



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2025년01월10일

(11) 등록번호 10-2752108

(24) 등록일자 2025년01월06일

(51) 국제특허분류(Int. Cl.)

C12Q 1/6883 (2018.01)

(52) CPC특허분류

C12Q 1/6883 (2022.01)

C12Q 2600/156 (2013.01)

(21) 출원번호 10-2020-7009809

(22) 출원일자(국제) 2018년09월05일

심사청구일자 2021년09월01일

(85) 번역문제출일자 2020년04월03일

(65) 공개번호 10-2020-0058435

(43) 공개일자 2020년05월27일

(86) 국제출원번호 PCT/US2018/049478

(87) 국제공개번호 WO 2019/050899

국제공개일자 2019년03월14일

(30) 우선권주장

62/554,857 2017년09월06일 미국(US)

(56) 선행기술조사문헌

Drug Discov Today Dis Mech. 2011 Winter;
8(3-4): e63-e69.

Front Immunol. 2012; 3: 322.

World J Gastroenterol. Jan 21, 2008; 14(3):
390-400

(73) 특허권자

리제너론 파마슈티칼스 인코포레이티드

미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리
버 로드 777

(72) 발명자

혼자가-자우리구이, 클라우디아 지.

미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리
버 로드 777

호로윙즈, 줄리에

미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리
버 로드 777

(74) 대리인

김진환, 박지하, 김민철

전체 청구항 수 : 총 44 항

심사관 : 박정웅

(54) 발명의 명칭 단일 면역글로불린 인터류킨-1 수용체 관련(SIGIRR) 변이체 및 이의 용도

(57) 요약

본 개시내용은 절두된 인간 단일 면역글로불린 인터류킨-1 수용체 관련(SIGIRR) 단백질을 암호화하는 변형을 포함하는, cDNA를 비롯한 핵산 분자를 제공한다. 본 개시내용은 215번 위치에 상응하는 위치에서 절두를 포함하는 단리된 및 제조된 인간 SIGIRR 단백질 변이체를 또한 제공한다. 상기 절두, 및 이 변형을 암호화하는 핵산 분자는 조기-발병 염증성 장 질환(EO-IBD)과 연관된다. 본 개시내용은 또한 SIGIRR을 암호화하는 핵산 분자에서의 이러한 변형의 확인에 기초하여 대상체가 EO-IBD를 갖거나 또는 EO-IBD를 발생시킬 위험을 갖는지의 여부를 결정하는 방법을 제공한다.

대표도

유전자	NT 변화	AA 변화	유전	보존	예측	대립유전자 빈도 ExAC
SIGIRR	c.557delA	p.K186fs*31	모 계	자연	손상	0.000471

명세서

청구범위

청구항 1

조기-발병 염증성 장 질환(early-onset inflammatory bowel disease) 또는 조기-발병 염증성 장 질환을 발생시킬 위험(risk)을 갖는 인간 대상체를 확인하는 방법으로서,

상기 대상체로부터 수득된 샘플에서

서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 SIGIRR 단백질; 및

서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 SIGIRR 단백질을 암호화하는 핵산 분자 중 적어도 하나의 존재 또는 부재를 검출하는 단계를 포함하되;

상기 절두된 SIGIRR 단백질 및 상기 절두된 SIGIRR 단백질을 암호화하는 핵산 분자 중 적어도 하나의 존재는 상기 대상체가 조기-발병 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타내는, 방법.

청구항 2

제1항에 있어서, 상기 절두된 SIGIRR 단백질이 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치 중 어느 하나에서 야생형 SIGIRR 단백질과 비교하여 상이한 아미노산을 포함하는, 방법.

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

제1항 또는 제2항에 있어서, 상기 샘플에서의 상기 절두된 SIGIRR 단백질을 암호화하는 상기 핵산 분자의 존재 또는 부재는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 생성시키는 상기 핵산 분자에서의 프레임쉬프트 돌연변이가 있는지의 여부를 결정하는 것에 의해 검출되는, 방법.

청구항 11

제1항 또는 제2항에 있어서, 상기 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 서열분석하는 단계를 포함하고, 서열분석된 핵산 분자는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는, 방법.

청구항 12

제11항에 있어서, 상기 서열분석된 핵산 분자의 일부가 서열번호 9에 따른 186번 위치에 상응하는 위치를 암호화하는 코돈을 포함하는 복수의 위치를 포함하는, 방법.

청구항 13

제11항에 있어서, 상기 검출 단계는 상기 SIGIRR 단백질을 암호화하는 전체 핵산 분자를 서열분석하는 단계를 포함하는, 방법.

청구항 14

제1항 또는 제2항에 있어서, 상기 검출 단계는,

SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키되, 증폭된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 아미노산을 암호화하는 코돈을 포함하는 것인 단계;

상기 증폭된 핵산 분자를 검출 가능한 표지로 표시하는 단계;

상기 표시된 핵산 분자를, 프로브를 포함하는 지지체와 접촉시키되, 상기 프로브는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 엄격한 조건하에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및

상기 검출 가능한 표지를 검출하는 단계를 포함하는, 방법.

청구항 15

제14항에 있어서, 상기 샘플에서의 핵산 분자는 mRNA이고, 검출 단계는 증폭 단계 전에 mRNA를 cDNA로 역전사시키는 단계를 추가로 포함하는, 방법.

청구항 16

제1항 또는 제2항에 있어서, 검출 단계는,

SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키되, 상기 프로브는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 엄격한 조건하에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및

상기 검출 가능한 표지를 검출하는 단계를 포함하는, 방법.

청구항 17

삭제

청구항 18

삭제

청구항 19

제1항 또는 제2항에 있어서, 상기 인간 대상체가 크론병을 갖거나 또는 크론병을 발생시킬 위험을 갖는 것으로 확인되는, 방법.

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

단리된 핵산 분자로서,

인간 단일 면역글로불린 인터류킨-1 수용체 관련단(Single Immunoglobulin Interleukin-1 Receptor Related: SIGIRR) 단백질을 암호화하는 핵산 서열 또는 상기 핵산 서열의 보체를 포함하되,

상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 것인, 단리된 핵산 분자.

청구항 41

삭제

청구항 42

삭제

청구항 43

제40항에 있어서, 상기 핵산 분자가 cDNA인, 단리된 핵산 분자.

청구항 44

삭제

청구항 45

제40항에 있어서, 상기 핵산 분자가 게놈 DNA이고 서열번호 2에 따른 9962번 위치에 상응하는 위치에서 구아닌을 포함하는, 단리된 핵산 분자.

청구항 46

제45항에 있어서, 상기 핵산 분자가 서열번호 2를 포함하는, 단리된 핵산 분자.

청구항 47

제40항에 있어서, 상기 핵산 분자가 mRNA이고 서열번호 4에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는, 단리된 핵산 분자.

청구항 48

제40항에 있어서, 상기 핵산 분자가 mRNA이고 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함하는, 단리된 핵산 분자.

청구항 49

제47항에 있어서, 상기 핵산 분자가 서열번호 4를 포함하는, 단리된 핵산 분자.

청구항 50

제40항에 있어서, 상기 절두된 SIGIRR 단백질이 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치 중 어느 하나에서 야생형 SIGIRR 단백질과 비교하여 상이한 아미노산을 포함하는, 단리된 핵산 분자.

청구항 51

삭제

청구항 52

삭제

청구항 53

백터로서,

제40항에 따른 단리된 핵산 분자를 포함하는 백터.

청구항 54

삭제

청구항 55

삭제

청구항 56

숙주 세포로서,

제40항에 따른 단리된 핵산 분자를 포함하는 숙주 세포.

청구항 57

숙주 세포로서,

제53항에 따른 백터를 포함하는 숙주 세포.

청구항 58

제56항에 있어서, 핵산 서열은 상기 숙주 세포에서 활성화된 프로모터에 작동 가능하게 연결된, 숙주 세포.

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

cDNA로서,

SIGIRR 단백질을 암호화하는 핵산 서열을 포함하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 것인, cDNA.

청구항 63

삭제

청구항 64

제62항에 있어서, 상기 절두된 SIGIRR 단백질이 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치 중 어느 하나에서 야생형 SIGIRR 단백질과 비교하여 상이한 아미노산을 포함하는, cDNA.

청구항 65

삭제

청구항 66

삭제

청구항 67

제62항에 있어서, 상기 cDNA가 서열번호 6에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는, cDNA.

청구항 68

제62항에 있어서, 상기 cDNA가 각각 서열번호 6에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CTA 및 AGC를 포함하는, cDNA.

청구항 69

제62항에 있어서, 상기 cDNA가 서열번호 6을 포함하는, cDNA.

청구항 70

백터로서,

제62항에 따른 cDNA를 포함하는 백터.

청구항 71

삭제

청구항 72

삭제

청구항 73

숙주 세포로서,

제62항에 따른 cDNA를 포함하는 숙주 세포.

청구항 74

숙주 세포로서,

제70항에 따른 백터를 포함하는 숙주 세포.

청구항 75

제73항에 있어서, 상기 cDNA가 상기 숙주 세포에서 활성인 프로모터에 작동 가능하게 연결된, 숙주 세포.

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

단리된 또는 재조합 폴리펩타이드로서,

절두된 SIGIRR 단백질을 포함하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 것인, 단리된 또는 재조합 폴리펩타이드.

청구항 80

삭제

청구항 81

제79항에 있어서, 상기 절두된 SIGIRR 단백질이 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치 중 어느 하나에서 야생형 SIGIRR 단백질과 비교하여 상이한 아미노산을 포함하는, 단리된 또는 재조합 폴리펩타이드.

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

제79항에 있어서, 상기 폴리펩타이드가 이중성 폴리펩타이드에 융합된, 단리된 또는 재조합 폴리펩타이드.

청구항 86

제85항에 있어서, 상기 이중성 폴리펩타이드가 펩타이드 정제 태그, 형광 단백질, 또는 펩타이드 정제 태그 및 형광 단백질 둘 다를 포함하는, 단리된 또는 재조합 폴리펩타이드.

청구항 87

삭제

청구항 88

삭제

청구항 89

조성물로서,

제79항에 따른 단리된 또는 재조합 폴리펩타이드 및 담체를 포함하는 조성물.

청구항 90

프로브 또는 프라이머로 사용하기 위한 핵산 분자로서,

서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상

응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 인간 SIGIRR 단백질을 암호화하는 핵산 서열을 갖는 핵산 분자에 특이적으로 혼성화하거나, 또는 상기 절두된 인간 SIGIRR 단백질을 암호화하는 핵산 서열의 보체에 특이적으로 혼성화하는, 적어도 15개의 뉴클레오타이드를 포함하는 핵산 서열을 포함하는, 핵산 분자.

청구항 91

삭제

청구항 92

삭제

청구항 93

제90항에 있어서, 상기 프로브 또는 프라이머가 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 분자의 일부에 특이적으로 혼성화하는, 핵산 분자.

청구항 94

제90항에 있어서, 상기 핵산 분자가 엄격한 조건하에 절두된 인간 SIGIRR 단백질을 암호화하는 핵산 서열 또는 그의 보체에 특이적으로 혼성화하는, 핵산 분자.

청구항 95

제90항에 있어서, 상기 핵산 분자가 표지를 포함하는, 핵산 분자.

청구항 96

삭제

청구항 97

지지체로서,

제90항에 따른 핵산 분자가 부착된 기질을 포함하는 지지체.

청구항 98

제97항에 있어서, 상기 지지체가 마이크로어레이인, 지지체.

청구항 99

프로브 또는 프라이머로 사용하기 위한 변경-특이적(alteration-specific) 핵산 분자로서,

서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 SIGIRR 단백질을 암호화하는 핵산 분자의 핵산 서열에 상보적인 핵산 서열을 포함하되,

상기 변경-특이적 핵산 분자가 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 분자의 일부에 상보적인 핵산 서열을 포함하는, 변경-특이적 핵산 분자.

청구항 100

제99항에 있어서, 상기 변경-특이적 핵산 분자가 적어도 15개의 뉴클레오타이드를 포함하는, 변경-특이적 핵산 분자.

발명의 설명

기술 분야

관련 출원에 대한 상호 참조

[0001]

[0002] 본 출원은 2017년 9월 6일로 제출된 미국 가출원 제62/554,857호의 우선권을 주장하고, 이는 본원에 그 전문이 참고로 인용된다.

[0003] **서열 목록에 대한 참조**

[0004] 본 출원은 56 킬로바이트의 크기로 2018년 8월 30일에 생성된 18923800602SEQ 명칭의 텍스트 파일로 전자로 제출된 서열 목록을 포함한다. 서열 목록은 본원에 참고로 인용된다.

[0005] **기술분야**

[0006] 본 개시내용은 일반적으로 유전학의 분야에 관한 것이다. 더 특히, 본 개시내용은, 예를 들어 조기-발병 염증성 장 질환(early-onset inflammatory bowel disease: EO-IBD)과 연관된 단일 면역글로불린 인터류킨-1 수용체 관련(Single Immunoglobulin Interleukin-1 Receptor Related: SIGIRR)에서의 유전자 변경 및 폴리펩타이드 변이체에 관한 것이다.

배경 기술

[0007] 다양한 참고문헌, 예를 들어 특허, 특허 출원, 수탁 번호, 기술 저널 및 전문 저널은 본 명세서에 걸쳐 인용된다. 각각의 참고문헌은 모든 목적을 위해 본원에 그 전문이 참고로 인용된다.

[0008] 염증성 장 질환(IBD)은 30세에 평균 발병 연령을 갖는 위장관에서의 편리공생균 미생물총에 대한 부적절한 면역 반응에 의해 개시되는 유전적으로 이질적인 만성 염증성 장애이다. IBD의 심각한 단일유전자 형태는 소아 발병 연령(18세 미만)에 나타날 수 있고, 약 50개의 '헨델' 유전자에서 희귀한 고도로 침투성인 변이체에 기인한다. 그러나, 조기-발병 염증성 장 질환(EO-IBD)의 유전적 구성은 잘 이해되지 않고, 대부분의 환자는 유전적으로 진단되지 않고 있다.

[0009] 본 개시내용은 SIGIRR의 생물학의 이해를 돕고, 조기-발병 염증성 장 질환을 갖는 아동의 진단 및 치료를 용이하게 하는 신규의 SIGIRR 변이체를 제공한다.

발명의 내용

[0010] 본 개시내용은 SIGIRR 변이체 폴리펩타이드를 암호화하는 신규의 핵산 분자(즉, 게놈 DNA, mRNA 및 cDNA) 및 SIGIRR 변이체 폴리펩타이드를 제공하는데, 이들은 염증성 장 질환, 예컨대, 조기-발병 염증성 장 질환과 연관된 것으로 본원에서 입증되었다.

[0011] 본 개시내용은 인간 SIGIRR 단백질을 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하는 단리된 핵산 분자를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0012] 본 개시내용은 또한 인간 SIGIRR 단백질의 적어도 일부를 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하는 게놈 DNA 분자를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0013] 본 개시내용은 또한 인간 SIGIRR 단백질의 적어도 일부를 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하는 cDNA 분자를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0014] 본 개시내용은 또한 인간 SIGIRR 단백질의 적어도 일부를 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하는 mRNA 분자를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0015] 본 개시내용은 또한 본원에 개시된 임의의 단리된 핵산 분자를 포함하는 벡터를 제공한다.

[0016] 본 개시내용은 또한 본원에 개시된 임의의 단리된 핵산 분자 또는 벡터 및 담체를 포함하는 조성물을 제공한다.

[0017] 본 개시내용은 또한 본원에 개시된 임의의 단리된 핵산 분자 또는 벡터를 포함하는 숙주 세포를 제공한다.

[0018] 본 개시내용은 또한 인간 SIGIRR 단백질의 적어도 일부를 포함하는 단리된 또는 재조합 폴리펩타이드를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0019] 본 개시내용은 또한 본원에 개시된 임의의 단리된 또는 재조합 폴리펩타이드 및 담체를 포함하는 조성물을 제공한다.

[0020] 본 개시내용은 또한 적어도 약 5개의 뉴클레오타이드를 포함하는 핵산 서열을 포함하는 프로브 또는 프라이머를

제공하되, 인간 SIGIRR 단백질을 암호화하는 핵산 서열에 혼성화하되, 여기서 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되거나, 또는 인간 SIGIRR 단백질을 암호화하는 핵산 서열의 보체에 혼성화하되, 여기서 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.

[0021] 본 개시내용은 또한 본원에 개시된 임의의 프로브가 혼성화하는 기질을 포함하는 지지체를 제공한다.

[0022] 본 개시내용은 또한 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 핵산 서열에 상보적인 핵산 서열을 포함하는 변경-특이적(alteration-specific) 프로브 또는 프라이머를 제공하되, 상기 변경-특이적 프로브 또는 프라이머는 서열번호 9에 따른 186번 내지 209번 또는 211번 내지 215번 위치에 상응하는 임의의 복수의 위치를 암호화하는 핵산 분자의 일부에 상보적인 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치를 암호화하는 핵산 분자의 일부 또는 그의 보체에 특이적으로 혼성화한다. 변경-특이적 프로브 또는 프라이머는 야생형 SIGIRR 단백질을 암호화하는 핵산 서열을 갖는 핵산 분자에 혼성화하지 않는다.

[0023] 본 개시내용은 또한 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험(risk)을 갖는 인간 대상체를 확인하는 방법을 제공하되, 상기 방법은 대상체로부터 취득된 샘플에서 절두된 SIGIRR 단백질; 및/또는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재 또는 부재를 검출하는 단계를 포함하고; 절두된 SIGIRR 단백질 및/또는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타낸다.

[0024] 본 개시내용은 또한 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는 인간 대상체를 확인하는 방법을 제공하되, 상기 방법은 대상체로부터 취득된 샘플에서 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질; 및/또는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재 또는 부재를 검출하는 단계를 포함하고; 절두된 SIGIRR 단백질 및/또는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타낸다.

[0025] 본 개시내용은 또한 인간 대상체에서 염증성 장 질환 또는 조기-발병 염증성 장 질환을 진단하기 위한 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 검출하기 위한 방법으로서, 인간 대상체로부터 취득된 SIGIRR 단백질을 암호화하는 핵산 분자에서의 변경을 검출하되, 상기 변경은 절두된 SIGIRR 단백질을 암호화하는 것인 단계; 및 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는 방법을 제공한다.

[0026] 본 개시내용은 또한 인간 대상체에서 염증성 장 질환 또는 조기-발병 염증성 장 질환을 진단하기 위한 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 검출하기 위한 방법으로서, 인간 대상체로부터 취득된 SIGIRR 단백질을 암호화하는 핵산 분자에서의 변경을 검출하되, 상기 변경은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 것인 단계; 및 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는 방법을 제공한다.

도면의 간단한 설명

[0027] 본 명세서의 부분에 통합되고 이를 구성하는 동반된 도면은 일부 양태를 예시하고, 설명과 함께 본 개시내용의 원칙을 설명하도록 작용한다.

도 1, 패널 a, b 및 c는 크론병을 갖는 가족에서 우성 분리를 갖는 SIGIRR에서의 절두 변이체를 보여준다.

도 2는, 비자극되거나 또는 72시간 동안 2 mg/ml의 LPS로 치료된, 건강한 대조군, SIGIRR LoF 환자 및 SIGIRR LoF를 보유하지 않는 4명의 E0 IBD 환자로부터 생성된 LCL로부터 취한 세포 배양 상청액에서 수행된 Mesoscale Discovery Pro-Inflammatory Cytokine 패널로부터의 결과를 보여준다.

도 3은, 비자극되거나 또는 16시간 동안 2 mg/ml의 항-IgM/항-CD40으로 치료된, 건강한 대조군, SIGIRR LoF 환자 및 SIGIRR LoF를 보유하지 않는 4명의 EO IBD 환자로부터 생성된 LCL로부터 취한 세포 배양 상청액에서 수행된 Mesoscale Discovery Pro-Inflammatory Cytokine 패널로부터의 결과를 보여준다.

본 개시내용의 추가적인 이점은 부분적으로는 하기 설명에 기재될 것이고 부분적으로는 그 설명으로부터 자명해 지거나 또는 본원에 개시된 실시형태의 실행에 의해 학습될 수 있다. 본 개시내용의 이점은 첨부된 청구범위에 특히 제시된 요소 및 조합에 의해 실현되고 획득될 것이다. 상기 일반 설명 및 하기 상세한 설명 둘 다가 오직 예시적이고 설명적이고, 청구된 바대로 실시형태를 제한하지 않는다고 이해되어야 한다.

발명을 실시하기 위한 구체적인 내용

- [0028] 본 개시내용의 양상에 관한 다양한 용어는 본 명세서 및 청구범위에 걸쳐 사용된다. 이러한 용어는, 달리 표시되지 않는 한, 당해 분야에서 이의 보통의 의미가 주어져야 한다. 다른 구체적으로 정의된 용어는 본원에 제공된 정의와 일치하는 방식으로 해석되어야 한다.
- [0029] 달리 명확히 기술되지 않는 한, 본원에 기재된 임의의 방법 또는 양상이 특정한 순서로 이의 단계가 수행될 것을 요하는 것으로 결코 의도되지 않는다. 따라서, 청구범위 또는 명세서에서 특정한 순서로 단계들이 제한되지 않는다고 방법 청구항이 구체적으로 기재하지 않는 경우, 어떤 점에서는 순서가 추론되는 것으로 결코 의도되지 않는다. 이는 단계 또는 조작 흐름의 배열과 관련한 논리 문제, 문법 조직화 또는 구두점으로부터 유래된 명백한 의미, 또는 본 명세서에 기재된 양상의 수 또는 유형을 포함하여, 해석을 위한, 임의의 가능한 비표현 기준을 유지한다.
- [0030] 본원에서 사용될 때, 단수 형태의 표현은, 문맥이 명확히 달리 기술하지 않는 한, 복수의 지시대상을 포함한다.
- [0031] 본원에서 사용될 때, "대상체" 및 "환자"의 용어는 상호 교환 가능하게 사용된다. 대상체는 포유류를 포함한 임의의 동물을 포함할 수 있다. 포유류는, 제한 없이, 농장 동물(예를 들어, 말, 소, 돼지), 반려 동물(예를 들어, 개, 고양이), 실험실 동물(예를 들어, 마우스, 래트, 토끼) 및 비인간 영장류를 포함한다. 일부 실시형태에서, 상기 대상체는 인간이다.
- [0032] 본원에서 사용될 때, "핵산", "핵산 분자", "핵산 서열", "폴리뉴클레오타이드" 또는 "올리고뉴클레오타이드"는 임의의 길이의 뉴클레오타이드의 중합체 형태를 포함할 수 있고, DNA 및/또는 RNA를 포함할 수 있고, 단일-가닥, 이중-가닥 또는 다중 가닥일 수 있다. 핵산의 하나의 가닥은 또한 그의 보체를 지칭한다.
- [0033] 본원에서 사용될 때, "에 상응하는"의 구절 또는 이의 문법 파생어는, 주어진 아미노산 또는 핵산 서열 또는 위치의 넘버링의 맥락에서 사용될 때, 주어진 아미노산 또는 핵산 서열이 기준 서열(예를 들어, (야생형 또는 전장) SIGIRR의 핵산 분자 또는 폴리펩타이드인 본원에서의 기준 서열)과 비교될 때의 기재된 기준 서열의 넘버링을 지칭한다. 다른 말로, 주어진 중합체의 잔기(예를 들어, 아미노산 또는 뉴클레오타이드) 수 또는 잔기(예를 들어, 아미노산 또는 뉴클레오타이드) 위치는 주어진 아미노산 또는 핵산 서열 내의 잔기의 실제 숫자 위치에 의하기보다는 기준 서열과 관련하여 지칭된다. 예를 들어, 주어진 아미노산 서열은 2개의 서열 사이의 잔기 일치를 최적화시키도록 갭을 도입함으로써 기준 서열로 정렬될 수 있다. 이 경우에, 갭이 존재하지만, 주어진 아미노산 또는 핵산 서열에서의 잔기의 넘버링은 이것이 정렬되는 기준 서열과 관련하여 이루어진다.
- [0034] 예를 들어, "서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질"의 구절(및 유사한 구절)은, SIGIRR 단백질의 아미노산 서열이 서열번호 9의 서열로 정렬되는 경우, SIGIRR 단백질이 서열번호 9의 215번 위치에 상응하는 위치에서 절두된다는 것을 의미한다(예를 들어, SIGIRR 단백질의 말단 아미노산은 215번 위치에서의 아미노산임). 아니면, 다른 말로, 이 구절은 서열번호 9의 215번 위치에 상동성인 위치에서 절두를 갖는 SIGIRR 단백질을 지칭한다. 본원에서, 이러한 단백질은 또한 "절두된 SIGIRR 단백질" 또는 "변이체 SIGIRR 단백질" 또는 "p.K186fs*31 변이체"라 칭해진다.
- [0035] 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질은 주어진 SIGIRR 단백질과 서열번호 9의 아미노산 서열 사이에 서열 정렬을 수행함으로써 쉽게 확인될 수 있다. 마찬가지로, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖는 SIGIRR 단백질은 주어진 SIGIRR 단백질과 서열번호 9의 아미노산 서열 사이에 서열 정렬을 수행함으로써 쉽게 확인될 수 있다. 서열번호 9에 따른 215번 위치에 상응하는 위치에서의 절두를 확인하기 위해서 또는 서열번호 9에 따른 186번 위치에 상응하는 위치에서의 세린을 확인하기 위해서 서열 정렬을 수행하기 위해 이용될 수 있는 다양한 컴퓨터 알고리즘이 존재한다. 예를 들어, NCBI BLAST 알고리즘(Altschul et al., 1997, Nucleic Acids Res., 25, 3389-3402) 또는 CLUSTALW 소프트웨어(Sievers et al.,

2014, Methods Mol. Biol., 1079, 105-116)를 사용함으로써 서열 정렬이 수행될 수 있다. 그러나, 서열은 또한 수동으로 정렬될 수 있다.

- [0036] 본 개시내용에 따르면 SIGIRR에서의 특정한 변이가 조기-발병 염증성 장 질환을 발생시킬 위험과 연관된다고 관찰되었다. 일반적으로, 이 단백질의 기능은 특히 18세보다 어린 아동에서 완전히 이해되지 않는다. SIGIRR 유전자 또는 단백질의 변이체가 인간에서 조기-발병 염증성 장 질환과 어떠한 알려진 연관성을 갖지 않는다고 믿어진다. SIGIRR 유전자 또는 단백질의 변이체가 특히 아동에서 조기-발병 염증성 장 질환과 어떠한 알려진 연관성을 갖지 않는다고 더 믿어진다. 이환된 가족 구성원에서 조기-발병 염증성 장 질환의 표현형으로 분리하는 SIGIRR 유전자에서의 희귀 변이체는 본 개시내용에 따라 확인되었다. 예를 들어, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된(예를 들어, 말단 아미노산은 215번 위치에 위치함) SIGIRR 단백질을 생산하는 프레임쉬프트를 생성시키는, 인간 SIGIRR mRNA 또는 cDNA의 557번 위치(예를 들어, 각각 야생형 서열번호 3 및 서열번호 5)에서 아데닌의 결실을 발생시키는 유전자 변형은 이러한 변형을 갖는 인간이 조기-발병 염증성 장 질환을 발생시킬 수 있다는 것을 나타낸다고 관찰된다. 종합하면, 본원에 기재된 유전자 분석은 SIGIRR 유전자 및 특히 SIGIRR 유전자에서의 절두 또는 기능 소실 변이체가 조기-발병 염증성 장 질환을 발생시키는 감수성의 증가와 연관된다는 것을 제안한다. 따라서, 조기-발병 염증성 장 질환과 연관된 SIGIRR 변형을 갖는 인간 대상체는 조기-발병 염증성 장 질환이 저해되고/되거나, 이의 증상이 감소되고/되거나, 증상의 발생이 억제되도록 치료될 수 있다. 따라서, 본 개시내용은 cDNA 및 mRNA를 포함하는 단리된 또는 재조합 SIGIRR 변이체 유전자, 및 단리된 또는 재조합 SIGIRR 변이체 폴리펩타이드를 제공한다. 추가적으로, 본 개시내용은, 위험에 있는 대상체 또는 활성 질환을 갖는 대상체가 치료될 수 있도록, 이러한 대상체에서의 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 확인하거나 계층화하기 위해, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 갖는 것으로 대상체를 진단하기 위해 대상체에서의 이러한 변이체의 확인을 발휘하는 방법을 제공한다.
- [0037] 2개의 야생형 SIGIRR 단백질에 대한 아미노산 서열은 서열번호 7 및 서열번호 8에 기재된다. 서열번호 7을 갖는 야생형 SIGIRR 단백질은 410개의 아미노산의 길이인 한편, 서열번호 8을 갖는 야생형 SIGIRR 단백질은 504개의 아미노산의 길이이다. 서열번호 7 및 서열번호 8 둘 다를 살펴보면, 야생형 단백질의 186번 내지 215번 위치는 하기 아미노산을 기재된 순서로 포함한다: Lys-Pro-Gln-Leu-Glu-Arg-Arg-Arg-Gly-Tyr-Lys-Leu-Phe-Leu-Asp-Asp-Arg-Asp-Leu-Leu-Pro-Arg-Ala-Glu-Pro-Ser-Ala-Asp-Leu-Leu(서열번호 10).
- [0038] 본 개시내용은 염증성 장 질환 또는 조기-발병 염증성 장 질환과 연관된 SIGIRR 변이체 단백질을 암호화하는 핵산 분자를 제공한다. 일부 실시형태에서, 핵산 분자는 절두된 SIGIRR 변이체 단백질을 암호화한다. 예를 들어, 본 개시내용은 인간 SIGIRR 단백질을 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하는 단리된 핵산 분자를 제공하되, 상기 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다.
- [0039] 일부 실시형태에서, 단리된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다.
- [0040] 일부 실시형태에서, 단리된 핵산 분자는 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함하는 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다.
- [0041] 일부 실시형태에서, 핵산 분자는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 아미노산 서열을 갖는 인간 SIGIRR 단백질을 암호화하는 핵산 서열 또는 그 핵산 서열의 보체를 포함하거나 또는 이들로 이루어진다. 본원에서, 퍼센트 서열 동일성을 언급하는 경우, 서열 동일성의 더 높은 백분율이 더 낮은 것에 비해 바람직하다.
- [0042] 일부 실시형태에서, 단리된 핵산 분자는 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어지고, 절두된 SIGIRR 단백질은 서열번호 9의 아미노산 서열, 또는 서열번호 9와 적어도 90%의 서열 동일성을 갖고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 포함한다.
- [0043] 야생형 SIGIRR 게놈 DNA의 핵산 서열은 서열번호 1에 기재된다. 서열번호 1을 포함하는 야생형 SIGIRR 게놈 DNA는 11,739개의 뉴클레오타이드의 길이이다. 서열번호 1을 살펴보면, 야생형 SIGIRR 게놈 DNA의 9962번 위치는 아데닌이다.
- [0044] 본 개시내용은 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자를 제공한다. 일부 실시형태에서, 게놈 DNA 분자는 절두된 SIGIRR 단백질을 암호화한다. 일부 실시형태에서, 게놈 DNA는 서열번호 9에 따른 215번 위치에 상

응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 게놈 DNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 게놈 DNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 게놈 DNA는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 게놈 DNA는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 게놈 DNA는 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어지고, 절두된 SIGIRR 단백질은 서열번호 9의 아미노산 서열, 또는 서열번호 9와 적어도 90%의 서열 동일성을 갖고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 포함한다.

[0045] 일부 실시형태에서, 게놈 DNA는 서열번호 2에 따른 9962번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 이에 반해, 야생형 SIGIRR 게놈 DNA는 서열번호 1에 따른 9962번 위치에 상응하는 위치에서 아데닌을 포함한다. 변이체 SIGIRR 게놈 DNA에서의 변경은 이 아데닌의 결실로 인한 것이고, 이는 1개의 뉴클레오타이드 염기 프레임쉬프트를 생성하여, 서열번호 2의 9962번 위치에 상응하는 위치에서 구아닌을 생성시킨다. 일부 실시형태에서, 게놈 DNA는 서열번호 2와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 게놈 DNA는 서열번호 2에 따른 핵산 서열을 포함하거나 또는 이들로 이루어진다.

[0046] 일부 실시형태에서, 단리된 핵산 분자는 전체보다 적은 게놈 DNA 서열을 포함한다. 일부 실시형태에서, 단리된 핵산 분자는 서열번호 2의 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개, 적어도 약 500개, 적어도 약 600개, 적어도 약 700개, 적어도 약 800개, 적어도 약 900개, 적어도 약 1000개, 적어도 약 2000개, 적어도 약 3000개, 적어도 약 4000개, 적어도 약 5000개, 적어도 약 6000개, 적어도 약 7000개, 적어도 약 8000개, 적어도 약 9000개, 적어도 약 10000개, 적어도 약 11000개, 또는 적어도 약 11500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 서열번호 2의 적어도 약 1000개 내지 적어도 약 2000개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다.

[0047] 일부 실시형태에서, 단리된 핵산 분자는 서열번호 2의 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개, 적어도 약 500개, 적어도 약 600개, 적어도 약 700개, 적어도 약 800개, 적어도 약 900개, 적어도 약 1000개, 적어도 약 1100개, 적어도 약 1200개, 적어도 약 1300개, 적어도 약 1400개, 적어도 약 1500개, 적어도 약 1600개, 적어도 약 1700개, 적어도 약 1800개, 적어도 약 1900개, 적어도 약 2000개, 적어도 약 2100개, 적어도 약 2200개, 적어도 약 2300개, 적어도 약 2400개, 또는 적어도 약 2500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 이러한 인접 뉴클레오타이드는 본원에 기재된 cDNA 분자를 제조하도록 인접 뉴클레오타이드의 다른 핵산 분자와 조합될 수 있다.

[0048] 이러한 단리된 핵산 분자는, 예를 들어 변이체 SIGIRR mRNA 및 단백질을 발현하도록 사용될 수 있거나 또는 외인성 도너 서열로서 사용될 수 있다. 집단 내에 유전자 서열이 다형, 예컨대 SNP로 인해 변할 수 있다고 이해된다. 본원에 제공된 예는 오직 예시적인 서열이고, 다른 서열이 또한 가능하다.

[0049] 일부 실시형태에서, 단리된 핵산 분자는 서열번호 9를 암호화하는 하나 이상의 비필수 분절(nonessential segments)이 상응하는 야생형 SIGIRR 게놈 DNA와 관련하여 결실된 변이체 SIGIRR 미니유전자(minigene)를 포함한다. 일부 실시형태에서, 결실된 비필수 분절(들)은 하나 이상의 인트론 서열을 포함한다. 일부 실시형태에서, SIGIRR 미니유전자는 서열번호 2의 일부와 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 100%의 서열 동일성을 갖고, 여기서 미니유전자는 서열번호 2에 따른 9962번 위치에 상응하는 위치에서 구아닌을 갖는 핵산

서열을 포함한다.

- [0050] 야생형 *SIGIRR* mRNA의 핵산 서열은 서열번호 3에 기재된다. 서열번호 3을 포함하는 야생형 *SIGIRR* mRNA는 1230개의 뉴클레오타이드의 길이이다. 서열번호 3을 살펴보면, 야생형 *SIGIRR* mRNA의 557번 위치는 아데닌이다.
- [0051] 본 개시내용은 또한 변이체 *SIGIRR* 단백질을 암호화하는 mRNA 분자를 제공한다. 일부 실시형태에서, mRNA 분자는 절두된 *SIGIRR* 단백질을 암호화한다. 일부 실시형태에서, mRNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, mRNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, mRNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, mRNA는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, mRNA는 서열번호 9를 갖는 변이체 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, mRNA는 절두된 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어지고, 절두된 *SIGIRR* 단백질은 서열번호 9의 아미노산 서열, 또는 서열번호 9와 적어도 90%의 서열 동일성을 갖고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 포함한다.
- [0052] 일부 실시형태에서, mRNA는 서열번호 4에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 이에 반해, 야생형 *SIGIRR* mRNA는 서열번호 3에 따른 557번 위치에 상응하는 위치에서 아데닌을 포함한다. 변이체 *SIGIRR* mRNA에서의 변경은 이 아데닌의 결실로 인한 것이고, 이는 1개의 뉴클레오타이드 염기 프레임쉬프트를 생성하여, 서열번호 4의 557번 위치에 상응하는 위치에서 구아닌을 생성시킨다. 일부 실시형태에서, mRNA는 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 이에 반해, 야생형 *SIGIRR* mRNA는 각각 서열번호 3에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AAG를 포함한다. 일부 실시형태에서, mRNA는 서열번호 4와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, mRNA는 서열번호 4에 따른 핵산 서열을 포함하거나 또는 이들로 이루어진다.
- [0053] 일부 실시형태에서, 단리된 핵산 분자는 전체 *SIGIRR* mRNA 서열보다 적은 뉴클레오타이드를 포함한다. 일부 실시형태에서, 단리된 핵산 분자는 서열번호 4의 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 12개, 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개, 적어도 약 500개 또는 적어도 약 600개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 서열번호 4의 적어도 약 200개 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자는 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 단리된 핵산 분자는 서열번호 4의 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 mRNA 분자는 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 이러한 mRNA 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함한다. 일부 실시형태에서, 이러한 mRNA 분자는 서열번호 4에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함한다. 일부 실시형태에서, 이러한 mRNA 분자는 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함한다.
- [0054] 야생형 *SIGIRR* cDNA의 핵산 서열은 서열번호 5에 기재된다. 서열번호 5를 포함하는 야생형 *SIGIRR* cDNA는 종결 코돈을 포함하여 1233개의 뉴클레오타이드의 길이이다. 서열번호 5를 살펴보면, 야생형 *SIGIRR* cDNA의 557번 위치는 아데닌이다.

- [0055] 본 개시내용은 또한 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자를 제공한다. 일부 실시형태에서, cDNA 분자는 절두된 SIGIRR 단백질을 암호화한다. 일부 실시형태에서, cDNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, cDNA는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA는 절두된 SIGIRR 단백질을 암호화하는 핵산 서열을 포함하거나 또는 이들로 이루어지고, 절두된 SIGIRR 단백질은 서열번호 9의 아미노산 서열, 또는 서열번호 9와 적어도 90%의 서열 동일성을 갖고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 포함한다.
- [0056] 일부 실시형태에서, cDNA는 서열번호 6에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 이에 반해, 야생형 SIGIRR cDNA는 서열번호 5에 따른 557번 위치에 상응하는 위치에서 아데닌을 포함한다. 변이체 SIGIRR cDNA에서의 변경은 이 아데닌의 결실로 인한 것이고, 이는 1개의 뉴클레오타이드 염기 프레임쉬프트를 생성하여, 서열번호 6의 557번 위치에 상응하는 위치에서 구아닌을 생성시킨다. 일부 실시형태에서, cDNA는 각각 서열번호 6에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CTA 및 AGC를 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 이에 반해, 야생형 SIGIRR cDNA는 각각 서열번호 5에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CTA 및 AAG를 포함한다. 일부 실시형태에서, cDNA는 서열번호 6과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA는 서열번호 6에 따른 핵산 서열을 포함하거나 또는 이들로 이루어진다.
- [0057] 일부 실시형태에서, cDNA 분자는 변이체 SIGIRR cDNA 분자의 전체보다 적은 서열을 포함한다. 일부 실시형태에서, cDNA 분자는 서열번호 6의 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 12개, 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개, 적어도 약 500개 또는 적어도 약 600개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, cDNA 분자는 서열번호 6의 적어도 약 200개 내지 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자는 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, cDNA 분자는 서열번호 6의 적어도 약 50개, 적어도 약 60개, 적어도 약 70개, 적어도 약 80개, 적어도 약 90개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개 또는 적어도 약 500개의 인접 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 cDNA 분자는 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 이러한 cDNA 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함한다. 일부 실시형태에서, 이러한 cDNA 분자는 서열번호 6에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함한다. 일부 실시형태에서, 이러한 cDNA 분자는 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CTA 및 AGC를 포함한다.
- [0058] 본 개시내용은 또한 변이체 SIGIRR 게놈 DNA(예컨대, 서열번호 2), 변이체 SIGIRR 미니유전자, 변이체 SIGIRR mRNA(예컨대, 서열번호 4), 및/또는 변이체 SIGIRR cDNA(예컨대, 서열번호 6)에 혼성화하는 단리된 핵산 분자를 제공한다. 일부 실시형태에서, 이러한 단리된 핵산 분자는 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 11개, 적어도 약 12개, 적어도 약 13개, 적어도 약 14개, 적어도 약 15개, 적어도 약 16개, 적어도 약 17개, 적어도 약 18개, 적어도 약 19개, 적어도 약 20개, 적어도 약 21개, 적어도 약 22개, 적어도 약 23개, 적어도 약 24개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적

어도 약 85개, 적어도 약 90개, 적어도 약 95개, 적어도 약 100개, 적어도 약 200개, 적어도 약 300개, 적어도 약 400개, 적어도 약 500개, 적어도 약 600개, 적어도 약 700개, 적어도 약 800개, 적어도 약 900개, 적어도 약 1000개, 적어도 약 2000개, 적어도 약 3000개, 적어도 약 4000개, 적어도 약 5000개, 적어도 약 6000개, 적어도 약 7000개, 적어도 약 8000개, 적어도 약 9000개, 적어도 약 10000개, 적어도 약 11000개 또는 적어도 약 11500개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 적어도 15개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 적어도 15개의 뉴클레오타이드 내지 적어도 약 35개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 이러한 단리된 핵산 분자는 엄격한 조건하에 변이체 *SIGIRR* 게놈 DNA(예컨대, 서열번호 2), 변이체 *SIGIRR* 미니유전자, 변이체 *SIGIRR* mRNA(예컨대, 서열번호 4), 및/또는 변이체 *SIGIRR* cDNA(예컨대, 서열번호 6)에 혼성화한다. 이러한 핵산 분자는, 본원에 기재된 또는 예시된 바와 같은, 예를 들어 프로브로서, 프라이머로서 또는 변경-특이적 프로브 또는 프라이머로서 사용될 수 있다.

[0059] 일부 실시형태에서, 단리된 핵산 분자는 변이체 *SIGIRR* 게놈 DNA(예컨대, 서열번호 2), 변이체 *SIGIRR* 미니유전자, 변이체 *SIGIRR* mRNA(예컨대, 서열번호 4), 및/또는 변이체 *SIGIRR* cDNA(예컨대, 서열번호 6)와 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 100% 동일한 핵산 분자의 적어도 약 15개의 인접 뉴클레오타이드에 혼성화한다. 일부 실시형태에서, 단리된 핵산 분자는 약 15개 내지 약 100개의 뉴클레오타이드, 또는 약 15개 내지 약 35개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 약 15개 내지 약 100개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단리된 핵산 분자는 약 15개 내지 약 35개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다.

[0060] 일부 실시형태에서, 본원에 개시된 임의의 핵산 분자, 게놈 DNA 분자, cDNA 분자 또는 mRNA 분자는 정제될 수 있고, 예를 들어 적어도 약 90% 순수하다. 일부 실시형태에서, 본원에 개시된 임의의 핵산 분자, 게놈 DNA 분자, cDNA 분자 또는 mRNA 분자는 정제될 수 있고, 예를 들어 적어도 약 95% 순수하다. 일부 실시형태에서, 본원에 개시된 임의의 핵산 분자, 게놈 DNA 분자, cDNA 분자 또는 mRNA 분자는 정제될 수 있고, 예를 들어 적어도 약 99% 순수하다. 정제는 인간 제조 정제 기법에 의해 인간 손에 따른다.

[0061] 본 개시내용은 또한 본원에 개시된 임의의 단리된 핵산 분자, 게놈 DNA 분자, cDNA 분자 또는 mRNA 분자의 단편을 제공한다. 일부 실시형태에서, 단편은 본원에 개시된 임의의 핵산 서열, 또는 그의 임의의 보체의 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 11개, 적어도 약 12개, 적어도 약 13개, 적어도 약 14개, 적어도 약 15개, 적어도 약 16개, 적어도 약 17개, 적어도 약 18개, 적어도 약 19개, 적어도 약 20개, 적어도 약 21개, 적어도 약 22개, 적어도 약 23개, 적어도 약 24개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적어도 약 85개, 적어도 약 90개, 적어도 약 95개 또는 적어도 약 100개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 단편이 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 단편은 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 11개, 적어도 약 12개, 적어도 약 13개, 적어도 약 14개, 적어도 약 15개, 적어도 약 16개, 적어도 약 17개, 적어도 약 18개, 적어도 약 19개, 적어도 약 20개, 적어도 약 21개, 적어도 약 22개, 적어도 약 23개, 적어도 약 24개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개 또는 적어도 약 50개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 이와 관련하여, 더 긴 단편이 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 단편은 적어도 약 20개, 적어도 약 25개, 적어도 약 30개 또는 적어도 약 35개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단편은 적어도 약 20개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단편은 적어도 약 25개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단편은 적어도 약 30개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 단편은 적어도 약 35개의 인접 잔기를 포함하거나 또는 이들로 이루어진다. 단편이 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하거나 또는 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치를 암호화하는 핵산 분자의 일부를 포함하거나 또는 이들로 이루어지는 것으로 고안된다. 이러한 단편은, 예를 들어 본원에 기재된 또는 예시된 바와 같은, 프로브로서, 프라이머로서 또는 대립유전자-특이적 프라이머로서 사용될 수 있다.

[0062] 본 개시내용은 또한 프로브 및 프라이머를 제공한다. 본 개시내용의 프로브 또는 프라이머는 본원에 개시된 임의의 핵산 분자 또는 그의 보체에 특이적으로 혼성화하는 핵산 서열을 갖는다. 일부 실시형태에서, 프로브 또는 프라이머는 엄격한 조건하에 본원에 개시된 임의의 핵산 분자에 특이적으로 혼성화한다. 본 개시내용은 또한 보

통의 조건(moderate conditions)하에 본원에 개시된 임의의 핵산 분자 또는 그의 보체에 혼성화하는 핵산 서열을 갖는 핵산 분자를 제공한다. 본 개시내용에 따른 프로브 또는 프라이머는 바람직하게는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 핵산 코돈 또는 그의 보체를 포함한다. 이와 같이, 바람직한 실시형태에서, 본 개시내용은 본원 상기에서 정의되어 있고 하기에 더 자세히 정의된 변경-특이적 프라이머를 제공한다.

[0063] 본 개시내용에 따른 프로브는 (예를 들어, 서열번호 9에 따른) 변이체 SIGIRR 단백질을 암호화하는 변이체 SIGIRR 핵산 분자(예를 들어, 게놈 DNA, mRNA, 및/또는 cDNA)를 검출하도록 사용될 수 있다. 게다가, 본 개시내용에 따른 프라이머는 변이체 SIGIRR 단백질 또는 그의 단편을 암호화하는 핵산 분자를 증폭시키도록 사용될 수 있다. 본 개시내용은 또한 상기에 기재된 프라이머 중 하나를 포함하는 프라이머의 쌍을 제공한다. 절두를 발생시키는 프레임쉬프트 변이체를 함유하는 SIGIRR 단편의 게놈 중합효소 연쇄 반응(PCR) 증폭에 대해, 적합한 프라이머 서열은 정방향 프라이머(5'→3'): TCAGTGGCTCTGAACTGCAC(서열번호 12) 및 역방향 프라이머(5'→3'): GGTCCTGTTGAGCAGAGGAG(서열번호 13)를 포함하지만, 이들로 제한되지는 않는다.

[0064] 본원에 개시된 핵산 분자는 자연 발생적 SIGIRR 게놈 DNA, cDNA 또는 mRNA 전사체의 핵산 서열을 포함할 수 있거나, 또는 비자연 발생적 서열을 포함할 수 있다. 일부 실시형태에서, 자연 발생적 서열은 동의 돌연변이 또는 암호화된 SIGIRR 폴리펩타이드에 영향을 갖지 않는 돌연변이로 인해 비자연 발생적 서열과 상이할 수 있다. 예를 들어, 서열은 동의 돌연변이 또는 암호화된 SIGIRR 폴리펩타이드에 영향을 갖지 않는 돌연변이를 제외하고 동일할 수 있다. 동의 돌연변이 또는 치환은 제조된 아미노산 서열이 변형되지 않도록 단백질을 암호화하는 유전자의 엑손에서 또 다른 뉴클레오타이드에 대한 하나의 뉴클레오타이드의 치환이다. 이는 하나 초과 3개 염기 쌍 코돈에 의해 몇몇 아미노산이 암호화되면서 유전 암호의 축퇴성 때문에 가능하다. 동의 치환은, 예를 들어 코돈 최적화의 과정에서 사용된다. 본원에 개시된 핵산 분자는 코돈 최적화될 수 있다.

[0065] 개시된 핵산 분자와 상호작용할 수 있는 기능적 폴리뉴클레오타이드가 본원에 또한 제공된다. 기능적 폴리뉴클레오타이드는 특이한 기능, 예컨대 표적 분자의 결합 또는 특이적 반응의 촉매화(catalyzing)를 갖는 핵산 분자이다. 기능적 폴리뉴클레오타이드의 예는 안티센스 분자, 압타머, 리보자임, 트리플렉스 형성 분자 및 외부 가이드 서열을 포함하지만, 이들로 제한되지는 않는다. 기능적 폴리뉴클레오타이드는 표적 분자가 보유하는 특이적 활성의 이펙터, 저해제, 조절제 및 자극제로서 작용할 수 있거나, 또는 기능적 폴리뉴클레오타이드는 임의의 다른 분자와 독립적인 신생 활성을 보유할 수 있다.

[0066] 안티센스 분자는 정준 또는 비정준 염기 짝짓기를 통해 표적 핵산 분자와 상호작용하도록 설계된다. 안티센스 분자 및 표적 분자의 상호작용은, 예를 들어 RNase-H 매개된 RNA-DNA 하이브리드 분해를 통해 표적 분자의 파괴를 촉진하도록 설계된다. 대안적으로, 안티센스 분자는 표적 분자에서 보통 발생하는 처리 기능, 예컨대 전사 또는 복제를 차단하도록 설계된다. 안티센스 분자는 표적 분자의 서열에 기초하여 설계될 수 있다. 표적 분자의 가장 접근 가능한 영역을 확인하는 것에 의해 안티센스 효율을 최적화하기 위한 많은 방법이 존재한다. 예시적인 방법은 시험관내 선택 실험 및 DMS 및 DEPC를 사용한 DNA 변형 연구를 포함하지만, 이들로 제한되지는 않는다. 안티센스 분자는 일반적으로 약 10^{-6} 이하, 약 10^{-8} 이하, 약 10^{-10} 이하 또는 약 10^{-12} 이하의 해리 상수(k_d)로 표적 분자에 결합한다. 안티센스 분자의 설계 및 사용을 돕는 방법 및 기법의 대표적인 샘플은 미국 특허의 하기 비제한적인 목록에서 발견될 수 있다: 제5,135,917호; 제5,294,533호; 제5,627,158호; 제5,641,754호; 제5,691,317호; 제5,780,607호; 제5,786,138호; 제5,849,903호; 제5,856,103호; 제5,919,772호; 제5,955,590호; 제5,990,088호; 제5,994,320호; 제5,998,602호; 제6,005,095호; 제6,007,995호; 제6,013,522호; 제6,017,898호; 제6,018,042호; 제6,025,198호; 제6,033,910호; 제6,040,296호; 제6,046,004호; 제6,046,319호; 및 제6,057,437호. 안티센스 분자의 예는 안티센스 RNA, 소간섭 RNA(small interfering RNA: siRNA) 및 짧은 헤어핀 RNA(short hairpin RNA: shRNA)를 포함하지만, 이들로 제한되지는 않는다.

[0067] 본원에 개시된 단리된 핵산 분자는 RNA, DNA, 또는 RNA 및 DNA 둘 다를 포함할 수 있다. 단리된 핵산 분자는 또한, 예컨대 벡터에서의 이중성 핵산 서열 또는 이중성 표지에 연결되거나 융합될 수 있다. 예를 들어, 본원에 개시된 단리된 핵산 분자는 단리된 핵산 분자를 포함하는 벡터 또는 외인성 도너 서열 및 이중성 핵산 서열에 있을 수 있다. 단리된 핵산 분자는 또한 이중성 표지, 예컨대 형광 표지에 연결되거나 또는 융합될 수 있다. 표지의 다른 예는 본원에서 다른 곳에 개시되어 있다.

[0068] 표지는 직접적으로 검출 가능(예를 들어, 형광단) 또는 간접적으로 검출 가능(예를 들어, 합텐, 효소 또는 형광단 켄처)할 수 있다. 이러한 표지는 분광학적, 광화학적, 생화학적, 면역화학적 또는 화학적 수단에 의해 검출 가능할 수 있다. 이러한 표지는, 예를 들어 방사선 계수 장치로 측정될 수 있는 방사선표지; 가시적으로 관찰될

수 있거나 또는 분광광도계로 측정될 수 있는 안료, 염료 또는 다른 발색체; 스핀 표지 분석기로 측정될 수 있는 스핀 표지; 및 출력 신호가 적합한 분자 부가물의 여기에 의해 생성되고, 염료에 의해 흡수된 광에 의한 여기에 의해 가시화될 수 있거나 또는 표준 형광측정계 또는 영상화 시스템으로 측정될 수 있는 형광 표지(예를 들어, 형판단)를 포함한다. 표지는 또한, 예를 들어 출력 신호가 신호 화합물의 화학 변형에 의해 생성된 화학 발광 물질; 금속 함유 물질; 또는 효소-의존적인 신호의 2차 생성, 예컨대 무색 기질로부터의 유색 생성물의 형성이 발생하는 효소일 수 있다. 용어 "표지"는 접합된 분자가, 후속하여 기질과 함께 첨가될 때, 검출 가능한 신호를 생성하도록 사용되도록 접합된 분자에 선택적으로 결합할 수 있는 "태그" 또는 합텐을 또한 지칭할 수 있다. 예를 들어, 태그로서 바이오틴을 사용하고 이어서 태그에 결합하는 겨자무과산화효소(horseradish peroxidase: HRP)의 아비딘 또는 스트렙타비딘 접합체를 사용하고, 이어서 HRP의 존재를 검출하는 열량측정 기질(예를 들어, 테트라메틸벤지딘(tetramethylbenzidine: TMB)) 또는 형광원성 기질을 사용할 수 있다. 정제를 용이하게 하도록 태그로서 사용될 수 있는 예시적인 표지는 myc, HA, FLAG 또는 3XFLAG, 6XHis 또는 폴리히스티딘, 글루타티온-S-전환효소(GST), 말토스 결합 단백질, 에피토프 태그 또는 면역글로불린의 Fc 부분을 포함하지만, 이들로 제한되지는 않는다. 많은 표지는 공지되어 있고, 예를 들어 입자, 형판단, 합텐, 효소 및 이의 열량 측정, 형광원성 및 화학발광 기질 및 다른 표지를 포함한다.

[0069] 개시된 핵산 분자는, 예를 들어 뉴클레오타이드 또는 비천연(non-natural) 또는 변형된 뉴클레오타이드, 예컨대 뉴클레오타이드 유사체 또는 뉴클레오타이드 치환체를 포함할 수 있다. 이러한 뉴클레오타이드는 변형된 염기, 당 또는 포스페이트 기를 함유하거나 또는 그 구조에 비천연 모이어티가 혼입된 뉴클레오타이드를 포함한다. 비천연 뉴클레오타이드의 예는 디데옥시뉴클레오타이드, 바이오틴화, 아미노화, 탈아미노화, 알킬화, 벤질화 및 형판단 표지된 뉴클레오타이드를 포함하지만, 이들로 제한되지는 않는다.

[0070] 본원에 개시된 핵산 분자는 또한 하나 이상의 뉴클레오타이드 유사체 또는 치환체를 포함할 수 있다. 뉴클레오타이드 유사체는 염기, 당 또는 포스페이트 모이어티에 대한 변형을 함유하는 뉴클레오타이드이다. 염기 모이어티에 대한 변형은 A, C, G 및 T/U의 천연 및 합성 변형, 및 상이한 퓨린 또는 피리미딘 염기, 예를 들어 슈도우리딘, 우라실-5-일, 하이포잔틴-9-일(I) 및 2-아미노아데닌-9-일 등을 포함하지만, 이들로 제한되지는 않는다. 변형된 염기는 5-메틸사이토신(5-me-C), 5-하이드록시메틸 사이토신, 잔틴, 하이포잔틴, 2-아미노아데닌, 아데닌 및 구아닌의 6-메틸 및 다른 알킬 유도체, 아데닌 및 구아닌의 2-프로필 및 다른 알킬 유도체, 2-티오우라실, 2-티오티민 및 2-티오사이토신, 5-할로우라실 및 사이토신, 5-프로피닐 우라실 및 사이토신, 6-아조 우라실, 사이토신 및 티민, 5-우라실(슈도우라실), 4-티오우라실, 8-할로, 8-아미노, 8-티올, 8-티오알킬, 8-하이드록실 및 다른 8-치환된 아데닌 및 구아닌, 5-할로, 특히 5-브로모, 5-트리플루오로메틸 및 다른 5-치환된 우라실 및 사이토신, 7-메틸구아닌 및 7-메틸아데닌, 8-아자구아닌 및 8-아자아데닌, 7-데아자구아닌 및 7-데아자아데닌 및 3-데아자구아닌 및 3-데아자아데닌을 포함하지만, 이들로 제한되지는 않는다. 소정의 뉴클레오타이드 유사체 예를 들어, 5-치환된 피리미딘, 6-아자피리미딘, 및 2-아미노프로필아데닌, 5-프로피닐우라실, 5-프로피닐사이토신 및 5-메틸사이토신을 비제한적으로 포함하는 N-2, N-6 및 O-6 치환된 퓨린 등은 듀플렉스 형성의 안정성을 증가시킬 수 있다. 대개, 염기 변형은 증가된 듀플렉스 안정성과 같은 고유한 특성을 달성하도록, 예를 들어 당 변형, 예컨대 2'-O-메톡시에틸과 조합될 수 있다.

[0071] 뉴클레오타이드 유사체는 또한 당 모이어티의 변형을 포함할 수 있다. 당 모이어티에 대한 변형은 리보스 및 데옥시 리보스의 천연 변형, 및 합성 변형을 포함하지만, 이들로 제한되지는 않는다. 당 변형은 2' 위치에서의 하기 변형을 포함하지만, 이들로 제한되지는 않는다: OH; F; O-, S- 또는 N-알킬; O-, S- 또는 N-알케닐; O-, S- 또는 N-알킬닐; 또는 O-알킬-O-알킬, 상기 알킬, 알케닐 및 알킬닐은 치환된 또는 비치환된 C₁₋₁₀알킬 또는 C₂₋₁₀알케닐 및 C₂₋₁₀알킬닐일 수 있다. 예시적인 2' 당 변형은 또한 -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-ONH₂, 및 -O(CH₂)_nON[(CH₂)_nCH₃]₂를 포함하지만, 이들로 제한되지는 않으며, 상기 n 및 m은 1 내지 약 10이다.

[0072] 2' 위치에서 다른 변형은 C₁₋₁₀알킬, 치환된 저급 알킬, 알크아릴, 아르알킬, 0-알크아릴 또는 0-아르알킬, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, 헤테로사이클로알킬, 헤테로사이클로알크아릴, 아미노알킬아미노, 폴리알킬아미노, 치환된 실릴, RNA 절단 기, 리포터 기, 삽입제(intercalator), 올리고뉴클레오타이드의 약동학적 특성을 개선하기 위한 기, 또는 올리고뉴클레오타이드의 약력학적 특성을 개선하기 위한 기, 및 유사한 특성을 갖는 다른 치환기를 포함하지만, 이들로 제한되지는 않는다. 유사한 변형은 또한 당 위의 다른 위치에서, 특히 3' 말단 뉴클레오타이드 위의 또는 2'-5' 연결된 올리고뉴클레오타이드에서의 당의 3' 위치 및 5' 말단 뉴클레오타이드의 5' 위치에서 이루어질 수 있다. 변형된 당은 또한 CH₂ 및 S와 같은 브

릿징(bridging) 고리 산소에서 변형을 함유하는 것을 포함할 수 있다. 뉴클레오타이드 당 유사체는 또한 펜토피라노실 당 대신에 당 모방체, 예컨대 사이클로부틸 모이어티를 가질 수 있다.

[0073] 뉴클레오타이드 유사체는 또한 포스페이트 모이어티에서 변형될 수 있다. 변형된 포스페이트 모이어티는 2개의 뉴클레오타이드 사이의 연결이 포스포로티오에이트, 키랄 포스포로티오에이트, 포스포로디티오에이트, 포스포트리에스테르, 아미노알킬포스포트리에스테르, 3'-알킬렌 포스포네이트 및 키랄 포스포네이트를 포함하는 메틸 및 다른 알킬 포스포네이트, 포스피네이트, 3'-아미노 포스포르아미데이트 및 아미노알킬포스포르아미데이트를 포함하는 포스포르아미데이트, 티오노포스포르아미데이트, 티오노알킬포스포네이트, 티오노알킬포스포트리에스테르 및 보라노포스페이트를 함유하도록 변형될 수 있는 것을 포함하지만, 이들로 제한되지는 않는다. 2개의 뉴클레오타이드 사이의 이 포스페이트 또는 변형된 포스페이트 연결은 3'-5' 연결 또는 2'-5' 연결을 통해서일 수 있고, 상기 연결은 3'-5'에서 5'-3'로 또는 2'-5'에서 5'-2'로와 같은 반대 극성을 함유할 수 있다. 다양한 염, 혼합 염 및 유리 산 형태가 또한 포함된다.

[0074] 뉴클레오타이드 치환체는 뉴클레오타이드와 유사한 기능 특성을 갖지만, 포스페이트 모이어티, 예컨대 펩타이드 핵산(peptide nucleic acid: PNA)을 함유하지 않는 분자를 포함한다. 뉴클레오타이드 치환체는 왓슨-클릭(Watson-Crick) 또는 후그스틴(Hoogsteen) 방식으로 핵산을 인식하지만 포스페이트 모이어티가 아닌 모이어티를 통해 함께 연결되는 분자를 포함한다. 뉴클레오타이드 치환체는 적절한 표적 핵산과 상호작용할 때 이중 나선 유형 구조에 부합할 수 있다.

[0075] 뉴클레오타이드 치환체는 또한 포스페이트 모이어티 또는 당 모이어티가 대체된 뉴클레오타이드 또는 뉴클레오타이드 유사체를 포함한다. 일부 실시형태에서, 뉴클레오타이드 치환체는 표준 인 원자를 함유하지 않을 수 있다. 포스페이트에 대한 치환체는, 예를 들어 짧은 사슬 알킬 또는 사이클로알킬 뉴클레오사이드간 연결, 혼합된 이중원자 및 알킬 또는 사이클로알킬 뉴클레오사이드간 연결, 또는 하나 이상의 짧은 사슬 이중원자성 또는 복소환식 뉴클레오사이드간 연결일 수 있다. 이들은 (부분적으로 뉴클레오사이드의 당 부분으로부터 형성된) 모르폴리노 연결을 갖는 것; 실록산 골격; 설퍼아이드, 설퍼사이드 및 설퍼 골격; 포름아세틸 및 티오포름아세틸 골격; 메틸렌 포름아세틸 및 티오포름아세틸 골격; 알켄 함유 골격; 설퍼메이트 골격; 메틸렌이미노 및 메틸렌 하이드라지노 골격; 설퍼네이트 및 설퍼아미드 골격; 아미드 골격; 및 혼합된 N, O, S 및 CH₂ 성분 부분을 갖는 기타를 포함한다.

[0076] 뉴클레오타이드 치환체에서 뉴클레오타이드의 당 및 포스페이트 모이어티 둘 다가, 예를 들어 아미드 유형 연결(아미노에틸글리신)(PNA)에 의해 대체될 수 있다고 또한 이해된다.

[0077] 예를 들어, 세포 흡수를 증대시키도록 뉴클레오타이드 또는 뉴클레오타이드 유사체에 분자(접합체)의 다른 유형을 또한 연결시킬 수 있다. 접합체는 뉴클레오타이드 또는 뉴클레오타이드 유사체에 화학적으로 연결될 수 있다. 이러한 접합체는, 예를 들어 지질 모이어티, 예컨대 콜레스테롤 모이어티, 콜산, 티오에테르, 예컨대 핵실-S-트리틸티올, 티오펜콜레스테롤, 지방족 사슬, 예컨대 도데칸다올 또는 운데실 잔기, 인지질, 예컨대 디-핵사데실-rac-글리세롤 또는 트리에틸암모늄 1,2-디-O-핵사데실-rac-글리세로-3-인-포스포네이트, 폴리아민 또는 폴리에틸렌 글리콜 사슬, 아다만탄 아세트산, 팔미틸 모이어티, 또는 옥타데실아민 또는 핵실아미노-카보닐-옥시콜레스테롤 모이어티를 포함한다.

[0078] 본 개시내용은 또한 본원에 개시된 핵산 분자의 임의의 하나 이상을 포함하는 벡터를 제공한다. 일부 실시형태에서, 벡터는 본원에 개시된 핵산 분자의 임의의 하나 이상 및 이중성 핵산을 포함한다. 벡터는 핵산 분자를 운반할 수 있는 바이러스성 또는 비바이러스성 벡터일 수 있다. 일부 실시형태에서, 벡터는 플라스미드 또는 코스미드(예를 들어, 추가 DNA 분절이 결합될 수 있는 원형 이중-가닥 DNA)이다. 일부 실시형태에서, 벡터는 추가 DNA 분절이 바이러스 게놈으로 결합될 수 있는 바이러스성 벡터이다. 일부 실시형태에서, 벡터는 이것이 도입되는 숙주 세포에서 자율적으로 복제할 수 있다(예를 들어, 박테리아 복제 기원을 갖는 박테리아 벡터 및 에피솜 포유류 벡터). 일부 실시형태에서, 벡터(예를 들어, 비-에피솜 포유류 벡터)는 숙주 세포로의 도입 시 숙주 세포의 게놈으로 통합될 수 있고, 이로써 숙주 게놈과 함께 복제된다. 더욱이, 특정 벡터는 이들이 작동 가능하게 연결된 유전자의 발현을 지시할 수 있다. 이러한 벡터는 본원에서 "재조합 발현 벡터" 또는 "발현 벡터"라 칭해진다. 이러한 벡터는 또한 표적화 벡터(즉, 외인성 도너 서열)일 수 있다.

[0079] 일부 실시형태에서, 본원에 개시된 다양한 유전자 변이체에 의해 암호화된 단백질은 개시된 유전자 변이체를 암호화하는 핵산 분자를 발현 벡터로 삽입하는 것에 의해 발현되어서, 유전자는 발현 제어 서열, 예컨대 전사 및 번역 제어 서열에 작동 가능하게 연결된다. 발현 벡터는 플라스미드, 코스미드, 레트로바이러스, 아데노바이러스, 아데노-연관된 바이러스(adeno-associated virus: AAV), 식물 바이러스, 예컨대 꽃양배추 모자이크 바이러스

스 및 담배 모자이크 바이러스, 효모 인공 염색체(yeast artificial chromosome: YAC), 엡스타인-바(Epstein-Barr: EBV)-유래 에피솜 및 당해 분야에 공지된 다른 발현 벡터를 포함하지만, 이들로 제한되지는 않는다. 일부 실시형태에서, 개시된 유전자 변이체를 포함하는 핵산 분자는 벡터에 결합될 수 있어서, 벡터 내의 전사 및 번역 제어 서열은 유전자 변이체의 전사 및 번역을 조절하는 의도된 기능을 제공한다. 발현 벡터 및 발현 제어 서열은 사용되는 발현 숙주 세포와 맞도록 선택된다. 개시된 유전자 변이체를 포함하는 핵산 서열은 별개의 벡터 내로 또는 변이체 유전 정보와 동일한 발현 벡터 내로 삽입될 수 있다. 개시된 유전자 변이체를 포함하는 핵산 서열은 표준 방법(예를 들어, 개시된 유전자 변이체를 포함하는 핵산 및 벡터에서의 상보적 제한 부위에서의 결합, 또는 제한 부위가 존재하지 않는 경우에는 평활 말단 결합(blunt end ligation))에 의해 발현 벡터로 삽입될 수 있다.

[0080] 개시된 유전자 변이체를 포함하는 핵산 서열 이외에, 재조합 발현 벡터는 숙주 세포에서 유전자 변이체의 발현을 제어하는 조절 서열을 보유할 수 있다. 조절 서열의 선택을 포함하여 발현 벡터의 설계는 형질전환되는 숙주 세포의 선택, 원하는 단백질의 발현의 수준 등과 같은 인자에 따라 달라질 수 있다. 포유류 숙주 세포 발현을 위한 원하는 조절 서열은, 예를 들어 포유류 세포에서 단백질 발현의 높은 수준을 지시하는 바이러스 요소, 예컨대 레트로바이러스 LTR, 사이토메갈로바이러스(cytomegalovirus: CMV)(예컨대, CMV 프로모터/인핸서), 유인원 바이러스 40(Simian Virus 40: SV40)(예컨대, SV40 프로모터/인핸서), 아데노바이러스(예를 들어, 아데노바이러스 주요 후기 프로모터(adenovirus major late promoter: AdMLP))로부터 유래된 프로모터 및/또는 인핸서, 폴리오마 및 강한 포유류 프로모터, 예컨대 천연 면역글로불린 및 액틴 프로모터를 포함할 수 있다. 박테리아 세포 또는 진균 세포(예를 들어, 효모 세포)에서 폴리펩타이드를 발현시키는 방법은 또한 널리 공지되어 있다.

[0081] 프로모터는, 예를 들어 구성적 활성 프로모터, 조건적 프로모터, 유도성 프로모터, 시간적으로 제한된 프로모터(예를 들어, 발생 조절된 프로모터), 또는 공간적으로 제한된 프로모터(예를 들어, 세포-특이적 또는 조직-특이적 프로모터)일 수 있다. 프로모터의 예는, 예를 들어 WO 제2013/176772호에서 발견될 수 있다.

[0082] 유도성 프로모터의 예는, 예를 들어 화학적으로 조절된 프로모터 및 물리적으로 조절된 프로모터를 포함한다. 화학적으로 조절된 프로모터는, 예를 들어 알코올-조절된 프로모터(예를 들어, 알코올 탈수소효소(alcohol dehydrogenase)(alcA) 유전자 프로모터), 테트라사이클린-조절된 프로모터(예를 들어, 테트라사이클린 반응성 프로모터, 테트라사이클린 조작자 서열(tetO), tet-On 프로모터 또는 tet-Off 프로모터), 스테로이드 조절된 프로모터(예를 들어, 랫트 글루코코르티코이드 수용체, 에스트로겐 수용체의 프로모터 또는 에크디손 수용체의 프로모터) 또는 금속-조절된 프로모터(예를 들어, 금속단백질 프로모터)를 포함한다. 물리적으로 조절된 프로모터는, 예를 들어 온도-조절된 프로모터(예를 들어, 열 충격 프로모터) 및 광-조절된 프로모터(예를 들어, 광-유도성 프로모터 또는 광-억제성 프로모터)를 포함한다.

[0083] 조직-특이적 프로모터는, 예를 들어 뉴런-특이적 프로모터, 교세포-특이적 프로모터, 근육 세포-특이적 프로모터, 심장 세포-특이적 프로모터, 신장 세포-특이적 프로모터, 골 세포-특이적 프로모터, 내피 세포-특이적 프로모터, 또는 면역 세포-특이적 프로모터(예를 들어, B 세포 프로모터 또는 T 세포 프로모터)일 수 있다.

[0084] 발생 조절된 프로모터는, 예를 들어 오직 배아 발생 단계 동안에만 또는 오직 성체 세포에서만 활성인 프로모터를 포함한다.

[0085] 개시된 유전자 변이체를 포함하는 핵산 서열 및 조절 서열 이외에, 재조합 발현 벡터는 추가 서열, 예컨대 숙주 세포에서 벡터의 복제를 조절하는 서열(예를 들어, 복제 기원) 및 선택 가능한 마커 유전자를 보유할 수 있다. 선택 가능한 마커 유전자는 벡터가 도입되는 숙주 세포의 선택을 용이하게 할 수 있다(예를 들어, 미국 특허 제 4,399,216호; 제4,634,665호; 및 제5,179,017호 참조). 예를 들어, 선택 가능한 마커 유전자는 벡터가 도입된 숙주 세포에서 약물, 예컨대 G418, 하이그로마이신 또는 메토틱세이트에 내성을 부여할 수 있다. 예시적인 선택 가능한 마커 유전자는 (dhfr-숙주 세포에서 메토틱세이트 선택/증폭과 사용하기 위한) 디히드로엽산 환원 효소(dihydrofolate reductase: DHFR) 유전자, (G418 선택을 위한) neo 유전자, 및 글루타메이트 합성효소(glutamate synthetase: GS) 유전자를 포함하지만, 이들로 제한되지는 않는다.

[0086] 추가 벡터는, 예를 들어 2016년 7월 28일에 제출된 미국 가출원 제62/367,973호에 기재되어 있으며, 상기 문헌은 본원에 그 전문이 참고로 인용된다.

[0087] 본 개시내용은 또한 본원에 개시된 단리된 핵산 분자, 게놈 DNA 분자, cDNA 분자 또는 mRNA 분자 중 임의의 하나 이상을 포함하는 조성물을 제공한다. 일부 실시형태에서, 조성물은 약제학적 조성물이다.

[0088] 본 개시내용은 또한 변이체 SIGIRR 폴리펩타이드를 제공한다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드

는 절두된다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함한다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 복수의 아미노산을 포함한다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는다. 일부 실시형태에서, 변이체 SIGIRR 폴리펩타이드는 서열번호 9에 따른 아미노산 서열을 포함하거나 또는 이들로 이루어진다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 서열번호 9에 따른 아미노산 서열, 또는 서열번호 9와 적어도 90%의 서열 동일성을 갖고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 포함하거나 또는 이들로 이루어진다.

[0089] 본 개시내용은 또한 본원에 개시된 임의의 폴리펩타이드의 단편을 제공한다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드(예컨대, 서열번호 9의 아미노산 서열을 갖는 폴리펩타이드)의 적어도 약 10개, 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적어도 약 85개, 적어도 약 90개, 적어도 약 95개, 적어도 약 100개, 적어도 약 150개 또는 적어도 약 200개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 것에 비해 바람직하다. 일부 실시형태에서, 단편은 암호화된 폴리펩타이드의 적어도 약 10개, 적어도 약 15개, 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적어도 약 85개, 적어도 약 90개, 적어도 약 95개 또는 적어도 약 100개의 인접 아미노산 잔기를 포함한다. 이와 관련하여, 더 긴 단편이 더 짧은 것에 비해 바람직하다.

[0090] 본 개시내용은 또한 변이체 SIGIRR 폴리펩타이드를 포함하는 단리된 폴리펩타이드를 포함하는 이합체를 제공하고, 여기서 폴리펩타이드는 본원에 개시된 임의의 폴리펩타이드로부터 선택된다.

[0091] 일부 실시형태에서, 본원에 개시된 단리된 폴리펩타이드는 이중성 폴리펩타이드 또는 이중성 분자 또는 표지에 연결되거나 또는 융합되고, 이의 많은 예는 본원 어딘가에 개시되어 있다. 예를 들어, 단백질은 이중성 폴리펩타이드에 융합될 수 있어서 증가된 또는 감소된 안정성을 제공할 수 있다. 융합된 도메인 또는 이중성 폴리펩타이드는 폴리펩타이드 내에서 N-말단에, C-말단에 또는 내부에 위치할 수 있다. 융합 파트너는, 예를 들어 T 헬퍼 에피토프를 제공하는 것을 도울 수 있거나(면역학적 융합 파트너), 또는 천연 재조합 폴리펩타이드보다 더 높은 수율로 단백질을 발현하는 것을 도울 수 있다(발현 인핸서). 소정의 융합 파트너는 면역학적 융합 파트너 및 발현 증대 융합 파트너 둘 다이다. 다른 융합 파트너는 폴리펩타이드의 용해도를 증가시키도록 또는 원하는 세포내 구획에 폴리펩타이드를 표적화하는 것을 용이하게 하도록 선택될 수 있다. 몇몇 융합 파트너는 폴리펩타이드의 정제를 용이하게 하는 친화도 태그를 포함한다.

[0092] 일부 실시형태에서, 융합 단백질은 이중성 분자에 직접 융합되거나, 링커, 예컨대 펩타이드 링커를 통해 이중성 분자에 연결된다. 적합한 펩타이드 링커 서열은, 예를 들어 하기 인자에 기초하여 선택될 수 있다: 1) 가요성 연장된 입체형태(conformation)를 채택하는 능력; 2) 제1 및 제2 폴리펩타이드에서 기능적 에피토프와 상호작용하는 2차 구조의 채택에 대한 저항; 및 3) 폴리펩타이드 기능적 에피토프와 반응하는 소수성 또는 하전된 잔기의 결여. 예를 들어, 펩타이드 링커 서열은 Gly, Asn 및 Ser 잔기를 함유할 수 있다. 거의 중성인 다른 아미노산, 예컨대 Thr 및 Ala는 또한 링커 서열에서 사용될 수 있다. 링커로서 유용하게 사용될 수 있는 아미노산 서열은, 예를 들어 문헌[Maratea *et al.*, *Gene*, 1985, 40, 39-46; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA*, 1986, 83, 8258-8262]; 및 미국 특허 제4,935,233호 및 제4,751,180호에 개시된 것을 포함한다. 링커 서열은 일반적으로, 예를 들어 1개 내지 약 50개의 아미노산의 길이일 수 있다. 제1 및 제2 폴리펩타이드가 기능적 도메인을 분리시키고 입체 장애를 방지하도록 사용될 수 있는 비필수 N 말단 아미노산 영역을 가질 때 링커 서열은 일반적으로 필요하지 않다.

[0093] 일부 실시형태에서, 폴리펩타이드는 세포-침투 도메인에 작동 가능하게 연결된다. 예를 들어, 세포-침투 도메인은 HIV-1 TAT 단백질, 인간 B형 간염 바이러스로부터의 TLM 세포-침투 모티프, MPG, Pep-1, VP22, 단순 포진

바이러스로부터의 세포-침투 펩타이드 또는 폴리아르기닌 펩타이드 서열로부터 유래될 수 있다. 예를 들어, WO 제2014/089290호를 참조한다. 세포-침투 도메인은 N 말단에, C 말단에 또는 단백질 내의 임의의 곳에 위치할 수 있다.

[0094] 일부 실시형태에서, 폴리펩타이드는 추적 또는 정제의 용이를 위해 이중성 폴리펩타이드, 예컨대 형광 단백질, 정제 태그 또는 에피토프 태그에 작동 가능하게 연결된다. 형광 단백질의 예는 녹색 형광 단백질(예를 들어, GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), 황색 형광 단백질(예를 들어, YFP, eYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), 청색 형광 단백질(예를 들어, eBFP, eBFP2, Azurite, mKalamal, GFPuv, Sapphire, T-sapphire), 청록색(cyan) 형광 단백질(예를 들어, eCFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), 적색 형광 단백질(예를 들어, mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), 오렌지색 형광 단백질(예를 들어, mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato), 및 임의의 다른 적합한 형광 단백질을 포함하지만, 이들로 제한되지는 않는다. 태그의 예는 글루타티온-S-전환효소(glutathione-S-transferase: GST), 키틴 결합 단백질(chitin binding protein: CBP), 말토스 결합 단백질, 티오레독신(thioredoxin: TRX), 폴리(NANP), 탠덤 친화도 정제(tandem affinity purification: TAP) 태그, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, 혈구응집소(hemagglutinin: HA), nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 히스티딘(histidine: His), 바이오틴 카복실 운반 단백질(biotin carboxyl carrier protein: BCCP) 및 칼모둘린을 포함하지만, 이들로 제한되지는 않는다. 일부 실시형태에서, 이중성 분자는 면역글로불린 Fc 도메인, 펩타이드 태그, 형질도입 도메인, 폴리(에틸렌 글리콜), 폴리시알산 또는 글리콜산이다.

[0095] 일부 실시형태에서, 단리된 폴리펩타이드는 비천연 또는 변형된 아미노산 또는 펩타이드 유사체를 포함한다. 예를 들어, 많은 D-아미노산 또는 자연 발생적 아미노산과 상이한 관능성 치환기를 갖는 아미노산이 있다. 자연 발생적 펩타이드의 반대의 입체이성질체, 및 펩타이드 유사체의 입체이성질체가 개시되어 있다. 이들 아미노산은, 부위-특이적 방식으로 유사체 아미노산을 펩타이드 사슬 내로 삽입하도록, tRNA 분자를 선택된 아미노산으로 하전시키고, 예를 들어 앰버 코돈을 사용하는 유전자 작제물을 조작함으로써, 폴리펩타이드 사슬 내로 용이하게 혼입될 수 있다.

[0096] 일부 실시형태에서, 단리된 폴리펩타이드는 펩타이드 모방체이며, 이들은 펩타이드를 닮도록 제조될 수 있지만 천연 펩타이드 연결을 통해 연결되어 있지 않다. 예를 들어, 아미노산 또는 아미노산 유사체를 위한 연결은 $-CH_2NH-$, $-CH_2S-$, $-CH_2-$, $-CH=CH-$ (시스 및 트랜스), $-COCH_2-$, $-CH(OH)CH_2-$ 및 $-CHHSO-$ 를 포함하지만, 이들로 제한되지는 않는다. 펩타이드 유사체, 예컨대 b-알라닌, g아미노부티르산 등은 결합 원자 사이에 1개보다 많은 원자를 가질 수 있다. 아미노산 유사체 및 펩타이드 유사체는 대개 증대된 또는 바람직한 특성, 예컨대 더 경제적인 생산, 더 큰 화학 안정성, 증대된 약물학적 특성(반감기, 흡수, 역가, 효능 등), 변경된 특이성(예를 들어, 생물학적 활성의 광역 스펙트럼), 감소된 항원성 및 다른 바람직한 특성을 갖는다.

[0097] 일부 실시형태에서, 단리된 폴리펩타이드는 D-아미노산을 포함하는데, D 아미노산은 펩티다제에 의해 인식되지 않기 때문에 더 안정한 펩타이드를 생성하도록 사용될 수 있다. 동일한 유형의 D-아미노산에 의한 공통 서열의 하나 이상의 아미노산의 시스템 치환(예를 들어, L-리신 대신에 D-리신)은 더 안정한 펩타이드를 생성하도록 사용될 수 있다. 시스템인 잔기는 2개 이상의 펩타이드를 함께 고리화하거나 또는 부착시키도록 사용될 수 있다. 이것은 특정한 입체형태로 펩타이드를 구속시키기 위해 유리할 수 있다(예를 들어, Rizo and Gierasch, *Ann. Rev. Biochem.*, 1992, 61, 387 참조).

[0098] 본 개시내용은 또한 본원에 개시된 임의의 폴리펩타이드를 암호화하는 핵산 분자를 제공한다. 이것은 특이적 폴리펩타이드 서열과 관련된 모든 축퇴성 서열(하나의 특정한 폴리펩타이드 서열을 암호화하는 서열을 갖는 모든 핵산, 및 단백질 서열의 개시된 변이체 및 유도체를 암호화하는 축퇴성 핵산을 포함하는 핵산)을 포함한다. 이와 같이, 각각의 특정한 핵산 서열이 본원에 세세히 기재되지 않을 수 있지만, 각각의 및 모든 서열은 사실 개시된 폴리펩타이드 서열에 걸쳐 본원에 개시되고 기재된다.

[0099] 핵산 내의 핵산 서열 또는 폴리펩타이드 내의 아미노산 서열의 특정한 스트레치 사이의 퍼센트 동일성(또는 퍼센트 상보성)은 스미스(Smith) 및 워터만(Waterman)(*Adv. Appl. Math.*, 1981, 2, 482-489)의 알고리즘을 이용하는 디폴트 설정을 이용하여 BLAST 프로그램(기본 국소 정렬 조사 도구) 및 PowerBLAST 프로그램(Altschul *et al.*, *J. Mol. Biol.*, 1990, 215, 403-410; Zhang and Madden, *Genome Res.*, 1997, 7, 649-656)을 사용하여 또는 Gap 프로그램(Wisconsin Sequence Analysis Package, 유닉스용 버전 8, Genetics Computer Group(위스콘신

주 매디슨 유니버시티 리서치 파크))을 사용함으로써 일상적으로 결정될 수 있다. 본원에서, 퍼센트 서열 동일성이 언급되는 경우, 서열 동일성의 더 높은 백분율이 더 낮은 것에 비해 바람직하다.

[0100] 본 개시내용은 또한 본원에 개시된 임의의 하나 이상의 핵산 분자 및/또는 임의의 하나 이상의 폴리펩타이드 및 담체 및/또는 부형제를 포함하는 조성물을 제공한다. 일부 실시형태에서, 담체는 핵산 분자 및/또는 폴리펩타이드의 안정성을 증가(예를 들어, 분해 생성물이 출발 핵산 또는 단백질의 0.5 중량% 미만과 같은 한계치 아래에 있는 주어진 저장 조건(예를 들어, -20℃, 4℃ 또는 주위 온도) 하에 기간의 연장; 또는 생체내 안정성의 증가)시킨다. 담체의 예는 폴리(락트산)(PLA) 마이크로구, 폴리(D,L-락트산-코-글리콜산)(PLGA) 마이크로구, 리포솜, 마이셀, 인버스 마이셀, 지질 코클레이트 및 지질 미세소관을 포함하지만, 이들로 제한되지는 않는다. 담체는 완충 염 용액, 예컨대 PBS, HBSS 등을 포함할 수 있다.

[0101] 본 개시내용은 또한 본원에 개시된 임의의 폴리펩타이드 또는 그의 단편을 제조하는 방법을 제공한다. 이러한 폴리펩타이드 또는 그의 단편은 임의의 적합한 방법에 의해 제조될 수 있다. 예를 들어, 폴리펩타이드 또는 그의 단편은 이러한 폴리펩타이드 또는 그의 단편을 암호화하는 핵산 분자(예를 들어, 재조합 발현 벡터)를 포함하는 숙주 세포로부터 제조될 수 있다. 이러한 방법은 폴리펩타이드 또는 그의 단편을 제조하기에 충분한 조건 하에 폴리펩타이드 또는 그의 단편을 암호화하는 핵산 분자(예를 들어, 재조합 발현 벡터)를 포함하는 숙주 세포를 배양하여서 폴리펩타이드 또는 그의 단편을 제조하는 단계를 포함할 수 있다. 핵산은 숙주 세포에서 활성인 프로모터에 작동 가능하게 연결될 수 있고, 배양은 핵산이 발현되는 조건하에 수행될 수 있다. 이러한 방법은 발현된 폴리펩타이드 또는 그의 단편을 회수하는 단계를 추가로 포함할 수 있다. 회수는 폴리펩타이드 또는 그의 단편을 정제하는 것을 추가로 포함할 수 있다.

[0102] 단백질 발현에 적합한 시스템의 예는 숙주 세포, 예를 들어 박테리아 세포 발현 시스템(예를 들어, 에스체리치아 콜라이(*Escherichia coli*), 락토코커스 락티스(*Lactococcus lactis*)), 효모 세포 발현 시스템(예를 들어, 사카로마이세스 세레비시아에(*Saccharomyces cerevisiae*), 피치아 파스토리스(*Pichia pastoris*)), 곤충 세포 발현 시스템(예를 들어, 바콜로바이러스 매개된 단백질 발현) 및 포유류 세포 발현 시스템을 포함한다.

[0103] 폴리펩타이드 또는 그의 단편을 암호화하는 핵산 분자의 예는 본원에 다른 곳에 더 자세히 개시되어 있다. 일부 실시형태에서, 핵산 분자는 숙주 세포에서 발현에 코돈 최적화된다. 일부 실시형태에서, 핵산 분자는 숙주 세포에서 활성인 프로모터에 작동 가능하게 연결된다. 프로모터는 이중성 프로모터(예를 들어, 자연 발생적 프로모터가 아닌 프로모터)일 수 있다. 에스체리치아 콜라이에 적합한 프로모터의 예는 아라비노스, *lac*, *tac* 및 T7 프로모터를 포함하지만, 이들로 제한되지는 않는다. 락토코커스 락티스에 적합한 프로모터의 예는 P170 및 니신 프로모터를 포함하지만, 이들로 제한되지는 않는다. 사카로마이세스 세레비시아에 적합한 프로모터의 예는 구성적 프로모터, 예컨대 알코올 탈수소효소(alcohol dehydrogenase: ADHI) 또는 엔올라제(enolase: ENO) 프로모터 또는 유도성 프로모터, 예컨대 PHO, CUP1, GAL1 및 G10을 포함하지만, 이들로 제한되지는 않는다. 피치아 파스토리스에 적합한 프로모터의 예는 알코올 옥시다제 I(AOX I) 프로모터, 글리세르알데하이드 3 포스페이트 탈수소효소(GAP) 프로모터 및 글루타티온 의존적 포름알데하이드 탈수소효소(FLDI) 프로모터를 포함하지만, 이들로 제한되지는 않는다. 바콜로바이러스-매개된 시스템에 적합한 프로모터의 예는 후기 바이러스 강한 폴리헤드린 프로모터이다.

[0104] 일부 실시형태에서, 핵산 분자는 단백질 정제를 용이하게 하도록 폴리펩타이드 또는 그의 단편과 인프레임으로 태그를 암호화한다. 태그의 예는 본원의 다른 곳에 개시되어 있다. 이러한 태그는, 예를 들어 파트너 리간드에 결합(예를 들어, 수지에 부동화)할 수 있어서, 태그화된 단백질은 모든 다른 단백질(예를 들어, 숙주 세포 단백질)로부터 분리될 수 있다. 친화도 크로마토그래피, 고성능 액체 크로마토그래피(high performance liquid chromatography: HPLC) 및 크기 배제 크로마토그래피(size exclusion chromatography: SEC)는 발현된 단백질의 순도를 개선하기 위해 이용될 수 있는 방법의 예이다.

[0105] 다른 방법은 또한 폴리펩타이드 또는 그의 단편을 제조하도록 이용될 수 있다. 예를 들어, 2개 이상의 펩타이드 또는 폴리펩타이드는 단백질 화학 기법에 의해 함께 연결될 수 있다. 예를 들어, 펩타이드 또는 폴리펩타이드는 Fmoc(9-플루오렌메틸옥시카보닐) 또는 Boc(*tert*-부틸옥시카르보노일) 화학물질을 사용하여 화학적으로 합성될 수 있다. 이러한 펩타이드 또는 폴리펩타이드는 표준 화학 반응에 의해 합성될 수 있다. 예를 들어, 펩타이드 또는 폴리펩타이드는 합성되고 이의 합성 수지로부터 절단되지 않을 수 있는 한편, 펩타이드 또는 단백질의 다른 단편은 합성되고 후속하여 수지로부터 절단될 수 있어서, 다른 단편에서 기능적으로 차단된 말단 기를 노출시킨다. 펩타이드 축합 반응에 의해, 이들 2개의 단편은 각각 이의 카복실 및 아미노 말단에서 펩타이드 결합을 통해 공유 연결될 수 있다. 대안적으로, 펩타이드 또는 폴리펩타이드는 본원에 기재된 바대로 생체내 독립적으

로 합성될 수 있다. 일단 단리되면, 이 독립적인 펩타이드 또는 폴리펩타이드는 유사한 펩타이드 축합 반응을 통해 펩타이드 또는 그의 단편을 형성하도록 연결될 수 있다.

- [0106] 일부 실시형태에서, 클로닝된 또는 합성 펩타이드 분절의 효소 결합은 비교적 짧은 펩타이드 단편이 더 큰 펩타이드 단편, 폴리펩타이드 또는 전체 단백질 도메인을 생성하도록 연결되게 한다(Abrahmsen *et al.*, *Biochemistry*, 1991, 30, 4151). 대안적으로, 합성 펩타이드의 천연 화학 결합은 더 짧은 펩타이드 단편으로부터 큰 펩타이드 또는 폴리펩타이드를 합성으로 삭제하도록 이용될 수 있다. 이 방법은 2-단계 화학 반응으로 구성될 수 있다(Dawson *et al.*, *Science*, 1994, 266, 776-779). 제1 단계는 초기 공유 생성물로서 티오에스테르 연결된 중간체를 생성하도록 아미노 말단 Cys 잔기를 함유하는 또 다른 비보호된 펩타이드 분절과의 비보호된 합성 펩타이드-티오에스테르의 화학선택적 반응일 수 있다. 반응 조건의 변화 없이는, 이 중간체는 결합 부위에서 천연 펩타이드 결합을 형성하도록 자발적인 빠른 분자내 반응을 겪을 수 있다.
- [0107] 일부 실시형태에서, 화학 결합의 결과가 비자연(비펩타이드) 결합이므로, 비보호된 펩타이드 분절은 화학적으로 연결될 수 있고, 여기서 펩타이드 분절 사이에 결합이 형성된다(Schnolzer *et al.*, *Science*, 1992, 256, 221).
- [0108] 일부 실시형태에서, 폴리펩타이드는 발현 후 변형, 예를 들어, 글라이코실화, 아세틸화 및 포스포릴화 등, 및 자연 발생적 및 비자연 발생적 둘 다인 당해 분야에 공지된 다른 변형을 보유할 수 있다. 폴리펩타이드는 전체 단백질 또는 이의 하위서열일 수 있다.
- [0109] 본 개시내용은 또한, 본원에 개시된 하나 이상의 폴리펩타이드를 암호화할 수 있는 폴리뉴클레오타이드를 포함하는 핵산 분자 또는 그의 보체를 포함하는 재조합 발현 벡터를 포함하는 숙주 세포를 배양하여서 폴리펩타이드를 제조하는 단계를 포함하는, 본원에 개시된 임의의 폴리펩타이드를 제조하는 방법을 제공한다.
- [0110] 본 개시내용은 또한, 핵산 분자를 포함하는 벡터를 포함하여, 본원에 개시된 임의의 하나 이상의 핵산 분자, 및/또는 임의의 하나 이상의 폴리펩타이드를 포함하는 세포(예를 들어, 재조합 숙주 세포)를 제공한다. 세포는 시험관내, 생체의 또는 생체내일 수 있다. 핵산 분자는 프로모터 및 다른 조절 서열에 연결될 수 있어서, 이들은 암호화된 단백질을 제조하도록 발현된다. 이러한 세포의 세포주는 추가로 제공된다.
- [0111] 일부 실시형태에서, 세포는 전능성 세포 또는 다능성 세포(예를 들어, 배아 줄기(embryonic stem: ES) 세포, 예컨대 설치류 ES 세포, 마우스 ES 세포 또는 래트 ES 세포)이다. 전능성 세포는 임의의 세포 유형을 생성시킬 수 있는 미분화 세포를 포함하고, 다능성 세포는 하나 초과와 분화 세포 유형으로 발생할 능력을 보유하는 미분화 세포를 포함한다. 이러한 다능성 및/또는 전능성 세포는, 예를 들어 ES 세포 또는 ES 유사 세포, 예컨대 유도 다능성 줄기(induced pluripotent stem: iPS) 세포일 수 있다. ES 세포는 배아로의 도입 시 배아를 발생시키는 임의의 조직에 기여할 수 있는 배아 유래 전능성 또는 다능성 세포를 포함한다. ES 세포는 배반포의 내부 세포 덩어리로부터 유래될 수 있고, 임의의 3개의 척추동물 배엽(내배엽, 외배엽 및 중배엽)의 세포로 분화할 수 있다. 본 개시내용에 따르면, 배아 줄기 세포는 비인간 배아 줄기 세포일 수 있다.
- [0112] 일부 실시형태에서, 세포는 1차 체세포 또는 1차 체세포가 아닌 세포이다. 체세포는 배우자, 생식 세포, 생식모 세포 또는 미분화 줄기 세포가 아닌 임의의 세포를 포함할 수 있다. 일부 실시형태에서, 세포는 또한 1차 세포일 수 있다. 1차 세포는 유기체, 장기 또는 조직으로부터 직접적으로 단리된 세포 또는 세포의 배양물을 포함한다. 1차 세포는 형질전환되지 않고 불멸이 아닌 세포를 포함한다. 1차 세포는 조직 배양에서 이전에 계대배양되지 않거나 조직 배양에서 이전에 계대배양되었지만 조직 배양에서 무기한으로 계대배양될 수 없는 유기체, 장기 또는 조직으로부터 수득된 임의의 세포를 포함한다. 이러한 세포는 종래의 기법에 의해 단리될 수 있고, 예를 들어 체세포, 조혈 세포, 내피 세포, 상피 세포, 섬유아세포, 간충직 세포, 각질세포, 멜라닌세포, 단핵구, 단핵 세포, 지방세포, 지방선구세포, 뉴런, 신경교 세포, 간세포, 골격 근원세포 및 평활근 세포를 포함한다. 예를 들어, 1차 세포는 연결 조직, 근육 조직, 신경계 조직 또는 상피 조직으로부터 유래될 수 있다.
- [0113] 일부 실시형태에서, 세포는 보통 무기한으로 증식하지 않을 수 있지만, 돌연변이 또는 변경으로 인해, 정상 세포 노화가 회피되고, 대신에 분열을 계속 겪을 수 있다. 이러한 돌연변이 또는 변경은 천연 발생할 수 있거나, 의도적으로 유도될 수 있다. 불활화된 세포의 예는 중국 햄스터 난소(Chinese hamster ovary: CHO) 세포, 인간 배아 신장 세포(예를 들어, HEK 293 세포) 및 마우스 배아 섬유아세포 세포(예를 들어, 3T3 세포)를 포함하지만, 이들로 제한되지는 않는다. 불활화된 세포의 많은 유형은 널리 공지되어 있다. 불활화된 또는 1차 세포는 재조합 유전자 또는 단백질을 배양하기 위해 또는 발현시키기 위해 전형적으로 사용되는 세포를 포함한다. 일부 실시형태에서, 세포는 분화 세포, 예컨대 간 세포(예를 들어, 인간 간 세포)이다.
- [0114] 세포는 임의의 소스 유래일 수 있다. 예를 들어, 세포는 진핵생물 세포, 동물 세포, 식물 세포 또는 진균(예를

들어, 효모) 세포일 수 있다. 이러한 세포는 어류 세포 또는 조류 세포일 수 있거나, 이러한 세포는 포유류 세포, 예컨대 인간 세포, 비인간 포유류 세포, 설치류 세포, 마우스 세포 또는 래트 세포일 수 있다. 포유류는 인간, 비인간 영장류, 원숭이, 유인원, 고양이, 개, 말, 황소, 사슴, 들소, 양, 설치류(예를 들어, 마우스, 래트, 햄스터, 기니아 피그), 가축(예를 들어, 소, 돼지, 양, 닭, 오리 등; 양, 돼지 등; 예컨대 양, 염소 등; 및 돼지 등, 예컨대 새끼돼지 및 수퇘지)을 포함하지만, 이들로 제한되지는 않는다. 조류는 닭, 칠면조, 타조, 거위, 오리 등을 포함하지만, 이들로 제한되지는 않는다. 길들인 동물 및 농업 동물이 또한 포함된다. "비인간 동물"의 용어는 인간을 배제한다.

[0115] 추가 숙주 세포는, 예를 들어 미국 특허 출원 공보 제US2018/0030114호에 기재되어 있으며, 이는 본원에 그 전문이 참고로 인용된다.

[0116] 본원에 개시된 핵산 분자 및 폴리펩타이드는 임의의 수단에 의해 세포로 도입될 수 있다. 형질주입 프로토콜, 및 세포로 핵산 또는 단백질을 도입하기 위한 프로토콜은 변할 수 있다. 비제한적인 형질주입 방법은 리포솜, 나노입자, 칼슘, 덴드리머, 및 양이온성 중합체, 예컨대 DEAE-텍스트란 또는 폴리에틸렌이민을 사용한 화학-기반 형질주입 방법을 포함한다. 비화학 방법은 전기천공, 초음파천공(sono-poration) 및 광학 형질주입을 포함한다. 입자 기반 형질주입은 유전자 총 또는 자석 보조 형질주입의 이용을 포함한다. 바이러스 방법은 형질주입에 또한 이용될 수 있다.

[0117] 세포로의 핵산 또는 단백질의 도입은 또한 전기천공에 의해, 세포질내 주사에 의해, 아데노바이러스에 의한, 아데노 연관된 바이러스에 의한, 렌티바이러스에 의한, 레트로바이러스에 의한, 바이러스 감염에 의해, 형질주입에 의해, 지질 매개된 형질주입에 의해 또는 뉴클레오펙션(nucleofection)에 의해 매개될 수 있다. 뉴클레오펙션은 핵산 기질이 세포질로뿐만 아니라 핵막을 통해 핵으로 전달되게 하는 개선된 전기천공 기술이다. 게다가, 본원에 개시된 방법에서의 뉴클레오펙션의 이용은 전형적으로 정기적인 전기천공보다 훨씬 더 적은 세포(예를 들어, 정기적인 전기천공에 의한 700 만개와 비교하여 불과 약 200 만개)를 요한다. 일부 실시형태에서, 뉴클레오펙션은 LONZA[®] NUCLEOFECTOR[™] 시스템을 이용하여 수행된다.

[0118] 세포로의 핵산 또는 단백질의 도입은 또한 미량주사에 의해 달성될 수 있다. mRNA의 미량주사는 보통 (예를 들어, 번역 기계로 mRNA를 직접 전달하기 위해) 세포질로 되지만, 단백질 또는 DNA의 미량주사는 보통 핵으로 된다. 대안적으로, 미량주사는 핵 및 세포질 둘 다로의 주사에 의해 수행될 수 있지만: 침은 처음에 핵으로 도입될 수 있고, 제1 양은 주사될 수 있고, 세포로부터 침을 제거하면서 제2 양이 세포질로 주사될 수 있다. 뉴클레아제 제제 단백질이 세포질로 주사되면, 단백질은 핵/전핵으로의 전달을 보장하도록 핵 국제화 신호를 포함할 수 있다.

[0119] 세포로 핵산 또는 단백질을 도입하기 위한 다른 방법은, 예를 들어 벡터 전달, 입자-매개된 전달, 엑소솜-매개된 전달, 지질-나노입자-매개된 전달, 세포-침투-펩타이드-매개된 전달, 또는 이식형-장치-매개된 전달을 포함할 수 있다. 생체내 세포를 변형시키기 위해 대상체에게 핵산 또는 단백질을 투여하는 방법은 본원 다른 곳에 개시되어 있다. 세포로의 핵산 및 단백질의 도입은 유체역학 전달(hydrodynamic delivery: HDD)에 의해 또한 달성될 수 있다.

[0120] 세포로 핵산 또는 단백질을 도입하기 위한 다른 방법은, 예를 들어 벡터 전달, 입자-매개된 전달, 엑소솜-매개된 전달, 지질-나노입자-매개된 전달, 세포-침투-펩타이드-매개된 전달 또는 이식형-장치-매개된 전달을 포함할 수 있다. 일부 실시형태에서, 핵산 또는 단백질은 담체, 예컨대 폴리(락트산)(PLA) 마이크로구, 폴리(D,L-락트산-코-글리콜산)(PLGA) 마이크로구, 리포솜, 마이셀, 인버스 마이셀, 지질 코클레이트 또는 지질 미세소관에서 세포로 도입될 수 있다.

[0121] 본 개시내용은 또한 프로브 및 프라이머를 제공한다. 프로브 및 프라이머의 예는, 예를 들어 상기에 개시되어 있다. 본 개시내용은 본원에 개시된 임의의 핵산 분자에 특이적으로 혼성화하는 핵산 서열을 포함하는 프로브 및 프라이머를 제공한다. 예를 들어, 프로브 또는 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 본원에 기재된 임의의 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함할 수 있다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 변이체 SIGIRR 단백질을 암호화하는 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 변이체 SIGIRR 폴리펩타이드를 암호화하고, 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두

되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 복수의 아미노산을 포함하는 변이체 SIGIRR 폴리펩타이드를 암호화하는 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 변이체 SIGIRR 폴리펩타이드를 암호화하는 핵산 분자에 혼성화하거나: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11), 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 아미노산 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 폴리펩타이드를 암호화하는 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 아미노산 서열을 포함하거나 또는 이들로 이루어지는 변이체 SIGIRR 폴리펩타이드를 암호화하는 핵산 분자에 혼성화하거나, 또는 이 핵산 분자의 보체에 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 분자의 일부에 특이적으로 혼성화한다.

[0122] 프로브 또는 프라이머는 임의의 적합한 길이를 포함할 수 있고, 이의 비제한적인 예는 적어도 약 5개, 적어도 약 8개, 적어도 약 10개, 적어도 약 11개, 적어도 약 12개, 적어도 약 13개, 적어도 약 14개, 적어도 약 15개, 적어도 약 16개, 적어도 약 17개, 적어도 약 18개, 적어도 약 19개, 적어도 약 20개, 적어도 약 21개, 적어도 약 22개, 적어도 약 23개, 적어도 약 24개 또는 적어도 약 25개의 뉴클레오타이드의 길이를 포함한다. 바람직한 실시형태에서, 프로브 또는 프라이머는 적어도 약 18개의 뉴클레오타이드의 길이를 포함한다. 프로브 또는 프라이머는 약 10개 내지 약 35개, 약 10개 내지 약 30개, 약 10개 내지 약 25개, 약 12개 내지 약 30개, 약 12개 내지 약 28개, 약 12개 내지 약 24개, 약 15개 내지 약 30개, 약 15개 내지 약 25개, 약 18개 내지 약 30개, 약 18개 내지 약 25개, 약 18개 내지 약 24개, 또는 약 18개 내지 약 22개의 뉴클레오타이드의 길이를 포함할 수 있다. 바람직한 실시형태에서, 프로브 또는 프라이머는 약 18개 내지 약 30개의 뉴클레오타이드의 길이이다.

[0123] 본 개시내용은 또한 변경-특이적 프로브 및 변경-특이적 프라이머를 제공한다. 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 핵산 서열에 상보적이고/이거나 혼성화하거나, 또는 특이적으로 혼성화하는 핵산 서열 또는 그의 보체를 포함한다. 본 개시내용의 맥락에서 "특이적으로 혼성화한다"는 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)가 야생형 SIGIRR 단백질을 암호화하는 핵산 분자에 혼성화하지 않는다는 것을 의미한다. 일부 실시형태에서, 변경-특이적 프로브는 서열번호 9에 따른 186번 위치 또는 그의 보체에 상응하는 위치에서 세린을 암호화하는 핵산 코돈에 특이적으로 혼성화한다. 일부 실시형태에서, 변경-특이적 프라이머 또는 프라이머 쌍은 변이체 SIGIRR 단백질을 암호화하는 핵산 분자의 영역(들)에 특이적으로 혼성화하여서, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈이 이로부터 제조된 임의의 전사체 내에 포함된다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 분자의 일부에 특이적으로 혼성화한다.

[0124] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 변이체 SIGIRR 단백질을 암호화하는 핵산 서열에 상보적이고/이거나 혼성화하거나, 또는 특이적으로 혼성화하는 핵산 서열 또는 그의 보체를 포함하고, 여기서 그 단백질은 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두를 포함한다.

[0125] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자에 상보적이고/이거나 혼성화하거나, 또는 특이적으로 혼성화하는 핵산 서열 또는 그의 보체를 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 핵산 서열을 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9와 적

어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 게놈 DNA의 일부에 특이적으로 혼성화한다.

[0126] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 2에 따른 9962번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 2와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 2에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다.

[0127] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 핵산 서열을 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 mRNA 분자의 일부에 특이적으로 혼성화한다.

[0128] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 4에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함하는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 4와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어지는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 4에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 mRNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다.

[0129] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는

서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 핵산 서열을 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 프로브 또는 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 cDNA 분자의 일부에 특이적으로 혼성화한다.

[0130] 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 6에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 각각 서열번호 6에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함하는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 6과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다. 일부 실시형태에서, 변경-특이적 프로브 또는 변경-특이적 프라이머는 서열번호 6에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자에 상보적이고/이거나, 혼성화하거나 특이적으로 혼성화하는 핵산 서열을 포함한다.

[0131] 본 개시내용의 프로브 또는 프라이머와 관련하여 상기에 기재된 길이는 또한 본 개시내용의 변경-특이적 프로브 또는 변경-특이적 프라이머에 대해 필요한 부분만 약간 수정하여 적용된다.

[0132] 본 개시내용은 또한 상기 기재된 바와 같은 변경-특이적 프라이머 중 2개를 포함하는 변경-특이적 프라이머의 쌍을 제공한다.

[0133] 일부 실시형태에서, 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)는 DNA를 포함한다. 일부 실시형태에서, 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)는 RNA를 포함한다. 일부 실시형태에서, 프로브 또는 프라이머(예를 들어, 변경-특이적 프로브 또는 변경-특이적 프라이머)는 엄격한 조건, 예컨대 높은 엄격한 조건하에 변이체 SIGIRR 단백질을 암호화하는 핵산 서열에 혼성화한다.

[0134] 일부 실시형태에서, 프로브는 표지를 포함한다. 일부 실시형태에서, 표지는 형광 표지, 방사선표지 또는 바이오틴이다. 일부 실시형태에서, 프로브의 길이는 상기에 기재되어 있다. 대안적으로, 일부 실시형태에서, 프로브는 적어도 약 20개, 적어도 약 25개, 적어도 약 30개, 적어도 약 35개, 적어도 약 40개, 적어도 약 45개, 적어도 약 50개, 적어도 약 55개, 적어도 약 60개, 적어도 약 65개, 적어도 약 70개, 적어도 약 75개, 적어도 약 80개, 적어도 약 85개, 적어도 약 90개, 적어도 약 95개 또는 적어도 약 100개의 뉴클레오타이드를 포함하거나 또는 이들로 이루어진다. 프로브(예를 들어, 대립유전자 특이적 프로브)는, 예를 들어 본원에 개시된 임의의 핵산 분자를 검출하기 위해 사용될 수 있다. 바람직한 실시형태에서, 프로브는 적어도 약 18개의 뉴클레오타이드의 길이를 포함한다. 프로브는 약 10개 내지 약 35개, 약 10개 내지 약 30개, 약 10개 내지 약 25개, 약 12개 내지 약 30개, 약 12개 내지 약 28개, 약 12개 내지 약 24개, 약 15개 내지 약 30개, 약 15개 내지 약 25개, 약 18개 내지 약 30개, 약 18개 내지 약 25개, 약 18개 내지 약 24개, 또는 약 18개 내지 약 22개의 뉴클레오타이드의 길이를 포함할 수 있다. 바람직한 실시형태에서, 프로브는 약 18개 내지 약 30개의 뉴클레오타이드의 길이이다.

- [0135] 본 개시내용은 또한 본원에 개시된 프로브의 임의의 하나 이상이 부착된 기질을 포함하는 지지체를 제공한다. 고체 지지체는 본원에 개시된 분자, 예컨대 임의의 프로브가 회합될 수 있는 고체 상태 기질 또는 지지체이다. 고체 지지체의 형태는 어레이이다. 고체 지지체의 또 다른 형태는 어레이 검출기이다. 어레이 검출기는 다수의 상이한 프로브가 어레이, 그리드 또는 다른 조직화된 패턴으로 커플링된 고체 지지체이다.
- [0136] 고체 지지체에서 사용하기 위한 고체-상태 기질은 분자가 커플링될 수 있는 임의의 고체 재료를 포함할 수 있다. 이것은 아크릴아미드, 아가로스, 셀룰로스, 니트로셀룰로스, 유리, 폴리스티렌, 폴리에틸렌 비닐 아세테이트, 폴리프로필렌, 폴리메타크릴레이트, 폴리에틸렌, 폴리에틸렌 옥사이드, 폴리실리케이트, 폴리카보네이트, 테플론, 플루오로카본, 나일론, 실리콘 고무, 폴리언하이드라이드, 폴리글리콜산, 폴리락트산, 폴리오르토에스테르, 폴리프로필푸마레이트, 콜라겐, 글리코사미노글리칸 및 폴리아미노산과 같은 재료를 포함한다. 고체 상태 기질은 박막, 막, 병, 접시, 섬유, 직조 섬유, 성형 중합체, 입자, 비드, 마이크로입자 또는 조합을 포함하는 임의의 유용한 형태를 가질 수 있다. 고체 상태 기질 및 고체 지지체는 다공성 또는 비다공성일 수 있다. 고체 상태 기질에 대한 형태는 미량정량 접시, 예컨대 표준 96웰 유형이다. 일부 실시형태에서, 웰마다 하나의 어레이를 보통 함유하는 멀티웰 유리 슬라이드가 사용될 수 있다. 이 특징은 분석법 재현성의 더 높은 제어, 증가된 쓰루풋 및 샘플 취급, 및 자동화의 용이를 허용한다. 일부 실시형태에서, 지지체는 마이크로어레이이다.
- [0137] 본원에 개시된 임의의 폴리펩타이드는 하나 이상의 치환(예컨대, 보존적 아미노산 치환), 삽입 또는 결실을 추가로 가질 수 있다. 삽입은, 예를 들어 아미노 또는 카복실 말단 융합, 및 단일 또는 다수의 아미노산 잔기의 서열내 삽입을 포함한다. 공지된 서열을 갖는 DNA에서 미리 결정된 부위에서 치환을 만드는 기법, 예를 들어 M13 프라이머 돌연변이유발 및 PCR 돌연변이유발은 널리 공지되어 있다. 아미노산 치환은 전형적으로 단일 잔기이지만, 한번에 다수의 상이한 위치에서 발생할 수 있고; 삽입은 보통 약 1개 내지 10개의 아미노산 잔기의 차수일 수 있고; 결실은 약 1개 내지 30개의 잔기의 범위일 것이다. 결실 또는 삽입은 인접한 쌍, 즉 2개의 잔기의 결실 또는 2개의 잔기의 삽입에서 이루어질 수 있다. 치환, 결실, 삽입 또는 임의의 이들의 조합은 최종 작제물에 도달하도록 조합될 수 있다. 일부 실시형태에서, 돌연변이는 리딩 프레임 밖에 서열을 위치시키지 않고, 2차 mRNA 구조를 생성하는 상보적 영역을 생성하지 않는다.
- [0138] 본 개시내용은 또한 조성물을 제조하고 본원에 기재된 방법을 이용하기 위한 키트를 제공한다. 본원에 기재된 키트는 대상체의 샘플에서 하나 이상의 유전자 변이체를 검출하기 위한 분석법 또는 분석법들을 포함한다.
- [0139] 일부 실시형태에서, SIGIRR 변이체의 인간 확인을 위한 키트는 상기에 기재된 조성물 및 방법을 사용한다. 일부 실시형태에서, 기본 키트는 본원에 개시된 임의의 핵산 분자(예를 들어, 서열번호 2, 서열번호 4, 및/또는 서열번호 6 등)에서의 유전좌위에 대한 적어도 한 쌍의 올리고뉴클레오타이드 프라이머 또는 프로브, 예컨대 변경-특이적 프로브 또는 변경-특이적 프라이머의 용기를 포함할 수 있다. 키트는 또한 사용 설명서를 선택적으로 포함할 수 있다. 키트는, 예를 들어 증폭된 유전좌위의 각각에 지향된 대립유전자 래더, 증폭을 위한 충분한 분량의 효소, 증폭을 용이하게 하는 증폭 완충제, 효소 활성을 용이하게 하는 2가 양이온 용액, 증폭 동안의 가닥연장을 위한 dNTP, 전기영동법을 위한 증폭된 재료의 제조를 위한 로딩 용액, 주형 대조군으로서의 게놈 DNA, 재료가 분리 매질에서 기대된 바대로 이동하도록 보장하는 크기 마커, 및 사용자를 가르치고 사용 시 오류를 제한하기 위한 프로토콜 및 매뉴얼 중 하나 이상과 같은 다른 선택적인 키트 성분을 또한 포함할 수 있다. 키트에서의 다양한 시약의 양은 또한 다수의 인자, 예컨대 공정의 최적 감수성에 따라 변할 수 있다. 매뉴얼 적용에서 사용하기 위한 시험 키트 또는 자동화 샘플 준비, 반응 셋업, 검출기 또는 분석기와 사용하기 위한 시험 키트를 제공하는 것은 이 교시내용의 범주 내에 있다.
- [0140] 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자 또는 그의 보체의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-

Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 키트는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0141] 일부 실시형태에서, 키트는 서열번호 2에 따른 9962번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 2와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 2에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 게놈 DNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0142] 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 키트는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0143] 일부 실시형태에서, 키트는 서열번호 4에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 각각 서열번호 4에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CUA 및 AGC를 포함하는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 4와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열

을 포함하거나 또는 이들로 이루어지는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 4에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 mRNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0144] 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 하기 아미노산 서열을 포함하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다: Ser-Arg-Ser-Trp-Ser-Gly-Val-Gly-Ala-Thr-Ser-Ser-Ser-Trp-Thr-Thr-Ala-Thr-Ser-Cys-Arg-Ala-Leu-Ser-Pro-Pro-Pro-Thr-Ser-Trp(서열번호 11). 일부 실시형태에서, 키트는 서열번호 9와 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 9를 갖는 변이체 SIGIRR 단백질을 암호화하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0145] 일부 실시형태에서, 키트는 서열번호 6에 따른 557번 위치에 상응하는 위치에서 구아닌을 포함하는 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 각각 서열번호 6에 따른 553번 내지 555번 및 556번 내지 558번 위치에 상응하는 위치에서 코돈 CTA 및 AGC를 포함하는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 6과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%의 서열 동일성을 갖는 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다. 일부 실시형태에서, 키트는 서열번호 6에 따른 핵산 서열을 포함하거나 또는 이들로 이루어지는 cDNA 분자의, 증폭을 위한 적어도 하나의 쌍의 올리고뉴클레오타이드 프라이머(예를 들어, 변경-특이적 프라이머), 또는 검출을 위한 적어도 하나의 표지된 올리고뉴클레오타이드 프로브(예를 들어, 변경-특이적 프로브)를 포함한다.

[0146] 일부 실시형태에서, 본원에 개시된 임의의 키트는 뉴클레오타이드 래더, 프로토콜, 효소(예컨대, 증폭, 예컨대 증합효소 연쇄 반응(polymerase chain reaction: PCR)에 사용되는 효소), dNTP, 완충제, 염 또는 염들, 및 조절 핵산 샘플 중 임의의 하나 이상을 추가로 포함할 수 있다. 일부 실시형태에서, 본원에 개시된 임의의 키트는 검출 가능한 표지, 어닐링 반응을 수행하는 데 필요한 제품 및 시약, 및 설명서 중 임의의 하나 이상을 추가로 포함할 수 있다.

[0147] 일부 실시형태에서, 본원에 개시된 키트는 서열번호 2의 9962번 위치에 상응하는 위치에서, 또는 서열번호 4 및/또는 서열번호 6의 557번 위치에 상응하는 위치에서 구아닌에 직접적으로 혼성화하는 3' 말단 뉴클레오타이드를 포함하는 프라이머 또는 프로브 또는 변경-특이적 프라이머 또는 변경-특이적 프로브를 포함할 수 있다.

- [0148] 당업자는 사용된 검출 기법이 일반적으로 제한이 아니라는 것을 이해한다. 오히려, 넓은 다양한 검출 수단은 개시된 방법 및 키트의 범주 내에 있고, 단 이들은 애플리곤의 존재 또는 부재가 결정되게 한다.
- [0149] 일부 양태에서, 키트는 본원에 개시된 프라이머 또는 프로브 중 하나 이상을 포함할 수 있다. 예를 들어, 키트는 개시된 유전자 변이체 중 하나 이상에 혼성화하는 하나 이상의 프로브를 포함할 수 있다.
- [0150] 일부 양태에서, 키트는 개시된 세포 또는 세포주 중 하나를 포함할 수 있다. 일부 양태에서, 키트는 형질전환 세포 또는 세포주를 생성하는 데 필요하는 재료를 포함할 수 있다. 예를 들어, 일부 양태에서, 키트는 개시된 유전자 변이체 중 하나 이상을 포함하는 핵산 서열을 포함하는 세포 및 벡터를 포함할 수 있다. 키트는 세포 배양에 대한 배지를 추가로 포함할 수 있다.
- [0151] 본 개시내용은 또한 대상체 인간으로부터의 생물학적 샘플에서 *SIGIRR* 변이체 게놈 DNA, mRNA, cDNA, 및/또는 폴리펩타이드의 존재를 검출하는 방법을 제공한다. 집단 내의 유전자 서열 및 이러한 유전자에 의해 암호화된 mRNA 및 단백질이 다형, 예컨대 단일 뉴클레오타이드 다형으로 인해 변할 수 있다고 이해된다. *SIGIRR* 게놈 DNA, mRNA, cDNA 및 폴리펩타이드에 대해 본원에 제공된 서열은 오직 예시적인 서열이다. *SIGIRR* 게놈 DNA, mRNA, cDNA, 및 폴리펩타이드에 대한 다른 서열이 또한 가능하다.
- [0152] 생물학적 샘플은 대상체로부터의 임의의 세포, 조직 또는 생체액으로부터 유래될 수 있다. 샘플은 임의의 임상적으로 관련된 조직, 예컨대 골수 샘플, 중앙 생검, 세침 흡입액, 또는 체액의 샘플, 예컨대 혈액, 치은열구액, 혈장, 혈청, 림프, 복수액, 낭종액 또는 소변의 샘플을 포함할 수 있다. 몇몇 경우에, 샘플은 협측 면봉을 포함한다. 본원에 개시된 방법에서 사용된 샘플은 분석법 포맷, 검출 방법의 성질, 및 샘플로서 사용되는 조직, 세포 또는 추출물을 기초하여 변할 수 있다. 생물학적 샘플은 이용되는 분석법에 따라 다르게 처리될 수 있다. 예를 들어, 변이체 *SIGIRR* 핵산 분자를 검출할 때, 게놈 DNA에 대해 샘플을 분리하거나 농후화하도록 설계된 예비처리가 이용될 수 있다. 다양한 공지된 기법은 이 목적에 이용될 수 있다. 변이체 *SIGIRR* mRNA의 수준을 검출할 때, 상이한 기법은 생물학적 샘플을 mRNA로 농후화시키도록 이용될 수 있다. mRNA의 존재 또는 수준 또는 특정한 변이체 게놈 DNA 유전자좌의 존재를 검출하기 위한 다양한 방법이 이용될 수 있다.
- [0153] 일부 실시형태에서, 본 개시내용은 단백질이 절두된 *SIGIRR* 단백질을 암호화하는 아미노산 서열을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 단백질의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 단백질의 존재 또는 부재를 검출하는 방법을 제공한다. 일부 실시형태에서, 본 개시내용은 단백질이 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하는 아미노산 서열을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 단백질의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 단백질의 존재 또는 부재를 검출하는 방법을 제공한다. 일부 실시형태에서, 본 개시내용은 단백질이 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하는 아미노산 서열을 포함하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 단백질의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 단백질의 존재 또는 부재를 검출하는 방법을 제공한다.
- [0154] 일부 실시형태에서, 본 개시내용은 핵산이 절두된 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 핵산의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 핵산 분자의 존재 또는 부재를 검출하는 방법을 제공한다. 일부 실시형태에서, 본 개시내용은 핵산이 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 핵산의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 핵산 분자의 존재 또는 부재를 검출하는 방법을 제공한다. 일부 실시형태에서, 본 개시내용은 핵산이 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 *SIGIRR* 단백질을 암호화하는 핵산 서열을 포함하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는지의 여부를 결정하기 위해 생물학적 샘플에서 핵산의 적어도 일부를 서열분석하는 단계를 포함하는 변이체 *SIGIRR* 핵산 분자의 존재 또는 부재를 검출하는 방법을 제공한다. 본원에 개시된 임의의 변이체 핵산 분자는 본원에 기재된 임의의 프로브 및 프라이머를 사용하여 검출될 수 있다.
- [0155] 일부 실시형태에서, 대상체에서 염증성 장 질환 연관된 변이체 *SIGIRR* 핵산 분자 또는 조기-발병 염증성 장 질환 연관된 변이체 *SIGIRR* 핵산 분자(예를 들어, 게놈 DNA, mRNA 또는 cDNA)의 존재 또는 부재를 검출하는 방법은 대상체로부터 획득된 생물학적 샘플에서 분석법을 수행하는 단계를 포함하고, 그 분석법은 생물학적 샘플에서의 핵산 분자가 본원에 개시된 임의의 변이체 *SIGIRR* 핵산 서열을 포함하는지의 여부를 결정한다(예를 들어, 절두된 *SIGIRR* 단백질을 암호화하는 핵산 분자, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된

SIGIRR 단백질을 암호화하는 핵산 분자, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 핵산 분자). 일부 실시형태에서, 생물학적 샘플은 세포 또는 세포 용해물을 포함한다. 이러한 방법은, 예를 들어 SIGIRR 게놈 DNA 또는 mRNA를 포함하는 대상체로부터의 생물학적 샘플을 수득하는 단계, 및 mRNA의 경우에 선택적으로 cDNA로 mRNA를 역전사시키는 단계, 및 SIGIRR 게놈 DNA, mRNA 또는 cDNA의 위치가 절두된 SIGIRR 단백질을 암호화하는지의 여부를 결정하는 생물학적 샘플에서의 분석법을 수행하는 단계를 추가로 포함할 수 있다. 이러한 방법은, 예를 들어, SIGIRR 게놈 DNA 또는 mRNA를 포함하는 대상체로부터의 생물학적 샘플을 수득하는 단계, 및 mRNA의 경우에 선택적으로 cDNA로 mRNA를 역전사시키는 단계, 및 SIGIRR 게놈 DNA, mRNA 또는 cDNA의 위치가 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는지의 여부를 결정하는 생물학적 샘플에서의 분석법을 수행하거나, 또는 SIGIRR 게놈 DNA, mRNA 또는 cDNA의 위치가 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되고 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 SIGIRR 단백질을 암호화하는지의 여부를 결정하는 생물학적 샘플에서의 분석법을 수행하는 단계를 추가로 포함할 수 있다. 이러한 분석법은, 예를 들어 특정한 SIGIRR 핵산 분자의 이들 위치의 식별을 결정하는 단계를 포함할 수 있다. 일부 실시형태에서, 대상체는 인간이다.

[0156] 일부 실시형태에서, 분석법은 대상체로부터의 생물학적 샘플에서 핵산 분자의 SIGIRR 게놈 DNA 서열의 적어도 일부를 서열분석하되, 서열분석된 일부는 서열번호 9에 따른 SIGIRR 단백질에서의 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치를 포함하는 것인 단계; 대상체로부터의 생물학적 샘플에서 핵산 분자의 SIGIRR mRNA 서열의 적어도 일부를 서열분석하되, 서열분석된 일부는 서열번호 9에 따른 SIGIRR 단백질에서의 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치를 포함하는 것인 단계; 또는 대상체로부터의 생물학적 샘플에서 핵산 분자의 SIGIRR cDNA 서열의 적어도 일부를 서열분석하되, 서열분석된 일부는 서열번호 9에 따른 SIGIRR 단백질에서의 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치를 포함하는 것인 단계를 포함한다.

[0157] 일부 실시형태에서, 검정법은 a) 생물학적 샘플을, i) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치에서 SIGIRR 게놈 서열의 위치에 가장 가까운 SIGIRR 게놈 DNA 서열의 일부; ii) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치에서 SIGIRR mRNA 서열의 위치에 가장 가까운 SIGIRR mRNA 서열의 일부; 또는 iii) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치에서 SIGIRR cDNA 서열의 위치에 가장 가까운 SIGIRR cDNA 서열의 일부에 혼성화하는 프라이머와 접촉시키는 단계; b) 적어도 i) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 코돈에 멀리 있는 뉴클레오타이드 위치에 상응하는 SIGIRR 게놈 DNA 서열의 위치; ii) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 코돈에 멀리 있는 뉴클레오타이드 위치에 상응하는 SIGIRR mRNA 서열의 위치; 또는 iii) 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 코돈에 멀리 있는 뉴클레오타이드 위치에 상응하는 SIGIRR cDNA 서열의 위치에 걸쳐 프라이머를 연장시키는 단계; 및 c) 프라이머의 연장 산물이 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 뉴클레오타이드를 포함하는지의 여부를 결정하는 단계를 포함한다. 일부 실시형태에서, 오직 SIGIRR 게놈 DNA가 분석된다. 일부 실시형태에서, 오직 SIGIRR mRNA가 분석된다. 일부 실시형태에서, SIGIRR mRNA로부터 수득된 SIGIRR cDNA가 오직 분석된다.

[0158] 일부 실시형태에서, 분석법은 a) 생물학적 샘플을, i) 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 뉴클레오타이드를 포함하는 SIGIRR 게놈 DNA 서열의 일부; ii) 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 뉴클레오타이드를 포함하는 SIGIRR mRNA 서열의 일부; 또는 iii) 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 뉴클레오타이드를 포함하는 SIGIRR cDNA 서열의 일부에 혼성화하는 변형-특이적 프라이머와 접촉시키는 단계; b) 변형-특이적 중합효소 연쇄 반응 기법을 이용하여 프라이머를 연장시키는 단계; 및 c) 연장이 발생하는지의 여부를 결정하는 단계를 포함한다. 변형-특이적 중합효소 연쇄 반응 기법은 핵산 서열에서의 결실과 같은 돌연변이를 검출하도록 이용될 수 있다. DNA 중합효소는 주형과의 미스매치가 존재할 때 연장되지 않으므로, 변형-특이적 프라이머가 사용된다. 기본적인 변형-특이적 중합효소 연쇄 반응 기법의 다수의 변형은 당업자의 재량에 따른다.

[0159] 변형-특이적 프라이머는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 SIGIRR 단백질을 암호화하는 핵산 서열에 상보적인 핵산 서열 또는 핵산 서열에 대한 보체를 포함할 수 있다. 예를 들어, 변형-특이적 프라이머는 서열번호 9를 암호화하는 핵산 서열에 상보적인 핵산 서열 또는 이 핵산 서열에 대한 보체를 포함할 수 있다. 변형-특이적 프라이머는 바람직하게는, 핵산 서열이 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화할 때, 변이체 SIGIRR 단백질을 암호화하는 핵산 서열에 특이적으로 혼성화한다.

[0160] 일부 실시형태에서, 분석법은 생물학적 샘플을 엄격한 조건하에 변이체 SIGIRR 게놈 DNA 서열, mRNA 서열 또는

cDNA 서열에 특이적으로 혼성화하고 상응하는 야생형 *SIGIRR* 서열에 혼성화하지 않는 프라이머 또는 프로브와 접촉시키는 단계, 및 혼성화가 발생하는지의 여부를 결정하는 단계를 포함한다.

[0161] 일부 실시형태에서, 분석법은 RNA 서열분석(RNA sequencing: RNA-Seq)을 포함한다. 일부 실시형태에서, 분석법은 또한 역전사효소 증합효소 연쇄 반응(reverse transcriptase polymerase chain reaction: RT-PCR)을 통해 cDNA로 mRNA를 역전사시키는 단계를 포함한다.

[0162] 일부 실시형태에서, 상기 방법은 표적 핵산 서열에 결합하고, 변이체 *SIGIRR* 게놈 DNA, mRNA 또는 cDNA를 포함하는 폴리뉴클레오타이드를 특이적으로 검출하고/하거나 확인하기에 충분한 뉴클레오타이드 길이의 프로브 및 프라이머를 사용한다. 혼성화 조건 또는 반응 조건은 이 결과를 달성하기 위해 조작자에 의해 결정될 수 있다. 이 뉴클레오타이드 길이는 본원에 기재되거나 예시된 임의의 분석법을 포함하는 선택의 검출 방법에서 사용하기에 충분한 임의의 길이일 수 있다. 일반적으로, 예를 들어 약 8개, 약 10개, 약 11개, 약 12개, 약 14개, 약 15개, 약 16개, 약 18개, 약 20개, 약 22개, 약 24개, 약 26개, 약 28개, 약 30개, 약 40개, 약 50개, 약 75개, 약 100개, 약 200개, 약 300개, 약 400개, 약 500개, 약 600개 또는 약 700개, 또는 이를 초과하는 뉴클레오타이드, 또는 약 11개 내지 약 20개, 약 20개 내지 약 30개, 약 30개 내지 약 40개, 약 40개 내지 약 50개, 약 50개 내지 약 100개, 약 100개 내지 약 200개, 약 200개 내지 약 300개, 약 300개 내지 약 400개, 약 400개 내지 약 500개, 약 500개 내지 약 600개, 약 600개 내지 약 700개, 또는 약 700개 내지 약 800개, 또는 이를 초과하는 뉴클레오타이드의 길이를 갖는 프라이머 또는 프로브를 사용한다. 바람직한 실시형태에서, 프로브 또는 프라이머는 적어도 약 18개의 뉴클레오타이드의 길이를 포함한다. 프로브 또는 프라이머는 약 10개 내지 약 35개, 약 10개 내지 약 30개, 약 10개 내지 약 25개, 약 12개 내지 약 30개, 약 12개 내지 약 28개, 약 12개 내지 약 24개, 약 15개 내지 약 30개, 약 15개 내지 약 25개, 약 18개 내지 약 30개, 약 18개 내지 약 25개, 약 18개 내지 약 24개, 또는 약 18개 내지 약 22개의 뉴클레오타이드의 길이를 포함할 수 있다. 바람직한 실시형태에서, 프로브 또는 프라이머는 약 18개 내지 약 30개의 뉴클레오타이드의 길이이다.

[0163] 이러한 프로브 및 프라이머는 높은 엄격성 혼성화 조건하에 표적 서열에 특이적으로 혼성화할 수 있다. 프로브 및 프라이머는 표적 서열과의 인접 뉴클레오타이드의 완전한 핵산 서열 동일성을 가질 수 있지만, 표적 핵산 서열과 상이하고, 표적 핵산 서열을 특이적으로 검출하고/하거나 확인하는 능력을 보유하는 프로브는 종래의 방법에 의해 설계될 수 있다. 따라서, 프로브 및 프라이머는 표적 핵산 분자와 약 80%, 약 85%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 또는 100%의 서열 동일성 또는 상보성을 공유할 수 있다.

[0164] 일부 실시형태에서, 특이적 프라이머는 생물학적 샘플에서 변이체 *SIGIRR* 유전좌위를 확인하기 위해 또는 특이적 *SIGIRR* mRNA 또는 cDNA의 수준을 결정하기 위해 특이적 프로브로서 사용될 수 있거나 자체가 검출될 수 있는 앰플리콘을 제조하도록 변이체 *SIGIRR* 유전좌위 및/또는 *SIGIRR* 변이체 mRNA 또는 cDNA를 증폭시키도록 사용될 수 있다. *SIGIRR* 변이체 유전좌위는 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치에 상응하는 위치를 포함하는 게놈 핵산 서열을 나타내도록 사용될 수 있다. 프로브가 핵산 분자에 대한 프로브의 결합을 허용하는 조건 하에 생물학적 샘플에서 핵산 분자와 혼성화될 때, 이 결합은 생물학적 샘플에서 검출될 수 있고, 변이체 *SIGIRR* 유전좌위의 존재 또는 변이체 *SIGIRR* mRNA 또는 cDNA의 존재 또는 수준의 표시를 허용한다. 결합된 프로브의 이러한 확인이 기재되어 있다. 특이적 프로브는 변이체 *SIGIRR* 유전자의 특이적 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다. 특이적 프로브는 변이체 *SIGIRR* mRNA의 특이적 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다. 특이적 프로브는 변이체 *SIGIRR* cDNA의 특이적 영역과 적어도 약 80%, 약 80% 내지 약 85%, 약 85% 내지 약 90%, 약 90% 내지 약 95%, 및 약 95% 내지 약 100% 동일한(또는 상보적인) 서열을 포함할 수 있다.

[0165] 일부 실시형태에서, 생물학적 샘플의 핵산 보체가 변이체 *SIGIRR* 단백질(예를 들어, 절두된 *SIGIRR* 단백질 또는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖는 변이체 *SIGIRR* 단백질)을 암호화하는 핵산 서열을 포함하는지의 여부를 결정하기 위해, 생물학적 샘플은, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 위치에서의 뉴클레오타이드의 존재에 대해 진단적인 앰플리콘을 제조하도록, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 위치에 인접한 5' 플랭킹 서열로부터 유래된 제1 프라이머, 및 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 위치에 인접한 3' 플랭킹 서열로부터 유래된 제2 프라이머를 포함하는 프라이머 쌍을 사용하여 핵산 증폭 방법으로 처리될 수 있다. 일부 실시형태에서, 앰플리콘은 프라이머 쌍과 하나의 뉴클레오타이드 염기 쌍의 합산 길이로부터 DNA 증폭 프로토콜

에 의해 제조 가능한 앰플리콘의 임의의 길이까지의 길이의 범위일 수 있다. 이 거리는 1개의 뉴클레오타이드 염기 쌍으로부터 증폭 반응의 한계, 또는 약 20000개의 뉴클레오타이드 염기 쌍의 범위일 수 있다. 선택적으로, 프라이머 쌍은 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치를 포함하는 영역 및 서열번호 9에 따른 186번 위치에서 세린을 암호화하는 위치의 각각의 측에서 적어도 1개, 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 또는 이를 초과하는 뉴클레오타이드를 측정시킨다. 유사한 앰플리콘은 mRNA 및/또는 cDNA 서열로부터 생성될 수 있다.

- [0166] 프로브 및 프라이머를 제조하고 사용하기 위한 대표적인 방법은, 예를 들어 *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Vol. 1-3, ed. Sambrook *et al.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 1989 (이하, "Sambrook *et al.*, 1989"); *Current Protocols in Molecular Biology*, ed. Ausubel *et al.*, Greene Publishing and Wiley-Interscience, New York, 1992(정기적으로 업데이트됨)(이하, "Ausubel *et al.*, 1992"); 및 Innis *et al.*, PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990)에 기재되어 있다. PCR 프라이머 쌍은, 예를 들어 그 목적을 위해 의도된 컴퓨터 프로그램, 예컨대 Vector NTI 버전 10(Informax Inc.(메릴랜드주 베테스다)); PrimerSelect(DNASTAR Inc.(위스콘신주 매디슨)); 및 Primer3(Version 0.4.0.COPYRG.T., 1991, Whitehead Institute for Biomedical Research(메사추세츠주 캠브리지))에서의 PCR 프라이머 분석 도구를 이용함으로써 공지된 서열로부터 유래될 수 있다. 추가적으로, 서열은 가시적으로 스캐닝될 수 있고, 프라이머는 공지된 가이드라인을 이용하여 수동으로 확인될 수 있다.
- [0167] 임의의 핵산 혼성화 또는 증폭 또는 서열분석 방법은 변이체 *SIGIRR* 유전자 유전좌위의 존재 및/또는 변이체 *SIGIRR* mRNA 또는 mRNA로부터 제조된 cDNA의 수준을 특이적으로 검출하도록 이용될 수 있다. 일부 실시형태에서, 핵산 분자는 *SIGIRR* 핵산의 영역을 증폭시키도록 프라이머로서 사용될 수 있거나, 핵산 분자는, 예를 들어 엄격한 조건하에, 변이체 *SIGIRR* 유전자 유전좌위를 포함하는 핵산 분자 또는 변이체 *SIGIRR* mRNA 또는 mRNA로부터 제조된 cDNA를 포함하는 핵산 분자에 특이적으로 혼성화하는 프로브로서 사용될 수 있다.
- [0168] 다양한 기법은, 예를 들어 핵산 서열분석, 핵산 혼성화 및 핵산 증폭을 포함하여 당해 분야에서 이용 가능하다. 핵산 서열분석 기법의 예시적인 예는 사슬 종결자 (Sanger) 서열분석 및 염료 종결자 서열분석을 포함하지만, 이들로 제한되지는 않는다.
- [0169] 다른 방법은 정제된 DNA, 증폭된 DNA 및 고정된 세포 준비에 지향된 표지된 프라이머 또는 프로브를 사용하는 것을 포함하는 서열분석 이외의 핵산 혼성화 방법(형광 동소 혼성화(fluorescence *in situ* hybridization)(FISH))을 수반한다. 몇몇 방법에서, 표적 핵산은 검출 전에 또는 검출과 동시에 증폭될 수 있다. 핵산 증폭 기법의 예시적인 예는 중합효소 연쇄 반응(PCR), 리가제 연쇄 반응(ligase chain reaction: LCR), 가닥 대체 증폭(strand displacement amplification: SDA) 및 핵산 서열 기반 증폭(nucleic acid sequence based amplification: NASBA)을 포함하지만, 이들로 제한되지는 않는다. 다른 방법은 리가제 연쇄 반응, 가닥 대체 증폭 및 호열성 SDA(thermophilic SDA)(tSDA)를 포함하지만, 이들로 제한되지는 않는다.
- [0170] 예를 들어, 혼성화 보호 분석법(Hybridization Protection Assay: HPA), 실시간의 증폭 과정의 정량적 평가 및 샘플에 초기에 존재하지만 실시간 증폭에 기초하지 않은 표적 서열의 분량의 결정을 포함하는, 임의의 방법은 비증폭된 또는 증폭된 폴리뉴클레오타이드를 검출하기 위해 이용될 수 있다.
- [0171] 반드시 서열 증폭을 요하지 않고, 예를 들어 염색체 물질의 썬던(DNA:DNA) 블롯 혼성화, 동소 혼성화(*in situ* hybridization: ISH) 및 형광 동소 혼성화(FISH)의 공지된 방법에 기초한, 핵산을 확인하는 방법이 또한 제공된다. 썬던 블로팅은 특이적 핵산 서열을 검출하도록 이용될 수 있다. 이러한 방법에서, 샘플로부터 추출된 핵산은 단편화되고, 매트릭스 겔에서 전기영동으로 분리되고, 막 필터로 이동된다. 필터 결합된 핵산은 관심 대상의 서열에 상보적인 표지된 프로브와의 혼성화로 처리된다. 필터에 결합된 혼성화된 프로브는 검출된다. 임의의 이러한 방법에서, 그 공정은 본원에 기재되거나 예시된 임의의 프로브를 사용한 혼성화를 포함할 수 있다.
- [0172] 혼성화 기법에서, 엄격한 조건은 프로브 또는 프라이머가 이의 표적에 특이적으로 혼성화하도록 이용될 수 있다. 일부 실시형태에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 엄격한 조건하에 다른 서열(예를 들어, 상응하는 야생형 *SIGIRR* 유전좌위, 야생형 mRNA 또는 야생형 cDNA)보다 검출 가능하게 더 높은 정도로, 예컨대 배경에 비해 적어도 2배, 적어도 3배, 적어도 4배, 또는 배경에 비해 10배 초과를 포함하여 그 초과로 이의 표적 서열(예를 들어, 변이체 *SIGIRR* 유전자 유전좌위, 변이체 *SIGIRR* mRNA 또는 변이체 *SIGIRR* cDNA)에 혼성화할 것이다. 일부 실시형태에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 엄격한 조건하에 다른 서열보다 검출 가능하게 더 높은 정도로 적어도 2배로 이의 표적 서열에 혼성화할 것이다. 일부 실시형태에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 엄격한 조건하에 다른 서열보다 검출 가능하게 더 높은 정도로 적어도 3배로 이의

표적 서열에 혼성화할 것이다. 일부 실시형태에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 엄격한 조건하에 다른 서열보다 검출 가능하게 더 높은 정도로 적어도 4배로 이의 표적 서열에 혼성화할 것이다. 일부 실시형태에서, 폴리뉴클레오타이드 프라이머 또는 프로브는 엄격한 조건하에 다른 서열보다 검출 가능하게 더 높은 정도로 배경에 비해 10배 초과로 이의 표적 서열에 혼성화할 것이다. 엄격한 조건은 서열 의존적이고, 상이한 환경에서 상이할 것이다. 혼성화의 엄격성 및/또는 세척 조건을 제어함으로써, 프로브와 100% 상보적인 표적 서열이 확인될 수 있다(동종성 프로빙). 대안적으로, 엄격성 조건은 더 낮은 정도의 동일성이 검출되도록(이종성 프로빙) 서열에서 약간의 미스매칭을 허용하도록 조정될 수 있다.

[0173] DNA 혼성화, 예를 들어 약 45°C에서의 6X 염화나트륨/시트르산나트륨(SSC), 이어서 50°C에서의 2X SSC의 세척을 촉진하는 적절한 엄격성 조건은 공지되거나, *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6에서 발견될 수 있다. 전형적으로, 혼성화 및 검출에 대한 엄격한 조건은 염 농도가 pH 7.0 내지 8.3에서 약 1.5 M 미만의 Na 이온, 전형적으로 약 0.01 내지 1.0 M Na 이온 농도(또는 다른 염)이고, 온도가 짧은 프로브(예를 들어, 10개 내지 50개의 뉴클레오타이드)에 대해 적어도 약 30°C 및 더 긴 프로브(예를 들어, 50개 초과 뉴클레오타이드)에 대해 적어도 약 60°C인 것일 것이다. 엄격한 조건은 탈안정화제, 예컨대 포름아미드의 첨가에 의해 또한 달성될 수 있다. 예시적인 낮은 엄격성 조건은 37°C에서의 30 내지 35% 포름아미드, 1 M NaCl, 1% SDS(황산 나트륨 도데실)의 완충제 용액에 의한 혼성화, 및 50 내지 55°C에서의 1X 내지 2X SSC(20X SSC = 3.0 M NaCl/0.3 M 시트르산삼나트륨) 중의 세척을 포함한다. 예시적인 보통의 엄격성 조건은 37°C에서의 40 내지 45% 포름아미드, 1.0 M NaCl, 1% SDS 중의 혼성화, 및 55 내지 60°C에서의 0.5X 내지 1X SSC 중의 세척을 포함한다. 예시적인 높은 엄격성 조건은 37°C에서의 50% 포름아미드, 1 M NaCl, 1% SDS 중의 혼성화, 및 60 내지 65°C에서의 0.1X SSC 중의 세척을 포함한다. 선택적으로, 세척 완충제는 약 0.1% 내지 약 1% SDS를 포함할 수 있다. 혼성화의 기간은 일반적으로 약 24시간 미만, 보통 약 4 내지 약 12시간이다. 세척 시간의 기간은 적어도 평형에 도달하기에 충분한 시간의 길이일 것이다.

[0174] 혼성화 반응에서, 특이성은 전형적으로 혼성화 후 세척의 함수이고, 중요한 인자는 최종 세척 용액의 이온 농도 및 온도이다. DNA-DNA 하이브리드에 대해, T_m 은 Meinkoth and Wahl, *Anal. Biochem.*, 1984, 138, 267-284의 식으로부터 근사치화될 수 있다: $T_m = 81.5^\circ\text{C} + 16.6(\log M) + 0.41(\% \text{ GC}) - 0.61(\% \text{ form}) - 500/L$; 여기서 M은 1가 양이온의 몰농도이고, %GC는 DNA에서의 구아노신 및 사이토신 뉴클레오타이드의 백분율이고, % form은 혼성화 용액 중의 포름아미드의 백분율이고, L은 염기 쌍에서의 하이브리드의 길이이다. T_m 은 상보적 표적 서열의 50%가 (한정된 이온 농도 및 pH 하에) 완벽하게 일치된 프로브에 혼성화하는 온도이다. T_m 은 미스매칭의 각각의 1%에 대해 약 1°C만큼 감소하고; 이에 의해, T_m , 혼성화, 및/또는 세척 조건은 원하는 동일성의 서열에 혼성화하도록 조정될 수 있다. 예를 들어, 90% 이상의 동일성을 갖는 서열이 추구되는 경우, T_m 은 10°C 감소할 수 있다. 일반적으로, 엄격한 조건은 한정된 이온 농도 및 pH에서 특이적 서열 및 그의 보체에 대한 열 융점(T_m)보다 약 5°C 더 낮도록 선택된다. 그러나, 심하게 엄격한 조건은 열 융점(T_m)보다 1°C, 2°C, 3°C 또는 4°C 낮은 온도에서의 혼성화 및/또는 세척을 이용할 수 있고; 보통의 엄격한 조건은 열 융점(T_m)보다 6°C, 7°C, 8°C, 9°C 또는 10°C 낮은 온도에서의 혼성화 및/또는 세척을 이용할 수 있고; 낮은 엄격성 조건은 열 융점(T_m)보다 11°C, 12°C, 13°C, 14°C, 15°C 또는 20°C 낮은 온도에서의 혼성화 및/또는 세척을 이용할 수 있다. 식, 혼성화 및 세척 조성물, 및 원하는 T_m 을 이용하여, 보통의 기술자는 혼성화의 엄격성 및/또는 세척 용액의 변동이 본질적으로 기재되어 있다는 것을 이해한다. 원하는 정도의 미스매칭이 45°C(수성 용액) 또는 32°C(포름아미드 용액) 미만의 T_m 을 발생시키는 경우, 더 높은 온도가 이용될 수 있도록 SSC 농도를 증가시키는 것이 최적이다.

[0175] 예를 들어, 단백질 서열분석 및 면역검정을 포함하여 생물학적 샘플에서 변이체 SIGIRR 폴리펩타이드의 존재를 검출하거나 수준을 정량화하는 방법이 또한 제공된다. 일부 실시형태에서, 인간 대상체에서 변이체 SIGIRR 단백질(예를 들어, 서열 번호 9)의 존재를 검출하는 방법은 생물학적 샘플에서 변이체 SIGIRR 단백질(예를 들어, 서열번호 4)의 존재를 검출하는 인간 대상체로부터의 생물학적 샘플에서 분석법을 수행하는 단계를 포함한다.

[0176] 단백질 서열분석 기법의 예시적인 비제한적인 예는 질량 분광법 및 Edman 분해를 포함하지만, 이들로 제한되지는 않는다. 면역검정의 예시적인 예는 면역침강, 웨스턴 블롯, 면역조직화학, ELISA, 면역세포화학, 유세포분석법 및 이뮤노-PCR을 포함하지만, 이들로 제한되지는 않는다. 다양한 공지된 기법(예를 들어, 열량측정, 형광, 화학발광 또는 방사성)을 이용하여 검출 가능하게 표지된 다중클론 또는 단일클론 항체는 면역검정에서 사용하기에 적합하다. 면역검정과 관련하여, 변이체 SIGIRR 단백질은 야생형 SIGIRR 단백질과 비교하여 상이한 크기를

갖고, 따라서 단백질 겔에서 상이한 분자량에서 흐른다. 이와 같이, 동일한 항체를 사용함으로써, 야생형 SIGIRR 단백질은, 예를 들어 웨스턴 블롯 검정에서 변이체 SIGIRR 단백질과 구별될 수 있다.

- [0177] 본 개시내용은 또한, 인간 대상체로부터 수득된 핵산 분자에서 본원에 기재된 임의의 SIGIRR 핵산 분자에서의 임의의 변형을 검출하는 단계; 및 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는, 인간 대상체에서 염증성 장 질환 또는 조기-발병 염증성 장 질환을 진단하기 위한 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 검출하기 위한 방법을 제공한다. 일부 실시형태에서, 인간 대상체는 이러한 진단을 필요로 한다. 일부 실시형태에서, 인간 대상체는 염증성 장 질환 또는 조기-발병 염증성 장 질환이 진단된 친척을 가질 수 있다.
- [0178] 염증성 장 질환 또는 조기-발병 염증성 장 질환의 증상은 설사, 열, 피로, 복통, 복부 경련, 구역, 구토, 혈변의 존재, 빈혈, 식욕 감소 및 의도하지 않은 체중 감소, 또는 임의의 이들의 조합을 포함하지만, 이들로 제한되지는 않는다.
- [0179] 일부 실시형태에서, 상기 방법은 대상체로부터 수득된 생물학적 샘플로부터 수득된 변이체 SIGIRR 게놈 DNA, mRNA, 또는 mRNA로부터 수득된 cDNA의 존재를 검출하는 단계를 포함한다. 집단 내의 유전자 서열 및 이러한 유전자에 의해 암호화된 mRNA가 다형, 예컨대 단일 뉴클레오타이드 다형(single nucleotide polymorphism)(SNP)으로 인해 변할 수 있다고 이해된다. SIGIRR 게놈 DNA, mRNA, cDNA 및 폴리펩타이드에 대한 본원에 제공된 서열은 오직 예시적인 서열이고, 추가적인 SIGIRR 대립유전자를 포함하는 다른 이러한 서열이 또한 가능하다.
- [0180] 일부 실시형태에서, 검출 단계는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 서열분석하는 단계를 포함한다. 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 서열분석하되, 서열분석된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치를 포함하는 아미노산 서열을 암호화하는 것인 단계를 포함한다. 본원에 개시된 임의의 핵산 분자(예를 들어, 게놈 DNA, mRNA 또는 cDNA)는 서열분석될 수 있다. 일부 실시형태에서, 검출 단계는 전체 핵산 분자를 서열분석하는 단계를 포함한다.
- [0181] 일부 실시형태에서, 검출 단계는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계; 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산을 프로브를 포함하는 지지체와 접촉시키되, 프로브는 엄격한 조건하에 절두된 SIGIRR 단백질을 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키되, 증폭된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치를 포함하는 아미노산 서열을 암호화하는 것인 단계; 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산을 프로브를 포함하는 지지체와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 아스파르트산을 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다. 본원에 개시된 임의의 핵산 분자는 증폭될 수 있다. 예를 들어, 본원에 개시된 임의의 게놈 DNA, cDNA 또는 mRNA 분자는 증폭될 수 있다. 일부 실시형태에서, 핵산 분자는 mRNA이고, 상기 방법은 증폭 단계 전에 cDNA로 mRNA를 역전사시키는 단계를 추가로 포함한다.
- [0182] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키되, 프로브는 엄격한 조건하에 절두된 SIGIRR 단백질을 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는 것인 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 암호화하는 핵산 서열에 혼성화하는 핵산 서열을 포함하는 것인 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 핵산 분자는 인간 대상체로부터 수득된 세포 내에 존재하여서, 검출은 인시츄 혼성화(*in situ* hybridization: ISH)기법에 따른다.
- [0183] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 변경-특이적 프라이머와 접촉시키는 단계, 및 변경-특이적 PCR 기법을 이용하여 핵산 분자를 증폭시키는 단계를 포함한다. 변경-특이적 프라이머는 본원에 기재된 임의의 이러한 프라이머일 수 있고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 변이체 SIGIRR 단백질에 특이적일 수 있다.

- [0184] 본원에 개시된 방법에서 이용될 수 있는 다른 분석법은, 예를 들어 역전사 중합효소 연쇄 반응(RT-PCR) 또는 정량적 RT-PCR(qRT-PCR)을 포함한다. 본원에 개시된 방법에서 이용될 수 있는 더 다른 분석법은, 예를 들어 RNA 서열분석(RNA-Seq), 이어서 생물학적 샘플에서의 변이체 mRNA 또는 cDNA의 존재 및 분량의 검출을 포함한다.
- [0185] 본 개시내용은 또한 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는 인간 대상체를 확인하는 방법을 제공한다. 상기 방법은 일반적으로 대상체로부터 수득된 샘플에서 변이체 SIGIRR 단백질의 존재 또는 부재; 및/또는 변이체 SIGIRR 단백질을 암호화하는 본원에 기재된 임의의 핵산 분자의 존재 또는 부재를 검출하는 단계를 포함한다. 절두된 SIGIRR 단백질의 존재는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타낸다. 각각 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 생생시키고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 함유하는, 서열번호 2에 따른 9962번 위치에 상응하는 위치(예를 들어, 게놈 DNA)에서의, 또는 서열번호 4에 따른 557번 위치에 상응하는 위치(예를 들어, mRNA)에서의, 또는 서열번호 6에 따른 557번 위치에 상응하는 위치(예를 들어, cDNA)에서의 (아데닌의 결실로 인한) 구아닌의 존재는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환, 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타낸다. 상기 방법은 시험관내, 인시츄 또는 생체내 수행될 수 있다.
- [0186] 본 개시내용은 또한 조기-발병 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는 인간 대상체를 확인하는 방법을 제공하고, 상기 방법은 대상체로부터 수득된 샘플에서 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖고, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질; 및/또는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖고, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재 또는 부재를 검출하는 단계를 포함하고; 여기서 절두된 SIGIRR 단백질 및/또는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 존재는 대상체가 조기-발병 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 갖는다는 것을 나타낸다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치 중 어느 하나에서 야생형 SIGIRR 단백질과 비교되는 상이한 아미노산을 포함한다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 11의 아미노산 서열을 포함한다.
- [0187] 일부 실시형태에서, 샘플에서의 절두된 SIGIRR 단백질의 존재 또는 부재는 절두된 SIGIRR에 특이적인 항체로 검출된다. 일부 실시형태에서, 절두된 SIGIRR에 특이적인 항체는 i) 서열번호 9에 따른 186번 위치에 상응하는 위치에서의 세린; 또는 ii) 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 생생시키는 프레임쉬프트 돌연변이(frameshift mutation)로 인해 SIGIRR 단백질에서 생성된 에피토프에 특이적이다. 일부 실시형태에서, 검출은 절두된 SIGIRR에 특이적인 항체의 반응을 야생형 SIGIRR에 특이적인 항체의 반응과 비교하는 것을 추가로 포함한다. 일부 실시형태에서, 상기 샘플에서의 상기 절두된 SIGIRR 단백질의 존재 또는 부재는 효소-결합 면역흡착 분석법(enzyme-linked immunosorbent assay: ELISA)에 의해 검출된다. 일부 실시형태에서, 샘플에서의 상기 절두된 SIGIRR 단백질을 암호화하는 상기 핵산 분자의 존재 또는 부재는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 생생시키는 핵산 분자에서 프레임쉬프트 돌연변이가 있는지의 여부를 결정함으로써 검출된다. 일부 실시형태에서, 서열분석된 핵산 분자의 일부는 서열번호 9에 따른 186번 위치에 상응하는 위치를 암호화하는 코돈을 포함하는 복수의 위치를 포함한다.
- [0188] 상기 방법의 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 서열분석하는 단계를 포함한다. 상기 방법의 일부 실시형태에서, 검출 단계는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 서열분석하는 단계를 포함한다. 서열분석된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치를 포함하는 아미노산 서열을 암호화할 수 있다. 각각 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 생생시키고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 함유하는, 서열번호 2에 따른 9962번 위치에 상응하는 위치(예를 들어, 게놈 DNA)에서의, 또는 서열번호 4에 따른 557번 위치에 상응하는 위치(예를 들어, mRNA)에서의, 또는 서열번호 6에 따른 557번 위치에 상응하는 위치(예를 들어, cDNA)에서의 (아데닌의 결실로 인한) 구아닌의 존재. 검출 단계는 전체 SIGIRR 단백질을 암호화하는 핵산 분자를 서열분석하는 단계를 포함할 수 있다.
- [0189] 상기 방법의 일부 실시형태에서, 검출 단계는 절두된 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계, 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계, 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키되, 프로브는, 예를 들어 엄격한 조건하에를 포함하여, 절두된 SIGIRR 단백질을 암호화

는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 상기 방법의 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키는 단계, 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계, 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키되, 프로브는, 예를 들어 엄격한 조건하에를 포함하여, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 핵산 서열(또는 각각 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 생생시키고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 함유하는, 서열번호 2에 따른 9962번 위치에 상응하는 위치(예를 들어, 게놈 DNA)에서의, 또는 서열번호 4에 따른 557번 위치에 상응하는 위치(예를 들어, mRNA)에서의, 또는 서열번호 6에 따른 557번 위치에 상응하는 위치(예를 들어, cDNA)에서의, (아데닌의 결실로 인해) 구아닌을 갖는 핵산 서열)에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계, 및 검출 가능한 표지를 검출하는 단계를 포함한다. 증폭된 핵산 분자는 바람직하게는 서열번호 9에 따른 186번 위치에 상응하는 위치를 포함하는 아미노산 서열을 암호화한다. 핵산이 mRNA를 포함하는 경우, 상기 방법은 증폭 단계 전에 cDNA로 mRNA를 역전사시키는 단계를 추가로 포함할 수 있다. 일부 실시형태에서, 결정 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키는 단계 및 검출 가능한 표지를 검출하는 단계를 포함한다. 프로브는 바람직하게는, 예를 들어 엄격한 조건하에를 포함하여, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 포함하는 아미노산 서열을 암호화하는 핵산 서열(또는 각각 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두된 변이체 SIGIRR 단백질을 생생시키고, 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 함유하는, 서열번호 2에 따른 9962번 위치에 상응하는 위치(예를 들어, 게놈 DNA)에서의, 또는 서열번호 4에 따른 557번 위치에 상응하는 위치(예를 들어, mRNA)에서의, 또는 서열번호 6에 따른 557번 위치에 상응하는 위치(예를 들어, cDNA)에서의, (아데닌의 결실로 인해) 구아닌을 갖는 핵산 서열)에 특이적으로 혼성화하는 핵산 서열을 포함한다. 핵산 분자는 인간 대상체로부터 수득된 세포 내에 존재할 수 있다.

[0190] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키되, 증폭된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 아미노산을 암호화하는 코돈을 포함하는 것인 단계; 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다.

[0191] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다. 일부 실시형태에서, 인간 대상체는 18세보다 어리다. 일부 실시형태에서, 인간 대상체는 크론병 또는 크론병을 발생시킬 위험을 갖는 것으로 확인된다.

[0192] 본 개시내용은 또한, 인간 대상체로부터 수득된 절두된 SIGIRR 단백질을 검출하는 단계; 및 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는, 인간 대상체에서 염증성 장 질환 또는 조기-발병 염증성 장 질환을 진단하기 위한 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 검출하기 위한 방법을 제공한다. 일부 실시형태에서, 인간 대상체는 이러한 진단을 필요로 한다. 본 개시내용은 또한, 인간 대상체로부터 수득된 서열번호 9를 포함하는 단백질과 같은 변이체 SIGIRR 단백질을 검출하는 단계; 및 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는, 인간 대상체에서 염증성 장 질환 또는 조기-발병 염증성 장 질환을 진단하기 위한 또는 염증성 장 질환 또는 조기-발병 염증성 장 질환을 발생시킬 위험을 검출하기 위한 방법을 제공한다. 일부 실시형태에서, 인간 대상체는 이러한 진단을 필요로 한다. 일부 실시형태에서, 인간 대상체는 염증성 장 질환 또는 조기-발병 염증성 장 질환이 진단된 친척을 가질 수 있다.

[0193] 본 개시내용은 또한, 인간 대상체로부터 수득된 SIGIRR 단백질을 암호화하는 핵산 분자를 검출하되, SIGIRR 단백질은 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖고, 서열번호 9에 따른 215번 위치에 상응

하는 위치에서 절두되는 것인 단계; 및/또는 인간 대상체로부터 수득된 SIGIRR 단백질을 검출하되, SIGIRR 단백질은 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 갖고, 서열번호 9에 따른 215번 위치에 상응하는 위치에서 절두되는 것인 단계; 및 대상체가 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖는 경우 조기-발병 염증성 장 질환을 갖는 인간 대상체를 진단하거나, 또는 대상체가 조기-발병 염증성 장 질환의 하나 이상의 증상을 갖지 않는 경우, 조기-발병 염증성 장 질환에 대한 위험에 있는 것으로 인간 대상체를 진단하는 단계를 포함하는, 인간 대상체에서 조기-발병 염증성 장 질환을 진단하기 위한 또는 조기-발병 염증성 장 질환의 위험을 검출하기 위한 방법을 제공한다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 서열번호 9에 따른 186번 내지 209번 및 211번 내지 215번 위치에 상응하는 위치의 어느 하나에서 야생형 SIGIRR 단백질과 비교되는 상이한 아미노산을 포함한다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 서열번호 9에 따른 186번 내지 215번 위치에 상응하는 위치에서 서열번호 13의 아미노산 서열을 포함한다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 절두된 SIGIRR에 특이적인 항체로 검출된다. 일부 실시형태에서, 절두된 SIGIRR이 특이적인 항체는 i) 서열번호 9에 따른 186번 위치에 상응하는 위치에서의 세린; 또는 ii) 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 생성시키는 프레임쉬프트 돌연변이로 인해 SIGIRR 단백질에서 생성된 에피토프에 특이적이다. 일부 실시형태에서, 검출은 절두된 SIGIRR에 특이적인 항체의 반응을 야생형 SIGIRR에 특이적인 항체의 반응과 비교하는 것을 추가로 포함한다. 일부 실시형태에서, 절두된 SIGIRR 단백질은 효소-결합 면역흡착 분석법(ELISA)에 의해 검출된다. 일부 실시형태에서, 상기 절두된 SIGIRR 단백질을 암호화하는 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 생성시키는 상기 핵산 분자에서의 프레임쉬프트 돌연변이를 검출함으로써 검출된다. 일부 실시형태에서, 서열분석된 핵산 분자의 일부는 서열번호 9에 따른 186번 위치에 상응하는 위치를 암호화하는 코돈을 포함하는 복수의 위치를 포함한다.

[0194] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자의 적어도 일부를 증폭시키되, 증폭된 핵산 분자는 서열번호 9에 따른 186번 위치에 상응하는 위치에서 아미노산을 암호화하는 코돈을 포함하는 것인 단계; 증폭된 핵산 분자를 검출 가능한 표지로 표지하는 단계; 표지된 핵산 분자를 프로브를 포함하는 지지체와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다.

[0195] 일부 실시형태에서, 검출 단계는 SIGIRR 단백질을 암호화하는 핵산 분자를 검출 가능한 표지를 포함하는 프로브와 접촉시키되, 프로브는 엄격한 조건하에 서열번호 9에 따른 186번 위치에 상응하는 위치에서 세린을 암호화하는 코돈을 포함하는 핵산 서열에 특이적으로 혼성화하는 핵산 서열을 포함하는 것인 단계; 및 검출 가능한 표지를 검출하는 단계를 포함한다.

[0196] 일부 실시형태에서, 샘플이 수득되는 본원에 기재된 인간 대상체(예를 들어, 진단되고/되거나 치료되는 인간 대상체)는 성인이 아니다. 일부 실시형태에서, 샘플이 수득되는 본원에 기재된 인간 대상체는 18세이거나 이보다 어리다. 일부 실시형태에서, 샘플이 수득되는 본원에 기재된 인간 대상체는 15세이거나 이보다 어리다. 일부 실시형태에서, 샘플이 수득되는 본원에 기재된 인간 대상체는 13세이거나 이보다 어리다. 일부 실시형태에서, 샘플이 수득되는 본원에 기재된 인간 대상체는 10세이거나 이보다 어리다. 일부 실시형태에서, 6세이거나 이보다 어린 인간 대상체는 매우 조기-발병인 염증성 장 질환을 가질 수 있다. 일부 실시형태에서, 인간 대상체는 신생아 피사성 장염을 갖지 않는다.

[0197] 일부 실시형태에서, 본원에 기재된 임의의 방법은 염증성 장 질환 또는 조기-발병 염증성 장 질환을 치료하기에 효과적인 제제로 대상체를 치료하는 단계를 추가로 포함할 수 있다. 일부 실시형태에서, 상기 방법은 변경이 대상체에서 관찰되고, 대상체가 염증성 장 질환 또는 조기-발병 염증성 장 질환을 갖는 것으로 진단될 때 염증성 장 질환 또는 조기-발병 염증성 장 질환을 치료하기에 효과적인 제제로 대상체를 치료하는 단계를 추가로 포함한다.

[0198] 본 개시내용은 또한, 조기-발병 염증성 장 질환의 진단 또는 조기-발병 염증성 장 질환을 발생시킬 위험의 진단에서의, 본원에 개시된 임의의 변이체 SIGIRR 게놈 DNA, mRNA, cDNA, 폴리펩타이드 및 혼성화 핵산 분자의 용도를 제공한다.

[0199] 상기에 또는 하기에 인용된 모든 특허 문헌, 웹사이트, 다른 공보, 수탁 번호 등은, 각각의 개별 항목이 참고로 그렇게 인용된다고 구체적으로 및 개별적으로 표시된 것과 동일한 정도로, 모든 목적을 위해 그 전문이 참고로 인용된다. 서열의 상이한 버전이 다른 시간에서의 수탁 번호와 연관되는 경우, 본 출원의 유효 출원일에 수탁 번호와 연관된 버전이 의도된다. 유효 출원일은 이용 가능한 경우 실제 출원일 또는 수탁 번호에 관한 선행 출

원의 출원일 중 이른 것을 의미한다. 마찬가지로, 공보, 웹사이트 또는 기타의 상이한 버전이 다른 시간에 공개되는 경우, 그 출원의 유효 출원일에 가장 최근에 공개된 버전이 달리 표시되지 않는 한 의도된다. 본 개시내용의 임의의 특징, 단계, 요소, 실시형태 또는 양상은, 구체적으로 달리 표시되지 않는 한, 임의의 다른 특징, 단계, 요소, 실시형태 또는 양상과 조합되어 사용될 수 있다. 본 개시내용이 명확성 및 이해의 목적을 위해 예시 및 예에 의해 약간 자세히 기재되어 있지만, 소정의 변화 및 변형이 첨부된 청구범위의 범주 내에 실행될 수 있다는 것이 명확할 것이다.

[0200] 본원에서 인용된 뉴클레오타이드 및 아미노산 서열은 뉴클레오타이드 염기에 대한 표준 철자 약어 및 아미노산에 대한 1철자 코드를 이용하여 표시된다. 뉴클레오타이드 서열은 서열의 5' 말단에서 시작하여 3' 말단으로 정방향으로(즉, 각 선에서 왼쪽으로부터 오른쪽으로) 진행되는 표준 관례를 따른다. 각각의 뉴클레오타이드 서열의 오직 하나의 가닥이 표시되지만, 상보적 가닥은 디스플레이된 가닥을 약간 참조하여 포함되는 것으로 이해된다. 아미노산 서열은 서열의 아미노 말단에서 시작하여 카복시 말단으로 정방향으로(즉, 각 선에서 왼쪽으로부터 오른쪽으로) 진행되는 표준 관례를 따른다.

[0201] 하기 실시예는 더 자세히 실시형태를 기재하도록 제공된다. 이들은 청구된 실시형태를 제한하지 않고 예시하도록 의도된다.

[0202] 실시예

[0203] 하기 실시예는 본원에 청구된 화합물, 조성물, 물품, 장치 및/또는 방법이 어떻게 제조되고 평가되는지의 완전한 개시내용 및 설명을 당해 분야의 보통의 기술자에게 제공하도록 제시되고, 순수히 예시적인 것으로 의도되고, 본 발명자들이 이 발명을 어떻게 여기는지의 범주를 제한하도록 의도되지 않는다. 수(예를 들어, 양, 온도 등)와 관련하여 정확성을 보장하려고 노력이 되지만, 약간의 오류 및 편차가 차지할 수 있다. 달리 나타내지 않는 한, 부는 중량부이고, 온도는 °C 단위이거나 상온이고, 압력은 대기압이거나 대기압에 가깝다.

[0204] 실시예 1: 환자 모집 및 표현형분석

[0205] 전장 엑솜 서열분석 및 트리오-기반 변이체 분석은 13세 EO-IBD 환자, IBD 이환된 엄마 및 비이환된 아빠에서 수행되었다. 이 가족은 환자에서 EO-IBD의 유전 평가에 대해 확인되었다.

[0206] 실시예 2: 게놈 샘플

[0207] 게놈 DNA는 말초 혈액 샘플로부터 추출되고, 전장 엑솜 서열분석을 위해 리제네론 유전학 센터(Regeneron Genetics Center: RGC)로 옮겨지고, -80°C에서 자동화 바이오뱅크에 저장되었다. 형광 기반 정량화는 서열분석 목적을 위해 적절한 DNA 분량 및 품질을 보장하도록 수행된다.

[0208] 1μg의 DNA는 150개의 염기 쌍의 평균 단편 길이(Covaris LE220)로 전단되고, Kapa Biosystems로부터 커스텀 시약 키트로 엑솜 포획을 위해 제조되었다. 샘플은 NimbleGen SeqCap VCRome 2.1 또는 Integrated DNA Technologies xGen 엑솜 표적 설계를 이용하여 포획되었다. 샘플은 바코드화되고, 혼주되고, v4 화학물질을 갖는 Illumina HiSeq 2500에서 75 bp 짝짓기-말단 서열분석을 이용하여 서열분석에 대해 다중화되었다. 포획된 단편은 20x 또는 더 높은 커버리지에서 커버된 표적 염기의 최소 85%를 달성하도록 서열분석되었다. 서열분석 이후에, 데이터는 샘플-수준 데이터 제조 및 분석을 위한 표준 도구를 실행하기 위해 DNAnexus 및 AWS를 이용하는 RGC에서 개발된 클라우드 기반 파이프라인을 사용하여 처리되었다. 간단히, 서열 데이터는 Illumina의 CASAVA 소프트웨어를 이용하여 생성되고 탈다중화되었다. 서열 판독은 맵핑되고, BWA-mem을 이용하여 GRCh37/hg19 인간 게놈 기준 어셈블리로 정렬되었다. 정렬 후, 이중 판독은 마킹되고, Picard 도구를 사용하여 플래깅되고, 인델은 GATK를 이용하여 재정렬되어서 변이체 콜 품질을 개선하였다. SNP 및 INDEL 변이체 및 유전자형은 GATK의 HaplotypeCaller를 이용하여 호출되고, GATK로부터의 변이체 품질 점수 재보정(Variant Quality Score Recalibration: VQSR)은 전체 변이체 품질 점수를 주석표시되도록 적용되었다. 서열분석 및 데이터 품질 미터 통계학은 포획 성능, 정렬 성능 및 변이체 호출을 평가하도록 각각의 샘플에 대해 포획되었다.

[0209] 실시예 3: 게놈 데이터 분석

[0210] 최소 판독 깊이(10 초과), 유전자형 품질(30 초과) 및 대립유전자 균형(20% 초과)에 대한 표준 품질-관리 필터는 호출된 변이체에 적용되었다. 통과 변이체는 RGC 개발된 주석표시 및 분석 파이프라인을 이용하여 이의 가능한 기능 효과(동의, 비동의, 스플라이싱, 프레임쉬프트 또는 비프레임쉬프트 변이체이든)에 기초하여 분류되고 주석표시되었다. 가족 관련성은 PRIMUS(Staples *et al.*, Amer. J. Human Genet., 2014, 95, 553-564)를 이용하여 코호트에서의 관련성 및 관계를 추론하도록 유전자 데이터로부터의 혈통에 의한 동일성 유래 미터법 및 이

가족에 대한 보고된 가계와의 상호참조에 의해 동일성을 통해 검증되었다.

[0211] 보고된 가족 병력을 고려하여 상염색체 우성 유전 패턴 하에 후보 질환 유전자를 확인하기 위해 가계-기반 변이체 분석 및 분리를 수행하였다. 이환된 발단자와 이환된 모계 간에 공유되지만 비이환된 아빠는 공유되지 않는 변이체는 후속하여 공통 다형 및 고빈도, 아마도 양성 변이체를 배제하도록 집단 대조군 데이터베이스, 예컨대 dbSNP, 1000 Genomes Project, NHLBI Exome Sequencing Project, Exome Aggregation Consortium Database(ExAc) 및 내부 RGC 데이터베이스에서 이의 관찰된 빈도에 의해 주석화되고 여과되었다. 다수의 종 정렬(즉, GERP, PhastCons, PhyloP)에 기초한 보존 점수와 함께 변이체의 기능 효과의 생물정보학 예측을 위한 알고리즘, 예컨대, LRT, Poly-phen2, SIFT, CADD 및 Mutation Taster는 변이체의 주석화 과정의 일부로서 포함되고, 확인된 후보 변이체의 잠재적인 해로움에 대해 알려주도록 사용된다.

[0212] IBD 이환된 엄마로부터 유전된 13세 환자에서 조기-발병 염증성 장 질환으로 분리하는 SIGIRR 유전자(SIGIRR: c.557delA; p.K186fs^{*}31)에서 회귀한 절두 인텔 변이체가 확인되었다. 도 1(패널 a, b 및 c)을 참조하면, 크론병(CD)을 갖는 가족에서 우성 분리를 갖는 SIGIRR 유전자에서의 절두 변이체의 확인이 도시되어 있다. 패널 a는 모계 유전을 갖는 c.557delA/p.K186fs^{*}31에서의 SIGIRR에서의 절두 변이체를 기재하는 표를 보여주고; 변이체 부위는 중에 걸친 중성 보존 값을 갖고, 단백질 기능을 손상시키는 것으로 예측되고; 이 변이체는 ExAC 브라우저에서 0.000471의 교대 대립유전자 빈도를 갖는다. 패널 b는 이환된 EO-IBD 환자(Utah81427), 크론병 이환된 엄마(Utah81428) 및 비이환된 아빠(Utah81429)로부터의 가계를 보여준다; 채워진 기호는 CD 이환된 개체를 나타내고, 채워지지 않은 부호는 비이환된 개체를 나타내고; 원은 여성을 나타내고, 정사각형은 남성을 나타낸다. 패널 c는 비이환된 아빠에서 관찰되지 않지만 CD 이환된 Utah81427 및 CD 이환된 엄마에서 분리하는 확인된 SIGIRR 절두 변이체의 시각적인 확인을 보여준다.

[0213] 실시예 4: 검출

[0214] 대상체에서의 소정의 유전자 변이체의 존재는 대상체가 조기-발병 염증성 장 질환을 갖거나 발생시킬 위험이 증가된다는 것을 나타낼 수 있다. 혈액 샘플과 같은 샘플은 대상체로부터 획득될 수 있다. 핵산은 혼한 핵산 추출 키트를 사용하여 샘플로부터 분리될 수 있다. 대상체로부터 획득된 샘플로부터 핵산을 분리시킨 후, 핵산은 유전자 변이체가 존재하는지의 여부를 결정하도록 서열분석된다. 핵산의 서열은 제어 서열(야생형 서열)과 비교될 수 있다. 대상체로부터 획득된 샘플로부터 획득된 핵산과 제어 서열 사이의 차이의 발견은 유전자 변이체의 존재를 나타낸다. 이 단계는 상기 실시예에서 및 본 개시내용에 걸쳐 기재된 바대로 수행될 수 있다. 하나 이상의 유전자 변이체의 존재는 조기-발병 염증성 장 질환을 갖거나 발생시킬 대상체의 증가된 위험을 나타낸다.

[0215] 실시예 5: 연구

[0216] 재료 및 방법

[0217] 세포주 및 배양 조건:

[0218] University of Utah Health Sciences Center에서 IRB 허가를 받은 소아 IBD 환자로부터 엡스타인-바 바이러스-형질전환된 림프아구성 세포주(LCL)를 생성하였다. 미국 균주 협회(ATCC; 버지니아주 매나사스)로부터 건강한 LCL을 구입하였다. 10% 태아 소 혈청(Gibco 제품 번호: 10438026), 1X pen-strep(Gibco 제품 번호: 15140122) 및 1X L-글루타민(Gibco 제품 번호: 25030081)이 보충된 RPMI Medium 1640(Gibco 제품 번호: 12633)에서 세포를 배양하였다.

[0219] 자극 조건:

[0220] 건강한 대조군, SIGIRR LoF 환자로부터 생성된 LCL, 및 SIGIRR LoF를 보유하지 않는 4명의 EO IBD 환자로부터의 LCL을 72시간 동안 2mg/ml의 LPS(InvivoGen(캘리포니아주 샌 디에고)) 또는 16시간 동안 2mg/ml의 αIgM/αCD40(Affymetrix(캘리포니아주 산타 클라라))으로 자극하였다. 자극 이후에, 세포 배양 상청액은 수집되고, Mesoscale Discovery V-Plex Human Pro-Inflammatory Cytokine 패널(K15049D)을 사용하여 전염증성 사이토카인의 분비를 정량화하도록 사용되고, Mesoscale Discovery QuickPlex SQ 120을 사용하여 정량화되고, 후속하여 RNA-서열분석으로 보완되었다.

[0221] 통계:

[0222] 유의도는 2방향 ANOVA 시험을 이용하여 결정되고, S.E.M.은 GraphPad PRISM(캘리포니아주 라 줄라)을 이용하여 3개의 독립 실험으로부터의 결과를 이용하여 계산되었다.

[0223]

결과:

[0224]

도 2를 참조하면, SIGIRR LoF 환자로부터 생성된 LCL은 건강한 LCL 또는 SIGIRR LoF를 보유하지 않는 EO IBD 환자로부터의 LCL보다 더 많은 IFN- γ , IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70 및 TNF- α 를 생성하였다. 추가로, SIGIRR LoF 환자로부터 생성된 비자극된 LCL은 건강한 LCL 또는 EO IBD 환자로부터의 일부 LCL보다 상승된 수준의 IFN- γ , IL-1 β , IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70 및 TNF- α 를 분비하여서, SIGIRR LoF LCL이 구성적으로 활성이고 LPS 자극에 불응성이라는 것을 나타낸다. 도 2를 참조하면, 청색 막대는 비자극된 세포로부터 단리된 상청액을 나타내고; 적색 막대는 LPS 자극된 세포로부터 단리된 상청액을 나타내고; "*"는 2방향 Anova에 의해 $p < 0.05$ 를 나타내고; 오차 막대는 3개의 독립 실험으로부터의 S.E.M.을 나타낸다.

[0225]

도 3을 참조하면, α IgM/ α CD40 자극 이후에, SIGIRR LoF 환자로부터 생성된 LCL이 건강한 LCL 또는 SIGIRR LoF를 보유하지 않는 EO IBD 환자 LCL로부터의 일부 LCL보다 더 많은 IFN- γ , IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70 및 TNF- α 를 생성하였다. 추가로, SIGIRR LoF 환자로부터 생성된 비자극된 LCL은 건강한 LCL 또는 EO IBD 환자로부터의 일부 LCL보다 더 상승된 수준의 IFN- γ , IL-2, IL-4, IL-6, IL-8, IL-10 및 TNF- α 를 분비하여서, SIGIRR LoF LCL이 구성적으로 활성이고 항-IgM/항-CD40 자극에 불응성이라는 것을 나타낸다. 도 3을 참조하면, 청색 막대는 비자극된 세포로부터 단리된 상청액을 나타내고; 적색 막대는 항-IgM/항-CD40-자극된 세포로부터 단리된 상청액을 나타내고; "*"는 2방향 Anova로 $p < 0.05$ 를 나타내고; 오차 막대는 3개의 독립 실험으로부터의 S.E.M.을 나타낸다.

[0226]

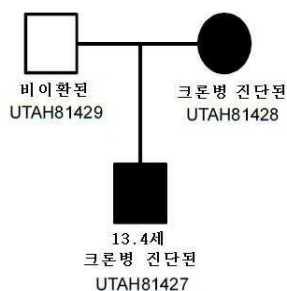
비자극된 세포에서, 건강한 대조군으로부터 생성된 LCL 및 SIGIRR 기능 소실(LoF) 변이체를 보유하지 않는 EO-IBD 환자로부터 생성된 LCL 둘 다에 비해 절두된 (c.557delA; p.K186fs*31) SIGIRR EO-IBD 환자에서 IL-1 β , IL-8 및 IL-6을 포함하는 면역 조절제의 상향조절이 관찰되었다. 이 관찰은 SIGIRR LoF 변이체를 보유하는 IBD 환자에서의 고유한 염증성 서명을 지지한다. 추가로, SIGIRR EO-IBD 환자 유래된 LCL은 IL-1 β 또는 TLR 자극에 의한 자극에 불응성인 것으로 관찰되어서, SIGIRR의 부재 하에 이 전염증성 경로의 구성적 활성을 지지한다.

도면

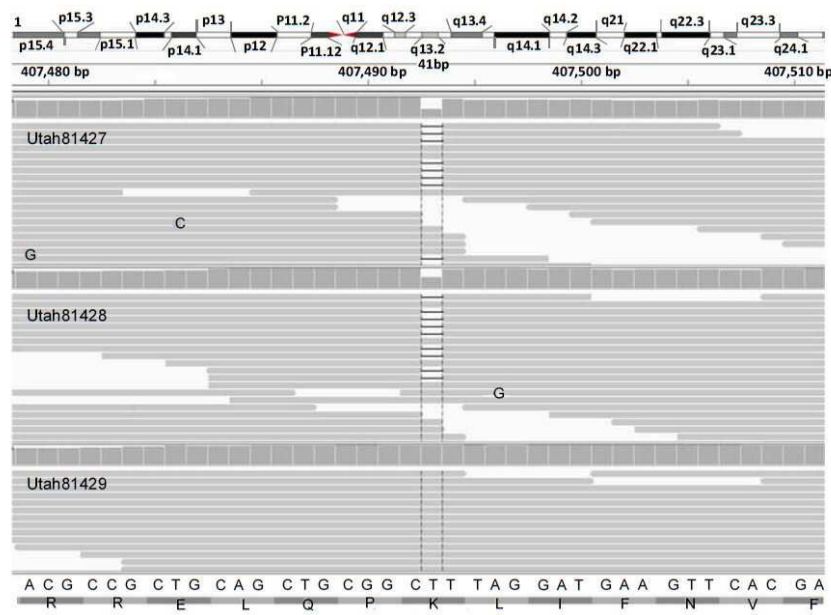
도면1a

유전자	NT 변화	AA 변화	유전	보존	예측	대립유전자 빈도 ExAC
SIGIRR	c.557delA	p.K186fs*31	모계	자연	손상	0.000471

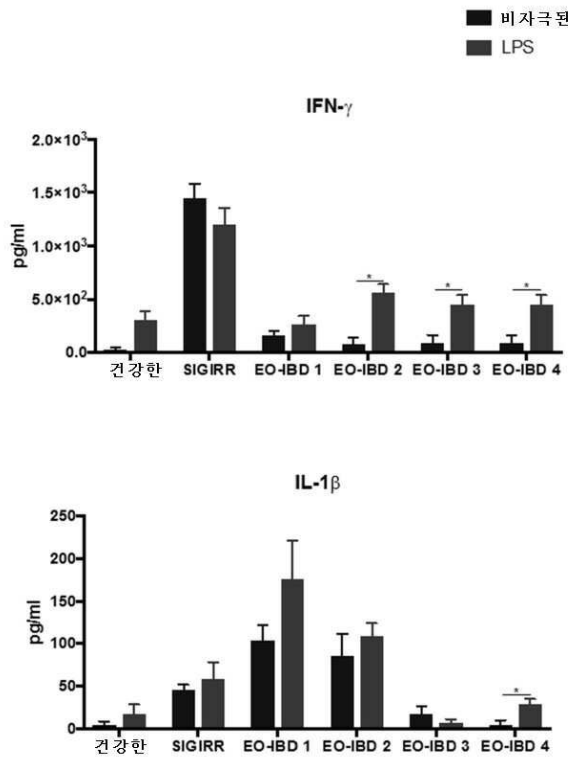
도면1b



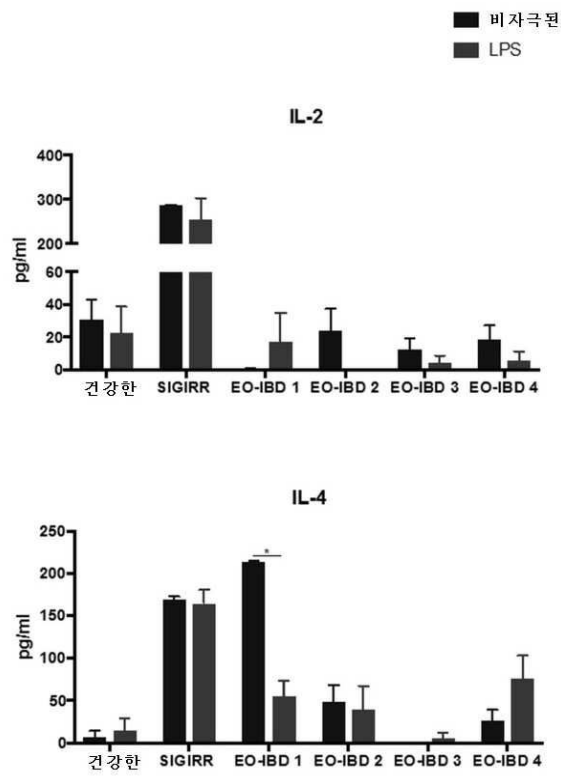
도면1c



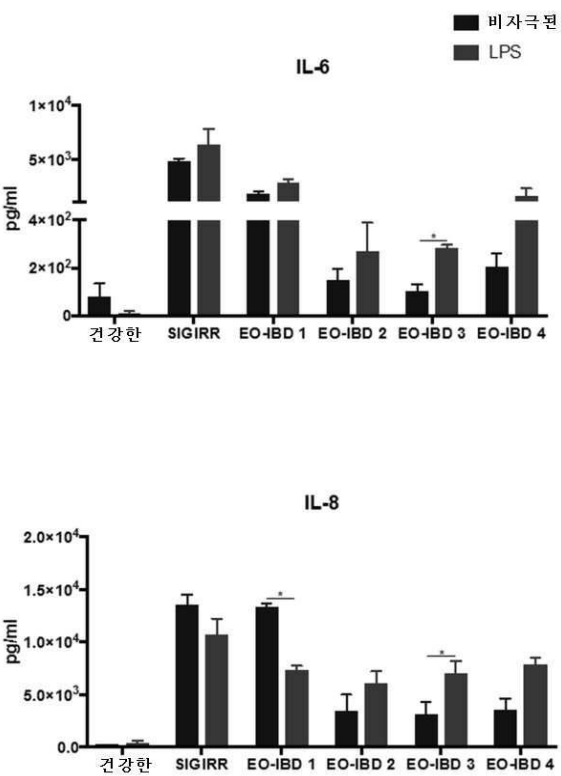
도면2a



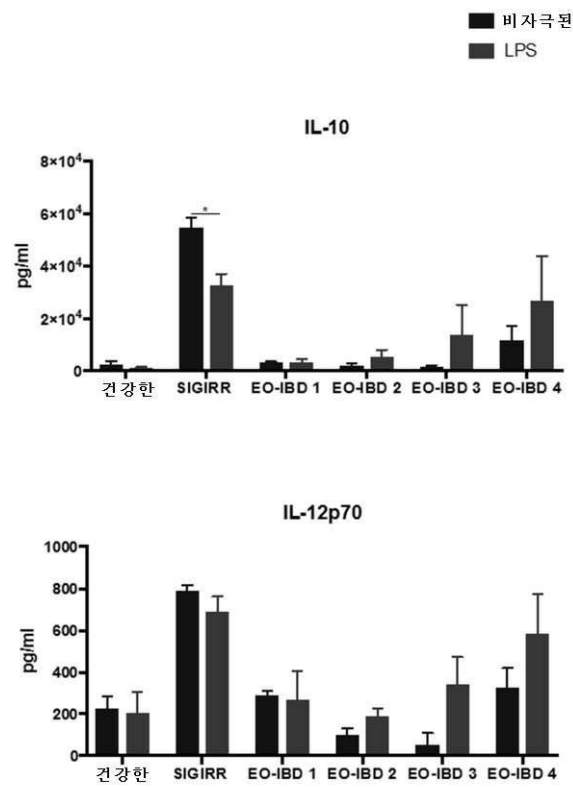
도면2b



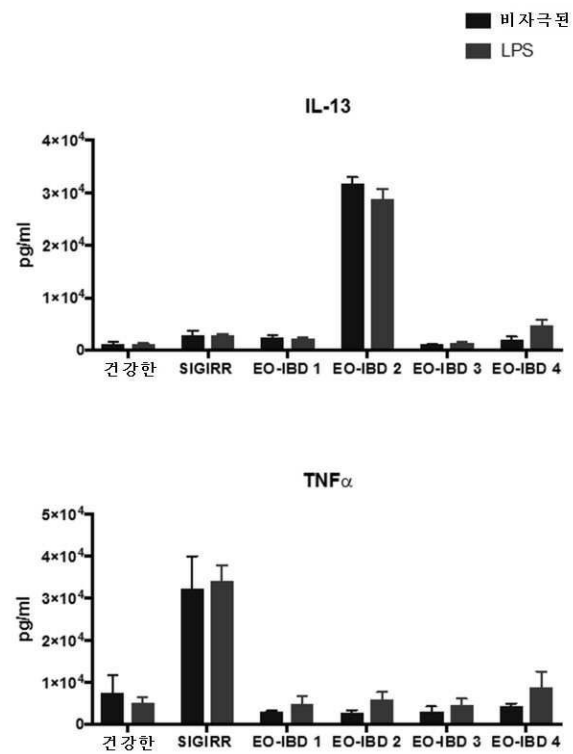
도면2c



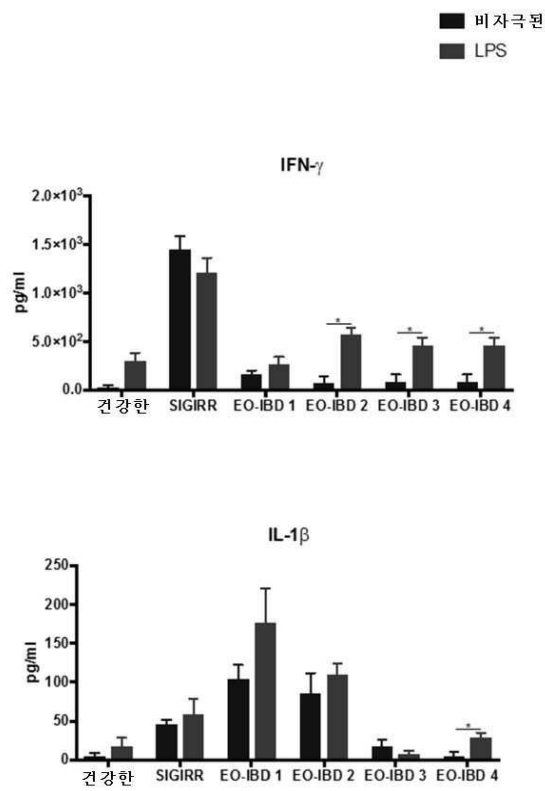
도면2d



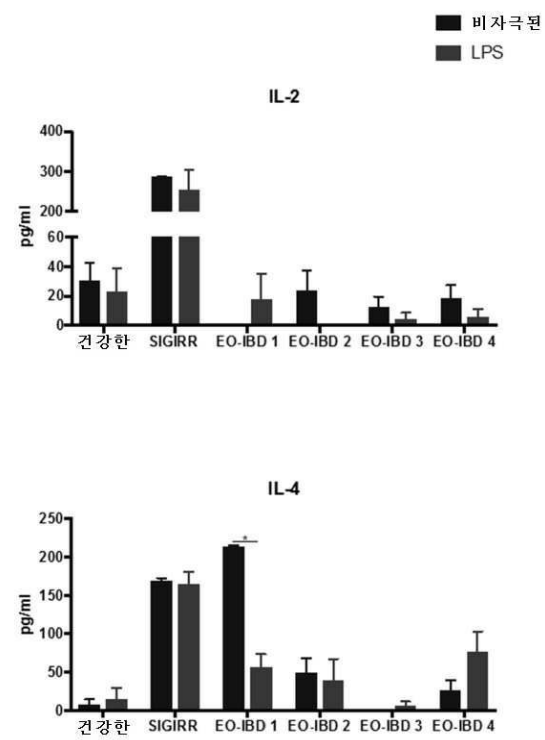
도면2e



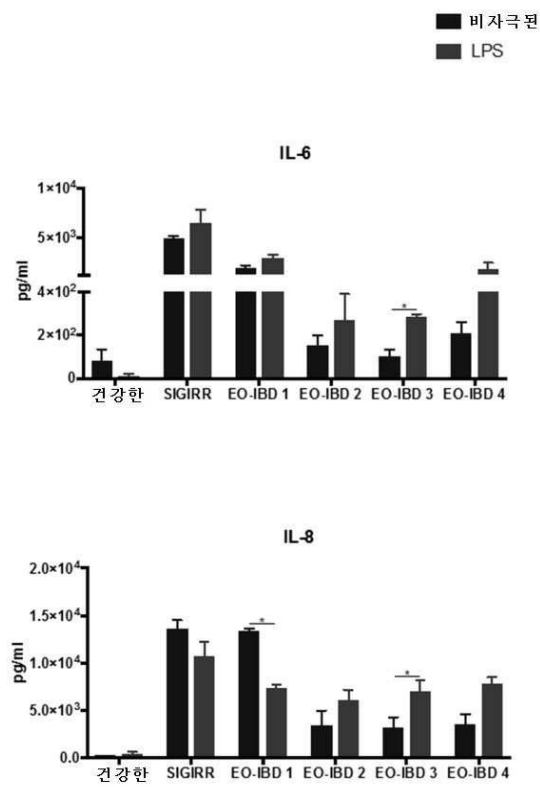
도면3a



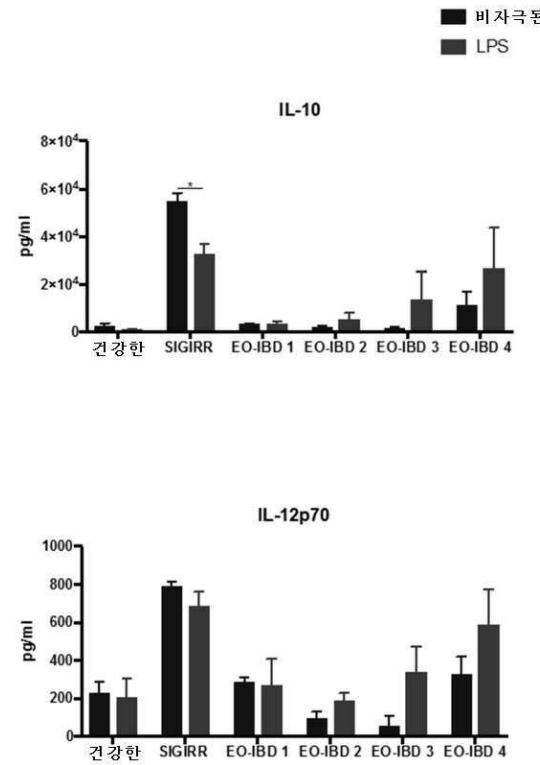
도면3b



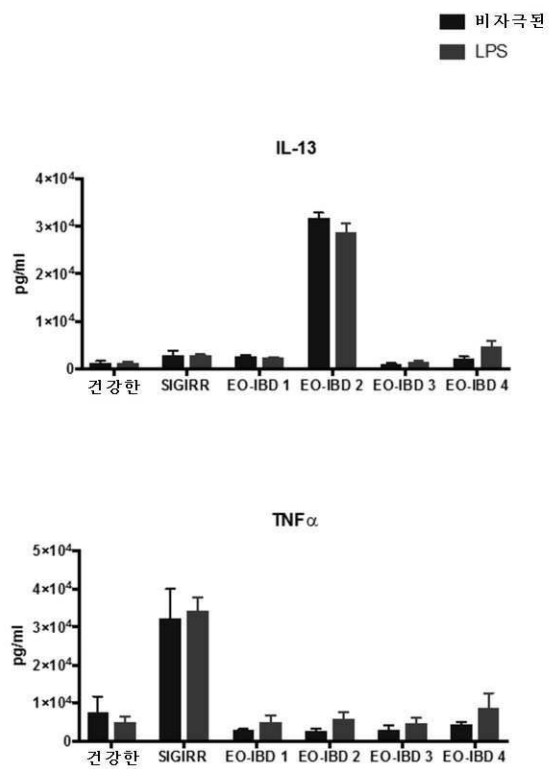
도면3c



도면3d



도면3e



서열목록

SEQUENCE LISTING

<110> Regeneron Pharmaceuticals, Inc.

<120> Single Immunoglobulin Interleukin-1 Receptor Related (SIGIRR)
Variants And Uses Thereof

<130> 189238.00602

<150> 62/554,857

<151> 2017-09-06

<160> 16

<170> PatentIn version 3.5

<210> 1

<211> 11739

<212> DNA

<213> Homo sapien

<220><223> genomic wild type SIGIRR

<400> 1

actgcggggc ggggccgcac ctgctgccgc tgccggggct acgccgccgc cggggctgtt

60

gccgctgcgc acctggctca ggtgagctgc cccgcccccg cccggcgcga gccccaggtc 120

ctggcagcag gtgagtggcg gtcgccatca gacccgcgaa ggtgagtgga ggtcgcagtc 180

ctcccagagc cgctgcaggc ctggagtcgg gggccggggc gctgcttgtc cctgggcgcg 240

gtctctcagg aaatggctcg gggccgggga ggttcccgcg tcgcccggaa gcccagctc 300

gcccgcgtcc tggcccttcc cagggtctgg ggaaggagg aaagtggcg aagggggcgg 360

gggccaccct ttctctgcc cgacgtggc agcggctttt tggtgaccct cgcgcagaac 420

agagcggggc taggagcgcc gaggcgcgca gggatccac ctcggtcgcc caagcctgga 480

tcccagctcg aggtttctg aggcgccgag ggcggaccaa ggtcaggct ggggcccgag 540

ctgccccggc gccgggtgga gcttccgcca tccgcgcgc tctgccccgg ccgagcgccc 600

cgagccgcac ccccgctgcc gcctcccgcg cgcgtcctcg ttctccgcc tcctcacttg 660

gagggcgccc agggcgagg tgggcgcggg gtaggggagc gaaccagagc ccctcgccct 720

gcggggccgc agccttttac ccacgttct gcttcaggac aggcgctggg gggccagcgc 780

gcacggggag cggagggggc gcgcgtccga gctgagcggg agtcgcccgc aggggtgctgg 840

gtgctgagcg tcgaaccac cgggcagggg cggccgcgga ccgaggacgc tggctgggtc 900

tgaggcgggg aggagctggc gctttctaag aacctcaagg tgcccagcag agtcggtgg 960

gcagtgcgag gcgaccagg ccaggcttgg actgagattc tgggcgcctc ggaagagctg 1020

tgcgggcgag ggtcggagac tttgtgggtg ttcatgtca ccctggacgg tggacctggg 1080

ggctcggggg tgcaggacgt tagcagagag aggggaaggag agggtgcaa gccagccctg 1140

agagatggtg acgtgagac cgggtgcaca gagcagccgc cgagccgggg ctacacagtt 1200

ggcggctgac ccaggggcct gtgctggagg ccctgtcct gcttctcg ggggtgccct 1260

gggccctgct gtgcctgctt cctgacctgc agtgtggga gaaccttctt gaggggcttc 1320

caaggaccca cttaaggag gggctggtac caggctctgt gggggaagg gtgctgtctt 1380

tgggacctg agcccaagg ccaccatcac ccaaggctga gtaccgggac aaaggaaggt 1440

cctagcagg ctgggactca gtcagacaag tacctgtaga aagcctgctg ctgagagggt 1500

gctcagccag ggccgggagg ctggatggca gcctccaca atgaagcct gccccatgt 1560

ggctgagggc agggccagg tccaaaggag acccttgtct gattctgtgg gcaataccag 1620

ccaccggtca ctgggtccc agtcccaggt caacctgggg agtgtcagtt tggtgacaac 1680

aggagatgcc atctgccct catctgctct gagccaggag cctcgccac ctgtgtttgc 1740

tccttatttg gaaaagggg cgctgccctc ccacccagc ctgaaggga ggtccacaag 1800

atgccatagt cccttgtaac tgctcaccca tcgggggagc tatgtctggg agagggacgg 1860

ggagagggcg gaagagttct gctggagccc cattgttgac cgggagtttc tgaaggccct 1920
 gttccccggg gtcagaggtc aaagacacat gggggagcaa accggcccct tggtcctga 1980
 gtgtgagctc tagtagcgtc tggaaggcat gccgcgcggt ccccggtcct ccttgtatgt 2040
 gtgcagggtg gccctggcgg gggctctcctg gctggaggac ccccatgca ttacccttgg 2100
 aacctgcccc tttagcacc cccacccca caacacacac gcacacacac aaacgcacgc 2160
 acacgcacat actgcacaca cattgcacac acgcagacac actgcacaca cacacacaca 2220

 cacacacaca cacacacaca cacacacgct gcacagagga aagtgccgtg cctgggtggg 2280
 ttggggcggg ctccagagga aggagccgct tgactcagag ctacaggaag agccggggca 2340
 ggcgggggtga gaaccaccaa ctgcccgtcc tcccggccgc ctagggaagg ggtgggtgcc 2400
 tgggtgggtg ttagggtag gagaggatgt gactgcaagg ccaggaagct acagggaag 2460
 ttctgatcaa aatacacaaa gacacaagcc aggtccccc cgcgcttggc gatgcatcca 2520
 tagcagcctt ggtgctgtgg acactccctg gggcccagcg aaggggagag ttgtctcca 2580
 aaggcgcacc aatgaccaac atttgcctcc cggaggaaag aactggaacc aggtgactgc 2640

 ctccccgct gctctgagtg agcctgtgtc agggcctggc tgagaggagg cccacggcgc 2700
 ccaagaccca tatgcatgtg tgtgcatgtg tgtgtgagac agagcgaggg gcagcaggag 2760
 gctgtgtcgc atcagcgccc aagggtgggt gtcagggtgca gtgtgtgtgc tgtgggtgc 2820
 ctgtggtcct aggtctgtg acaagcgttg caatgggtgt tgtatatact caacagtgt 2880
 gtgaggggtc tgcagcctaa gctgtgtgca gtggtaatgg gaggtgtgtg tggtaggaag 2940
 tgtgtgtgtg aggaggtgca agtgtgcaca cgggctgaga ctgaagctct gtgcagggtc 3000
 acatgggcca gagggctgtg tgtatgtgat gtgtgcaggg gaggaggggc agaggagagg 3060

 gtgcagggga agttaggag gggaggggtg acaaggaaag gtgtgcaggg gagagggtgc 3120
 agggaaaggt ggggagggga gaggtgacag gggaaggtag gcagaggaga ggtgacagg 3180
 gagagggtgc aggggaaggt gtgcaggga aggtggggag gggagagggt gcagggaag 3240
 gtaggcagag gagagggtgc aggggaaggt gtgcaggga gaggtgacag ggaaggttg 3300
 ggaggggaga ggtgacagg gaaggtagc agaggagagg gtgcaggga aggtgtgacag 3360
 ggaaaggttg ggaggagag ggtgacaggg aaggtaggca gagagagag tgcagggaag 3420
 ggtgtgcagg ggagagggtg cagggaaggt tggggagggg agagggtgca ggggaaggt 3480

 ggcagacgag agtacaggga aaggtaggga ggagagaagg tgcaggagga ggcaggcaga 3540
 agagagggtg cagggaaggt tggggagggg agagggtgca gggggaggca ggcagaggag 3600
 agggttcagg ggagagggtg cagggaaggt tgggcagggg aaggtgggga gtggagagcg 3660
 tgcaggggga ggcaggcaga ggagagggtg cagggaaggt tgcgtagggg agaggagca 3720

ggggagaggg tgcgggaaga ggggtgcaggg gagggtttgc agggagaggg tgcaggggag 3780
 aggggtcaag gagaggcagg cagcggagag ggtgcagagg agggtttgca ggggagggca 3840
 cctgctttct tgcgtgtctc gtatcctggc agtctgccct tgacttccag aaggctctga 3900

 ggctgctttt ggagggagga ctttgaggtc tgagcttggg gcggtcacgg aagcactgca 3960
 ggagatgccca ggggttctgc tgggctggcg gtgtgacctc ccgggcttgc tgtgaggggt 4020
 tggtttcttg ggaagtgtgt gctttaagga cctcagtggt gtggttcccc tgggtcaagg 4080
 actttccttg gagtgtcctc caccgcaaag gacttcttca aggacccctt ggccatcccc 4140
 catgggtgtg tgcctcatgc tgtatcggtt cccgaggaaa ctgccaaaca gtcaaacctg 4200
 gggcctctgg ctgtgcttca cccatggaga gggaataggt ggcaggctaa gaggactgtc 4260
 cctgccactg cctctggggc cttagcgggg aggaatcctg agggagctgg ggtgcagaaa 4320

 cttagccttg atccccacat cccacttcc actgtgtccc tgcagcaggg aggggagtag 4380
 acagtcatgt ccgtttggat aaacaggggtg accctgtggc atccagttag gagaggcccc 4440
 tggggccctg ggctgccaaag gaagaggctc tgggcttctc cctgcagcca cctcactgag 4500
 gagcttttggg agacttgagg tgtgtgtggc tgcgtctacg ggctgggaaa ggacatggga 4560
 cccgagacct ccatccccac cgggcgctgt tcacctccg gcctcacaac cgcagggcga 4620
 ggctaagctg agggttcccg aggtgtttcc cgaggtgttt cgaggtgtct gtgtcacgt 4680
 gcaactctga ctgtaaccag acgtgggcca gagctgggca cccccctcg actgcatcca 4740

 gacatggtta gagctgggca ccccccccc ccgactgtat ccagacatgg tcagagctga 4800
 gcacccctcc ccgactgtat ccagacatgt tcagagctgg gcacccccc ccgactgtat 4860
 ccagacatgt tcagagctgg gcacccccc ccgactgtat ccagacatgg tcagagctgg 4920
 gcacccccc ccgactgtat ccagacatgg tcagagctgg gcacccccg actgcatcca 4980
 gacatggtca gagctgagca ccccccccg ccgactgta tccagacatg gtcagagctg 5040
 ggcacccccc gactgcatcc agacatggtc agagctgggc accccctcc ccgactgcat 5100
 tcagacatgg tcagagctgg gcacccccgc cgactgcac cagacatggt cagagtggg 5160

 ccccccccc ccgactgtat ccagacatgg tcagagctgg gcacccccg actgcatcca 5220
 gacatggtca gagctgggca cccctccccg actgtatcca gacatggtca gagctgggca 5280
 ccccccccg actgtatcca gacatggtca gagctgggca cctcccccg actgtatcca 5340
 gacatggtca gagctgggca cctccccga ctgcatccag acatggtcag agctgggcac 5400
 ccccccgac tgcattcaga catggtcaga gctgggcacc cccccccc actgtatcca 5460
 gacatggtca gagctgggca cccccctccc cgactgcac cagacatggt cagagctgag 5520

cacccccccc ccccgactgt atccagacat ggtcagagct gggcaccccc ctccccgact 5580

glatccagac atggtcagag ctgagccccc ccccccccg actgtatcca gacatggtca 5640

gagctgggca ccccccgact gcatccagac atggtcagag ctgagcacc ccccccgact 5700

gcatccagac atggtcagag ctgggcaccc cccccccca ccgactgtat ccagacatgg 5760

tcagagctgg gcaaccccca ccgactgtat ccagacatgg tcagagctgg gcaccccccg 5820

actgcatcca gacatggtca gagctgggca ccccccccg actgtatcca gacatggtca 5880

gagctgggca cccccctccc cgactgcatc cagacatggt cagagctggg cccccctc 5940

cccgactgca tccagacatg gtcagagctg ggcacccctc cccgactgta tccagacatg 6000

gtcagagctg ggcaccccc tccccgactg tatccagaca tggtcagagc tgggcacccc 6060

cgccgactgt atccagacat ggtcagagct gggcaccccc cgactgcatc cagacatggt 6120

cagagctggg cccccctccc cgactgtatc cagacatggt cagagctggg ccccccccc 6180

ccccgactgt atccagacat ggtcagagct gagcacccct ccccgactgt atccagacat 6240

gttcagagct gggcaccccc cccccgact gtatccagac atggtcagag ctgggcaccc 6300

ccctccccga ctgcatccag acatggtcag agctgagcac ccccccgcc cgactgcatc 6360

cagacatggt cagagctggg cccccccga ctgcatccag acatggtcag agctgggcac 6420

ccccctccc gactgcatc agacatggtc agagctgggc accccctcc ccgactgcat 6480

ccagacatgg tcagagctgg gcacccccgc cgactgtatc cagacatggt cagagctggg 6540

cacccccctc cccgactgta tccagacatg gtcagagctg ggcacccccg ccgactgtat 6600

ccagacatgg tcagagctgg gcacccccct ccccgactgc atccagacat gttcagagct 6660

gggcaccccc ccccgactgt atccagacat ggtcagagct gggcaccccc cgactgcatc 6720

cagacatggt cagagctggg caacccccct cctgactgca tccagacatg gtcagagctg 6780

ggcaccccc cgggctgtat ccagacatgg tcagagctgg gcaccccccc cccccgactg 6840

catccagaca tggtcagagc taagcaaagc cccccccga ctgtatccag acatggtcag 6900

agctgggcac cccacccca actgcatcca gacatggtca gagctgagca cccccaccg 6960

actgtatcca gacatggtca gagctgggca tctcccatgg ttaggttac atcactcaca 7020

cccctgcccc aaggccaggc tgtgtctctg gcactgcagc acccctgca gtggcagctc 7080

tgttgaaggg gacagccagg agtgaggcca actgccatgt cgggagacag ggtccacct 7140

agtgttcgtc cactgcaccc tggggccgtc atggtggccc caggcccggt ggaggatgc 7200

ctggacttga tggccccctc ccgtggctct gggtgcgtg aagtgggtcc cattctggcc 7260

acgtgtgagg ggctcaaaga gacaccctgc tggcctccga gaactctgcc agcatccatg 7320
cggctgctca gggctgcacg tggtcagggc cagtgtctgc tgtggccggc agttactggt 7380
taactcatcc tgttctctgt taacttgttc cctggggatg gcggcagcct ctgacctgtc 7440
caggtgccct gtccagctga ctgcaaggac agagaggagt cctgcccagc tcttgatca 7500
gtctgtctggc cgaggagccc ggtggagcca ggggtgacct tggagcccag cctgcccga 7560
ggaggccccc gctcagagcc atgccaggta ggaggtcag agaggatgg gcttgaattc 7620
cccactccca cctctgccc ctgaggtgg ctccacattg tgcatcctgc tcttatccca 7680

ggcactctgc ctggccacat gccgggggta caaaggcccc tgctacccc tcccactggg 7740
cccttggccc tgaaccagg gcgagtggg aggatgagcc ctgcacctc tctgcccga 7800
gaccaggctc atgagggtca gtaaagccca ggaatgctag gagtgggcac tctggggcc 7860
cacagcaggg tgtggccgat gtgagctggc aggagacct accctactga gaagatggtg 7920
gtttgacctc ccctagggc tcatcaccct gatgtaaaga tgaacaagt caggcctcag 7980
gtggtgggga gggcttgct gtggccagga gcagagggcc ccaagacccc caaggcagga 8040
atgaggccaa ctgctccagg ctcccagctg tgcccatcca ggtggaaggt gctggcacac 8100

tgggcatctg ggtcccaaaa cccactggtg cccacactct ctgcagattc cccagcagag 8160
gcaggccagg ctgagggtg cagtccccg cccctctat gagacaggca tcatggtgtc 8220
accaggggtg ggcaacaggt gcatctggga aatggggggc tggactcaca tctgggtgg 8280
acaagcagcc agatggggtg tgagcgtggg cctggggcag ctcaactgtg gccactgtc 8340
tctgtccct cagcacaggg tgggcggaac acccaaagcc tggccactac tggggaaga 8400
cgtgccaaagt gctaaagctg tcccatggg actcgtggac agacatggtg tgactgcaa 8460
gccacacaga ggccttattt tttctgcac cacggaccac tgtggaggtg gtggggccag 8520

gcacctggc acagctgacc cactgtgtcc tctctctca ggtgtctgtg ataggcccc 8580
tgacttctc tcccgtctg aagaccaggt gctgaggcct gccttggga gctcagtggc 8640
tctgaactgc acggcttggg tagtctctgg gcccactgc tccctgctt cagtccagt 8700
gctgaaagac gggcttccat tgggaattgg gggccactac agcctccag agtactctg 8760
glaagagacc cgggtatcca gggtcagagt gacctctcag acatgccaag ggcagtggc 8820
tccctgacct atgcctgggt ctgtaaccac ctggccagac atgtcacgag cagagtgcct 8880
ctgtggacct agacggaggg tcagtaattg caggacaga ccttgggtc tgcatggccc 8940

ctgtaggcac acacctgta actgcccagg gctacgggtg tcagcaaagt ctagggtctg 9000
aatgaacacc gaccagctgg ccaagtctgt gatcaaagga cctccaggcc ccagaccag 9060
gactgggacg gctgttctga aacaatctg catcccctcc aggttcaggt cacttctctt 9120

ggcttaaggc ccaagaacca cccaggaggc ccaattctgc agcctccctg ggtgaaccca 9180
 gacaggtggg gttcgggggt gtccactggt atccccagtc gacctgacc ctggcttggt 9240
 atccccaggg tcaaggccaa cctgtcagag gtgcttgtgt ccagtgtcct gggggtaaac 9300
 gtgaccagca ctgaagtcta tggggccttc acctgtccta tccagaacat cagcttctcc 9360

 tccttcactc ttcagagagc tggatgatggg ggttcccca ggtggagggt agaaggggac 9420
 ctagacctag tgaggccag ggatcttgag gcttgggggc tgctggaggg gcagcggctc 9480
 aggcaaccca tccgcaggcc ctacaagcca cgtggctgcg gtgctggcct ccctcctggt 9540
 cctgtcggcc ctgctgtggt ccgccctgct ctatgtcaag tgccgtctca acgtgtctgt 9600
 ctggtaccag gacgcgtatg gggagggtgga gataaacggt gcgtggggcc cgcgtgggcg 9660
 cgaggagggg tcgccacggc ctccctgaga gtgcgctgct gagctcggct ttggaggcgt 9720
 gcggcgggga ggggttagcc cccgggtgct ctgtgcggcc cggctgggggt taaagtccgg 9780

 ggccgggtctg cccctgtgca cgtgggagat ggttaggggt gttggggacc ccgagggtg 9840
 gtgtgaggcc tcgggagcca tcgggggtgac tgccgccct cgcagacgg gaagctctac 9900
 gacgcctacg tctcctacag cgactgcccc gaggaccgca agttcgtgaa cttcctccta 9960
 aagccgcagc tggagcggcg tcggggctac aagctcttcc tggacgaccg cgacctcctg 10020
 ccgcgcgtg gtatccggg cccacccccg tgcgccccc acccgaggagg cccgccccgc 10080
 cccgccccat gtccgtccc acccggggccc ccacccccac catcgagctc cgcccccatc 10140
 cccgccccacc cgaggccccg cctcaaaaac ccgccacccc cgcaggcccc gccctccct 10200

 tagagctctg cgtccggcac cgccccgag cctccggccc cgcctcccc cgccccgcc 10260
 cctgcgccgc cgacgccgc cctccgcag agccctccgc cgacctcttg gtgaacctga 10320
 gccgtgccg acgcctcctc gtggtgcttt cggacgcctt cctgagccgg gcctggtgca 10380
 gccacagctt ccggtgggtc ccgcgcgggg ttgggtgggc cccagcgtag caccacccc 10440
 cctgacggtc ccgccccga gggagggcct gtcccggtg ctggagctca cccgcagacc 10500
 catcttcac accctcagg gccagaggcg cgaccccgcg caccggcgcc tccgctgct 10560
 gcgccagcac cgccacctgg tgaccttct gtctggagg cccggctccg tgggtcggag 10620

 caggcgcggg agggtcggg gctagcggcg ggttagagat gggcgggtgcc cgggtccag 10680
 gctgggacc ctccgtggg agctctgcgg caccacgctt tgtgaatggg ccctgggggg 10740
 aggttccgct gcctggggcc ccgatcggg gagccgcct tgaggcccc ggagccacgg 10800
 aatagctgtc gcaggcggtg gaaccctgg gcagccgag gtgtgctctt gggggccagg 10860
 acgccagggg cttccgaggt gttcacacct gcaaacgcc ccgacctggc cccagactc 10920
 cttcctccga tttttgaaa gaagtgcagc tggcgctgcc gcggaaggtg cagtacaggc 10980

ctgtggaagg agacccccag acgcagctgc aggacgacaa ggaccccatg ctgattcttc 11040

gaggccgagt ccttgagggc cgggccctgg actcagaggt ggaccgagc cctgagggcg 11100

acctgggtat gcccgccag cccactccc caactggaga agctcagcac agggcggagt 11160

gggggcaggc acagggcaca gggcctggag gggctccagg tgttgaggac tcttcccggc 11220

accgggagcc cctgcacggc ctctgccctg gaggtgctcg gccctcggtc tgcctgggaa 11280

cttctgggc ctcacaggcc atcacagcag ggggtgagca ggggcagccc ctggcagtgg 11340

gtctgggcca aggctgtggg tggccacctc aggcgtctcg gtctccccc cccaggtgtc 11400

cgggggcctg tctttgaga gccatcagct ccaccgcaca ccagtggggt ctgctggga 11460

gagagccgga gcagcgaagt ggacgtctcg gatctcggct cgcgaaacta cagtggccgc 11520

acagacttct actgcctggt gtccaaggat gatatgtagc tcccaccca gagtgcagga 11580

tcatagggac agcgggggcc agggcagcgg cgtcgtcct ctgctcaaca ggaccacaac 11640

ccctgccagc agccctggga ccctgccagc agccctggga aaaggctgtg gcctcagggc 11700

gcctcccagt gccagaaaat aaagtccttt tggattctg 11739

<210> 2

<211> 11738

<212> DNA

<213> Homo sapien

<220><223> Variant SIGIRR genomic (c.557delA)

<400> 2

actgcggggc ggggccgcac ctgctgccgc tgcggggct acgcgcgcgc cggggctgtt 60

gccgctgcgc acctggctca ggtgagctgc cccgcccccg cccggcgcga gccccaggtc 120

ctggcagcag gtgagtggcg gtcccatca gaccgcgaa ggtgagtgga ggtcgcagtc 180

ctcccagagc cgctgcaggc ctggagtcgg gggccggggc gctgcttgct cctgggcgcg 240

gtctctcagg aaatggctcg gggccgggga ggttcccgcg tcgcccgaa gcccagctc 300

gcccgcgtcc tggccctttc cagggtggg ggaaggagg aaagtggcg aaggggcg 360

gggccacctt ttctctgcc cgactgggtc agcggctttt tggtagacct cgcgcagaac 420

agagcggggc taggagcgc gaggcgcga gggatccac ctcggtcgcc caagcctgga 480

tcccagctcg aggttctctg aggcgccgag ggcggaccaa gattcaggct ggggcccag 540

ctgccccgc gccgggtgga gcttcgccca tccgcgcgc tctgccccgg ccgagcgccc 600

cgagccgcac ccccggtgcc gcctcccgc cgcgtctctg ttctccgcc tctcacttg 660

gagggcgccc agggcgagg tgggcgcggg gtaggggagc gaaccagagc ccctcgccct 720

gcggggccgc agccttttac ccacgttcct gcttcaggac aggcgctggg gggccagcgc 780

gcacggggag cggaggggag gcgcgtccga gctgagcggg agtcgcccgc aggggtgctgg 840

gtgctgagcg tcgaaccac cgggcagggg cggccgcgga ccgaggacgc tggctgggtc 900

tgaggcgagg aggagctggc gctttctaag aacctcaagg tgcccagcag agtcggtgg 960

gcagtgcgag gcgaccagg ccaggcttgg actgagattc tgggcgcctc ggaagagctg 1020

tgccggcgag ggtcggagac tttgtgggtg ttcattgctca ccctggacgg tggacctggg 1080

ggctcggggg tgcaggacgt tagcagagag aggggaaggag aggggtgcaa gccagccctg 1140

agagatgggt acgtgagac cgggtgcaca gagcagccgc cgagccgggg cttcacagtt 1200

ggcggctgac cccagggcct gtgctggagg cctgtctcct gcttcctgcg ggggtgccct 1260

gggccctgct gtgcctgctt cctgacctgc agtgtgggga gaaccttctt gaggggcttc 1320

caaggaccca cttaaggag gggctggtac caggctctgt gggggaaggg gtgctgtctt 1380

tgggacctg agcccaagg ccaccatcac ccaaggctga gtaccgggac aaaggaaggt 1440

cctagcaggg ctgggactca gtcagacaag tacctgtaga aagcctgctg ctgagagggt 1500

gctcagccag ggccgggagg ctggatggca gccctccaca atgaagccct gccccatgt 1560

ggctgagggc agggccagg tccaaaggag acccttgtct gattctgtgg gcaataccag 1620

ccaccggtca ctgggtcccc agtcccagg caacctgggg agtgtcagtt tggtgacaac 1680

aggagatgcc atctgccct catctgctct gagccaggag cctcgccac ctgtgtttgc 1740

tccttatttg gaaaaggggg cgtgccctc ccacccagc ctgaaggga ggtccacaag 1800

atgcatagt ccttctaac tgctcacca tcgggggagc tatgtctggg agagggacgg 1860

ggagaggcg gaagagtct gctggagccc cattgttgac cgggagtctc tgaaggccct 1920

gttccccggg gtcagagtc aaagacacat gggggagcaa accggccct tggctcctga 1980

gtgtgagtc tagtagctc tggaaggcat gccgcgcgg ccccggtcct ccttgatgt 2040

gtgcagggtg gccctggcgg gggctcctg gctggaggac ccccatgca ttaccctgg 2100

aacctgcccc tttagcacc cccaccca caacacacac gcacacacac aaacgcacgc 2160

acacgcatat actgcacaca cattgcacac acgcagacac actgcacaca cacacacaca 2220

cacacacaca cacacacaca cacacagct gcacagagga aagtccctg cctgggtggg 2280

ttggggcggg ctccagagga aggagccgt tgactcagag ctacaggaag agccggggca 2340

ggcgggtga gaaccacaa ctgcccgtcc tcccggcgc ctagggaagg ggtgggtgcc 2400

tgggtgggt ttaggttag gagaggatgt gactgcaagg ccaggaagct acagggaaag 2460

ttctgatcaa aataacaaa gacacaagcc aggtccccc cgcgttggc gatgcatcca 2520

tagcagcctt ggtgctgtgg acactccctg gggcccagcg aaggggagag tttgctccca	2580
aaggcgcacc aatgaccaac atttgccccc cggaggaaag aactggaacc aggtgactgc	2640
cttccccgct gctctgagtg agcctgtgtc agggcctggc tgagaggagg cccacggcgc	2700
ccaagaccca tatgcatgtg tgtgcatgtg tgtgtgagac agagcgaggg gcagcaggag	2760
gctgtgtcgc atcagcgccc aagggtgggt gtcaggtgca gtgtgtgtgc tgtggggtgc	2820
ctgtggtcct aggtctcttg acaagcgttg caatgggtgt tgtatatact caacagtgt	2880
gtgaggggtc tgcagcctaa gctgtgtgca gtggtaatgg gaggtgtgtg tggtaggaag	2940
tgtgctgtgt aggaggtgca agtgtgcaca cgggctgaga ctgaagctct gtgcaggtgc	3000
acatgggcca gagggtctgt tgtatgtgat gtgtgcaggg gaggaggggc agaggagagg	3060
gtgcagggga agttaggag gggagggggt acaaggaaag gtgtgcaggg gagagggtgc	3120
agggaaaagt ggggaggga gagggtgcag gggaaggtag gcagaggaga ggggtgcaggg	3180
gagagggtgc aggggaagt gtgcagggaa aggtggggag gggagagggt gcaggggaag	3240
gtaggcagag gagagggtgc aggggaaggt gtgcagggga gagggtgcag ggaaaggtgg	3300
ggaggggaga ggggtgcaggg gaaggtaggt agaggagagg gtgcagggga aggtgtgcag	3360
ggaaaggtgg ggagggagag ggtgcagggg aaggtaggca gaggagagag tgcaggggaa	3420
ggtgtgcagg ggagagggtg cagggaagg tggggagggg agagggtgca ggggaaggtg	3480
ggcagacgag agtacaggga aaggtaggga ggagagaagg tgcaggagga ggcaggcaga	3540
agagagggtg cagggaagg tggggagggg agagggtgca gggggaggca ggcagaggag	3600
agggttcagg ggagagggtg cagggaagg tgggcagggg aaggtgggga gtggagagcg	3660
tgcaggggga ggcaggcaga ggagagggtg cagggaagg tgcgtagggg agaggagca	3720
ggggagaggg tgcgggaaga ggggtcaggg gaggtttgc agggagaggg tgcaggggag	3780
aggggtcaag gagaggcagg cagcggagag ggtgcagagg agggtttgca ggggagggca	3840
cctgctttct tgctgtgtc gtatcctggc agtctgccct tgacttcag aaggctctga	3900
ggctgctttt ggagggagga ctttgaggtc tgagcttggg gcggtcacgg aagcactgca	3960
ggagatgccca ggggtttctgc tgggctggcg gtgtgacctc ccgggcttgc tgtgaggggt	4020
tggtttcttg ggaagtgtgt gctttaagga ccctcagtgg gtggttcccc tgggtcaagg	4080
actttccttg gagtgtcctc caccgcaaag gacttttca aggacccct ggccatcccc	4140
catgggtgtg tgcctcatgc tgtatcggtt cccagggaaa ctgccaaaca gtcaaacctg	4200
gggcctcttg ctgtgcttca cccatggaga gggaataggt ggcaggctaa gaggactgtc	4260

cctgccactg cctctggggc cttagcgggg aggaatcctg agggagctgg ggtgcagaaa 4320
 cttagccttg atccccacat cccacttcc actgtgtccc tgcagcaggg aggggagtag 4380
 acagtcaagt cctgttggat aaacaggggtg accctgtggc atccagttag gagaggcccc 4440
 tggggccctg ggctgccaag gaagaggctc tgggcttctc cctgcagcca cctcactgag 4500
 gagctttggg agacttgagg tgtgtgtggc tgcgtctacg ggctgggaaa ggacatggga 4560
 cccgagacct ccattccccc cgggcgctgt tcacctccg gcctcacaac cgcagggcga 4620
 ggctaagctg agggttcccg aggtgtttcc cgaggtgttt cgaggtgtct gtgctcacgt 4680

gcactcctga ctgtaaccag acgtgggcca gagctgggca cccccctcg actgcatcca 4740
 gacatggtta gagctgggca ccccccccc cgcactgtat ccagacatgg tcagagctga 4800
 gcacccctcc cgcactgtat ccagacatgt tcagagctgg gcacccccc cgcactgtat 4860
 ccagacatgt tcagagctgg gcacccccc cgcactgtat ccagacatgg tcagagctgg 4920
 gcacccccc cgcactgtat ccagacatgg tcagagctgg gcacccccg actgcatcca 4980
 gacatggtca gagctgagca ccccccccg cccgactgta tccagacatg gtcagagctg 5040
 ggcaccccc gactgcatcc agacatggtc agagctgggc accccctcc cgcactgcat 5100

tcagacatgg tcagagctgg gcacccccg cgactgcatc cagacatggt cagagttggg 5160
 ccccccccc cgcactgtat ccagacatgg tcagagctgg gcacccccg actgcatcca 5220
 gacatggtca gagctgggca cccctccccg actgtatcca gacatggtca gagctgggca 5280
 ccccccccg actgtatcca gacatggtca gagctgggca cctcccccg actgtatcca 5340
 gacatggtca gagctgggca cctccccga ctgcatccag acatggtcag agctgggcac 5400
 ccccccgac tgcattcaga catggtcaga gctgggcacc ccccccccc actgtatcca 5460
 gacatggtca gagctgggca cccccctcc cgactgcatc cagacatggt cagagctgag 5520

caccccccc ccccgactgt atccagacat ggtcagagct gggcaccccc ctccccgact 5580
 glatccagac atggtcagag ctgagcccc ccccccccg actglatcca gacatggtca 5640
 gagctgggca cccccgact gcatccagac atggtcagag ctgagcacc ctccccgact 5700
 gcatccagac atggtcagag ctgggcacc cccccccca cgcactgtat ccagacatgg 5760
 tcagagctgg gcaacccca cgcactgtat ccagacatgg tcagagctgg gcacccccg 5820
 actgcatcca gacatggtca gagctgggca ccccccccg actgtatcca gacatggtca 5880
 gagctgggca cccccctcc cgactgcatc cagacatggt cagagctggg cccccctc 5940

cccgactgca tccagacatg gtcagagctg ggcacccctc cccgactgta tccagacatg 6000
 gtcagagctg ggcaccccc tccccgactg tatccagaca tggtcagagc tgggcacccc 6060
 cgccgactgt atccagacat ggtcagagct gggcaccccc cgactgcatc cagacatggt 6120

cagagctggg caccctccc cgactgtatc cagacatggt cagagctggg ccccccccc 6180
ccccgactgt atccagacat ggtcagagct gaggaccct ccccgactgt atccagacat 6240
gttcagagct gggcaccccc cccccgact gtatccagac atggtcagag ctgggcaccc 6300
ccctccccga ctgcatccag acatgggtcag agctgagcac ccccccgcc cgactgcatc 6360

cagacatggt cagagctggg cccccccga ctgcatccag acatgggtcag agctgggcac 6420
ccccctccc gactgcatc agacatggtc agagctgggc accccctcc ccgactgcat 6480
ccagacatgg tcagagctgg gcacccccgc cgactgtatc cagacatggt cagagctggg 6540
ccccccctc cccgactgta tccagacatg gtcagagctg ggcacccccg ccgactgtat 6600
ccagacatgg tcagagctgg gcacccccct ccccgactgc atccagacat gttcagagct 6660
gggcaccccc ccccgactgt atccagacat ggtcagagct gggcaccccc cgactgcatc 6720
cagacatggt cagagctggg caaccctc cctgactgca tccagacatg gtcagagctg 6780

ggcaccccc ccggtgtat ccagacatgg tcagagctgg gcaccccc ccccgactg 6840
catccagaca tggtcagagc taagcaaagc cccccccga ctgtatccag acatgggtcag 6900
agctgggcac cccacccca actgcatcca gacatggtca gagctgagca cccccaccg 6960
actgtatcca gacatggtca gagctgggca tctccatgg ttaggttac atcactaca 7020
ccctgcccc aaggccaggc tgtgtctctg gcactgcagc acccctgca gtggcagctc 7080
tgttgaagg gacagccagg agtgaggcca actgcatgt cgggagacag ggtccacct 7140
agtgttcgtc cactgcacc tggggccgtc atggtggccc cagggccggt ggaggtgatc 7200

ctggacttga tggccctcc ccgtgggtcct gggctgcgtg aagtgggtcc cattctggcc 7260
acgtgtgagg ggctcaaaga gacaccctgc tggcctccga gaactctgcc agcatccatg 7320
cggctgtca gggctgcac tggtcagggc cagtgtgtc tgtggccggc agttactggt 7380
taactcatcc tgttctgtt taacttgttc cctggggatg gcggcagcct ctgacctgtc 7440
caggtgccct gtccagctga ctgcaaggac agagaggagt cctgcccagc tcttgatca 7500
gtctgtggc cgaggagccc ggtggagcca ggggtgacc tggagcccag cctgccccga 7560
ggaggccccg gtcagagcc atgccaggta ggcaggctcag agagggatgg gcttgaattc 7620

cccactccca cctcctgccc ctgaggtgg ctccacattg tgcactctgc tcttatccca 7680
ggcactctgc ctggccacat gccgggggta caaaggcccc tgcacccc tcccactggg 7740
cccttggccc tgaaccagg gcgagtgtg aggatgagcc ctgcaccctc tcctgcccc 7800
gaccaggctc atgagggtca gtaaagccca ggaatgctag gagtgggcac tgctggggcc 7860
cacagcaggg tgtggccgat gtgagctggc aggagacat accctactga gaagatggtg 7920
gtttgacctc ccgtagggc tcatcacct gatgtaaaga tgaacaagtt caggcctcag 7980

gtggtgggga gggcttgcct gtggccagga gcagagggcc ccaagacccc caaggcagga 8040

atgaggccaa ctgctccagg ctcccagctg tgcccatcca ggtggaaggt gctggcacac 8100

tgggcatctg ggtcccaaaa cccactggtg cccacactct ctgcagattc cccagcagag 8160

gcaggccagg ctgaggggtg cagtgtcccg ccctcctat gagacaggca tcatggtgtc 8220

accaggggtg ggcaacaggt gcatctggga aatggggggc tggactcaca tcttgggtgg 8280

acaagcagcc agatggggct tgagcgtggg cctggggcag ctcaactgtg gccactgtc 8340

tctgtccct cagcacaggg tgggcggaac acccaaagcc tggccactac tggggaaaga 8400

cgtgccaagt gctaagcctg tccccatggg actcgtggac agacatggtg tgactgccaa 8460

gccacacaga ggccttattt ctctctgcac cacggaccac tgttgagggtg gtggggccag 8520

gcacctggc acagctgacc cactgtgtcc tctctcttca ggtgtctgtg atagggcccc 8580

tgacttctc tccccgtctg aagaccaggt gctgaggcct gccttgggca gctcagtggc 8640

tctgaactgc acggcttggg tagtctctgg gcccactgc tccctgcctt cagtccagtg 8700

gctgaaagac gggcttccat tgggaattgg gggccactac agcctccacg agtactcctg 8760

gtaagagacc cgggtatcca gggtcagagt gacctctcag acatgccaag ggcagtgtgc 8820

tccctgacct atgcctgggt ctgtaaccac ctggccagac atgtcacgag cagagtgcct 8880

ctgtggacc agacggaggg tcagtaatgt caggacaga cccttgggtc tgcattggccc 8940

ctgtaggcac acacctgta actgcccagg gctacgggtg tcagcaaagt ctagggtctg 9000

aatgaacacc gaccagctgg ccaagtctgt gatcaaagga cctccaggcc ccagaccag 9060

gactgggacg gctgttctga aacaaatctg catccctcc aggttcaggt cacttctctt 9120

ggcttaaggc ccaagaacca cccaggaggc ccaattctgc agcctccctg ggtgaacca 9180

gacaggtggg gttcgggggt gtccactggt atccccagtc gacctgacc ctggcttgg 9240

atccccaggg tcaaggccaa cctgtcagag gtgcttgtgt ccagtgtcct gggggtcaac 9300

gtgaccagca ctgaagtcta tggggccttc acctgtcca tccagaacat cagcttctcc 9360

tccttctc ttcagagagc tggatgagg ggttcccaaa ggtggagggt agaaggggac 9420

ctagacctag tgaggccag ggatcttgag gcttgggggc tgctggagg gcagcggctc 9480

aggcaacca tccgaggcc ctacaagcca cgtggctgcg gtgtggcct cctcctggt 9540

cctgtggcc ctgtgtgg ccgccctgct ctatgtcaag tgccgtctca acgtgtgtct 9600

ctggtaccag gacgcgtatg gggagggtga gataaacggt gcgtggggcc cgcgtgggcg 9660

cgaggagggg tgcacaggc ctccctgaga gtgcgtgtgt gagctcggct ttggaggcgt 9720

gcggcgggga ggggttagcc cccgggtgct ctgtgcggcc cggctggggt taaagtccgg 9780
ggcgggtctg cccctgtgca cgtgggagat ggtaggggt gttggggacc ccgagggtg 9840
gtgtgaggcc tcgggagcca tcgggtgac tggccgccgt ccgcagacgg gaagctctac 9900
gacgcctacg tctcctacag cgactgcccc gaggaccgca agttcgtgaa cttcatccta 9960
agccgcagct ggagcggcgt cggggctaca agctcttctt ggacgaccgc gacctcctgc 10020
cgcgctggtg tatcccgggc cccaccccggt gccccgcca cccggagggc ccgccccgcc 10080
ccgccccatg ctccgtccca cccggggccc caccaccacc atcgagctcc gcccccatcc 10140

ccgccccacc gagggccgcg ctcaaaaacc cgcccccc gcaggcccg cccctccctt 10200
agagctctgc gtccggcacc gcccttgagc ctccggcccc gccctcccc gccccgccc 10260
ctgcgccgcc gacgcccccc ctccgcaga gccctccgcc gacctcttgg tgaacctgag 10320
ccgctgccga cgcctcatcg tgggtctttc ggacgccttc ctgagccggg cctggtgcag 10380
ccacagcttc cgggtgggtcc cgcgcggggt tgggtgggcc ccagcgtagc accaccccc 10440
ctgacggtcc cgccccgag ggagggcctg tgccggctgc tggagctcac ccgcagacc 10500
atcttcatca ctttcgaggg ccagaggcgc gaccccgccg acccggcgct ccgctgctg 10560

cgccagcacc gccacctggt gacctgtctg ctctggaggc ccggtccgt ggtgcggagc 10620
aggcgcgga ggggtccggg ctagcggcgg gttagagatg ggcggtgcc gggtccagg 10680
ctgggacccc tcgtgggga gctctgcggc accacgttt gtgaatgggc cctgggggga 10740
ggttccgctg cctggggccc cgatgcgggg agccgccctt gagggcccg gagccacgga 10800
atagctgtcg cagggcgtgg aaccgtggg cagccgcagg tgtctcttg ggggccagga 10860
cgccaggggc ttccgaggtg ttacacctg caaacgccc cgacctggcc ccagactcc 10920
ttctccgat ttttgaaag aagtgcagct ggcgtgccg cggaaggtgc agtacaggcc 10980

tgtggaagga gacccccaga cgcagctgca ggacgacaag gacccatgc tgattcttcg 11040
aggccgagtc cctgagggcc gggccctgga ctcagagtg gaccggacc ctgaggcgga 11100
cctgggtatg cccgccagc cccactcccc aactggagaa gctcagcaca gggcgagtg 11160
ggggcaggca cagggcacag ggcctggagg ggtccaggt gttgaggact cttcccgca 11220
ccgggagccc ctgcacggcc tctgcctgg aggtgctcgg ccctcggtct gctgggaac 11280
ttctgggcc tcacaggcca tcacagcagg gggtagcag gggcagcccc tggcagtggg 11340
tctgggcaa ggtgtgggt ggccacctca ggcgtctcgg tctccccacc ccaggtgtcc 11400

gggggcctgt ctttgagag ccatacgtc caccgcacac cagtggggtc tcgctgggag 11460
agagccggag cagcgaagtg gacgtctcgg atctcggtc gcgaaactac agtccccga 11520
cagacttcta ctgcctggtg tccaaggatg atatgtagct cccaccag agtgcaggat 11580

catagggaca gcgggggcca gggcagcggc gtcgctcctc tgctcaacag gaccacaacc 11640
cctgccagca gccctgggac cctgccagca gccctgggaa aaggctgtgg cctcagggcg 11700
cctcccagtg ccagaaaata aagtcctttt ggattctg 11738

<210> 3

<211> 1230

<212> DNA

<213> Homo sapien

<220><223> wild type SIGIRR mRNA

<400> 3

augccaggug ucugugauag gggcccugac uuccucuccc cgucugaaga ccaggugcug 60
aggccugccu ugggcagcuc aguggcucug aacugcacgg cuuggguagu cucuggggccc 120
cacugcuccc ugccuucagu ccaggugcug aaagacgggc uuccauuggg aauggggggc 180
cacuacagcc uccacgagua cuccuggguc aaggccaacc ugucagaggu gcuugugucc 240
aguguccugg gggucaacgu gaccagcacu gaagucuaug gggccuucac cugcuccauc 300
cagaacauc gcuuccuccu cuucacucuu cagagagcug gcccuacaag ccaggugcuc 360
gcggugcugg ccuccuccu gguccugcug gccugcugc ugcccgcccu gcucuauguc 420

aagugccguc ucaacgugcu gcucugguac caggacgcgu auggggaggu ggagauaaac 480
gacgggaagc ucuacgacgc cuacgucucc uacagcgacu gcccagagga ccgcaaguuc 540
gugaacuua uccuaaagcc gcagcuggag cggcgucggg gcuacaagcu cuuccuggac 600
gaccgcgacc uccugccgag cgucagagcc uccgcccacc ucuuggugaa ccugagccgc 660
ugccgacgcc ucaucguggu gcuuucggac gccuuccuga gccgggcccug gugcagccac 720
agcuuccggg agggccugug ccggcugcug gacgucaccc gcagaccgau cuucaucacc 780
uucgagggcc agaggcgcga ccccgccgac ccggcgucucc gccugcugcg ccagaccgcg 840

caccugguga ccuugcugcu cuggaggccc ggcuccguga cuccuuccuc cgauuuuugg 900
aaagaagugc agcuggcgcu gccgcggaag gugcaguaca ggccugugga aggagacccc 960
cagacgcagc ugacggacga caaggacccc augcugauuc uucgagggcg agucccugag 1020
ggccggggccc uggacucaga gguggacccg gaccucagag ggcaccuggg uguccggggg 1080
ccugucuuug gagagccauc agcuccaccg cacaccagug gggucucgcu gggagagagc 1140
cggagcagcg aaguggacgu cucggaucuc ggucgcgaa acuacagugc ccgcacagac 1200
uucuacugcc ugguguccaa ggaugauaug 1230

<210> 4

<211> 645

<212> DNA

<213> Homo sapien

<220><223> variant (c.557delA) SIGIRR mRNA

<400> 4

```

augccaggug ucugugauag ggccccugac uuccucuccc cgucugaaga ccaggugcug      60
aggccugccu ugaggcagcuc aguggcucug aacugcacgg cuuggguagu cucuggggccc      120
cacugcuccc ugccuucagu ccaguggcug aaagacgggc uuccauuggg aauugggggc      180
cacuacagcc uccacgagua cuccuggguc aaggccaacc ugucagaggu gcuugugucc      240
aguguccugg gggucaacgu gaccagcacu gaagucuaug gggccuucac cugcuccauc      300
cagaacauca gcuucuccuc cuucacucuu cagagagcug gcccuacaag ccacugggcu      360

gcggugcugg ccuccuccu gguccugcug gccucgucg ugcccgcccu gcucuauguc      420
aagugccguc ucaacgugcu gcucugguac caggacgcu auggggaggu ggagauaaac      480
gacgggaagc ucuacgacgc cuacgucucc uacagcgacu gccccgagga ccgcaaguuc      540
gugaacuuca uccuaagccg cagcuggagc ggcgucgggg cuacaagcuc uuccuggacg      600
accgcgaccu ccugccgcgc gcugagcccu ccgccgaccu cuugg      645

```

<210> 5

<211> 1233

<212> DNA

<213> Artificial Sequence

<220><223> wild type SIGIRR cDNA

<400> 5

```

atgccagggtg tctgtgatag ggccccgac ttctctccc cgtctgaaga ccagggtctg      60

aggcctgcct tgggcagctc agtggctctg aactgcacgg cttgggtagt ctctgggccc      120
cactgtctcc tgccttcagt ccagtggctg aaagacgggc ttccattggg aattgggggc      180
cactacagcc tccacgagta ctctgggtc aaggccaacc tgtcagaggt gcttgtgtcc      240
agtgtcctgg gggtaacgt gaccagcact gaagtctatg gggccttcac ctgtccatc      300
cagaacatca gcttctctc cttaactctt cagagagctg gccctacaag ccacgtggct      360
gcggtgctgg cctccctctt ggtctgtctg gccctgtctg tggccgcct gctctatgtc      420
aagtgccgtc tcaactgtct gctctgttac caggacgcgt atggggaggt ggagataaac      480

gacgggaagc tctacgacgc ctacgtctcc tacagcgact gccccgagga ccgcaagttc      540
gtgaacttca tcctaaagcc gcagctggag cggcgtcggg gctacaagct cttcctggac      600

```

gaccgcgacc tctgccgcg cgctgagccc tccgccgacc tcttggtgaa cctgagccgc 660
tgccgacgcc tcatcgtggt gctttcggac gccttcctga gccgggcctg gtgcagccac 720
agcttccggg agggcctgtg ccggctgctg gagctcacc gcagacccat cttcatcacc 780
ttcgagggcc agaggcgca ccccgccac ccggcgctcc gcctgctgcg ccagcaccgc 840
cacctggtga ccttgctgct ctggaggccc ggctccgtga ctccttcctc cgatttttgg 900

aaagaagtgc agctggcgct gccgcggaag gtgcagtaca ggctgtgga aggagacccc 960
cagacgcagc tgcaggacga caaggacccc atgctgattc ttcgaggccg agtccctgag 1020
ggccggggccc tggactcaga ggtggaccgc gacctgagg gcgacctggg tgtccggggg 1080
cctgtctttg gagagccatc agctccaccg cacaccagtg gggctcgcct gggagagagc 1140
cggagcagcg aagtggacgt ctcgatctc ggctcgcga actacagtgc ccgcacagac 1200
ttctactgcc tgggtgtcaa ggatgatatg tag 1233

<210> 6

<211> 648

<212> DNA

<213> Artificial Sequence

<220><223> variant (c.557delA) SIGIRR cDNA

<400> 6

atgccaggtg tctgtgatag ggccctgac ttctctccc cgtctgaaga ccagtgctg 60
aggcctgcct tgggcagctc agtggctctg aactgcacgg cttgggtagt ctctgggccc 120
cactgtctcc tgccttcagt ccagtggctg aaagacgggc ttccattggg aattgggggc 180
cactacagcc tccagagta ctctgggtc aaggccaacc tgtcagaggt gcttgtgtcc 240
agtgtcctgg gggtaacgt gaccagcact gaagtctatg gggccttcac ctgtccatc 300
cagaacatca gcttctctc cttactctt cagagagctg gccctacaag ccacgtggct 360

gcggtgctgg cctccctcct ggtcctgctg gcctgctgc tggccgcct gctctatgtc 420
aagtgccgtc tcaactgtct gctctggtac caggacgcgt atggggaggt ggagataaac 480
gacgggaagc tctacagc ctactctcc tacagcact gccccgagga ccgcaagtcc 540
gtgaacttca tctaagccg cagctggagc ggctcgggg ctacaagctc ttctggagc 600
accgcgacct cctgccgcgc gctgagccct ccgccgacct cttggtga 648

<210> 7

<211> 410

<212> PRT

<213> Homo sapien

<220><223> wild type SIGIRR protein

<400> 7

Met Pro Gly Val Cys Asp Arg Ala Pro Asp Phe Leu Ser Pro Ser Glu

1 5 10 15
Asp Gln Val Leu Arg Pro Ala Leu Gly Ser Ser Val Ala Leu Asn Cys

20 25 30
Thr Ala Trp Val Val Ser Gly Pro His Cys Ser Leu Pro Ser Val Gln

35 40 45
Trp Leu Lys Asp Gly Leu Pro Leu Gly Ile Gly Gly His Tyr Ser Leu

50 55 60
His Glu Tyr Ser Trp Val Lys Ala Asn Leu Ser Glu Val Leu Val Ser

65 70 75 80
Ser Val Leu Gly Val Asn Val Thr Ser Thr Glu Val Tyr Gly Ala Phe

85 90 95
Thr Cys Ser Ile Gln Asn Ile Ser Phe Ser Ser Phe Thr Leu Gln Arg

100 105 110
Ala Gly Pro Thr Ser His Val Ala Ala Val Leu Ala Ser Leu Leu Val

115 120 125
Leu Leu Ala Leu Leu Leu Ala Ala Leu Leu Tyr Val Lys Cys Arg Leu

130 135 140
Asn Val Leu Leu Trp Tyr Gln Asp Ala Tyr Gly Glu Val Glu Ile Asn

145 150 155 160
Asp Gly Lys Leu Tyr Asp Ala Tyr Val Ser Tyr Ser Asp Cys Pro Glu

165 170 175
Asp Arg Lys Phe Val Asn Phe Ile Leu Lys Pro Gln Leu Glu Arg Arg

180 185 190
Arg Gly Tyr Lys Leu Phe Leu Asp Asp Arg Asp Leu Leu Pro Arg Ala

195 200 205
Glu Pro Ser Ala Asp Leu Leu Val Asn Leu Ser Arg Cys Arg Arg Leu

210 215 220
Ile Val Val Leu Ser Asp Ala Phe Leu Ser Arg Ala Trp Cys Ser His

225 230 235 240
 Ser Phe Arg Glu Gly Leu Cys Arg Leu Leu Glu Leu Thr Arg Arg Pro
 245 250 255
 Ile Phe Ile Thr Phe Glu Gly Gln Arg Arg Asp Pro Ala His Pro Ala

 260 265 270
 Leu Arg Leu Leu Arg Gln His Arg His Leu Val Thr Leu Leu Leu Trp
 275 280 285
 Arg Pro Gly Ser Val Thr Pro Ser Ser Asp Phe Trp Lys Glu Val Gln
 290 295 300
 Leu Ala Leu Pro Arg Lys Val Gln Tyr Arg Pro Val Glu Gly Asp Pro
 305 310 315 320
 Gln Thr Gln Leu Gln Asp Asp Lys Asp Pro Met Leu Ile Leu Arg Gly

 325 330 335
 Arg Val Pro Glu Gly Arg Ala Leu Asp Ser Glu Val Asp Pro Asp Pro
 340 345 350
 Glu Gly Asp Leu Gly Val Arg Gly Pro Val Phe Gly Glu Pro Ser Ala
 355 360 365
 Pro Pro His Thr Ser Gly Val Ser Leu Gly Glu Ser Arg Ser Ser Glu
 370 375 380
 Val Asp Val Ser Asp Leu Gly Ser Arg Asn Tyr Ser Ala Arg Thr Asp

 385 390 395 400
 Phe Tyr Cys Leu Val Ser Lys Asp Asp Met
 405 410

 <210> 8
 <211> 504
 <212> PRT
 <213> Homo sapien
 <220><223> wild type SIGIRR isoform protein
 <400> 8
 Met Pro Gly Val Cys Asp Arg Ala Pro Asp Phe Leu Ser Pro Ser Glu
 1 5 10 15
 Asp Gln Val Leu Arg Pro Ala Leu Gly Ser Ser Val Ala Leu Asn Cys

20	25	30	
Thr Ala Trp Val Val Ser Gly Pro His Cys Ser Leu Pro Ser Val Gln			
35	40	45	
Trp Leu Lys Asp Gly Leu Pro Leu Gly Ile Gly Gly His Tyr Ser Leu			
50	55	60	
His Glu Tyr Ser Trp Val Lys Ala Asn Leu Ser Glu Val Leu Val Ser			
65	70	75	80
Ser Val Leu Gly Val Asn Val Thr Ser Thr Glu Val Tyr Gly Ala Phe			
85	90	95	
Thr Cys Ser Ile Gln Asn Ile Ser Phe Ser Ser Phe Thr Leu Gln Arg			
100	105	110	
Ala Gly Pro Thr Ser His Val Ala Ala Val Leu Ala Ser Leu Leu Val			
115	120	125	
Leu Leu Ala Leu Leu Leu Ala Ala Leu Leu Tyr Val Lys Cys Arg Leu			
130	135	140	
Asn Val Leu Leu Trp Tyr Gln Asp Ala Tyr Gly Glu Val Glu Ile Asn			
145	150	155	160
Asp Gly Lys Leu Tyr Asp Ala Tyr Val Ser Tyr Ser Asp Cys Pro Glu			
165	170	175	
Asp Arg Lys Phe Val Asn Phe Ile Leu Lys Pro Gln Leu Glu Arg Arg			
180	185	190	
Arg Gly Tyr Lys Leu Phe Leu Asp Asp Arg Asp Leu Leu Pro Arg Ala			
195	200	205	
Glu Pro Ser Ala Asp Leu Leu Val Asn Leu Ser Arg Cys Arg Arg Leu			
210	215	220	
Ile Val Val Leu Ser Asp Ala Phe Leu Ser Arg Ala Trp Cys Ser His			
225	230	235	240
Ser Phe Arg Glu Gly Leu Cys Arg Leu Leu Glu Leu Thr Arg Arg Pro			
245	250	255	
Ile Phe Ile Thr Phe Glu Gly Gln Arg Arg Asp Pro Ala His Pro Ala			
260	265	270	

Leu Arg Leu Leu Arg Gln His Arg His Leu Val Thr Leu Leu Leu Trp
 275 280 285

Arg Pro Gly Ser Val Thr Pro Ser Ser Asp Phe Trp Lys Glu Val Gln
 290 295 300

Leu Ala Leu Pro Arg Lys Val Gln Tyr Arg Pro Val Glu Gly Asp Pro
 305 310 315 320

Gln Thr Gln Leu Gln Asp Asp Lys Asp Pro Met Leu Ile Leu Arg Gly
 325 330 335

Arg Val Pro Glu Gly Arg Ala Leu Asp Ser Glu Val Asp Pro Asp Pro
 340 345 350

Glu Gly Asp Leu Gly Met Pro Ala Gln Pro His Ser Pro Thr Gly Glu
 355 360 365

Ala Gln His Arg Ala Glu Trp Gly Gln Ala Gln Gly Thr Gly Pro Gly
 370 375 380

Gly Ala Leu Gly Val Glu Asp Ser Ser Arg His Arg Glu Pro Leu His
 385 390 395 400

Gly Leu Cys Pro Gly Gly Ala Arg Pro Ser Val Cys Leu Gly Thr Ser
 405 410 415

Trp Ala Ser Gln Ala Ile Thr Ala Gly Gly Glu Gln Gly Gln Pro Leu
 420 425 430

Ala Val Gly Leu Gly Gln Gly Cys Gly Trp Pro Pro Gln Ala Ser Arg
 435 440 445

Ser Pro His Pro Arg Cys Pro Gly Ala Cys Phe Trp Arg Ala Ile Ser
 450 455 460

Ser Thr Ala His Gln Trp Gly Leu Ala Gly Arg Glu Pro Glu Gln Arg
 465 470 475 480

Ser Gly Arg Leu Gly Ser Arg Leu Ala Lys Leu Gln Cys Pro His Arg
 485 490 495

Leu Leu Leu Pro Gly Val Gln Gly
 500

<210> 9

<211> 215

<212> PRT

<213> Homo sapien

<220><223> Variant SIGIRR (c.557delA)

<400> 9

Met Pro Gly Val Cys Asp Arg Ala Pro Asp Phe Leu Ser Pro Ser Glu

1 5 10 15

Asp Gln Val Leu Arg Pro Ala Leu Gly Ser Ser Val Ala Leu Asn Cys

20 25 30

Thr Ala Trp Val Val Ser Gly Pro His Cys Ser Leu Pro Ser Val Gln

35 40 45

Trp Leu Lys Asp Gly Leu Pro Leu Gly Ile Gly Gly His Tyr Ser Leu

50 55 60

His Glu Tyr Ser Trp Val Lys Ala Asn Leu Ser Glu Val Leu Val Ser

65 70 75 80

Ser Val Leu Gly Val Asn Val Thr Ser Thr Glu Val Tyr Gly Ala Phe

85 90 95

Thr Cys Ser Ile Gln Asn Ile Ser Phe Ser Ser Phe Thr Leu Gln Arg

100 105 110

Ala Gly Pro Thr Ser His Val Ala Ala Val Leu Ala Ser Leu Leu Val

115 120 125

Leu Leu Ala Leu Leu Leu Ala Ala Leu Leu Tyr Val Lys Cys Arg Leu

130 135 140

Asn Val Leu Leu Trp Tyr Gln Asp Ala Tyr Gly Glu Val Glu Ile Asn

145 150 155 160

Asp Gly Lys Leu Tyr Asp Ala Tyr Val Ser Tyr Ser Asp Cys Pro Glu

165 170 175

Asp Arg Lys Phe Val Asn Phe Ile Leu Ser Arg Ser Trp Ser Gly Val

180 185 190

Gly Ala Thr Ser Ser Ser Trp Thr Thr Ala Thr Ser Cys Arg Ala Leu

195 200 205

Ser Pro Pro Pro Thr Ser Trp

210 215

<210> 10

<211> 30

<212> PRT

<213> Homo sapien

<220><223> Wild type SIGIRR amio acids 186 to 215

<400> 10

Lys Pro Gln Leu Glu Arg Arg Arg Gly Tyr Lys Leu Phe Leu Asp Asp

1 5 10 15

Arg Asp Leu Leu Pro Arg Ala Glu Pro Ser Ala Asp Leu Leu

20 25 30

<210> 11

<211> 30

<212> PRT

<213> Homo sapien

<220><223> Variant SIGIRR amio acids 186 to 215

<400> 11

Ser Arg Ser Trp Ser Gly Val Gly Ala Thr Ser Ser Ser Trp Thr Thr

1 5 10 15

Ala Thr Ser Cys Arg Ala Leu Ser Pro Pro Pro Thr Ser Trp

20 25 30

<210> 12

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> forward PCR primer for SIGIRR

<400> 12

tcagtggctc tgaactgcac 20

<210> 13

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> reverse PCR primer for SIGIRR

<400> 13

ggtcctgttg agcagaggag 20

<210> 14

<211> 910

<212> PRT

<213> Artificial Sequence

<220

><223> rilonacept fusion protein

<400> 14

Met Val Leu Leu Trp Cys Val Val Ser Leu Tyr Phe Tyr Gly Ile Leu

1 5 10 15

Gln Ser Asp Ala Ser Glu Arg Cys Asp Asp Trp Gly Leu Asp Thr Met

20 25 30

Arg Gln Ile Gln Val Phe Glu Asp Glu Pro Ala Arg Ile Lys Cys Pro

35 40 45

Leu Phe Glu His Phe Leu Lys Phe Asn Tyr Ser Thr Ala His Ser Ala

50 55 60

Gly Leu Thr Leu Ile Trp Tyr Trp Thr Arg Gln Asp Arg Asp Leu Glu

65 70 75 80

Glu Pro Ile Asn Phe Arg Leu Pro Glu Asn Arg Ile Ser Lys Glu Lys

85 90 95

Asp Val Leu Trp Phe Arg Pro Thr Leu Leu Asn Asp Thr Gly Asn Tyr

100 105 110

Thr Cys Met Leu Arg Asn Thr Thr Tyr Cys Ser Lys Val Ala Phe Pro

115 120 125

Leu Glu Val Val Gln Lys Asp Ser Cys Phe Asn Ser Pro Met Lys Leu

130 135 140

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

145 150 155 160

Pro Asn Val Asp Gly Tyr Phe Pro Ser Ser Val Lys Pro Thr Ile Thr

165 170 175

Trp Tyr Met Gly Cys Tyr Lys Ile Gln Asn Phe Asn Asn Val Ile Pro

180 185 190

Glu Gly Met Asn Leu Ser Phe Leu Ile Ala Leu Ile Ser Asn Asn Gly
 195 200 205

 Asn Tyr Thr Cys Val Val Thr Tyr Pro Glu Asn Gly Arg Thr Phe His
 210 215 220
 Leu Thr Arg Thr Leu Thr Val Lys Val Val Gly Ser Pro Lys Asn Ala
 225 230 235 240
 Val Pro Pro Val Ile His Ser Pro Asn Asp His Val Val Tyr Glu Lys
 245 250 255
 Glu Pro Gly Glu Glu Leu Leu Ile Pro Cys Thr Val Tyr Phe Ser Phe
 260 265 270
 Leu Met Asp Ser Arg Asn Glu Val Trp Trp Thr Ile Asp Gly Lys Lys

 275 280 285
 Pro Asp Asp Ile Thr Ile Asp Val Thr Ile Asn Glu Ser Ile Ser His
 290 295 300
 Ser Arg Thr Glu Asp Glu Thr Arg Thr Gln Ile Leu Ser Ile Lys Lys
 305 310 315 320
 Val Thr Ser Glu Asp Leu Lys Arg Ser Tyr Val Cys His Ala Arg Ser
 325 330 335
 Ala Lys Gly Glu Val Ala Lys Ala Ala Lys Val Lys Gln Lys Val Pro
 340 345 350

 Ala Pro Arg Tyr Thr Val Ser Gly Gly Ala Pro Met Leu Ser Glu Ala
 355 360 365
 Asp Lys Cys Lys Glu Arg Glu Glu Lys Ile Ile Leu Val Ser Ser Ala
 370 375 380
 Asn Glu Ile Asp Val Arg Pro Cys Pro Leu Asn Pro Asn Glu His Lys
 385 390 395 400
 Gly Thr Ile Thr Trp Tyr Lys Asp Asp Ser Lys Thr Pro Val Ser Thr
 405 410 415
 Glu Gln Ala Ser Arg Ile His Gln His Lys Glu Lys Leu Trp Phe Val

 420 425 430
 Pro Ala Lys Val Glu Asp Ser Gly His Tyr Tyr Cys Val Val Arg Asn

435 440 445
 Ser Ser Tyr Cys Leu Arg Ile Lys Ile Ser Ala Lys Phe Val Glu Asn
 450 455 460
 Glu Pro Asn Leu Cys Tyr Asn Ala Gln Ala Ile Phe Lys Gln Lys Leu
 465 470 475 480
 Pro Val Ala Gly Asp Gly Gly Leu Val Cys Pro Tyr Met Glu Phe Phe
 485 490 495

 Lys Asn Glu Asn Asn Glu Leu Pro Lys Leu Gln Trp Tyr Lys Asp Cys
 500 505 510
 Lys Pro Leu Leu Leu Asp Asn Ile His Phe Ser Gly Val Lys Asp Arg
 515 520 525
 Leu Ile Val Met Asn Val Ala Glu Lys His Arg Gly Asn Tyr Thr Cys
 530 535 540
 His Ala Ser Tyr Thr Tyr Leu Gly Lys Gln Tyr Pro Ile Thr Arg Val
 545 550 555 560
 Ile Glu Phe Ile Thr Leu Glu Glu Asn Lys Pro Thr Arg Pro Val Ile

 565 570 575
 Val Ser Pro Ala Asn Glu Thr Met Glu Val Asp Leu Gly Ser Gln Ile
 580 585 590
 Gln Leu Ile Cys Asn Val Thr Gly Gln Leu Ser Asp Ile Ala Tyr Trp
 595 600 605
 Lys Trp Asn Gly Ser Val Ile Asp Glu Asp Asp Pro Val Leu Gly Glu
 610 615 620
 Asp Tyr Tyr Ser Val Glu Asn Pro Ala Asn Lys Arg Arg Ser Thr Leu
 625 630 635 640

 Ile Thr Val Leu Asn Ile Ser Glu Ile Glu Ser Arg Phe Tyr Lys His
 645 650 655
 Pro Phe Thr Cys Phe Ala Lys Asn Thr His Gly Ile Asp Ala Ala Tyr
 660 665 670
 Ile Gln Leu Ile Tyr Pro Val Thr Asn Ser Gly Asp Lys Thr His Thr
 675 680 685
 Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe

690 695 700
 Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro

 705 710 715 720
 Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val

 725 730 735
 Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr

 740 745 750
 Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val

 755 760 765
 Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys

 770 775 780

 Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
 785 790 795 800
 Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro

 805 810 815
 Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val

 820 825 830
 Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly

 835 840 845
 Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp

 850 855 860
 Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
 865 870 875 880
 Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His

 885 890 895
 Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys

 900 905 910

 <210> 15
 <211> 116
 <212> PRT
 <213> Artificial Sequence

<220><223> heavy chain variable region of antibody directed to IL-6R

<400> 15

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg

1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Arg Phe Thr Phe Asp Asp Tyr

20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35 40 45

Ser Gly Ile Ser Trp Asn Ser Gly Arg Ile Gly Tyr Ala Asp Ser Val

50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Glu Asn Ser Leu Phe

65 70 75 80

Leu Gln Met Asn Gly Leu Arg Ala Glu Asp Thr Ala Leu Tyr Tyr Cys

85 90 95

Ala Lys Gly Arg Asp Ser Phe Asp Ile Trp Gly Gln Gly Thr Met Val

100 105 110

Thr Val Ser Ser

115

<210> 16

<211> 107

<212> PRT

<213> Artificial Sequence

<220><223> light chain variable region of antibody directed to IL-6R

<400> 16

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Val Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Trp

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

Tyr Gly Ala Ser Ser Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Ser Tyr Tyr Cys Gln Gln Ala Asn Ser Phe Pro Tyr

85 90 95

Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys
100 105