



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0040343  
(43) 공개일자 2020년04월20일

(51) 국제특허분류(Int. Cl.)

C12Q 1/6883 (2018.01)

(52) CPC특허분류

C12Q 1/6883 (2018.05)

C12Q 2600/156 (2013.01)

(21) 출원번호 10-2018-0119891

(22) 출원일자 2018년10월08일

심사청구일자 2018년10월08일

(71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)

(72) 발명자

이상열

서울특별시 강남구 역삼로 307 205동 1301호 (역삼동, 역삼아이파크아파트)

우정택

서울특별시 강남구 압구정로 151 신현대 아파트 125동 704호

(74) 대리인

특허법인한얼

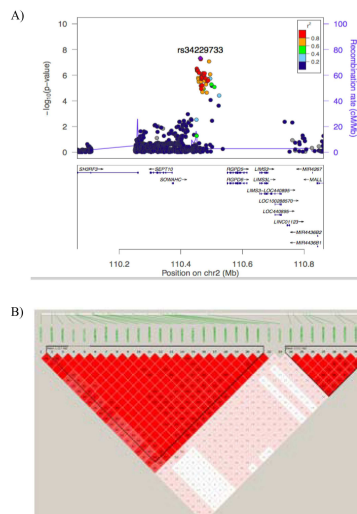
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 당뇨 망막증 진단용 마커 및 이의 용도

(57) 요약

본 발명은 당뇨 망막증 진단용 SNP 마커, 상기 SNP 마커를 검출 또는 증폭할 수 있는 재제를 포함하는, 당뇨 망막증 진단용 조성물, 상기 조성물을 포함하는 당뇨 망막증 진단용 키트, 및 (a) 생체 시료의 DNA로부터 제1항의 당뇨 망막증 진단용 SNP 마커의 다형성 부위의 서열번호 1의 201번째 염기를 증폭하거나 검출하는 단계; 및 (b) 상기 (a) 단계의 증폭 또는 검출된 다형성 부위의 염기를 결정하는 단계를 포함하는, 당뇨 망막증 진단 방법에 관한 것이다. 본 발명에서 제공되는 당뇨 망막증 진단용 조성물은 당뇨 망막증 감수성 유전자의 위치와 아무 밀접하게 연관된 마커를 이용하여 효율적으로 당뇨 망막증 위험도를 진단할 수 있으므로, 당뇨 망막증 조기 진단 및 치료를 위해 유용할 수 있다.

대표도 - 도1



(52) CPC특허분류

C12Q 2600/158 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 HI10C2020  
부처명 보건복지부  
연구관리전문기관 보건산업진흥원  
연구사업명 보건의료기술진흥산업  
연구과제명 제2형 당뇨병 임상연구센터  
기 여 율 1/2  
주관기관 경희의료원  
연구기간 2005.05.01 ~ 2014.03.31

이 발명을 지원한 국가연구개발사업

과제고유번호 HI16C2048  
부처명 보건복지부  
연구관리전문기관 보건산업진흥원  
연구사업명 포스트게놈다부처유전체사업  
연구과제명 대사성질환의 새로운 유전인자 발굴을 위한 유전자-환경 상호작용 연구  
기 여 율 1/2  
주관기관 경희의료원  
연구기간 2016.07.01 ~ 2019.04.30

---

## 명세서

### 청구범위

#### 청구항 1

서열번호 1의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기가 A 또는 G이고, 상기 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 것인, 당뇨 망막증(diabetic retinopathy) 진단용 SNP(Single Nucleotide Polymorphism) 마커 조성물.

#### 청구항 2

제1항에 있어서,

상기 조성물은 서열번호 2 내지 23의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 추가로 포함하는 것인, 당뇨 망막증 진단용 SNP 마커 조성물.

#### 청구항 3

제1항에 있어서,

상기 조성물은 서열번호 2 내지 31의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 추가로 포함하는 것인, 당뇨 망막증 진단용 SNP 마커 조성물.

#### 청구항 4

제1항에 있어서,

서열번호 1로 기재된 rs17197569의 201번째 염기가 A인 경우, 당뇨 망막증 발병 위험도가 높은 것으로 진단하는 것인, SNP 마커 조성물.

#### 청구항 5

제2항에 있어서,

서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우;  
 서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우;  
 서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우;  
 서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우;  
 서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우;  
 서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우;  
 서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우;  
 서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우;

서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우;  
 서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우;  
 서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우;  
 서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우;  
 서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우;  
 서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우;  
 서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우;  
 서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우;  
 서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우;  
 서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우;  
 서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우;  
 서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우;  
 서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우; 및  
 서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우, 당뇨 망막증 발병 위험도가 높은 것으로 진단하는 것인, SNP 마커 조성물.

## 청구항 6

제3항에 있어서,

서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우;  
 서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우;  
 서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우;  
 서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우;  
 서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우;  
 서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우;  
 서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우;  
 서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우;  
 서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우;  
 서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우;  
 서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우;  
 서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우;  
 서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우;  
 서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우;  
 서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우;  
 서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우;  
 서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우;  
 서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우;

서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우;  
 서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우;  
 서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우;  
 서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우;  
 서열번호 24로 기재된 rs34229733의 201번째 염기가 T인 경우;  
 서열번호 25로 기재된 rs12603633의 201번째 염기가 G인 경우;  
 서열번호 26으로 기재된 rs7359572의 201번째 염기가 G인 경우;  
 서열번호 27로 기재된 rs34845949의 201번째 염기가 C인 경우;  
 서열번호 28로 기재된 rs2734363의 201번째 염기가 A인 경우;  
 서열번호 29로 기재된 rs200789의 201번째 염기가 G인 경우;  
 서열번호 30으로 기재된 rs72624872의 201번째 염기가 A인 경우; 및  
 서열번호 31로 기재된 rs7764150의 201번째 염기가 T인 경우, 당뇨 망막증 발병 위험도가 높은 것으로 진단하는 것인, SNP 마커 조성물.

#### 청구항 7

제1항 내지 제3항 중 어느 한 항의 당뇨 망막증 진단용 SNP 마커를 검출 또는 증폭할 수 있는 제제를 포함하는 것인, 당뇨 망막증 진단용 조성물.

#### 청구항 8

제7항에 있어서,

상기 제제는 상기 SNP 마커를 검출 또는 증폭할 수 있는 프라이머 세트 또는 프로브인 것인, 당뇨 망막증 진단용 조성물.

#### 청구항 9

제7항 또는 제8항의 조성물을 포함하는 것인, 당뇨 망막증 진단용 키트.

#### 청구항 10

- (a) 생체 시료의 DNA로부터 서열번호 1의 201번째 염기를 증폭하거나 검출하는 단계; 및
- (b) 상기 (a) 단계의 증폭 또는 검출된 다형성 부위의 염기를 결정하는 단계를 포함하는, 당뇨 망막증 진단을 위한 정보 제공 방법.

#### 청구항 11

제 10항에 있어서,

상기 (a) 단계는 서열번호 2 내지 23의 201번째 염기를 증폭하거나 검출하는 단계를 추가로 포함하는 것인, 당뇨 망막증 진단을 위한 정보 제공 방법.

## 청구항 12

제 10항에 있어서,

상기 (a) 단계는 서열번호 2 내지 31의 201번째 염기를 증폭하거나 검출하는 단계를 추가로 포함하는 것인, 당뇨병 망막증 진단을 위한 정보 제공 방법.

## 청구항 13

제10항에 있어서,

상기 (b) 단계에서 결정된 염기서열 중,

서열번호 1로 기재된 rs17197569의 201번째 염기가 A인 경우, 당뇨병 망막증 발병 위험도가 높은 것으로 진단하는 것인, 진단을 위한 정보 제공 방법.

## 청구항 14

제11항에 있어서,

상기 (b) 단계에서 결정된 염기서열 중,

서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우;

서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우;

서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우;

서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우;

서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우;

서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우;

서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우;

서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우;

서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우;

서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우;

서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우;

서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우;

서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우;

서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우;

서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우;

서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우;

서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우;

서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우;

서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우;

서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우;

서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우; 및

서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우, 당뇨병 망막증 발병 위험도가 높은 것으로 진단

하는 것인, 진단을 위한 정보 제공 방법.

## 청구항 15

제12항에 있어서,

상기 (b) 단계에서 결정된 염기서열 중,

서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우;

서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우;

서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우;

서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우;

서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우;

서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우;

서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우;

서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우;

서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우;

서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우;

서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우;

서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우;

서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우;

서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우;

서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우;

서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우;

서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우;

서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우;

서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우;

서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우;

서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우;

서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우;

서열번호 24로 기재된 rs34229733의 201번째 염기가 T인 경우;

서열번호 25로 기재된 rs12603633의 201번째 염기가 G인 경우;

서열번호 26으로 기재된 rs7359572의 201번째 염기가 G인 경우;

서열번호 27로 기재된 rs34845949의 201번째 염기가 C인 경우;

서열번호 28로 기재된 rs2734363의 201번째 염기가 A인 경우;

서열번호 29로 기재된 rs200789의 201번째 염기가 G인 경우;

서열번호 30으로 기재된 rs72624872의 201번째 염기가 A인 경우; 및

서열번호 31로 기재된 rs7764150의 201번째 염기가 T인 경우, 당뇨 망막증 발병 위험도가 높은 것으로 진단하는 것인, 진단을 위한 정보 제공 방법.

## 청구항 16

서열번호 1 내지 31의 염기서열로 구성되는, 당뇨 망막증 감수성 형질을 나타내는 폴리뉴클레오티드.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 당뇨 망막증 진단용 SNP 마커, 상기 SNP 마커를 검출 또는 증폭할 수 있는 체제를 포함하는, 당뇨 망막증 진단용 조성물, 상기 조성물을 포함하는 당뇨 망막증 진단용 키트, 및 상기 SNP 마커를 이용한 당뇨 망막증 진단 방법에 관한 것이다.

### 배경 기술

[0003] 당뇨병은 인슐린의 생성 혹은 이용 과정의 이상으로 혈당이 증가하고 그에 따른 다양한 급성, 만성 합병증을 동반하는 심각한 질환이다. 미국의 경우, 20 세 이상 인구의 9.6% (2000만 명 이상)가 당뇨를 가지고 있고, 발병 위험이 높은 당뇨 전단계(pre-diabetes) 환자가 5000만 명 이상으로 추정 된다(2005년 미국 national diabetes fact sheet). 2002년 직접, 간접적인 의료비용으로 1320억 달러가 당뇨와 관련되어 지출되었다.

[0004] 우리나라의 경우, 2005년 질병관리본부에서 시행한 국민건강영양조사에 따르면 30세 이상 남자의 9.0%, 여자의 7.2% 가 당뇨병 환자로 나타났다. 대한 당뇨병학회가 발간한 '2007년 한국인 당뇨병 연구보고서'에 따르면, 당뇨의 유병률은 약 8%이고 매년 10%의 새로운 환자가 발생하고 있으며 당뇨 관련 의료비가 건강보험 비용의 약 20% (약 3조원)를 차지한다. 현재 400-500만 명으로 추산되는 당뇨 환자 수의 최근의 빠른 증가 속도를 고려할 때 10-20년 후 1000만 명에 이를 수 있다.

[0005] 당뇨병의 이환 기간이 길어짐에 따라 전신의 다양한 합병증을 동반하게 되는데, 대표적으로 심혈관계 질환, 당뇨신증, 당뇨신경병증, 당뇨 망막증이 발생하게 된다. 당뇨 망막증 (Diabetic Retinopathy, DR) 은 당뇨 환자에게서 당뇨 진단 10년 내에 60% 이상에서, 20년 내에 90% 이상에서 나타난다.

[0006] 당뇨 망막증은 당뇨병의 미세혈관 합병증 (microangiopathy)의 하나로 망막 혈관의 투과성 변화와 혈관 폐쇄, 허혈 변화(ischemia), 신생혈관 생성 (neovascularization), 그리고 이에 따른 섬유혈관 증식(fibrovascular proliferation) 이 특징이다.

[0007] 당뇨 망막증은 성인의 최대의 후천적 실명 원인으로, 미국의 경우 매년 12,000명에서 24,000명이 당뇨로 인해 실명에 이르고 있다. 미국의 연구에서, 당뇨 망막증의 유병률은 당뇨 환자의 40% 정도로 추산되며 8% 정도는 실명에 이를 수 있는 심각한 상태인 것으로 보고되고 있다.

[0008] 당뇨 망막증 진행 정도에 따라 초기의 비증식성 당뇨망병증(NPDR, non-proliferative diabetic retinopathy) 과 후기의 증식성 당뇨 망막증(PDR, proliferative diabetic retinopathy)으로 구분한다.

[0009] 비증식성 당뇨 망막증(NPDR)은 망막 모세혈관의 폐쇄 및 투과성 변화 등으로 망막출혈, 미세혈관류(microaneurysm), 삼출물(exudate), 망막 부종(edema) 등이 나타나면서 조금씩 시력이 떨어지게 된다. 또한 황반부의 부종 (DME, 당뇨황반부종) 을 동반하게 되면 이 시기에서도 심각한 시력 저하를 보일 수 있다.

[0010] 증식성 당뇨 망막증(PDR)은 망막의 광범위한 혈관 폐쇄에 따르는 허혈 (ischemia) 상태로 인해 신생혈관(neovascularization) 이 증식하는 단계이다. 이러한 증식은 망막에서 유리체로 진행되고 섬유혈관 증식이 일어나 견인막에 의해 유리체출혈(vitreous hemorrhage)이나 망막이 원래 부착 부위에서 떨어지는 견인망막박리(tractional retinal detachment), 신생혈관녹내장 등의 합병증이 발생해 실명이 진행되는 단계이다.

[0011] 레이저 치료 혹은 유리체 수술적 치료에도 불구하고 병증이 계속 진행되어 실명에 이르는 환자가 아직도 많아, 당뇨망막의 조기발견과 진행 억제, 고위험군에 대한 조기 치료에 대한 필요성이 증가하고 있으나 당뇨 망막증의 병인은 아직 정확하게 밝혀지지 않는다고, 당뇨에 의한 망막증의 진행 정도를 판단하는 바이오마커도 매우 한정적이다.

- [0012] 현재까지 당뇨 망막증 연구는 유리체의 개별 단백질에 대한 생화학 및 분자생물학적 연구를 중심으로 주로 이루어지고 있다. 또한 당뇨 망막증의 단백질체 연구도 환자 유리체에서 단백질을 2-DE 및 Mass spectrometry로 동정하는 유리체 단백질체의 Profiling(Discovery) 단계의 연구이다. 이들 유리체 단백질들이 혈액에서 발현이 되는지 혹은 이들을 임상적인 바이오마커로 이용할 수 있는 지에 대한 확인(verification) 및 검증(validation) 연구는 거의 이루어지지 않은 상태이다.
- [0013] 따라서, 임상적으로 특이성과 민감도가 높은 신규 진단 마커와 이를 검출할 수 있는 키트를 개발하여, 당뇨 망막증의 조기 진단 및 진행 정도를 손쉽게 미리 예측할 필요가 있다.
- [0014] 이러한 배경 하에, 본 발명자들은 당뇨 망막증의 조기 진단에 유용한 마커를 개발하기 위해 예의 노력한 결과, 당뇨 망막증 연관 유전자를 확인하고, 당뇨 망막증과 연관성이 높은 SNP를 발굴하여, 본 발명을 완성하였다.

## 발명의 내용

### 해결하려는 과제

- [0016] 본 발명의 하나의 목적은 서열번호 1의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기가 A 또는 G이고, 상기 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 것인, 당뇨 망막증(diabetic retinopathy) 진단용 SNP(Single Nucleotide Polymorphism) 마커 조성물을 제공하는 것이다.
- [0017] 본 발명의 다른 하나의 목적은 상기 당뇨 망막증 진단용 SNP 마커를 검출 또는 증폭할 수 있는 재제를 포함하는 것인, 당뇨 망막증 진단용 조성물을 제공하는 것이다.
- [0018] 본 발명의 또 다른 하나의 목적은 상기 당뇨 망막증 진단용 조성물을 포함하는 것인, 당뇨 망막증 진단용 키트를 제공하는 것이다.
- [0019] 본 발명의 또 다른 하나의 목적은 (a) 생체 시료의 DNA로부터 상기 당뇨 망막증 진단용 SNP 마커의 다형성 부위인 서열번호 1의 201번째 염기를 증폭하거나 검출하는 단계; 및 (b) 상기 (a) 단계의 증폭 또는 검출된 다형성 부위의 염기를 결정하는 단계를 포함하는, 당뇨 망막증 진단 방법을 제공하는 것이다.
- [0020] 본 발명의 또 다른 하나의 목적은 서열번호 1 또는 2의 염기서열로 구성되는, 당뇨 망막증 감수성 형질을 나타내는 폴리뉴클레오티드를 제공하는 것이다.

### 과제의 해결 수단

- [0022] 상기의 목적을 달성하기 위한 본 발명의 하나의 양태는, 서열번호 1의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기가 A 또는 G이고, 상기 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 것인, 당뇨 망막증 진단용 SNP 마커 조성물을 제공한다.
- [0023] 상기 마커는 *SOWAHC*과 *RGPD5*유전자 사이에 위치하는 SNP인 rs17197569 부위를 포함하는 전체 또는 일부 폴리뉴클레오티드일 수 있고, 구체적으로 서열번호 1로 구성된 폴리뉴클레오티드의 전체 또는 일부 폴리뉴클레오티드일 수 있으며, 보다 구체적으로 서열번호 1의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기가 A 또는 G이고, 상기 201번째 염기를 포함하는 5 내지 200개, 구체적으로 5 내지 150개, 보다 구체적으로 5 내지 120개, 보다 더 구체적으로 5 내지 100개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 당뇨 망막증 진단용 SNP 마커일 수 있다. 또한, 상기 마커는 201번째 염기가 A인 경우 당뇨 망막증 발병 위험도가 높은 것일 수 있다.
- [0024] 상기 조성물은 서열번호 2 내지 23의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 추가로 포함할 수 있다.
- [0025] 구체적으로, 상기 조성물은 서열번호 2 내지 31의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기를 포함하는 5 내지 200개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를

추가로 포함할 수 있다.

[0026]

상기 서열번호 2의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A, C 또는 G이고, 서열번호 3의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 4의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 5의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A, C 또는 T이고, 서열번호 6의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C 또는 T이고, 서열번호 7의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A, G 또는 T이고, 서열번호 8의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A, C 또는 G이고, 서열번호 9의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C 또는 T이고, 서열번호 10의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 C이고, 서열번호 11의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 C이고, 서열번호 12의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 13의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 14의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 15의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 16의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 17의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C, G 또는 T이고, 서열번호 18의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 19의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 20의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C 또는 T이고, 서열번호 21의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 G 또는 T이고, 서열번호 22의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 C이고, 서열번호 23의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A, C 또는 G이고, 서열번호 24의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 G 또는 T이고, 서열번호 25의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 26의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 27의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C 또는 T이고, 서열번호 28의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 29의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C 또는 T이고, 서열번호 30의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 A 또는 G이고, 서열번호 31의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기는 C, G 또는 T이다.

[0027]

상기 서열번호 2 내지 31의 염기서열로 구성되는 폴리뉴클레오티드는 SNP 부위를 포함하는 전체 또는 일부 폴리뉴클레오티드일 수 있고, 구체적으로 서열번호 2 내지 31로 구성된 폴리뉴클레오티드의 전체 또는 일부 폴리뉴클레오티드일 수 있으며, 보다 구체적으로 서열번호 2 내지 31의 염기서열로 구성되는 폴리뉴클레오티드에서 201번째 염기에 SNP부위를 포함하고, 상기 201번째 염기를 포함하는 5 내지 200개, 구체적으로 5 내지 150개, 보다 구체적으로 5 내지 120개, 보다 더 구체적으로 5 내지 100개의 연속적인 염기로 구성되는 폴리뉴클레오티드, 또는 이에 상보적인 폴리뉴클레오티드를 포함하는 당뇨 망막증 진단용 SNP 마커일 수 있다. 또한 상기 마커는 서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우; 서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우; 서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우; 서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우; 서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우; 서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우; 서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우; 서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우; 서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우; 서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우; 서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우; 서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우; 서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우; 서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우; 서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우; 서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우; 서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우; 서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우; 서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우; 서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우; 서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우; 서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우; 서열번호 24로 기재된 rs34229733의 201번째 염기가 T인 경우; 서열번호 25로 기재된 rs12603633의 201번째 염기가 G인 경우; 서열번호 26으로 기재된 rs7359572의 201번째 염기가 G인 경우; 서열번호 27로 기재된 rs34845949의 201번째 염기가 C인 경우; 서열번호 28로 기재된 rs2734363의 201번째 염기가 A인 경우; 서열번호 29로 기재된 rs200789의 201번째 염기가 G인 경우; 서열번호 30으로 기재된 rs72624872의 201번째 염기가 A인 경우; 및 서열번호 31로 기재된 rs7764150의 201번째 염기가 T인 경우, 당뇨 망막증 발병 위험도가 높은 것일 수 있다.

- [0028] 본 발명에서 용어, "당뇨 망막증(diabetic retinopathy)"은 당뇨병의 미세혈관 합병증 (microangiopathy)의 하나로 망막 혈관의 투과성 변화와 혈관 폐쇄, 허혈 변화 (ischemia), 신생혈관 생성 (neovascularization), 그리고 이에 따른 섬유혈관 증식 (fibrovascular proliferation)을 특징으로 하며, 당뇨 망막증 진행 정도에 따라 초기의 비증식성 당뇨망병증(NPDR, non-proliferative diabetic retinopathy)과 후기의 증식성 당뇨 망막증(PDR, proliferative diabetic retinopathy)으로 구분한다.
- [0029] 비증식성 당뇨 망막증(NPDR)은 망막 모세혈관의 폐쇄 및 투과성 변화 등으로 망막출혈, 미세혈관류 (microaneurysm), 삼출물(exudate), 망막 부종(edema) 등이 나타나면서 조금씩 시력이 떨어지게 된다. 또한 황반부의 부종 (DME, 당뇨황반부종)을 동반하게 되면 이 시기에서도 심각한 시력 저하를 보일 수 있다.
- [0030] 증식성 당뇨 망막증(PDR)은 망막의 광범위한 혈관 폐쇄에 따르는 허혈 (ischemia) 상태로 인해 신생혈관 (neovascularization)이 증식하는 단계이다. 이러한 증식은 망막에서 유리체로 진행되고 섬유혈관 증식이 일어나 견인막에 의해 유리체출혈(vitreous hemorrhage)이나 망막이 원래 부착 부위에서 떨어지는 견인망막박리 (tractional retinal detachment), 신생혈관녹내장 등의 합병증이 발생해 실명이 진행되는 단계이다.
- [0031] 본 발명에서 용어, "진단"은 병리 상태의 존재 또는 특징을 확인하는 일련의 행위를 의미한다.
- [0032] 본 발명의 목적상, 상기 진단은 당뇨 망막증의 발병여부를 확인하는 행위를 의미하는 것으로 해석될 수 있다.
- [0033] 본 발명에서 용어, "SNP 마커"란, 개체 또는 종을 식별하기 위해 이용되는 DNA 서열 상의 단일 염기 다형성 대립유전자 염기쌍을 의미한다. SNP는 비교적 그 빈도가 높고 안정하며 유전체 전체에 분포되어 있고 이에 의하여 개체의 유전적 다양성이 발생하므로, SNP 마커는 개체 간의 유전적 근접성을 알려주는 지표의 역할을 할 수 있다. SNP 마커는 일반적으로 단일 염기 다형성에 수반되는 표현형의 변화를 포함하지만 경우에 따라 그렇지 않을 수도 있다. 본 발명의 SNP 마커의 경우 아미노산 서열의 변이 또는 당뇨 망막증 감수성과 같은 개체의 표현형의 차이를 나타낼 수 있다.
- [0034] 본 발명에서 용어 "다형성(polymorphism)"이란, 하나의 유전자 좌위(locus)에 두 가지 이상의 대립유전자 (allele)가 존재하는 경우를 의미하며 다형성 부위 중에서, 개체에 따라 단일 염기만이 다른 것을 단일 염기 다형성이라 한다. 바람직한 다형성 마커는 선택된 집단에서 1% 이상, 더욱 바람직하게는 5% 또는 10% 이상의 발생 빈도를 나타내는 두 가지 이상의 대립유전자를 가진다.
- [0035] 본 발명에서 용어, "대립유전자(allele)"란, 상동염색체의 동일한 유전자좌위에 존재하는 한 유전자의 여러 타 입을 의미한다. 대립유전자는 다형성을 나타내는데 사용되기도 하며, 예컨대, SNP는 두 종류의 대립인자 (biallele)를 갖는다.
- [0036] 본 발명에서 용어, "개체"란, 당뇨 망막증 위험도를 확인하고자 하는 대상을 의미하며, 상기 대상으로부터 얻어진 검체를 이용하여, 상기 SNP 마커의 유전자형을 분석함으로써 상기 당뇨 망막증을 진단할 수 있다. 상기 검체로는 분리된 조직, 또는 분리된 세포와 같은 시료 등이 될 수 있으나, 이에 특별히 제한되지는 않는다.
- [0038] 본 발명의 다른 하나의 양태는, 상기 상기 당뇨 망막증 진단용 SNP 마커를 검출 또는 증폭할 수 있는 체제를 포함하는 것인, 당뇨 망막증 진단용 조성물을 제공한다.
- [0039] 본 발명에서 용어, "SNP 마커를 검출 또는 증폭할 수 있는 체제"란, 상기과 같은 유전자의 다형성 부위를 검출 또는 증폭을 통해 확인하여 당뇨 망막증을 진단할 수 있는 조성물을 의미하며, 바람직하게는 상기 SNP 마커의 폴리뉴클레오티드를 특이적으로 검출 또는 증폭할 수 있는 프라이머 세트 또는 프로브를 의미한다.
- [0040] 상기 SNP 마커 증폭에 사용되는 프라이머는, 적절한 버퍼 중의 적절한 조건(예를 들면, 4개의 다른 뉴클레오타이드 트리포스페이트 및 DNA, RNA 폴리머라제 또는 역전사 효소와 같은 중합제) 및 적당한 온도 하에서 주형-지시 DNA 합성의 시작점으로서 작용할 수 있는 단일가닥 올리고뉴클레오티드가 될 수 있는데, 상기 프라이머의 적절한 길이는 사용 목적에 따라 달라질 수 있으나, 통상 15 내지 30 뉴클레오티드이다. 짧은 프라이머 분자는 일반적으로 주형과 안정한 혼성체를 형성하기 위해서는 더 낮은 온도를 필요로 한다. 상기 프라이머 서열은 상기 SNP 마커와 완전하게 상보적일 필요는 없으나, 상기 SNP 마커와 혼성화 할 정도로 충분히 상보적이어야 한다.
- [0041] 본 발명에서 용어 "프라이머"란, 짧은 자유 3' 말단 수산화기(free 3' hydroxyl group)를 가지는 염기 서열로 상보적인 템플레이트(template)와 염기쌍(base pair)을 형성할 수 있고 주형 가닥 복사를 위한 시작 지점으로 기능을 하는 짧은 서열을 의미하며, 주로 특정 구간을 증폭하는 프라이머 세트의 형태로 사용된다. 프라이머는

적절한 완충용액 및 온도에서 중합반응(즉, DNA 폴리머레이즈 또는 역전사효소)을 위한 시약 및 상이한 4가지 뉴클레오타이드 트리포스페이트의 존재하에서 DNA 합성을 개시할 수 있다. 이때, PCR 조건, 센스 및 안티센스 프라이머의 길이는 당업계에 공지된 것을 기초로 변형할 수 있다.

[0042] 본 발명의 용어, "프로브"는 다양한 길이의 DNA 또는 RNA 염기서열로 방사성 표지 또는 형광 표지 등을 통해 표지되어 있으며, 단일가닥의 염기서열로 구성되어 상보적인 서열을 가진 단일가닥 DNA 또는 RNA에 특이적으로 결합하여 혼성화하는 것이 특징이며, 혼성화한 프로브의 표지 신호를 탐지하는 방식을 통해 타겟을 검출할 수 있다.

[0043] 본 발명의 프라이머 또는 프로브는 포스포아미다이트 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 뉴클레오타이드 간의 변형, 예를 들면, 하전되지 않은 연결체(예: 메틸 포스포네이트, 포스포트리에스테르, 포스포아미데이트, 카바메이트 등) 또는 하전된 연결체(예: 포스포로티오에이트, 포스포로디티오에이트 등)로의 변형이 있다.

[0044] 본 발명의 또 다른 하나의 양태는, 상기 당뇨 망막증 진단용 조성물을 포함하는 것인, 당뇨 망막증 진단용 키트를 제공한다.

[0045] 본 발명의 키트는 전기영동 키트 또는 DNA 칩 키트일 수 있다.

[0046] 본 발명의 전기영동 키트는 전해질 중에 존재하는 하전(荷電) 입자에 직류전압을 걸면 정의 하전 입자는 음극으로, 부의 하전 입자는 양극으로 향하여 이동하는 특성을 통해 크기가 다른 성분들이 분리되어 분석하는 성분을 파악할 수 있으며, 특히 DNA의 경우, 음전하를 띄기 때문에 양극으로 이동하면서 그 크기에 따라 이동하는 속도가 다르기 때문에 본 발명에서 사용되는 폴리뉴클레오타이드 크기 차이에 따라 이동하는 속도에 변화가 생기는 점을 이용하여 당뇨 망막증을 진단할 수 있다.

[0047] 본 발명의 DNA 칩 키트는 당뇨 망막증 진단용 마커인 SNP 마커를 포함하는 폴리뉴클레오타이드의 서열과 상보적인 염기서열을 포함하는 올리고뉴클레오타이드가 부착된 칩을 이용하여 특이적으로 타겟 DNA와 혼성화하여 결합하는 것을 특징으로 하며, SNP의 염기 변이에 따라 혼성화 수준의 변화가 나타나는 것을 이용하여 당뇨 망막증을 진단할 수 있다. 구체적으로, 본 발명에서 제공하는 당뇨 망막증 진단용 DNA 칩 키트는 서열번호 1의 염기서열로 구성되는 폴리뉴클레오타이드의 201번째 염기를 포함하는 5 내지 200개의 염기로 구성되는 폴리뉴클레오타이드 서열과 상보적인 올리고뉴클레오타이드를 표지자와 함께 칩에 부착하고 타겟 DNA를 칩에 반응시켜 혼성화 수준을 통해 유전자형을 판단하는 키트인 것인, 당뇨 망막증 진단용 키트일 수 있다.

[0049] 본 발명의 또 다른 하나의 양태는, (a) 생체 시료의 DNA로부터 서열번호 1의 201번째 염기를 증폭하거나 검출하는 단계; 및 (b) 상기 (a) 단계의 증폭 또는 검출된 다형성 부위의 염기를 결정하는 단계를 포함하는, 당뇨 망막증 진단을 위한 정보 제공 방법을 제공한다.

[0050] 상기 (a) 단계는 서열번호 2 내지 23의 201번째 염기를 증폭하거나 검출하는 단계를 추가로 포함할 수 있다.

[0051] 구체적으로, 상기 (a) 단계는 서열번호 2 내지 31의 201번째 염기를 증폭하거나 검출하는 단계를 추가로 포함할 수 있다.

[0052] 상기 (a) 단계의 DNA 시료로부터 폴리뉴클레오타이드를 증폭하는 단계는 당업자에게 알려진 어떠한 방법이든 사용 가능하다. 예를 들면, 표적 핵산을 PCR을 통하여 증폭하고 이를 정제하여 얻을 수 있다. 그 외 리가제 연쇄 반응(LCR), 전사증폭(transcription amplification) 및 자가유지 서열 복제 및 핵산에 근거한 서열 증폭(NASBA)이 사용될 수 있다.

[0053] 상기 방법 중 (b)단계의 증폭된 SNP 마커 부위의 염기를 결정하는 방법은 전기영동, 시퀀싱 분석, 마이크로어레이(microarray)에 의한 혼성화, 대립유전자 특이적인 PCR(allele specific PCR), 다이내믹 대립유전자 혼성화 기법(dynamic allele-specific hybridization, DASH), PCR 연장 분석, PCR-SSCP, PCR-RFLP 분석, HRM 분석 또는 TaqMan 기법, SNPlex 플랫폼(Applied Biosystems), 질량 분석법(예를 들면, Sequenom의 MassARRAY 시스템), 미니-시퀀싱(mini-sequencing) 방법, Bio-Plex 시스템(BioRad), CEQ and SNPstream 시스템(Beckman), Molecular Inversion Probe 어레이 기술(예를 들면, Affymetrix GeneChip), 및 BeadArray Technologies(예를 들면, Illumina GoldenGate 및 Infinium 분석법) 등을 이용하여 수행될 수 있으나, 특별히 이에 제한되지는 않

는다. 상기 방법들 또는 본 발명이 속하는 기술분야의 당업자에게 이용 가능한 다른 방법에 의해, 상기 SNP 마커에서의 하나 이상의 대립유전자를 확인할 수 있다.

[0054] 상기 시퀀싱 분석은 염기서열 결정을 위한 통상적인 방법을 사용할 수 있으며, 자동화된 유전자분석기를 이용하여 수행될 수 있다. 또한, 대립유전자 특이적 PCR은 변이 부위가 위치하는 염기를 3' 말단으로 하여 고안한 프라이머를 포함한 프라이머 세트로 상기 변이가 위치하는 DNA 단편을 증폭하는 PCR 방법을 의미한다. 상기 방법의 원리는, 예를 들어, 특정 염기가 A에서 G로 치환된 경우, 상기 A를 3' 말단염기로 포함하는 프라이머 및 적당한 크기의 DNA 단편을 증폭할 수 있는 반대 방향 프라이머를 고안하여 PCR 반응을 수행할 경우, 상기 변이 위치의 염기가 A인 경우에는 증폭반응이 정상적으로 수행되어 원하는 위치의 밴드가 관찰되고, 상기 염기가 G로 치환된 경우에는 프라이머는 주형 DNA에 상보결합할 수 있으나, 3' 말단 쪽이 상보결합을 하지 못함으로써 증폭반응이 제대로 수행되지 않는 점을 이용한 것이다. DASH는 통상적인 방법으로 수행될 수 있고, 바람직하게는 프린스 등에 의한 방법에 의하여 수행될 수 있다.

[0055] 상기 TaqMan 방법은 (1) 원하는 DNA 단편을 증폭할 수 있도록 프라이머 및 TaqMan 탐침을 설계 및 제작하는 단계; (2) 서로 다른 대립유전자의 탐침을 FAM 염료 및 VIC 염료로 표지(Applied Biosystems)하는 단계; (3) 상기 DNA를 주형으로 하고, 상기의 프라이머 및 탐침을 이용하여 PCR을 수행하는 단계; (4) 상기의 PCR 반응이 완성된 후, TaqMan 분석 플레이트를 핵산 분석기로 분석 및 확인하는 단계; 및 (5) 상기 분석결과로부터 단계 (1)의 폴리뉴클레오타이드의 유전자형을 결정하는 단계를 포함한다.

[0056] 한편, PCR 연장 분석은 먼저 단일염기 다형성이 위치하는 염기를 포함하는 DNA 단편을 프라이머 쌍으로 증폭한 다음, 반응에 첨가된 모든 뉴클레오타이드를 탈인산화시킴으로써 불활성화시키고, 여기에 특이적 연장 프라이머, dNTP 혼합물, 디디옥시뉴클레오타이드, 반응 완충액 및 DNA 중합효소를 첨가하여 프라이머 연장반응을 수행함으로써 이루어진다. 이때, 연장 프라이머는 변이 부위가 위치하는 염기의 5' 방향의 바로 인접한 염기를 3' 말단으로 삼으며, dNTP 혼합물에는 디디옥시뉴클레오타이드와 동일한 염기를 갖는 핵산이 제외되고, 상기 디디옥시뉴클레오타이드는 변이를 나타내는 염기 종류 중 하나에서 선택된다. 예를 들어, A에서 G로의 치환이 있는 경우, dGTP, dCTP 및 TTP 혼합물과 ddATP를 반응에 첨가할 경우, 상기 치환이 일어난 염기에서 프라이머는 DNA 중합효소에 의하여 연장되고, 몇 염기가 지난 후 A 염기가 최초로 나타나는 위치에서 ddATP에 의하여 프라이머 연장반응이 종결된다. 만일 상기 치환이 일어나지 않았다면, 그 위치에서 연장반응이 종결되므로, 상기 연장된 프라이머의 길이를 비교함으로써 변이를 나타내는 염기 종류를 판별할 수 있게 된다.

[0057] 이때, 검출방법으로는 연장 프라이머 또는 디디옥시뉴클레오타이드를 형광 표지한 경우에는 일반적인 염기서열 결정에 사용되는 유전자 분석기(예를 들어, ABI사의 Model 3700 등)를 사용하여 형광을 검출함으로써 상기 변이를 검출할 수 있으며, 무-표지된 연장 프라이머 및 디디옥시뉴클레오타이드를 사용할 경우에는 MALDI-TOF(matrix assisted laser desorption ionization-time of flight) 기법을 이용하여 분자량을 측정함으로써 상기 변이를 검출할 수 있다.

[0058] 구체적으로, 상기 (b) 단계에서 결정된 염기서열 중, 서열번호 1로 기재된 rs17197569의 201번째 염기가 A인 경우; 서열번호 2로 기재된 rs6108152의 201번째 염기가 A인 경우; 서열번호 3으로 기재된 rs9671386의 201번째 염기가 G인 경우; 서열번호 4로 기재된 rs35774187의 201번째 염기가 G인 경우; 서열번호 5로 기재된 rs7763869의 201번째 염기가 C인 경우; 서열번호 6으로 기재된 rs8181760의 201번째 염기가 C인 경우; 서열번호 7로 기재된 rs6035258의 201번째 염기가 A인 경우; 서열번호 8로 기재된 rs17741348의 201번째 염기가 A인 경우; 서열번호 9로 기재된 rs8077896의 201번째 염기가 T인 경우; 서열번호 10으로 기재된 rs7730344의 201번째 염기가 A인 경우; 서열번호 11로 기재된 rs200787의 201번째 염기가 T인 경우; 서열번호 12로 기재된 rs200788의 201번째 염기가 T인 경우; 서열번호 13으로 기재된 rs7581991의 201번째 염기가 G인 경우;

[0059] 서열번호 14로 기재된 rs10998731의 201번째 염기가 A인 경우; 서열번호 15로 기재된 rs2007025의 201번째 염기가 G인 경우; 서열번호 16으로 기재된 rs7559515의 201번째 염기가 A인 경우; 서열번호 17로 기재된 rs34273433의 201번째 염기가 T인 경우; 서열번호 18로 기재된 rs10027363의 201번째 염기가 G인 경우; 서열번호 19로 기재된 rs10998725의 201번째 염기가 A인 경우; 서열번호 20으로 기재된 rs1406844의 201번째 염기가 T인 경우; 서열번호 21로 기재된 rs242171의 201번째 염기가 T인 경우; 서열번호 22로 기재된 rs8049793의 201번째 염기가 C인 경우; 서열번호 23으로 기재된 rs41306471의 201번째 염기가 G인 경우; 서열번호 24로 기재된 rs34229733의 201번째 염기가 T인 경우; 서열번호 25로 기재된 rs12603633의 201번째 염기가 G인 경우; 서열번호 26으로 기재된 rs7359572의 201번째 염기가 G인 경우; 서열번호 27로 기재된 rs34845949의 201번째 염기가 C인 경우; 서열번호 28로 기재된 rs2734363의 201번째 염기가 A인 경우; 서열번호 29로 기재된 rs200789의 201번째 염기가 G인

경우; 서열번호 30으로 기재된 rs72624872의 201번째 염기가 A인 경우; 및 서열번호 31로 기재된 rs7764150의 201번째 염기가 T인 경우, 당뇨 망막증 발병 위험도가 높은 것으로 진단할 수 있다. 상기 다형성 부위는 위험 대립 유전자로 판단하며, 상기 대립 유전자의 빈도 가 높을수록 당뇨 망막증 발병 위험이 높다고 판단할 수 있다. 보다 구체적으로, 서열번호 1 내지 23 중 다형성 부위 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 10개 이상, 11개 이상, 12개 이상, 13개 이상, 14개 이상, 15개 이상, 16개 이상, 17개 이상, 18개 이상, 19개 이상, 20개 이상, 21개 이상, 22개 이상, 23개 이상, 24개 이상, 25개 이상, 26개 이상, 27개 이상, 28개 이상, 29개 이상, 30개 이상이 위험 대립 유전자인 경우, 보다 더 구체적으로는 31개 모두 위험 대립 유전자일 경우 당뇨 망막증 발병 위험도가 높다고 판단할 수 있다.

[0061] 본 발명의 또 다른 하나의 양태는, 서열번호 1 내지 서열번호 31의 염기서열로 구성되는, 당뇨 망막증 감수성 형질을 나타내는 폴리뉴클레오티드를 제공한다.

[0062] 상기 당뇨 망막증 감수성 형질을 갖는 SNP 마커를 포함하는 폴리뉴클레오티드는 당뇨 망막증 감수성을 갖는다고 판단할 수 있으며, 상기 폴리뉴클레오티드를 포함하는 유전자 또한 당뇨 망막증 감수성을 가진다고 판단할 수 있다. 게다가 본 발명에서 제공하는 폴리뉴클레오티드는 당뇨 망막증과 연관된 유전자의 주변에 존재하여 상기 폴리뉴클레오티드와 당뇨 망막증 감수성 형질과의 관계는 더욱 밀접한 것으로 판단할 수 있다.

[0063] 따라서, 본 발명의 폴리뉴클레오티드를 포함하는 염기서열을 갖는 개체는 당뇨 망막증 감수성 형질을 나타낼 수 있음을 시사한다.

### 발명의 효과

[0065] 본 발명에서 제공되는 당뇨 망막증 진단용 조성물은 당뇨 망막증 감수성 유전자의 위치와 아무 밀접하게 연관된 마커를 이용하여 효율적으로 당뇨 망막증 위험도를 진단할 수 있으므로, 당뇨 망막증 조기 진단 및 치료를 위해 유용할 수 있다.

### 도면의 간단한 설명

[0067] 도 1은 *SOWAHC-RGPD5*의 염색체 위치와 p-값을 나타내는 지역 연관성 그래프이다; A) *SOWAHC-RGPD5*의 가장 유의적인 rs34229733 ( $p = 3.62 \times 10^{-9}$ ) 및 rs17197569 ( $p = 7.32 \times 10^{-8}$ )는 2 번 염색체에 위치함, B) *SOWAHC-RGPD5*의 연쇄 불균형 그래프.

도 2는 DR 예측을 위한 ROC 곡선 분석 결과를 나타낸 그래프이다; 파란색 곡선은 모든 유의적인 SNP (AUC = 71 %)를 가진 모델에 의해 구현된 ROC를 나타내었고, 녹색은 rs17197569, rs9671386, rs35774187, rs7763869 및 rs6035258 (AUC = 63 %) 모델에 의해 구현된 ROC를 나타내었으며, 마지막으로 붉은색의 ROC는 가장 유의적인 SNP (rs17197569, AUC = 54 %)만을 가진 모델을 나타내었다.

### 발명을 실시하기 위한 구체적인 내용

[0068] 이하, 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.

[0070] 실시예 1. 전장-유전체 연관성 분석을 통한 당뇨 망막증 (diabetic retinopathy, DR) 연관 유전 변이 탐색

[0071] 1-1. 당뇨 망막증 연관 SNP의 초기 탐색

[0073] 한국인에서 비증식성 당뇨 망막증 (non-proliferative DR, NPDR), 증식성 당뇨 망막증 (proliferative DR, PDR) 및 제2형 당뇨 (type 2 diabetes mellitus, T2D)를 포함하는 DR의 유전 변이를 확인하기 위해 전장-유전체 연관성 분석 (genome-wide association study, GWAS)을 수행하였다. 당뇨병 임상연구센터(Korean National Diabetes Program, KNDP)에서 모집한 환자 중 조건에 맞는 대상을 선별하여 676명의 T2D 환자와 703명의 DR 환

자 (588명의 NPDR 및 115 명의 PDR 환자)를 분석하였고, 분석에 참여한 환자들의 임상 특성은 표 1에 나타내었다.

표 1

|           | 초기 집단          |                 |                 | 검증 집단           |                  |                  |
|-----------|----------------|-----------------|-----------------|-----------------|------------------|------------------|
|           | Case           | Control         | Total           | Case            | Control          | Total            |
| 집단 크기 (n) | 703            | 676             | 1379            | 553             | 421              | 974              |
| 남성/여성 (n) | 379/324        | 404/272         | 783/596         | 224/329         | 182/239          | 406/568          |
| 나이 (년)    | 55.90(9.53)    | 52.06<br>(8.95) | 53.97<br>(9.43) | 57.91<br>(9.92) | 55.05<br>(10.53) | 56.68<br>(10.28) |
| HbA1c (%) | 7.97<br>(1.80) | 7.53<br>(1.73)  | 7.76<br>(1.78)  | 7.67<br>(1.57)  | 6.98<br>(1.43)   | 7.37<br>(1.55)   |

GWAS 분석은 Affymetrix Axiom Genome-Wide ASI 어레이를 사용하였으며, 분석에 사용된 SNP 및 삽입/결실 정보는 HapMap, 단일 염기 다형성 데이터베이스 (Single Nucleotide Polymorphism Database, dbSNP) 및 dbSNP의 1000 게놈 프로젝트 정보를 통해 확인한 유전 변이 600,307개 중 필터링을 거친 426,593개의 변이 ( $MAF \geq 0.01$ ,  $call\ rate \geq 0.98$ ,  $HWE\ p \geq 1 \times 10^{-6}$ ,  $MAF\ 0.2 \leq 1000G\ 패널$ )를 사용하였다.

분석 결과, 초기 집단 내 T2D와 T2D 당뇨병 합병증 그룹 (NPDR과 PDR) 사이에서 DR ( $p < 1.0 \times 10^{-5}$ )과 연관된 적어도 하나의 유전 변이를 갖는 21개의 게놈 부위를 확인하였다 (표 2). 표 2에는 각 SNP에 대한 유전 변이 정보가 기재되어 있으며, 해당 SNP 위치에서 원조 서열이 변이 서열(Alt)로 변이되었을 경우, 당뇨 망막증 발병에 대한 위험도가 증가하는 정도를 OR값을 통해 확인할 수 있으며, 해당 위험도가 유의성을 갖는다는 것을  $p$ 값을 통해 확인할 수 있다. 따라서, 표 2에 기재된 23개의 SNP는 모두 충분한 유의성 ( $p < 1.0 \times 10^{-5}$ ) 이상의  $p$ 값을 가지고 있으므로, 각 SNP가 변이 서열을 갖는 경우, 당뇨 망막증에 대한 위험도가 증가하는 것으로 판단할 수 있다.

이 중 가장 유의적인 연관성을 나타낸 변이는 rs17741348 (STX8, 1st  $p = 1.45 \times 10^{-7}$ ) 및 rs17197569 (SOWAHC-RGPD5, 1st  $p = 4.76 \times 10^{-7}$ )였다. 또한, rs17197569 (SOWAHC-RGPD5, 혼합  $p = 9.41 \times 10^{-8}$ )는 NPDR 및 PDR과 비교하여 NPDR과 통계적으로 유의한 전장-유전체 연관성을 나타내었다. rs35774187 (SASH1)과 rs7730344 (ANKRD31-HMGCR)의 두가지 추가 좌위는 DR과의 연관성에 대해 의미있는 수준 ( $p < 2.52 \times 10^{-6}$ )을 나타내었다.

표 2

| Gene                             | Variant    | Chr | Alt | location   | Seq ID | 1st_OR | 1st_p    | 2nd_OR | 2nd_p    | combined_OR | combined_p |
|----------------------------------|------------|-----|-----|------------|--------|--------|----------|--------|----------|-------------|------------|
| <i>SOWAHC-RGPD5</i>              | rs17197569 | 2   | A   | Intergenic | 1      | 0.65   | 4.76E-07 | 0.80   | 2.86E-02 | 0.71        | 7.32E-08   |
| <i>PLCB1</i>                     | rs6108152  | 20  | A   | Intron     | 2      | 0.67   | 5.39E-05 | 0.85   | 1.88E-01 | 0.63        | 1.36E-05   |
| <i>AREL1</i>                     | rs9671386  | 14  | G   | Intron     | 3      | 0.78   | 2.93E-03 | 0.99   | 9.44E-01 | 0.59        | 1.42E-05   |
| <i>SASH1</i>                     | rs35774187 | 6   | G   | Intron     | 4      | 0.68   | 2.52E-06 | 0.93   | 4.76E-01 | 0.77        | 4.03E-05   |
| <i>SASH1</i>                     | rs7763869  | 6   | C   | Intron     | 5      | 0.69   | 5.14E-05 | 0.85   | 1.17E-01 | 0.75        | 4.24E-05   |
| <i>MED13L</i>                    | rs8181760  | 12  | C   | Intron     | 6      | 0.46   | 7.76E-05 | 0.65   | 5.72E-03 | 0.78        | 4.52E-05   |
| <i>SLC24A3</i>                   | rs6035258  | 20  | A   | intron     | 7      | 0.65   | 3.51E-04 | 0.71   | 1.41E-02 | 0.74        | 6.79E-05   |
| <i>STX8</i>                      | rs17741348 | 17  | A   | intron     | 8      | 0.46   | 1.45E-07 | 0.92   | 6.25E-01 | 0.68        | 2.63E-04   |
| <i>STX8</i>                      | rs8077896  | 17  | T   | intron     | 9      | 0.51   | 4.40E-06 | 0.97   | 8.35E-01 | 0.80        | 1.01E-03   |
| <i>ANKRD31-HMG<br/>CR</i>        | rs7730344  | 5   | A   | intergenic | 10     | 1.42   | 9.66E-06 | 1.01   | 9.58E-01 | 1.23        | 5.59E-04   |
| <i>ANKS1B</i>                    | rs200787   | 12  | T   | intron     | 11     | 0.70   | 1.55E-05 | 0.92   | 3.64E-01 | 1.30        | 2.44E-03   |
| <i>ANKS1B</i>                    | rs200788   | 12  | T   | intron     | 12     | 0.71   | 1.91E-05 | 0.88   | 1.66E-01 | 0.79        | 1.15E-04   |
| <i>CNTNAP5</i>                   | rs7581991  | 2   | G   | intron     | 13     | 1.43   | 2.77E-05 | 0.94   | 5.27E-01 | 1.21        | 4.24E-03   |
| <i>HK1</i>                       | rs10998731 | 10  | A   | intron     | 14     | 1.61   | 3.32E-05 | 0.97   | 8.38E-01 | 1.30        | 2.19E-03   |
| <i>F11-AS1</i>                   | rs2007025  | 4   | G   | intron     | 15     | 1.47   | 3.43E-05 | 0.84   | 1.03E-01 | 1.16        | 3.56E-02   |
| <i>CNTNAP5</i>                   | rs7559515  | 2   | A   | intron     | 16     | 1.43   | 3.44E-05 | 0.95   | 6.03E-01 | 1.21        | 3.74E-03   |
| <i>LOC101926928-S<br/>LCO3A1</i> | rs34273433 | 15  | T   | intergenic | 17     | 1.42   | 3.95E-05 | 0.98   | 8.20E-01 | 0.86        | 1.62E-02   |
| <i>IGFBP7-AS1-LP<br/>HNS</i>     | rs10027363 | 4   | G   | intergenic | 18     | 0.63   | 3.99E-05 | 1.04   | 7.42E-01 | 0.79        | 3.74E-03   |
| <i>HK1</i>                       | rs10998725 | 10  | A   | intron     | 19     | 1.60   | 4.02E-05 | 0.98   | 8.87E-01 | 1.54        | 1.26E-03   |
| <i>COL9A1</i>                    | rs1406844  | 6   | T   | intron     | 20     | 1.47   | 7.02E-05 | 0.92   | 4.57E-01 | 1.20        | 1.04E-02   |
| <i>FHIT</i>                      | rs242171   | 3   | T   | intron     | 21     | 0.73   | 7.67E-05 | 0.92   | 3.64E-01 | 0.81        | 3.18E-04   |
| <i>RBFOX1</i>                    | rs8049793  | 16  | C   | intron     | 22     | 0.72   | 8.19E-05 | 0.94   | 5.10E-01 | 0.80        | 3.79E-04   |
| <i>GRIN3A</i>                    | rs41306471 | 9   | G   | intron     | 23     | 2.01   | 8.35E-05 | 1.05   | 8.25E-01 | 3.39        | 4.58E-04   |

[0082]

[0084]

## 1-2. 당뇨 망막증 연관 SNP의 검증

[0086]

실시에 1-1에서 확인한 SNP의 검증을 위해, 1000G 페이즈 3 변이의 45,953,042개로부터 필터링된 6,198,263개의 전가된(imputed) 유전 변이 ( $r^2 \geq 0.7$  및  $MAF \geq 0.01$ )의 연관성을 분석하였으며, T2D와 유의한 연관성이 있는 472 개의 좌위 ( $p < 1 \times 10^{-5}$ )를 확인하였다.

[0087]

상기 472개의 좌위에서 연관성이 높게 나타난 31개의 좌위(표 2의 23개 좌위 + 표 3의 8개 좌위)를 선정하여 독립적인 KNDP 집단에서 각각 T2D 환자 552 명과 T2D 및 DR 환자 421 명 사이의 추가 후속 복제(follow up replication)를 수행하였다. 그 결과, 도 1에 나타낸 바와 같이, DR과 통계적으로 유의미한 전장-유전체 연관성을 보인 31개의 좌위 내에서 하나의 새로운 유전 위험 좌위 (rs34229733)를 확인하였고  $p$  값의 유의 수준은  $< 5 \times 10^{-8}$ 이었다. *SOWAHC-RGPD5*의 유전자간 영역에 존재하는 rs34229733은 DR과 가장 유의한 연관성을 보였다 ( $p = 3.62 \times 10^{-9}$ ). 따라서, rs34229733은 T2D 및 DR 환자에서 질병 좌위 대립 유전자와 인접한 마커의 대립 유전자 사이의 연관 불균형에 대한 집단 기반 연구에서 가장 유의한 것으로 나타났다. 추가된 8개의 좌위에 대한 정보는 표 3에 나타내었다.

표 3

| Gene                | Variant    | Chr | Alt | location   | Seq ID | 1st_OR | 1st_p    | 2nd_OR | 2nd_p   | combined_OR | combined_p |
|---------------------|------------|-----|-----|------------|--------|--------|----------|--------|---------|-------------|------------|
| <i>SOWAHC-RGPD5</i> | rs34229733 | 2   | T   | intergenic | 24     | 0.6299 | 4.66E-08 | 0.774  | 0.01169 | 0.6823      | 3.62E-09   |
| <i>STX8</i>         | rs12603633 | 17  | G   | intron     | 25     | 0.452  | 8.48E-08 | 0.8887 | 0.3875  | 0.6613      | 2.39E-05   |
| <i>STX8</i>         | rs7359572  | 17  | G   | intron     | 26     | 0.4892 | 6.61E-07 | 0.9357 | 0.681   | 0.6513      | 4.59E-05   |
| <i>SASH1</i>        | rs34845949 | 6   | C   | intron     | 27     | 0.6782 | 2.20E-06 | 0.9665 | 0.7207  | 0.7895      | 1.28E-04   |
| <i>FHIT</i>         | rs2734363  | 3   | A   | intron     | 28     | 0.7028 | 9.56E-06 | 0.9181 | 0.3459  | 0.7891      | 7.11E-05   |
| <i>ANKS1B</i>       | rs200789   | 12  | G   | intron     | 29     | 0.7027 | 1.18E-05 | 0.8789 | 0.168   | 0.7757      | 2.94E-05   |
| <i>FHIT</i>         | rs72624872 | 3   | A   | intron     | 30     | 0.7135 | 2.00E-05 | 0.978  | 0.8078  | 0.816       | 6.51E-04   |
| <i>SASH1</i>        | rs7764150  | 6   | T   | intron     | 31     | 0.6812 | 2.31E-05 | 0.864  | 0.1678  | 0.7518      | 3.16E-05   |

[0089]

[0091] 실시예 2. DR 예측을 위한 ROC 곡선 분석

[0093] 본 발명에서 발굴한 SNP의 DR을 예측 수준을 확인하기 위해, SNP들의 GRS를 분석하고 이에 대한 ROC 곡선을 나타내었다. GRS는 초기 및 검증 집단에 의해 얻어진 결합된 OR ( $\beta = \log(\text{OR})$ )의 대수에 대한 각 좌위 (0, 1, 2)에서의 위험 대립 유전자의 수를 곱함으로써 측정되었다. 이러한 가중치는 DR 개시 당시 중요한 위치에 각 위험 대립 유전자의 영향을 제공하였다. 그런 다음 SAS 소프트웨어를 사용하여 측정된 ROC 곡선에서 면적을 측정(area under the ROC curves, AUC)하는 방법으로 유의한 SNP의 효과를 평가하였다.

[0094] 그 결과, 도 2에 나타난 바와 같이, 표 1에 나타난 23개의 모든 SNP가 GRS에서 고려된 경우, 최상의 예측 수준(AUC = 71%,  $\text{OR}_{\text{GRS}} = 1.23$ ,  $\text{CI}_{95\%} = 1.19\text{--}1.27$ )을 얻을 수 있었다. 위험 점수에서 가장 유의한 5 개의 SNP만 유지할 경우, 더 낮은 예측 수준(AUC = 63%,  $\text{OR}_{\text{GRS}} = 1.04$ ,  $\text{CI}_{95\%} = 1.03\text{--}1.05$ )을 나타냈으며, 가장 유의한 SNP만 유지할 경우 더욱 낮은 예측 수준(AUC = 57%,  $\text{OR}_{\text{GRS}} = 1.01$ ,  $\text{CI}_{95\%} = 1.00\text{--}1.01$ )을 얻을 수 있었다.

[0095] 따라서, *SOWAHC-RGPD5*에 위치한 SNP(rs17197569)가 더 큰 연관성(혼합  $p$  값  $\sim 10^{-8}$ , 표 2)을 보였음에도 불구하고 3가지 GRS에 대한 ROC 곡선을 비교할 때 DR 예측에 대한 모든 유의한 SNP를 고려할수록 더욱 정확한 예측을 할 수 있음을 확인하였다.

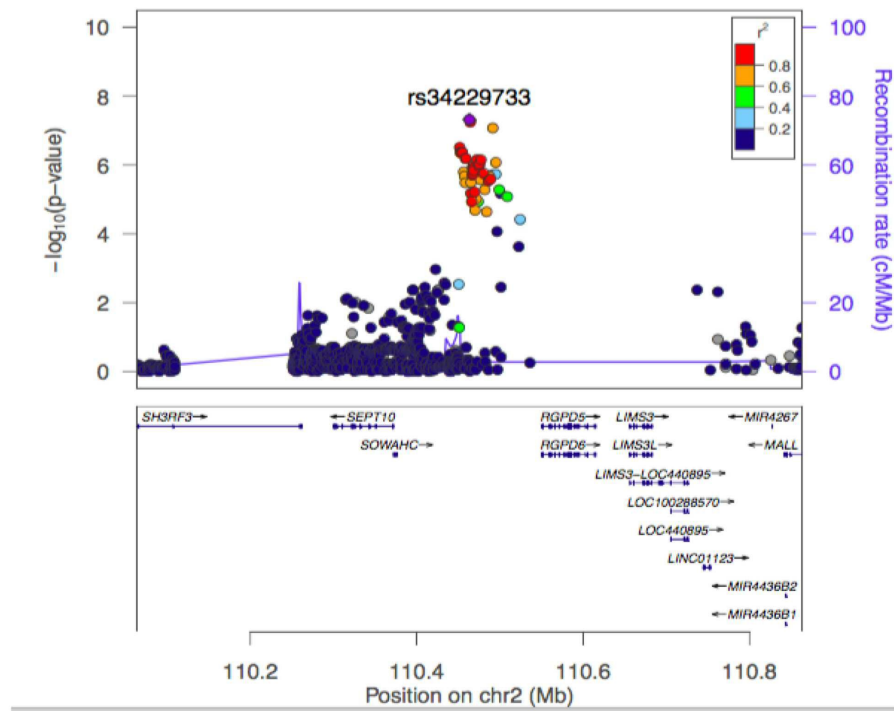
[0097] 종합하면, 본 발명은 특정 유전자에 위치한 변이체가 DR과 관련이 있음을 발견하였으며, 특히 *SOWAHC-RGPD5*의 새로운 유전자형 rs17197569와 전가된 rs34229733가 당뇨 망막증과 큰 연관성을 갖는 SNP임을 발견하였다.

[0099] 본 명세서는 본 발명의 기술 분야에서 통상의 지식을 가진 자이면 충분히 인식하고 유추할 수 있는 내용은 그 상세한 기재를 생략하였으며, 본 명세서에 기재된 구체적인 예시들 이외에 본 발명의 기술적 사상이나 필수적 구성을 변경하지 않는 범위내에서 보다 다양한 변형이 가능하다. 따라서 본 발명은 본 명세서에서 구체적으로 설명하고 예시한 것과 다른 방식으로 실시될 수 있으며, 이는 본 발명의 기술 분야에 통상의 지식을 가진 자이면 이해할 수 있는 사항이다.

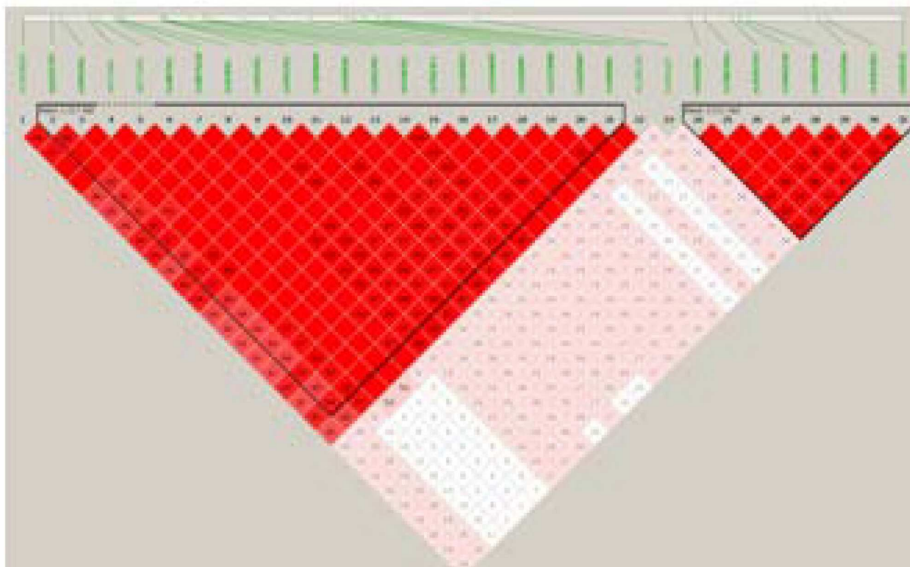
도면

도면1

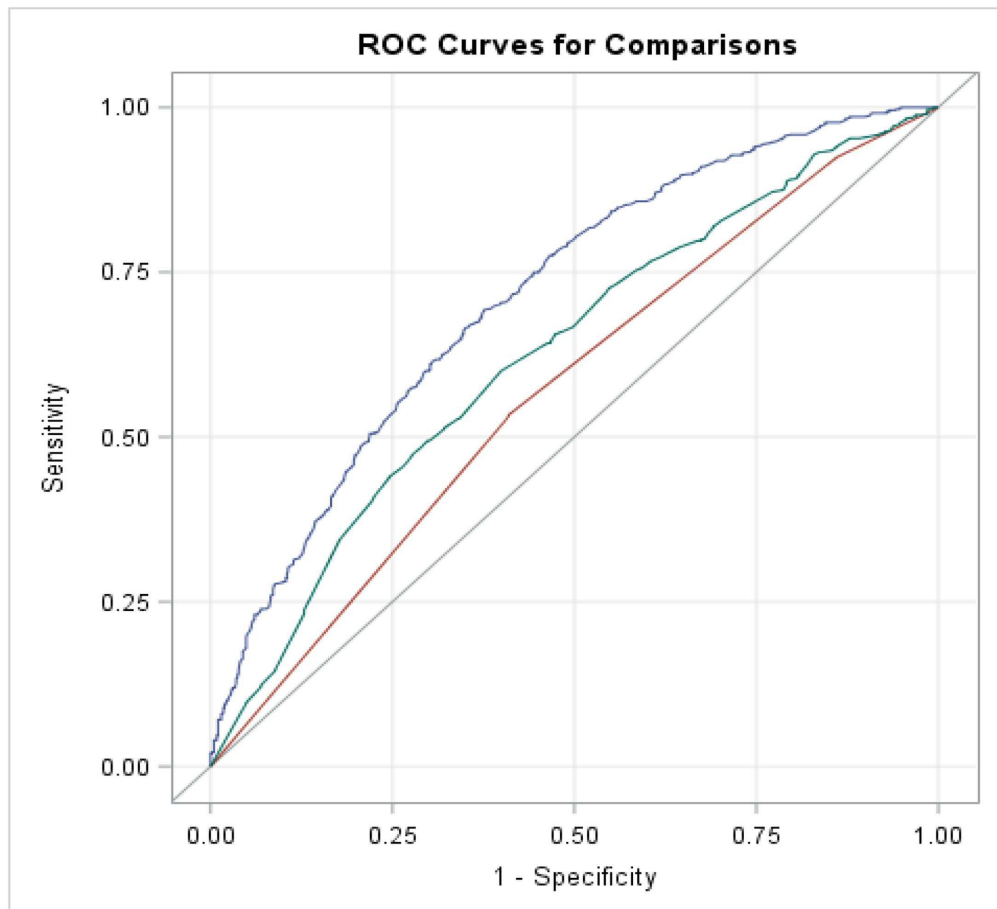
A)



B)



도면2



서열목록

<110> University-Industry Cooperation Group of Kyung Hee University  
 <120> Marker For Diagnosing Diabetic Retinopathy and Use Thereof  
 <130> KPA180935-KR  
 <160> 31  
 <170> KoPatent In 3.0  
 <210> 1  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 1

agggtgggaag gtgatgtggg aggaaagcgg ggagtttagg gaaaggcaag ggcaaaagag 60

cgtttgtgaa tgacacaaac tttagaaact acaaaggagg gaggttcagc ctacacgcct 120

cttgtggctg agggaaagca gtggagctga agatgcagga gacgcgttca caggagatgg 180

gctccaggct cagccttgcc rcttgcttta cagaaggaga gactgggaag atctgtcccc 240

ttgtgcagag gcagcatgca agcacacgag cccaaacccc atgaggagct cctgctgcaa 300  
ctgggggcag tttcttcata aaaggcaagt cggaaaaaaa cacattatct tctctcatag 360  
ggattttccc ttcgaacctt gttaaccag aatatctaag a 401

<210> 2  
<211> 401  
<212> DNA  
<213> Homo sapiens  
<400> 2

attgatccat atgtttatga ttgcatat ttgtatataa attatactc aataaaatgc 60  
tgttaacata aaaaaattgc ccagcttctg agttcttgct taataatatt agtatcacct 120

gttgctgttc tacgttaact caggagctaa attcaaattc agtgtatgac aaaatcttat 180  
ttgtgtgaaa ttactgggac vtgtactttt gcatttaaga aaatgctcca aaaaactgaa 240  
gattctccct tcatgttaga aacttaaaaa taaaataaaa taaaataaag atttttttgg 300  
gaggaaggga atcttaagtt aaaatgtcac aaatctctcc tatcatgtta gtcacttttt 360  
tcttgacttg gcatttgctt gggtgcctta agcctttgtc t 401

<210> 3  
<211> 401  
<212> DNA  
<213> Homo sapiens  
<400> 3

gacaaatgta tttaaagaaa cagaagggtg gcctgatttg ctaacccta tctataatac 60

ttactctttt ctgtaatatt tttcttattt tgaaggta acattaagct ttatacttct 120  
tcaaagctaa ggacctatta aatcaaaatg acagagaaat ttgtgactct tcagatagta 180  
gcaattcact agttcctgga rggagttaga aaacaagaga aaatgttaac attccaggcc 240  
agcacctggg ctctggaaag caagtactc agaaaccca atgggtatgg agacagaaac 300  
attcccttac atgcactcgg aaggccatgc tgtggccac ctcatagggg tccttccaat 360  
cccaggagac ttgcaagac cggggatcca ggttaatttc c 401

<210> 4  
<211> 401  
<212> DNA

<213> Homo sapiens  
<400> 4

gaagagtgag ttgccagttc attaaattct tgiccattaa aaccatttc tctttctggc 60  
cgcttgctca cagtcacccc attctcagat gtactcatcc cagcacattc ttttttttc 120  
attgtggcaa aaaaaaaaaa aaaaaaaaaac caaaacaaac aaaaaaaaaa accacgtaac 180  
ataaaattca cactcccaac rgtttaccba agtgtgcggt gcagcgttgt cagccacatg 240  
cacattgttg agcaacagat tcccagaaca ctccccaca atcctgcatg aaagtgcatt 300  
cattccgata ctctatatc gtgctggatt cctgttcctt atttaaaagt ataactttgc 360  
tctgagttat acttatatta atactttatt gctactttta t 401

<210> 5

<211> 401

<212> DNA

<213> Homo sapiens

<400> 5

cctgctcact aatatctcag ccacccaaac tgccattca tataacataa ttcaccactt 60  
cctcttgacc cttctctaac taactgctct ttgtccacac ttcctctct ctatccttga 120  
gaaacttttc tgcccttcac ccagactttt tcatcatgct gtctctgttt ttttcttga 180  
cccttgattg tcagttgcca hactctgttt ttcaacttct atagtcttcc ctttgtttcc 240  
ctctgtttta gtccattttg tgttgctgtg atagaatacc cgagactggc taatttcatt 300  
tcatttcttt tttttttttt ttggagacag agtctcacc aggctggagt gcagtgggtgc 360

aaatctcggc tcaactgcaac ctctgccttc tgggttcaag t 401

<210> 6

<211> 401

<212> DNA

<213> Homo sapiens

<400> 6

tatccagatc aacaacataa caaagaaaat tataacctaa tggataatct aactttcaag 60  
tagcaagaaa aaaaattact ggttgaactc atctagctta atttttaaat tatatttttt 120  
actgttgacc tcactttaga gggccaaatt gtgtacagaa tacatgctac ccaagaacac 180  
cacagaggtc ctctgggtgt ytccacagag caaaattgaa aacagggcca ataagtaagg 240  
tgcaattaac aaaaaaaaaa aaagtcgtac tggcccagtc ttcggcttcg gctgcaactt 300

aatagtcctt taaattcagt ttctaacc cagagaaatg acacatgaga taagagtgt 360

ttctttccaa aacgggtggc ataaaacttg tgattggggt a 401

<210> 7

<211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 7  
 acttcttgtg acaaaaaaga gatttgttta atcagaaaca ctggctgttg gaatgctcct 60  
 aggttgcttt agaagagagt ttgcattgtt gggtcaccac cttctacctg ggttgtaacc 120  
 tcatataacc attaggatga ttgaccgttt tgaaaatgag tggccattaa ttcattttgt 180  
 taggggtcct tgagcttacg dgttttcaac ttttttcac cttcttgttg gaattctggg 240  
  
 ttgtgatttg aagatacttc caaagaaata aaatattcct aggaccttaa gaaccaggc 300  
 tttgtaaagc tgatatgaag gacctcttga atggcattcc acagtttggt tgtttgaaat 360  
 cacagtggtt tgttgggcca tgtgcaggaa gaaagtcatt g 401  
 <210> 8  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 8  
 ttgacaaaga cgtttctgtt ctgagggggt gtgtgggacc aacagcagca gctggggtgg 60  
 ccaaatgacc gccttctgtt agaatagaga gggattcaaa gaatgatctc tgccagattg 120  
 ccgggctttg ctgtgtcaca tcagctgac acagatggcc ttggccacta aggacgtgc 180  
  
 cttcttcctc acagaagccc vaggaaatgc cctgagaagg aacaaggacc agataactac 240  
 ccggcaggga gtggctggtg ggatttcagc ttgaaatgtg atctctggat gcccagagc 300  
 cagtgtggca tttatttcca ccgcatcagg ggtcccaacc ttggataaca tttatatttc 360  
 tagagcacct gaagaacaaa gattccagat tgtaaggtgc a 401  
 <210> 9  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 9  
 ttggaaggga taagaagagt caagagaggc atagaaggag tggtcagcag agagggagaa 60  
 gctgagaaaa tgtggcaaga aggccaatga gtaagaaaat tcaaaaagtg gccaatgcca 120  
  
 acccacaata ctacagagac aagatagaag aatgctcttc cacagaagag ctgcagagac 180  
 ctttcatagt acagttttag ygcattggtg gtgggtctag tggaactatt attaatacag 240

tggtggctaa gctgctaatt tttctaccac agatgctgtg aattagacaa actgtgcccc 300  
aggaaaggca tgacagggca cacattatct ataaaaatg aaaatgggcc aggcgctgtg 360  
gctcacgcct gtaatccag cactttggga ggccaaggcg g 401

<210> 10

<211> 401

<212> DNA

<213> Homo sapiens

<400> 10

ggccacaagg ggttttatgc cctgagcctt ggattccatc cagccacaa ggggttgat 60

tccctgtgct tagatttggg tgcagcaggg cagccttcca cctttggca cagagcttgg 120

tgttccaaag gccacgaggg ggttttagacc ctggaccag gacatgttcc aagactcttt 180

tacataatga cagacaagcc mgtcttgcac cagctcttct accaacagtt agcaattcct 240

ctaacctttt ttaaggttct tagcttctt gcattgggtt agaacatgct cctttagctt 300

ggaggagtct gttattacce atcttctgaa gcctacttct gtcaattcat caaatccatt 360

ctctgtccag ttttgttgcc ttgctggtaa ggagttgtga t 401

<210> 11

<211> 401

<212> DNA

<213> Homo sapiens

<400> 11

gcagatcact tacattacat gaggaagagg ttaaaatgta gaccattaaa gtcattctta 60

aggaaaaatg aaaaagactt ctttgaggca aatttagaaa agggaaaaag acctgacctc 120

aattcaaaga ggtgggaaga aatttttacc agatgagaaa ggaagaaatg aaatggtggc 180

accttcctg cccctactcc magtctgaat aatcagagga atgacacaag cctgcacaga 240

atatgtctcc tcttacaaaa tcatctataa gggaccttcc ttgttactgt cgggcaaagg 300

cattttgtaa gcaccatcag agcctgggtac ggcataggat gtagcacaaa gaaggtttgt 360

cgctctgggc ctgcctcctg gatgctgaca tagatgcaga a 401

<210> 12

<211> 401

<212> DNA

<213> Homo sapiens

<400> 12

tagtgtttga actaatggta taactggact ccagagtcac ttgacttaac taaaaggata 60  
tatgatgagc tcttatctgc ccagaaaact attacaatgt ttacattta taccaaatat 120  
tccaaatatt attataaaca tcttagattt cctgtgtagc tctgtttcga atgaaaatat 180  
tatagagctc gtcactcatt rcatactaaa ctttgttcct aagggcctgg gacatgtgac 240  
agtgtacaag ctttttcagt actgcattga aagctacatc acattcattg tcattgctaa 300  
agtgatgaga ttacacgttg cttaaattcc agtaattigt tcccactgta gaagataaat 360

tcttgattta gatggatcat caatctacat tatctagatg t 401

<210> 13

<211> 401

<212> DNA

<213> Homo sapiens

<400> 13

tatgtttatg cagaatgtgt tatacatag ggaagtgatt ttaaaatatac attaatgat 60  
gagacagagt acttcttata tgtttaatgt tatatcaact ctttaaaaat gtttgcctat 120  
gtataaaaag aaaaaataat tctggaggaa aaatgccaga atgtttacag ctatttcctg 180  
gttacagtgg tctttgcagg rgcctgtaa tcttgacatc atccggctgc cagtctccct 240  
acccccaact tctaccact caccctcccc tacacccag ttaaacatg tgctttctat 300

tccccatctc tcagcttagc atctttgtgc cagctgatgc cactgccttt ccctgtattc 360

aactgcttag ttaattccct tacccttca agtttgacta a 401

<210> 14

<211> 401

<212> DNA

<213> Homo sapiens

<400> 14

aagaaccttc ctcggtggac tctacaggga tagtgccatc aggacaggtg tctagctcaa 60  
ttctaggcac ctagctgggtg ccagtgat ccttttttgg gagaataaat gcgaagatgc 120  
catgaaggac aaacaattag tgatagccag tgacaacca gggtgcctaa ggtaggagc 180  
ctgggcagct tgaggagaac rgggaactca ggaagaggag caggttgggt caggcctggt 240

gaagggtggg agtagtttga ctcggagggtg cagggcccat ggcctggcac acaggcccgg 300

cacatgggta tgggagccca ggggctgtct ccattcattg aggtctgggt taaggtttct 360

gtgctctggg cctgccacc tgtgttctca gaagggtcct a 401

<210> 15

<211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 15  
 atttctgaaa tttcttttac tttatttgac atattttaca ctaaaaaaaa tcaatgtaac 60  
 tgataggact tcttaaattg cctatittta aaaaatgtat taccctaca attgtaagta 120  
 atttgggaagc cctggtacaa cctatacagt gtttttgaat ttgattctga gtggacagct 180  
  
 gtaggaggga gaattactga rtcacaaaat gaaacagatc acagcactga ttgggccta 240  
 ctctttatca aaagaggtc ttcatacact taactcagca agtcactgtt gagcactgga 300  
 aatcgtggag ctaggtccat cctcactccc ctccctgtcc aaatgagagg aggaaatgcc 360  
 caaactatta tctttccaaa cctgagccgt cttttgtcg c 401  
 <210> 16  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 16  
 caattccata aaacaaatga ttttttaata gcaaagagag tatgaatagc ataaaaatga 60  
 gtagcgtaaa aatctgagca aagtatttta gcacatagg aaaagtaaaa aaaaaaaaaa 120  
  
 atctgaataa agttttaaga tgttcaaata cctagacact gaatggtctt gaaattaata 180  
 aagcaaaactg ttaaatttca rgtaaccct tggttcatca tgctttacca cctggtcacc 240  
 ttgtgtgatg ttagtttga aaataactgt ggacatccac tgtttctcct tctcttagg 300  
 cctctgatca tcacaggcaa agggaaaacc taacagctca ttctgccag gtctcatctc 360  
 tcctacctga attgctgttg ttgggtccag aagaaactaa t 401  
 <210> 17  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 17  
 tttataattc cctaaagcta caattacatg caaactggat tcttgaaata aaacacaatt 60  
  
 gaatttaaat gacatctaca tgaaacccaa taactcgatt acattgtttt aatgttctta 120  
 atgttggtcaa aggcagccca tggctttatc atcaattagg taagcttttg agggaaattc 180  
 tgaaaaatgg caagtcctca bgtgtggact gacaactgag caattgcct gctgtgcttc 240

agaaagagag gccagctctg tgaggtgctg tcgaaatgac tgaaaaccac cactgcgga 300  
 tttctacttc cctctatgaa aaccagcctt ctccaggatt atcaccgcca ataagagggg 360  
 aaaattgcct gcacagtgg tgattgtggt taactggtca g 401

<210> 18  
 <211> 401  
 <212> DNA

<213> Homo sapiens

<400> 18  
 cactttaaat ttgcagtgc ttctaactta tttgtatat ttaaagatac tgattaaagg 60  
 aaaaatttct ttgtgaggac aacgctcttt atatgcctta gttacggat gaggaaacct 120  
 cagcagaggg ttaagacac ttgctaaatt tccttctctt cagagatagt caatggtgtt 180  
 tagattcggt ccgtttgtca rtggaaacat ttgtttctgc aattaagcct tattaatatag 240  
 cagaaaagtg tactttattg ggtgtaactt tttttcttc tgtgttttac agctatattt 300  
 tatttcctta tatattaat atttgaatga attttctta atcatgacac agccagacta 360  
 tattttttca atactataaa aggggttaaat atatttccat a 401

<210> 19  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 19

ctgcctggct gggcagtgtg gcagatacca gggagctgac ttctcccctg gcagcccagc 60  
 cccaggccta gggatattgc tggttcagtg gccccttctg atctcccag gaccagtga 120  
 aagaaagggc ttcccttcatt tgtctcatt catgtaaaat ccaggaggcc ccagtttcag 180  
 ttccggcctt gcctcaattg rgcctccctt ttggccccag tattcatgtc tcttctgaga 240  
 ggggttgtag ctcatgacga ctcatgaacc tgaacaaaag gaggaaaacc ctccctccat 300  
 agcagcttga ggaccaggcc gttctctgtg tctgcagcct gctttttgat ttgagcttct 360

aagctttatt tggatttagc ttggttgctc tgaacgttg g 401

<210> 20  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 20

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| gacatagtgc tagggatgta ggcagcactc atatttaacg caatactgta ccaaaggcta  | 60  |
| cagcagaatg gtgaaaggag aaagcagtaa acttggagtt agtagaccgg agacatgctc  | 120 |
| atatttttat tcaacgtatt ttcataaac acattcatgt gctttgggat agacaccatg   | 180 |
| aaagacacaa aaatgtctac yatataaggc agtatctaag cagaaaatga ttgatgata   | 240 |
| aatactatag cagtggagat gacagaaata tttttctggc tcaggtcatc aatgaaggga  | 300 |
|                                                                    |     |
| tcataataca cgcgcatata tgtgcatgca catatatggg tggggccctg aagccccact  | 360 |
| tctgttatta taagcagaga cacggagggt aaagattagg g                      | 401 |
| <210> 21                                                           |     |
| <211> 401                                                          |     |
| <212> DNA                                                          |     |
| <213> Homo sapiens                                                 |     |
| <400> 21                                                           |     |
| aaagtttctg gagggttgtg atttaatagc cgtatagcac atgccttagc aagtgtcgtc  | 60  |
| aacactcacc tctctatcat tcttagctag tactttcttc atcttgctc agttattatt   | 120 |
| tgaaatgtgc ttgtttcctt agtcattaca ttttctaatt tgccaaatgt taaacaaggg  | 180 |
| caagagttgg cactttctgt ktctctctgc ttctccac caccactgta agccccaat     | 240 |
|                                                                    |     |
| aaaataattg tttttacat ttctgaaagc ttttactct gttacaaatc aaggggctcc    | 300 |
| aacatgccct tgtttctaac aattctcaat gcagtttaat gctaggggtg aggtgggctc  | 360 |
| attgactttt gctcttggt gacatttcca ttagcacatg t                       | 401 |
| <210> 22                                                           |     |
| <211> 401                                                          |     |
| <212> DNA                                                          |     |
| <213> Homo sapiens                                                 |     |
| <400> 22                                                           |     |
| gcctgtaatt ccagctacgt gggaggctga ggcaggagaa tcaactgaac ccaggaagca  | 60  |
| gtggttgtgg tgagccaaga tcatgccatt gcattccagc ctggacaaca agagcaaac   | 120 |
| tcctttcaaa aaaaaaaaaa aaaaaaaaaag atagggaata tctgagaaac tatcacagac | 180 |
|                                                                    |     |
| aagaagaaac ttcagagtca mggcccacca gatgcatgt ggtgtctcag atgggatcct   | 240 |
| ggagccgaaa aaggacatta gatgaaaact aaagcaatcg aagtatagc ttagttaat    | 300 |
| cgtgatgtat tcatagtgtt tccttagttg tgaacatgta ccatgctact gtaagatgtt  | 360 |
| aataatagtg aaaattggct ggacaccgtc gctcacggct g                      | 401 |
| <210> 23                                                           |     |

<211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 23  
 tgatctgaat aattatacag ttagatctct aaagattatt ttacaaatat aaagacagtt 60  
 gccattggcc ctttcttacc tcttttccac acgtttggtc ttgaaatgct gctgcttttc 120  
  
 ctctataaat gatgtattta ttgctctgtg taatctctga tggagaaaac agattcacag 180  
 gtagaaatta tgagaaacac vtaggtcagg gaacaaagaa ccacttgaga tgtaacagta 240  
 tatgtagtgc ttacaaacct tggcctcttt aaacaaattc cagaaatgcc taaaaataga 300  
 tcaacattca gcctcaaadc tcaacagaac aaaaattcta attagagggtt attattccac 360  
 ctggagaaaac aaaaaagtgt tctgatggaa agctcagtat t 401  
 <210> 24  
 <211> 401  
 <212> DNA  
 <213> Homo sapiens  
 <400> 24  
 tggctttctt accatcttca ggagtggcgt taaagcttgc tctggctcag tccctgtgtg 60  
  
 catccatgag gtcatagttc catgagggtc catccagctc aagctctgtg gccgggcctg 120  
 agccccaggg cgttgtgtgg agccactcca gagccaggac tcaggcagaa ccatccctgg 180  
 gcctcaggca gagccattct kgttcagtgc agtgaactcc aaaatacga cttgggggtg 240  
 ggagttcagg tctcagtcac caaggagctc ttggactctt acaagaatgg agccaccgcc 300  
 tccacagctc tgccaaccca ggtttttagga aatggttatc taatggacat ctcaaatcca 360  
 aacttctaag gtggaaatct tgatctcagc cccccacaaa a 401  
 <210> 25  
 <211> 401  
 <212> DNA  
  
 <213> Homo sapiens  
 <400> 25  
 ggaggcagag gttgcagtga gccaaagattg tgccactgca ctccagcctg gctatagagc 60  
 aagactccat ctcaaaaaaa aaaaaaaaaa aaagaaagaa agaaattcaa cttcaagggt 120  
 cctacctaga aacctgtttt aacaagcctc catgtgactc ctctgggtgt gaagaaattc 180  
 ttcacacat acaataaaat rtcactgtt accatgatag aagaaccgt ggctttctct 240

ggcgactata agtacggata atgggccagg gctccctagc tggaggaaa gtcactgtcc 300  
tgggggcaaa gcaggtaccc agtgactgtc cctgctgatg atgcaccccg agaagtggag 360  
caggctggga tgtgtttgac acaaagccct cttgctgtcc t 401

<210> 26

<211> 401

<212> DNA

<213> Homo sapiens

<400> 26

gcgcttcaact cagctcctgc cctgggcccg ctggcacctg gcgaggcaga ttttccttc 60  
caatactcaa gagtcccttt caggggcagg catgggatgc cagggttaca agagtggaaa 120  
aaaatgtcca tttcctttgg tcagtttacc agatgagtat tactcaacca cttacaatat 180  
ttacaagtcc agctcatggc rattaaaaaa agatgtaaga aaacaaactc cttttcaaac 240  
tgggagagga tattaaacac agtctctcag gaatggtgtt ttctgaagac tcgtcatctg 300  
gagtcatggt ccatactttt ttccagtga acactggaaa cgagcaaat acggagcttg 360

tgctgccaac gcccttgcca caagttgggt gtaggtaaca g 401

<210> 27

<211> 401

<212> DNA

<213> Homo sapiens

<400> 27

aagccgaaga tcgtgccatt gcgctccagc ctgggagaca gagccagact ctgactcaaa 60  
aaaaaaaaa aaaaaaggat agtatattta cacttaggaa acaatagtga atacttttct 120  
aataaactct aaagttacat tcaacttcag aatgcatttc ttcatgtccc ttgcaacat 180  
ttagatggag tatctaaatt ycataaatac gatttttctt tatgaccaa gaaagtcctc 240  
cagtaccagg aaggagtcaa attcttgggt gttgagccac gtgctcaaga ttgagcaaca 300

gttaaggaac tcigatagct aattcttaac ttagaaatgg tgatatttca tgatttgatg 360

tctgggattt tcttcaaagt aaaatgagat gaacatgaag t 401

<210> 28

<211> 401

<212> DNA

<213> Homo sapiens

<400> 28

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| caaataaggaa tggatgtgca catcgctgct gtcagggtc agaaagataa tacccaaaat  | 60  |
| gaaggcctca gaaacaaaag actcttggac cttctccagc cctcctctct ctagecctcta | 120 |
| atcttctccc cataggccag ccacaaaacc taaaaatatt attctaattcc ctaccacctt | 180 |
| cctgcataat aactggccat raagaaacta tctgacctac cttgcttaat tataggatcat | 240 |
|                                                                    |     |
| tagaccccca ttccagaaag cgtctctgcc cataccaga agcaaggaat gctgctgaga   | 300 |
| ggcaaagatg aatgcagaca caccttgctg ggttttctca gtctaccagg attagatcac  | 360 |
| atcttttctc caatcatatt tctatatggc tctctgtact t                      | 401 |
| <210> 29                                                           |     |
| <211> 401                                                          |     |
| <212> DNA                                                          |     |
| <213> Homo sapiens                                                 |     |
| <400> 29                                                           |     |
| ccatgccag ctaatctttg tatttttagt agagacaggg tttcactgta ttgccaggc    | 60  |
| tggctctaaa ctctgacct caagcaatct gccctctctg gcctccaaa gtgctgggct    | 120 |
| tacaagcgtg agctaccatg ccagccatt aattagctct tttaaagccc tgtcttcaac   | 180 |
|                                                                    |     |
| tacagtctca ttcataggta ytgggaatta gggcttcaac atatgaattt ggggagatgc  | 240 |
| aattcagcct gtaagagctt gtatttagtc tttggtagtg tttgaactaa tggatataact | 300 |
| ggactccaga gtcacttgac ttaactaaaa ggatatatga tgagctctta tctgcccaga  | 360 |
| aaactattac aatgttttac atttatacca aatattccaa a                      | 401 |
| <210> 30                                                           |     |
| <211> 401                                                          |     |
| <212> DNA                                                          |     |
| <213> Homo sapiens                                                 |     |
| <400> 30                                                           |     |
| gacacaaagt cagcattaga cacctaggag attactgggg gaggaagca gggtaggaag   | 60  |
| gaagaacttc cagaccatga tgctaattgt tgtgaagcag aggaagagat aggaggtttg  | 120 |
|                                                                    |     |
| ggaagaagga gtcttaggct ccagggtact tccaagagag gtctggcagg ccaactagga  | 180 |
| gtccttgagc ccatgttcca rtgagaaagt ctacatgta aaaggaagga gtctatgcta   | 240 |
| acagctcatt cactgactat aggaagtttg gacttgccat gagacagcag tggatcaagg  | 300 |
| atagcagtag ctgagttctt cagccagtta tgtcccaca acaagagacc tcagtggccc   | 360 |
| atctttatgg tgccacatca gcctactgat agcgacagga g                      | 401 |
| <210> 31                                                           |     |

<211> 401

<212> DNA

<213> Homo sapiens

<400> 31

tacctgctct ttgtccacac ttccctctct ctatccttga gaaacttttc tgcccttcac 60

ccagactttt tcatcatgct gtectgcttt ttttcttga cccttgattg tcagttgcca 120

aatcctgttt ttcaacttct atagtctttc ctttgtttcc ctctgtttta gtccattttg 180

tgttgctgtg atagaatacc bgagactggc taatttcatt tcatttcttt tttttttttt 240

ttggagacag agtctcacc aggctggagt gcagtgggtgc aaatctcggc tcactgcaac 300

ctctgccttc tgggttcaag tgattctcct gcctcagcct ctcgagtagc tgggactaca 360

ggcatgtgcc accacgcca actaattttt gtatttttag t 401