

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2023-0039213
(43) 공개일자 2023년03월21일

(51) 국제특허분류(Int. Cl.)

G16B 20/20 (2019.01) G16B 40/00 (2019.01)

G16H 50/50 (2018.01)

(52) CPC특허분류

G16B 20/20 (2019.02)

C12Q 1/6883 (2022.01)

(21) 출원번호 10-2021-0122235

(22) 출원일자 2021년09월14일

심사청구일자 2021년09월14일

(71) 출원인

한림대학교 산학협력단

강원도 춘천시 한림대학길 1, 한림대학교(옥천동)

(72) 발명자

전진평

경기도 용인시 수지구 동천로 75(동천동)

홍은표

서울특별시 노원구 노원로1길 103, 3동 205호(공릉동, 대명아파트)

(뒷면에 계속)

(74) 대리인

김남혁

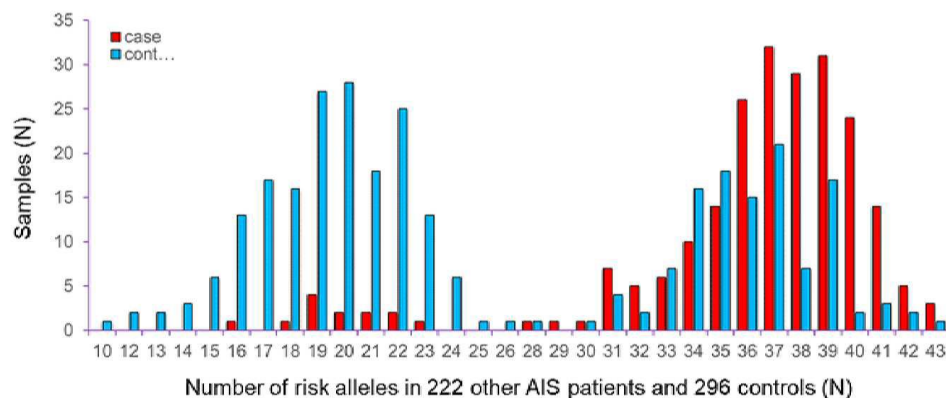
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 뇌졸중 진단용 단일염기다형성 마커 및 이를 이용한 뇌졸중 진단 방법

(57) 요약

본 발명은 뇌졸중 진단용 단일염기다형성 마커 및 이를 이용한 뇌졸중 진단 방법에 관한 것으로, 보다 자세하게는 복수개의 단일염기다형성 유전자 정보를 확인하여 뇌졸중을 진단하는 방법에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

G16B 40/00 (2019.02)

G16H 50/50 (2018.01)

C12Q 2600/156 (2013.01)

(72) 발명자

김봉준

강원도 춘천시 교동길 50

윤동혁

충청북도 단양군 어상천면 어상천로 1589-8

이 발명을 지원한 국가연구개발사업

과제고유번호 1345322048

과제번호 2020R1I1A3070726

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 지역대학우수과학자지원사업

연구과제명 리아 표적 치료 기술 개발
마이토파지 조절기전에 근거한 지주막하 출혈 후 뇌손상 규명 및 나노기반 미토콘드

기 여 율 1/1

과제수행기관명 한림대학교 산학협력단

연구기간 2020.06.01 ~ 2025.05.31

명세서

청구범위

청구항 1

rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 단계; 및

상기 유전자 정보를 하기 수학식 1에 입력하는 단계;를 포함하는, 뇌졸중 예측 또는 진단을 위한 정보 제공 방법;

[수학식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

상기 수학식 1에서,

상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고,

상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고,

상기 w_i 는, 상기 Var_i 에 부여되는 가중치이다.

청구항 2

제1항에 있어서,

상기 유전자 정보는 목적하는 개체로부터 분리된 생물학적 시료에서 확인하는 것인, 방법.

청구항 3

제1항에 있어서,

상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것인, 방법.

청구항 4

제1항에 있어서,

상기 w_i 는, 뇌혈관 질환과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학식 2를 만족하는 것인, 방법:

[수학식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

상기 수학식 2에서,

상기 k 는, 0이 아닌 실수이고,
 상기 OR_i 는, 상기 상관관계의 오즈비(odds ratio)이고,
 상기 P_i 는, 상기 상관관계의 P값이다.

청구항 5

제4항에 있어서,
 상기 k 는 음수이고,
 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것인, 방법.

청구항 6

제5항에 있어서,
 상기 점수는,
 상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs112859779-C인 경우 rs112859779-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs79134766-G인 경우 rs79134766-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs371331393-A인 경우 rs371331393-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs138525217-T인 경우 rs138525217-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs74115822-A인 경우 rs74115822-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs75861150-T인 경우 rs75861150-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs116969723-G인 경우 rs116969723-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs6741819-C인 경우 rs6741819-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs59626274-C인 경우 rs59626274-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs17688188-G인 경우 rs17688188-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs56942085-A인 경우 rs56942085-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs72835045-G인 경우 rs72835045-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs78458145-G인 경우 rs78458145-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs118107419-C인 경우 rs118107419-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs727333-C인 경우 rs727333-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs79461840-C인 경우 rs79461840-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs61775135-C인 경우 rs61775135-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs76855873-C인 경우 rs76855873-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs55800589-G인 경우 rs55800589-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs117398778-C인 경우 rs117398778-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs3737926-C인 경우 rs3737926-T인 경우에 비하여 높은 것;

상기 유전자형이 rs3744644-G인 경우 rs3744644-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs2440154-A인 경우 rs2440154-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs4979583-T인 경우 rs4979583-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs11646803-C인 경우 rs11646803-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs7964241-G인 경우 rs7964241-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs12350582-G인 경우 rs12350582-A인 경우에 비하여 높은 것; 또는
 상기 유전자형이 rs1052270-C인 경우 rs1052270-T인 경우에 비하여 높은 것;이고,
 양수인 것인, 방법.

청구항 7

제1항에 있어서,
 상기 n은 20 이상 40 이하인 것인, 방법.

청구항 8

제1항에 있어서,
 상기 뇌졸중은 허혈성 뇌졸중인 것인, 방법.

청구항 9

rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및

상기 유전자 정보를 하기 수학식 1에 입력하여 뇌졸중 위험도를 산출하는 연산부;를 포함하는, 뇌졸중 예측 또는 진단을 위한 장치:

[수학식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

상기 수학식 1에서,

상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고,

상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고,

상기 w_i 는, 상기 Var_i 에 부여되는 가중치이다.

청구항 10

제9항에 있어서,

목적하는 개체로부터 분리된 생물학적 시료로부터 유전자 서열을 검출하는 검출부를 더 포함하는 것인, 장치.

청구항 11

제9항에 있어서,

상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것인, 장치.

청구항 12

제9항에 있어서,

상기 w_i 는, 뇌혈관 질환과 상기 i 번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학식 2를 만족하는 것인, 장치:

[수학식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

상기 수학식 2에서,

상기 k 는, 0이 아닌 실수이고,

상기 OR_i 는, 상기 상관관계의 오즈비(odds ratio)이고,

상기 P_i 는, 상기 상관관계의 P값이다.

청구항 13

제12항에 있어서,

상기 k 는 음수이고,

상기 연산부는 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것인, 장치.

청구항 14

제9항에 있어서,

상기 점수는,

상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 높은 것;

상기 유전자형이 rs112859779-C인 경우 rs112859779-T인 경우에 비하여 높은 것;

상기 유전자형이 rs79134766-G인 경우 rs79134766-A인 경우에 비하여 높은 것;

상기 유전자형이 rs371331393-A인 경우 rs371331393-G인 경우에 비하여 높은 것;

상기 유전자형이 rs138525217-T인 경우 rs138525217-C인 경우에 비하여 높은 것;

상기 유전자형이 rs74115822-A인 경우 rs74115822-G인 경우에 비하여 높은 것;

상기 유전자형이 rs75861150-T인 경우 rs75861150-C인 경우에 비하여 높은 것;

상기 유전자형이 rs116969723-G인 경우 rs116969723-A인 경우에 비하여 높은 것;

상기 유전자형이 rs6741819-C인 경우 rs6741819-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs59626274-C인 경우 rs59626274-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs17688188-G인 경우 rs17688188-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs56942085-A인 경우 rs56942085-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs72835045-G인 경우 rs72835045-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs78458145-G인 경우 rs78458145-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs118107419-C인 경우 rs118107419-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs727333-C인 경우 rs727333-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs79461840-C인 경우 rs79461840-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs61775135-C인 경우 rs61775135-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs76855873-C인 경우 rs76855873-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs55800589-G인 경우 rs55800589-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs117398778-C인 경우 rs117398778-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs3737926-C인 경우 rs3737926-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs3744644-G인 경우 rs3744644-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs2440154-A인 경우 rs2440154-G인 경우에 비하여 높은 것;
 상기 유전자형이 rs4979583-T인 경우 rs4979583-C인 경우에 비하여 높은 것;
 상기 유전자형이 rs11646803-C인 경우 rs11646803-T인 경우에 비하여 높은 것;
 상기 유전자형이 rs7964241-G인 경우 rs7964241-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs12350582-G인 경우 rs12350582-A인 경우에 비하여 높은 것;
 상기 유전자형이 rs1052270-C인 경우 rs1052270-T인 경우에 비하여 높은 것;이고,
 양수인 것인, 장치.

청구항 15

rs371331393, rs75822236, rs79134766, rs112859779, rs6841581, rs116969723, rs3737926 및 rs3744644rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589 및 rs117398778의 단일염기다형성(SNP) 유전자를 검출 또는 증폭할 수 있는 제제를 포함하는, 뇌졸중 예측 또는 진단용 조성물.

청구항 16

제15항에 있어서,

상기 유전자는 rs138525217, rs118107419, rs78458145, rs727333, rs6741819, rs55800589, rs61775135, rs72835045, rs75861150, rs79461840, rs2440154, rs1052270, rs56942085, rs700651, rs59626274, rs17688188, rs76855873, rs74115822, rs11646803, rs117398778, rs4979583, rs12350582 및 rs7964241rs6841581 및 rs700651에서 선택된 어느 하나 이상을 더 포함하는 것인, 조성물.

청구항 17

제15항 내지 제16항 중 어느 한 항의 조성물을 포함하는, 뇌졸중 예측 또는 진단용 키트.

발명의 설명

기술 분야

[0001] 본 발명은 뇌졸중 진단용 단일염기다형성 마커 및 이를 이용한 뇌졸중 진단 방법에 관한 것으로, 보다 자세하게는 복수개의 단일염기다형성 유전자 정보를 확인하여 뇌졸중을 진단하는 방법에 관한 것이다.

배경 기술

[0002] 뇌졸중(stroke)은 세계적으로 사망원인 중 2위를 차지한다. 뇌졸중은 얼굴근육의 마비, 팔의 무력 및 언어장애 등의 증상을 보인다. 뇌졸중은 대부분 혈전에 의해 뇌의 한 부분으로 가는 혈류가 차단되어 발생하는데, 혈류가 조속히 재개되지 않으면 뇌세포가 사멸한다. 특히 뇌졸중 발생에서 90% 빈도를 차지하는 급성 허혈성 뇌졸중(acute ischemic stroke; AIS)은 혈전에 의해 뇌동맥 혈류가 차단될 때 발생한다. 이 경우 신속한 혈액 희석 치료를 통하여 혈액의 흐름을 개선하여, 뇌에 공급되는 산소와 영양물질을 증가시키고, 결국 사멸할 위기에 처했던 뇌세포를 살릴 필요가 있다. 따라서 상기 치료를 적용할지 여부와 같은 의사결정을 위하여, 본 발명은 상기 급성 허혈성 뇌졸중(AIS)에 정밀 의학 모델을 개발하였고, 이는 개인화된 진단 및 치료를 위하여 제공될 수 있다.

[0003] 한편, 핵산 분자 검출은 인간, 동물, 식물 등이 병원체에 감염되었는지 여부를 검출할 수 있을뿐만 아니라 다양한 질환 위험을 검출할 수 있다. 현재에는 Realtime PCR, FISH 혼성화 기술 등 많은 핵산 검출 방법이 개발되어 있고, 이를 이용한 질병 진단 마커가 개발되어 있으나 단일 마커를 이용하여 암을 진단하는 방법이 주를 이루고 있다. 반면, 다유전자 마커 세트를 이용할 경우에는 단일 마커에 비하여 높은 정확도로 질병의 예후를 진단할 수 있으나, 복잡도가 상승하는 문제로 기술 개발이 더딘 실정이다.

발명의 내용

해결하려는 과제

[0004] 본 발명의 일 목적은 뇌졸중 진단용 단일염기다형성 마커 및 이를 이용한 뇌졸중 진단 방법을 제공하고자 한다.

[0005] 또한, 본 발명이 해결하고자 하는 과제들은 이상에서 언급된 과제로 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0006] 상기와 같은 목적을 달성하기 위하여 본 발명의 일 실시예에 따르면, rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 단계; 및 상기 유전자 정보를 하기 수학식 1에 입력하는 단계;를 포함하는, 뇌졸중 예측 또는 진단을 위한 정보 제공 방법을 제공하는 것일 수 있다:

[0007] [수학식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

[0008]

[0009] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i 는, 상기 Var_i 에 부

여되는 가중치이다.

[0010] 또한 본 발명에서, 상기 유전자 정보는 목적하는 개체로부터 분리된 생물학적 시료에서 확인하는 것일 수 있다.

[0011] 또한 본 발명에서, 상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것일 수 있다.

[0012] 또한 본 발명에서, 상기 w_i 는, 뇌혈관 질환과 상기 i 번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학적 식 2를 만족하는 것일 수 있다:

[0013] [수학적 식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

[0015] 상기 수학적 식 2에서, 상기 k 는, 0이 아닌 실수이고, 상기 OR_i 는, 상기 상관관계의 오즈비(odds ratio)이고, 상기 P_i 는, 상기 상관관계의 P값이다.

[0016] 또한 본 발명에서, 상기 k 는 음수이고, 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것일 수 있다.

[0017] 또한 본 발명에서, 상기 점수는, 상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 높은 것; 상기 유전자형이 rs112859779-C인 경우 rs112859779-T인 경우에 비하여 높은 것; 상기 유전자형이 rs79134766-G인 경우 rs79134766-A인 경우에 비하여 높은 것; 상기 유전자형이 rs371331393-A인 경우 rs371331393-G인 경우에 비하여 높은 것; 상기 유전자형이 rs138525217-T인 경우 rs138525217-C인 경우에 비하여 높은 것; 상기 유전자형이 rs74115822-A인 경우 rs74115822-G인 경우에 비하여 높은 것; 상기 유전자형이 rs75861150-T인 경우 rs75861150-C인 경우에 비하여 높은 것; 상기 유전자형이 rs116969723-G인 경우 rs116969723-A인 경우에 비하여 높은 것; 상기 유전자형이 rs6741819-C인 경우 rs6741819-T인 경우에 비하여 높은 것; 상기 유전자형이 rs59626274-C인 경우 rs59626274-T인 경우에 비하여 높은 것; 상기 유전자형이 rs17688188-G인 경우 rs17688188-A인 경우에 비하여 높은 것; 상기 유전자형이 rs56942085-A인 경우 rs56942085-G인 경우에 비하여 높은 것; 상기 유전자형이 rs72835045-G인 경우 rs72835045-A인 경우에 비하여 높은 것; 상기 유전자형이 rs78458145-G인 경우 rs78458145-A인 경우에 비하여 높은 것; 상기 유전자형이 rs118107419-C인 경우 rs118107419-A인 경우에 비하여 높은 것; 상기 유전자형이 rs727333-C인 경우 rs727333-A인 경우에 비하여 높은 것; 상기 유전자형이 rs79461840-C인 경우 rs79461840-T인 경우에 비하여 높은 것; 상기 유전자형이 rs61775135-C인 경우 rs61775135-A인 경우에 비하여 높은 것; 상기 유전자형이 rs76855873-C인 경우 rs76855873-T인 경우에 비하여 높은 것; 상기 유전자형이 rs55800589-G인 경우 rs55800589-C인 경우에 비하여 높은 것; 상기 유전자형이 rs117398778-C인 경우 rs117398778-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3737926-C인 경우 rs3737926-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3744644-G인 경우 rs3744644-C인 경우에 비하여 높은 것; 상기 유전자형이 rs2440154-A인 경우 rs2440154-G인 경우에 비하여 높은 것; 상기 유전자형이 rs4979583-T인 경우 rs4979583-C인 경우에 비하여 높은 것; 상기 유전자형이 rs11646803-C인 경우 rs11646803-T인 경우에 비하여 높은 것; 상기 유전자형이 rs7964241-G인 경우 rs7964241-A인 경우에 비하여 높은 것; 상기 유전자형이 rs12350582-G인 경우 rs12350582-A인 경우에 비하여 높은 것; 또는 상기 유전자형이 rs1052270-C인 경우 rs1052270-T인 경우에 비하여 높은 것;이고, 양수인 것일 수 있다.

[0018] 또한 본 발명에서, 상기 n 은 20 이상 40 이하인 것일 수 있다.

[0019] 또한 본 발명에서, 상기 뇌졸중은 허혈성 뇌졸중인 것일 수 있다.

[0020] 본 발명의 다른 실시예에 따르면, rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및 상기 유전자 정보를 하기 수학적 식 1에 입력하여 뇌졸중 위험도를 산출하는 연산부;를 포함하는, 뇌졸중 예측 또는 진단을 위한 장치인 것일 수 있다:

[0021] [수학식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

[0022]

[0023] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i 는, 상기 Var_i 에 부여되는 가중치이다.

[0024] 또한 본 발명에서, 목적하는 개체로부터 분리된 생물학적 시료로부터 유전자 서열을 검출하는 검출부를 더 포함하는 것일 수 있다.

[0025] 또한 본 발명에서, 상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것일 수 있다.

[0026] 또한 본 발명에서, 상기 w_i 는, 뇌혈관 질환과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학식 2를 만족하는 것일 수 있다:

[0027] [수학식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

[0028]

[0029] 상기 수학식 2에서, 상기 k는, 0이 아닌 실수이고, 상기 OR_i 는, 상기 상관관계의 오즈비(odds ratio)이고, 상기 P_i 는, 상기 상관관계의 P값이다.

[0030] 또한 본 발명에서, 상기 k는 음수이고, 상기 연산부는 상기 PRS가 대조군에 비해 높은 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것일 수 있다.

[0031] 또한 본 발명에서, 상기 점수는, 상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 높은 것; 상기 유전자형이 rs112859779-C인 경우 rs112859779-T인 경우에 비하여 높은 것; 상기 유전자형이 rs79134766-G인 경우 rs79134766-A인 경우에 비하여 높은 것; 상기 유전자형이 rs371331393-A인 경우 rs371331393-G인 경우에 비하여 높은 것; 상기 유전자형이 rs138525217-T인 경우 rs138525217-C인 경우에 비하여 높은 것; 상기 유전자형이 rs74115822-A인 경우 rs74115822-G인 경우에 비하여 높은 것; 상기 유전자형이 rs75861150-T인 경우 rs75861150-C인 경우에 비하여 높은 것; 상기 유전자형이 rs116969723-G인 경우 rs116969723-A인 경우에 비하여 높은 것; 상기 유전자형이 rs6741819-C인 경우 rs6741819-T인 경우에 비하여 높은 것; 상기 유전자형이 rs59626274-C인 경우 rs59626274-T인 경우에 비하여 높은 것; 상기 유전자형이 rs17688188-G인 경우 rs17688188-A인 경우에 비하여 높은 것; 상기 유전자형이 rs56942085-A인 경우 rs56942085-G인 경우에 비하여 높은 것; 상기 유전자형이 rs72835045-G인 경우 rs72835045-A인 경우에 비하여 높은 것; 상기 유전자형이 rs78458145-G인 경우 rs78458145-A인 경우에 비하여 높은 것; 상기 유전자형이 rs118107419-C인 경우 rs118107419-A인 경우에 비하여 높은 것; 상기 유전자형이 rs727333-C인 경우 rs727333-A인 경우에 비하여 높은 것; 상기 유전자형이 rs79461840-C인 경우 rs79461840-T인 경우에 비하여 높은 것; 상기 유전자형이 rs61775135-C인 경우 rs61775135-A인 경우에 비하여 높은 것; 상기 유전자형이 rs76855873-C인 경우 rs76855873-T인 경우에 비하여 높은 것; 상기 유전자형이 rs55800589-G인 경우 rs55800589-C인 경우에 비하여 높은 것; 상기 유전자형이 rs117398778-C인 경우 rs117398778-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3737926-C인 경우 rs3737926-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3744644-G인 경우 rs3744644-C인 경우에 비하여 높은 것; 상기 유전자형이 rs2440154-A인 경우 rs2440154-G인 경우에 비하여 높은 것; 상기 유전자형이 rs4979583-T인 경우 rs4979583-C인 경우에 비하여 높은 것; 상기 유전자형이 rs11646803-C인 경우 rs11646803-T인 경우에 비하여 높은 것; 상기 유전자형이 rs7964241-G인 경우 rs7964241-A인 경우에 비하여 높은 것; 상기 유전자형이 rs12350582-G인 경우 rs12350582-A인 경우에 비하여 높은 것; 또는 상기 유전자형이 rs1052270-C인 경우 rs1052270-T인 경우에 비하여 높은 것;이고, 양수인 것일 수 있다.

[0032] 본 발명의 또 다른 실시예에서, rs371331393, rs75822236, rs79134766, rs112859779, rs6841581, rs116969723, rs3737926 및 rs3744644rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145,

rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589 및 rs117398778의 단일염기다형성(SNP) 유전자를 검출 또는 증폭할 수 있는 체제를 포함하는, 뇌졸중 예측 또는 진단용 조성물을 제공하는 것일 수 있다.

[0033] 또한 본 발명에서, 상기 유전자는 rs138525217, rs118107419, rs78458145, rs727333, rs6741819, rs55800589, rs61775135, rs72835045, rs75861150, rs79461840, rs2440154, rs1052270, rs56942085, rs700651, rs59626274, rs17688188, rs76855873, rs74115822, rs11646803, rs117398778, rs4979583, rs12350582 및 rs7964241rs6841581 및 rs700651에서 선택된 어느 하나 이상을 더 포함하는 것일 수 있다.

[0034] 본 발명의 또 다른 실시예에서, 본 발명의 조성물을 포함하는, 뇌졸중 예측 또는 진단용 키트를 제공하는 것일 수 있다.

발명의 효과

[0035] 본 발명에 따른 뇌졸중 진단용 단일염기다형성 마커 및 이를 이용한 뇌졸중 진단을 위한 정보 제공 방법, 뇌졸중 진단을 위한 장치, 뇌졸중 진단용 조성물 및 키트는, 뇌졸중 위험도를 예측하여 개인화된 진단 및 치료를 제공하는데 도움이 될 수 있다.

도면의 간단한 설명

[0036] 도 1은 본 발명의 실시예에 따른 222명의 급성 허혈성 뇌졸중(AIS) 환자와 296명의 대조군 각각에서 합산된 31개의 합산된 위험 대립 유전자의 수의 분포를 나타낸 것이다.

도 2는 본 발명의 실시예에 따른 wPRS의 급성 허혈성 뇌졸중(AIS) 위험 예측 효과를 검증하여, 그 결과를 나타낸 것이다.

도 3은 본 발명의 실시예에 따른 wPRS의 급성 허혈성 뇌졸중(AIS) 위험 예측 효과를 검증하여, 그 결과를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

[0037] 이하, 본 발명에 대하여 상세히 설명하기로 한다. 이에 앞서, 본 명세서 및 특허청구범위에 사용된 용어 또는 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서, 본 명세서에 기재된 실시예에 기재된 구성은 본 발명의 가장 바람직한 일 실시예에 불과할 뿐이고 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원 시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.

[0038] 한편, 본 명세서에서 개시된 각각의 설명 및 실시형태는 각각의 다른 설명 및 실시 형태에도 적용될 수 있다. 즉, 본 명세서에서 개시된 다양한 요소들의 모든 조합이 본 출원의 범주에 속한다. 또한, 하기 기술된 구체적인 서술에 의하여 본 발명의 범주가 제한된다고 볼 수 없다.

[0040] 본 발명에서 용어 "다형성(polymorphism)"이란 하나의 유전자 좌위(locus)에 두 가지 이상의 대립 유전자(allele)가 존재하는 경우를 말하며, 다형성 부위 중에서, 사람에 따라 단일 염기만이 다른 것을 단일염기다형성(single nucleotide polymorphism; SNP)이라 한다. 진단용 바이오 마커로 사용되기에 바람직한 단일염기다형성(SNP)은 선택된 집단에서 1% 이상, 더욱 바람직하게는 5% 또는 10% 이상의 발생 빈도를 나타내는 두 가지 이상의 대립 유전자를 가진다.

[0041] 본 발명에서 용어 "대립 유전자"는 상동 염색체의 동일한 유전자 좌위에 존재하는 한 유전자의 여러 타입을 의미한다. 대립 유전자는 다형성을 나타내는데 사용되기도 하며, 예컨대, 단일염기다형성(SNP)은 두 종류의 대립 인자(biallele)를 갖는다.

[0042] 본 발명에서 용어, "rsID"란 1998년부터 단일염기다형성(SNP) 정보를 축적하기 시작한 NCBI가 초기에 등록되는 모든 단일염기다형성(SNP)에 대하여 부여한 독립된 표지자를 의미한다. 본 발명에서는 rs627363과 같은 형태로 기재하였다. 이와 같은 표에 기재된 rsID는 본 발명에서 바이오 마커로 사용되는 단일염기다형성(SNP)을 의미한다. 당업자라면 상기 rsID를 이용하여 단일염기다형성(SNP)의 위치 및 서열을 용이하게 확인할 수 있을 것이다.

NCBI의 dbSNP(The Single Nucleotide Polymorphism Database) 번호인 rsID에 해당하는 구체적인 서열은 시간이 지남에 따라 약간 변경될 수 있으나, 본 발명의 범위가 상기 변경된 서열에도 미치는 것은 당업자에게 자명할 것이다. 상기 rsID의 유전자형은 “rsID-염기”와 같은 방식으로 나타낼 수 있다.

- [0043] 본 발명에서 진단 또는 예측의 대상이 되는 “뇌졸중(stroke)”은 뇌기능의 부분적 또는 전체적으로 급속히 발생한 장애가 상당 기간 이상 지속되는 것으로, 뇌혈관의 병 이외에는 다른 원인을 찾을 수 없는 상태를 일컫는다. 상기 뇌졸중은 뇌혈관이 막혀서 발생하는 뇌경색인 허혈성 뇌졸중(ischemic stroke)과 뇌혈관의 파열로 인해 뇌 조직 내부로 혈액이 유출되어 발생하는 뇌출혈인 출혈성 뇌졸중(cerebral hemorrhage)을 통틀어 일컫는 말로, 본 발명의 뇌졸중은 허혈성 뇌졸중일 수 있고, 바람직하게는 급성 허혈성 뇌졸중일 수 있다. 특히 상기 급성 허혈성 뇌졸중은 증상이 뚜렷하더라도 CT나 MRI검사 결과는 정상으로 나올 수 있기 때문에, 적절한 진단 및 치료를 위하여 본 발명의 진단 또는 예측 방법, 장치, 조성물 및 키트가 제공될 수 있다.
- [0044] 본 발명에서 용어 “예측”은, 특정 개체에 대하여 뇌졸중이 발병할 가능성이 있는지, 발병할 가능성이 있다면 뇌졸중이 발병할 가능성이 불특정 다수인에 비하여 상대적으로 높은 지 여부를 판별하는 것 또는 뇌졸중의 예후를 예측하는 것을 의미한다.
- [0045] 본 발명에서 “예후”는 상기 질환이 중증도로 진행할지 여부 및 질병에 기인한 사망 등의 병의 경과 및 완치 여부를 의미한다. 본 발명의 목적상 예후는 뇌졸중 환자의 질환 경과 또는 생존 예후를 의미하고, 상기 예후 예측은 질병의 경과 및 사망 또는 생존의 결과를 미리 예측하는 행위를 말한다. 상기 예후 예측 또는 예후 진단이란 질환의 경과가 환자의 생리적 또는 환경적 상태에 따라 달라질 수 있으며, 이러한 환자의 상태를 종합적으로 고려하여 치료 전/후 질병의 경과를 예측하는 모든 행위를 의미하는 것으로 해석될 수 있다.
- [0046] 본 발명에서 용어 “진단”은, 특정 개체에 대하여 뇌졸중이 이미 발병하였는지 여부를 판별하는 것을 의미한다.
- [0047] 본 발명에서 용어, “개체”는 뇌졸중이 발병되거나 발병될 가능성이 있는 모든 생물체를 의미하며, 구체적인 예로, 마우스, 원숭이, 소, 돼지, 미니돼지, 가축, 인간 등을 포함하는 포유동물, 양식어류 등을 포함할 수 있으며, 이에 제한되는 것은 아니다.
- [0048] 본 발명에서 용어, “시료” 내지 “생물학적 시료”는 뇌졸중이 발병되거나 발병될 가능성이 있는 개체로부터 유래된 임의의 물질, 생물학적 체액, 조직 또는 세포를 의미하는 것으로, 예를 들면, 전혈(whole blood), 백혈구(leukocytes), 말초혈액 단핵 세포(peripheral blood mononuclear cells), 백혈구 연층(buffy coat), 혈장(plasma), 혈청(serum), 객담(sputum), 눈물(tears), 점액(mucus), 세비액(nasal washes), 비강 흡인물(nasal aspirate), 호흡(breath), 소변(urine), 정액(semen), 침(saliva), 복강 세척액(peritoneal washings), 복수(ascites), 낭종액(cystic fluid), 뇌척수막 액(meningeal fluid), 양수(amniotic fluid), 선액(glandular fluid), 췌장액(pancreatic fluid), 림프액(lymph fluid), 흉수(pleural fluid), 유두 흡인물(nipple aspirate), 기관지 흡인물(bronchial aspirate), 활액(synovial fluid), 관절 흡인물(joint aspirate), 기관 분비물(organ secretions), 세포(cell), 세포 추출물(cell extract) 또는 뇌척수액(cerebrospinal fluid)을 포함할 수 있다.
- [0049] 또한, 이들 시료로부터 유전자 시료를 얻을 수 있으며, 유전자 시료는 핵산, 예를 들어, DNA, mRNA, 또는 mRNA로부터 합성되는 cDNA 등을 포함할 수 있으며, 이로부터 단일염기다형성(SNP)의 염기 정보를 확인할 수 있는 한, 그 종류는 이에 제한되는 것은 아니다.
- [0050] 본 발명에서, 단일염기다형성(SNP)을 검출 또는 증폭하는 체계는, RNA 시퀀싱(RNAseq) 또는 상기 유전자에 특이적으로 결합하는 프라이머, 프로브 및 안티센스 뉴클레오티드로 이루어진 군에서 선택된 1종 이상을 포함할 수 있다.
- [0051] 본 발명에서 상기 “프라이머”는 표적 유전자 서열을 인지하는 단편으로서, 정방향 및 역방향의 프라이머 쌍을 포함하나, 바람직하게는, 특이성 및 민감성을 가지는 분석 결과를 제공하는 프라이머 쌍이다. 프라이머의 핵산 서열이 시료 내 존재하는 비-표적 서열과 불일치하는 서열이어서, 상보적인 프라이머 결합 부위를 함유하는 표적 유전자 서열만 증폭하고 비특이적 증폭을 유발하지 않는 프라이머일 때, 높은 특이성이 부여될 수 있다.
- [0052] 본 발명에서, 상기 단일염기다형성(SNP)의 증폭에 사용되는 프라이머는, 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오티드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합체) 및 적당한 온도 하에서 주형-지시 DNA 합성의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오티드를 말한다. 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 bp 내지 30 bp 길이의 뉴클레오티드 일 수 있다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를

필요로 한다. 프라이머 서열은 주형과 완전하게 상보적일 필요는 없으나, 주형과 혼성화 할 정도로 충분히 상보적이어야 한다.

[0053] 본 발명에서 상기 "프로브"란 시료 내의 검출하고자 하는 표적 물질과 특이적으로 결합할 수 있는 물질을 의미하며, 상기 결합을 통하여 특이적으로 시료 내의 표적 물질의 존재를 확인할 수 있는 물질을 의미한다. 프로브의 종류는 당 업계에서 통상적으로 사용되는 물질로서 제한은 없으나, 바람직하게는 PNA(peptide nucleic acid), LNA(locked nucleic acid), 펩타이드, 폴리펩타이드, 단백질, RNA 또는 DNA일 수 있으며, 가장 바람직하게는 PNA이다. 보다 구체적으로, 상기 프로브는 바이오 물질로서 생물에서 유래되거나 이와 유사한 것 또는 생체 외에서 제조된 것을 포함하는 것으로, 예를 들어, 효소, 단백질, 항체, 미생물, 동식물 세포 및 기관, 신경세포, DNA, 및 RNA일 수 있으며, DNA는 cDNA, 게놈 DNA, 올리고뉴클레오타이드를 포함하며, RNA는 게놈 RNA, mRNA, 올리고뉴클레오타이드를 포함하며, 단백질의 예로는 항체, 항원, 효소, 펩타이드 등을 포함할 수 있다.

[0054] 본 발명에서 상기 "Nucleic Acid)"는 인공적으로 합성된, DNA 또는 RNA와 비슷한 중합체를 가리키며, 1991년 덴마크 코펜하겐 대학교의 Nielsen, Egholm, Berg와 Buchardt 교수에 의해 처음으로 소개되었다. DNA는 인산-리보스당 골격을 갖는데 반해, PNA는 펩타이드 결합에 의해 연결된 반복된 N-(2-아미노에틸)-글리신 골격을 가지며, 이로 인해 DNA 또는 RNA에 대한 결합력과 안정성이 크게 증가되어 분자 생물학, 진단 분석 및 안티센스 치료법에 사용되고 있다.

[0055] 본 발명에서 상기 "nucleic acids)"란, 2'-O, 4'-C 메틸렌 브릿지를 포함하는 핵산 아날로그를 의미한다. LNA 뉴클레오타이드는 DNA와 RNA의 일반적 핵산 염기를 포함하며, Watson-Crick 염기 쌍 규칙에 따라 염기 쌍을 형성할 수 있다. 하지만, 메틸렌 브릿지로 인한 분자의 'locking'으로 인해, LNA는 Watson-Crick 결합에서 이상적 형상을 형성하지 못하게 된다. LNA가 DNA 또는 RNA 올리고뉴클레오타이드에 포함되면, LNA는 보다 빠르게 상보적 뉴클레오타이드 사슬과 쌍을 이루어 이중 나선의 안정성을 높일 수 있다. 본 발명에서 상기 "안티센스"는 안티센스 올리고머가 왓슨-크릭 염기쌍 형성에 의해 RNA 내의 표적 서열과 혼성화되어, 표적서열 내에서 전형적으로 mRNA와 RNA: 올리고머 헤테로이중체의 형성을 허용하는, 뉴클레오타이드 염기의 서열 및 서브유닛간 백본을 갖는 올리고머를 의미한다. 올리고머는 표적 서열에 대한 정확한 서열 상보성 또는 근사 상보성을 가질 수 있다.

[0056] 본 발명에서, 상기 프로브는 대립 유전자 특이적 프로브로서, 같은 종의 두 구성원으로부터 유래한 핵산 단편 중에 단일염기다형성(SNP)이 존재하여, 한 구성원으로부터 유래한 DNA 단편에는 혼성화하나, 다른 구성원으로부터 유래한 단편에는 혼성화하지 않는다. 이 경우 혼성화 조건은 대립 유전자간의 혼성화 강도에 있어서 유의한 차이를 보여, 대립 유전자 중 하나에만 혼성화 하도록 충분히 엄격해야 한다. 이렇게 함으로써 다른 대립 유전자 형태 간에 좋은 혼성화 차이를 유발할 수 있다.

[0057] 본 발명에서, 상기 단일염기다형성(SNP)의 유전자형을 확인하는 것, 시퀀싱 분석, 예를 들어, 자동염기서열분석기를 사용한 시퀀싱 분석, 파이로시퀀싱(pyrosequencing), 마이크로어레이에 의한 혼성화, PCR-RELP(restriction fragment length polymorphism)법, PCR-SSCP(single strand conformation polymorphism)법, PCR-SSO(specific sequence oligonucleotide)법, PCR-SSO법과 도트 하이브리드화법을 조합한 ASO(allele specific oligonucleotide) 하이브리드화법, TaqMan-PCR법, MALDI-TOF/MS법, RCA(rolling circle amplification)법, HRM(high resolution melting)법, 프라이머 신장법, 서던 블롯 하이브리드화법, 도트 하이브리드화법 등의 방법에 의하여 수행될 수 있다.

[0058] 본 발명에서, 상기 단일염기다형성(SNP)의 염기를 결정하는 단계는, 예를 들어, 대상으로부터 얻은 시료로부터 분리한 핵산으로부터 단일염기다형성(SNP)을 포함하는 염기서열을 시퀀싱(sequencing)하고, 단일염기다형성(SNP)의 대립형질을 직접적으로 확인함으로써 수행될 수 있다. 염기서열의 시퀀싱은 당 업계에 공지된 방법을 통하여 수행될 수 있다. 예를 들어, 염기서열의 시퀀싱은 게체로부터 얻은 시료로부터 분리한 핵산을 주형으로 하고, 각 단일염기다형성(SNP)을 포함하는 유전자 부위를 증폭할 수 있는 프라이머를 사용하여 PCR을 수행하는 단계, 및 상기 PCR 반응물을 제한효소 절단방법으로 그 서열을 확인하는 단계를 순차적으로 수행함으로써 이뤄질 수 있다.

[0059] 본 발명에서, 상기 시퀀싱 분석은 차세대 염기서열 분석에 의할 수 있다. 상기 차세대 염기서열 분석(Next Generation Sequencing; NGS)은 주형 DNA를 대상으로 짧은 길이의 염기서열을 대용량으로 빠르게 생성시킬 수 있는 방식으로, 저렴한 비용과 빠른 데이터 생산으로 인해 전통적인 생거(Sanger) 시퀀싱 방식을 빠르게 대체하고 있는 분석 방법이다.

[0060] 본 발명에 따른 유전자의 서열 정보는 알려져 있으므로, 당업자라면 이를 바탕으로 상기 유전자에 특이적으로

결합하는 프라이머, 프로브 또는 안티센스 뉴클레오타이드를 용이하게 디자인할 수 있을 것이며 또한 특정 디자인 없이도 일반적인 RNA 시퀀싱 방법을 통해 정성 또는 정량 분석이 가능하다.

[0061] 본 발명에서 용어 “키트”는 바이오 마커 성분에 특이적으로 결합하는 프로브 또는 항체를 검출 가능한 표지로 표지하여 바이오 마커의 발현 수준을 평가할 수 있는 도구를 말한다. 프로브 또는 항체 관련하여 검출 가능한 물질을 기질과의 반응에 의해서 직접적으로 표지하는 것뿐만 아니라, 직접적으로 표지된 다른 시약과의 반응성에 의한 발색하는 표지체가 접합된 간접적 표지도 포함한다. 상기 표지체와 발색 반응할 발색 기질 용액, 세척액 및 기타 다른 용액 등을 포함할 수 있으며, 사용되는 시약 성분을 포함하여 제작될 수 있다. 본 발명에서 키트는 RT-PCR을 수행하기 위해 필요한 필수 요소를 포함하는 키트일 수 있으며, 마커 유전자에 대한 특이적인 각각의 프라이머 쌍 외에도 테스트 튜브, 반응 완충액, 테옥시뉴클레오타이드(dNTPs), Taq-중합효소, 역전사효소, DNase, RNase 억제제, 멸균수 등을 포함할 수 있다. 또한, 키트는 DNA 칩을 수행하기 위해 필요한 필수 요소를 포함하는 예후 예측용 유전자를 검출하기 위한 키트일 수 있다. DNA 칩 키트는 유전자 또는 그의 단편에 해당하는 cDNA가 프로브로 부착되어 있는 기판을 포함하고 기판은 정량 대조군 유전자 또는 그의 단편에 해당하는 cDNA를 포함할 수 있다. 본 발명의 키트는 당 업계에 공지되어 있는 것이라면, 이에 제한되지 않는다.

[0062] 본 발명에서 상기 키트는 RT-PCR 키트 또는 DNA 칩 키트일 수 있으나, DNA를 분석하기 위한 것이라면 제한없이 포함될 수 있다.

[0063] 본 발명의 상기 키트는 분석 방법에 적합한 한 종류 또는 그 이상의 다른 구성 성분 조성물, 용액 또는 장치를 더 포함할 수 있다. 예를 들면, 본 발명에서 상기 키트는 역전사 중합효소반응을 수행하기 위해 필요한 필수 요소를 더 포함할 수 있다. 역전사 중합효소반응 키트는 마커 단백질을 암호화하는 유전자에 대해 특이적인 프라이머 쌍을 포함한다. 프라이머는 상기 유전자의 핵산 서열에 특이적인 서열을 가지는 뉴클레오타이드로써, 약 7 bp 내지 50 bp의 길이, 보다 바람직하게는 약 10 bp 내지 30 bp의 길이를 가질 수 있다. 또한 대조군 유전자의 핵산 서열에 특이적인 프라이머를 포함할 수 있다. 그 외 역전사 중합효소반응 키트는 테스트 튜브 또는 다른 적절한 용기, 반응 완충액(pH 및 마그네슘 농도는 다양), 테옥시뉴클레오타이드(dNTPs), Taq-폴리머라아제 및 역전사효소와 같은 효소, DNase, RNase 억제제 DEPC-수(DEPC-water), 멸균수 등을 포함할 수 있다.

[0064] 또한, 본 발명의 상기 키트는 DNA 칩을 수행하기 위해 필요한 필수 요소를 포함할 수 있다. DNA 칩 키트는 유전자 또는 그의 단편에 해당하는 cDNA 또는 올리고뉴클레오타이드(oligonucleotide)가 부착되어 있는 기판, 및 형광 표지 프로브를 제작하기 위한 시약, 제제, 효소 등을 포함할 수 있다. 또한 기판은 대조군 유전자 또는 그의 단편에 해당하는 cDNA 또는 올리고뉴클레오타이드를 포함할 수 있다.

[0065] 본 발명의 상기 키트는 고정체를 더 포함할 수 있고, 상기 고정체로는 니트로셀룰로오즈 막, PVDF 막, 폴리비닐(polyvinyl) 수지 또는 폴리스티렌(polystyrene) 수지로 합성된 웰 플레이트(Well plate), 유리로 된 슬라이드 글래스 등이 사용될 수 있으나, 이에 제한되는 것은 아니다.

[0066] 또한, 본 발명의 상기 키트에서 2차 항체의 표지체는 발색 반응을 하는 통상의 발색제가 바람직하며, HRP(horseradish peroxidase), 염기성 탈인산화효소(alkaline phosphatase), 콜로이드 골드(colloid gold), FITC(폴리 L-라이신-플루오르세인 아이소티오시아네이트), RITC(로다민-B-아이소티오시아네이트) 등의 형광물질(fluorescein) 및 색소(dye) 등의 표지체가 사용될 수 있으나, 이에 제한되는 것은 아니다.

[0067] 또한, 본 발명의 상기 키트에서 발색을 유도하기 위한 발색 기질은 발색 반응을 하는 표지체에 따라 사용하는 것이 바람직하며, TMB(3,3',5,5'-테트라메틸 베지딘), ABTS[2,2'-아지노-비스(3-에틸벤조티아졸린-6-설폰산)], OPD(o-페닐렌디아민) 등을 사용할 수 있다. 이때, 발색 기질은 완충 용액(0.1 M NaAc, pH 5.5)에 용해된 상태로 제공되는 것이 더욱 바람직하다. TMB와 같은 발색기질은 이차 항체 접합체의 표지체로 사용된 HRP에 의해 분해되어 발색 침적체를 생성하고, 이 발색 침적체의 침적 정도를 육안으로 확인함으로써 상기 마커 단백질들의 존재 유무를 검출한다.

[0068] 본 발명의 상기 키트에서 세척액은 인산염 완충 용액, NaCl 및 트윈 20을 포함하는 것이 바람직하며, 0.02 M 인산염 완충용액, 0.13 M NaCl, 및 0.05% 트윈 20으로 구성된 완충 용액(PBST)이 더욱 바람직하다. 세척액은 항원-항체 결합 반응 후 항원-항체 결합체에 2차 항체를 반응시킨 다음 적당량을 고정체에 첨가하여 3 내지 6회 세척한다. 반응 정지 용액은 황산 용액(H₂SO₄)이 바람직하게 사용될 수 있다.

[0069] 본 발명에서, 용어 “모듈”, “유닛”, “부(part)” 등과 같은 용어는 적어도 하나의 기능이나 동작을 수행하는 구성요소를 지칭하기 위한 것이며, 이러한 구성요소는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 “모듈”, “유닛”, “부(part)” 등은 각각이 개별적인

특정한 하드웨어로 구현될 필요가 있는 경우를 제외하고는, 적어도 하나의 모듈이나 칩으로 일체화되어 적어도 하나의 프로세서로 구현될 수 있다.

[0070] 본 발명에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적인 연결뿐 아니라, 다른 매체를 통한 간접적인 연결의 경우도 포함한다. 또한, 어떤 부분이 어떤 구성요소를 포함한다는 의미는, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0071] 본 발명에서, 상기 장치는 하나 이상의 컴퓨터를 포함하는 다양한 전자기기 내지는 시스템으로 구현될 수 있다. 상기 장치는 적어도 하나의 단말 장치와 통신을 수행하기 위한 서버로 구현될 수 있다. 상기 장치는 의료용 로봇, 의료용 컴퓨터, 의료용 모니터링 시스템 등으로 구현될 수도 있다. 이 경우, 상기 장치는 다양한 바이오 마커를 측정하기 위한 다른 장치와 통신을 수행할 수 있다. 일 예로, 상기 장치는 스마트폰, 데스크탑 PC, 태블릿 PC, 노트북 PC, 웨어러블 장치, VR/AR 기기, PDA 등 다양한 단말 장치로 구현될 수도 있다.

[0072] 본 발명에서, 상기 장치는 프로세서를 포함할 수 있고, 상기 프로세서는 머신 러닝 모델 훈련 및 바이오 마커에 대한 정보를 분석하는 기능을 수행할 수 있다.

[0073] 본 발명에서, 상기 장치는 메모리를 포함할 수 있고, 상기 메모리는 구성요소들의 전반적인 동작을 제어하기 위한 운영체제(Operating System; OS), 적어도 하나의 머신 러닝 모델 및 데이터를 저장하기 위한 구성이다. 일 예로, 상기 메모리는 ROM, 플래시 메모리 등의 비휘발성 메모리를 포함할 수 있으며, DRAM 등으로 구성된 휘발성 메모리를 포함할 수 있다. 또한, 상기 메모리는 하드 디스크, SSD(Solid state drive) 등을 포함할 수도 있다.

[0075] 본 발명에서, 검출의 대상이 되는 단일염기다형성(SNP) 유전자 마커에 대한 정보는 하기 표 1과 같고, 본 발명에서 “단일염기다형성(SNP) 유전자 정보”는 상기 단일염기다형성 유전자형(SNP genotype)에 관한 정보일 수 있고, 상기 정보의 분석은 단일염기다형성 유전자형 분석(SNP genotyping)이라고 한다.

표 1

[0077]

서열 번호	유전자	염색체	SNP	비-위험 대립유전자	위험 대립유전자
1	GBA	1q22	rs75822236	C	T
2	TCF24	8q13.1	rs112859779	T	C
3	OLFML2A	9q33.3	rs79134766	A	G
4	ARHGAP32	11q24.3	rs371331393	G	A
5	CD163L1	12p13.31	rs138525217	C	T
6	CUL4A	13q34	rs74115822	G	A
7	LOC102724084	16q23.2	rs75861150	C	T
8	LRRC3	21q22.3	rs116969723	A	G
9	RNF144A	2p25.1	rs6741819	T	C
10	FLJ45964	2q37.3	rs59626274	T	C
11	SPCS3	4q34.2	rs17688188	A	G
12	LINGO2	9p21.1	rs56942085	G	A
13	MINK1	17p13.2	rs72835045	A	G
14	LINC01237	2q37.3	rs78458145	A	G
15	SLC2A14	12p13.31	rs118107419	A	C
16	DSCAM	21q22.2	rs727333	A	C
17	SUSD1	9q31.3	rs79461840	T	C
18	PRDM2	1p36.21	rs61775135	A	C
19	MYEOV	11q13.3	rs76855873	T	C
20	NAPA-AS1	19q13.32	rs55800589	C	G
21	SLC5A4-AS1	22q12.3	rs117398778	T	C
22	FMO4	1q24.3	rs3737926	T	C
23	SCARF1	17p13.3	rs3744644	C	G
24	SLC47A1	17p11.2	rs2440154	G	A
25	LINC00474	9q33.1	rs4979583	C	T

26	LINC02130	16p13.12	rs11646803	T	C
27	DRAM1	12q23.2	rs7964241	A	G
28	C9orf92	9p22.3	rs12350582	A	G
29	TMOD1	9q22.33	rs1052270	T	C
30	EDNRA	4q31.22	rs6841581	A	G
31	BOLL	2q33.1	rs700651	A	G

[0079] 본 발명의 일 구현예에 따르면, rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 단계; 및 상기 유전자 정보를 하기 수학적 식 1에 입력하는 단계;를 포함하는, 뇌졸중 예측 또는 진단을 위한 정보 제공 방법을 제공한다:

[0080] [수학적 식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

[0081]

[0082] 상기 수학적 식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i 는 상기 Var_i 에 부여되는 가중치이다. 또한 상기 w_i 는 n개의 단일염기다형성(SNP) 간 부여되는 대한 상대값이므로, 상기 n이 상기 w_i 가중치의 계산에 선 반영될 수 있어, 상기 n에 의하여 본 발명의 권리범위가 제한되지는 않는다.

[0083] 본 발명에서, 상기 유전자 정보는 목적하는 개체로부터 분리된 생물학적 시료에서 확인하는 것일 수 있다.

[0084] 본 발명에서, 상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것인, 방법.

[0085] 본 발명에서, 상기 w_i 는, 혈관 관련 질환인 심뇌혈관 질환(cardio cerebrovascular disease), 바람직하게는 뇌혈관 질환(cerebrovascular diseases; CVDs)과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학적 식 2를 만족하는 것인, 방법일 수 있다:

[0086] [수학적 식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

[0087]

[0088] 상기 수학적 식 2에서, 상기 k는, 0이 아닌 실수이고, 상기 OR_i 는, 상기 상관관계의 오즈비(odds ratio)이고, 상기 P_i 는, 상기 상관관계의 P값이다. 상기 뇌혈관 질환은 뇌동맥류일 수 있다.

[0089] 본 발명에서, 상기 k는 음수이고, 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것일 수 있다. 상기 PRS는 다중유전자위험점수(polygenic risk score; PRS)일 수 있고, 바람직하게는 가중치가 부여된 다중유전자위험점수(weighted polygenic risk score; wPRS)인 것일 수 있다.

[0090] 본 발명에서, 상기 점수는, 상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 높은 것; 상기 유전자형이 rs112859779-C인 경우 rs112859779-T인 경우에 비하여 높은 것; 상기 유전자형이 rs79134766-G인 경우 rs79134766-A인 경우에 비하여 높은 것; 상기 유전자형이 rs371331393-A인 경우 rs371331393-G인 경우에 비하여 높은 것; 상기 유전자형이 rs138525217-T인 경우 rs138525217-C인 경우에 비하여 높은 것; 상기 유전자형이 rs74115822-A인 경우 rs74115822-G인 경우에 비하여 높은 것; 상기 유전자형이 rs75861150-T인 경우 rs75861150-C인 경우에 비하여 높은 것; 상기 유전자형이 rs116969723-G인 경우 rs116969723-A인 경우에 비하

여 높은 것; 상기 유전자형이 rs6741819-C인 경우 rs6741819-T인 경우에 비하여 높은 것; 상기 유전자형이 rs59626274-C인 경우 rs59626274-T인 경우에 비하여 높은 것; 상기 유전자형이 rs17688188-G인 경우 rs17688188-A인 경우에 비하여 높은 것; 상기 유전자형이 rs56942085-A인 경우 rs56942085-G인 경우에 비하여 높은 것; 상기 유전자형이 rs72835045-G인 경우 rs72835045-A인 경우에 비하여 높은 것; 상기 유전자형이 rs78458145-G인 경우 rs78458145-A인 경우에 비하여 높은 것; 상기 유전자형이 rs118107419-C인 경우 rs118107419-A인 경우에 비하여 높은 것; 상기 유전자형이 rs727333-C인 경우 rs727333-A인 경우에 비하여 높은 것; 상기 유전자형이 rs79461840-C인 경우 rs79461840-T인 경우에 비하여 높은 것; 상기 유전자형이 rs61775135-C인 경우 rs61775135-A인 경우에 비하여 높은 것; 상기 유전자형이 rs76855873-C인 경우 rs76855873-T인 경우에 비하여 높은 것; 상기 유전자형이 rs55800589-G인 경우 rs55800589-C인 경우에 비하여 높은 것; 상기 유전자형이 rs117398778-C인 경우 rs117398778-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3737926-C인 경우 rs3737926-T인 경우에 비하여 높은 것; 상기 유전자형이 rs3744644-G인 경우 rs3744644-C인 경우에 비하여 높은 것; 상기 유전자형이 rs2440154-A인 경우 rs2440154-G인 경우에 비하여 높은 것; 상기 유전자형이 rs4979583-T인 경우 rs4979583-C인 경우에 비하여 높은 것; 상기 유전자형이 rs11646803-C인 경우 rs11646803-T인 경우에 비하여 높은 것; 상기 유전자형이 rs7964241-G인 경우 rs7964241-A인 경우에 비하여 높은 것; 상기 유전자형이 rs12350582-G인 경우 rs12350582-A인 경우에 비하여 높은 것; 또는 상기 유전자형이 rs1052270-C인 경우 rs1052270-T인 경우에 비하여 높은 것;이고, 양수인 것일 수 있다. 일례로 상기 유전자형이 rs75822236-T인 경우 rs75822236-C인 경우에 비하여 상기 점수가 높은 것은, 상기 rs75822236의 유전자형이 TT인 것이 TC 또는 CC인 것에 비하여 상기 점수가 높은 것; 또는 상기 rs75822236의 유전자형이 TC인 것이 CC인 것에 비하여 상기 점수가 높은 것;과 같이 해석될 수 있다. 일례로 상기 점수는 상기 단일염기다형성(SNP) 유전자형에 따라 0, 1 또는 2인 것일 수 있다.

[0091] 본 발명에서, 상기 점수는 상기 유전자형이 rs6841581-G인 경우 rs6841581-A인 경우에 비하여 높은 것; 또는 상기 유전자형이 rs700651-G인 경우 rs700651-A인 경우에 비하여 높은 것;일 수 있다.

[0092] 본 발명의 방법에 사용되는 단일염기다형성(SNP) 유전자는 상기 rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270를 모두 사용하는 것일 수 있고, 다른 단일염기다형성(SNP)이 추가되는 것일 수 있으며, 상기 n은 20 이상 40 이하인 것일 수 있고, 바람직하게는 25 이상 35 이하인 것일 수 있다.

[0094] 본 발명의 다른 구현예에 따르면, rs75822236, rs112859779, rs79134766, rs371331393, rs138525217, rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589, rs117398778, rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및 상기 유전자 정보를 상기 수학적 식 1에 입력하여 뇌졸중 위험도를 산출하는 연산부;를 포함하는, 뇌졸중 예측 또는 진단을 위한 장치를 제공하는 것일 수 있다.

[0095] 본 발명에서, 상기 장치는 목적하는 개체로부터 분리된 생물학적 시료로부터 유전자 서열을 검출하는 검출부를 더 포함하는 것일 수 있고, 상기 검출부는 상기 수신부에 상기 유전자 서열 내지 상기 유전자 정보를 전달할 수 있다.

[0096] 본 발명에서, 상기 유전자 정보는 rs6841581 및 rs700651로 이루어진 군에서 선택된 어느 하나 이상을 더 포함하는 것일 수 있다.

[0097] 본 발명에서, 상기 연산부는 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 뇌졸중 위험이 높은 것으로 판단하는 것일 수 있다.

[0098] 본 발명에서, 생략된 나머지 기재들은 앞서 기재된 정보 제공 방법에 관한 기재와 마찬가지로 해석될 수 있다.

[0100] 본 발명의 또 다른 구현예에 따르면, rs371331393, rs75822236, rs79134766, rs112859779, rs6841581, rs116969723, rs3737926 및 rs3744644rs75822236, rs112859779, rs79134766, rs371331393, rs138525217,

rs74115822, rs75861150, rs116969723, rs6741819, rs59626274, rs17688188, rs56942085, rs72835045, rs78458145, rs118107419, rs727333, rs79461840, rs61775135, rs76855873, rs55800589 및 rs117398778의 단일염기다형성(SNP) 유전자를 검출 또는 증폭할 수 있는 제제를 포함하는, 뇌졸중 예측 또는 진단용 조성물을 제공하는 것일 수 있다,

[0101] 본 발명에서, 상기 유전자는 rs138525217, rs118107419, rs78458145, rs727333, rs6741819, rs55800589, rs61775135, rs72835045, rs75861150, rs79461840, rs2440154, rs1052270, rs56942085, rs700651, rs59626274, rs17688188, rs76855873, rs74115822, rs11646803, rs117398778, rs4979583, rs12350582 및 rs7964241rs3737926, rs3744644, rs2440154, rs4979583, rs11646803, rs7964241, rs12350582 및 rs1052270을 더 포함하는 것일 수 있다.

[0102] 본 발명에서, 상기 유전자는 rs6841581 및 rs700651에서 선택된 어느 하나 이상을 더 포함하는 것일 수 있고, 상기 유전자 모두를 포함하는 것일 수 있다.

[0103] 본 발명에서, 생략된 나머지 기재들은 앞서 기재된 정보 제공 방법 및 장치에 관한 기재와 마찬가지로 해석될 수 있다.

[0105] 본 발명의 또 다른 구현예에 따르면, 본 발명의 조성물을 포함하는, 뇌졸중 예측 또는 진단용 키트를 제공할 수 있다.

[0106] 본 발명에서, 생략된 나머지 기재들은 앞서 기재된 정보 제공 방법, 장치 및 조성물에 관한 기재와 마찬가지로 해석될 수 있다.

[0108] 이하 본 발명을 실시예에 의하여 더욱 상세하게 설명한다. 이들 실시예는 단지 본 발명을 보다 구체적으로 설명하기 위한 것으로, 본 발명의 범위가 이들 실시예에 국한되지 않는다는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.

[0110] 실시예

[0112] 1. 코호트 획득

[0114] 본 발명은 2015년 3월부터 2020년 12월까지 5곳의 대학 병원에서 수집한 다양한 전향적 뇌혈관 질환(cerebrovascular diseases; CVDs) 환자의 데이터를 기반으로 하였다. 또한 상기 CVDs와 관련되어, 게놈 전체 연관성 연구(genome-wide association study; GWAS), 전체-엑솜/전사체 시퀀싱, 후성 유전체 스크리닝과 같은 다양한 유전자 분석이 수행되었다. 병원 기반 관찰 코호트 연구는 1) 18세 이상의 성인 환자, 2) 급성 허혈성 뇌졸중(AIS), 출혈성 뇌졸중, 뇌동맥류, 모야모야병 및 혈관 기형과 같은 CVDs 있는 환자를 등록했다. 또한 본 발명의 실시예에서는 치매 및 파킨슨병을 포함한 신경 퇴행성 질환을 가진 CVDs 환자는 제외시켰고, 임상 특성, 의료 정보, 방사선학적 데이터가 수집 및 업데이트되었으며, 기관 검토위원회 및 윤리위원회의 승인을 받았다.

[0116] 2. 유전형 및 품질 관리

[0118] AxiomTH Asian Precision Medicine Research Array(APMRA, Thermo Fisher Scientific, MA, USA)를 사용하여, 환자의 말초 혈액에서 유래된 게놈 DNA에 대한 유전자형을 분석한 결과가 GWAS로 보고되었다. 고품질 플레이트는 샘플에 대해 95%를 초과하는 플레이트 통과율을 보였고 평균 호출 속도는 99%를 초과하여 통과되었으며, 798,148개의 SNP 중 512,575개의 SNP가 품질 관리 필터를 통과하였다. 구체적으로는 유전형 호출율은 95% 이상, 부대립 유전자 빈도(minor allele frequency; MAF)는 1% 이상, Hardy-Weinberg 평형 P값은 1×10^{-6} 이상이었다.

[0120] 3. 통계 분석

[0122] 다중유전자위험점수(polygenic risk score; PRS)는 총 31개의 단일염기다형성(single nucleotide polymorphisms; SNP)에 위험 대립 유전자(risk alleles) 및 효과 크기로 가중치를 부여하여 산출되었고, 그 결과를 하기 표 2 및 표 3에 나타내었다. 상기 31개의 SNP는 하기 표 2와 같이 29개의 SNP는 다른 연구에서 확인된 게놈 전체 신호(genome-wide signals)가 포함되었고($p < 5.0 \times 10^{-8}$), 하기 표 3과 같이 추가로 보고된 2개의 SNP가 포함($p < 0.01$)되었다. 단, 여기서 OR는 오즈비(odds ratio); SNP는 단일염기다형성(single nucleotide polymorphism), NR 및 R은 각각 비-위험 대립유전자 및 위험 대립유전자를 의미한다.

표 2

[0124]

유전자	염색체	SNP	NR	R	ln(OR)	OR	P값
<i>PRDM2</i>	1p36.21	rs61775135	A	C	1.153597	0.3155	3.59E-13
<i>GBA</i>	1q22	rs75822236	C	T	5.083886	161.4	1.09E-19
<i>FMO4</i>	1q24.3	rs3737926	T	C	1.014731	0.3625	1.83E-10
<i>RNF144A</i>	2p25.1	rs6741819	T	C	1.382302	0.251	4.05E-14
<i>FLJ45964</i>	2q37.3	rs59626274	T	C	1.342852	0.2611	5.78E-14
<i>LINC01237</i>	2q37.3	rs78458145	A	G	1.407722	0.2447	3.14E-15
<i>SPCS3</i>	4q34.2	rs17688188	A	G	1.312301	0.2692	5.99E-13
<i>TCF24</i>	8q13.1	rs112859779	T	C	1.686319	0.1852	3.33E-16
<i>C9orf92</i>	9p22.3	rs12350582	A	G	0.888714	2.432	1.49E-08
<i>LINGO2</i>	9p21.1	rs56942085	G	A	1.809108	6.105	3.90E-08
<i>TMOD1</i>	9q22.33	rs1052270	T	C	0.893063	0.4094	2.71E-08
<i>SUSD1</i>	9q31.3	rs79461840	T	C	1.933693	6.915	1.40E-08
<i>LINC00474</i>	9q33.1	rs4979583	C	T	0.896496	2.451	4.61E-09
<i>OLFML2A</i>	9q33.3	rs79134766	A	G	1.972562	0.1391	1.70E-19
<i>MYEOV</i>	11q13.3	rs76855873	T	C	1.107451	0.3304	1.23E-11
<i>ARHGAP32</i>	11q24.3	rs371331393	G	A	3.774369	43.57	9.32E-27
<i>CD163L1</i>	12p13.31	rs138525217	C	T	4.33047	75.98	6.20E-23
<i>SLC2A14</i>	12p13.31	rs118107419	A	C	1.395537	0.2477	1.44E-14
<i>DRAM1</i>	12q23.2	rs7964241	A	G	0.910675	2.486	1.95E-08
<i>CUL4A</i>	13q34	rs74115822	G	A	1.831781	6.245	1.12E-10
<i>LINC02130</i>	16p13.12	rs11646803	T	C	0.790319	0.4537	4.76E-10
<i>LOC102724084</i>	16q23.2	rs75861150	C	T	1.934476	0.1445	1.27E-08
<i>SCARF1</i>	17p13.3	rs3744644	C	G	0.943405	0.3893	6.02E-09
<i>MINK1</i>	17p13.2	rs72835045	A	G	1.382701	0.2509	1.69E-12
<i>SLC47A1</i>	17p11.2	rs2440154	G	A	0.981704	2.669	1.42E-08
<i>NAPA-AS1</i>	19q13.32	rs55800589	C	G	0.963644	0.3815	3.35E-13
<i>DSCAM</i>	21q22.2	rs727333	A	C	1.364533	0.2555	3.69E-14
<i>LRRC3</i>	21q22.3	rs116969723	A	G	1.447744	0.2351	3.83E-15
<i>SLC5A4-AS1</i>	22q12.3	rs117398778	T	C	1.318818	3.739	2.00E-09

표 3

[0126]

유전자	염색체	SNP	NR	R	ln(OR)	OR	P값
<i>BOLL</i>	2q33.1	rs700651	A	G	0.34713	1.415	0.00789

EDNRA	4q31.22	rs6841581	A	G	0.630548	0.5323	0.00065
-------	---------	-----------	---	---	----------	--------	---------

[0127]

[0128]

상기 표 2 및 표 3에 나타난 것처럼, 총 31개의 선택된 SNP가 0.8 미만의 연쇄불균형(linkage disequilibrium; LD)을 보였으며, 베타계수는 오즈비(OR)를 자연로그 변환한 $\ln(OR)$ 를 사용하였다. 개별 위험 점수는 추가 상속 모델(즉, 위험 대립 유전자의 0, 1 또는 2 개 사본)하에 유전자형 대치 방법을 추가하고 변이체의 효과 크기를 곱한 PLINK v1.9 점수 옵션으로 계산되었다. 예측 모델은 하기 수학식 1 및 수학식 2를 기반으로 계산되었다.

[0129]

[수학식 1]

$$PRS = \frac{\sum_{i=1}^n w_i Var_i}{n}$$

[0130]

[0131]

[수학식 2]

$$w_i = k \log(OR_i) \cdot \log(P_i)$$

[0132]

[0133]

상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i 는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이다.

[0134]

또한 상기 수학식 1 및 수학식 2에서, 상기 w_i 는, 상기 Var_i 에 부여되는 가중치로, 뇌혈관 질환과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이다.

[0135]

또한 상기 수학식 1 및 수학식 2에서, 상기 k는 $-e^{-1}$ 이고, 상기 OR_i 는 상기 상관관계의 오즈비(odds ratio; OR)이고, 상기 P_i 는 상기 상관관계의 P값이다.

[0137]

상기 PRS에 가중치가 부여된 wPRS의 위험 평가는 급성 허혈성 뇌졸중(AIS) 환자 222명의 데이터 세트 값에서 연령, 성별, 고혈압, 당뇨병, 고지혈증, 흡연상태 및 4가지 주요 성분 분석(PCA) 값을 조정하여, 일반화된 선형 모델에서 수행되었다. 296명의 피험자로 구성된 비-CVD 대조군은 상기 급성 허혈성 뇌졸중(AIS) 환자의 데이터 세트에서 공유되었다. PCA 값은 SNP를 공유하는 한국인 768명과 1000-게놈 참조 패널 v3(<ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release>)의 샘플 2,504개를 포함하는 통합 데이터 세트에서 추정되었고, 5% 이상의 부대립 유전자 빈도(minor allele frequency; MAF)로 나타났다. 상기 위험 모델의 예측 가능성, 민감도 및 특이성은 수신기 작동 특성 곡선(area under the receiver operating characteristic curve; AUROC)의 아래 영역을 계산하여 평가하였고, 교차 검증되었다. PRS를 기반으로 한 최적의 위험 모델은 최저, 중간 및 최고 위험 그룹의 3분위수로 계층화되었다. 로지스틱 회귀 분석은 10개의 교란 요인을 조정한 후 오즈비(OR) 및 95% 신뢰구간(CI)을 추정하기 위해 수행되었다. 모든 단변량 분석 및 다변량 분석은 STATA 소프트웨어 패키지 v.11.2 (Stata Corp., TX, USA)를 사용하여 수행되었다.

[0139]

4. 기준특성 및 다유전자 위험 점수 확인

[0141]

222명의 급성 허혈성 뇌졸중(AIS) 환자는 평균 연령이 69.5 ± 12.9 세(평균±표준편차[SD])인 남성 118명, 여성 104명이 포함되었고, 대조군으로는 평균 연령이 52.1 ± 16.6 세인 남성 142명, 여성 154명이 포함되었다. 상기 222명의 급성 허혈성 뇌졸중(AIS) 환자와 296명의 대조군 각각에서 합산된 31개의 합산된 위험 대립 유전자의 수의 분포를 도 1에 나타내었다. 단, 여기서 Y축은 표본 크기를 나타내고, X축은 위험 대립유전자의 수를 나타내며, 빨간색과 파란색 막대는 각각 급성 허혈성 뇌졸중(AIS)이 발생한 케이스와 대조군을 나타낸 것이다.

[0142]

도 1에 나타난 것처럼, 31개의 예측 변이체에 의해 구성된 합산된 위험 대립유전자의 수는 급성 허혈성 뇌졸중(AIS) 환자에서 16개 내지 43개 카피 분포가 나타났고, 그 중 위험 대립 유전자의 평균은 상기 급성 허혈성 뇌

졸중(AIS)이 발생한 환자의 케이스에서 37개의 사본을 보이는데 비하여, 대조군에서는 26개 사본을 보이는 바, 본 발명의 급성 허혈성 뇌졸중(AIS) 위험 예측 효과를 확인할 수 있었다.

5. 다중 대립유전자에 가중치를 부여한 모델의 뇌졸중 예측 능력 확인

대립 유전자 스코어링 시스템에 기반하고, 효과 크기에 따라 유전자별로 순위 및 가중치를 부여한 결과를 하기 하기 표 4 및 표 5에 나타내었다.

표 4

순위	SNP	BP	Function	Effect	P값
1	rs371331393	128839272	stopgain	3.774369	9.32E-27
2	rs75822236	155204793	R535H	5.083886	1.09E-19
3	rs79134766	127561723	A208T	1.972562	1.70E-19
4	rs112859779	67860369	G141S	1.686319	3.33E-16
5	rs6841581	148401190	Near 3-UTR	0.630548	0.00065
6	rs116969723	45876716	P63P	1.447744	3.83E-15
7	rs3737926	171303565	F281F	1.014731	1.83E-10
8	rs3744644	1538628	E639E	0.943405	6.02E-09
9	rs138525217	7528192	splicing	4.33047	6.20E-23
10	rs118107419	7980394	intronic	1.395537	1.44E-14
11	rs78458145	242877035	ncRNA_intronic	1.407722	3.14E-15
12	rs727333	41702947	intronic	1.364533	3.69E-14
13	rs6741819	7147973	intronic	1.382302	4.05E-14
14	rs55800589	47990889	ncRNA_intronic	0.963644	3.35E-13
15	rs61775135	14148407	intronic	1.153597	3.59E-13
16	rs72835045	4780795	intronic	1.382701	1.69E-12
17	rs75861150	80193070	ncRNA_intronic	1.934476	1.27E-08
18	rs79461840	114822158	ncRNA_intronic	1.933693	1.40E-08
19	rs2440154	19444987	intronic	0.981704	1.42E-08
20	rs1052270	100275424	intronic	0.893063	2.71E-08
21	rs56942085	28518911	intronic	1.809108	3.90E-08
22	rs700651	198631714	intronic	0.34713	0.00789
23	rs59626274	240500249	intergenic	1.342852	5.78E-14
24	rs17688188	177471059	intergenic	1.312301	5.99E-13
25	rs76855873	69154455	intergenic	1.107451	1.23E-11
26	rs74115822	113931854	intergenic	1.831781	1.12E-10
27	rs11646803	14463359	intergenic	0.790319	4.76E-10
28	rs117398778	32609178	intergenic	1.318818	2.00E-09
29	rs4979583	118571428	intergenic	0.896496	4.61E-09
30	rs12350582	16135142	intergenic	0.888714	1.49E-08
31	rs7964241	102338235	intergenic	0.910675	1.95E-08

표 5

순위	SNP	특징 1-엑손	특징 2-세부기능	특징 3-유의수준(P값)
1	rs371331393	엑손, 기능변이	기능과발현, 단백질변형	단백질변형 중 Top1 SNP
2	rs75822236	엑손, 기능변이	단백질변형	단백질변형 중 Top2 SNP
3	rs79134766	엑손, 기능변이	단백질변형	단백질변형 중 Top3 SNP
4	rs112859779	엑손, 기능변이	단백질변형	단백질변형 중 Top4 SNP
5	rs6841581	UTR프로모토근처	유전자생성에간접적영향	아시아/유럽메타분석 재검증됨
6	rs116969723	엑손, 기능변이	단백질미변형, 기능변이	단백질미변형 중 Top1

7	rs3737926	엑솜,기능변이	단백질미변형,기능변이	단백질미변형 중 Top2
8	rs3744644	엑솜,기능변이	단백질미변형,기능변이	단백질미변형 중 Top3
9	rs138525217	유전자내	mRNA작동시절단변이	GWAS p-value < 5E-8
10	rs118107419	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
11	rs78458145	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
12	rs727333	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
13	rs6741819	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
14	rs55800589	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
15	rs61775135	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
16	rs72835045	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
17	rs75861150	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
18	rs79461840	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
19	rs2440154	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
20	rs1052270	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
21	rs56942085	유전자내	제거되는부분기능없음	GWAS p-value < 5E-8
22	rs700651	유전자내	메타분석검증,역할모름	아시아인연구에서 재검증됨
23	rs59626274	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
24	rs17688188	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
25	rs76855873	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
26	rs74115822	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
27	rs11646803	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
28	rs117398778	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
29	rs4979583	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
30	rs12350582	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8
31	rs7964241	유전자 간 사이	후보유전자표식자,기능없음	GWAS p-value < 5E-8

[0152] 다음으로, 상기 결과에 따라 가중치를 부여한 wPRS의 급성 허혈성 뇌졸중(AIS) 위험 예측 효과를 검증하여, 그 결과를 도 2, 도 3 및 하기 표 6에 나타내었다.

[0153] 단, 여기서 Case는 222명의 급성 허혈성 뇌졸중(AIS) 환자; Control은 296명의 대조군; T1은 계층화된 가장 다 유전자 위험 모델의 삼분위수에서 가장 낮은 급성 허혈성 뇌졸중(AIS) 위험을 가진 경우; T2는 급성 허혈성 뇌졸중(AIS) 위험이 중간 정도인 경우; T3는 급성 허혈성 뇌졸중(AIS) 위험이 가장 높은 경우; OR는 오즈비; CI는 신뢰 구간; AUROC는 수신 작동 특성 곡선 아래 영역을 의미한다. 또한 상기 오즈비, 95% 신뢰 구간 및 P값은 연령, 성별, 고혈압, 당뇨병, 고지혈증, 흡연 여부 및 4가지 주성분 값을 조정한 후 다변량 로지스틱 회귀 분석으로 추정되었다. 또한 민감도, 특이도 및 AUROC는 STATA 소프트웨어의 "roctab" 패키지로 추정되었다.

표 6

[0155]

모델	Case, N (%)	Control, N (%)	OR (95% CI)	P값	민감도	특이도	AUROC
T1: 0.290-0.712	36 (16.2)	213 (72.0)	1.00				
T2: 0.712-0.789	123 (55.4)	72 (24.3)	11.39 (6.43-20.17)	7.7×10^{-17}	0.838	0.720	
T3: 0.789-1.126	63 (28.4)	11 (3.7)	39.76 (16.91-93.46)	3.1×10^{-17}	0.284	0.963	0.803

[0157] 도 2에서 나타난 것처럼 상기 wPRS는 급성 허혈성 뇌졸중(AIS) 위험을 유의한 정확도로 예측한 것을 확인하였다 (AUROC = 0.842, 95% CI : 0.808-0.876). 또한 도 3 및 상기 표 6에서 나타난 것처럼 각 PRS 모델에 따라 개별 위험을 3분위수로 계층화 하였을 때, T3 그룹의 급성 허혈성 뇌졸중(AIS) 위험이 T1 그룹보다 현저하게 높았고 (OR = 39.76), 특히 계층화 후에도 여전히 높은 임상 민감도, 특이성 및 예측 가능성을 나타낸 바(각각 0.728, 0.963, 0.930), 본 발명의 wPRS 시스템은 급성 허혈성 뇌졸중(AIS) 환자와 대조군 간의 차이를 식별하는 데 도

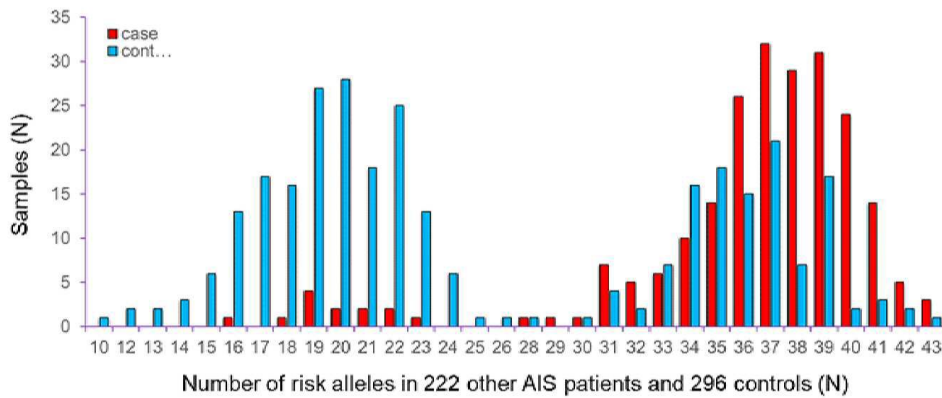
움을 주는 것을 확인하였다(정확도 = 67.2%).

[0159]

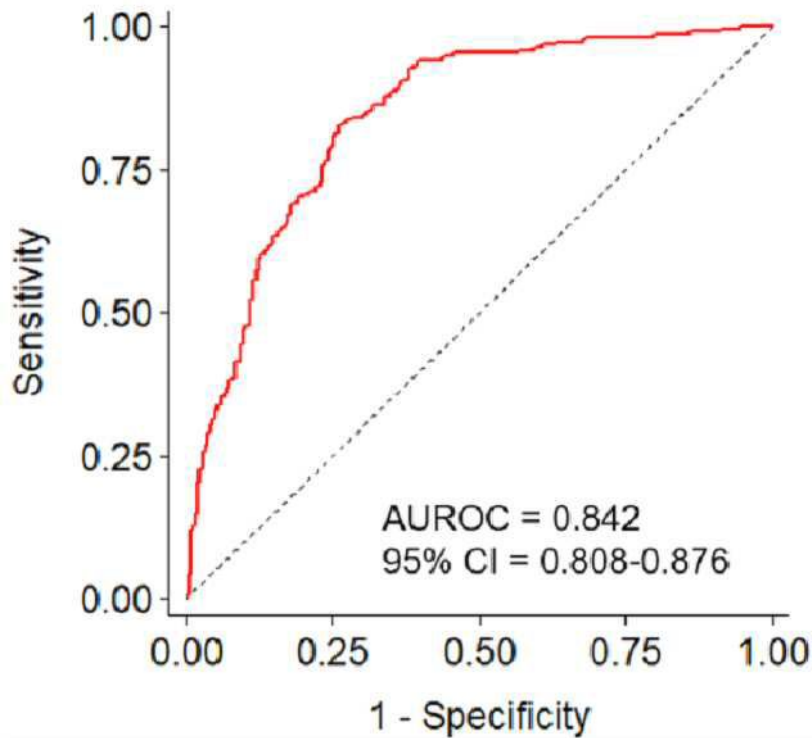
이상으로 본 발명의 특정한 부분을 상세히 기술 하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시예일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서, 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

도면

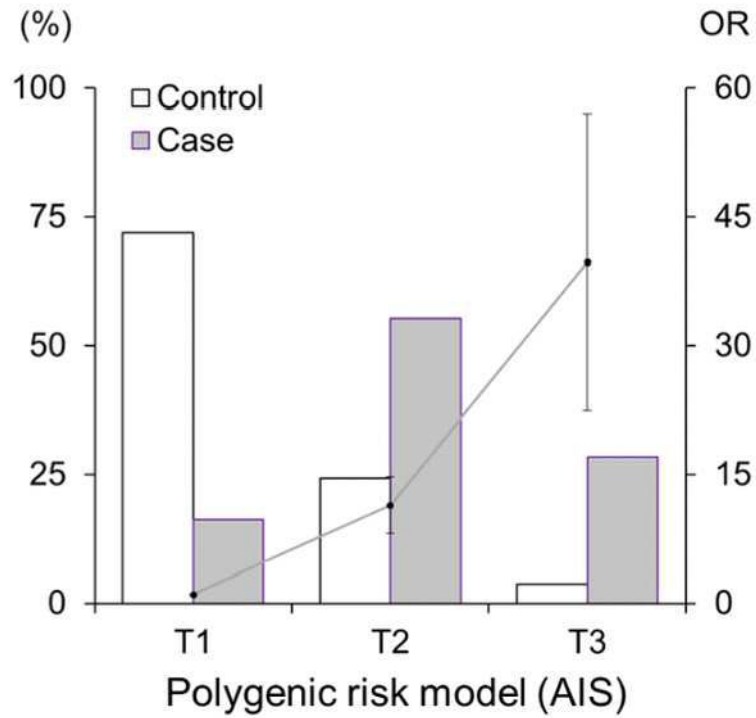
도면1



도면2



도면3



서열 목록

- <110> Industry Academic Cooperation Foundation Hallym University
 - <120> Single nucleotide polymorphism marker for stroke diagnosis and stroke diagnosis method using the same
 - <130> P212100
 - <160> 31
 - <170> KoPatentIn 3.0
 - <210> 1
 - <211> 51
 - <212> DNA
 - <213> Homo sapiens
 - <220><221> allele
 - <222> (26)
 - <223> y = c or t
 - <400> 1
- ccttgagtat ctgtccatc actggygacg ccacaggtag gtgtgaatgg a
- <210> 2
 - <211> 51

51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> y = c or t

<400> 2

tgcttcagaa actgaccagt agcacysgatg tacaatcttg atcgcatggg c 51

<210> 3

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 3

cctccagctg ctgcagaagg atgcccrcgc cgcccctgcc acccctgcca c 51

<210> 4

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 4

ttgtggttca gtcttaaaga ttcttrcccg gatgcagcat attttactcc a 51

<210> 5

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> y = c or t

<400> 5

ttcacaaagtc ggacatctgt atatcytagg aggagacaag gccatagaag a 51

<210> 6

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 6

gaatccacgt tgaaaaatct gtccarggct cacccagtat caacaatagt g 51

<210> 7

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> y = c or t

<400> 7

atatctttct gtgcctttat ttattytgca tatgttaaaa gtggatgagg c 51

<210> 8

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 8

ttcaggaggt ccccgaggac atccrgcca acaccgtgct cctgaagctc g 51

<210> 9

<211> 51

<212> DNA

<213> Homo sapiens

<220><221>

> allele

<222> (26)
 <223> y = c or t
 <400> 9
 tattgaacag atagaacaaa cgggcygcgc tgccacacc accgtttctc t 51
 <210> 10
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> y = c or t
 <400> 10
 gaggcccgag aaccaggcac gccgaygaca gcacacttcc agcctgagag c 51
 <210> 11
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> r = a or g
 <400> 11
 acagagaaaa tgtgggaggc aatacrgtga cagtgggtga attaaaatgt t 51
 <210> 12
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> r = a or g
 <400> 12
 aagttttgac tattctactg gcctgratit atctaacttg tgttgactt t 51
 <210> 13
 <211> 51
 <212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 13

ctcagttctca gggggactgg gaagarggaa ggacaaaagg atggtggccc t 51

<210> 14

<211> 51

<212> DNA

<213> Homo sapiens

<220>

><221> allele

<222> (26)

<223> r = a or g

<400> 14

gtgggagatg gatttgggca ggaagrtcca ggctgtacca ggctccttct t 51

<210> 15

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> m = a or c

<400> 15

ggtgaggtga tgggtagcat ccattmctga acctgccttc tcacctcatt c 51

<210> 16

<211> 51

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> m = a or c

<400> 16

cagtgtatgg ccacacacaa aagacmccag caaactggag cttaagttaa c 51

<210> 17
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> y = c or t
 <400> 17

tctgtcaggc ttgatagcag ggatcygaac gactgacttg tattcttgca a 51

<210> 18
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> m = a or c
 <400> 18

agatgctgga gtgagtgagc ctgacmtggg ttcaagtctc acctctgctg c 51

<210> 19
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> y = c or t
 <400> 19

catgcactgc ttgtgttca ctgccygccc agagttgtct caaacatgct g 51

<210> 20
 <211> 51
 <212> DNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> s = g or c

<400> 20
 tgggtagcag aggtgaggag gaagahagac tcatgcaaaa gtgtgtttt t 51

<210> 21
 <211> 51
 <212> DNA
 <213> Homo sapiens

<220
 ><221> allele
 <222> (26)
 <223> y = c or t

<400> 21
 tgttgggcac ctctagtcag tcaccytgaa gctcttctcc aagtgtcatt t 51

<210> 22
 <211> 51
 <212> DNA
 <213> Homo sapiens

<220><221> allele
 <222> (26)
 <223> y = c or t

<400> 22
 tgtgtcaaag gaaaaaagca aaattyattg tgaatgatga gctgccaac t 51

<210> 23
 <211> 51
 <212> RNA
 <213> Homo sapiens

<220><221> allele
 <222> (26)
 <223> s = g or c

<400> 23
 gaaaggactc gggggcttct gcttctctg ggcctgtgga ctcttcggct t 51

<210> 24
 <211> 51
 <212> RNA
 <213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 24

gctacctcca gctcctcctg tctacrgact tccttgctgg aatgtgttct t 51

<210> 25

<211> 51

<212> RNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> y = c or t

<400> 25

atgataaacc atctcaaaat ttagtygtgc aaaacaactt tttctgtttg c 51

<210> 26

<211> 51

<212> RNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> y = c or t

<400> 26

ctgcctgtgt cctctgcaga cagcgyctgc tcctccaca catccaaagc c 51

<210> 27

<211> 51

<212> RNA

<213> Homo sapiens

<220><221> allele

<222> (26)

<223> r = a or g

<400> 27

atataggagg gccacatctt aggggrcaga taaatcaca gactgtagac t 51

<210> 28

<211> 51

<212> RNA
 <213> Homo sapiens
 <220>
 ><221> allele
 <222> (26)
 <223> r = a or g
 <400> 28
 ggaaaccagg atagtgggttg cttatrgggg agaagaccga cttgggggag a 51
 <210> 29
 <211> 51
 <212> RNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> y = c or t
 <400> 29
 tgatttatgc tgggtctaca gtcacyggag agccccacat ctttctcatg g 51
 <210> 30
 <211> 51
 <212> RNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> r = a or g
 <400> 30

 cacaagtctt tgtggagaga cgcacrtgga gaaaaactaa cactcaacac c 51
 <210> 31
 <211> 51
 <212> RNA
 <213> Homo sapiens
 <220><221> allele
 <222> (26)
 <223> r = a or g
 <400> 31

ggaaacagtt aatgtttata atagtrttta gaactgaatg aatagtcgta t

51