(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0166278
(43) 공개일자 2023년12월07일

(51) 국제특허분류(Int. Cl.)
G16B 20/20 (2019.01) *A61B 5/00* (2021.01)
C12Q 1/6883 (2018.01) *G16H 50/50* (2018.01)
(52) CPC특허분류
G16B 20/20 (2019.02)
A61B 5/4842 (2013.01)
(21) 출원번호 10-2022-0066025
(22) 출원일자 2022년05월30일
심사청구일자 2022년05월30일

(71) 출원인
한림대학교 산학협력단
강원도 춘천시 한림대학길 1, 한림대학교(옥천동)
(72) 발명자
전진평
경기도 용인시 수지구 동천로 75(동천동)
홍은표
강원도 춘천시 영서로 2317, 102동 801(온의동, 럭키아파트)
윤동혁
충청북도 단양군 어상천면 어상천로 1589-8
(74) 대리인
특허법인본

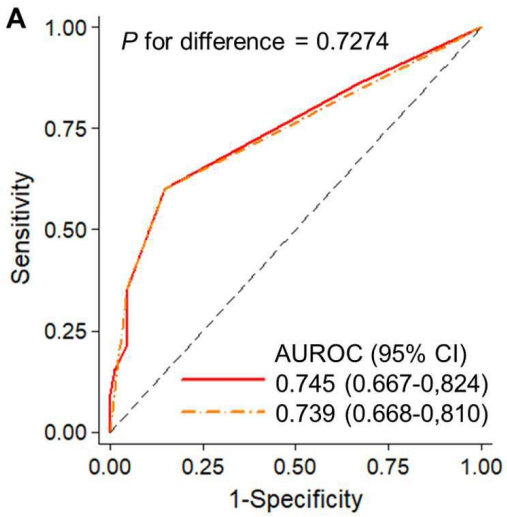
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **복수개의 SNP 마커를 이용한 지주막하 출혈의 예후 예측용 모델**

(57) 요 약

본 발명은 유전자 재조합율이 높은 영역을 포함하는 복수개의 SNP를 확인하여 완성된 지주막하 출혈의 예후 예측용 모델에 관한 것이다. 구체적으로는 상기 SNP 유전자 정보를 수신하는 단계; 및 상기 유전자 정보를 하기 수학식 1에 입력하는 단계;를 포함하는, 지주막하 출혈의 예후 예측용 모델 및 이를 이용한 진단 장치에 관한 것이다.

대 표 도 - 도7



(52) CPC특허분류
C12Q 1/6883 (2022.01)
G16H 50/50 (2018.01)
C12Q 2600/156 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1465036270
과제번호	HR21C0198040022
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	연구중심병원육성R&D
연구과제명	유전체 · 인공지능 기반 환자맞춤형 신치료 개발 및 관리 기술 개발
기 여 율	1/1
과제수행기관명	한림대학교 산학협력단
연구기간	2021.07.01 ~ 2029.12.31

명 세 서

청구범위

청구항 1

rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 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,

지주막하 출혈의 예후 예측용 정보제공 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_i는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i는, 상기 Var_i에 부여되는 가중치이고, 상기 PRS는 다중유전자위험점수이다.

청구항 2

제1항에 있어서,

상기 유전자 정보는 아포지단백 E(Apolipoprotein E; APOE), 합토글로빈(Haptoglobin; Hp)으로 이루어진 군에서 선택된 어느 하나 이상의 유전자의 정보를 더 포함하는 것인,

정보제공 방법.

청구항 3

제1항에 있어서,

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

[수학식 2]

$$w_i = k \log(HR_i)$$

상기 수학식 2에서, 상기 k는, 0이 아닌 실수이고, 상기 HR_i는, 상기 상관관계의 위험비(Hazard ratio)이다.

청구항 4

제1항에 있어서,

상기 유전자 정보는 아포지단백 E(Apolipoprotein E; APOE), 합토글로빈(Haptoglobin; Hp)으로 이루어진 군에서 선택된 어느 하나 이상의 유전자의 정보를 더 포함하는 것인,

정보제공 방법.

청구항 5

제1항에 있어서,

상기 k는 음수이고, 상기 연산부는 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 지주막하 출혈 후 인지기능 저하 위험이 높은 것으로 판단하는 것인,

정보제공 방법.

청구항 6

제1항에 있어서,

상기 n은 5 이상 100 이하인 것인,

정보제공 방법.

청구항 7

제1항에 있어서,

상기 점수는 상기 유전자형이 rs10503670-A인 경우, rs10503670-G인 경우에 비하여 높은 것인,

정보제공 방법.

청구항 8

제1항에 있어서,

상기 점수는 상기 유전자형이 rs56823384-C인 경우 rs56823384-T인 경우에 비하여 높은 것인,

정보제공 방법.

청구항 9

제1항에 있어서,

상기 점수는 상기 유전자형이 rs76507772-C인 경우, rs76507772-A인 경우에 비하여 높은 것인,

정보제공 방법.

청구항 10

제1항에 있어서,

상기 점수는 상기 유전자형이 rs138753053-A인 경우, rs138753053-G인 경우에 비하여 높은 것인,

정보제공 방법.

청구항 11

제1항에 있어서,

상기 점수는 상기 유전자형이 rs145397166-G인 경우, rs145397166-C인 경우에 비하여 높은 것인,

정보제공 방법.

청구항 12

제1항에 있어서,

상기 예후는 지주막하 출혈 후 인지기능이상 발생 여부에 관한 것인,

정보제공 방법.

청구항 13

rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및

상기 유전자 정보를 하기 수학식 1에 입력하여 인지기능이상의 발생가능성을 산출하는 연산부;를 포함하는,

지주막하 출혈 후 인지기능이상의 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_i는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i는, 상기 Var_i에 부여되는 가중치이고, 상기 PRS는 다중유전자위험점수(polygenic risk score; PRS)이다.

발명의 설명

기 술 분 야

[0001] 본 발명은 유전자 재조합율이 높은 영역을 포함하는 복수개의 SNP를 확인하여 완성된 지주막하 출혈의 예후 예측용 모델에 관한 것이다. 구체적으로는 상기 SNP 유전자 정보를 수신하는 단계; 및 상기 유전자 정보를 하기 수학식 1에 입력하는 단계;를 포함하는, 지주막하 출혈의 예후 예측용 모델 및 이를 이용한 진단 장치에 관한 것이다.

[0002] 본 발명은 대한민국 보건복지부(과제번호: HR21C0198)와 한림연구기금의 지원을 받아 한국보건산업진흥원(KHIDI)을 통한 한국보건기술 R&D 사업의 지원으로 완성되었다.

배 경 기 술

[0003] 맞춤의학(Personalized medicine)은 질병의 위험성에 대해 예상되는 반응에 따라 각각의 환자에 대해 어떠한 약이 효과가 있는지 결정해 처방을 내리는 의학 모형으로, 정밀의학 및 층화의학과 유사한 의미를 갖고 있으며, 유전자 분석 기술의 발전으로 관심이 높아지고 있는 영역이다.

[0004] 한편, 개인별 유전자의 단일염기다형성(Single Nucleotide Polymorphisms; SNPs)이란, DNA 염기서열에서 하나의 염기서열(A,T,G,C)의 차이를 보이는 유전적 변화 또는 변이를 의미하는 것으로, 단일 염기 다형현상은 각 개인마다 많은 변이를 보이는 부분이므로 DNA 지문 분석에 주로 이용된다. SNP는 개인의 현재 특정 질병을 진단하고 합병증의 발생을 예측하며 미래 질병의 발병 가능성을 예측하는 데 이용될 수 있다. 최근 생물정보학 분석의 발달로 많은 양의 임상 및 게놈 데이터를 처리하고 복잡한 표현형을 가진 질병을 연구할 수 있게 되었으며, 뇌혈관질환을 비롯한 복합질환에 대한 유전자 분석이 증가하고 있다.

[0005] 지주막하 출혈(Subarachnoid hemorrhage; SAH)은 두개강내 출혈의 아형(subtype)으로 혈관내의 혈액이 밖으로 유출(extravasation)되면서 지주막하공간(subarachnoid space)으로 들어가서 뇌실질(Brain parenchyma)의 주변부에 위치하는 뇌 조직에 다양한 염증 반응과 뇌세포 손상을 일으키는 질병이다.

[0006] 지주막하 출혈은 뇌 주변을 전반적으로 감싸고 있는 지주막하공간에서 발생하는 출혈이기 때문에 뇌 전반에 걸쳐 손상을 일으키고, 이로 인해 발생 즉시 사망을 일으키거나 심각한 영구적 뇌손상을 초래하게 된다. 지주막하 출혈의 원인으로는 두부의 외부적 충격으로 인한 외상(trauma), 뇌의 혈관벽의 변화로 인해 발생하는 뇌동맥류(aneurysm) 등을 들 수 있다. 또한, 예후 측면에 있어서 현대의학에서조차도 지주막하 출혈은 여전히 사망률이 높고 예후가 나쁜 질환이다. 지주막하 출혈 환자 중 10~15%는 병원에 도착하기도 전에 사망하며, 25%의 환자는 출혈이 발생한지 24시간 이내에 사망하게 된다. 또한, 발병 초기에 사망하지 않고 병원에 입원하더라도 1개월 이내에 평균 40%의 환자가 사망하며, 지주막하 출혈 중 50%의 환자는 6개월 내에 사망하고, 종합적으로 환자 중 51~80%가 결국 사망하게 된다. 생존을 하게 되더라도 2주 이내에 재출혈(rebleeding), 뇌수종(hydrocephalus), 혈관연축(vasospasm)으로 인해 신경학적 합병증을 일으키고 대다수가 영구적인 뇌손상으로 인해 인지능력이 떨어져서 일상 생활이 불가능하고, 평생 침상 생활을 하게 되거나, 마비가 발생하여 걷지 못하게 된다.

[0007] 지주막하 출혈의 진단은 파열된 뇌동맥류에 대하여 CT 혈관조영술 및 뇌혈관조영술을 시행한다. 뇌동맥류 파열에 의한 지주막하 출혈을 진단하기 위해서는 뇌 전산화단층촬영(brain CT)이 가장 먼저 시행하는 검사방법이다.

지주막하 출혈 후 48시간이내에 전산화단층촬영을 하면 95% 이상에서 진단을 할 수 있으며, 만약에 지주막하 출혈이 강력하게 의심되면서 전산화단층촬영상 음성으로 나오면 요추 천자를 시행하게 된다. 다만 이러한 진단 방법은 값비싼 검사 장비가 필요하여, 그 과정 또한 복잡할 수 있어, 신속하게 지주막하 출혈의 상태를 진단하는 데에는 불편하다는 단점이 있다.

[0008] 유전자 분석 결과는 인종에 따라 다를 수 있으며, 한국인에 대한 결과도 동북아 국가인 중국, 일본과 다를 수 있다. 따라서, 본 발명자들은 지주막하 출혈 환자군의 GWAS에서 업데이트된 GWAS를 수행하여 한국인의 지주막하 출혈 이후 인지장애 발생과 연관된 SNP 마커를 기반으로 한 모델을 수립하여 본 발명을 완성하였다.

선행기술문헌

비특허문헌

[0009] (비특허문헌 0001) Yoo SK, Kim CU, Kim HL, Kim S, Shin JY, Kim N, et al. Nard: Whole-genome reference panel of 1779 northeast asians improves imputation accuracy of rare and low-frequency variants. Genome Med. 2019;11:64

(비특허문헌 0002) Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526:68-74

발명의 내용

해결하려는 과제

[0010] 본 발명자들은 국내 지주막하 출혈 후 인지기능이상 변화를 시간이 지남에 따라 추적한 자료를 기반으로 수행한 전장유전체연관성 연구를 통해서 발굴된 SNP 유전자들이 한국인의 지주막하 출혈 이후 인지장애 발생과 연관됨을 확인하여, 본 발명을 완성하였다.

[0011] 이에, 본 발명이 이루고자 하는 기술적 과제는 한국인 및 아시아인을 대상으로 한 유전체기반 SNP 마커를 이용한 예측모델을 제공하는 것이다.

[0012] 그러나 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당해 기술분야의 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 상기 과제를 해결하기 위하여, 본 발명의 일 예시에 따르면, rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 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, 지주막하 출혈의 예후 예측용 정보제공 방법을 제공한다.

[0014] [수학식 1]

[0015] $$PRS = \frac{\sum_{i=1}^{n} w_i Var_i}{n}$$

[0016] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Vari는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 wi는, 상기 Vari에 부여되는 가중치이고, 상기 PRS는 다중유전자위험점수이다.

[0017] 또한, 본 발명에서, 상기 유전자 정보는 아포지단백 E(Apolipoprotein E; APOE), 합토글로빈(Haptoglobin; Hp)으로 이루어진 군에서 선택된 어느 하나 이상의 유전자의 정보를 더 포함하는 것정보제공 방법을 제공할 수 있다.

[0018] 또한, 본 발명에서, 상기 w_i는, 뇌혈관 질환과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포

함하는 데이터베이스에서 산출된 것이고, 하기 수학식 2를 만족하는 것인 정보제공 방법을 제공할 수 있다.

[0019] [수학식 2]

$$w_i = k \log(HR_i)$$

[0020]

[0021] 상기 수학식 2에서, 상기 k는, 0이 아닌 실수이고, 상기 HR_i는, 상기 상관관계의 위험비(Hazard ratio)이다.

[0022] 또한, 본 발명에서, 상기 유전자 정보는 아포지단백 E(Apolipoprotein E; APOE), 합토글로빈(Haptoglobin; Hp)으로 이루어진 군에서 선택된 어느 하나 이상의 유전자의 정보를 더 포함하는 것인, 정보제공 방법을 제공할 수 있다.

[0023] 또한, 본 발명에서, 상기 k는 음수이고, 상기 연산부는 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 지주막하 출혈 후 인지기능 저하 위험이 높은 것으로 판단하는 것일 수 있다.

[0024] 또한, 본 발명에서, 상기 n은 5 이상 100 이하인 것일 수 있다.

[0025] 또한, 본 발명에서, 상기 점수는 상기 유전자형이 rs10503670-A인 경우, rs10503670-G인 경우에 비하여 높은 것일 수 있다.

[0026] 또한, 본 발명에서, 상기 점수는 상기 유전자형이 rs56823384-C인 경우 rs56823384-T인 경우에 비하여 높은 것일 수 있다.

[0027] 또한, 본 발명에서, 상기 점수는 상기 유전자형이 rs76507772-C인 경우, rs76507772-A인 경우에 비하여 높은 것일 수 있다.

[0028] 또한, 본 발명에서, 상기 점수는 상기 유전자형이 rs138753053-A인 경우, rs138753053-G인 경우에 비하여 높은 것일 수 있다.

[0029] 또한, 본 발명에서, 상기 점수는 상기 유전자형이 rs145397166-G인 경우, rs145397166-C인 경우에 비하여 높은 것일 수 있다.

[0030] 또한, 본 발명에서, 상기 예후는 지주막하 출혈 후 인지기능이상 발생 여부에 관한 것일 수 있다.

[0031] 본 발명의 다른 예시에 따르면, rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및 지주막하 출혈 후 인지기능이상의 예측 또는 진단을 위한 장치를 제공한다.

[0032] [수학식 1]

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

[0033]

[0034] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Vari는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 wi는, 상기 Vari에 부여되는 가중치이고, 상기 PRS는 다중유전자위험점수(polygenic risk score; PRS)이다.

발명의 효과

[0035] 본 발명은 유전자형 교정 및 결측값 대체를 통해, 한국인을 대상으로 하는 업데이트된 GWAS 유전정보 데이터를 기반으로 지주막하 출혈 후 인지기능이상 발생 위험도가 높은 예측 모형을 제공할 수 있다. 그 결과, 지주막하 출혈 발생 후 예측된 예후에 따라 인지기능이상에 따른 빠른 대처가 가능하도록 하는 장점이 있다.

도면의 간단한 설명

[0036] 도 1은 LINC00676-IRS2 영역(13q34.3 chr13:109882882-110882882; rs76507772 SNP ± 500kb)에서 출혈성 출혈

환자의 인지장애와 관련된 유전자 지역 연관성 플롯을 확인한 결과를 나타낸 것이다.

도 2는 rs76507772 SNP의 유전자형 차이를 확인하기 위해 Kaplan-Meier 생존 곡선을 확인한 결과를 나타낸 것이다.

도 3은 PDCD6IP-LOC101928135 유전자좌의 ± 500kb의 지역 연관성 플롯을 확인한 것이다.

도 4는 LINC00499 유전자좌의 ± 500kb의 지역 연관성 플롯을 확인한 것이다.

도 5는 CASC15 유전자좌의 ± 500kb의 지역 연관성 플롯을 확인한 것이다.

도 6은 LPL-SLC18A1 유전자좌의 ± 500kb의 지역 연관성 플롯을 확인한 것이다.

도 7은 지주막하 출혈(SAH) 후 발달 인지장애(CI)에 따른 가중 다유전자 위험 점수(wPRS) 모델에 대한 수신기 작동 곡선(AUROC) 아래 영역을 확인한 것이다.

도 8은 Kaplan-Meier 생존 곡선은 추적 기간 동안 3개의 다유전자 위험 그룹을 기반으로 한 인지장애 없는 생존 확률을 확인한 것이다.

도 9는 APOE ε3/ε3 지주막하 출혈 환자의 가중 플로발생 위험 점수 모델을 기반으로 한 인지 손상 없는 Kaplan-Meier 생존 곡선을 확인한 것이다.

도 10은 APOE ε4 보유자인 지주막하 출혈 환자의 가중 플로발생 위험 점수 모델을 기반으로 한 인지 손상 없는 Kaplan-Meier 생존 곡선을 확인한 것이다.

도 11은 Hp2-1을 갖는 지주막하 출혈 환자의 가중 플로발생 위험 점수 모델을 기반으로 한 인지 손상 없는 Kaplan-Meier 생존 곡선을 확인한 것이다.

도 12는 Hp2-2를 갖는 지주막하 출혈 환자의 가중 플로발생 위험 점수 모델을 기반으로 한 인지 손상 없는 Kaplan-Meier 생존 곡선을 확인한 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, "지주막하 출혈(subarachnoid hemorrhage; SAH)" 은 뇌의 지주막 아래 공간에 뇌출혈이 일어나는 질환을 의미하는 것으로서, 뇌동맥류 파열과 같은 심각한 원인이 있을 수 있고, 이 외에도 뇌혈관의 기형이나 외상 등에 의해서 지주막하 공간에 출혈이 발생하는 모든 경우를 의미하는 것일 수 있고, 상기 "지주막하 출혈"은 "SAH"와 혼용되어 명명될 수 있다.

[0041] 본 발명에서, “진단”은 질병 발생의 예측 및 발병 위험도(risk) 또는 발병 감수성(susceptibility)를 결정하거나 도출하는데 사용되는 모든 유형의 분석을 포함한다. 일례로, 지주막하 출혈에 대한 예후를 예측하기 위한 것일 수 있고, 특히 지주막하 출혈 후 인지기능이상 변화를 예측하기 위한 것일 수 있다.

[0042] 본 발명에서, "인지기능"은 지각, 사고, 추리, 기억 등 뇌의 사고처리능력을 의미하고, 상기 인지기능이 저하 내지 이상 발생한 상태를 "인지기능이상(cognitive dysfunction)", "인지장애(cognitive impairment)"라고 하며, 본 발명에서는 지주막하 출혈 등으로 인한 뇌 손상에 의해 발생한 것을 의미할 수 있고, 인지기능을 평가하기 위해 간이정신상태검사(Mini-Mental State Examination; MMSE)가 역학적 연구 및 임상 분야에서 세계적으로 널리 이용되고 있다.

[0043] 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.

[0044] 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.

[0045] 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.

[0046] 본 발명에서, 상기 장치는 프로세서를 포함할 수 있고, 상기 프로세서는 머신 러닝 모델 훈련 및 바이오 마커에 대한 정보를 분석하는 기능을 수행할 수 있다.

[0047] 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.

[0049] 본 발명의 일 구현 예에 따르면, 지주막하 출혈 후 인지기능이상을 예측하기 위한 SNP 마커로서 rs145397166, rs10503670, rs56823384, rs76507772, rs138753053에서 이루어진 군으로부터 선택된 어느 하나 이상을 더 포함할 수 있고, 추가로 SNP 상기 마커를 모두 포함하는 것일 수 있다.

[0050] 본 발명의 상기 바이오마커의 구체적인 정보는 하기 표 1에 나타내었다.

표 1

[0051]

서열번호	유전자	염색체	SNP	좌위	서열(염기서열위치[ ] 표기)
1	LPL, SLC18A1	8p21.3	rs10503670	19985382	TTTTTAGCAAATTAAATTTTCTAATTTGCTTTTACTCAGTACATTGATAT GCTTAGTATAAATGCACTAACTGTGCTACTTTTAAACGCATCCAACACAC [G/A] TCATTTTAAAGTGTTTTATGAGAAATTCTGAACCTCCAGTAGACAACGAC CCTTCTCCTACAAACTACAGTTATTCCAAGGGCATTTCACTCTGTCTTCC
2	LINC00499	4q28.3	rs56823384	139255670	ATACATGAAGACTTATTTCAAGAATTTAAATACTTTTTAAAGAGTTAACC ATAAATACAGAGTCCTGGCATGTTCCACCCATACCTGACGCCAGGAATTA [T/C] GAAATAGGAAGCTGGCAGTGAGTTACGGTGCCTACTCTGGTTGTAAAGTA GTTTACTTTCTGCCTGAGTAAAGCATTAAGACGTGCGTCAGGTATCGTTA

3	IRS2	13q34	rs76507772	110382882	GCCCGTGTCACTCTCAGACTCCATGTCTCAATTTGAATGTTTCCATGTCA CTTAAAAAGTAGAATATTCATAGTGATGTTGAAGATACAAAAGATATACC [A/C] GGATTAATTTTCTAGTCAATTTTATGTGTCTGTGACCCAAGACAATAAAA GTTAAGTCATCGTTAGTTCCTTAGTTGCTGGACACTTTAGCCCAATATAC
4	PDCD6IP, LOC101928135	3p22.3	rs138753053	33961435	GTTGACCCCCACTGGCTGGGTACCACAAGTTTTCTGTTTCTGAGCTCTTA GCCAGGGTGGAGCCCTACCCTTCCAGAGATGTGCCTTTGTGCTTAAGCTG [G/A] CAAGAATGCAATCAGCCTAGTTAAAATTTTGATAATCCACTTGATCCATC CCAGTTATAGGCTCAAATCCCAGGAGGCATCTTATAGACCTGAGCCTTAT
5	CASC15	6p22.3	rs145397166	22067476	CAAGATTGAGAAAAACATGTCTTTAGGCAGGCTAATCTAATCACACACAG AAGGCAGCCTGTTTTCTTACTTGCTGGATCGCTTCCACTTTATATTCTCA [C/G] AGGGAACAATATATTAAGTACAGTATTTAGAAATATTTGTTGAATCCTAC TGTCTGATTTTGGCAAGTCCCAAGGAAAAAAAGAATATGTCTTGGACTGC

[0052] 한편, 일배체형(Haplotype) 분석은 SNP와 표현형과의 연관성을 보다 강하게 보장한다. 같은 블록 내의 유전자형은 함께 유전되는 경향이 있기 때문에 LD가 발생하는데 일배체형은 상대적으로 유전되는 인접 유전자좌들의 조합을 포함하여 LD 값이 높을 수록 질병 연관 가능성이 높음을 의미한다.

[0053] 본 발명은 종래의 한국인 대상의 GWAS 결과에서 동북아 인구를 기반으로 한 대체(imputation)을 추가하여 통계적 검정력을 향상시켰다.

[0054] 본 명세서에서 "다형성"이라는 용어는 군집 내에서 변하는 유전자의 서열에서의 배치를 지칭한다. 다형성은 상이한 "대립유전자"로 구성된다. 이러한 다형성의 배치는 유전자에서의 그의 위치 및 그에서 발견되는 상이한 아미노산 또는 염기에 의해 확인될 수 있다. 이러한 아미노산 변이는 2개의 상이한 대립유전자인, 2개의 가능한 변이체 염기, C 및 T의 결과이다. 유전자형은 2개의 다른 별개의 대립유전자로 구성되기 때문에, 여러 가능한 변이체 중 임의의 변이체가 어느 한 개체에서 관찰될 수 있다(예를 들어, 이 예에서, CC, CT 또는 TT). 개개의 다형성은 또한 당업자에게 공지되어 있고, 예를 들면, NCBI 웹사이트 상에서 이용가능한 뉴클레오타이드 염기 변이의 단일 뉴클레오타이드 다형성 데이터베이스(Single Nucleotide Polymorphism Database(dbSNP) of Nucleotide Sequence Variation)에서 사용되는 것인, 지정된 독특한 식별자("기준 SNP", "refSNP" 또는 "rs #")일 수 있다.

[0055] 본 명세서에서 "유전적 다형성(genetic polymorphism)"은 하나의 유전자 좌위(locus)에 두가지 이상의 대립 유전자(allele)가 존재하는 경우를 말하며, 인구집단에서 적어도 1% 이상의 빈도로 유전자 변이가 나타나는 경우를 말한다. DNA에 서 한 개의 뉴클레오타이드의 삽입, 소실, 또는 치환이 일어나는 것을 단일염기다형성(single nucleotide polymorphism, SNP) 이라고 한다.

[0056] 본 명세서에서 "단일염기다형성(single nucleotide polymorphism, SNP)"은 게놈에서 단일염기(A, T, C 또는 G)가 종의 멤버들 간 또는 한 개체(individual)의 쌍 염색체 간에 다른 경우에 발생하는 DNA 서열의 다양성을 의미한다. 예를 들어, 서로 다른 개체의 세 개의 DNA 단편들(예: AAGT[A/A]AG, AAGT[A/G]AG,AAGT[G/G]AG)처럼 단일염기에서 차이를 포함하는 경우, 두 개의 대립 유전자(C 또는 T)라고 부르며, 일반적으로 거의 모든 SNPs는 두 개의 대립 유전자를 가진다. 한 집단(population)내에서, SNP는 소수 대립인자 빈도(minor allele frequency, MAF; 특정 집단에서 발견되는 유전자위치(locus)에서 가장 낮은 대립인자 빈도)로 할당될 수 있다. 단일염기는 폴리뉴클레오타이드 서열에 변화(대체), 제거(결실) 또는 첨가(삽입)될 수 있다. SNP는 번역 프레임

의 변화를 유발할 수 있다.

[0057] 한편 본 명세서에서 유전자의 변이를 표기할 때 [염기일문자/염기일문자]는 좌측에 쓰인 염기일문자가 우측에 쓰인 염기일문자로 치환되는 것을 의미한다.

[0058] 본 명세서에서 “폴리뉴클레오티드(polynucleotide)” 또는 “핵산”은 단일-또는 이중-가닥의 형태로 된 데옥시리보뉴클레오티드(DNA) 또는 리보뉴클레오티드(RNA)를 말한다. 다른 제한이 없는한, 자연적으로 생성되는 뉴클레오티드와 비슷한 방법으로 핵산에 혼성화되는 자연적 뉴클레오티드의 공지된 아날로그도 포함된다. 일반적으로 DNA는 아데닌(adenine, A), 구아닌(guanine, G), 사이토신(cytosine, C), 티민(thymine, T) 등 네 가지 염기로 구성되어 있으며, RNA는 티민 대신 우라실(Uracil, U)을 가지고 있다. 핵산 이중 가닥에서 A는 T 또는 U, C는 G 염기와 수소결합을 이루는데, 이러한 염기의 관계를 '상보적(complementary)'이라고 한다. 한편, 본 명세서에서 '폴리뉴클레오티드'는 다른 언급이 없으면 본 발명의 SNP가 위치하는 유전체 영역에서 상기 SNP를 포함하는 8~100 개의 연속된 염기서열을 갖는 폴리뉴클레오티드를 의미한다.

[0059] 본 명세서에서 “게놈 DNA(genomic DNA: gDNA)”는 한 개체의 유전자의 총 염기서열로서, 거의 완전한 유전정보를 포함하고 있는 DNA를 의미한다.

[0061] 본 발명의 다른 구현 예에 따르면, rs145397166, rs10503670, rs56823384, rs76507772, rs138753053에서 이루어진 군으로부터 선택된 어느 하나 이상을 더 포함하는, 지주막하 출혈의 예후 예측용 조성물을 제공할 수 있다.

[0062] 본 발명의 상기 조성물은 한국인 또는 아시아인을 대상으로 지주막하 출혈 후 예후 예측용인 것일 수 있다.

[0063] 본 발명에서, 지주막하 출혈 후 인지기능이상을 예측하기 위한 조성물로서 rs145397166, rs10503670, rs56823384, rs76507772, rs138753053에서 이루어진 군으로부터 선택된 어느 하나 이상을 더 포함하는 것일 수 있고, 추가로 SNP의 유전자형을 검출하는 조성물을 모두 포함하는 것일 수 있다.

[0065] 본 발명의 또 다른 일 구현 예에 따르면, 상기 SNP 중에서 선택되는 1종 이상의 SNP를 검출할 수 있는 제제를 포함하는 지주막하 출혈의 예후 예측용 키트를 제공한다.

[0066] 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, 데옥시뉴클레오티드(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.

[0067] 본 발명에서 상기 키트는 RT-PCR 키트 또는 DNA 칩 키트일 수 있으나, DNA를 분석하기 위한 것이라면 제한없이 포함될 수 있다.

[0068] 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 뉴클레오티드로써, 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), 데옥시뉴클레오타이드(dNTPs), Taq-폴리머라아제 및 역전사효소와 같은 효소, DNase, RNase 억제제 DEPC-수(DEPC-water), 멸균수 등을 포함할 수 있다.

[0069] 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 올리고뉴클레오티드(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 올리고뉴클레오티드를 포함할 수 있다.

[0070] 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.

[0071] 또한, 본 발명의 상기 키트에서 2차 항체의 표지체는 발색 반응을 하는 통상의 발색제가 바람직하며, HRP(horseradish peroxidase), 염기성 탈인산화효소(alkaline phosphatase), 콜로이드 골드(coloid 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.

[0072] 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-페닐렌다이아민) 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.

[0073] 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(H_2SO_4)이 바람직하게 사용될 수 있다.

[0075] 본 발명의 또 다른 구현 예에 따르면, rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 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, 지주막하 출혈의 예후 예측용 정보제공 방법인 것일 수 있다.

[0076] [수학식 1]

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

[0077]

[0078] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i는, 상기 Var_i에 부여되는 가중치이다. 또한, 상기 PRS는 다중유전자위험점수(polygenic risk score; PRS) 내지 가중치가 부여된 다중유전자위험점수(weighted polygenic risk score; wPRS)인 것일 수 있다.

[0079] 본 발명에서, 상기 유전자 정보는 아포지단백 E(Apolipoprotein E; APOE), 합토글로빈(Haptoglobin; Hp)으로 이루어진 군에서 선택된 어느 하나 이상의 유전자의 정보를 더 포함하는 것일 수 있다.

[0080] 본 발명에서, 상기 w_i는, 뇌혈관 질환과 상기 i번째 단일염기다형성(SNP)의 유전자형 간의 상관관계를 포함하는 데이터베이스에서 산출된 것이고, 하기 수학식 2를 만족하는 것일 수 있다:

[0081] [수학식 2]

$$w_i = k \log(HR_i)$$

[0082]

[0083] 상기 수학식 2에서, 상기 k는, 0이 아닌 실수이고, 상기 HR_i는, 상기 상관관계의 위험비(Hazard ratio)이다.

[0084] 본 발명에서, 상기 k는 음수이고, 상기 연산부는 상기 PRS가 대조군에 비해 높을 경우, 상기 유전자 정보를 가진 개체의 지주막하 출혈 후 인지기능 저하 위험이 높은 것으로 판단하는 것일 수 있다.

[0085] 본 발명의 방법에 사용되는 단일염기다형성(SNP) 유전자의 개수는 추가될 수 있고, 상기 n은 5 이상 100 이하인 것일 수 있다.

[0086] 본 발명에서, 상기 점수는 상기 유전자형이 rs10503670-A인 경우, rs10503670-G인 경우에 비하여 높은 것일 수 있다.

[0087] 본 발명에서, 상기 점수는 상기 유전자형이 rs56823384-C인 경우 rs56823384-T인 경우에 비하여 높은 것일 수 있다.

[0088] 본 발명에서, 상기 점수는 상기 유전자형이 rs76507772-C인 경우, rs76507772-A인 경우에 비하여 높은 것일 수 있다.

[0089] 본 발명에서, 상기 점수는 상기 유전자형이 rs138753053-A인 경우, rs138753053-G인 경우에 비하여 높은 것일 수 있다.

[0090] 본 발명에서, 상기 점수는 상기 유전자형이 rs145397166-G인 경우, rs145397166-C인 경우에 비하여 높은 것일 수 있다.

[0091] 본 발명에서, 상기 예후는 지주막하 출혈 후 인지기능이상을 예측하기 위한 것일 수 있다.

[0092] 본 발명에서, 개체의 생물학적 시료를 준비하는 단계; 상기 생물학적 시료로부터 게놈(genomic) DNA를 얻는 단계; 상기 게놈 DNA에서 GenBank SNP database에서 rs145397166의 단일염기다형성 위치의 염기 타입을 검출하는 단계; 및 상기 게놈 DNA에서 GenBank SNP database 단일염기다형성(single nucleotide polymorphism; SNP) 위치의 염기 타입을 검출하는 단계; 및 상기 검출한 염기 타입으로부터 개체의 지주막하 출혈에 대한 예후를 결정하는 단계;를 더 포함할 수 있다.

[0093] 본 명세서에서 "개체"란 지주막하 출혈 후 예후 예측 등의 진단이 필요한 포유류라면 제한되지 않는다.

[0094] 한편, 본 발명에서 염기 타입을 검출하는 단계는 당업계에 염기서열을 분석하고 결정하기 위하여 통상적으로 사용되는 것이라면 제한없이 사용할 수 있다. 보다 구체적으로는 시퀀싱(sequencing), 엑솜 시퀀싱(exome sequencing), 마이크로어레이에 의한 혼성화(microarray hybridization), 대립유전자 특이적인 PCR(allele specific PCR), 다이나믹 대립유전자 혼성화 기법(dynamic allele-specific hybridization, DASH), 형광 융해곡선 기법(Fluorescence Melting Curve Analysis, FMCA) 및 Taqman 기법으로 이루어진 군으로부터 선택되는 어느 한 방법으로 수행할 수 있다.

[0095] 본 발명에서 "SNP를 검출할 수 있는 제제"란 상기 SNP가 위치하는 유전체 영역에서 상기 SNP를 포함하는 8개 내지 100개의 연속된 염기서열을 갖는 폴리뉴클레오티드를 특이적으로 증폭할 수 있는 프라이머 또는 상기 폴리뉴클레오티드와 특이적으로 결합할 수 있는 프로브를 의미한다. 상기 SNP를 포함하는 폴리뉴클레오티드 증폭에 사용되는 프라이머는, 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오시드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성(template-directed DNA synthesis)의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오타이드가 될 수 있는데, 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오타이드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 상기 프라이머 서열은 상기 SNP 영역의 폴리뉴클레오티드와 완전하게 상보적일 필요는 없으나, 상기 폴리뉴클레오티드와 혼성화할 정도로 충분히 상보적이어야 한다.

[0096] 본 명세서에서 "프라이머"란, 짧은 자유 3' 말단 수산화기(free 3' hydroxyl group)를 가지는 염기 서열로 상보적인 주형(template)과 염기쌍(base pair)을 형성할 수 있고 주형 가닥 복사를 위한 시작 지점으로 기능을 하는 짧은 서열을 의미한다. 프라이머는 적절한 완충용액 및 온도에서 중합반응(즉, DNA 폴리머레이즈 또는 역전사효소)을 위한 시약 및 상이한 4가지 뉴클레오사이드 트리포스페이트의 존재 하에서 DNA 합성을 개시할 수 있다. PCR 증폭을 실시하여 원하는 생성물의 생성 여부를 통해 아토피 피부염을 예측할 수 있다. PCR 조건, 센스 및 안티센스 프라이머의 길이는 당업계에 공지된 것을 기초로 변형할 수 있다.

[0097] 본 발명의 프라이머는 포스포르아미다이트(phosphoramidite) 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 비-제한적인 예로는 메틸화, "캡화", 천연 뉴클레오타이드 하나 이상의 동족체로의 치환, 및 뉴클레오타이드간의 변형, 예를 들면, 하전되지 않은 연결체(예: 메틸포스포네이트, 포스포트리에스테르, 포스포로아미데이트, 카바메이트 등) 또는 하전된 연결체(예: 포스포로티오에이트, 포스포로디티오에이트 등)로의 변형이 있을 수 있다.

[0098] 본 발명에서 상기 시료를 준비하는 단계의 개체는 지주막하 출혈 후 예후 예측이 요구되는 인간이라면 제한되지 아니하나, 바람직하게는 한국인 또는 아시아인일 수 있다.

[0100] 본 발명의 또 다른 구현 예에 따르면, rs76507772, rs56823384, rs76507772, rs138753053 및 rs145397166으로 이루어진 군에서 선택된 어느 2 이상을 포함하는 복수개의 단일염기다형성(SNP) 유전자 정보를 수신하는 수신부; 및 상기 유전자 정보를 하기 수학식 1에 입력하여 인지기능이상의 발생가능성을 산출하는 연산부;를 포함하는, 지주막하 출혈 후 인지기능이상의 예측 또는 진단을 위한 장치인 것일 수 있다:

[0101] [수학식 1]

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

[0102]

[0103] 상기 수학식 1에서, 상기 n은, 상기 유전자 정보가 포함하는 단일염기다형성(SNP)의 개수이고, 상기 Var_i는, 상기 유전자 정보가 포함하는 i번째 단일염기다형성(SNP)의 유전자형에 따른 점수이고, 상기 w_i는, 상기 Var_i에 부여되는 가중치이다. 또한, 상기 PRS는 다중유전자위험점수(polygenic risk score; PRS) 내지 가중치가 부여된 다중유전자위험점수(weighted polygenic risk score; wPRS)인 것일 수 있다.

[0104] 본 발명에서, 목적하는 개체로부터 분리된 생물학적 시료로부터 유전자 서열을 검출하는 검출부를 더 포함하는 것일 수 있다.

[0106] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 이하 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0107] [실험 방법]

[0108] 1. 연구 대상 특정

[0109] 본 연구는 2016년 3월부터 2020년 6월까지 진행된 "제1회 대한뇌졸중 유전학회 연구" 연구 데이터베이스의 지주막하 출혈(subarachnoid hemorrhage; SAH) 환자를 대상으로 하였다. 상기 SAH 데이터베이스에서 1) 18세 이상의 성인 환자, 2) 동맥류 파열로 인한 SAH 환자, 3) GWAS 결과 및 인지 테스트를 받은 환자, 4) 발작 후 최소 6개월의 추적 관찰된 환자를 포함하였고, 1) 외상 또는 감염으로 인한 SAH 환자, 2) GWAS 결과가 없는 환자, 3) 방사선 소견 및 인지 검사를 포함한 기록이 불충분한 환자, 4) 추적 관찰기록이 소실된 환자, 5) 인지기능 검사를 수행할 수 없는 환자를 제외하였다. 본 연구는 기관심사위원회 및 윤리위원회의 승인을 받았고(No. 2016-3 및 2019-06-006), 환자 또는 그 친척으로부터 사전 동의를 받았다.

[0111] 2. 연구 수행 단계

[0112] 먼저, SAH 후 종단 GWAS를 사용하여 인지장애와 관련된 변이체를 식별하였다. 다음으로, APOE 및 합토글로빈(Hp)과 같은 알려진 유전적 위험 인자에 따른 유전적 변이체를 비교하였다. 또한, GWAS에 의해 유도된 감수성

유전자좌를 기반으로 인지장애를 예측하기 위해 가중 다유전자 위험 점수(wPRS) 평가를 수행하였다. 인지기능 검사는 한국어판 Mini-mental State Examination(K-MMSE)을 사용하여 진행되었다. 검사는 SAH 발생 후 6개월 후에 일상적으로 수행되었으며 그 이후에는 매년 정기적으로 수행되었는데, MMSE 점수가 27 미만이면 인지장애로 정의했다. APOE 일배체형(haplotypes, ε2: 388T-526T, ε3: 388T-526C, ε4: 388C-526C)은 현재 GWAS의 rs429358 (388T>C) 및 rs7412 (526C>T) 변이에 의해 평가되었다. Hp 표현형은 웨스턴 블롯 분석을 사용하여 분류되었다.

[0114] 3. 유전자형 및 품질 관리

[0115] AxiomTH Asian Precision Medicine Research Array(APMRA)(Thermo Fisher Scientific, MA, USA)에 의해 유전형이 분석된 SAH 환자의 말초 혈액에서 유래한 게놈 DNA를 사용하였다. 고품질 플레이트는 샘플에 대해 > 95%의 플레이트 통과율과 > 99%의 통과된 샘플의 평균 호출율을 보였다. 802,688개의 SNP 중 512,503개의 SNP가 유전자형 호출 비율 ≥ 95%, 소수 대립 유전자 빈도 ≥ 1%, 하디-바인베르크 평형 P 값 ≥ 1×10^{-6}과 같은 품질 관리 필터를 통과했다.

[0117] 4. 고처리량 결측값 대체 및 품질 관리

[0118] Eagle v2.4 및 Minimac4 프로그램을 사용하여 개별 SNP의 유전자형을 사전 단계화하고 490,038개의 SNP가 있는 153명의 대상에 대해 대량의 SNP 및 누락된 유전자형 값을 대체하였다. NIH(National Institutes of Health) National Heart, Lung, and Blood Institute(NHLBI)의 지원을 받은 GenomeAsia 100K 프로젝트에서 생성한 아시아인 관련 참조 패널(GRCh37hg19)을 적용했다. 고처리량 결측값 대체를 위한 서버는 Michigan Imputation Server v1.5.7(https://imputationserver.sph.umich.edu/index.html#!run/minimac4)에서 수행되었다. 단일형 영역을 포함한 292,174,934개 부위 중 21,494,648개 부위가 SAH 환자 153명에서 대체 점수 R^2 임계값 0.5 이상으로 남아 있었다. 마지막으로, 5,971,372개의 SNP가 이 종단 GWAS에서 분석되었으며, 이는 유전자형 호출률 ≥ 95%, 소수 대립 유전자 빈도 ≥ 1%, 하디-바인베르크 평형 P 값 ≥ 0.001로 품질 관리 테스트를 통과했다.

[0120] 5. 통계 분석

[0121] 이산 변수는 피험자 수와 백분율로 표시되었다. 연속변수는 평균과 표준편차(SD)로 나타내었다. STATA 소프트웨어 v17.0(Stata Corp., College Station, TX, USA). 주성분 분석(PCA)은 153명의 SAH 환자와 2,504명의 1000-게놈(1KG) 참조 패널(3상, 버전 5)을 기반으로 한 샘플 이질성 및 분포를 평가하기 위해 수행되었다. PCA 값은 4개의 다차원 척도(MDS)를 사용하여 변환되었다. GWAS 기반 CPH 회귀 분석은 성별, 연령 및 4가지 유전적 가계 MDS를 조정한 후 추적 기간 동안 인지장애를 기반으로 수행되었다. 추가 GWAS 하위 그룹 분석은 APOE 대립 유전자 및 Hp 표현형을 기반으로 수행되었으며, 이는 SAH 이후 인지장애에 대한 알려진 위험 요소이다. 이 종단 GWAS의 데이터는 "coxph" 기능으로 구현된 R 패키지 소프트웨어의 "생존" 구성 요소에서 분석되었다(https://cran.r-project.org/web/packages/survival/index.html). 맨해튼과 지역 협회는 R 패키지 v3.6.2(https://cran.r-project.org/web/packages/qqman)의 "qqman"과 수정된 Python 및 R 스크립트(http://cran.r-project.org/web/packages/qqman)와 LocusZoom v1.3으로 작성되었다(https://locuszoom.org/).

[0122] 마지막으로, wPRS는 SAH 후 인지장애를 예측하는 데 이 위험 점수 시스템이 가능한지 여부를 결정하기 위해 GWAS 기반 유전자좌를 기반으로 평가되었다. 각 SNP의 wPRS 값은 로그 변환된 HR에 위험 대립 유전자의 수(즉, 0, 1 또는 2개 사본)를 곱하여 생성되었다. 최종 wPRS 모델은 GWA 신호에 표시된 개별 SNP의 가중 유전자형을 합산하여 구성되었다($p < 5x10-8$). 모델의 예측 가능성은 AUROC(수신기 작동 곡선 아래 영역)로 평가되었다. 또한 tertiles(T1, 낮은 위험, T2, 중간 위험, T3, 높은 위험)로 계층화한 후 위험 그룹 간의 추가 연관성을 테스트하였다.

[0123] [실험 결과]

[0124] 1. IRS2 유전자 지역 연관성 플롯 확인

[0125] LINC00676-IRS2 영역(13q34.3 chr13:109882882-110882882; rs76507772 SNP ± 500kb)에서 출혈성 출혈 환자의

인지장애와 관련된 유전자 지역 연관성 플롯을 확인하여 도 1 및 하기 표 2에 나타내었다. 이때, 삼각형은 양수 효과 크기를, 역삼각형은 음수 효과 크기를 나타내며, 각 색상은 쌍별 연결 불균형을 보여준다. 구체적으로, 보라색 위쪽 삼각형은 rs76507772의 상위 SNP를 나타내고, 다른 색상은 쌍별 연결 불균형을 나타낸다(LD, r2): 네이비, 0-0.2; 녹색, 0.2-0.4; 하늘색, 0.4-0.6; 주황색, 0.6-0.8; 빨간색, 0.8-1.

표 2

[0127]

서열번호	유전자	염색체	SNP	좌위	P 값	가설
1	IRS2	13q34	rs76507772	110382882	3.5×10^{-8}	Haplotype 구조 및 변이간 연관불균형이 강한 지역으로서 (Tagging SNP = rs76507772), 대사질환관련 유전자 IRS2 근처에 위치하여 인지장애 발생에 영향을 줄 가능성이 높음

[0128] 그 결과, 일배체형(Haplotype) 구조 및 변이간 연관불균형이 강한 지역으로서 (Tagging SNP = rs76507772)의 경우, 대사질환관련 유전자 IRS2 근처에 위치하여 인지장애 발생에 영향을 줄 가능성이 높은 것을 확인하였다.

[0130] 2. rs76507772 SNP 유전자형에 따른 Kaplan-Meier 생존 곡선

[0131] SAH 후 인지장애에서 rs76507772 SNP의 유전자형 차이를 확인하기 위해 Kaplan-Meier 생존 곡선을 확인하여, 도 2와 같이 나타내었다. 구체적으로는 위험도가 낮은 동형 접합 유전자형을 가진 환자의 생존확률을 검은색 실선으로 나타내었고, 이형 유전자형을 가진 환자의 생존확률을 주황색 긴 대시 선으로 나타내었으며, 유전자형에 따른 인지장애가 없는 생존 확률의 차이를 P 값으로 나타내었다.

[0132] 그 결과, 도 2에 나타난 것처럼, rs76507772 SNP가 이형 유전자형을 가질 경우, 환자의 생존확률이 낮은 것을 확인하였다.

[0134] 3. LOC101928135, LINC00499, CASC15, LPL-SLC18A1 유전자좌 지역 연관성 플롯 확인

[0135] PDCD6IP-LOC101928135, LINC00499, CASC15 및 LPL-SLC18A1 4개의 유전자좌의 ± 500kb의 지역 연관성 플롯을 확인하여 도 3, 도 4, 도 5, 도 6 및 하기 표 3에 나타내었다. 이때, 삼각형은 양수 효과 크기를, 역삼각형은 음수 효과 크기를 나타내며, 각 색상은 쌍별 연결 불균형을 보여준다. 구체적으로, 보라색 위쪽 삼각형은 각 지역의 상위 SNP를 나타내고, 다른 색상은 쌍별 연결 불균형을 나타낸다(LD, r2): 네이비, 0-0.2; 녹색, 0.2-0.4; 하늘색, 0.4-0.6; 주황색, 0.6-0.8; 빨간색, 0.8-1.

표 3

[0137]

서열번호	유전자	염색체	SNP	좌위	P 값	가설
2	LINC00499	4q28.3	rs56823384	139255670	2.8×10^{-9}	ncRNA site
3	IRS2	13q34	rs76507772	110382882	3.5×10^{-8}	Haplotype 구조 및 변이간 연관불균형이 강한 지역으로서 (Tagging SNP = rs76507772), 대사질환관련 유전자 IRS2 근처에 위치하여 인지장애 발생에 영향을 줄 가능성이 높음
4	PDCD6IP, LOC101928135	3p22.3	rs138753053	33961435	3.4×10^{-8}	인지장애 발생에 절대 위험도(Hazard ratio, R)가 가장 높은 site (HR=28.33)
5	CASC15	6p22.3	rs145397166	22067476	1.7×10^{-8}	유전자 재조합율이 높은 영역

[0138] 그 결과, 각 유전자좌의 지역에서 rs138753053(PDCD6IP-LOC101928135), rs56823384(LINC00499), rs145397166(CASC15) 및 rs10503670(LPL-SLC18A1)의 SNP에서 SAH 후 인지장애 발생에 영향을 줄 가능성이 높은 것을 확인하였다.

[0140] 4. wPRS 모델에 대한 AUROC 및 Kaplan-Meier 생존 곡선 확인

[0141] 지주막하 출혈(SAH) 후 발달 인지장애(CI)에 따른 가중 다유전자 위험 점수(wPRS) 모델에 대한 수신기 작동 곡선(AUROC) 아래 영역을 확인하여 도 7에 나타내었다. 여기서, 빨간색 실선 및 주황색 점선은 wPRS 및 해당 모델의 정확도를 각각 저위험(T1), 중간 위험(T2) 및 고위험(T3) 그룹과 같은 삼분위수로 계층화하였다. 또한, Kaplan-Meier 생존 곡선은 추적 기간 동안 3개의 다유전자 위험 그룹을 기반으로 한 인지장애 없는 생존 확률을 확인하여 도 8 및 하기 표 4에 나타내었다. 또한, APOE ε3/ε3, APOE ε4 보유자, 합토글로빈 2-1, 합토글로빈 2-2를 갖는 지주막하 출혈 환자의 가중 플로발생 위험 점수 모델을 기반으로 한 인지 손상 없는 Kaplan-Meier 생존 곡선을 도 9, 도 10, 도 11, 도 12, 표 5, 표 6, 표 7 및 표 8에 나타내었다. 여기서, 환자는 인지장애의 위험을 낮음(T1), 중간(T2), 높음(T3) 3분위로 계층화되었고, CI은 신뢰구간, AUROC는 ROC 아래 영역을, 검정색 실선은 저위험을, 주황색 점선은 중위험을, 빨간색 점선은 고위험을 나타내는 것이다.

표 4

[0143]

All subjects 모델	인지장애 (n=65)	비인지장애 (n=88)	위험률 (95% CI)	P 값	AUROC	민감도	특이도
T1: < 1.06	26 (40.0%)	75 (85.2%)	Reference				
T2: 1.06-2.12	16 (24.6%)	9 (10.2%)	5.81 (2.93-11.48)	4.4×10^{-7}		0.6	0.852
T3: 2.12 <	23 (35.4%)	6 (4.6%)	18.22 (9.12-36.41)	2.1×10^{-16}	0.739	0.354	0.955

표 5

[0144]

APOE ε3/ε3 모델	인지장애 (n=40)	비인지장애 (n=62)	위험률 (95% CI)	P 값	AUROC	민감도	특이도
T1: < 1.06	12 (30.0)	55 (88.7)	Reference				
T2: 1.06-2.12	10 (25.0)	3 (4.8)	12.67 (4.77-33.70)	3.6×10^{-7}		0.7	0.887
T3: 2.12 <	18 (45.0)	4 (6.5)	26.32 (9.95-69.67)	4.5×10^{-11}	0.796	0.45	0.936

표 6

[0145]

APOE ε4 carrier 모델	인지장애 (n=18)	비인지장애 (n=15)	위험률 (95% CI)	P 값	AUROC	민감도	특이도
T1: < 1.06	12 (66.6)	12 (80.0)	Reference				
T2: 1.06-2.12	3 (16.7)	13 (20.0)	1.50 (0.39-5.75)	0.55		0.333	0.8
T3: 2.12 <	3 (16.7)	0 (0.0)	18.02 (3.03-107.33)	0.002	0.583	0.167	1

표 7

[0146]

Hp2-1 모델	인지장애 (n=18)	비인지장애 (n=39)	위험률 (95% CI)	P 값	AUROC	민감도	특이도
T1: < 1.06	6 (33.3)	34 (87.2)	Reference				
T2: 1.06-2.12	6 (33.3)	3 (7.7)	6.36 (1.63-24.60)	0.007		0.667	0.872
T3: 2.12 <	6 (33.3)	2 (5.1)	44.59 (8.61-231.08)	6.1×10^{-6}	0.774	0.333	0.949

표 8

[0147]

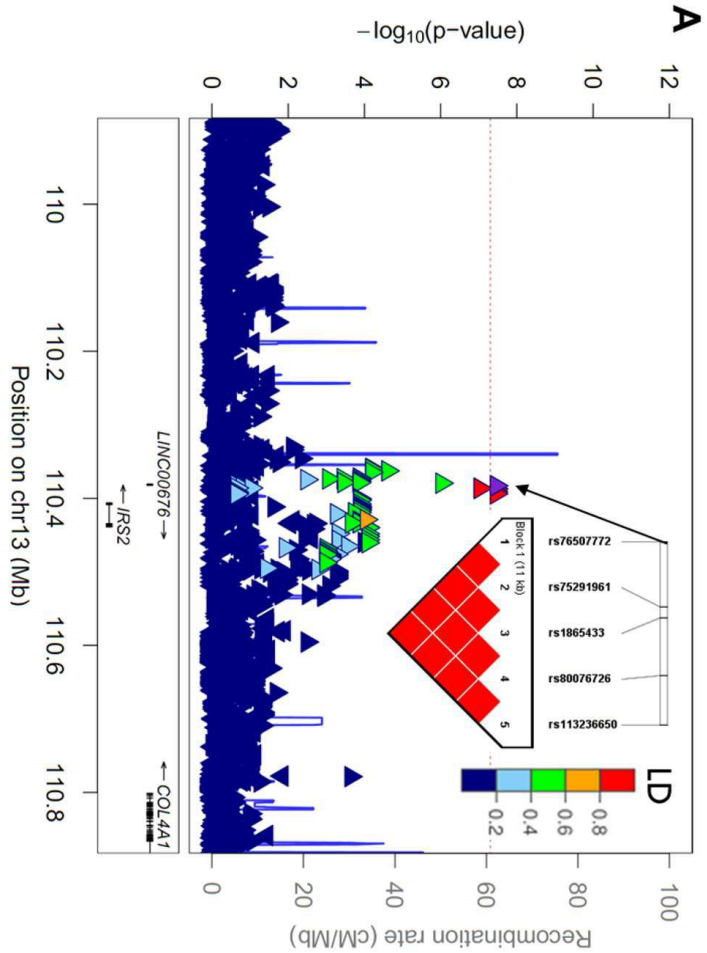
Hp2-2 모델	인지장애 (n=46)	비인지장애 (n=36)	위험률 (95% CI)	P 값	AUROC	민감도	특이도
T1: < 1.06	20 (43.5)	31 (86.1)	Reference				
T2: 1.06-2.12	10 (21.7)	5 (13.9)	7.86 (2.99-20,65)	2.9×10^{-5}		0.565	0.861
T3: 2.12 <	16 (34.8)	0 (0.0)	13.13 (5.99-28.78)	1.2×10^{-10}	0.737	0.348	1

[0148] 그 결과, 상기 wPRS 모델에 의할 경우, 지주막하 출혈(SAH) 후 발달 인지장애(CI)를 예측할 수 있었다.

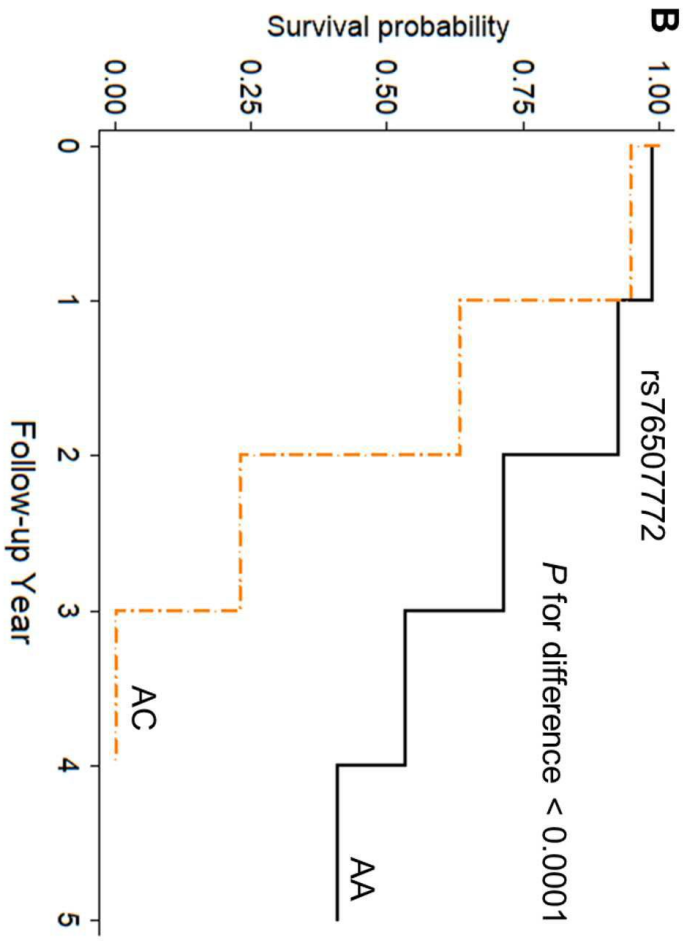
[0150] 이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시 양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

도면

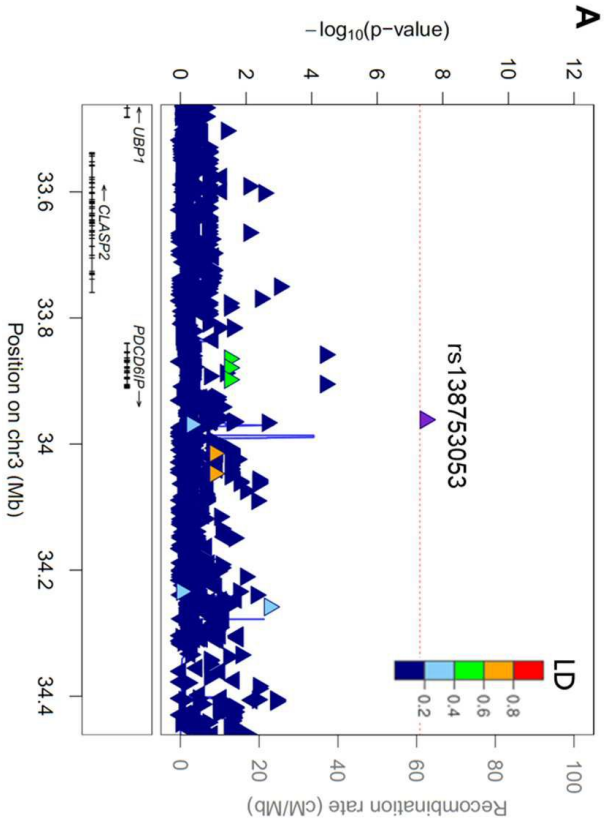
도면1


A
−log10(p−value)
0
2
4
6
8
10
12
Recombination rate (cM/Mb)
0
20
40
60
80
100
Position on chr13 (Mb)
110
110.2
110.4
110.6
110.8
LINC00676
IRS2
COL4A1
Block 1 (11 kb)
1
2
3
4
5
rs76507772
rs75291961
rs1865433
rs80076726
rs113236650
LD
0.2
0.4
0.6
0.8

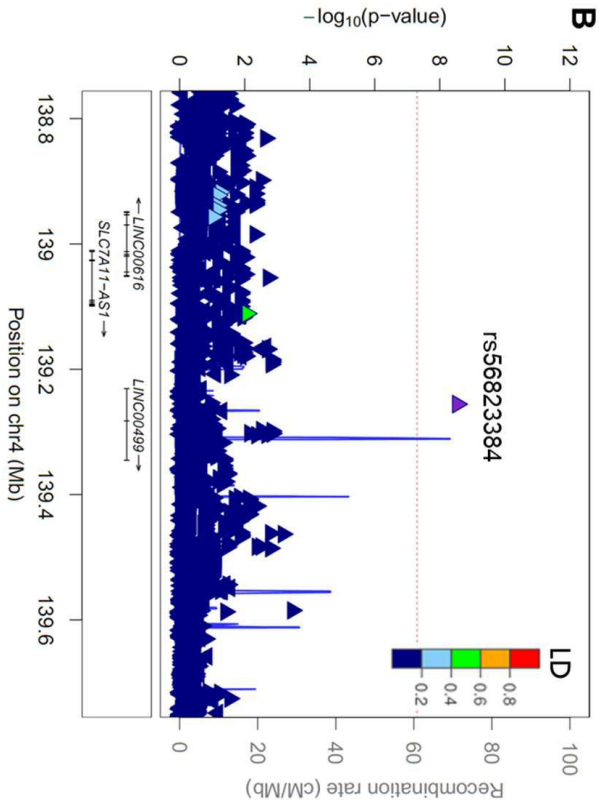
도면2


B
Survival probability
0.00
0.25
0.50
0.75
1.00
0
1
2
3
4
5
Follow-up Year
rs76507772
P for difference < 0.0001
AA
AC

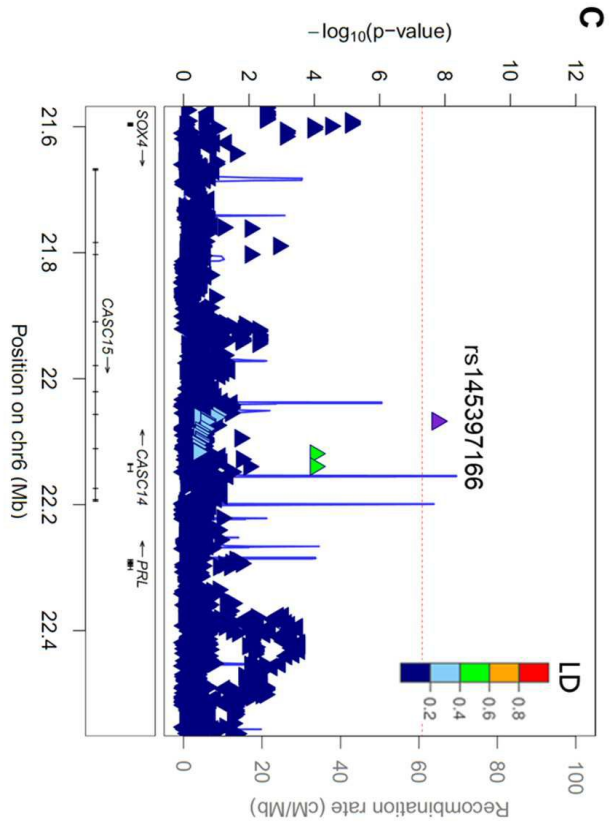
도면3


A
−log10(p-value)
rs138753053
LD
Recombination rate (cM/Mb)
Position on chr3 (Mb)
UBP1
CLASP2
PDCD6IP

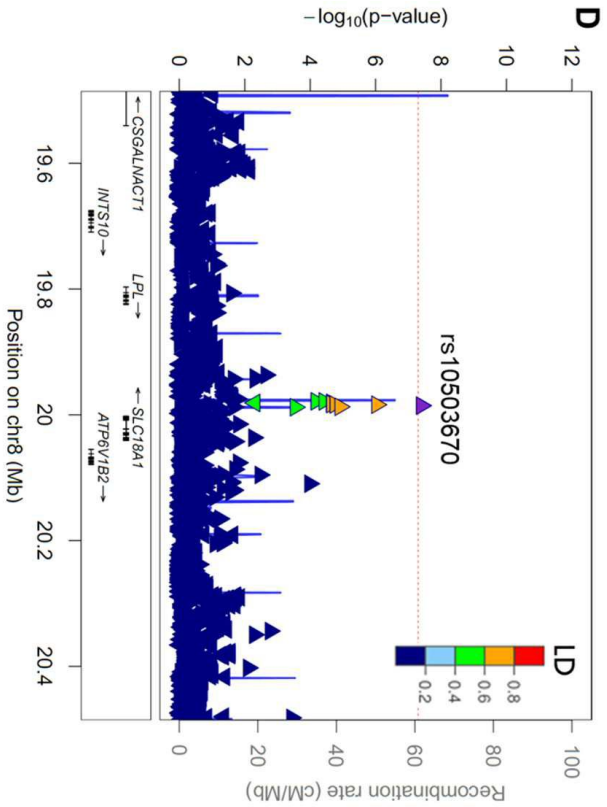
도면4


B
−log10(p-value)
rs56823384
LD
Recombination rate (cM/Mb)
Position on chr4 (Mb)
LINC00616
SLC7A11-AS1
LINC00499

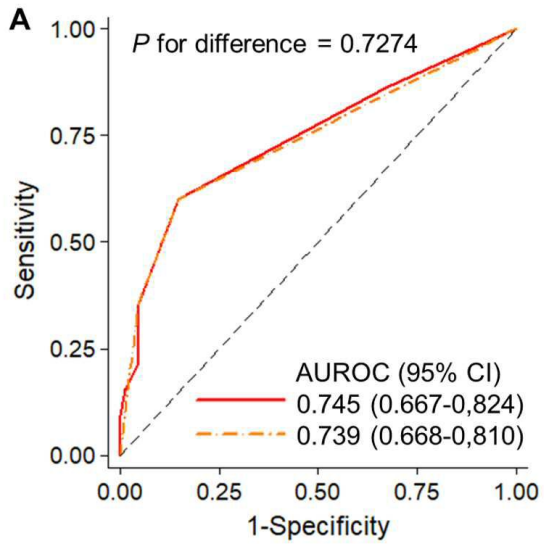
도면5


C
−log10(p-value)
rs145397166
LD
Recombination rate (cM/Mb)
Position on chr6 (Mb)
SOX4
CASC15
CASC14
PRL

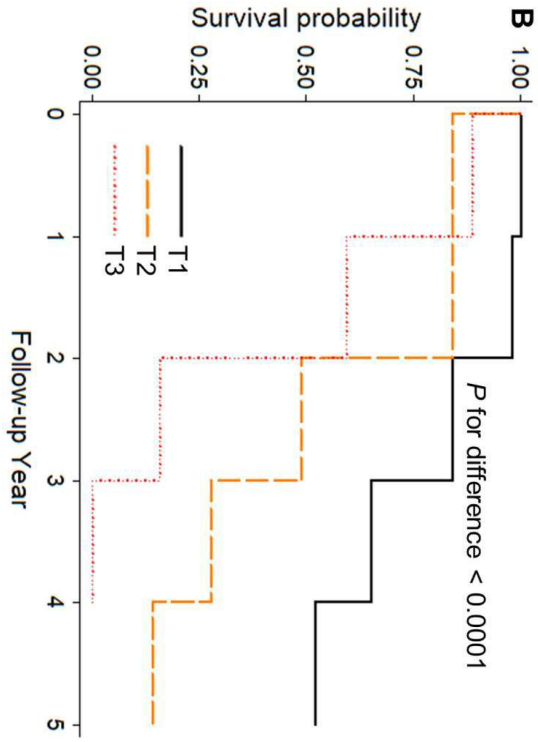
도면6


D
−log10(p-value)
rs10503670
LD
Recombination rate (cM/Mb)
Position on chr8 (Mb)
CSGALNACT1
INTS10
LPL
SLC18A1
ATP6V1B2

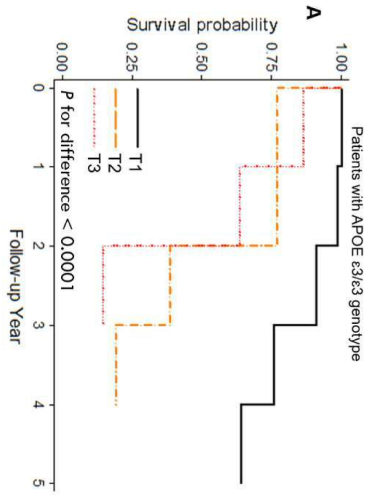
도면7


A
P for difference = 0.7274
Sensitivity
1.00
0.75
0.50
0.25
0.00
AUROC (95% CI)
0.745 (0.667-0,824)
0.739 (0.668-0,810)
0.00
0.25
0.50
0.75
1.00
1-Specificity

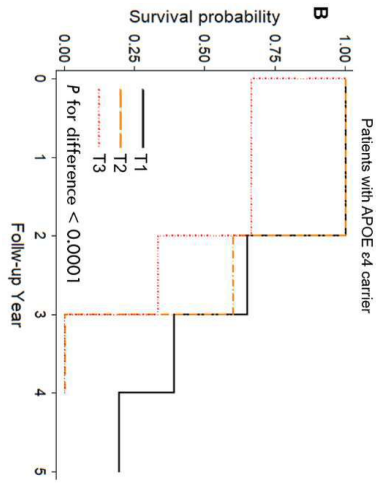
도면8


Survival probability
B
0.00
0.25
0.50
0.75
1.00
0
1
2
3
4
5
Follow-up Year
T1
T2
T3
P for difference < 0.0001

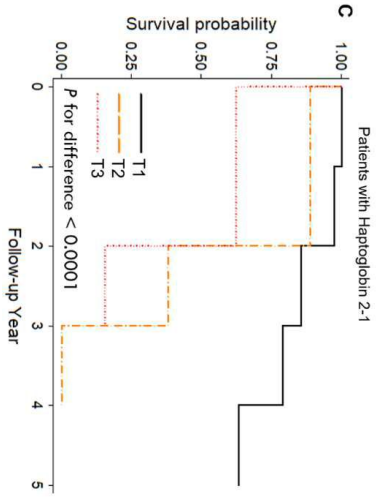
도면9



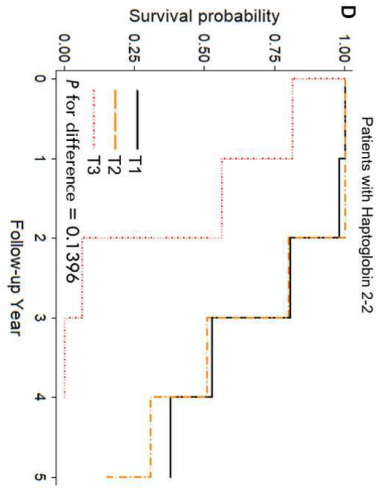
도면10



도면11



도면12


D
Survival probability
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
Follow-up Year
T1
T2
T3
P for difference = 0.1396
Patients with Haptoglobin 2-2