(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0117876
(43) 공개일자 2023년08월10일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6883 (2018.01) *C12Q 1/6827* (2018.01)
(52) CPC특허분류
C12Q 1/6883 (2022.01)
C12Q 1/6827 (2018.05)
(21) 출원번호 10-2022-0014158
(22) 출원일자 2022년02월03일
심사청구일자 2022년02월03일

(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
특허법인본

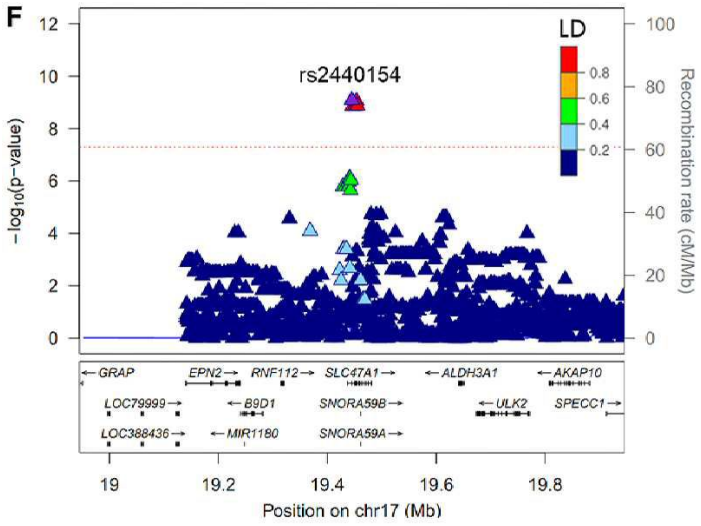
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **뇌동맥류 진단용 rs2440154 마커 조성물 및 이의 이용**

(57) 요 약

본 발명은 유전자형 교정 및 고처리량의 대체를 통해 업데이트된 IA GWAS를 구축하고 뇌동맥류 발병과 가장 유의한 관련성을 갖는 6종 SNP를 발굴하여 뇌동맥류 진단을 위한 SNP 마커로 제공한다. 본 발명의 SNP 마커는 높은 신뢰도를 갖는 뇌동맥류 진단 결과를 제공할 수 있으며, 뇌동맥류의 조기 진단을 통해 발병의 예방과 빠른 처치가 가능하게 할 수 있다.

대 표 도 - 도9



(52) CPC특허분류
C12Q 2600/156 (2013.01)
(72) 발명자
김봉준
강원도 춘천시 교동길 50

윤동혁
충청북도 단양군 어상천면 어상천로 1589-8

이 발명을 지원한 국가연구개발사업

과제고유번호	1465034383
과제번호	HR21C0198040021
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

rs2440154 단일염기다형성(single nucleotide polymorphism; SNP)을 포함하는, 뇌동맥류 진단을 위한 SNP 마커 조성물.

청구항 2

rs2440154 단일염기다형성(single nucleotide polymorphism; SNP)을 검출할 수 있는 제제를 포함하는, 뇌동맥류 진단용 조성물.

청구항 3

제2항에 있어서,

rs3120004, rs1522095, rs7779989, rs12935558, 및 rs3826442에서 이루어진 군으로부터 선택된 어느 하나 이상의 SNP를 검출할 수 있는 제제를 더 포함하는, 뇌동맥류 진단용 조성물.

청구항 4

제2항에 있어서,

상기 SNP를 검출할 수 있는 제제는 상기 SNP가 위치하는 유전체 영역에서 상기 SNP를 포함하는 8~50개의 연속된 염기서열 또는 이와 상보적인 염기서열을 갖는 폴리뉴클레오티드에 특이적으로 결합하는 프로브 또는 프라이머인 것을 특징으로 하는, 진단용 조성물.

청구항 5

제2항에 있어서,

상기 조성물은 한국인 또는 아시아인을 대상으로 뇌동맥류를 진단하기 위한 것인, 진단용 조성물.

청구항 6

제2항의 조성물을 포함하는, 뇌동맥류 진단용 키트.

청구항 7

하기 단계를 포함하는 뇌동맥류 진단을 위한 정보제공방법:

(1) 개체의 생물학적 시료를 준비하는 단계;

(2) 상기 생물학적 시료로부터 게놈(genomic) DNA를 얻는 단계;

(3) 상기 게놈 DNA에서 GenBank SNP database rs2440154 단일염기다형성(single nucleotide polymorphism; SNP) 위치의 염기 타입을 검출하는 단계; 및

(4) 상기 검출한 염기 타입으로부터 개체의 뇌동맥류 발병 위험도 또는 발병 감수성을 결정하는 단계.

청구항 8

제7항에 있어서,

상기 (1) 단계의 개체는 한국인 또는 아시아인인 것을 특징으로 하는, 정보제공방법.

청구항 9

제7항에 있어서,

상기 (3) 단계의 염기 타입 검출은 시퀀싱(sequencing), 마이크로 어레이에 의한 혼성화(microarray hybridization), 대립유전자 특이적인 PCR(allele specific PCR), 다이나믹 대립유전자 혼성화 기법(dynamic allele-specific hybridization; DASH), 형광 융해 곡선 기법(Fluorescence Melting Curve Analysis; FMCA)및 Taqman 기법으로 이루어진 군으로부터 선택되는 어느 한 방법으로 수행되는 것을 특징으로 하는, 정보제공방법.

청구항 10

제7항에 있어서,

상기 (3) 단계에서 rs2440154 위치의 염기 타입이 구아닌(Guanine)으로 검출되는 경우,

상기 (4) 단계의 개체는 뇌동맥류 발병 위험도 또는 발병 감수성이 높은 것으로 판단하는 것을 특징으로 하는, 정보제공방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 대한민국 보건복지부(과제번호: HR21C0198)와 한림연구기금의 지원을 받아 한국보건산업진흥원(KHIDI)을 통한 한국보건기술 R&D 사업의 지원으로 완성되었다.

[0002] 본 발명은 뇌동맥류 진단을 위한 바이오마커로서 rs2440154 SNP 마커 및 이를 포함하는 6종의 SNP를 발굴하여 개체의 뇌동맥류 발병가능성을 미리 진단할 수 있도록 상기 SNP의 용도를 제공하는 것이다.

배 경 기 술

[0003] 개인 맞춤의학의 등장으로 유전자 분석 기술에 대한 관심이 높아지고 있다. 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.

[0004] 뇌동맥류(Intracranial aneurysm; IA)는 대뇌 혈관의 국소 풍선 팽창을 나타낸다. 혈관이 파열되지 않은 IA는 일반적으로 무증상이지만 파열되면 발작 후 1개월 이내에 최대 50%의 사망률을 나타낸다. 이에, IA 발병 가능성이 높거나 혈관 파열 가능성이 높은 환자의 특성을 확인하기 위해 많은 연구가 진행중에 있다. 2010년 이후로 해외에서는 IA와 관련된 인과적 변이를 확인하기 위한 IA의 유전적 연구 수행되었다. 기존의 국내 연구는 후향적 연구 설계에 대한 단순 임상적 요인 분석만을 수행하는 한계가 있었으나, 후보유전자의 SNP에 대한 단순 분석, 전장유전체연관분석(Genome-Wide Association Study; GWAS), 전체 게놈 연관(whole-genome association), 차세대 염기서열 분석(next-generation sequencing; NGS) 및 가능한 원인 변이의 기능으로 연구가 확대되고 있

다.

[0005] 유전자 분석 결과는 인종에 따라 다를 수 있으며, 한국인에 대한 결과도 동북아 국가인 중국, 일본과 다를 수 있다. 예를들어, lysyl oxidase(LOX) 유전자에 위치한 rs3900446의 C 대립유전자는 한국인에서 IA 위험과 관련이 있는 것으로 알려져 있지만 일본 코호트 결과 IA와 LOX 유전자의 대립형질 또는 일배체형은 연관성이 없는 것으로 확인되었다. 따라서 질병과 관련된 SNP 발굴 연구에 있어 인종은 필수적으로 고려되어야 할 요인이다.

[0006] 2019년 1781명의 동북아시아인을 기반으로 한 게놈 참조 패널(A genome-reference panel)이 공개되었다. 백인(Caucasians) 기반의 1000개 게놈 프로젝트 컨소시엄의 참조 패널과 비교하여, 동북아 인구에 대한 강력한 전가 참조 패널(imputation reference panel)은 보다 개선된 결측값 대체 정확도(imputation accuracy)와 낮은 빈도의 변이를 나타내었다.

[0007] 본 발명자들은 이전의 IA GWAS에서 한국인의 결측값 대체(imputation)가 없는 문제를 해결하기 위해 업데이트된 GWAS를 수행하여 유전자형 교정 및 high-throughput imputation 후 IA 감수성 변이를 식별하여 IA와 유관한 총 6종의 SNP를 발굴하여 본 발명을 완성하였다.

선행기술문헌

비특허문헌

[0008] (비특허문헌 0001) 1. 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) 2. 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

[0009] 본 발명자들은 한국인을 대상으로 하는 첫번째 IA GWAS를 기반으로 뇌동맥류와 연관이 있는 단일염기다형성(Single Nucleotide Polymorphisms; SNPs) 후보군을 선정하고, 상기 GWAS 유전정보 데이터를 유전자형 교정 및 high-throughput imputation을 수행하여 업데이트된 GWAS 데이터를 기반으로 상기 SNP 후보군 중에서 최종적으로 한국인에서 뇌동맥류에 대하여 감수성을 갖는 SNP를 확보하여 본 발명을 완성하였다.

[0010] 이에, 본 발명이 이루고자 하는 기술적 과제는 뇌동맥류 진단을 위한 SNP 마커를 제공하고, 상기 SNP 마커를 검출할 수 있는 제제를 이용한 뇌동맥류 진단용 조성물 및 키트를 제공하는 것이다.

[0011] 또한, 본 발명은 개체의 생물학적 시료로부터 상기 SNP를 검출하여 뇌동맥류 진단을 위한 정보제공방법을 제공하고자 한다.

[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 발명은 뇌동맥류를 진단하기 위한 rs2440154 SNP 마커를 제공한다. 상기 SNP는 서열번호 6에서 나타난 염기 서열에 포함되는 것일 수 있다.

[0014] 또한, 본 발명은 뇌동맥류를 진단하기 위한 SNP 마커로서 rs3120004, rs1522095, rs7779989, rs12935558, 및 rs3826442에서 이루어진 군으로부터 선택된 어느 하나 이상의 SNP를 검출할 수 있는 제제를 더 포함할 수 있다. 일례로, 뇌동맥류를 진단하기 위한 SNP 마커로서 rs3120004, rs1522095, rs7779989, rs12935558, rs3826442 및 rs2440154을 모두 제공하는 것일 수 있다.

[0015] 또한, 본 발명은 상기 SNP 중에서 선택되는 1종 이상의 SNP를 검출할 수 있는 제제를 포함하는 뇌동맥류 진단용 조성물 및 진단용 키트를 제공한다.

[0016] 본 발명의 일 구현예로서, 상기 SNP 마커 조성물은 한국인 또는 아시아인을 대상으로 뇌동맥류를 진단하기 위한 것일 수 있다.

[0017] 본 발명의 다른 구현예로서, 상기 SNP를 검출할 수 있는 제제는 상기 SNP가 위치하는 유전체 영역에서, 상기 SNP를 포함하는 8-50개의 연속된 염기서열 또는 이와 상보적인 염기서열을 갖는 폴리뉴클레오티드에 특이적으로 결합하는 프로브 또는 프라이머이거나, 상기 폴리뉴클레오티드를 증폭할 수 있도록 상기 유전체에서 SNP의 위치를 기준으로 ± 4~100의 연속된 염기서열과 상보적인 8-20bp 길이의 프라이머일 수 있다.

[0018] 본 발명의 또 다른 구현예로서, 상기 SNP를 검출할 수 있는 제제는 rs3120004 위치에서 티민(Thymine)을 검출할 수 있거나, rs1522095 위치에서 구아닌(Guanine)을 검출할 수 있거나, rs7779989 위치에서 구아닌을 검출할 수 있거나, rs12935558 위치에서 아데닌(Adenine)을 검출할 수 있거나, rs3826442 위치에서 티민을 검출할 수 있거나, rs2440154 위치에서 구아닌을 검출할 수 있는 것일 수 있다.

[0019] 또한, 본 발명은 하기 단계를 포함하는 뇌동맥류 진단방법 내지 진단을 위한 정보제공방법을 제공한다:

[0020] (1) 개체의 생물학적 시료를 준비하는 단계;

[0021] (2) 상기 생물학적 시료로부터 게놈(genomic) DNA를 얻는 단계;

[0022] (3) 상기 게놈 DNA에서 GenBank SNP database rs3120004, rs1522095, rs7779989, rs12935558, rs3826442, 및 rs2440154로 이루어진 군으로부터 선택되는 어느 하나 이상의 단일염기다형성 위치의 염기 타입을 검출하는 단계; 및

[0023] (4) 상기 검출한 염기 타입으로부터 개체의 뇌동맥류 발병 위험도 및/또는 발병 감수성을 결정하는 단계.

[0024] 본 발명의 일 구현예로서, 상기 (1) 단계의 개체는 뇌동맥류 발병 가능성 예측이 요구되는 인간이라면 제한되지 아니하나, 바람직하게는 한국인 또는 아시아인일 수 있다.

[0025] 본 발명의 다른 구현예로서, 상기 생물학적 시료는 개체로부터 분리되어 게놈 DNA를 포함하는 것이라면 제한되지 아니하나, 간이한 비침습적 진단을 위하여 바람직하게는 혈액일 수 있다.

[0026] 본 발명의 또 다른 구현예로서, 상기 (3) 단계의 염기 타입 검출은 시퀀싱(sequencing), 마이크로 어레이에 의한 혼성화(microarray hybridization), 대립유전자 특이적인 PCR(allele specific PCR), 다이나믹 대립유전자 혼성화 기법(dynamic allele-specific hybridization; DASH), 형광 융해 곡선 기법(Fluorescence Melting Curve Analysis; FMCA) 및 Taqman 기법으로 이루어진 군으로부터 선택되는 어느 한나 이상의 방법을 이용하여 수행될 수 있다.

[0027] 본 발명의 또 다른 구현예로서, 상기 진단방법 내지 정보제공방법은 (3) 단계에서 rs3120004 위치의 염기 타입이 티민으로 검출되거나, rs1522095 위치의 염기 타입이 구아닌으로 검출되거나, rs7779989 위치의 염기 타입이 구아닌으로 검출되거나, rs12935558 위치의 염기 타입이 아데닌으로 검출되거나, rs3826442 위치의 염기 타입이 티민으로 검출되거나, rs2440154 위치의 염기 타입이 구아닌으로 검출되는 경우, (4) 단계의 개체는 뇌동맥류 발병 위험도 및/또는 발병 감수성이 높은 것으로 판정할 수 있다.

발명의 효과

[0028] 본 발명은 유전자형 교정 및 결측값 대체를 통해, 한국인을 대상으로 하는 업데이트된 IA GWAS 유전정보 데이터를 기반으로 뇌동맥류에 감수성을 갖는 SNP를 제공할 수 있다. 그 결과, 높은 신뢰도를 갖는 뇌동맥류 진단 결과를 제공할 수 있으며, 뇌동맥류의 조기 진단을 통해 발병의 예방과 빠른 처치가 가능하도록 하는 장점이 있다.

도면의 간단한 설명

[0029] 도 1은 Axiom™Genotyping Array에 의해 권장되는 유전자형 교정 프로토콜이다. Axiom™Analysis Suite v1.1.1, Applied Biosystems™ (APT) v1.18, 및 SNPolisher™module of R package v3.0. APT를 통해 실행되는 각 단계에 사용된 모든 명령은 적절한 QC 메트릭스를 생성하였다. 3단계 및 5단계의 샘플 필터링은 SNPolisher™module에 의해 구현되는 R script로 수행하였다.

도 2는 뇌동맥류에 대한 업데이트된 GWAS 구축을 위한 방법과 각 단계에서 도출된 결과를 요약한 모식도이다. 도 2에서 R^2은 imputation score를 의미하고, GCR은 genotype call rate의 약자이며, MAF은 소수 대립유전자 빈

도(minor allele frequency)를 의미하고, HWE p는 Hardy-Weinberg equilibrium p-value를 의미한다.

도 3(A)는 1000 Genome(1KG) reference panel(Phase 3 version 5)에서 546명의 한국 성인과 2,504명의 피험자의 multi-dimensional scaled principal component(MDS PC)의 플롯이다. K는 1KG reference panel에서 2,504명의 피험자를 나타낸다(유럽인 수 = 503[빨간색 원], 히스패닉 또는 혼혈 미국인 = 347[주황색 원], 아프리카인 = 661[회색 원], 동아시아인 = 504[진녹색]). 빨간색 원은 패널 A에서 문자 "H"로 표시된 한국 성인의 영역을 나타낸다.

도 3(B)는 MDS PC에서 한국 성인 546명의 MDS PC 플롯이다. 한국 성인 546명 중 PC1과 PC2 사이의 유전적 가계 클러스터는 1KG 참조 패널의 504명의 동아시아인 대상과 비교할 때 546명의 성인의 우수한 분산을 나타낸다.

도 3(C)는 250명의 환자 및 296명의 대조군의 뇌동맥류에 대한 GWAS 맨해튼 플롯이다. X축과 Y축은 각각 염색체 수와 -log10으로 변환된 p값을 나타낸다. 빨간색 하위 라인은 게놈 전체의 유의성 임계값을 나타낸다($p = 5\times 10^{-8}$).

도 4는 본 발명의 rs3120004(1q31.2) SNP ± 500kb 영역의 지역 연관 플롯으로,삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

도 5는 본 발명의 rs1522095(4q12) SNP ± 500kb 영역의 지역 연관 플롯으로, 삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

도 6은 본 발명의 rs7779989(7p15.3) SNP ± 500kb 영역의 지역 연관 플롯으로,삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

도 7은 본 발명의 rs12935558(18q) SNP ± 500kb 영역의 지역 연관 플롯으로, 삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

도 8은 본 발명의 rs3826442(17p13.1) SNP ± 500kb 영역의 지역 연관 플롯으로, 삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

도 9는 본 발명의 rs2440154(17p11.2) SNP ± 500kb 영역의 지역 연관 플롯으로, 삼각형 또는 역삼각형은 각각 양의 효과 크기와 음의 효과 크기를 나타내며, 각 색상은 pairwise LD을 나타내고, 보라색 위쪽 삼각형은 IA 형성과 가장 중요한 연관성을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0030] 유럽, 일본, 중국과 달리 한국인을 대상으로 하는 뇌동맥류(IA)에 대한 유전적 연구는 부족한 실정이다. 본 발명자들은 한국인을 대상으로 하는 최초의 GWAS보다 유전자형 교정 및 대체를 통해 업데이트된 GWAS를 구축하고 이를 기반으로 뇌동맥류의 발병 가능성을 예측할 수 있는 최종 6종 SNP를 발굴하여 본 발명을 완성하였다.

[0031] 구체적으로, 본 발명자들은 250명의 IA 환자와 296명의 건강한 대조군의 유전정보 데이터를 이용하여 임상변수와 4가지 주요성분분석(PCA) 값으로 조정하고 SNP 및 유전자형 결측값의 고처리량 대체(high-throughput imputation)를 수행하였다. imputation R^2 score of ≥ 0.5의 총 73,33,746개 사이트 중에서 6,105,212개의 SNP가 분석되었으며, genome-wide significant threshold($p = 5\times 10^{-8}$)에 도달한 총 38개 SNP를 선별하였으며, pairwise LD(linkage disequilibrium)($r^2 < 0.8$)에 의한 가지치기 기법(pruning) 수행 후 11개의 SNP가 일관되게 IA와 관련성을 나타냄을 확인하였다. 나아가 rs3120004, rs1851347, rs1522095, rs7779989, rs12935558, rs3826442, 및 rs2440154의 6개 표지된 SNP는 강하게 LD tower를 갖는 일배체형(haplotype) 구조를 형성하며, 특히 rs3120004는 *LOC440704* 및 *RGS18* 유전자 사이에 1q31.2(OR = 2.34, 95% CI: 1.74-3.14; p = 1.4×10-8) 위치에서 크고 강하게 일배체형 구조를 형성함을 확인하였다. rs2440154 SNP는 IA 발병 위험 증가와 가장 높은 관련성을 나타내었으며, rs3826442(*MYH13*, 17p13.1)을 포함하는 영역은 약 70 cM/Mbp.의 높은 재조합 비율을 나타냄을 확인하였다.

[0032] 상기로부터, 본 발명자들은 고처리량 대체(high-throughput imputation) 접근법을 사용한 업데이트된 GWAS로부터 한국인에서 IA와 관련된 11개의 새로운 SNP를 도출하였으며, SNP 중에서 일배체형 구조를 구성하는 6종 SNP(rs3120004, rs1522095, rs7779989, rs12935558, rs3826442 및 rs2440154)를 IA 진단을 위한 최종 SNP 마커로 선정하였다. 상기 최종 SNP 마커 중에서 rs2440154(*SLC47A1*, 17p11.2)가 IA의 발생과 가장 강하게 관련된다.

[0033] 이에, 본 발명은 상기 6종 SNP를 IA 진단을 위한 바이오마커로 제공한다.

[0034] 본 발명의 IA 진단용 바이오마커의 구체적인 정보는 하기 표 1에 나타내었다.

표 1

[0035]

서열 번호	유전자	염색체	SNP	Position	P-value	SNP 염기서열 정보 (염기서열위치[] 표기)
1	LOC440704, RGS18	1q31.2	rs3120004	191221672	1.4E-08	AAGTATTTGGTAAAATTCAGCAGTAAATTCACCAGGTCCTTGGCTATTCTTTGATGGGAGACATTTTATTCATAGTTCTATCTTGTTACTTAATATTGAT[C/T]GATTCAGGTTTTATATTTCTTCATGGTTCATTCTTGGTAGGTGGTATGTGTTTAGAAATTTATCCATTACCTCTAAGTTTTCAAATTTAGTGGTACATAG
2	LOC101928851, NONE	4q12	rs1522095	59222703	2.6E-09	TCCCTAAAAAGGGAGAAATTACAAACATTTTTTCCTAGAAATGGAGAGGACATAATATTTAATGCTATTCTGATGGCTTTGCATTGAAATTGGTCAAGCT[A/G]ACTAGGCTGAGATTATTTTCCAAAATTTTCTTTCAAGTTACAGTGGGACACAAGAGAAATTCTTGTGTGAGTTTTGGAGGGCAGAGGTGATACTCATACT
3	LINC01162, SP4	7p15.3	rs7779989	21265984	4.6E-09	TTCTGCTCTTTGAACCCATCTCTCCAGGTTGAAACCATCTTGGCAAAGTATTATCAGCCGGTGTCTTCGGCTAGATGGACCTTTCAAGCCCTCCTAAGGT[A/G]AGAAAATGTTATTTTTCCCTTTTATGTGTTGTGGTTCATTGGATAGTATTTGTCATCTTGTCACATAATAGGAATGACGTTCACAATGTTGAAATACGCC
4	LOC101928880, ZNF469	16q24.2	rs12935558	88336978	2.6E-08	CAAGGTAACAATTCACTTCATCCCATGTATCATTGCATAAGTTCTCCTGGGCCCCGTTGGCTGGGGACGTGGACACGTGTGATTTCTGCTCTGCACCCTC[G/A]CCAGCGCTGTTCTTTCCAGTGGGCCTTTCCAGCCGGGATCCTGCCCACTTCCTTGTGCCCAGCTTTCCAGAGCCCCTCAGCCCCTCCAAGGGCCCACCA
5	MYH13	17p13.1	rs3826442	10227708	3.1E-09	TTGATTGGCCTCTTTGAAAATGCTACAAAGAAAGGGTATCACCCTCTAAAAGCTTATTATCTACACCTAAGCCAAGAAATGGGTCATGCACACACATGCA[C/T]GCACATGTGCACACACTGCACACACACGTGTACATAAACTTAGATATCACCCATAAACACATTTTTTCTTTTTAGATGGAGTTTCTCTCTGTGGCCCAGG
6	SLC47A1	17p11.2	rs2440154	19444987	8.2E-10	AGGCGTTCAGCAGGCCTGTAATGTTTCACCCTGAGAGTCCGGATCCCGCTGGGACTCCAAGTGGACGATCTCATCGCTACCTCCAGCTCCTCCTGTCTAC[A/G]GACTTCCTTGCTGGAATGTGTTCTTGTTGTTGAAGGCTGTATTCTCTCCAGTCCTCACTAGGAATCTTGGATAACAAGGGGGTTATCCCAGAAGGCAGCC

[0036] IA 발병과 가장 강한 관련성을 나타낸 rs2440154 SNP는 *SLC47A1* 유전자 상에 위치하며, *SLC47A1* 유전자는 solute carrier family 47 member 1로 불리우며 담즙과 소변을 통한 유기 양이온의 ATP 유출에 관여하여 약물 반응 및 독성과 관련이 있는 것으로 알려져 있으나, 뇌혈관 질환의 발병과의 관련성에 대해서는 알려진 바가 없다.

[0037] 한편, 일배체형(Haplotype) 분석은 SNP와 표현형과의 연관성을 보다 강하게 보장한다. 같은 블록 내의 유전자형

은 함께 유전되는 경향이 있기 때문에 LD가 발생하는데 일배체형은 상대적으로 유전되는 인접 유전자좌들의 조합을 포함하여 LD 값이 높을 수록 질병 연관 가능성이 높음을 의미한다. 본 발명의 6종 SNP 마커는 모두 일배체형 구조를 나타내었으며, 그 중에서 rs3120004는 0.8 이상의 높은 LD로 크고 강한 일배체형 구조를 나타내었다. 또한, rs3826442는 이를 둘러싸고 있는 영역에서 높은 재조합률을 보여 다음 세대로의 유전 가능성을 시사한다.

[0038] 본 발명은 종래의 한국인 대상의 IA GWAS 결과에서 동북아 인구를 기반으로 한 대체(imputation)을 추가하여 통계적 검정력을 향상시켰으며, 1000명 이상의 뇌혈관 질환 환자를 통합한 유전자형 교정을 수행하여 결측값의 수를 감소시켰고, 본 발명의 6종 SNP는 pairwise LD 구축하에서 높은 확률로 다음 세대로 유전됨을 알 수 있다.

[0039] 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.

[0040] 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.

[0041] 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 프레임의 변화를 유발할 수 있다.

[0042] 한편 본 명세서에서 유전자의 변이를 표기할 때 [염기일문자/염기일문자]는 좌측에 쓰인 염기일문자가 우측에 쓰인 염기일문자로 치환되는 것을 의미한다.

[0043] 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.

[0044] 본 명세서에서 “게놈 DNA(genomic DNA: gDNA)”는 한 개체의 유전자의 총 염기서열로서, 거의 완전한 유전정보를 포함하고 있는 DNA를 의미한다.

[0045] 한편, 뇌동맥류(Intracranial aneurysm: IA)란 뇌에 혈액을 공급하는 뇌동맥 일부가 약해져서 그 부분이 풍선이나 꽈리처럼 부풀어 오르는 질환을 의미한다.

[0046] 본 명세서에서 “진단”은 질병 발생의 예측 및 발병 위험도(risk) 또는 발병 감수성(susceptibility)를 결정하거나 도출하는데 사용되는 모든 유형의 분석을 포함한다.

[0047] 본 명세서에서 "개체"란 뇌동맥류의 발병 가능성 예측 등의 진단이 필요한 포유류라면 제한되지 아니하나, 특히 유전적 소인에 의해 뇌동맥류 발병이 우려되는 개체일 수 있다.

[0048] 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.

[0049] 본 발명에서 "SNP를 검출할 수 있는 제제"란 상기 SNP가 위치하는 유전체 영역에서 상기 SNP를 포함하는 8~100개의 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.

[0050] 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.

[0051] 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 있다.

[0052] 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.

[0053] **[실험 방법]**

[0054] **1. 연구 대상 및 GWA 데이터 세트**

[0055] 본 연구는 2019년 한국인을 대상으로 하는 첫번째 IA GWAS를 기반으로 한 생물정보학 데이터를 이용하였다(J Clin Med. 2019 Feb; 8(2): 275.). 보다 구체적으로 250명의 IA 환자와 296명의 건강한 대조군의 유전정보 데이터를 이용하였으며, 본 연구에서 분류한 IA 환자의 기준은 아래와 같다:

[0056] 1) 낭상 뇌동맥류(saccular cerebral aneurysm);

[0057] 2) 18세 이상의 성인 환자;

[0058] 3) 뇌에 다른 혈관 기형이 없음.

[0059] Eagle v2.4 및 Minimac4 프로그램을 사용하여 각 개체별로 SNP의 사전 단계 유전자형을 분석하고, Axiom™ Asian Precision Medicine Research Array(APMRA)(Thermo Fisher Scientific, MA, USA)을 이용하여 유전자형이 지정된 512,575개의 SNP를 포함하는 250명의 환자 및 296명의 대조군에 대한 대량의 SNP 및 유전자형 결측값을 추정하였다.

[0060] **2. 유전자형 교정, high-throughput imputation, 및 quality control(QC)**

[0061] 상기 방법 1에서 확보한 첫번째 GWAS는 무작위 대조연구 데이터가 아니므로, "Batch Allele Consistency(BAC)" 기능을 통해 게놈 전체의 유전자형 데이터를 수정하였다. 각 개체의 유전자형 분석에 있어서 각각의 샘플을 별도로 처리함에 따라 발생하는 강도 변화로 데이터에 불규칙한 SNP가 포함될 수 있으며, BAC는 유전자형에서 불규칙한 SNP를 식별하고 제거하는데 사용될 수 있다.

[0062] 이에, 상기 방법 1에서 확보한 첫번째 유전자형 데이터를 Axiom™Analysis Suite v1.1.1, Applied Biosystems™ (APT) v1.18, 및 SNPolisher™R package v3.0 을 이용하여 교정하였다(도 1).

[0063] 이어서, NHLBI(National Institutes of Health(NIH) National Heart, Lung and Blood Institute)에서 지원하는 GenomeAsia 100K 프로젝트에서 제작한 Asian-specific reference panel(GRCh37hg19)을 적용하였다.

[0064] 또한, 높은 처리량(high-throughput)의 변이(variation)는 Michigan Imputation Server v1.5.7 (https://imputationserver.sph.umich.edu/index.html#!run/minimac4)으로 보충(imputation)하였다.

[0065] 본 연구는 기관생명윤리위원회(The Institutional Review Board and Ethics Committee)의 승인을 받아 수행하였다(No. 2016-3 및 2019-06-006).

[0066] **3. 통계 분석**

[0067] 고처리량 GWAS(high-throughput GWAS)는 성별, 연령, 기저질환 및 4가지 주요성분분석(PCA) 값과 같은 변수를 조정한 후 수행하였다. 가법 유효 모델(the additive effective model)에서 상속 모델을 수립하고, 대규모 GWAS는 PLINK v1.9 program (https://www.cog-genomics.org/plink/)으로 수행되었다. 맨해튼 프롯 및 지역 관련 플롯은 R package v3.6.0 (https://cran.r-project.org/web/packages/qqman)의 "qqman"과 수정된 Python and R scripts (http://locuszoom.org/)의 LocusZoom v1.3로 제작하였다.

[0068] **[실험 결과]**

[0069] 실험의 모식도와 결과의 요약은 도 2와 같다.

[0070] 첫번째 GWAS로 29개의 SNP를 확인하였으며, 이는 멘델 변이(Mendelian variants)에 의한 것일 수 있다. LD(Linkage Disequilibrium) tower의 결핍에도 불구하고 *GBA*, *TCF24*, *OLFML2A*, *ARHGAP32*, *CD163L1*, *CUL4A*, *LOC102724084*, 및 *LRRC3*의 8개 유전자위(loci)에서 80%의 충분한 통계적 검증력을 나타내었으며, 이전에 보고된 *BOLL* 및 *EDNRA* 유전자위도 검증되었다. high-throughput imputation를 기반으로 업데이트된 GWAS는 총 7,333,746개 사이트 중 6,105,212개의 SNP를 제공하여 imputation R^2 score of ≥ 0.5를 나타내었으며, genotyping call rate of ≥ 95%, minor allele frequency(MAF) of ≥ 0.01, 및 Hardy-Weinberg equilibrium(HWE) p-value of ≥ 0.001의 QC 테스트를 통과하였다.

[0071] 546명의 연구대상(250명의 IA 환자 및 296명의 대조군)의 PC1 및 PC2는 동아시아 인구(진한 녹색)에서 양호한 분산을 나타내었고(도 3A), 대상간의 이질성 없이 확대된 분포를 나타내었다(도 3B). 6M 마커 중에서 39개 SNP가 genome-wide significant threshold($p = 5\times10^{-8}$)에 도달하였다(도 3C).

[0072] LD 쌍($r^2 < 0.8$)에 의한 가지치기(pruning) 처리 이후에 MAF 및 HWE p-value(≥0.05)로 정의된 IA와 일관되게 관련된 11개의 SNP를 확인할 수 있었다(표 2).

표 2

[0073]

Gene	Function	Chr	SNP	Position	M	m	MAF	HWE p	OR	L95	U95	P
LOC440704, RGS18	intergenic	1q31.2	rs3120004	191221672	C	T	0.374/0.218	0.865	2.34	1.74	3.14	1.4E-08
LOC101928851, NONE	intergenic	4q12	rs1851347	59220726	A	G	0.438/0.279	0.564	2.26	1.70	3.01	1.9E-08
LOC101928851, NONE	intergenic	4q12	rs1522095	59222703	A	G	0.344/0.188	0.848	2.66	1.93	3.66	2.6E-09
LINC01162, SP4	intergenic	7p15.3	rs7779989	21265984	A	G	0.378/0.221	0.065	2.49	1.83	3.37	4.6E-09
IGSF9B	intronic	11q25	rs191300	133800017	T	C	0.192/0.074	1.000	3.14	2.08	4.72	4.4E-08
C12orf5	Near UTR-5	12p13.32	rs4766232	4430018	A	G	0.292/0.451	0.196	0.44	0.33	0.59	1.1E-08
LINC00433, LINC00440	intergenic	13q31.2	rs1571056	89306589	T	C	0.18/0.323	0.425	0.42	0.31	0.57	3.7E-08
ATP11A	intronic	13q34	rs2297801	113532224	C	T	0.48/0.323	1.000	2.22	1.68	2.95	3.0E-08
LOC101928880, ZNF469	intergenic	16q24.2	rs12935558	88336978	A	G	0.34/0.186	0.254	2.47	1.80	3.40	2.6E-08
MYH13	intronic	17p13.1	rs3826442	10227708	C	T	0.242/0.416	0.634	0.42	0.32	0.56	3.1E-09
SLC47A1	intronic	17p11.2	rs2440154	19444987	G	A	0.282/0.125	0.791	2.90	2.07	4.08	8.2E-10

[0074] [a] M and m indicate major and minor allele types, respectively. Minor allele frequency (MAF) shows in case (left) and control (right). [b] Hardy-Weinberg equilibrium p-value (HWE p) was estimated under the control group composed of 296 subjects.

[0075] [c] Odds ratio (OR), 95% confidence intervals (Cis: L95, lower level of 95% CI; U95, upper level of 95% CI), and p-value were estimated `by multivariate generalized linear model.

[0076] [d] One tagging SNP is presented, which shows a strong pairwise linkage disequilibrium (LD, r^2 > 0.8) in multiple SNPs.

[0077] 상기 11개 SNP 중에서 6개의 SNP가 LD tower를 갖는 일배체형(haplotype) 구조를 형성하였다(도 4, 도 5, 도 6, 도 7, 도 8 및 도 9). 표지된 rs3120004 SNP는 1q31.2 위치의 *LOC440704 및 RGS18* 유전자 사이에서 커다란 일배체형 구조를 형성하였다((odds ratio (OR) = 2.34, 95% confidence interval (CI): 1.74-3.14; p = 1.4×10^{-8}). 그러나, 유전자간 SNP(*LOC101928851*, 4q12)인 rs1522095는 rs1851347 및 rs1996807 SNP로 완전한 LD를 나타내었다(OR = 2.28, 95% CI: 1.70-3.01; p = 1.9×10^{-8}). 그러나, rs1522095는 rs1851347와 표지된 rs1522095와 함께 0.64의 LD(r^2) 쌍을 나타내는 단일 공통 SNP임을 확인하였다(OR = 2.66 95% CI: 1.93-3.66; p = 2.8×10^{-9}). 상술한 4개의 SNP는 유전자 사막 영역(gene desert region)에 위치하였다.

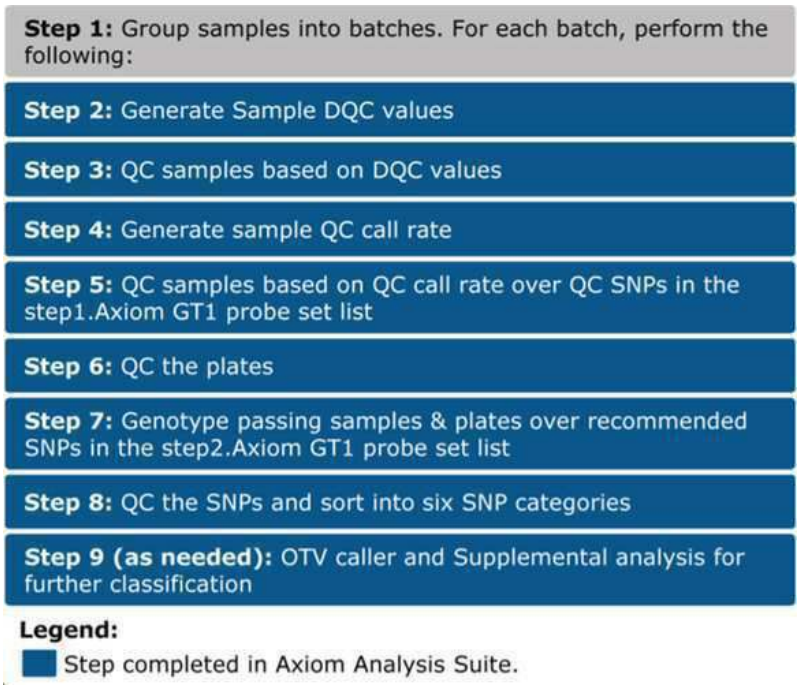
[0078] 한편, rs2440154 SNP(*SLC47A1*, 17p11.2)는 IA 발생 위험 증가와 가장 유의한 관련성을 나타내었다((OR = 2.90, 95% CI: 2.07-4.08; p = 8.2×10^{-10}). 추가로 7p15.3, 16q24.2 및 17p13.1의 3 영역도 IA 발생의 위험을 예측하는데 유의한 관련성을 나타내었다. 특히, rs3826442 SNP(*MYH13*, 17p13.1) 주변 영역은 약 70 centi-morgan(cM)/mega base-pair(Mbp)의 높은 재조합율을 나타내었다.

[0079] rs191300(*IGSF9B*, 11q25), rs4766232(*C12orf5*, 12p13.32), rs1571056(*LINC00433-LINC00440*, 13q31.2), 및 rs2297801(*ATP11A*, 13q34)의 다른 4개 SNP는 현재 GWAS에서 단일 분리된 SNP를 나타내었으며 표지된 SNP 주변에는 LD를 확인할 수 없었다.

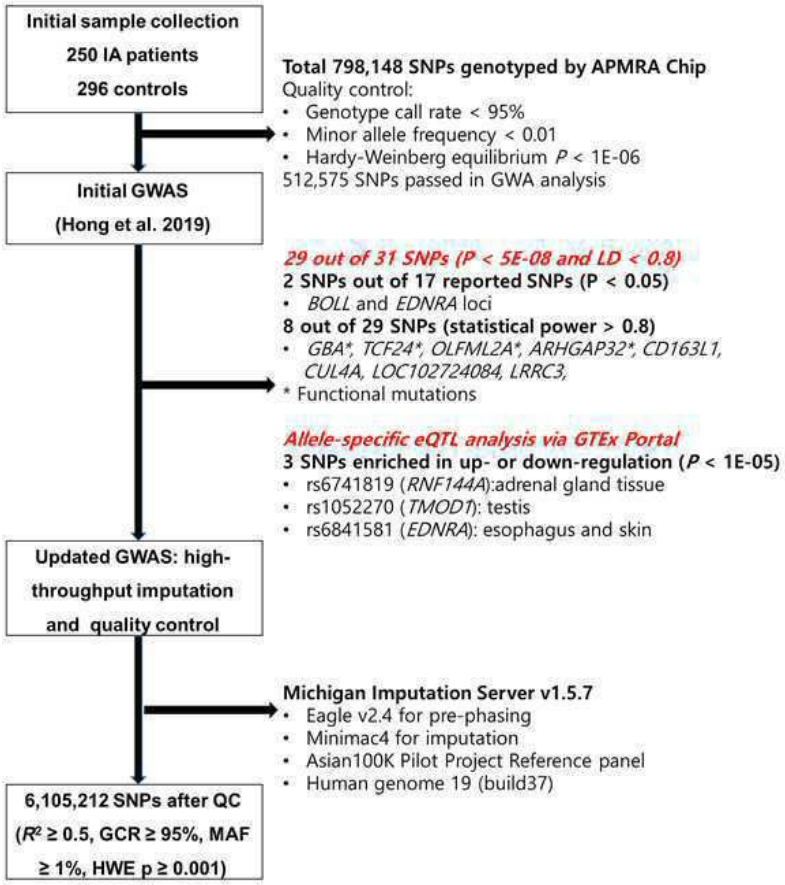
[0080] 이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시 양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

도면

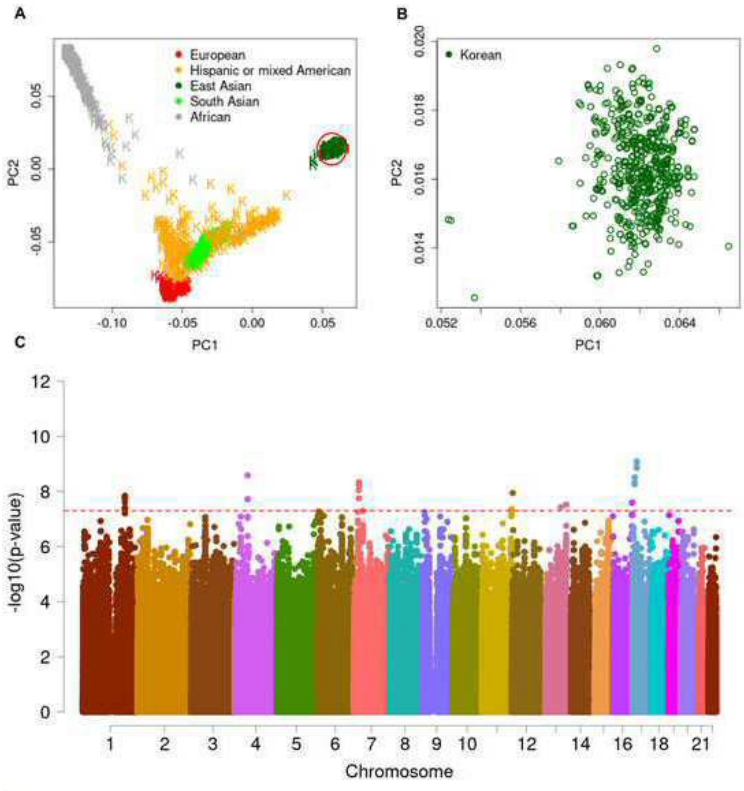
도면1



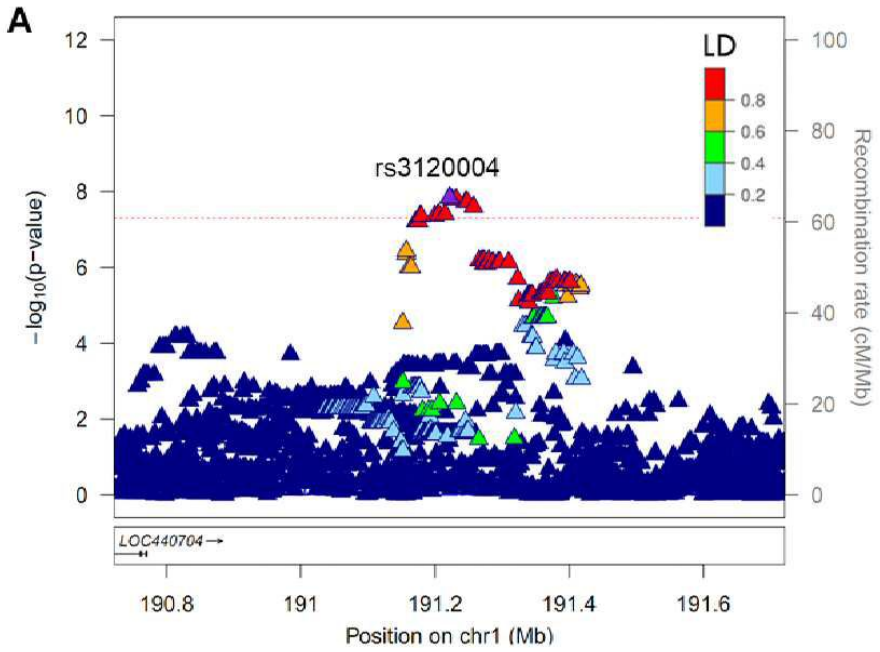
도면2


Initial sample collection
250 IA patients
296 controls
Total 798,148 SNPs genotyped by APMRA Chip
Quality control:
• Genotype call rate < 95%
• Minor allele frequency < 0.01
• Hardy-Weinberg equilibrium P < 1E-06
512,575 SNPs passed in GWA analysis
Initial GWAS
(Hong et al. 2019)
29 out of 31 SNPs (P < 5E-08 and LD < 0.8)
2 SNPs out of 17 reported SNPs (P < 0.05)
• BOLL and EDNRA loci
8 out of 29 SNPs (statistical power > 0.8)
• GBA*, TCF24*, OLFML2A*, ARHGAP32*, CD163L1, CUL4A, LOC102724084, LRRC3,
* Functional mutations
Allele-specific eQTL analysis via GTEx Portal
3 SNPs enriched in up- or down-regulation (P < 1E-05)
• rs6741819 (RNF144A):adrenal gland tissue
• rs1052270 (TMOD1): testis
• rs6841581 (EDNRA): esophagus and skin
Updated GWAS: high-throughput imputation and quality control
Michigan Imputation Server v1.5.7
• Eagle v2.4 for pre-phasing
• Minimac4 for imputation
• Asian100K Pilot Project Reference panel
• Human genome 19 (build37)
6,105,212 SNPs after QC
(R² ≥ 0.5, GCR ≥ 95%, MAF ≥ 1%, HWE p ≥ 0.001)

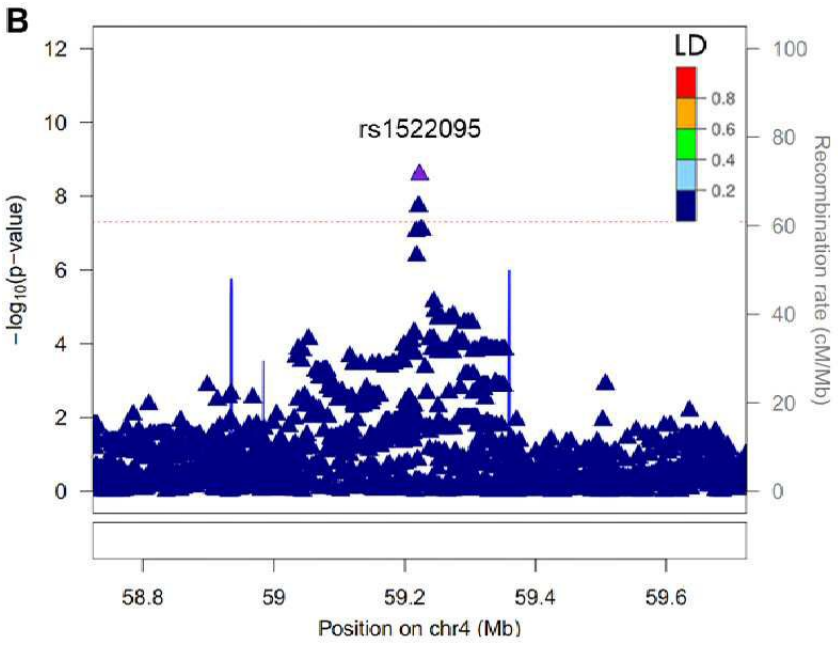
도면3


A
European
Hispanic or mixed American
East Asian
South Asian
African
PC1
PC2
B
Korean
PC1
PC2
C
-log10(p-value)
Chromosome

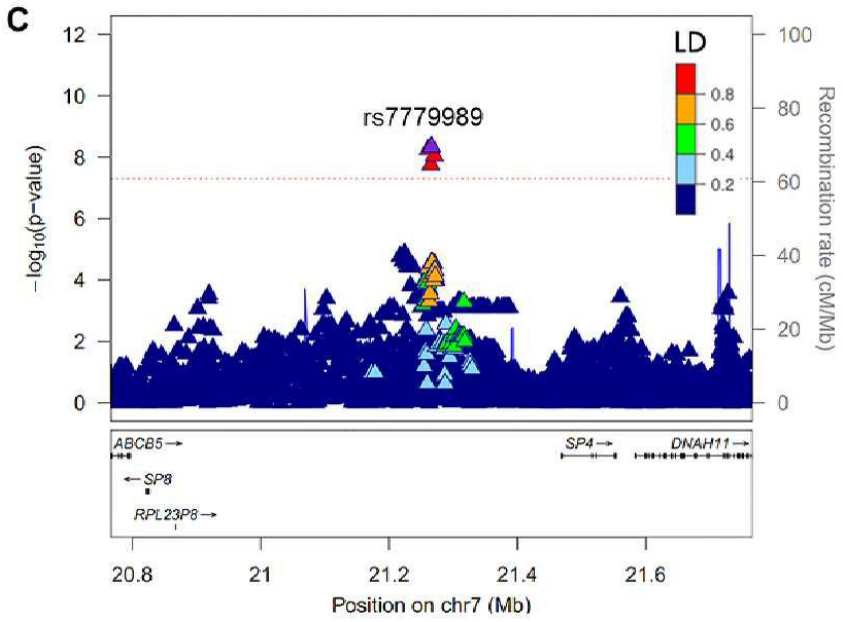
도면4


A
LD
0.8
0.6
0.4
0.2
rs3120004
−log10(p−value)
Recombination rate (cM/Mb)
LOC440704
190.8
191
191.2
191.4
191.6
Position on chr1 (Mb)

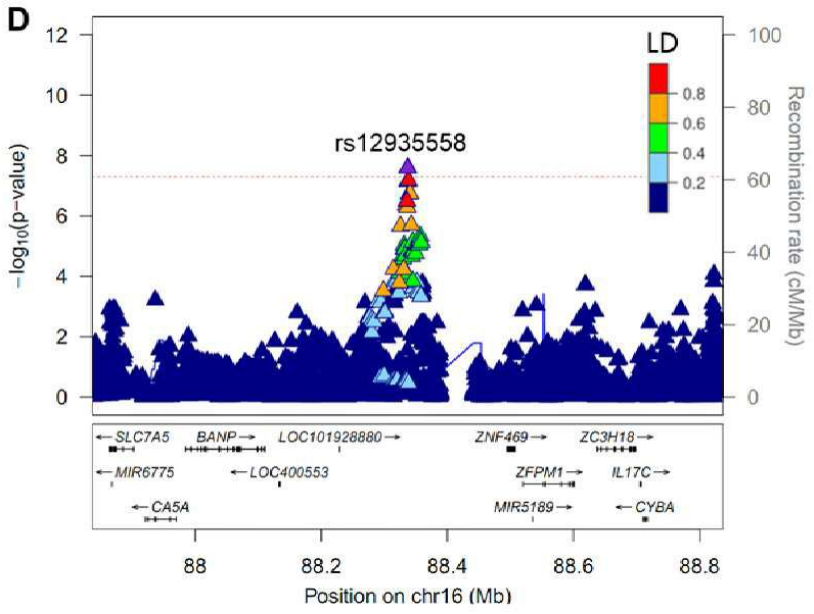
도면5


B
LD
0.8
0.6
0.4
0.2
rs1522095
−log10(p−value)
Recombination rate (cM/Mb)
58.8
59
59.2
59.4
59.6
Position on chr4 (Mb)

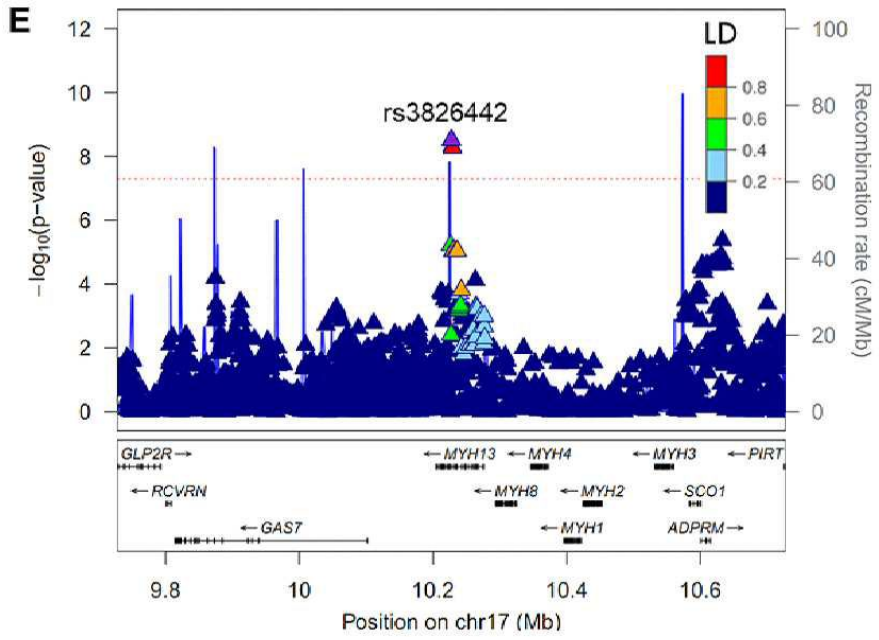
도면6


C
LD
rs7779989
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
ABCB5
SP4
DNAH11
SP8
RPL23P8
20.8
21
21.2
21.4
21.6
Position on chr7 (Mb)

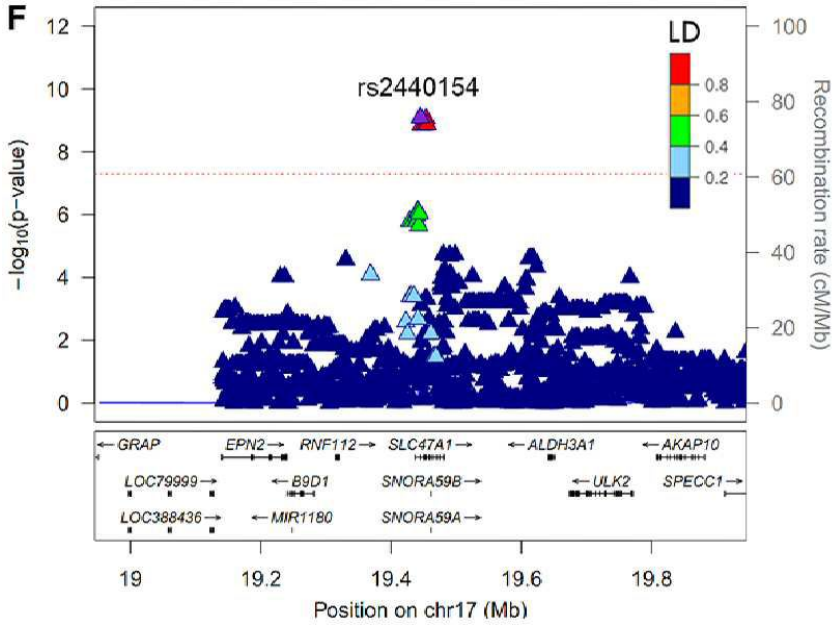
도면7


D
LD
rs12935558
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
SLC7A5
BANP
LOC101928880
ZNF469
ZC3H18
MIR6775
LOC400553
ZFPM1
IL17C
CA5A
MIR5189
CYBA
88
88.2
88.4
88.6
88.8
Position on chr16 (Mb)

도면8



도면9



서 열 목 록

<110> Industry Academic Cooperation Foundation Hallym University

<120> rs2440154 marker composition for diagnosing cerebral aneurysm and method of use thereof

<130> P214560

<160> 6

<170> KoPatentIn 3.0

<210> 1

<211> 201

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> y = c or t

<400> 1

aagtatttgg taaaattcag cagtaaattc accaggtcct tggctattct ttgatgggag 60

acattttatt catagttcta tcttgttact taatattgat ygattcaggt tttatatttc 120

ttcatggttc attcttggta ggtggtatgt gtttagaaat ttatccatta cctctaagtt 180

ttcaaattta gtggtacata g 201

<210> 2

<211> 201

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> r = a or g

<400> 2

tccctaaaaa gggagaaatt acaaacattt tttcctagaa atggagagga cataatattt 60

aatgctattc tgatggcttt gcattgaaat tggtcaagct ractaggctg agattatttt 120

ccaaaatttt ctttcaagtt acagtgggac acaagagaaa ttcttgtgtg agttttggag 180

ggcagaggtg atactcatac t 201

<210> 3

<211> 201

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> r = a or g

<400> 3

ttctgctctt tgaacccatc tctccaggtt gaaaccatct tggcaaagta ttatcagccg 60

gtgtcttcgg ctagatggac ctttcaagcc ctcctaaggt ragaaaatgt tatttttccc 120

ttttatgtgt tgtggttcat tggatagtat ttgtcatctt gtcacataat aggaatgacg 180

ttcacaatgt tgaaatacgc c 201

<210> 4

<211> 200

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> r = a or g

<400> 4

caaggtaaca attcacttca tcccatgtat cattgcataa gttctcctgg gccccgttgg 60

ctggggacgt ggacacgtgt gatttctgct ctgcaccctc rccagcgctg ttctttccag 120

tgggcctttc cagccgggat cctgcccact tccttgtgcc cagctttcca gagcccctca 180

gcccctccaa gggcccacca 200

<210> 5

<211> 201

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> y = c or t

<400> 5

ttgattggcc tctttgaaaa tgctacaaag aaagggtatc accctctaaa agcttattat 60

ctacacctaa gccaagaaat gggtcatgca cacacatgca ygcacatgtg cacacactgc 120

acacacacgt gtacataaac ttagatatca cccataaaca catttttct ttttagatgg 180

agtttctctc tgtggcccag g 201

<210> 6

<211> 201

<212> DNA

<213> Homo sapiens

<220><221> allele

<222> (101)

<223> r = a or g

<400> 6

aagtatttgg taaaattcag cagtaaattc accaggtcct tggctattct ttgatgggag 60

acattttatt catagttcta tcttgttact taatattgat ygattcaggt tttatatttc 120

ttcatggttc attcttggta ggtggtatgt gtttagaaat ttatccatta cctctaagtt 180

ttcaaattta gtggtacata g 201