



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0089240
(43) 공개일자 2020년07월24일

(51) 국제특허분류(Int. Cl.)

C12Q 1/6883 (2018.01)

(52) CPC특허분류

C12Q 1/6883 (2018.05)

C12Q 2600/156 (2013.01)

(21) 출원번호 10-2020-0019086(분할)

(22) 출원일자 2020년02월17일

심사청구일자 2020년02월17일

(62) 원출원 특허 10-2019-0005899

원출원일자 2019년01월16일

심사청구일자 2019년01월16일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상학

서울특별시 강남구 압구정로29길 71, 33동 102호 (압구정동, 현대아파트)

이지현

서울특별시 종로구 송월길 99, 206동 1303호(홍파동, 경희궁자이 2단지)

이찬주

서울특별시 은평구 은평로8길 7, 5층(응암동)

(74) 대리인

리엔목특허법인

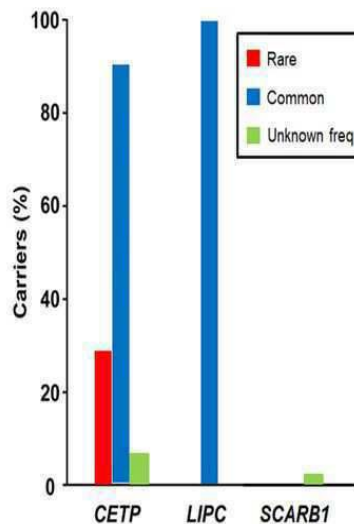
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 고밀도지질단백질콜레스테롤 수준을 예측하기 위한 조성물, 키트, 및 이를 이용한 방법

(57) 요약

유전자의 변이를 검출하여 고밀도지질단백질콜레스테롤의 수준을 예측하기 위한 조성물 및 키트, 및 이를 이용한 개체의 고밀도지질단백질콜레스테롤 수준을 예측하기 위한 정보를 제공하는 방법에 관한 것이다. 일 양상에 따른 조성물 또는 키트, 및 이를 이용한 개체의 고밀도지질단백질콜레스테롤의 수준을 예측하기 위한 정보를 제공하는 방법에 의하면, 개체의 고밀도지질단백질콜레스테롤의 수준을 예측할 수 있어, 상기 고밀도지질단백질콜레스테롤의 예측을 통하여 최종적인 개체의 심혈관 질환을 예측 또는 측정할 수 있다. 예측 또는 측정된 개체의 심혈관 질환의 발병 위험 고밀도지질단백질콜레스테롤의 수준을 조기에 측정하여 심혈관 질환의 발병을 예방하기 위한 정보로 이용될 수 있고, 개인 맞춤형의학 치료에 적용하여 치료 효과를 증대시킬 수 있다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 2015M3A9B6029138

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 바이오의료기술개발사업

연구과제명 융합가상 인체모델을 활용한 심혈관 예측기술 개발 사업

기 여 율 1/1

주관기관 서울대학교

연구기간 2015.06.01 ~ 2020.05.31

명세서

청구범위

청구항 1

서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오타이드, 1084번째 뉴클레오타이드, 1195번째 뉴클레오타이드, 974번째 뉴클레오타이드, 또는 537번째 뉴클레오타이드를 검출할 수 있는 제1 폴리뉴클레오타이드를 포함하는, 고밀도지질단백질콜레스테롤(HDL-C) 수준을 예측하기 위한 조성물.

청구항 2

청구항 1에 있어서,

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오타이드는 A가 G로 치환되거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오타이드는 G가 A로 치환되거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오타이드는 G가 T로 치환되거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오타이드는 T가 C로 치환되거나; 또는

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오타이드는 G가 A로 치환된 것인 조성물.

청구항 3

청구항 1에 있어서, 상기 검출할 수 있는 제1 폴리뉴클레오타이드는 프라이머 또는 프로브인 것인 조성물.

청구항 4

청구항 1에 있어서, 상기 검출할 수 있는 제1 폴리뉴클레오타이드는 길이가 5 내지 100bp인 것인 조성물.

청구항 5

서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산, 362번째 아미노산, 325번째 아미노산, 또는 179번째 아미노산을 검출할 수 있는 제1 폴리펩티드를 포함하는, 고밀도지질단백질콜레스테롤(HDL-C) 수준을 예측하기 위한 조성물.

청구항 6

청구항 5에 있어서,

상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산은 D가 G로 치환되거나;

상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 362번째 아미노산은 V가 I로 치환되거나;

상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산은 D가 Y로 치환되거나;

상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 325번째 아미노산은 V가 A로 치환되거나; 또는

상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 179번째 아미노산인 W가 넌센스 변이로 치환되는 것인 조성물.

청구항 7

청구항 5에 있어서, 상기 검출할 수 있는 제1 폴리펩티드 중 어느 하나 이상은 항체 또는 항원 결합 단편인 것인 조성물.

청구항 8

서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 또는 537번째 뉴클레오티드를 검출할 수 있는 제1 폴리뉴클레오티드를 포함하는,

고밀도지질단백질콜레스테롤(HDL-C) 수준을 예측하기 위한 키트.

청구항 9

하기 단계를 포함하는 개체의 고밀도지질단백질콜레스테롤(HDL-C) 수준을 예측하기 위한 정보를 제공하는 방법으로서,

개체의 생물학적 시료로부터 핵산 시료를 획득하는 단계; 및

수득된 핵산 시료로부터 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에 존재하는 변이 위치의 유전자형 (genotype)을 분석하는 단계를 포함하고,

상기 변이 위치는,

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 또는 537번째 뉴클레오티드인 것인 방법.

청구항 10

청구항 9에 있어서,

상기 유전자형을 분석하는 단계는,

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드는 A 또는 G이거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드는 G 또는 A이거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드는 G 또는 T이거나;

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드는 T 또는 C이거나; 또는

상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드는 G 또는 A인지 여부를 결정하는 단계를 더 포함하는 것인 방법.

청구항 11

청구항 9 또는 10에 있어서,

상기 유전자형을 분석하는 단계는 시퀀싱, 마이크로어레이에 의한 혼성화, 대립유전자 특이적인 PCR (allele specific PCR), 다이내믹 대립유전자 혼성화 (dynamic allele-specific hybridization), PCR 연장 분석, PCR-SSCP (PCR-single strand conformation polymorphism) 및 TaqMan 방법으로 이루어진 군으로부터 선택된 하나 이상의 기법에 의해 수행되는 것인 방법.

발명의 설명

기술 분야

[0001] 유전자의 변이를 검출하여 고밀도지질단백질콜레스테롤의 수준을 예측하기 위한 조성물 및 키트, 및 이를 이용한 개체의 고밀도지질단백질콜레스테롤 수준을 예측하기 위한 정보를 제공하는 방법에 관한 것이다.

배경 기술

[0002] 고밀도지질단백질콜레스테롤(High High-Density Lipoprotein-Cholesterol : HDL-C)의 수준이 높은 경우 심혈관 질환의 위험이 증가한다는 것이 최근 연구를 통하여 알려졌다 (Van der Steeg WA, Holme I, Boekholdt SM, et al. J Am Coll Cardiol 2008;51:634-642). 전 세계적으로 연간 1,700만 명이 심혈관계 질환으로 사망하며, 이는 전체 사망 원인의 30%에 달할 정도로 심혈관계 질환은 위험도가 높다. 우리나라에서도 심혈관계 질환은 주요 사망원인으로 2012년 관상동맥 질환 사망률은 남자는 10만 명당 31명, 여자는 10만 명당 27명이었으며, 뇌혈관 질환 사망률은 남자는 10만 명당 49명, 여자는 10만 명당 53명이었다.

[0003] 상기 심혈관질환은 유전적 및 환경적 요인의 상호작용에 의한 결과로 알려져 있다. 환경적 요인으로는 음주, 흡연 등에 의해서 심혈관질환이 발생할 수 있고, 유전적 요인으로는 유전에 의한 비만 또는 당뇨병, 또는 혈액 내 특정 지질이 증가됨으로써 심혈관질환이 발생할 수 있다.

[0004] 다만 HDL-C 수준과 심혈관질환의 관련성이 높고, 유전적 요인이 HDL-C의 수치에 영향을 미치기 때문에, 유전적 특징과 개체 내 고 수준의 HDL-C의 발현으로 인한 심혈관질환과의 연관성이 클 것으로 여겨진다. 이에, 고밀도지질단백질콜레스테롤 발현 수준을 예측한다면 이와 관련된 심혈관질환을 발병을 예측할 수 있을 것으로 보여진다.

[0005] 개인 맞춤의학적 견지에서, 최근에 이루어진 전장유전체연구 (GWAS, genome-wide association study)를 통해서, 비만, 고혈압, 당뇨병 관련 표현형과 연관성이 있는 다수의 SNP (단일염기다형성, single nucleotide polymorphism)들이 밝혀지고 있다. 그러나, 아직 HDL-C에 관한 유전적 특징을 이용하여 심혈관질환의 발병을 예측할 수 있는 유전적 지표에 대한 연구는 미흡한 실정이다. 따라서 지속적인 연구를 토대로 유전자 검사를 통한 심혈관질환의 고위험군을 조기 진단하여 예방 및 관리할 필요가 있다.

발명의 내용

해결하려는 과제

[0006] 일 양상은 고밀도지질단백질콜레스테롤(HDL-C) 수준을 예측하기 위한 조성물을 제공한다.

[0007] 다른 양상은 고밀도지질단백질콜레스테롤 수준을 예측하기 위한 키트를 제공한다.

[0008] 다른 양상은 개체의 고밀도지질단백질콜레스테롤 수준을 예측하기 하기 위한 정보를 제공하는 방법을 제공한다.

과제의 해결 수단

- [0009] 일 양상은 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드, 또는 이들의 조합을 검출할 수 있는 제1 폴리뉴클레오티드; 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합을 검출할 수 있는 제 2 폴리뉴클레오티드; 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드를 포함하는 HDL-C 수준을 예측하기 위한 조성물을 제공한다.
- [0010] 고밀도지질단백질-콜레스테롤. 혈장의 HDL(고밀도지질단백질, high density lipoprotein)에 함유된 콜레스테롤의 농도를 지칭한다. 혈장의 HDL-C의 농도는 보통 40~60mg/dL이다. 여성이 남성보다 약간 높은 값을 나타내며, 식사, 운동, 각종 질병, 흡연, 약물투여 등도 그 농도에 영향을 주는 것으로 알려져 있다.
- [0011] 상기 심혈관질환(cardiovascular disease)은 심장과 주요 동맥에 발생하는 질환을 말하며, 주요 혈관계 질환은 대동맥, 허파동맥, 목동맥, 뇌혈관, 신장동맥, 하지 동맥(온영덩(장골)동맥, 넓다리(대퇴)동맥 등) 등의 주요 동맥이 막히거나 늘어나거나 터지는 출혈이 일어나는 질환이다. 주된 원인으로 (죽상)동맥경화증, 고혈압, 퇴행성 변화, 유전 등이 있으나, 최근 연구에 의하여 HDL-C의 수준이 높아질수록 심혈관계 질환 위험이 증가한다는 것이 알려져 있다.
- [0012] 변이는 유전체에서 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산의 변경 (alteration)을 의미한다.
- [0013] 변이 위치는 유전체 상에서 상기 변이가 일어난 위치를 의미한다.
- [0014] 변이는 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산의, 치환 (substitution), 삽입 (insertion), 결실 (deletion) (삽입, 결실을 Insertion, Deletion: 'InDel'이라고도 함) 등을 포함할 수 있다. 치환은 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산이, 다른 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산으로 바뀌는 변경을 의미한다. 삽입은 다른 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산이 추가되는 변경을 의미한다. 결실은 염기, 뉴클레오티드, 폴리뉴클레오티드 또는 핵산이 제거되는 변경을 의미한다.
- [0015] 단일염기 변이 또는 단일뉴클레오티드 변이 (single nucleotide variant: SNV)는 유전체 상에서 하나의 염기 또는 뉴클레오티드의 차이를 보이는 서열의 변경 또는 변이를 의미한다. 단일염기 변이는 단일염기 다형성 (single nucleotide polymorphism: SNP, 이하 'SNP'라고 함)과 혼용될 수 있다. SNP는 유전체 상에서 단일 염기 또는 뉴클레오티드 (A, T, C 또는 G, 뉴클레오티드 A는 아데닌, 뉴클레오티드 T는 티민, 뉴클레오티드 C는 시토신, 뉴클레오티드 G는 구아닌을 의미함)가 종의 멤버들 간 또는 한 개체 (individual)의 쌍 염색체 간에 다른 경우에 발생하는 염기 또는 뉴클레오티드 서열의 다양성을 의미한다. SNP는 한 집단 (population)에서 소수, 1% 이상 또는 5% 이상의 빈도로 존재하는 2개 이상의 대립 염기 또는 뉴클레오티드 서열일 수 있다. SNP는 인간 유전체 상에 가장 많이 존재하는 유전적 다형성으로, 유전학적으로 SNP 위치에 따라 각 개체에 큰 차이를 야기할 수 있다. 예를 들면, SNP가 단백질을 암호화하고 있는 위치에 존재할 경우, 단백질의 구조에 영향을 미쳐 단백질 기능이 달라질 수 있고, 질병을 유발할 수 있다. SNP가 단백질을 암호화하지 않는 비암호화 영역에 존재할 경우, 즉 프로모터 (promoter) 또는 인트론 (intron)에 존재할 경우, 각각에 대하여 단백질의 발현 수준에 차이를 가져와 그 단백질의 전체적인 활성이 증가 또는 감소할 수 있고, 선택적 이어맞추기 (alternative splicing)를 통하여 비정상적 단백질이 발현될 수도 있다.
- [0016] 상기 제1 폴리뉴클레오티드는 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드, 또는 이들의 조합을 포함하는 폴리뉴클레오티드 서열과 동일하거나 또는 상보적일 수 있다.
- [0017] 상기 제2 폴리뉴클레오티드는 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합을 포함하는 폴리뉴클레오티드 서열과 동일하거나 또는 상보적일 수 있다.
- [0018] 상기 제3 폴리뉴클레오티드는 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드를 포함하는 폴리뉴클레오티드 서열과 동일하거나 또는 상보적일 수 있다.
- [0019] 상기 제1 폴리뉴클레오티드는 콜레스테롤에스테르수송단백질 (cholesterol ester transfer protein: CETP) 유전

자의 단일염기 변이 또는 단일염기 다형성의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다. 콜레스테롤에스테르수송단백질 (이하, 'CETP'이라고 함)은 혈장지방내 콜레스테롤 에스테르 또는 트리아실글리세롤 등인 비극성지질(및 지방단백질 표층인 인지질)을 지방단백질 간에서 수송하는 활성을 가지고 있는데, 혈장내에서는 주로 콜레스테롤 에스테르를 고밀도지방단백질(HDL)에서 초(超)저밀도지방단백질(VLDL, LDL)방향으로 수송하고 역방향으로는 트리아실글리세롤을 수송하는데 기여한다. 상기 CETP는 인간의 경우 CETP 유전자에 의해 암호화되는 단백질일 수 있다. CETP는 인간의 경우, MIM No. 118470, UCSC genome browser (GRCh37/hg19): chr16: 57,017,292, (GRCh37/hg19):chr16: 57,016,092, (GRCh37/hg19):chr16: 57,017,291, (GRCh37/hg19):chr16: 57,015,077, (GRCh37/hg19):chr16: 57,004,954, 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에 의해 암호화되는 폴리펩티드 또는 서열번호 4의 아미노산 서열을 갖는 폴리펩티드일 수 있다.

[0020] 상기 제2 폴리뉴클레오티드는 hepatocyte triacylglyceride lipase (Hepatic triglyceride lipase :LIPC) 유전자의 단일염기 변이 또는 단일염기 다형성의 유전자형, 또는 염기 치환, 삽입 및/또는 결실에 의해 유발된 격자이동의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다. hepatocyte triacylglyceride lipase (이하, 'LIPC'이라고 함)은 중성지방 (triacylglyceride: TG) 의 가수분해를 촉매하는 리파아제의 일종이다. 상기 LIPC 유전자는 HDL-C의 혈장 내 수준과 밀접한 연관성을 가진다. 상기 LIPC 는 인간의 경우 LIPC 유전자에 의해 암호화되는 단백질일 수 있다. LIPC 는 인간의 경우, MIM No. 151670, UCSC genome browser (GRCh37/hg19):chr15: 58,853,079, (GRCh37/hg19):chr15: 58,838,010, (GRCh37/hg19):chr15: 58,833,993, 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에 의해 암호화되는 폴리펩티드 또는 서열번호 5의 아미노산 서열을 갖는 폴리펩티드일 수 있다.

[0021] 상기 제3 폴리뉴클레오티드는 scavenger receptor class B member 1 (Scavenger reCETPor class B member 1:SCARB1) 유전자의 단일염기 변이 또는 단일염기 다형성의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다. scavenger receptor class B member 1 (이하, 'SCARB1'이라고 함)은 주로 간세포에 존재하면서, 고밀도지질단백질로부터 콜레스테롤 에스테르의 선택적 흡수를 매개하는 단백질로 알려져 있다. 상기 SCARB1은 인간의 경우 SCARB1 유전자에 의해 암호화되는 단백질일 수 있다. SCARB1은 인간의 경우, MIM 601040, UCSC genome browser (SCARB1):Chr12: 125,294,817, 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에 의해 암호화되는 폴리펩티드 또는 서열번호 6의 아미노산 서열을 갖는 폴리펩티드일 수 있다.

[0022] 통상의 기술자라면 상기 등록번호를 이용하여 변이의 위치 및 서열을 용이하게 확인할 수 있을 것이다. UCSC genome browser 또는 GenBank에 등록되어 있는 번호에 해당하는 구체적인 서열은 시간이 지남에 따라 다소 변경될 수 있다. 본 발명의 범위가 상기 변경된 서열에도 미치는 것은 통상의 기술자에게 자명할 것이다.

[0023] 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이는, HDL-C의 수준을 예측하기 위한 개체들로 구성되는 한 집단에서 드문 빈도로 나타날 수 있으며, 10% 이하, 9% 이하, 8% 이하, 7% 이하, 6% 이하, 5% 이하, 또는 4% 이하로 나타날 수 있다.

[0024] 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 또는 변이는, 수준을 예측하기 위한 개체들로 구성되는 한 집단에서 드문 빈도로 나타날 수 있으며, 10% 이하, 9% 이하, 8% 이하, 7% 이하, 6% 이하, 5% 이하, 또는 4% 이하로 나타날 수 있다.

[0025] 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이는, HDL-C의 수준을 예측하기 위한 개체들로 구성되는 한 집단에서 드문 빈도로 나타날 수 있으며, 10% 이하, 9% 이하, 8% 이하, 7% 이하, 6% 이하, 5% 이하, 또는 4% 이하로 나타날 수 있다.

[0026] 상기 조성물에 있어서, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드는 A가 G로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드는 G가 T로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드는 T가 C로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드는 C가 A로 치환되거나; 상기 서열번호 2의 5' 말단으로부터 644번째 뉴클레오티드는 A가 G로 치환되거나; 상기 서열번호 2의 5' 말단으로부터 551번째 뉴클레오티드는 C가 G로 치환되거나; 상기 서열번호 2

의 5' 말단으로부터 283번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 107번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드는 G가 A로 치환되거나; 또는 이들의 조합인 것일 수 있다.

- [0027] 상기 제1 폴리뉴클레오티드는 CETP 유전자의 단일염기 변이 또는 단일염기 다형성의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다.
- [0028] 상기 제2 폴리뉴클레오티드는 LIPC 유전자의 단일염기 변이 또는 단일염기 다형성의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다.
- [0029] 상기 제3 폴리뉴클레오티드는 SCARB1 유전자의 단일염기 변이 또는 단일염기 다형성의 유전자형을 분석하기 위한 폴리뉴클레오티드일 수 있다.
- [0030] 상기 제1 폴리뉴클레오티드 내지 제3 폴리뉴클레오티드 중 어느 하나 이상 (이하, '제n 폴리뉴클레오티드'라고 함)은 단수개 또는 복수개, 또는 1쌍 또는 복수의 쌍일 수 있고, 각각의 폴리뉴클레오티드는 5 내지 100 뉴클레오티드 (이하, 'nt'라고도 함)의 연속 뉴클레오티드를 포함할 수 있다. 상기 제n 폴리뉴클레오티드는 DNA 또는 RNA가 될 수 있다.
- [0031] 상기 제n 폴리뉴클레오티드는 프라이머 또는 프로브일 수 있다. 상기 프라이머 또는 프로브는, 상기 변이를 포함하는 서열에 완전하게 (perfectly) 상보적인 서열이 이용될 수 있으나, 특이적으로 혼성화를 방해하지 않는 범위 내에서 실질적으로 (substantially) 상보적인 서열이 이용될 수도 있다.
- [0032] 용어 프라이머 (primer)는 중합효소에 의한 뉴클레오티드의 중합반응에서, 개시점으로 작용할 수 있는 단일 가닥의 올리고뉴클레오티드를 의미한다. 예를 들면, 상기 프라이머는 적합한 온도 및 적합한 완충액 내에서 적합한 조건, 즉, 4종의 다른 뉴클레오티드 트리포스페이트 및 중합효소의 존재 하에서 주형-지시 (template-directed) DNA 합성의 개시점으로 작용할 수 있는 단일가닥의 올리고뉴클레오티드일 수 있다. 프라이머의 적합한 길이는 다양한 인자, 예를 들면, 온도와 프라이머의 용도에 따라 달라질 수 있다. 상기 프라이머는 길이가 5 내지 100nt, 5 내지 70nt, 10 내지 50nt, 또는 15 내지 30nt인 것일 수 있다. 예를 들면, 프라이머의 길이가 짧을수록, 낮은 어닐링 (annealing) 온도에서 주형과 충분히 안정된 혼성화 복합체를 형성할 수 있다.
- [0033] 상기 프라이머는 포스포로티오에이트 (phosphorothioate), 알킬포스포로티오에이트와 같은 뉴클레오티드 유사체 (analogue), 펩티드 핵산 (peptide nucleic acid) 또는 삽입 물질 (intercalating agent)을 더 포함할 수 있다. 또한, 형광, 인광 또는 방사성을 발하는 표지 물질을 더 포함할 수 있다. 상기 형광 표지 물질은 VIC, NED, FAM, PET, 또는 이들의 조합일 수 있다. 상기 표지 물질은 상기 폴리뉴클레오티드의 5' 말단에 표지될 수 있다. 또한, 방사성 표지 물질은, ³²P 또는 ³⁵S 와 같은 방사성 동위원소가 첨가된 중합효소 연쇄 반응 (polymerase chain reaction: PCR, 이하 'PCR'이라고 함) 반응액을 이용한 PCR 반응을 통해 증폭 산물에 혼입될 수 있다.
- [0034] 상기 프라이머는 대립유전자 특이적인 PCR (allele specific PCR), PCR 연장 분석, PCR-SSCP (PCR-single strand conformation polymorphism), TaqMan 방법 및 시퀀싱 등을 이용한 방법에 이용될 수 있다.
- [0035] 용어 프로브 (probe)는 상보적인 폴리뉴클레오티드 가닥에 서열 특이적으로 결합할 수 있는 폴리뉴클레오티드를 말한다. 상기 프로브는 길이가 5 내지 100nt, 10 내지 90nt, 15 내지 80nt, 20 내지 70nt, 또는 30 내지 50nt인 것일 수 있다. 상기 프로브는 혼성화 방법, 예를 들면 마이크로어레이 (microarray), 서던 블로팅, 다이내믹 대립유전자 혼성화 (dynamic allele-specific hybridization) 및 DNA 칩 등을 이용한 방법에 이용될 수 있다. 마이크로어레이는 당업계에서 알려진 의미로 사용되며, 예를 들면, 기관 상의 복수 개의 구분된 영역에 프로브 또는 프로브의 집단이 고정화되어 있는 것일 수 있다. 상기 기관은 적합한 견고성 또는 반-견고성 지지체로서, 예를 들면, 막, 필터, 칩, 슬라이드, 웨이퍼, 파이버, 자기성 비드 또는 비자기성 비드, 젤, 튜빙, 플레이트, 고분자, 미소입자 및 모세관을 포함할 수 있다. 상기 프로브 또는 그에 상보적인 프로브는 개체로부터 수득된 핵산과 혼성화되고 그로부터 얻어지는 혼성화 정도를 측정할 수 있는 방법에 이용될 수 있다.
- [0036] 다른 양상은 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산, 362번째 아미노산, 325번째 아미노산, 179번째 아미노산, 또는 이들의 조합을 검출할 수 있는 제1 폴리펩티드; 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 5의 N 말단으로부터 356번째 아미노산, 215번째 아미노산, 95번째 아미노산, 또는 이들의 조합을 검출할 수 있는 제2 폴리펩티드; 서열번호 6의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환

위치인 상기 서열번호 6의 N 말단으로부터 249번째 아미노산을 검출할 수 있는 제3 폴리펩티드; 또는 이들의 조합을 포함하는 HDL-C의 수준을 예측하기 위한 조성물을 제공한다.

[0037] 아미노산 치환은 하나 이상의 염기 또는 뉴클레오티드의 변경에 의하여 아미노산 서열이 변경된 것을 의미한다. 아미노산 치환은, 미스센스 변이(missense mutation, 과오 변이), 잠재성 변이(silent mutation), 넌센스 변이(nonsense mutation), 중립 변이(neutral mutation), 인트론에서 일어나는 변이, 격자이동 변이(Frame shift mutation, 프레임 쉬프트) 등을 포함할 수 있다. 미스센스 변이는 염기 또는 뉴클레오티드가 바뀌어 다른 아미노산이 암호화되는 변경을 의미한다. 잠재성 변이는 염기 또는 뉴클레오티드가 바뀌었지만, 암호화되는 아미노산이 같은 변이를 의미한다. 넌센스 변이는 염기 또는 뉴클레오티드가 바뀌어 종결코돈이 됨으로써, 아미노산을 더이상 생성하지 않는 변경을 의미한다. 중립 변이는 염기 또는 뉴클레오티드가 바뀌어 다른 아미노산이 암호화되었으나, 본래 지정된 아미노산과 성질이 같은 변경을 의미한다. 인트론에서 일어나는 변이는 암호화하지 않는 인트론 영역에서 일어나는 변경을 의미한다. 격자이동 변이는 염기, 뉴클레오티드 또는 폴리뉴클레오티드의, 치환, 삽입, 결실 등에 의하여, 유전자를 암호화하는 해독틀이 이동하여 번역되는 아미노산이 달라지는 변경을 의미한다. 따라서, 상기 아미노산 치환에 따라, 단백질의 기능에 변화가 없거나, 내성을 갖거나, 양성을 갖거나, 해롭거나, 손상을 주거나, 또는 질병이 유발될 수 있다.

[0038] 아미노산 G는 글리신, 아미노산 A는 알라닌, 아미노산 V는 발린, 아미노산 L은 류신, 아미노산 I는 이소류신, 아미노산 M은 메티오닌, 아미노산 F는 페닐알라닌, 아미노산 W는 트립토판, 아미노산 P는 프롤린, 아미노산 S는 세린, 아미노산 T는 트레오닌, 아미노산 C는 시스테인, 아미노산 Y는 타이로신, 아미노산 N은 아스파라긴, 아미노산 Q는 글루타민, 아미노산 D는 아스파르테이트, 아미노산 E는 글루타메이트, 아미노산 K는 리신, 아미노산 R은 알지닌, 아미노산 H는 히스티딘을 의미한다.

[0039] 상기 제1 폴리펩티드는 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산, 362번째 아미노산, 325번째 아미노산, 179번째 아미노산, 또는 이들의 조합을 포함하는 아미노산 서열과 특이적으로 결합할 수 있다. 상기 399번째 아미노산은 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드인 A가 G로 치환되어 생성된 아미노산일 수 있다. 상기 362번째 아미노산은 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드인 G가 A로 치환되어 생성된 아미노산일 수 있다. 상기 399번째 아미노산은 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드인 G가 T로 치환되어 생성된 아미노산일 수 있다. 상기 325번째 아미노산은 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드인 T가 C로 치환되어 생성된 아미노산일 수 있다. 상기 179번째 아미노산은 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드인 G가 A로 치환되어 생성된 아미노산일 수 있다.

[0040] 상기 제2 폴리펩티드는 상기 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 5의 N 말단으로부터 356번째 아미노산, 215번째 아미노산, 95번째 아미노산, 또는 이들의 조합을 포함하는 아미노산 서열과 특이적으로 결합할 수 있다. 상기 356번째 아미노산은 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드인 C가 A로 치환되어 생성된 아미노산일 수 있다. 상기 215번째 아미노산은 서열번호 2의 5' 말단으로부터 644번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 2의 5' 말단으로부터 644번째 뉴클레오티드인 A가 G로 치환되어 생성된 아미노산일 수 있다. 상기 95번째 아미노산은 서열번호 2의 5' 말단으로부터 283번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 2의 5' 말단으로부터 283번째 뉴클레오티드인 G가 A로 치환되어 생성된 아미노산일 수 있다.

[0041] 상기 제3 폴리펩티드는 상기 서열번호 6의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 6의 N 말단으로부터 249번째 아미노산을 포함하는 아미노산 서열과 특이적으로 결합할 수 있다. 상기 356번째 아미노산은 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드를 포함하는 폴리뉴클레오티드를 암호화한 것일 수 있고, 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드인 G가 A로 치환되어 생성된 아미노산일 수 있다.

[0042] 상기 조성물에 있어서, 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상

기 서열번호 4의 N 말단으로부터 399번째 아미노산은 D가 G로 치환되거나; 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 362번째 아미노산은 V가 I로 치환되거나; 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산은 D가 Y로 치환되거나; 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 325번째 아미노산은 V가 A로 치환되거나; 상기 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 179번째 아미노산은 W가 X로 치환되거나; 상기 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 5의 N 말단으로부터 356번째 아미노산은 F가 L로 치환되거나; 상기 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 5의 N 말단으로부터 215번째 아미노산은 N이 S로 치환되거나; 상기 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 서열번호 5의 N 말단으로부터 95번째 아미노산은 V가 M으로 치환되거나; 상기 서열번호 6의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 6의 N 말단으로부터 249번째 아미노산은 D가 N으로 치환되거나; 또는 이들의 조합인 것일 수 있다.

[0043] 상기 검출할 수 있는 제1 내지 제3 폴리펩티드 중 어느 하나 이상(이하, '제n 폴리펩티드'라고 함)은 항체 또는 항원 결합 단편인 것일 수 있고, 단수개 또는 복수개일 수 있다.

[0044] 항체는 전체 항체 형태일 뿐 아니라 항체 분자의 기능적인 단편을 포함할 수 있다. 전체 항체는 2개의 전체 길이의 경쇄 및 2개의 전체 길이의 중쇄를 가지는 구조이며 각각의 경쇄는 중쇄와 다이설파이드 결합으로 연결되어 있다. 항체 분자의 기능적인 단편이란 항원 결합 기능을 보유하고 있는 단편을 의미한다.

[0045] 항원 결합 단편은 면역글로불린 전체 구조에 대한 그의 단편으로, 항원이 결합할 수 있는 부분을 포함하는 폴리펩티드의 일부를 의미한다. 예를 들면, scFv 단편, (scFv)₂ 단편, Fab 단편, Fab' 단편, F(ab')₂ 단편 등을 포함할 수 있다. 상기 항원 결합 단편은 단백질 가수분해 효소를 이용해서 얻을 수 있고, 예를 들면, 전체 항체를 파파인으로 제한 절단하면 Fab를 얻을 수 있고 펩신으로 절단하면 F(ab')₂ 단편을 얻을 수 있으며, 유전자 재조합 기술을 통하여 제작할 수 있다.

[0046] 상기 제n 폴리펩티드는 검출 대상이 되는 아미노산 서열을 포함하는 폴리펩티드에 항원-항체 결합으로 결합하는 것일 수 있다. 상기 제n 폴리펩티드는 검출 대상이 되는 폴리펩티드에 약 $1 \times 10^7 \text{ M}^{-1}$, 약 $1 \times 10^8 \text{ M}^{-1}$, 약 $1 \times 10^9 \text{ M}^{-1}$, 약 $1 \times 10^{10} \text{ M}^{-1}$, 또는 약 $1 \times 10^{11} \text{ M}^{-1}$ 의 친화도 상수 (KA)를 가지면서 결합하는 것일 수 있다.

[0047] 상기 제n 폴리펩티드는 면역세포화학 및 면역조직화학, 방사선 면역 분석법 (radioimmunoassays), 효소결합면역법 (ELISA: Enzyme Linked Immunoabsorbent assay), 면역 블롯 (immunoblotting), 파아르 분석법 (Farr assay), 면역침강, 라텍스 응집, 적혈구 응집, 비탁계법, 면역확산법, 카운터-전류 전기영동법, 단일 라디칼 면역확산법, 면역크로마토그래피법, 단백질 칩 및 면역형광법 등을 이용한 방법에 이용될 수 있다. 즉 항원과 항체의 결합을 측정할 수 있는 방법에 이용될 수 있다.

[0048] 다른 양상은 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드, 또는 이들의 조합을 검출할 수 있는 제1 폴리뉴클레오티드; 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합을 검출할 수 있는 제2 폴리뉴클레오티드; 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드를 검출할 수 있는 제3 폴리뉴클레오티드; 또는 이들의 조합을 포함하는 고 밀도지질단백질콜레스테롤 의 수준을 예측하기 위한 키트를 제공한다.

[0049] 상기 키트에 있어서, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드는 A가 G로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드는 G가 T로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드는 T가 C로 치환되거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 2

의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드는 C가 A로 치환되거나; 상기 서열번호 2의 5' 말단으로부터 644번째 뉴클레오티드는 A가 G로 치환되거나; 상기 서열번호 2의 5' 말단으로부터 551번째 뉴클레오티드는 C가 G로 치환되거나; 상기 서열번호 2의 5' 말단으로부터 283번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 107번째 뉴클레오티드는 G가 A로 치환되거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드는 G가 A로 치환되거나; 또는 이들의 조합일 수 있다.

[0050] 상기 키트 (kit)는 HDL-C의 수준을 예측하기 위한 키트일 수 있다. 키트는 당업계에 알려진 의미로 사용된다. 상기 키트는 예를 들면, 상기한 바와 같은 폴리뉴클레오티드와 그의 특정 용도에 필요한 항목들을 포함하는 것일 수 있다. 상기한 바와 같은 폴리뉴클레오티드와 함께 그의 사용 방법에 필요한 시약을 포함하는 것일 수 있다.

[0051] 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드 내지 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드, 각각의 변이, 각각의 변이 위치, 각각의 변이 위치에서의 뉴클레오티드 변경, 제1 폴리뉴클레오티드 내지 제3 폴리뉴클레오티드에 대하여는 상기한 바와 같다.

[0052] 다른 양상은 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 4의 N 말단으로부터 399번째 아미노산, 362번째 아미노산, 325번째 아미노산, 179번째 아미노산, 또는 이들의 조합을 검출할 수 있는 제1 폴리펩티드; 서열번호 5의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 5의 N 말단으로부터 356번째 아미노산, 215번째 아미노산, 95번째 아미노산, 또는 이들의 조합을 검출할 수 있는 제2 폴리펩티드; 서열번호 6의 아미노산 서열을 포함하는 폴리펩티드에서 아미노산 치환 위치인 상기 서열번호 6의 N 말단으로부터 249번째 아미노산을 검출할 수 있는 제3 폴리펩티드; 또는 이들의 조합을 포함하는 HDL-C 수준을 예측하기 위한 키트를 제공한다.

[0053] 상기 키트는 서열번호 4의 아미노산 서열을 포함하는 폴리펩티드 내지 서열번호 6의 아미노산 서열을 포함하는 폴리펩티드 중 어느 하나 이상의 발현 수준을 측정하기 위한, 폴리펩티드의 분석 방법에 적합한 하나 또는 그 이상의 다른 구성 성분 또는 장치가 포함될 수 있다. 상기 아미노산 서열 중 어느 하나 이상의 발현 수준을 측정하는 것은 유전자에서 발현되는 폴리펩티드 또는 단백질의 존재 여부와 발현 정도를 확인하는 과정으로서, 제1 폴리펩티드 내지 제3 폴리펩티드 중 어느 하나 이상을 이용하여 단백질의 양을 확인할 수 있다.

[0054] 다른 양상은 개체의 고밀도지질단백질콜레스테롤 수준을 예측하기 하기 위한 정보를 제공하는 방법으로서, 개체의 생물학적 시료로부터 핵산 시료를 획득하는 단계; 및 획득된 핵산 시료로부터 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드 내지 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드 중 어느 하나 이상의 폴리뉴클레오티드에 존재하는 변이 위치의 유전자형 (genotype)을 분석하는 단계를 포함한다. 상기 방법에 있어서, 상기 변이 위치는, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드 또는 이들의 조합; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드; 또는 이들의 조합인 것이다.

[0055] 상기 방법은 개체의 생물학적 시료로부터 핵산 시료를 획득하는 단계를 포함한다.

[0056] 상기 개체는 인간을 포함한 포유동물일 수 있다.

[0057] 상기 생물학적 시료는 생물로부터 획득된 시료를 말한다. 상기 생물학적 시료는 예를 들면 혈액, 조직, 소변, 점액, 타액, 눈물, 혈장, 혈청, 객담, 척수액, 흉수, 유두 흡인물, 림프액, 기도액, 장액, 비뇨생식관액, 모유, 림프계 체액, 정액, 뇌척수액, 기관계내 체액, 복수, 낭성 종양 체액, 양수액 또는 이들의 조합일 수 있다. 상기 조직은 혈액일 수 있다.

[0058] 상기 생물학적 시료로부터 핵산 시료를 획득하는 단계는 통상의 DNA 분리방법에 의하여 수행될 수 있다. 예를 들면, 표적 핵산을 증폭효소 연쇄 반응 (polymerase chain reaction: PCR), 리가제 연쇄 반응 (ligase chain reaction: LCR), 전사 증폭 (transcription amplification), 또는 실시간-핵산 서열 기초 증폭 (realtime-nucleic acid sequence based amplification: NASBA)을 통하여 증폭하고 이를 정제하여 얻을 수 있다.

- [0059] 상기 방법은 수득된 핵산 시료로부터 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드 내지 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드 중 어느 하나 이상의 폴리뉴클레오티드에 존재하는 변이 위치의 유전자형을 분석하는 단계를 포함한다.
- [0060] 상기 유전자형을 분석하는 단계는 상기 변이 위치에서의 특정 대립유전자의 유무 또는 특정 유전자형의 유무에 근거하여, 개체 또는 피검체가 고중성지방혈증에 걸릴 위험성 또는 고중성지방혈증을 나타내는 환자인지를 판별할 수 있다.
- [0061] 상기 변이 위치는, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드, 또는 이들의 조합, 또는 이들의 조합; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드; 또는 이들의 조합인 것일 수 있다.
- [0062] 상기 방법에 있어서, 상기 유전자형을 분석하는 단계는, 수득된 핵산 시료로부터 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드 내지 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드 중 어느 하나 이상의 폴리뉴클레오티드에 존재하는 변이 위치의 유전자형을 분석하는 단계를 더 포함할 수 있다.
- [0063] 상기 변이 위치는, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드, 1084번째 뉴클레오티드, 1195번째 뉴클레오티드, 974번째 뉴클레오티드, 537번째 뉴클레오티드, 또는 이들의 조합; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드, 644번째 뉴클레오티드, 283번째 뉴클레오티드, 또는 이들의 조합; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드; 또는 이들의 조합인 것일 수 있다.
- [0064] 상기 방법에 있어서, 상기 유전자형을 분석하는 단계는, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드는 A 또는 G이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드는 G 또는 A이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드는 G 또는 T이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드는 T 또는 C이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드는 G 또는 A이거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드는 C 또는 A이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 644번째 뉴클레오티드는 A 또는 G이거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 283번째 뉴클레오티드는 G 또는 A이거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오티드에서 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오티드는 G 또는 이거나; 또는 이들의 조합인지 여부를 결정하는 단계를 더 포함하는 것일 수 있다.
- [0065] 상기 결정하는 단계에 따라, 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1196번째 뉴클레오티드가 G이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1084번째 뉴클레오티드가 A이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 1195번째 뉴클레오티드가 T이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 974번째 뉴클레오티드가 C이거나; 상기 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 1의 5' 말단으로부터 537번째 뉴클레오티드가 A이거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 1068번째 뉴클레오티드가 A이거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오티드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 644번째

뉴클레오타이드가 G이거나; 상기 서열번호 2의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 2의 5' 말단으로부터 283번째 뉴클레오타이드가 A이거나; 상기 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오타이드에서 단일염기 변이 위치인 상기 서열번호 3의 5' 말단으로부터 745번째 뉴클레오타이드가 A이거나; 또는 이들의 조합인 경우, 개체의 HDL-C의 수준이 높은 것으로 판단할 수 있다.

[0066] 상기 개체의 HDL-C의 수준을 예측하기 위한 정보를 제공하는 것은 HDL-C의 수준과 관련성이 높은 심혈관질환의 발병의 상대적인 위험을 예측 또는 진단하기 위한 것일 수 있다. 예를 들면, 상기 위험은 참조 유전체 서열을 갖고 있는 군에 비하여 HDL-C의 수준이 높으면 심혈관 질환의 발병 가능성이 증가되어 있는지를 예측 또는 진단하기 위한 것일 수 있다. 또한, 상기 위험은 다른 특정 대립유전자 또는 유전자형에 대한 특정 대립유전자 또는 유전자형을 갖는 개체에서의 위험일 수 있다. 상기 변이 위치에서 변이의 수가, 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 또는 10개 이상인 경우, HDL-C의 수준이 높아 이와 관련성이 높은 심혈관 질환의 발병 위험이 높다고 판단할 수 있으며, 예측에 대한 정확도가 높아질 수 있다.

[0067] 상기 결정하는 단계에서, 개체의 HDL-C의 수준이 높은 것은 심혈관질환이 발병할 위험이 높은 것으로 판단할 수 있다. 예를 들면, 개체의 혈장에서 HDL-C 농도가, HDL-C > 100 mg/dL, HDL-C > 150 mg/dL, HDL-C > 200 mg/dL, HDL-C > 250 mg/dL, HDL-C > 300 mg/dL, HDL-C > 350 mg/dL, HDL-C > 400 mg/dL, HDL-C > 450 mg/dL, HDL-C > 500 mg/dL, HDL-C > 550 mg/dL 일 가능성이 높은 것으로 판단할 수 있다.

[0068] 서열번호 1의 핵산 서열을 포함하는 폴리뉴클레오타이드 내지 서열번호 3의 핵산 서열을 포함하는 폴리뉴클레오타이드, 각각의 변이, 각각의 변이 위치, 각각의 변이 위치에서의 뉴클레오타이드에 대하여는 상기한 바와 같다.

[0069] 상기 방법에 있어서, 상기 유전형을 분석하는 단계는 뉴클레오타이드 또는 염기 서열을 결정하는 방법으로서, 시퀀싱, 마이크로어레이에 의한 혼성화, 대립유전자 특이적인 PCR (allele specific PCR), 다이내믹 대립유전자 혼성화 (dynamic allele-specific hybridization), PCR 연장 분석, PCR-SSCP (PCR-single strand conformation polymorphism) 및 TaqMan 방법으로 이루어진 군으로부터 선택된 하나 이상의 기법에 의해 수행되는 것일 수 있다. 상기 유전형을 분석하는 단계는 상기 조성물 또는 키트를 이용할 수 있다.

[0070] 시퀀싱 분석은 염기서열 결정을 위한 통상적인 방법을 사용할 수 있으며, 자동화된 유전자 분석기를 이용하여 수행될 수 있다. 대립유전자 특이적 PCR은 단일염기 변이가 위치하는 염기를 3' 말단으로 하여 고안한 프라이머를 포함한 프라이머 세트로 상기 단일염기 변이가 위치하는 DNA 단편을 증폭하는 PCR 방법을 의미한다. PCR 연장 분석은 먼저 단일염기 변이가 위치하는 염기를 포함하는 DNA 단편을 프라이머 쌍으로 증폭한 다음, 반응에 첨가된 모든 뉴클레오타이드를 탈인산화시킴으로써 불활성화시키고, 여기에 단일염기 변이에 특이적인 연장 프라이머, dNTP 혼합물, 디디옥시뉴클레오타이드, 반응 완충액 및 DNA 중합효소를 첨가하여 프라이머 연장반응을 수행함으로써 이루어진다. TaqMan 방법은 원하는 DNA 단편을 증폭할 수 있도록 프라이머 및 TaqMan 탐침을 설계 및 제작하고, 서로 다른 대립유전자의 탐침을 FAM 및 VIC로 표지 (Applied Biosystems)하여, 증폭 및 분석하는 단계로 수행된다.

[0071] 상기 조성물, 키트 또는 방법은, 아시아인, 한국인, 40세 이상의 한국인 또는 45세 이상의 고밀도지질단백질콜레스테롤의 수준을 측정하여 상기 고밀도지질단백질콜레스테롤과 관련성이 깊은 개체의 심혈관질환의 위험도를 측정할 수 있어, 상기 질환의 위험을 예방할 수 있다. 따라서, 상기 조성물, 키트 또는 방법은, 아시아인, 한국인, 40세 이상의 한국인 또는 45세 이상의 한국인의 고밀도지질단백질콜레스테롤의 수준의 측정을 통하여 심혈관 질환의 발병 위험을 예측 또는 진단하는데 유용하게 사용될 수 있다.

발명의 효과

[0072] 일 양상에 따른 조성물 또는 키트, 및 이를 이용한 개체의 고밀도지질단백질콜레스테롤의 수준을 예측하기 위한 정보를 제공하는 방법에 의하면, 개체의 고밀도지질단백질콜레스테롤의 수준을 예측할 수 있어, 상기 고밀도지질단백질콜레스테롤의 예측을 통하여 최종적인 개체의 심혈관 질환을 예측 또는 측정할 수 있다. 예측 또는 측정된 개체의 심혈관 질환의 발병 위험 고밀도지질단백질콜레스테롤의 수준을 조기에 측정하여 심혈관 질환의 발병을 예방하기 위한 정보로 이용될 수 있고, 개인 맞춤형의학 치료에 적용하여 치료 효과를 증대시킬 수 있다.

도면의 간단한 설명

[0073] 도 1은 3가지 유전자(CETP, LIPC 및 SCARB1)에서 확인한 변이의 비율을 확인한 결과를 나타낸 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0074] 본 발명은 하기 실시예에 의하여 더욱 구체적으로 설명한다. 그러나, 하기 실시예는 본 발명의 이해를 돕기 위한 것일 뿐, 어떤 의미로든 본 발명의 범위가 이러한 실시예에 의하여 한정되는 것은 아니다.

[0075] 실시예 1. 실험에 참여한 피험자의 선정 및 상기 피험자의 임상적인 특징의 확인

[0076] 높은 수준의 고밀도지질단백질-콜레스테롤(high-density lipoprotein-cholesterol : HDL-C)을 가진 피험자를 피험자로 하여 상기 피험자의 서면 동의를 받은 후 실험을 수행하였다. 2000 년 11 월부터 2011 년 3 월까지 연세대학교 의과 대학 심혈관 유전체센터 코호트를 방문한 사람 중 13545 명을 모집단 피험자로 등록하였다. 피험자들은 자신의 병력에 대한 인터뷰를 수행한 후 신체 검사를 받았고, 이중 HDL-C 수치가 100 mg / dL 이 넘는 42 명의 피험자를 최종적으로 표본으로 선택하였다. 상기 피험자들은 본 연구에 등록하기 전에 지질 강하 요법을 받지 않았으며, 갑상선, 간 또는 신장 질환, 암을 앓고 있거나 임신을 하거나 또는 약물 요법 (지질 개질제, 코르티코스테로이드(corticosteroid) 또는 경구 에스트로겐과 같은)을 받고 있어 본 연구에 영향을 줄 수 있는 사람들은 피험자에서 제외하였다. 총 콜레스테롤, 트리글리세라이드(triglyceride), HDL-C 및 LDL-C의 수준을 모든 피험자로부터 측정하였다. 참가자들은 측정 전에 금식을 하고 적어도 12시간 동안 금주하였다. 이후 상기 피험자들로부터 수득한 샘플을 한국 임상 시험 학회에서 인증한 실험실에서 4시간 이내에 분석하였다.

[0077] 42명의 피험자의 임상적인 특징을 표 1에 나타내었다.

[0078] [표 1]

	Total population (n=13,545)	Patients with very high HDL-C (n=42)	P
Age, years	60.4 ± 10.6	54.1 ± 12.8	<.001
Male	6,722 (50)	16 (38)	0.12
Medical history			
Hypertension	7,234 (53)	9 (21)	<.001
DM	2,293 (17)	2 (5)	<.001
Smoking	2,004 (15)	11 (26)	0.09
CAD	4,747 (35)	3 (7)	<.001
BMI, kg/m ²	24.8 ± 3.1	22.6 ± 4.2	0.002
Lab values, mg/dL			
TC	189 ± 43	219.9 ± 38.1	<.001
TG	117 (83)	62.5 (27)	<.001
HDL-C	48.8 ± 14.7	110.1 ± 12.6	<.001
LDL-C	115 ± 38	101 ± 35	0.02

[0079]

[0080] 상기 피험자의 평균 연령은 54 였고, 남성은 38 % 였다. 피험자의 병력에서 당뇨병이 2 명 (5 %), 관상 동맥 질환이 있는 당뇨병 환자가 3 명 (7 %)이었다. HDL-C의 평균 수준은 110.1 ± 12.6 mg / dL이었다. 종합적으로, HDL-C 수준이 높은 피험자들은 모집단 보다 더 젊었고 심장 혈관 위험 관련 질환에 대한 유병률은 낮았고, 콜레스테롤 수치는 표본 피험자군이 높았으나, 트리글리세라이드 수치는 모집단 보다 낮았음을 확인할 수 있었다.

[0081] 실시예 2. 타겟 시퀀싱 및 변이 분석

[0082] CETP (MIM 118470), LIPC (MIM 151670) 및 SCARB1 (MIM 601040)의 3가지 표적 유전자를 서열분석하였다. Qiagen DNeasy 키트 (Qiagen, Valencia, CA, USA)를 사용하여 피험자의 혈액으로부터 게놈 DNA를 추출하였다. 돌연변이 분석을 위해, 셀레믹스(서울, 한국)에 의하여 개발된 패널을 표적 DNA 캡처 및 시퀀싱을 위하여 사용하였다. 표적 서열 분석 및 변이 분석은 하기와 같이 수행하였다. 간단하게, 모든 코딩 엑손과 엑손 - 인트론

접합부가 포함된 DNA 단편은 솔루션 기반 하이브리드화 캡처를 통하여 강화되었으며, 이후 Illumina HiSeq 2000 platform (Illumina, Inc., San Diego, CA, USA)로 시퀀싱을 수행하였다.

[0083] 시퀀싱 데이터 분석은 자체 분석 파이프 라인을 사용하여 수행되었다. 간단하게, HiSeq 2000의 미가공 데이터의 시퀀싱 리딩을 인덱스 및 바코드 시퀀스별로 정렬하였다. 정렬된 fastq 파일은 Burrows-Wheeler Aligner (BWA; ver 0.7.12) BWA-MEM 알고리즘을 사용하여 hg19 참조 게놈에 맞춰서 정렬하였다. 출력 SAM 파일은 SAMtools (버전 1.1)를 사용하여 BAM 파일로 변환되고 정렬되었다. 중복 제거는 Picard tools(버전 1.128)에 의하여 수행되었다. 알려진 BAT 파일을 만들기 위해 GATK (v3.3.0)를 사용하여 알려진 indel 사이트 및 BQSR (Base Quality Score Recalibration)에 대한 재배치 작업을 수행하였다. 변이는 50X 이상의 시퀀싱 깊이를 갖는 loci에 대해 GATK v3.3.0 Unified Genotyper 알고리즘을 사용하여 결정되었다. 충분한 인트론 염기를 포함한 스플라이스 영역의 분석을 Human Splicing Finder를 사용하여 수행하였다.

[0084] 유전 변이체의 기능적 주석은 ANNOVAR (ver.2014-11-12)에 의해 수행되었다. SIFT, PolyPhen-2 및 MutationTaster를 사용하여 한국인 exome 데이터 (n = 476) 및 공개형 데이터베이스인 dsSNP 138(Exome 변이 서버 및 1000 게놈 프로젝트)에서 단일 뉴클레오타이드 변이체에 대한 기능적 효과 예측을 수행하였다. 이후 다음과 같은 기준에 따라 변이에 대한 우선 순위를 결정하였다.

[0085] 1) 인간 유전자 변이 데이터베이스에서 질병을 일으킨 것으로 보고된 변이형; 2) 새롭거나 드문 유전자의 기능을 파괴하는 변이(년센스, 스플라이스 부위 [인트론 / 엑손 경계의 양측에 2 개의 뉴클레오타이드] 및 프레임 쉬프트); 3) 세 가지 예측 프로그램 중 하나 이상에서 유전자형 기능을 파괴할 것으로 예측되었던 새롭거나 드문 미스센스 변이.

[0086] 상기 기준을 충족시키는 변이는 PCR 증폭 산물의 양방향 Sanger 시퀀싱에 의해 검증되었다. 변종의 개체 및 빈도를 측정하기 위하여 사용된 데이터베이스에는 1000 Genome Project, Exome Sequencing Project 6500 및 gnomAD 브라우저 (<http://gnomad.broadinstitute.org/>)가 포함되었다. 변이는 공공 데이터베이스에서 대립유전자형빈도(minor allele frequency : MAF)가 1 % 미만인 경우 희소한 것으로 분류하였고, MAF가 5 % 이상인 경우 흔한 것으로 분류하였다.

[0087] 상기 분석을 통하여 피험자 군에서 확인한 세가지 후보군 유전자에 대한 변이를 표 2에 나타내었으며, 3가지 유전자에서 확인한 변이의 비율을 확인한 결과를 도 1에 나타내었다.

[0088] [표 2]

Gene	Genomic coordinate	Nucleotide change	Mutation type	Amino acid change (rs number in dbSNP)	Allele frequency	Frequency in gnomAD Database	Affected patients (homo/hetero)	Report	SIFT/Polyphen/Mutation taster prediction (Clinical significance based on ClinVar)
<i>CETP</i>	Rare								
	chr16: 57,017,292	c.A1196G	nonsynonymous SNV	p.D399G	12 (0.286)	0.0002-0.0298	1/11	Yes	Deleterious/ possibly damaging/ polymorphism
	Common								
	chr16: 57,016,092	c.G1084A	nonsynonymous SNV	p.V362I	38 (0.881)	0.5339-0.564	15/23	Yes	Tolerated/ benign/ polymorphism_automatic
	Unknown frequency								
	chr16: 57,017,291	c.G1195T	nonsynonymous SNV	p.D399Y	1 (0.024)	NA	0/1	No	Deleterious/ probably damaging/ polymorphism
	chr16: 57,015,077	c.T974C	nonsynonymous SNV	p.V325A	1 (0.024)	NA	0/1	No	Tolerated/ benign/ polymorphism
	chr16: 57,004,954	c.G537A	stopgain SNV	p.W179X	1 (0.024)	NA	0/1	No	
<i>LIPC</i>	Common								
	chr15: 58,853,079	c.C1068A	nonsynonymous SNV	p.F356L	42 (1.000)	0.9361-0.9857	32/10	Yes	Tolerated/ benign/ polymorphism_automatic
	chr15: 58,838,010	c.A644G	nonsynonymous SNV	p.N215S	42 (1.000)	0.4697-0.792	30/12	Yes	Deleterious/ benign/ polymorphism_automatic
	chr15: 58,833,993	c.G283A	nonsynonymous SNV	p.V95M	16 (0.381)	0.0436-0.2384	0/16	Yes	Tolerated/ benign/ polymorphism_automatic
	Unknown frequency								
<i>SCARB1</i>									
	Chr12: 125,294,817	c.G745A	nonsynonymous SNV	p.D249N	1 (0.024)	NA	0/1	No	

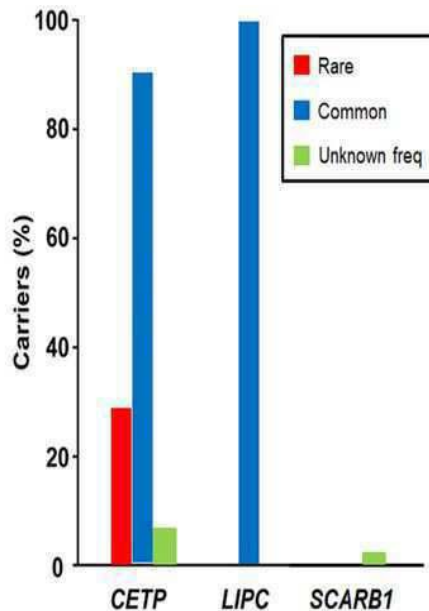
[0089] 상기 표 2에서 확인한 바와 같이, 변이를 분석한 결과 HDL-C 수준이 높은 피험자들에게서는 CETP 유전자에서 흔한 변이와 LIPC 유전자에서 일반적으로 변이가 빈번하게 발생하는 것을 확인하였다. CETP 유전자에서는 두 가지 희귀 변이가 13 명(31 %)에게서 확인되었으며, 상기 희귀 변이는 c.A1196G (p.D399G)로 이는 12 명이었고, c.G1195T (p.D399Y)는 1명이었다. c.G1084A (p.V362I)의 변이는 흔한 변이로서, 38 명의 피험자에서 확인되었다. 변이 운반체 중 15 개는 동형 접합 형태로 운반되는 것을 확인할 수 있었다. CETP 유전자에서 나타난 2 개의 변이는 임상적으로 유의한 변이로서, 유전자 산물의 기능을 변형시키는 변이인 것으로 확인되었다.

[0091] 종합적으로, HDL-C 수준이 높은 피험자들에서 3개의 유전자의 4개의 흔한 유전자 변이가 확인되었으며, 총 42명

의 피험자 중에서 38, 42 및 42 명의 피험자에게서 확인한 흔한 유전자 변이는 CETP 유전자의 cG1084A [p.V362I]의 변이와 LIPC 유전자의 c.C1068A [p.F356L] 및 c.A644G [p.N215S] 변이인 것을 확인하였다. 따라서, 상기와 같이 CETP, LIPC 및 SCARB1 유전자에서 확인한 단일형성다염기를 확인하면 HDL-C 수준이 높아 심혈관질환의 발병 위험이 높은 고위험군을 손쉽게 진단할 수 있음을 확인할 수 있었다.

도면

도면1



서열 목록

<110> Industry-Academic Cooperation Foundation, Yonsei University

<120> Compositions, kits for predicting high density lipoprotein cholesterol levels, and method using the same

<130> PN131761

<160> 6

<170> KoPatentIn 3.0

<210> 1

<211> 1302

<212> DNA

<213> Homo sapiens

<400> 1

atgctggctg ccacagtcct gaccctggcc ctgctgggca atgcccattg ctgctccaaa 60

ggcacctgc acgaggcagg catcgtgtgc cgcattacca agcctgcct cctggtgttg 120

aaccacgaga ctgccaaggt gatccagacc gccttcagc gagccagcta cccagatatc 180

acgggcgaga aggccatgat gctccttggc caagtcaagt atgggttgca caacatccag 240
atcagccact tgtccatcgc cagcagccag gtggagctgg tggaagccaa gtccattgat 300
gtctccattc agaacgtgtc tgtgtgtcttc aaggggaccc tgaagtatgg ctacaccact 360
gcctgggtggc tgggtattga tcagtcatt gacttcgaga tcgactctgc cattgacctc 420
cagatcaaca cacagctgac ctgtgactct ggtagagtgc ggaccgatgc cctgactgc 480
tacctgtctt tccataagct gctcctgcat ctccaagggg agcgagagcc tgggtggatc 540
aagcagctgt tcacaaattt catctccttc accctgaagc tggctctgaa gggacagatc 600

tgcaaagaga tcaacgtcat ctctaacatc atggccgatt ttgtccagac aagggtgccc 660
agcatccttt cagatggaga cattgggggtg gacatttccc tgacaggtga tcccgtcatc 720
acagcctcct acctggagtc ccatcacaag gcagtgtctg agacctgggg cttcaacacc 780
aaccaggaaa tcttccaaga ggttgtcggc ggcttcccca gccaggccca agtcaccgtc 840
cactgcctca agatgcccaa gatctcctgc caaaacaagg gagtctgtgt caattcttca 900
gtgatggatga aattcctctt tccagccca gaccagcaac attctgtagc ttacacattt 960
gaagaggata tcgtgactac cgtccaggcc tctatttcta agaaaaagct cttcttaagc 1020

ctcttggatt tccagattac accaaagact gtttccaact tgactgagag cagctccgag 1080
tccgtccaga gcttctgca gtcaatgatc accgctgtgg gcatccctga ggtcatgtct 1140
cggtcagagg tagtgtttac agcctcatg aacagcaaag gcgtgagcct cttcgacatc 1200
atcaaccctg agattatcac tcgagatggc ttctgtctgc tgcagatgga ctttggtctc 1260
cctgagcacc tgctggtgga tttctccag agcttgagct ag 1302

<210> 2

<211> 1500

<212> DNA

<213> Homo sapiens

<400> 2

atggacacaa gtccctgtg tttctccatt ctgttggttt tatgcatctt tatccaatca 60

agtgccttg gacaaagcct gaaaccagag ccatttggaa gaagagctca agctgttgaa 120
acaaacaaaa cgctgcatga gatgaagacc agattcctgc tctttggaga aaccaatcag 180
ggctgtcaga ttcgaatcaa tcatccggac acgttacagg agtgcggctt caactcctcc 240
ctgcctctgg tgatgataat ccacgggtgg tcggtggacg gcgtgctaga aaactggatc 300
tggcagatgg tggccgcgt gaagtctcag ccggcccagc cagtgaacgt ggggctggtg 360
gactggatca ccttgccca cgaccactac accatcgccg tccgcaacac ccgccttgtg 420

ggcaaggagg tcgcggctct tctccggtgg ctggaggaat ctgtgcaact ctctcgaagc 480

catgttcacc taattgggta cagcctgggt gcacacgtgt caggatttgc cggcagttcc 540

atcggtgga cgcacaagat tgggagaatc acagggtctg atgccgcggg acctttgttt 600

gagggaagtg cccccagcaa tcgtctttct ccagatgatg ccaattttgt ggatgccatt 660

cataccttta cccgggagca catgggcctg agcgtgggca tcaaacagcc cataggacac 720

tatgacttct atcccaacgg gggctccttc cagcctggct gccacttctt agagctctac 780

agacatattg ccagcacagg cttcaatgcc atcaccaga ccaataaatg ctcccacgag 840

cgatcggtgc accttttcat cgactccttg ctgcacgccg gcacgcagag catggcctac 900

ccgtgtggtg acatgaacag cttcagccag ggcctgtgcc tgagctgcaa gaagggccgc 960

tgcaacacgc tgggctacca cgtccgccag gagccgcgga gcaagagcaa gaggtcttcc 1020

ctcgtaacgc gagcccagtc ccccttcaaa gtttatcatt accagttcaa gatccagttc 1080

atcaacaaa ctgagacacc aatacaaca acttttacca tgtcactact cggaacaaaa 1140

gagaaaatgc agaaaattcc catcactctg ggcaaaggaa ttgctagtaa taaaacgtat 1200

tcctttctta tcacgttga tgtggatgc ggcgagctga tcatgatcaa gttcaagtgg 1260

gaaaacagtg cagtgtgggc caatgtctgg gacacggtcc agaccatcat cccatggagc 1320

acagggccgc gccactcagg cctcgttctg aagacgatca gagtcaaagc aggagaaacc 1380

cagcaaaaga tgacattttg ttcagaaaac acagatgacc tactacttcg cccaaccag 1440

gaaaaaatct tcgtgaaatg tgaataaag tctaaaacat caaagcgaaa gatcagatga 1500

1500

<210> 3

<211> 1530

<212> DNA

<213> Homo sapiens

<400> 3

atgggctgct ccgccaagc gcgctgggct gccggggcgc tgggcgtcgc ggggctactg 60

tcgctgtgc tgggcgctgt catgatctg atggtgccgt cgctcatcaa gcagcaggtc 120

cttaagaacg tgcgcatga cccagtagc ctgtccttca acatgtggaa ggagatccct 180

atccccctct atctctccgt ctacttcttt gacgtcatga accccagcga gatcctgaag 240

ggcgagaagc cgcaggtgcg ggagcgcggg ccctacgtgt acaggagatt caggcacaaa 300

agcaacatca ccttcaaca caacgacacc gtgtccttcc tcgagtaccg caccttcag 360

ttccagccct ccaagtccca cggctcggag agcgactaca tcgtcatgcc caacatcctg 420

gtcttgggtg cggcgggtgat gatggagaat aagcccatga ccttgaagct catcatgacc 480
ttggcattca ccacctcgg cgaacgtgcc ttcattgaacc gcactgtggg tgagatcatg 540

tggggtaca aggacccct tgtgaatctc atcaacaagt actttccagg catgttcccc 600
ttcaaggaca agttcggatt atttctgag ctcaacaact ccgactctgg gctcttcacg 660
gtgttcacgg ggttcagaa catcagcagg atccacctcg tggacaagtg gaacgggctg 720
agcaaggttg acttctggca ttccgatcag tgcaacatga tcaatggaac ttctgggcaa 780
atgtggccgc ccttcatgac tctgagtcc tcctggagt tctacagccc ggaggcctgc 840
cgatccatga agctaata caaggagtca ggggtgtttg aaggcatccc cacctatcgc 900
ttctgggtc caaaaacct gtttgccaac gggtccatct acccacccaa cgaaggcttc 960

tgcccggtgcc tggagtctgg aattcagaac gtcagcacct gcaggttcag tgcccccttg 1020
tttctctccc atctcactt cctcaacgt gaccgggttc tggcagaagc ggtgactggc 1080
ctgcacccta accaggaggc acatccttg ttcttgga tccacccggt cacgggaatc 1140
cccatgaact gctctgtgaa actgcagctg agcctctaca tgaaatctgt cgcaggcatt 1200
ggacaaactg ggaagattga gcctgtggtc ctgccgtgc tctggtttgc agagagcggg 1260
gccatggagg gggagactct tcacacattc tacactcagc tgggtttgat gccaaggtg 1320
atgcactatg cccagtacgt cctctggcg ctgggtgcg tctgctgct ggtccctgct 1380

atctgcaaaa tccggagcca agagaaatgc tatttatatt ggagtagtag taaaaagggc 1440
tcaaaggata aggaggccat tcaggcctat tctgaatccc tgatgacatc agctcccaag 1500
ggctctgtgc tgcaggaagc aaaactgtag 1530

<210> 4

<211> 433

<212> PRT

<213> Homo sapiens

<400> 4

Met Leu Ala Ala Thr Val Leu Thr Leu Ala Leu Leu Gly Asn Ala His

1 5 10 15

Ala Cys Ser Lys Gly Thr Ser His Glu Ala Gly Ile Val Cys Arg Ile

20 25 30

Thr Lys Pro Ala Leu Leu Val Leu Asn His Glu Thr Ala Lys Val Ile

35 40 45

Gln Thr Ala Phe Gln Arg Ala Ser Tyr Pro Asp Ile Thr Gly Glu Lys

50 55 60
 Ala Met Met Leu Leu Gly Gln Val Lys Tyr Gly Leu His Asn Ile Gln
 65 70 75 80
 Ile Ser His Leu Ser Ile Ala Ser Ser Gln Val Glu Leu Val Glu Ala
 85 90 95
 Lys Ser Ile Asp Val Ser Ile Gln Asn Val Ser Val Val Phe Lys Gly

 100 105 110
 Thr Leu Lys Tyr Gly Tyr Thr Thr Ala Trp Trp Leu Gly Ile Asp Gln
 115 120 125
 Ser Ile Asp Phe Glu Ile Asp Ser Ala Ile Asp Leu Gln Ile Asn Thr
 130 135 140
 Gln Leu Thr Cys Asp Ser Gly Arg Val Arg Thr Asp Ala Pro Asp Cys
 145 150 155 160
 Tyr Leu Ser Phe His Lys Leu Leu Leu His Leu Gln Gly Glu Arg Glu
 165 170 175

 Pro Gly Trp Ile Lys Gln Leu Phe Thr Asn Phe Ile Ser Phe Thr Leu
 180 185 190
 Lys Leu Val Leu Lys Gly Gln Ile Cys Lys Glu Ile Asn Val Ile Ser
 195 200 205
 Asn Ile Met Ala Asp Phe Val Gln Thr Arg Ala Ala Ser Ile Leu Ser
 210 215 220
 Asp Gly Asp Ile Gly Val Asp Ile Ser Leu Thr Gly Asp Pro Val Ile
 225 230 235 240
 Thr Ala Ser Tyr Leu Glu Ser His His Lys Ala Val Leu Glu Thr Trp

 245 250 255
 Gly Phe Asn Thr Asn Gln Glu Ile Phe Gln Glu Val Val Gly Gly Phe
 260 265 270
 Pro Ser Gln Ala Gln Val Thr Val His Cys Leu Lys Met Pro Lys Ile
 275 280 285
 Ser Cys Gln Asn Lys Gly Val Val Val Asn Ser Ser Val Met Val Lys
 290 295 300
 Phe Leu Phe Pro Arg Pro Asp Gln Gln His Ser Val Ala Tyr Thr Phe

305	310							315							320			
Glu	Glu	Asp	Ile	Val	Thr	Thr	Val	Gln	Ala	Ser	Tyr	Ser	Lys	Lys	Lys			
325				330							335							
Leu	Phe	Leu	Ser	Leu	Leu	Asp	Phe	Gln	Ile	Thr	Pro	Lys	Thr	Val	Ser			
340				345							350							
Asn	Leu	Thr	Glu	Ser	Ser	Ser	Glu	Ser	Val	Gln	Ser	Phe	Leu	Gln	Ser			
355				360							365							
Met	Ile	Thr	Ala	Val	Gly	Ile	Pro	Glu	Val	Met	Ser	Arg	Leu	Glu	Val			
370				375							380							
Val	Phe	Thr	Ala	Leu	Met	Asn	Ser	Lys	Gly	Val	Ser	Leu	Phe	Asp	Ile			
385				390							395					400		
Ile	Asn	Pro	Glu	Ile	Ile	Thr	Arg	Asp	Gly	Phe	Leu	Leu	Leu	Gln	Met			
405				410							415							
Asp	Phe	Gly	Phe	Pro	Glu	His	Leu	Leu	Val	Asp	Phe	Leu	Gln	Ser	Leu			
420				425							430							
Ser																		
<210>	5																	
<211>	499																	
<212>	PRT																	
<213>	Homo sapiens																	
<400>	5																	
Met	Asp	Thr	Ser	Pro	Leu	Cys	Phe	Ser	Ile	Leu	Leu	Val	Leu	Cys	Ile			
1	5				10							15						
Phe	Ile	Gln	Ser	Ser	Ala	Leu	Gly	Gln	Ser	Leu	Lys	Pro	Glu	Pro	Phe			
20				25							30							
Gly	Arg	Arg	Ala	Gln	Ala	Val	Glu	Thr	Asn	Lys	Thr	Leu	His	Glu	Met			
35				40							45							
Lys	Thr	Arg	Phe	Leu	Leu	Phe	Gly	Glu	Thr	Asn	Gln	Gly	Cys	Gln	Ile			
50				55							60							
Arg	Ile	Asn	His	Pro	Asp	Thr	Leu	Gln	Glu	Cys	Gly	Phe	Asn	Ser	Ser			
65	70				75							80						

Leu Pro Leu Val Met Ile Ile His Gly Trp Ser Val Asp Gly Val Leu

85 90 95

Glu Asn Trp Ile Trp Gln Met Val Ala Ala Leu Lys Ser Gln Pro Ala

100 105 110

Gln Pro Val Asn Val Gly Leu Val Asp Trp Ile Thr Leu Ala His Asp

115 120 125

His Tyr Thr Ile Ala Val Arg Asn Thr Arg Leu Val Gly Lys Glu Val

130 135 140

Ala Ala Leu Leu Arg Trp Leu Glu Glu Ser Val Gln Leu Ser Arg Ser

145 150 155 160

His Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ser Gly Phe

165 170 175

Ala Gly Ser Ser Ile Gly Gly Thr His Lys Ile Gly Arg Ile Thr Gly

180 185 190

Leu Asp Ala Ala Gly Pro Leu Phe Glu Gly Ser Ala Pro Ser Asn Arg

195 200 205

Leu Ser Pro Asp Asp Ala Asn Phe Val Asp Ala Ile His Thr Phe Thr

210 215 220

Arg Glu His Met Gly Leu Ser Val Gly Ile Lys Gln Pro Ile Gly His

225 230 235 240

Tyr Asp Phe Tyr Pro Asn Gly Gly Ser Phe Gln Pro Gly Cys His Phe

245 250 255

Leu Glu Leu Tyr Arg His Ile Ala Gln His Gly Phe Asn Ala Ile Thr

260 265 270

Gln Thr Ile Lys Cys Ser His Glu Arg Ser Val His Leu Phe Ile Asp

275 280 285

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

290 295 300

Met Asn Ser Phe Ser Gln Gly Leu Cys Leu Ser Cys Lys Lys Gly Arg

305 310 315 320

Cys Asn Thr Leu Gly Tyr His Val Arg Gln Glu Pro Arg Ser Lys Ser

325 330 335
 Lys Arg Leu Phe Leu Val Thr Arg Ala Gln Ser Pro Phe Lys Val Tyr
 340 345 350
 His Tyr Gln Phe Lys Ile Gln Phe Ile Asn Gln Thr Glu Thr Pro Ile
 355 360 365
 Gln Thr Thr Phe Thr Met Ser Leu Leu Gly Thr Lys Glu Lys Met Gln

370 375 380
 Lys Ile Pro Ile Thr Leu Gly Lys Gly Ile Ala Ser Asn Lys Thr Tyr
 385 390 395 400
 Ser Phe Leu Ile Thr Leu Asp Val Asp Ile Gly Glu Leu Ile Met Ile
 405 410 415
 Lys Phe Lys Trp Glu Asn Ser Ala Val Trp Ala Asn Val Trp Asp Thr
 420 425 430
 Val Gln Thr Ile Ile Pro Trp Ser Thr Gly Pro Arg His Ser Gly Leu
 435 440 445

Val Leu Lys Thr Ile Arg Val Lys Ala Gly Glu Thr Gln Gln Arg Met
 450 455 460
 Thr Phe Cys Ser Glu Asn Thr Asp Asp Leu Leu Leu Arg Pro Thr Gln
 465 470 475 480
 Glu Lys Ile Phe Val Lys Cys Glu Ile Lys Ser Lys Thr Ser Lys Arg
 485 490 495
 Lys Ile Arg

<210> 6

<211> 509

<212> PRT

<213> Homo sapiens

<400> 6

Met Gly Cys Ser Ala Lys Ala Arg Trp Ala Ala Gly Ala Leu Gly Val

1 5 10 15
 Ala Gly Leu Leu Cys Ala Val Leu Gly Ala Val Met Ile Val Met Val
 20 25 30

Pro Ser Leu Ile Lys Gln Gln Val Leu Lys Asn Val Arg Ile Asp Pro
35 40 45

Ser Ser Leu Ser Phe Asn Met Trp Lys Glu Ile Pro Ile Pro Phe Tyr
50 55 60

Leu Ser Val Tyr Phe Phe Asp Val Met Asn Pro Ser Glu Ile Leu Lys
65 70 75 80

Gly Glu Lys Pro Gln Val Arg Glu Arg Gly Pro Tyr Val Tyr Arg Glu
85 90 95

Phe Arg His Lys Ser Asn Ile Thr Phe Asn Asn Asn Asp Thr Val Ser
100 105 110

Phe Leu Glu Tyr Arg Thr Phe Gln Phe Gln Pro Ser Lys Ser His Gly
115 120 125

Ser Glu Ser Asp Tyr Ile Val Met Pro Asn Ile Leu Val Leu Gly Ala
130 135 140

Ala Val Met Met Glu Asn Lys Pro Met Thr Leu Lys Leu Ile Met Thr
145 150 155 160

Leu Ala Phe Thr Thr Leu Gly Glu Arg Ala Phe Met Asn Arg Thr Val
165 170 175

Gly Glu Ile Met Trp Gly Tyr Lys Asp Pro Leu Val Asn Leu Ile Asn
180 185 190

Lys Tyr Phe Pro Gly Met Phe Pro Phe Lys Asp Lys Phe Gly Leu Phe
195 200 205

Ala Glu Leu Asn Asn Ser Asp Ser Gly Leu Phe Thr Val Phe Thr Gly
210 215 220

Val Gln Asn Ile Ser Arg Ile His Leu Val Asp Lys Trp Asn Gly Leu
225 230 235 240

Ser Lys Val Asp Phe Trp His Ser Asp Gln Cys Asn Met Ile Asn Gly
245 250 255

Thr Ser Gly Gln Met Trp Pro Pro Phe Met Thr Pro Glu Ser Ser Leu
260 265 270

Glu Phe Tyr Ser Pro Glu Ala Cys Arg Ser Met Lys Leu Met Tyr Lys
275 280 285

Glu Ser Gly Val Phe Glu Gly Ile Pro Thr Tyr Arg Phe Val Ala Pro

290

295

300

Lys Thr Leu Phe Ala Asn Gly Ser Ile Tyr Pro Pro Asn Glu Gly Phe

305

310

315

320

Cys Pro Cys Leu Glu Ser Gly Ile Gln Asn Val Ser Thr Cys Arg Phe

325

330

335

Ser Ala Pro Leu Phe Leu Ser His Pro His Phe Leu Asn Ala Asp Pro

340

345

350

Val Leu Ala Glu Ala Val Thr Gly Leu His Pro Asn Gln Glu Ala His

355

360

365

Ser Leu Phe Leu Asp Ile His Pro Val Thr Gly Ile Pro Met Asn Cys

370

375

380

Ser Val Lys Leu Gln Leu Ser Leu Tyr Met Lys Ser Val Ala Gly Ile

385

390

395

400

Gly Gln Thr Gly Lys Ile Glu Pro Val Val Leu Pro Leu Leu Trp Phe

405

410

415

Ala Glu Ser Gly Ala Met Glu Gly Glu Thr Leu His Thr Phe Tyr Thr

420

425

430

Gln Leu Val Leu Met Pro Lys Val Met His Tyr Ala Gln Tyr Val Leu

435

440

445

Leu Ala Leu Gly Cys Val Leu Leu Leu Val Pro Val Ile Cys Gln Ile

450

455

460

Arg Ser Gln Glu Lys Cys Tyr Leu Phe Trp Ser Ser Ser Lys Lys Gly

465

470

475

480

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

485

490

495

Ser Ala Pro Lys Gly Ser Val Leu Gln Glu Ala Lys Leu

500

505