] 본 발명에 있어서, 상기 조성물은 선택적 세로토닌 재흡수 억제제 투여 후 적어도 일주일 내 치료 효과 예측이 가능한 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0068] 본 발명에 있어서, 상기 제제는 상기 바이오마커 단백질 또는 이의 단편에 특이적으로 결합하는 항체 또는 이의 항원 결합단편, 또는 압타머인 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0069] 본 발명에 있어서, 상기 제제는 상기 바이오마커 단백질을 암호화하는 폴리뉴클레오티드에 특이적으로 결합하는 프라이머, 프로브 또는 안티센스 뉴클레오티드인 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0070] 본 발명에서, 상기 바이오마커 단백질의 단편은 면역원성 단편일 수 있다. 즉, 상기 바이오마커 단백질의 단편은 상기 바이오마커 단백질이 항체에 의해 인식될 수 있도록 하는 하나 이상의 에피토프(epitope)를 가지고 있

는 단편일 수 있다.

[0071] 본 발명에서, 상기 "단백질의 발현 수준 측정"은 대상체의 시료에서 바이오마커 단백질의 존재 여부와 발현정도를 확인하는 과정으로, 바이오마커 단백질의 양을 측정하는 것일 수 있다. 또는 바이오마커 단백질을 암호화하는 폴리뉴클레오티드, 예를 들어 mRNA 발현 수준을 측정함으로써 확인 가능한 것일 수 있다.

[0072] 일 양태로서 바이오마커 단백질의 발현 수준 측정은 상기 바이오마커 단백질 또는 이의 단편에 대해 특이적으로 결합하는 항체 또는 항원 결합 단편 또는 압타머를 이용하여 단백질을 확인할 수 있다.

[0073] 바이오마커 단백질의 발현 수준을 측정하기 위한 다른 양태의 분석방법으로는 웨스턴블랏팅, ELISA (enzyme linked immunosorbent assay), 방사선 면역분석 (RIA: Radioimmunoassay), 방사 면역 확산법 (radioimmunodiffusion), 오우크테로니 (Ouchterlony) 면역 확산법, 로케트 (rocket) 면역전기영동, 조직면역염색, 면역침전 분석법 (Immunoprecipitation Assay), 보체 고정 분석법 (Complement Fixation Assay), 유세포 분석 (Fluorescence Activated Cell Sorter, FACS), 및 단백질 칩 (protein chip), 질량 분석(Mass Spectrometry) 등이 있으나, 이에 한정되는 것은 아니다.

[0074] 본 명세서에서 용어 "항체"는 당해 분야에서 공지된 용어로서 항원성 부위에 대해서 지시되는 특이적인 단백질 분자를 의미할 수 있다. 일 양상에 따른 항체의 형태는 특별히 한정되지 않으며 폴리클로날 항체, 모노클로날 항체 또는 항원 결합성을 갖는 것이면 그것의 일부도 포함되고 모든 면역글로불린 항체가 포함될 수 있으며, 나아가, 인간화 항체 등의 특수항체도 포함될 수 있다. 일 양태에 따른 항체는 2개의 전체 길이의 경쇄 및 2개 의 전체 길이의 중쇄를 가지는 완전한 형태뿐만 아니라 항체 분자의 기능적인 단편을 포함할 수 있다. 항체 분자의 기능적인 단편은 적어도 항원 결합기능을 보유하고 있는 단편을 의미하며, 예를 들어, Fab, F(ab'), F(ab')2, Fv, scFv 등일 수 있다.

[0075] 상기 "유전자의 발현 수준"은 해당 유전자의 mRNA의 발현 수준 또는 해당 유전자에 의해 코딩되는 단백질의 발현 수준을 측정함으로써 확인 가능한 것일 수 있다. 일 양태로서, 발현 수준에 차이가 나는 유전자를 바이오마커로 활용할 수 있으며, 구체적으로는, 해당 유전자의 발현 수준에 따라 대상체에 대한 항우울제의 치료 효과가 다르게 판단되는 것일 수 있다.

[0076] 본 명세서에서 용어 "유전자"는 단백질 코딩 또는 전사 시에 또는 다른 유전자 발현의 조절시에 기능적 역할을 하는 임의의 핵산 서열 또는 그의 일부를 의미할 수 있다. 유전자는 기능적 단백질을 코딩하는 모든 핵산이거나, 단백질을 코딩 또는 발현하는 핵산의 일부만으로 이루어질 수 있다. 핵산 서열은 엑손, 인트론, 개시 또는 종료 영역, 프로모터 서열, 다른 조절 서열 또는 유전자에 인접한 특유한 서열 내에 유전자 이상을 포함할 수 있다.

[0077] 유전자 발현 수준 측정을 위한 분석방법으로는 RT-PCR, 경쟁적 RT-PCR(competitive RT-PCR), 실시간 RT-PCR(Real-time RT-PCR), RNase 보호 분석법(RPA: RNase protection assay), 노던 블랏팅(Northern blotting), 또는 DNA 칩 등이 있으나, 이에 한정되는 것은 아니다.

[0078] 상기 단백질을 암호화하는 mRNA 발현양을 측정하는 제제는 일 양태에 따르면 유전자의 mRNA에 대해 특이적으로 결합하는 프라이머, 프로브, 또는 안티센스 뉴클레오티드이며, 상기 유전자의 핵산 서열이 알려져 있으므로, 당업자는 상기 서열을 바탕으로 이들 유전자의 mRAN에 대해 특이적으로 결합하는 프라이머 또는 프로브를 디자인할 수 있다.

[0079] 본 명세서에서 용어 "프라이머"는 짧은 자유 3-말단 수산화기(free 3' hydroxyl group)를 가지는 핵산 서열로 상보적인 템플레이트(template)와 염기쌍을 형성할 수 있고 템플레이트 가닥 복사를 위한 시작 지점으로서 작용하는 짧은 핵산 서열을 말한다. 프라이머는 적절한 완충용액 및 온도에서 중합반응을 위한 시약(즉, DNA 폴리머라아제 또는 역전사효소) 및 상이한 4 가지의 뉴클레오사이드 트리포스페이트의 존재 하에서 DNA 합성을 개시할 수 있다. 또한, 프라이머는, 7개 내지 50개의 뉴클레오타이드 서열을 가진 센스 및 안티센스 핵산으로서, DNA 합성의 개시점으로 작용하는 프라이머의 기본 성질을 변화시키지 않는 추가의 특징을 혼입할 수 있다. 프라이머의 서열은 반드시 주형의 서열과 정확히 동일할 필요는 없으며, 충분히 상보적이어서 주형과 혼성화될 수 있으면 된다. 프라이머의 위치 혹은 프라이머 결합부위는 프라이머가 혼성화하는 표적 DNA 절편을 의미할 수 있다. 본 발명의 일 구체예에 따르면 상기 유전자의 mRNA의 센스 및 안티센스 프라이머를 이용하여 PCR 증폭을 실시하여 원하는 생성물의 생성량의 측정을 통해 진단할 수 있다. PCR 조건, 센스 및 안티센스 프라이머의 길이는 당업계에 공지된 기술에 따라 적절히 선택될 수 있다.

[0080] 본 명세서에서 용어 "프로브"란 mRNA외 특이적으로 결합을 이룰 수 있는 짧게는 수 염기 내지 길게는 수백 염기

에 해당하는 RNA 또는 DNA 등의 핵산 단편을 의미하며 라벨링되어 있어서 특정 mRNA의 존재 유무, 발현양을 확인할 수 있다.

[0081] 프로브는 올리고뉴클레오타이드(oligonucleotide) 프로브, 단쇄 DNA(single strand DNA) 프로브, 이중쇄 DNA(double strand DNA)프로브, RNA 프로브 등의 형태로 제작될 수 있다. 일 양상에 따른 유전자의 폴리뉴클레오티드와 상보적인 프로브를 이용하여 혼성화를 실시하여, 혼성화 여부 및 정도를 통해 방사선 피폭 여부 및 정도를 판단할 수 있다. 적당한 프로브의 선택 및 혼성화 조건은 당업계에 공지된 것을 기초로 변형할 수 있다

[0082] 상기 프라이머 또는 프로브는 포스포르아미다이트 고체 지지체 방법 또는 기타 널리 공지된 방법을 사용하여 화학적으로 합성할 수 있다. 이러한 핵산 서열은 또한 당해 분야에 공지된 많은 수단을 이용하여 변형시킬 수 있다. 이러한 변형의 비제한적인 예로는 메틸화, 캡화, 천연 뉴클레오티드 하나 이상의 동족체로의 치환 및 뉴클레오티드 간의 변형, 예를 들어, 메틸 포스포네이트, 포스포트리에스테르, 포스포로아미데이트, 카바메이트 등과 같은 하전되지 않은 연결체 또는 포스포로티오에이트, 포스포로디티오에이트 등과 같은 하전된 연결체로의 변형이 있다.

[0084] 또 다른 관점에서, 본 발명은 상기 조성물을 포함하는, 선택적 세로토닌 재흡수 억제제 치료 효과 예측용 키트에 관한 것이다.

[0085] 상기 키트에 사용되는 상기 유전자, 상기 단백질, 상기 단백질의 단편, 발현 수준, 및 발현 수준의 측정은 상술한 바와 같으며, 중복 기재를 생략한다.

[0086] 상기 키트는 세포 용해 완충액(lysis buffer)을 더 포함할 수 있다. 상기 세포 용해 완충액은 상용화된 제품 등 다양한 물질을 포함할 수 있다. 상기 세포 용해 완충액은 HEPES, $MgCl_2$, NaCl, 이미다졸, DNAse 및 β-mercaptoethanol이 포함될 수 있고, 단백질의 분해를 방지하기 위해 프로테이즈 저해제인 PMSF (phenylmethanesulfonylfluoride)가 첨가될 수 있으나, 이에 한정되지 않는다.

[0087] 상기 키트는 유전자 또는 단백질을 분해 (Digestion)하는 시약을 더 포함할 수 있다. 상기 유전자 분해 시약은 제한효소 등을 포함할 수 있다. 상기 단백질 분해 시약은 트립신, 트립신/EDTA, 또는 TrypLE 등을 포함할 수 있으나, 상기 유전자 또는 단백질 분해 시약은 이에 한정되지 않는다.

[0088] 상기 키트는 동위원소로 라벨링된 표준 펩타이드(Isotope labeled standard peptide) 혼합물을 더 포함할 수 있다.

[0089] 상기 키트는 데이터 분석 또는 통계처리를 할 수 있는 프로그램을 더 포함할 수 있다. 상기 프로그램은 유전자 또는 단백질의 발현 수준을 분석하는 것일 수 있다. 상기 데이터 분석 또는 통계처리 프로그램은 Skyline Software, ProteoWizard Software, SCIEX OS Software, SPSS Statistics Software, MedCalc Software, MultiQuant Software, MasterView Software 또는 Cliquid Software 중에서 선택된 하나 이상일 수 있으나, 이에 한정되지 않는다.

[0090] 상기 키트는 진단에 필요한 시료를 더 포함할 수 있다. 상기 키트는 항체, 항체의 면역학적 검출을 위하여 기질, 적합한 완충용액, 발색 효소 또는 형광물질로 표지된 2차 항체, 또는 발색 기질을 포함할 수 있다. 상기 키트는 압타머, 프라미어, 프로브 또는 올리고뉴클레오티드를 포함할 수 있다.

[0091] 상기 키트는 RT-PCR(Reverse transcription polymerase chain reaction) 키트, DNA 칩 키트, ELISA(Enzyme linked immunosorbent assay) 키트, 단백질 칩 키트, 래피드(rapid) 키트 또는 MRM(Multiple reaction monitoring) 키트일 수 있다.

[0093] 또 다른 관점에서, 본 발명은 선택적 세로토닌 재흡수 억제제 치료 효과 예측을 위한 정보 제공 방법에 관한 것이다:

[0094] (a) 선택적 세로토닌 재흡수 억제제 치료제가 투여된 대상체에서 분리된 샘플에서 PHOX2B, SH3BGRL3 및 YWHAE로 구성된 군에서 선택되는 하나 이상의 바이오마커 단백질, 이의 단편 또는 상기 단백질을 코딩하는 유전자의 발현 수준을 측정하는 단계; 및

[0095] (b) 상기 측정된 발현 수준을 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하는 단계.

[0096] 본 발명에 있어서, 상기 샘플은 혈액, 혈장, 혈청, 소변, 점액, 타액, 눈물, 객담, 척수액, 흉수, 유두 흡인물, 림프액, 기도액, 장액, 비뇨생식관액, 모유, 림프계 체액, 정액, 뇌척수액, 기관계내 체액, 복수, 낭성 종양 체액, 양수액 또는 이들의 조합인 것을 특징으로 할 수 있으며, 바람직하게는 혈액, 혈장 또는 혈청 샘플일 수 있으나, 이에 한정되지는 않는다.

[0097] 본 발명에 있어서, 상기 선택적 세로토닌 재흡수 억제제는 에스시탈로프람인 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0098] 본 발명에 있어서, 상기 샘플은 선택적 세로토닌 재흡수 억제제 투여 후 4주 이내, 3주 이내, 2주 이내 또는 1주 이내 대상체에서 분리되는 것을 특징으로 하나, 이에 한정되지는 않는다. 예컨대, 본 발명에서는 에스시탈로프람 투여 후 1일 이내, 2일 이내, 3일 이내, 4일 이내, 5일 이내, 6일 이내, 7일 이내, 8일 이내, 9일 이내, 10일 이내, 11일 이내, 12일 이내, 13일 이내 14일 이내, 15일 이내, 16일 이내, 17일 이내, 18일 이내, 19일 이내, 20일 이내, 21일 이내, 22일 이내, 23일 이내, 24일 이내, 25일 이내, 26일 이내, 27일 이내 또는 28일 이내 분석을 위한 샘플이 대상체에서 분리되는 것을 특징으로 할 수 있다.

[0099] 바람직하게는, 상기 샘플은 선택적 세로토닌 재흡수 억제제 투여 후 3 내지 14일에 대상체에서 분리되는 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0100] 본 발명에 있어서, 상기 단백질 또는 그의 단편의 발현 수준을 측정하는 단계는 전기영동, 면역블로팅, 효소결합 면역흡착 분석법(Enzyme-Linked Immunosorbent Assay: ELISA), 면역 조직 화학 염색, 단백질 칩, 면역침강, 질량분석, 또는 이들의 조합으로 수행; 및/또는

[0101] 상기 유전자의 발현 수준을 측정하는 단계는 RT-PCR, 경쟁적 RT-PCR(competitive RT-PCR), 실시간 RT-PCR(Real-time RT-PCR), RNase 보호 분석법 (RPA: RNase protection assay), 노던 블랏팅(Northern blotting), DNA 칩, 또는 이들의 조합으로 수행되는 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0102] 본 발명에 있어서, 상기 방법은, 상기 분리된 샘플에서 측정된 바이오마커 발현 수준이 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여

[0103] (i) PHOX2B의 발현이 증가;

[0104] (ii) SH3BGRL3 발현이 감소; 및/또는

[0105] (iii) YWHAE 발현이 감소; 하는 경우,

[0106] 상기 선택적 세로토닌 재흡수 억제제는 치료 효과가 있는 것으로 예측하는 단계를 추가로 포함할 수 있다.

[0107] 본 발명에 있어서, 상기 치료 효과는 상기 선택적 세로토닌 재흡수 억제제가 투여된 후 약 4주 및/또는 약 10주에 치료 효과가 있는 것으로 예측하는 것을 특징으로 할 수 있으나, 이에 한정되지는 않는다.

[0108] 본 발명에서, 용어 '선택적 세로토닌 재흡수 억제제가 투여되기 전'은 기준일(baseline)과 상호 교환적으로 사용될 수 있다.

[0109] 본 발명에서, 상기 (i) PHOX2B의 발현은 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 통계적으로 약물반응군에서만 유의미하게 그 발현양이 변화하였고, 약 1.28배 그 양이 증가하였다. 따라서, 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 PHOX2B의 발현양이 약 20% 이상 증가하는 경우, 상기 선택적 세로토닌 재흡수 억제제는 치료 효과를 발휘하는 것으로 예측될 수 있다. 한편, 무반응군에서는 통계적으로 유의미한 변화가 확인되지 않았다.

[0110] 본 발명에서, 상기 (ii) SH3BGRL3의 발현은 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 통계적으로 약물반응군에서만 유의미하게 그 발현양이 변화하였고, 약 0.61배 그 양이 감소하였다. 따라서, 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 SH3BGRL3의 발현양이 약 40% 이상 감소하는 경우, 바람직하게는 약 50% 이상 감소하는 경우, 더욱 바람직하게는 약 60% 이상 감소하는 경우, 상기 선택적 세로토닌 재흡수 억제제는 치료 효과를 발휘하는 것으로 예측될 수 있다. 한편, 무반응군에서는 통계적으로 유의미한 변화가 확인되지 않았다.

[0111] 본 발명에서, 상기 (iii) YWHAE의 발현은 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 통계적으로 약물반응군에서만 유의미하게 그 발현양이 변화하였고, 0.64배 그 양이 감소하였다. 따라서, 상기 선택적 세로토닌 재흡수 억제제가 투여되기 전과 비교하여 YWHAE의 발현양이 약 40% 이상 감소하는 경우, 바람직하

게는 약 50% 이상 감소하는 경우, 더욱 바람직하게는 약 60% 이상 감소하는 경우, 상기 선택적 세로토닌 재흡수 억제제는 치료 효과를 발휘하는 것으로 예측될 수 있다. 한편, 무반응군에서는 통계적으로 유의미한 변화가 확인되지 않았다.

[0113] 본 발명은 환자의 혈장 단백질체에서 시간 의존적 종단 변화에 따른 신뢰할 수 있는 후보 바이오마커를 도출하였다는 것에 의의가 있다. 본 발명에서는 초기 약물 반응을 예측하는 바이오마커를 1차로 확인한 후, 19명의 반응자와 5명의 무반응자에서 액체 크로마토그래피 다중 반응 모니터링/질량 분석(liquid chromatography multiple reaction monitoring/mass spectrometry, LC-MRM/MS) 기술에 의한 추가 검증을 통하여 본 발명의 신뢰도를 검증하였다.

[0114] 따라서, 본 발명에 따른 바이오마커는 약물 투여 초기에, 바람직하게는 일주일 이내에, 해당 약물이 주요 우울 장애에 치료 효과를 발휘하는지 여부를 정확하고 객관적으로 확인할 수 있는 지표를 제공하고, 이를 기초로 약물의 지속 투여 또는 교체 여부를 신속하고 정확하게 결정할 수 있도록 함으로써, 주요 우울 장애 치료 효과를 조기에 예측하기 위한 예후 진단 마커의 사용이 가능하다.

[0116] **실시예**

[0117] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 예시하기 위한 것으로서, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지는 않는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.

[0119] **실시예 1. 임상 시험 피험자 선정**

[0120] MADRS 점수는 약물 투여에 대한 반응을 결정하는 기준으로 간주되므로, 2013년 2월부터 2016년 2월까지 서울대학교병원에서 수행하는 에스시탈로프람 용량 증량의 효능 및 안전성을 테스트하는 임상 시험에 참여한 10명의 주요 우울 장애 환자로부터 혈장 샘플을 수집하였다. 모든 참가자는 한국인이었다.

[0121] 이 시험에는 2단계가 포함되어 있는데, 표준 용량(10-20mg/일)의 에스시탈로프람으로 4주 동안 공개 치료한 후 20mg/일 또는 30mg/일 에스시탈로프람으로 6주간 무작위 이중 맹검 치료(double-blinded treatment)를 진행하였다. 정신 장애 진단 및 통계 매뉴얼(Diagnostic and Statistical Manual of Mental Disorders, 4th edition)에 정의된 대로 주요 우울 장애의 1차 진단을 받은 18-65세의 환자가 포함되었다. 모든 환자는 초기 스크리닝 및 기준일 방문에서 총 MADRS 점수가 18점 이상이었다. 한편, 에스시탈로프람에 과민증을 경험한 경우, 항정신병제, 기분 안정제 또는 선택적 모노아민 산화효소 억제제와 같은 향정신성 약물을 투여받은 적이 있는 경우, 우울증 증상이 있고 2가지 이상의 항우울제 치료에 내성이 있는 것으로 간주되는 경우, 주요 우울 장애 이외의 정신 질환이 있거나 조증 또는 경조증 삽화, 정신 분열증, 분열 정동 장애 또는 약물 남용 장애와 같은 정신 장애의 이전 병력이 조사자에 의해 평가되거나 MADRS의 항목 10에서 5점 이상의 점수로 자살의 상당한 위험이 있는 경우, 또는 신경학적 장애 또는 의학적으로 불안정한 상태(예: 신장 또는 간 장애, 또는 심혈관, 폐 또는 위장 장애)의 병력이 있는 경우에는 피험자에서 제외시켰다.

[0122] 임상시험에 참가한 환자 중, 5명의 반응자(남성 1명 및 여성 4명) 및 5명의 무반응자(남성 1명 및 여성 4명)가 선택되었으며, 기준일(baseline), 1주차, 4주차(무작위화) 및 10주차(무작위화 후 6주)의 4개의 시점에서 각각의 혈장 샘플을 수득하였다.

[0123] 연구 프로토콜은 서울대학교병원 기관심의위원회에 의해 승인되었다 (승인번호: 1008-116-329, 2010년 12월 2일 승인). 본 연구는 헬싱키 선언문 및 국제 조화 우수 임상 관행 지침에 명시된 윤리 원칙에 따라 수행되었으며, 모든 환자는 서면 동의서를 제공하였고, 원하는 경우 언제든지 연구를 중단할 수 있도록 하였다.

[0125] 10명의 피험자, 즉 5명의 반응자 및 5명의 무반응자의 기준일 특징(baseline characteristics)은 표 1과 같다.

*표 1*

[0127] 피험자의 인구통계학적 및 임상적 변수

| 변수 | 반응자 (N=5) | 무반응자 (N=5) | *p*-Value |
|---|---|---|---|
| 나이 (표준편차) | 44.2 (14.2) | 42.8 (16.4) | 0.841 |
| 남성 (%) | 1 (20) | 1 (20) | 1.000 |
| 발병 시기 (표준편차) | 41.8 (11.9) | 33.4 (9.8) | 0.093 |
| 체질량 지수($kg/m^2$) (표준편차) | 23.1 (3.7) | 24.7 (4.6) | 0.309 |
| 기준선(baseline)에서 임상 특성 | | | |
| MADRS (표준편차) | 31.0 (4.6) | 28.8 (2.5) | 0.599 |
| CGI-S (표준편차) | 5.0 (0.7) | 4.2 (1.3) | 0.344 |
| BDI (표준편차) | 32.6 (7.3) | 26.8 (3.6) | 0.206 |
| HRM-D (표준편차 | 21.6 (3.4) | 21.0 (2.9) | 1.000 |
| CUDOS (표준편차) | 38.4 (12.8) | 40.4 (3.0) | 0.917 |
| 세계보건기구 삶의 질 축약 버전 | | | |
| 육체적 삶의 질 (표준편차) | 8.8 (1.7) | 8.5 (0.9) | 0.831 |
| 심리적 삶의 질 (표준편차) | 8.0 (0.8) | 8.5 (0.9) | 0.827 |
| 사회적 삶의 질 (표준편차) | 10.1 (1.5) | 11.7 (2.4) | 0.193 |
| 환경적 삶의 질 (표준편차) | 10.1 (1.7) | 10.1 (1.1) | 0.914 |

[0128] *p*-value는 *Mann-Whitney U test* 또는 *Fisher's exact test*를 사용하여 적절히 계산됨.

[0130] 표 1에서 볼 수 있듯이, 피험자의 평균(표준편차) 연령은 각각 반응자(44.2(14.2)세)와 무반응자(42.8(16.4)세)로 유사하였고, Montgomery and Asberg Depression Rating Scale (MADRS), Clinical Global Impression-Severity (CGI-S), Beck's Depression Inventory (BDI), Hamilton Rating Scale for Depression (HAM-D), Clinically Useful Depression Outcome Scale (CUDOS)에 대한 점수 및 신체적, 심리적, 사회적, 환경적 삶의 질을 포함한 세계보건기구 삶의 질 축약 버전 점수를 포함하여 정서적 증상 및 질병 중증도에서도 두 그룹 간 유의한 차이가 없었다.

[0132] 검증 실험을 위하여, 이후 19명의 반응자 및 5명의 무반응자의 24명의 환자가 추가로 선택되었으며, 이들로부터 혈장 샘플이 기준일 및 1주차에 각각 획득되어 추가 검증 실험이 진행되었다. 이들 환자 중 누구도 비스테로이드성 항염증제 또는 스테로이드와 같은 관련 인자의 혈중 농도를 변경할 수 있는 약물을 복용하지 않았고, 심혈관 질환, 폐 질환, 고혈압, 내분비 이상, 류마티스 질환 또는 뇌혈관 질환과 같은 급성 또는 만성 질환을 가지고 있지 않았다.

[0134] **실시예 2. 혈장 샘플 준비 및 LC-MS/MS 분석**

[0136] 2-1. 혈액 수집 및 혈장 준비

[0137] 혈장은 Human Proteome Organization Plasma Proteome Project에서 제안한 대로 준비되었다. 혈액 샘플 (3mL)을 기준일, 1주차, 4주차 (무작위화) 및 10주차 (무작위화 후 6주차)에 에틸렌디아민테트라아세트산 함유 튜브에 수집하였고, 혈액 샘플은 밤새 금식한 후 (최소 12시간) 오전 9시 30분부터 11시 30분 사이 피험자로부터 채취하였다. 혈액 샘플은 샘플 수집 직후 실온 (RT)에서 15분 동안 2000×g로 원심분리하고, 분리된 혈장은 0.5mL 튜브로 옮긴 후, 원심분리 후 20분 이내에 동결시켰다. 이후 샘플을 얼음 위에 놓고 실험실로 운반한 후, 분석할 때까지 -80 ℃에서 즉시 동결시켜 보관하였다.

[0139] 2-2. 혈장 샘플 준비

[0140] 혈장 샘플에서 단백질체 분석을 방해할 수 있는 다량의 혈장 단백질을 제거하고자 하였다. 혈장 내 존재하는 14가지 가장 풍부한 혈장 단백질(알부민, IgA, IgG, IgM, α1-항트립신, α1-산 당단백질, 아포지단백질 A1, 아포지단백질 A2, 보체 C3, 트랜스페린, α2-마크로글로불린, 트랜스티레틴, 합토글로빈 및 피브리노겐)을 제거하기 위하여, Buffer A로 4배 희석된 혈장의 40 μL 분취량을 바이너리 HPLC 시스템 (20A Prominence, Shimadzu, Tokyo, Japan)의 MARS14 고갈 컬럼 (Agilent Technology, Palo Alto, CA, USA)에 주입하였다. 결합되지 않은 분획을 50mM Tris (pH 8) 중 8M 요소를 포함하는 완충액으로 교환하고, Vivaspin 500 3kDa 컷오프 필터 (Sartorius, Goettingen, Germany)를 사용하여 한외여과 한 후 약 50μL로 농축하고, 새로운 필터 장치 (Nanosep, 30kDa; Pall Corporation, NY, USA)로 옮겼주었다. 50mM Tris (pH 8.5) 중 8M 우레아를 포함하는 용액에 200μL 분취량을 첨가하고 혼합물을 14,000 x g에서 15분 동안 원심분리하는 절차를 두 번 반복하였다. 수집 튜브에서 통과액 (flow-through)을 제거하고 0.05M 요오도아세트아미드 용액 100μL를 첨가하여 제제를 열혼합기에서 600rpm으로 1분 동안 혼합한 후, 20분 동안 혼합 과정없이 인큐베이션하였다. 상기 필터 유닛을 14,000 x g에서 10분 동안 원심분리하고, 50mM Tris(pH 8.5)에 용해된 8M 요소 100㎕를 첨가하고, 필터 유닛을 다시 14,000×g에서 15분 동안 원심분리하고, 이 단계를 2회 반복하였다. 0.05M 중탄산암모늄의 100μL 분취량을 필터 유닛에 첨가하고, 상기 유닛을 10분 동안 14,000 × g에서 원심분리하고, 이 단계도 2회 반복하였다. 2.5㎍ Lys-C/트립신을 함유하는 0.05M 중탄산암모늄의 40㎕ 분취량을 첨가하고, 제제를 열-혼합기에서 600rpm으로 1분 동안 혼합하였다. 상기 유닛을 37 ℃의 습식 챔버에서 12시간 동안 인큐베이션하고 새로운 수집 튜브로 옮겨주었다. 상기 필터 유닛은 14,000 x g에서 10분 동안 원심분리되었으며, 40 μL의 0.5 M NaCl을 첨가하고, 필터 유닛을 다시 14,000 x g에서 10분 동안 원심분리하였다. 포름산을 0.3%의 최종 농도로 첨가하여 분해 (digestion) 반응을 중단시켰다. 상기 펩타이드 혼합물을 Sep Pak C-18 카트리지 (Waters, Milford, MA, USA)로 탈염하고, 콜드 트랩 (CentriVap Cold Traps, Labconco, Kansas City, MO, USA)으로 동결건조한 후, 사용할 때까지 -80 ℃에서 보관하였다.

[0142] 2-3. 나노-LC-ESI-MS/MS 분석

[0143] Dionex UltiMate 3000 RSLCnano 시스템(Thermo Fisher Scientific, Waltham, MA, USA)을 사용하여 펩타이드를 분리하였다. 비드 컬럼의 트립신 펩타이드는 0.1% 포름산에서 재구성되고 200분(250nL/min)에 걸쳐 2μm C18 수지(Thermo Fisher Scientific)로 패킹된 75μm 내부 직경의 50cm Easy-Spray 컬럼에서 분리되었다. 상기 컬럼은 50 ℃에서 0.1% 포름산 및 5% DMSO에서 0-45% 아세토니트릴 구배를 사용하여 전개(develop)되었다. LC를 나노 ESI 소스가 있는 Q Exactive 질량 분석기에 연결하고, 전체 스캔과 20개의 데이터 종속 MS/MS 스캔 사이의 자동 전환을 사용하여 데이터 종속 모드에서 질량 스펙트럼을 획득하였다. 전체 스캔 MS 스펙트럼의 목표 값은 3,000,000이었고 최대 주입 시간은 120ms이고 분해능은 m/z 400에서 70,000이었다. 반복되는 펩타이드는 20초 동안 동적으로 배제되었다. 모든 MS 데이터는 Project PXD017211로 PRIDE 아카이브 (www.ebi.ac.uk/pride/archive/projects/PXD017211)에 기탁하였다.

[0145] 2-4. 데이터베이스 검색 및 비표지 정량화

[0146] 실시예 2-3.에 따른 MS/MS 스펙트럼 분석 결과를 SwissProt 인간 데이터베이스(2017년 5월) Proteome Discoverer(버전 2.2, Thermo Fisher Scientific)에서 SequestHT를 사용하여 검색하였다. 검색 매개변수는 고정 변형으로 시스테인 카르바미도메틸화를 포함하고 두 개의 오류절단(miscleavage)이 있는 가변 변형으로 N-말단 아세틸화 및 메티오닌 산화를 포함하는 기본값으로 설정되었다. 펩타이드는 최대 10ppm의 전구체 이온의 초기 질량 편차(initial mass deviation)와 허용된 단편 질량 편차(allowed fragment mass deviation)를 20ppm으로 설정한 검색을 기반으로 식별되었다. 단백질을 펩타이드에 할당할 때 고유한 펩타이드(unique peptide)와 레이저 펩티드(razor peptide)를 모두 사용하였다. 각 단백질의 고유한 펩타이드에 대한 피크 강도를 사용하여 비표지 정량(Label-free quantitation, LFQ)을 수행하였다.

[0148] 2-5. LC-MS/MS 프로파일링 결과

[0149] 10명의 대상자 각각으로부터 4개의 혈장 샘플을 얻어, 총 40개의 샘플에 대해 LC-MS/MS로 단백질을 프로파일링한 결과, 총 1159개의 단백질이 확인되었고, 이 중 절반 이상인 684개의 단백질에 대해 정량 분석이 진행되었으며, 104회의 LC-MS/MS 분석을 통해, 총 206개의 단백질이 완전히 정량화되었다. 상기 분석은 duplicate 또는

triplicate로 진행하였다.

[0150] LFQ (label-free quantification) 방법으로 혈장 단백질 풍부도 (plasma protein abundance)를 비교하기 전 6개의 상대적으로 안정적이고 풍부한 내인성 단백질(BTD, C8B, C1S, ITIH2, IGFALS 및 SERPINA3)을 이용하여 다른 단백질의 풍부도 데이터를 정규화하였다. 즉, 모든 실험에서 단백질 풍부도를 정규화 한 후 샘플 간 차이를 수정하였다. 정규화된 풍부도는 Pearson의 상관 계수가 0.677로, 혈장 단백질체 데이터베이스의 혈장 단백질 농도(ng/mL)와 통계적으로 유의한 상관 관계를 보여주었다 (조정된 $p$-value < 0.001).

[0151] 기술적 변이(technical variation)가 매우 작다고 판단되는 346개의 단백질중 혈장 고갈 표적 단백질(plasma depletion target protein) 14개, 면역글로불린 관련 단백질(immunoglobulin-related protein) 10개, 정상화 인자(normalization factor) 6개를 제외시키고 총 316개의 단백질에 대해 추가 분석을 진행하였다.

[0153] 2-6. 배치(batch) 간 보정, 누락 데이터 대치 및 정규화 방법

[0154] 샘플 준비 날짜를 기준으로 S15 (무반응자) 및 S29 (반응자)로 구성된 배치 1; S54(비응답자) 및 S52 (반응자)의 배치 2; 및 S6(무반응자), S11(무반응자), S32(반응자), S34(무반응자), S38(반응자) 및 S46(반응자)의 배치 3의 3개의 배치를 준비하였다. 배치 간 분석 오차 보정을 위하여, 104개 LC-MS/MS 분석의 원시 데이터(raw data)에 단백질 당 평균 중심 보정(mean-centering correction per protein)을 적용하였다.

[0155] 이후 누락 데이터 대치(missing data imputation)를 수행하였다. 각 개별 샘플에서 한 번에 측정된 316개의 정량된 단백질 중 180개가 완전히 정량화되었고, 나머지 136개의 단백질에 대한 누락 데이터는 국소 최소 제곱 대치법(local least-squares imputation method)(Kim, H. et al., Bioinformatics 2005, 21, 187-198 참조)에 의해 결정되었다. 이 방법을 사용하여 완전히 정량화된 180개의 단백질을 Pearson의 상관 분석에 의해 15개의 그룹으로 클러스터링하고 15개의 선택된 클러스터의 선형 최적 조합으로 누락 값을 추정하였다.

[0156] 이들 데이터는 스파이크 인 표준(spike-in standard) 없이 내인성 정규화 단백질에 대해 정규화되었다. 전체 데이터에서 다음 기준에 따라 210개 단백질 중 6개 단백질이 최종적으로 LFQ 정규화에 적합한 것으로 선택되었다: (1) 혈장 농도가 NormFinder 안정성 값 (NormFinder stability value)에 의해 결정된 바와 같이 모든 샘플에서 거의 일정하게 유지됨; (2) 혈장 농도가 LMM 분석에 의해 나타난 바와 같이 5명의 반응자와 5명의 무반응자에서 유의한 차이가 없음 ($p$-value > 0.05); (3) 우울증에 대한 보고가 없음. 각 샘플에서 6개의 선택된 정규화 단백질 BTD, C8B, C1S, ITIH2, IGFALS 및 SERPINA3의 원시 풍부도(raw abundance)를 모든 샘플에서 6개의 원시 풍부도의 기하 평균(geometric mean)으로 나누었다. 샘플에서 이러한 6가지 비율의 중간값(median)이 해당 샘플에 대한 정규화 크기 조정 인자 (normalization scaling factor, NSF)로 정의되었다. 샘플 s에 대한 NSF는 다음 방정식을 사용하여 계산할 수 있다:

$$NSF_s = geomean\left(\frac{N_{1,s}}{\hat{N}_1}, \frac{N_{2,s}}{\hat{N}_2}, \cdots, \frac{N_{6,s}}{\hat{N}_6}\right)$$

[0157]

[0158] 여기서 $N_{i,s}$는 샘플 s에서 정규화 단백질 i의 원시 단백질 풍부도, $\hat{N}_i$는 모든 샘플에서 단백질 i의 풍부도 중간값이다.

[0159] 샘플에서 각 바이오마커 후보의 강도의 정규화된 풍부도(normalized abundance of the intensity of each biomarker candidate)는 원시 피크 강도(raw peak intensity)를 NSF로 나누어 계산하였다:

$$\check{PA}_{j,s} = \frac{PA_{j.s}}{NSF_s}$$

[0160]

[0161] 여기서, $\check{PA}_{j,s}$는 샘플 s에서 j번째 바이오마커 후보의 정규화된 풍부도이고, $PA_{j,s}$는 그 상응하는 단백질의

원시 풍부도이다.

[0163] **실시예 3. 반응자와 무반응자에서 시간에 따른 혈장 단백질 변화 분석**

[0164] 보정 없는 Mann-Whitney 테스트를 통하여, 기준일과 1주, 4주, 10주간의 치료 후에서의 혈장 단백질 풍부도를 통계적으로 비교해 본 결과, 1주, 4주, 10주차에 각각 7개, 4개 및 6개의 단백질이 반응자에서 상향 조절되었고, 각각 16개, 17개 및 19개의 단백질이 무반응자에서 상향 조절되었다 ($p$-value < 0.05; 도 1A).

[0165] $T_1$, $T_4$ 및 $T_{10}$의 세 시점에서의 벤 다이어그램은 도 1B와 같다. 무반응자에서 상향조절된 단백질은 유전자 온톨로지(GO)에서 자극에 대한 반응, 상처에 대한 반응 및 튜브 형태 형성과 관련이 있었다 (도 1C). 이는 무반응자에서 약물이 뇌의 신경 회로와 함께 많은 수의 염증 경로를 활성화 시킨다는 것을 의미할 수 있다.

[0166] 유전자 온톨로지로 구분할 때, 세포 외 구조 구성, 보체 활성화 조절 및 트리글리세리드가 풍부한 지질단백질 입자 리모델링과 관련된 단백질은 두 그룹 모두에서 풍부하게 발현되었다.

[0167] 선형 혼합 모델은 종단 단백질체 데이터를 기반으로 차등적으로 풍부한 혈장 단백질을 식별하는 데 적합한데, 반응/시간 상호작용 항(response/time interaction term)이 선택적 세로토닌 재흡수 억제제에 대한 시간 의존적 반응성의 그룹 간 차이를 측정하는 데 중요하다. 선형 혼합 모델 다중 비교 분석을 통해 SGoF 방법(Carvajal-Rodriguez, A. et al., *BMC Bioinform.* 2009, 10, 209. 참조)으로 보정된 37개의 유의한 단백질을 확인하였다 (조정된 $p$-value < 0.05, 반응/시간 상호작용 항). 시간 경과에 따른 이러한 단백질의 그룹 간 차이는 표 2에서 가장 낮은 조정된 $p$-value로 표시된다.

*표 2*

[0169] 반응자와 무반응자 사이에 상이하게 존재하는 반응/시간 상호작용에 상응하는 37개 단백질

| UNIPROT 접근번호 | 조정된 $p$-Value | 유전자 명 | 단백질 명 | 클러스터 번호 | COR |
|---|---|---|---|---|---|
| P04278 | $2.70\times10^{-3}$ | SHBG | Sex hormone-binding globulin | 5 | 0.12 |
| P05090 | $2.95\times10^{-3}$ | APOD | Apolipoprotein D | 6 | -0.34 |
| Q06033 | $4.01\times10^{-3}$ | ITIH3 | Inter-alpha-trypsin inhibitor heavy chain H3 | 5 | -0.21 |
| P08567 | $4.36\times10^{-3}$ | PLEK | Pleckstrin | 4 | 0.04 |
| P04275 | $4.69\times10^{-3}$ | VWF | von Willebrand factor | 6 | -0.19 |
| P52566 | $5.11\times10^{-3}$ | ARHGDIB | Rho GDP-dissociation inhibitor 2 | 1 | -0.15 |
| P02656 | $6.89\times10^{-3}$ | APOC3 | Apolipoprotein C-III | 6 | -0.06 |
| P06276 | $7.13\times10^{-3}$ | BCHE | Cholinesterase | 6 | -0.11 |
| P27169 | $8.89\times10^{-3}$ | PON1 | Serum paraoxonase/arylesterase 1 | 5 | -0.08 |
| P22105-4 | $9.85\times10^{-3}$ | TNXB | Tenascin-X | 6 | -0.25 |
| P02774-3 | $1.09\times10^{-2}$ | GC | Vitamin D-binding protein | 5 | 0.04 |
| P0C0L5 | $1.10\times10^{-2}$ | C4B | Complement C4-B | 5 | -0.21 |
| P02649 | $1.15\times10^{-2}$ | APOE | Apolipoprotein E | 6 | -0.04 |
| P07339 | $1.45\times10^{-2}$ | CTSD | Cathepsin D | 4 | 0.02 |
| Q92820 | $1.50\times10^{-2}$ | GGH | Gamma-glutamyl hydrolase | 5 | -0.01 |
| P09172 | $1.69\times10^{-2}$ | DBH | Dopamine beta-hydroxylase | 2 | 0.03 |
| P40189 | $1.75\times10^{-2}$ | IL6ST | Interleukin-6 receptor subunit beta | 4 | 0.15 |

| | | | | | |
|---|---|---|---|---|---|
| Q8NBP7 | $1.81\times10^{-2}$ | PCSK9 | Proprotein convertase subtilisin/kexin type 9 | 3 | -0.14 |
| Q16610 | $1.83\times10^{-2}$ | ECM1 | Extracellular matrix protein 1 | 2 | 0.02 |
| P62258 | $1.83\times10^{-2}$ | YWHAE | 14-3-3 protein epsilon | 6 | 0.18 |
| P80188 | $1.83\times10^{-2}$ | LCN2 | Neutrophil gelatinase-associated lipocalin | 6 | -0.11 |
| Q9H299 | $1.99\times10^{-2}$ | SH3BGRL3 | SH3 domain-binding glutamic acid-rich-like protein 3 | 1 | 0.13 |
| P27918 | $2.05\times10^{-2}$ | CFP | Properdin | 6 | 0.08 |
| P08571 | $2.12\times10^{-2}$ | CD14 | Monocyte differentiation antigen CD14 | 2 | 0.24 |
| P08697 | $2.33\times10^{-2}$ | SERPINF2 | Alpha-2-antiplasmin | 5 | 0.22 |
| P36980 | $2.34\times10^{-2}$ | CFHR2 | Complement factor H-related protein 2 | 4 | 0.16 |
| P08253 | $2.57\times10^{-2}$ | MMP2 | 72 kDa type IV collagenase | 6 | 0.13 |
| P13671 | $2.65\times10^{-2}$ | C6 | Complement component C6 | 5 | 0.13 |
| O43852-3 | $2.80\times10^{-2}$ | CALU | Calumenin | 3 | 0.10 |
| P14543 | $2.93\times10^{-2}$ | NID1 | Nidogen-1 | 4 | -0.05 |
| P35579 | $2.95\times10^{-2}$ | MYH9 | Myosin-9 | 1 | 0.05 |
| P05160 | $3.70\times10^{-2}$ | F13B | Coagulation factor XIII B chain | 6 | -0.17 |
| P02765 | $3.75\times10^{-2}$ | AHSG | Alpha-2-HS-glycoprotein | 2 | 0.20 |
| Q99453 | $3.85\times10^{-2}$ | PHOX2B | Paired mesoderm homeobox protein 2B | 2 | -0.12 |
| O43157 | $4.01\times10^{-2}$ | PLXNB1 | Plexin-B1 | 5 | -0.01 |
| P06396 | $4.26\times10^{-2}$ | GSN | Gelsolin | 6 | 0.04 |
| O43866 | $4.41\times10^{-2}$ | CD5L | CD5 antigen-like | 3 | -0.27 |

[0171] 풍부도 패턴으로부터 유사한 패턴으로 단백질을 클러스터링하기 위하여 세 가지 다른 시간($T_1$, $T_4$ 및 $T_{10}$)에서 단백질 풍부도를 기준일(T0)에서의 단백질 풍부도에서 빼준 후, t-SNE 및 친화성 전파(affinity propagation)를 진행하여(도 2A), 고유한 패턴의 6개 클러스터를 확보하였다(도 2B).

[0172] 클러스터 1 및 4에서 각각 3개 및 5개의 단백질이 각각 반응자에서 시간이 지남에 따라 감소하고 무반응자에서 시간이 지남에 따라 증가하였다. 5개의 단백질을 포함하는 클러스터 2는 반응자에서 시간이 지남에 따라 거의 변화를 나타내지 않았지만 무반응자에서 시간이 지남에 따라 감소하였다. 클러스터 3에서 3개의 단백질은 반응자에서 1주차부터 증가하였으나 4주차 이후에는 균일하게 나타났다. 이와 대조적으로, 무반응자에서 이들 단백질들은 4주차에 급격히 감소한 다음 10주차에 균일하게 나타났다. 클러스터 5에서 10개의 단백질은 반응자에서 시간이 지남에 따라 거의 변화하지 않았지만 무반응자에서는 시간이 지남에 따라 증가하였다. 클러스터 6에서 11개의 단백질은 반응자에서 4주차에 감소하고 10주차에 증가하였으나, 무반응자에서는 시간이 지남에 따라 거의 변화하지 않았다.

[0173] 이러한 단백질이 항우울제 반응 및 정신 장애와 관련이 있는지 여부를 평가하기 위하여 PsyGeNet에서 37개의 단백질을 검색한 결과 (도 2C), 이들 단백질 중 14개인 APOD, APOE, BCHE, DBH, GGH, GSN, ITIH3, LCN2, MMP2, PHOX2B, PON1, TNXB, VWF, YWHAE가 조현병, 양극성 장애, 코카인 사용 장애, 물질 유발 정신병, 알코올 사용 장애, 우울증 등과 같은 정신과적 증상과 관련이 있는 것으로 확인되었다.

[0174] 또한, Enrichr의 DrugMatrix 범주에서 선택적 세로토닌 재흡수 억제제 (SSRI)에 대한 쥐 조직 및 세포에서의 반응을 검색하여, 이 37개 단백질이 에스시탈로프람의 유사체인 시탈로프람과 관련이 있는지 여부를 평가해 본 결

과, ITIH3, PON1, MMP2, MYH9, APOE, GC, CD14, LCN2 및 CTSD의 9개 단백질이 선택적 세로토닌 재흡수 억제제 처리 및 대조군인 옥수수유 처리한 쥐의 간에서 유의하게 상이한 발현 양상을 나타낸다는 것을 확인하였고, ITIH3, PON1, LCN2, APOE, GC, CLU 및 CTSD의 7가지 단백질이 선택적 세로토닌 재흡수 억제제 및 옥수수유 처리된 쥐의 간세포에서 유의하게 상이한 발현을 나타내었으며, PLXNB1, MMP2 및 CTSD의 3가지 단백질이 선택적 세로토닌 재흡수 억제제 및 옥수수유 처리된 쥐의 심장에서 유의하게 상이한 발현 양상을 나타내었다 (도 2D).

[0175] 상기와 같은 결과로부터, 이들 단백질들은 혈액 뿐 아니라 간과 심장의 장기에서도 약물에 반응하여 양적 변화를 일으킨다는 것을 확인할 수 있었다.

[0177] **실시예 4. 조기 반응을 예측하는 후보 혈장 단백질의 LC-MRM/MS 검증**

[0178] 기준일, 초기 치료 단계(기준선에서 1주까지; Mann-Whitney U 테스트: $p$-value < 0.05) 및 LMM의 반응/시간 상호작용에서 반응자와 무반응자 그룹 간에서 유의한 발현 양상 차이를 보이는 10개 단백질을 선택하고 (도 3A), 연속 동위원소 희석-MRM/MS (serial isotope dilute-MRM/MS)로 검증하였다. 이 중, 신뢰도가 낮은 CFHR2는 약물 반응 조기 예측용 바이오마커에서 제외시켰다.

[0179] 액체 크로마토그래피(LC)는 50 ℃의 온도에서 역상 초고성능 LC (UHPLC) 컬럼 (Agilent ZORBAX Eclipse Plus C18 Column, 95 A, 2.1 mm i.d. × 100 mm, packed with 1.8 μm C18 resin)이 있는 Agilent 1290 Infinity UHPLC 시스템 상에서 수행되었다. 이동상은 증류수 중 0.1% 포름산 (용매 A) 및 아세토니트릴 중 0.1% 포름산 (용매 B)이었다. 컬럼은 0.3 mL/min의 유속으로 5분 동안 0-2% 용매 B, 5분 동안 2-3% 용매 B, 10분 동안 3-50% 용매 B, 4분 동안 50-50% 용매 B, 1분 동안 50-0% 용매 B, 9분 동안 0-0% 용매 B로 전개 (develop)되었다. 주입된 샘플은 분해된 혈장 펩타이드(초기 혈장 부피: 40 μL)와 동위원소 표지된 내부 표준 펩타이드의 혼합물로 구성되었다. UHPLC 시스템은 양이온 모드에서 작동되는 표준 흐름 Jet Stream 전자분무 소스에 의해 삼중 사중극자 질량 분석기(Agilent 6495 QQQ)에 연결되었다. 추가 매개변수는 모세관 전압, 3.5kV; 노즐 전압, 1kV; 가스 온도, 290 ℃; 건조 가스 유량, 350 ℃에서 11L/min; 분무기 가스 압력, 350 ℃에서 40 PSI; 및 Q1 및 Q3에 대한 단위 해상도를 포함한다. MRM 전환이 선택되었고 충돌 에너지는 Skyline (64비트, 버전 19.1.0.193) 소프트웨어에 의해 최적화되었다. 세포 가속기 전압은 5V로 설정되었다. 정량화 실험은 동적 MRM (델타 유지 시간: 3분)을 사용하여 수행되었으며 총 주기 시간은 약 1.5초였다. 질량 분석기는 MRM/MS 데이터(*.d)를 생성하는 MassHunter 소프트웨어(버전 B.08.00, Agilent)로 작동되었다. 추출된 이온 크로마토그램의 MRM 결과를 Skyline으로 분석하고, 그 상응하는 안정 동위원소 표지 펩타이드(SpikeTides™; JPT Peptdie Technologies Berlin, Germany)와 비교하여 정량화하였다.

[0180] 역표준 검량선(reverse standard calibration curves)를 기반으로 9개의 단백질을 나타내는 15개의 대리 펩타이드(surrogate peptide)를 단백질 정량을 위해 선택하고, LC-MRM/MS 결과를 기반으로 각 단백질에 대해 강한 신호를 가지는 대표적인 펩타이드를 선택하였다.

[0181] 19명의 반응자와 5명의 무반응자를 비교한 검증 MRM 결과 중, 중경량 비율이 정량 한계 미만인 MYH9는 정량화할 수 없었다. 나머지 8개 단백질 중 3개 단백질인 PHOX2B, SH3BGRL3, YWHAE이 반응자(Wilcoxon signed-rank test: FDR-adjusted $p$-value < 0.05)에서 기준일과 1주차에 유의미한 차이를 보였으나 무반응자에서는 이와 같은 차이가 확인되지 않았다 (Wilcoxon signed-rank test: FDR-adjusted $p$-value > 0.05, 도 4B).

[0182] 즉, 약물 투여 1 주일 후, PHOX2B 단백질 발현 수준은 반응자에서 유의하게 증가한 반면 SH3BGRL3 및 YWHAE 단백질 발현 수준은 반응자에서 유의하게 감소하는 것으로 확인되었다. 그러나 무반응자에서는 이들 3개 단백질의 유의미한 변화가 관찰되지 않았다. 한편, 다른 5개 단백질의 발현 수준은 두 그룹에서 크게 차이가 나지 않은 것으로 확인되었다.

[0184] **실시예 5. 통계 분석**

[0185] 데이터는 R(버전 3.6.0)을 포함한 RStudio(버전 1.1.456)를 사용하여 분석되었다. 종단 혈장 단백질 풍부도는 약물 반응(무반응 또는 반응), 샘플링 시간(기준일, 1주, 4주 및 10주), 고정 변수로서의 반응/시간 상호 작용 및 기술적 복제, 고정 변수에 대한 개별 환자 중첩, 및 개인을 랜덤 변수로 사용하여 선형 혼합 모델 분석(lme4 패키지)에 의해 평가되었다. GEE 분석(geesmv 패키지)에서는, 혈장 풍부도를 2~3개의 기술적 복제의 평균값으로 병합한 다음 약물 반응, 샘플링 시간 및 약물/샘플링 시간을 분석하였다. 작업 상관 구조(working correlation

structure)를 독립적으로 설정하고 가우스 추정을 수행하였다. 클러스터링 분석은 4개의 시점에서 각 그룹(반응자 및 무반응자)의 중간 단백질 농도를 기반으로 하였으며, t-확률적 이웃 임베딩(t-stochastic neighbor embedding, t-SNE)(perplexity = 2, theta=0, dims=2) 및 선호도 전파(방법=상관 대칭 행렬 및 Spearman)는 각각 Rtsne 및 apcluster 패키지를 사용하여 계산하였다.

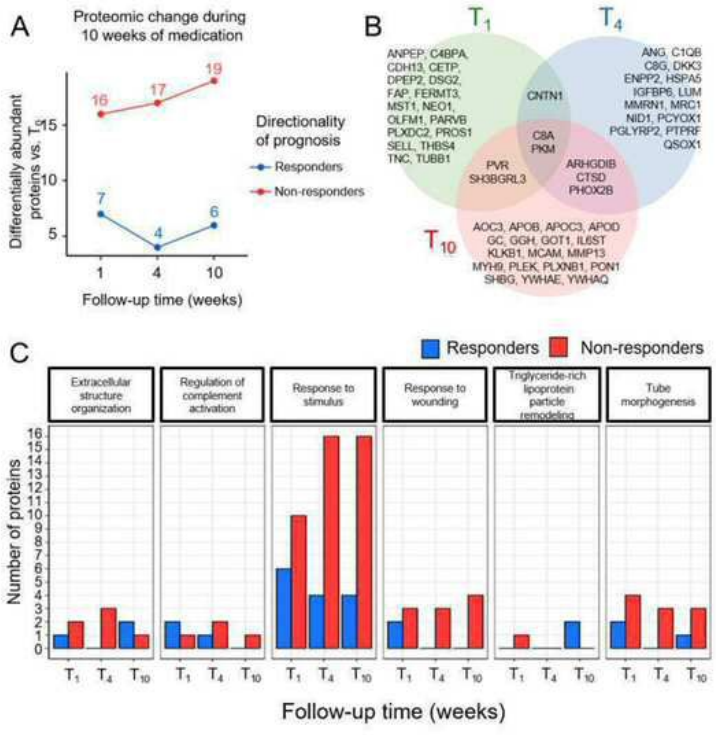
[0186] 다른 소프트웨어 패키지에는 데이터베이스 = "ALL"에서 정신 장애 유전자 연관 네트워크(PsyGeNet)의 산포도(scatter plots)를 위한 ggline과 단백질 매핑을 위한 psygenet2r이 포함되어 있다. 다중 비교를 통해 타입 I 에러를 컨트롤하고자, 선형 혼합 모델 분석에 의한 반응/시간 상호작용의 *p*-value에 대해 SGoF R 패키지에서 기본 옵션(알파 = 0.05, 감마 = 0.05, P0 = 0.5, a0 = 1, b0 = 1)의 베이지안 순차 적합도 메타 테스트(SGoF) 방법을 적용하고, MRM 쌍 분석의 *p*-value에 대한 Benjamini-Hochberg 절차를 적용한 후, 상관 분석을 위해 치환된 *p*-value을 계산하였다.

[0187]

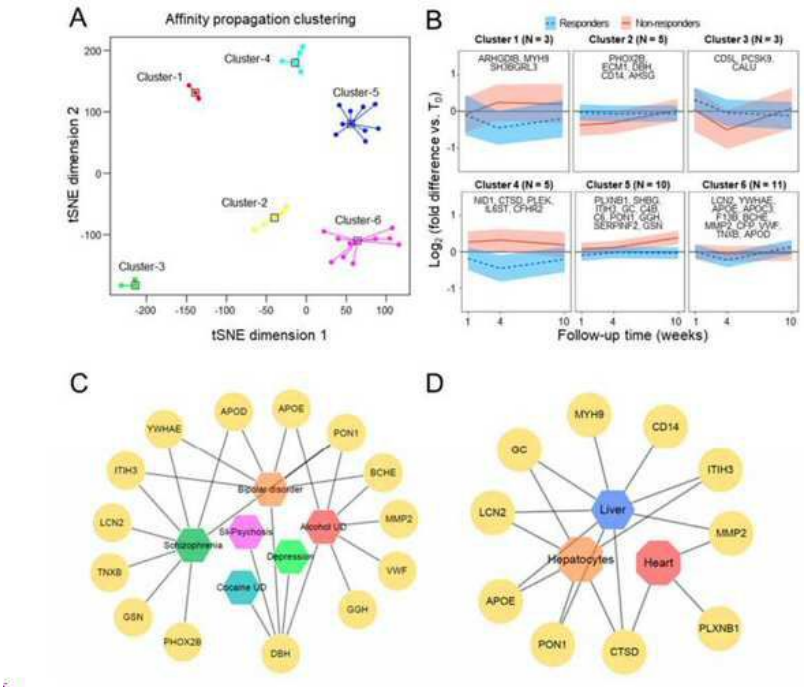
[0188] 이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

## 도면

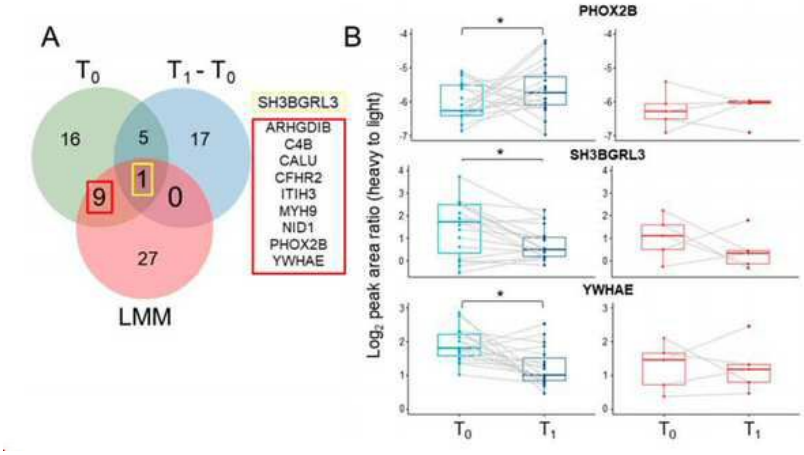
### 도면1



도면2




도면3




## 서 열 목 록

<110> Seoul National University R&DB Foundation

THE ASAN FOUNDATION

UNIVERSITY OF ULSAN FOUNDATION FOR INDUSTRY COOPERATION

<120> Biomarker for early prediction of the treatment effect in major depressive disorder

<130> 1069408

<150> KR 10-2020-0117021

<151> 2020-09-11

<160> 3

<170> KoPatentIn 3.0

<210> 1

<211> 314

<212> PRT

<213> Homo sapiens

<400> 1

Met Tyr Lys Met Glu Tyr Ser Tyr Leu Asn Ser Ser Ala Tyr Glu Ser

1 5 10 15

Cys Met Ala Gly Met Asp Thr Ser Ser Leu Ala Ser Ala Tyr Ala Asp

20 25 30

Phe Ser Ser Cys Ser Gln Ala Ser Gly Phe Gln Tyr Asn Pro Ile Arg

35 40 45

Thr Thr Phe Gly Ala Thr Ser Gly Cys Pro Ser Leu Thr Pro Gly Ser

50 55 60

Cys Ser Leu Gly Thr Leu Arg Asp His Gln Ser Ser Pro Tyr Ala Ala

65 70 75 80

Val Pro Tyr Lys Leu Phe Thr Asp His Gly Gly Leu Asn Glu Lys Arg

85 90 95

Lys Gln Arg Arg Ile Arg Thr Thr Phe Thr Ser Ala Gln Leu Lys Glu

100 105 110

Leu Glu Arg Val Phe Ala Glu Thr His Tyr Pro Asp Ile Tyr Thr Arg

115 120 125

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

130 135 140

Trp Phe Gln Asn Arg Arg Ala Lys Phe Arg Lys Gln Glu Arg Ala Ala

145 150 155 160

Ala Ala Ala Ala Ala Ala Ala Lys Asn Gly Ser Ser Gly Lys Lys Ser

165 170 175

Asp Ser Ser Arg Asp Asp Glu Ser Lys Glu Ala Lys Ser Thr Asp Pro

180 185 190

Asp Ser Thr Gly Gly Pro Gly Pro Asn Pro Asn Pro Thr Pro Ser Cys
195 200 205
Gly Ala Asn Gly Gly Gly Gly Gly Gly Pro Ser Pro Ala Gly Ala Pro
210 215 220

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

290 295 300
Lys Ala Ala Leu Val Lys Ser Ser Met Phe
305 310
<210> 2
<211> 93
<212> PRT
<213> Homo sapiens
<400> 2
Met Ser Gly Leu Arg Val Tyr Ser Thr Ser Val Thr Gly Ser Arg Glu
1 5 10 15
Ile Lys Ser Gln Gln Ser Glu Val Thr Arg Ile Leu Asp Gly Lys Arg
20 25 30
Ile Gln Tyr Gln Leu Val Asp Ile Ser Gln Asp Asn Ala Leu Arg Asp

35 40 45
Glu Met Arg Ala Leu Ala Gly Asn Pro Lys Ala Thr Pro Pro Gln Ile
50 55 60
Val Asn Gly Asp Gln Tyr Cys Gly Asp Tyr Glu Leu Phe Val Glu Ala
65 70 75 80
Val Glu Gln Asn Thr Leu Gln Glu Phe Leu Lys Leu Ala

85 90

<210> 3

<211> 255

<212> PRT

<213> Artificial Sequence

<220><223> Homo sapiens

<400> 3

Met Asp Asp Arg Glu Asp Leu Val Tyr Gln Ala Lys Leu Ala Glu Gln
1 5 10 15

Ala Glu Arg Tyr Asp Glu Met Val Glu Ser Met Lys Lys Val Ala Gly
20 25 30

Met Asp Val Glu Leu Thr Val Glu Glu Arg Asn Leu Leu Ser Val Ala
35 40 45

Tyr Lys Asn Val Ile Gly Ala Arg Arg Ala Ser Trp Arg Ile Ile Ser
50 55 60

Ser Ile Glu Gln Lys Glu Glu Asn Lys Gly Gly Glu Asp Lys Leu Lys

65 70 75 80

Met Ile Arg Glu Tyr Arg Gln Met Val Glu Thr Glu Leu Lys Leu Ile
85 90 95

Cys Cys Asp Ile Leu Asp Val Leu Asp Lys His Leu Ile Pro Ala Ala
100 105 110

Asn Thr Gly Glu Ser Lys Val Phe Tyr Tyr Lys Met Lys Gly Asp Tyr
115 120 125

His Arg Tyr Leu Ala Glu Phe Ala Thr Gly Asn Asp Arg Lys Glu Ala
130 135 140

Ala Glu Asn Ser Leu Val Ala Tyr Lys Ala Ala Ser Asp Ile Ala Met
145 150 155 160

Thr Glu Leu Pro Pro Thr His Pro Ile Arg Leu Gly Leu Ala Leu Asn
165 170 175

Phe Ser Val Phe Tyr Tyr Glu Ile Leu Asn Ser Pro Asp Arg Ala Cys
180 185 190

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

195 200 205

Leu Ser Glu Glu Ser Tyr Lys Asp Ser Thr Leu Ile Met Gln Leu Leu

210 215 220

Arg Asp Asn Leu Thr Leu Trp Thr Ser Asp Met Gln Gly Asp Gly Glu

225 230 235 240

Glu Gln Asn Lys Glu Ala Leu Gln Asp Val Glu Asp Glu Asn Gln

245 250 255