



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년02월04일
(11) 등록번호 10-2762920
(24) 등록일자 2025년01월31일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6883 (2018.01) C12Q 1/6827 (2018.01)
(52) CPC특허분류
C12Q 1/6883 (2022.01)
C12Q 1/6827 (2018.05)
(21) 출원번호 10-2020-7012372
(22) 출원일자(국제) 2018년10월01일
심사청구일자 2021년09월23일
(85) 번역문제출일자 2020년04월28일
(65) 공개번호 10-2020-0058527
(43) 공개일자 2020년05월27일
(86) 국제출원번호 PCT/GB2018/052808
(87) 국제공개번호 WO 2019/069067
국제공개일자 2019년04월11일
(30) 우선권주장
62/566,756 2017년10월02일 미국(US)
(56) 선행기술조사문헌
W02016207653 A1*
(뒷면에 계속)

(73) 특허권자
옥스포드 바이오다이나믹스 피엘씨
영국, 오엑스4 2더블유비 옥스퍼드, 옥스퍼드 비
즈니스 파크, 빌딩 3140
(72) 발명자
아쿨리츠체프, 알렉산드르
영국, 옥스포드 옥스포드셔 오엑스1 2엔피, 보몬
트 스트리트 26, 옥스포드 바이오다이나믹스 리미
티드 내
라마다스, 어로울 셀밤
영국, 옥스포드 옥스포드셔 오엑스1 2엔피, 보몬
트 스트리트 26, 옥스포드 바이오다이나믹스 리미
티드 내
(뒷면에 계속)
(74) 대리인
이희숙, 김석만

전체 청구항 수 : 총 6 항

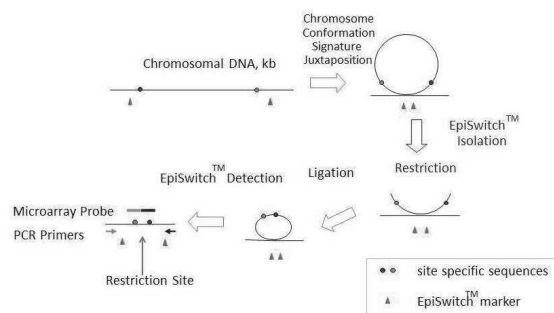
심사관 : 김승범

(54) 발명의 명칭 바이오마커

(57) 요약

ALS 및 헌팅턴병과 관련된 염색체 영역 및 상호 작용을 분석하는 공정.

대표도 - 도1



(52) CPC특허분류

C12Q 2600/106 (2013.01)

C12Q 2600/118 (2013.01)

(72) 발명자

헌터, 이완

영국, 옥스포드 옥스포드셔 오엑스1 2엔피, 보몬트 스트리트 26, 옥스포드 바이오다이내믹스 리미티드 내

솔터, 매튜

영국, 옥스포드 옥스포드셔 오엑스1 2엔피, 보몬트 스트리트 26, 옥스포드 바이오다이내믹스 리미티드 내

(56) 선행기술조사문헌

W02016207647 A1

W02009147386 A1

Journal of the Neurological Sciences
(2017.10.15) 381:561-756

EBioMedicine (2018) 33:6-7

Nature (2016.1.7) 529:110-114

Scientific Reports (2017.5.19) 7:2186

EBioMedicine (2017.6.23) 33:169-184

Biomarkers in Medicine (2010) 4(4):611-629

*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

염색체 상태와 관련된 염색체 상호 작용들이 인간 개체에 존재하는지 여부를 결정하는 단계를 포함하는, 집단에서의 질병 서브그룹을 나타내는 염색체 상태를 검출하는 공정으로서,

상기 질병 서브그룹은 근위축성측삭경화증(ALS) 서브그룹이고;

적어도 이하의 모든 염색체 상호작용의 존재 또는 부존재가 결정되며:

(a) 염색체 위치 번호(chromosome position number) 127483440 부터 127483469, 및 127600019 부터 127600048에 의해 형성된 6번 염색체의 염색체 상호작용;

(b) 염색체 위치 번호 29442590 부터 29442619, 및 29482083 부터 29482112에 의해 형성된 22번 염색체의 염색체 상호작용;

(c) 염색체 위치 번호 112765788 부터 112765817, 및 112810767 부터 112810796에 의해 형성된 2번 염색체의 염색체 상호작용;

(d) 염색체 위치 번호 98282669 부터 98282698, 및 98346932 부터 98346961에 의해 형성된 13번 염색체의 염색체 상호작용;

(e) 염색체 위치 번호 66441278 부터 66441307, 및 66481281 부터 66481310에 의해 형성된 17번 염색체의 염색체 상호작용;

(f) 염색체 위치 번호 6163839 부터 6163868, 및 6195954 부터 6195983에 의해 형성된 17번 염색체의 염색체 상호작용;

(g) 염색체 위치 번호 95420909 부터 95420938, 및 95465339 부터 95465368에 의해 형성된 7번 염색체의 염색체 상호작용;

(h) 염색체 위치 번호 230738574 부터 230738603, 및 230752059 부터 230752088에 의해 형성된 1번 염색체의 염색체 상호작용;

(i) 염색체 위치 번호 105657479 부터 105657508, 및 105741523 부터 105741552에 의해 형성된 7번 염색체의 염색체 상호작용;

(j) 염색체 위치 번호 146728708 부터 146728737, 및 146792792 부터 146792821에 의해 형성된 7번 염색체의 염색체 상호작용;

(k) 염색체 위치 번호 66302476 부터 66302505, 및 66512972 부터 66513001에 의해 형성된 10번 염색체의 염색체 상호작용;

(l) 염색체 위치 번호 105638873 부터 105638902, 및 105814875 부터 105814904에 의해 형성된 8번 염색체의 염색체 상호작용;

(m) 염색체 위치 번호 58334020 부터 58334049, 및 58485551 부터 58485580에 의해 형성된 15번 염색체의 염색체 상호작용;

(n) 염색체 위치 번호 58334020 부터 58334049, 및 58540854 부터 58540883에 의해 형성된 15번 염색체의 염색체 상호작용;

(o) 염색체 위치 번호 79009348 부터 79009377, 및 79284592 부터 79284621에 의해 형성된 7번 염색체의 염색체 상호작용;

(p) 염색체 위치 번호 123969419 부터 123969448, 및 124085755 부터 124085784에 의해 형성된 8번 염색체의 염색체 상호작용;

(q) 염색체 위치 번호 56538371 부터 56538400, 및 56570116 부터 56570145에 의해 형성된 X 염색체의 염색체 상호작용;

상기 각 염색체 상호작용의 존재 또는 부존재는 이하의 단계를 포함하는 방법에 의해 결정되는, 집단에서의 질병 서브그룹을 나타내는 염색체 상태를 검출하는 공정:

(1) 염색체 상호작용으로 함께 모인 상기 개체의 염색체 영역의 시험관 내 가교(in vitro cross-linking) 단계;

(2) 상기 가교된 DNA를 절단하는 단계;

(3) 연결된(ligated) 핵산을 형성하기 위해 상기 가교되고 절단된 DNA 말단을 연결하는(ligating) 단계로서, 상기 연결된 핵산의 존재 또는 부존재의 검출은 염색체 상호작용의 존재 또는 부존재의 결정에 사용될 수 있는 것인, 상기 가교되고 절단된 DNA 말단을 연결하는 단계.

청구항 2

제1항에 있어서, ALS를 진단하거나 ALS에 대한 예후를 결정하기 위해 수행되는 것을 특징으로 하는 공정.

청구항 3

제1항에 있어서, 상기 연결된 핵산의 존재 또는 부존재의 검출은, PCR 반응 동안 연결된 핵산을 증폭시킬 수 있는 프라이머 및 연결 부위에 결합하는 프로브를 사용하는 정량적 PCR (qPCR)에 의한, 연결된 핵산의 특정 검출을 포함하고,

상기 프로브는 염색체 상호 작용에서 함께 모인 각각의 염색체 영역으로부터의 서열에 상보적인 서열을 포함하는 것을 특징으로 하는 공정.

청구항 4

제3항에 있어서, 상기 프로브는 다음을 포함하는 것을 특징으로 하는 공정:

상기 프로브의 5' 말단에 공유 결합된 형광 단(fluorophore), 또는

상기 프로브의 3' 말단에 공유 결합된 소광제를 포함.

청구항 5

제4항에 있어서,

상기 형광단은 HEX, 텍사스 레드 또는 FAM인 것을 특징으로 하는 공정.

청구항 6

제4항에 있어서,

상기 프로브는 길이 10 내지 40개 뉴클레오티드 염기의 핵산 서열을 포함하는 것을 특징으로 하는 공정.

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

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

발명의 설명

기술 분야

[0001] 본 발명은 염색체 상호 작용을 검출하는 것에 관한 것이다.

배경 기술

[0002] 바이오마커를 통해 질병 특성을 식별할 수 있다. 본 발명에 일반적으로 사용되는 바이오마커는 RNA 발현 패턴 및 단백질 마커를 포함한다.

발명의 내용

해결하려는 과제

[0003] 유전자좌에서의 특정 염색체 변형 시그니처 (CCS)는 병리 또는 치료와 관련된 조절 후성 유전적 조절 설정으로 인해 존재하거나 부재한다. CCS는 경미한 오프율(mild off-rates)을 가지고 있으며 특정 표현형 또는 병리를 나타낼 때 새로운 표현형으로의 생리학적 신호 전이 또는 외부 개입의 결과로만 변경된다. 또한, 이러한 이벤트의 측정은 바이너리(binary) 하므로, 이 판독 값은 다양한 수준의 DNA 메틸화, 히스톤 변형 및 대부분의 넌 코딩(non-coding) RNA의 연속 판독 값과 완전히 대조된다. 현재까지 대부분의 분자 바이오 마커에 사용된 연속체 판독은 특정 바이오 마커에 대한 변화의 크기가 환자마다 크게 다르기 때문에 데이터 분석에 어려움을 주며, 이는 환자의 집단을 계층화하는데 사용될 때 분류 통계에 문제를 일으킨다. 이러한 분류 통계는 규모가 없는 바이오 마커를 사용하는 것이 더 적합하고 표현형 차이의 “예 또는 아니오” 라는 바이너리(binary) 점수만 제공한다. - 염색체 형태(EpiSwitch™) 바이오 마커는 잠재적 진단, 예후 및 예측 바이오마커를 위한 탁월한 자원임을 나타낸다.

[0004] 본 발명자들은 한 집단내에서 서브그룹을 식별 할 수 있는 접근법을 사용하여 근 위축성 측삭 경화증 (ALS) 또는 헌팅턴병과 관련이 있는 염색체 상호 작용이 있는 게놈 영역을 확인하였다. 따라서 본 발명은 염색체 상호 작용이 게놈의 정의된 질환 관련 영역 내에 존재하는지 여부를 결정하는 것을 포함하는, 집단에서의 서브그룹을 나타내는 염색체 상태를 검출하는 방법을 제공하며, 여기서 상기 질환은 ALS 또는 헌팅턴병이다. 염색체 상호 작용은 임의의 염색체 상호 작용이 집단의 ALS 또는 헌팅턴병 서브그룹에 상응하는 염색체 상태와 관련이 있는지를 결정하는 방법에 의해 선택적으로 식별되거나 식별이 가능 (또는 유도 가능)하며, 이는 제 2 인덱스 핵산 세트를 갖는 상이한 상태의 염색체를 갖는 서브 그룹과 제 1 핵산 세트가 접촉하는 것을 포함하고, 상보적 서열이 혼성화 되도록 허용하고, 핵산 제 1 및 제 2 세트에서 핵산은 염색체 상호 작용에 함께 모인 염색체 영역 둘로부터의 서열을 포함하는 연결된 생성물을 나타내고, 핵산 제 1 및 제 2 세트 사이의 혼성화 패턴은 ALS 또는 헌팅턴병 서브그룹에 특이적인 염색체 상호 작용의 결정을 허용한다. ALS 또는 헌팅턴병 서브그룹은 진단 (ALS 또는 헌팅턴병의 존재) 또는 예후 (예 : ALS 또는 헌팅턴병의 진행률)와 관련 될 수 있다. 본원 발명에 기재된 임의의 특정 관련 염색체 상호 작용 (마커)은 마커의 조합을 포함하여 본 발명의 기초로서 사용될 수 있다.

[0005] 본 발명은 염색체 상태와 관련된 염색체 상호 작용이 게놈의 정의된 영역 내에 존재하는지 여부를 결정하는 것을 포함하는, 집단에서 질병 서브 그룹을 나타내는 염색체 상태를 검출하는 공정을 제공하며, 여기서 상기 질병

서브 그룹은 근 위축성 측삭 경화증(ALS) 서브그룹이고;

[0006] -상기 염색체 상호 작용은 임의의 염색체 상호 작용이 집단의 ALS 서브 그룹에 상응하는 염색체 상태와 관련이 있는지를 결정하는 방법에 의해 선택적으로 식별되었으며, 서브 그룹으로부터의 제 1 핵산 세트를 제 2 인덱스 핵산 세트와 함께 상이한 상태의 염색체와 접촉시키는 단계를 포함하고, 상보적 서열이 혼성화하는 것을 허용하며, 핵산 제 1 및 제 2 세트의 핵산은 염색체 상호 작용에서 함께 모이는 염색체 영역 둘 다로부터의 서열을 포함하는 연결된 생성물을 나타내고, 제1 핵산세트와 제2 핵산 세트 사이의 혼성화 패턴은 ALS 서브그룹에 특이적인 염색체 상호 작용의 결정을 허용하며;

[0007] -상기 염색체 상호 작용은

[0008] (i) 표 1 또는 5에 열거된 영역 또는 유전자 중 어느 하나에 존재하고/하거나;

[0009] (ii) 표 1 또는 5에 나타난 임의의 프로브로 표시되는 염색체 상호 작용 중 어느 하나에 상응하고/하거나,

[0010] (iii) 표 10 또는 11에 나타난 염색체 상호 작용 중 어느 하나에 상응하고/하거나,

[0011] (iv) (i), (ii) 또는 (iii)을 포함하거나 측면(flanks)에 위치하는 4,000개의 염기 영역에 존재한다.

[0012] 본 발명은 또한 상기 염색체 상태와 관련된 염색체 상호 작용이 게놈의 정의 된 영역 내에 존재하는지 여부를 결정하는 것을 포함하는, 집단에서 질병 서브 그룹을 나타내는 염색체 상태를 검출하는 방법을 제공하며, 여기서 상기 질병 서브 그룹은 헌팅턴병 서브그룹과;

[0013] -상기 염색체 상호 작용은 임의의 염색체 상호 작용이 집단의 헌팅턴병 서브그룹에 상응하는 염색체 상태와 관련이 있는지를 결정하는 방법에 의해 선택적으로 식별되었으며, 서브그룹으로부터의 제 1 핵산 세트를 제 2 인덱스 핵산 세트와 함께 상이한 상태의 염색체와 접촉시키는 단계를 포함하고, 상보적 서열이 혼성화하는 것을 허용하며, 여기서 제 1 및 제 2 핵산 세트의 핵산은 염색체 상호 작용에서 함께 모이는 염색체 영역 둘 다로부터의 서열을 포함하는 연결된 생성물을 나타내고, 제 1 핵산 세트와 제 2 핵산 세트 사이의 혼성화 패턴은 헌팅턴병 서브그룹에 특이적인 염색체 상호 작용의 결정을 허용하고;

[0014] -상기 염색체 상호 작용은 :

[0015] (i) 표 12에 열거 된 영역 또는 유전자 중 어느 하나에 존재하고/하거나;

[0016] (ii) 표 12에 나타난 임의의 프로브로 표시되는 염색체 상호 작용 중 어느 하나에 상응하고/하거나,

[0017] (iii) 표 12에 나타난 염색체 상호 작용 중 어느 하나에 상응하고/하거나,

[0018] (iv) (i), (ii) 또는 (iii)을 포함하거나 측면(flanks)에 위치하는 4,000개의 염기 영역에 존재한다.

과제의 해결 수단

[0019] 발명의 공정

[0020] 본 발명의 공정은 ALS 또는 헌팅턴병과 관련된 염색체 상호 작용을 검출하기위한 타이핑(typing) 시스템을 포함한다. 이러한 타이핑은 염색체 상호 작용에 함께 모인 염색체의 가교(cross-linking) 영역을 기초로 하는 본원 발명에 언급된 EpiSwitch™ 시스템을 사용하고, 염색체 DNA를 절단하는 단계, 염색체 상호 작용을 형성한 두 영역으로부터의 서열을 갖는 연결된 핵산을 유도하기 위해 가교된 개체에 존재하는 핵산을 연결시키는 단계에 의해 수행 될 수 있다. 이 연결된 핵산의 검출은 특정 염색체 상호 작용의 존재 또는 부재의 결정을 허용한다.

[0021] 염색체 상호 작용은 제 1 및 제 2 핵산 집단이 사용되는 상기 기재된 방법을 사용하여 식별될 수 있다. 이러한 핵산은 EpiSwitch™ 기술을 사용하여 생성 될 수도 있다.

[0022] 발명과 관련된 후생유전학적 상호 작용

[0023] 본원 발명에서 사용되는 용어 "후생유전학적" 및 "염색체" 상호 작용은 일반적으로 염색체의 원위 영역 사이의 상호 작용을 말하며, 상기 상호 작용은 염색체 영역의 상태에 따라 동적이고 변화되고, 형성되거나 또는 파괴된다.

[0024] 본 발명의 특정 공정에서, 염색체 상호 작용은 먼저 상호 작용의 일부인 염색체의 양쪽 영역으로부터 서열을 포함하는 연결된 핵산을 생성함으로써 검출된다. 이러한 공정에서, 상기 영역은 임의의 적합한 수단에 의해 가교될 수 있다. 바람직한 구체화된 예에서, 상호 작용은 포름 알데히드(formaldehyde)를 사용하여 가교되거나, 임

의의 알데히드(aldehyde), 또는 D - 바이오틴 - e - 아미노카프로산 - N - 히드록시숙신이미드 에스터 (D-Biotinoyl-e- aminocaproic acid-N-hydroxysuccinimide ester) 또는 디곡시제닌 - 3 - 0 - 메틸카르보닐 - e - 아미노카프로산 - N - 하이드록시숙신이미드 에스터(Digoxigenin - 3 - 0 - methylcarbonyl-e-aminocaproic acid-N-hydroxysuccinimide este)에 의해 가교 될 수 있다. 파라-포름알데히드는 4 옹스트롬 떨어진 DNA 사슬을 가교시킬 수다. 바람직하게는 염색체 상호 작용은 동일한 염색체 상에 있고 선택적으로 2 내지 10 옹스트롬 떨어져 있다.

[0025] 염색체 상호 작용은 예를 들어 생리학적 조건의 변화에 반응하여 전사되거나 억제되는 경우에 염색체 영역의 상태를 반영 할 수 있다. 본원 발명에 정의된 바와 같은 서브그룹에 특이적인 염색체 상호 작용은 안정한 것으로 밝혀져서, 2개의 서브그룹 사이의 차이를 측정하는 신뢰할 수 있는 수단을 제공한다.

[0026] 또한, 한 특성(질병 상태와 같은)에 특이적인 염색체 상호 작용은 일반적으로 예를 들어 메틸화 또는 히스톤 단백질의 결합 변화와 같은 다른 후생유전학적 마커와 비교하여 생물학적 공정에서 초기에 발생할 것이다. 따라서 본 발명의 공정은 생물학적 공정의 초기 단계를 검출 할 수 있다. 이것은 결과적으로 더 효과적 일 수 있는 조기 개입(예 : 치료)을 가능하게 한다. 더우기 동일한 서브 그룹 내에서 개체 간의 관련 염색체 상호 작용에는 거의 변화가 없다. 염색체 상호 작용을 탐지하는 것은 유전자 당 최대 50 가지의 서로 다른 가능한 상호 작용에 대해 매우 유익한 정보이므로, 본 발명의 공정은 500,000개의 다른 상호 작용을 조사 할 수 있다.

[0027] *신호 마커 세트*

[0028] 본 명세서에서 용어 "마커" 또는 "바이오 마커"는 본 발명에서 검출 될 수 있는 특정 염색체 상호 작용을 지칭한다. 특정 마커가 본원 발명에 개시되어 있으며, 이들 중 임의의 것이 본 발명에 사용될 수 있다. 추가의 마커 세트는 예를 들어 본원 발명에 개시된 조합 또는 수로 사용될 수 있다. 본원 발명의 표에 개시된 마커가 바람직하다. 이들은 임의의 적합한 방법, 예를 들어 qPCR 방법을 포함하여 본원 발명에 개시된 PCR 또는 프로브 기반 방법에 의해 유형화(typed) 될 수 있다. 상기 마커는 본원 발명에서 위치 또는 프로브 및 / 또는 프라이머 서열에 의해 정의된다.

[0029] *후생 유전학적 상호 작용의 위치와 원인*

[0030] 후생 유전적 염색체 상호 작용은 중복되고, 관련되거나 설명되지 않은 유전자를 암호화하는 것으로 나타난 염색체 영역을 포함 할 수 있지만, 유전자 간 영역에도 동일하게 존재할 수 있다. 또한, 본 발명자들은 모든 영역에서 후생유전학적 상호 작용이 염색체 유전자좌의 상태를 결정하는데 똑같이 중요하다는 것을 발견하였다. 이들 상호 작용은 유전자좌에 위치한 특정 유전자의 암호화(coding) 영역에 반드시 존재하는 것은 아니며 유전자 간 영역에 존재할 수 있다.

[0031] 본 발명에서 검출되는 염색체 상호 작용은 환경적 요인, DNA 메틸화, 넌코딩 안티센스 RNA 전사체, 비 돌연변이성 발암 물질, 히스톤 변형, 염색질 리모델링 및 특정 국소 DNA 상호 작용에 의한 기본 DNA 서열의 변화에 의해 야기 될 수 있다. 염색체 상호 작용을 야기하는 변화는 기본 핵산 서열의 변화에 의해 야기 될 수 있으며, 그 자체는 유전자 산물 또는 유전자 발현 방식에 직접 영향을 미치지 않는다. 이러한 변화는 예를 들어 유전자 내부 및 / 또는 외부의 SNP, 유전자 간 융합 및 / 또는 유전자 간 DNA, 마이크로 RNA 및 넌코딩 RNA의 결실 일 수 있다. 예를 들어, 대략 20%의 SNP가 넌코딩 영역에 있는 것으로 알려져 있으므로, 기술된 바와 같은 공정은 넌코딩 상황에서도 유익하다. 일 구체 예에서, 상호 작용을 형성하기 위해 함께 모인 염색체의 영역은 동일한 염색체에서 5kb, 3kb, 1kb, 500개 염기쌍 또는 200개 염기쌍 미만으로 떨어져 있다.

[0032] 검출되는 염색체 상호 작용은, 바람직하게는, 표 1 또는 5에 언급된 임의의 유전자 내에 있다. 그러나 이는 또한 유전자의 상류 또는 하류, 예를 들어 최대 50,000, 최대 30,000, 최대 20,000, 최대 10,000 또는 최대 5000 염기 유전자 또는 코딩 서열로부터의 상류 또는 하류에, 존재할 수 있다.

[0033] *서브그룹, 진단 및 맞춤 치료*

[0034] 본 발명의 목적은 ALS 또는 헌팅턴병 서브그룹과 관련된 염색체 상호 작용의 검출을 허용하는 것이다. 따라서 이 공정은 ALS 또는 헌팅턴병의 진단에 사용되거나 사용되지 않을 수 있다. 이 공정은 ALS 또는 헌팅턴병의 예후에 사용되거나 사용되지 않을 수 있다. 이 공정이 ALS의 진단에 사용되는 경우, 표 1과 관련된 마커의 타이핑이 바람직하다(즉, 구체적으로 개시된 마커 및 표 1에 개시된 유전자, 영역 및 측면 영역의 마커). 진단과 관련된 일 실시 형태에서, 표 1과 관련된 마커만이 유형화(typed)되고 다른 마커는 없다. 진단과 관련된 다른 구체적 예에서, (예를 들어, 나열된 프로브 서열로 표시되는) 표 5와 관련된 적어도 1, 2, 3, 4, 5 개 이상의 마커

는 유형화(typed)되지 않는다.

- [0035] 공정이 예후에 사용되는 경우, 표 5와 관련된 마커의 타이핑이 바람직하다(즉, 구체적으로 개시된 마커 및 표 5에 개시된 유전자, 영역 및 측면 영역의 마커). 예후와 관련된 일 실시 형태에서, 표 5와 관련된 마커 만이 유형화되고 다른 마커는 없다. 예후에 관한 다른 구체 예에서(예를 들어, 나열된 프로브 서열로 표시되는) 표 1에 관련된 적어도 1, 2, 3, 4, 5 개 이상의 마커는 유형화(typed)되지 않는다.
- [0036] 전형적으로 '예후'는 ALS의 진행과 관련이 있으며, 개체를 진행률 서브그룹으로 나눌 수 있다. 진행은 ALS-FRS-R 점수(ALS 기능 평가 척도)를 사용하여 측정 될 수 있으며, 개체는 특정 값 이상 또는 이하, 예를 들어 30 일당 0.5 포인트 감소 이상 또는 이하로, 분류 될 수 있다. 이를 통해 예측 예후를 수행 할 수 있다.
- [0037] 본원 발명에 사용된 "서브그룹"은, 바람직하게는, 집단 서브그룹(집단의 서브 그룹), 보다 바람직하게는, 특정 진행 생물 또는 포유 동물과 같은 특정 동물 집단의 서브그룹(예 : 인간, 비인간, 비인간 영장류, 또는 설치류, 예 : 마우스 또는 쥐)을 지칭한다. 가장 바람직하게는, "서브그룹"은 인간 집단의 서브그룹을 지칭한다.
- [0038] 본 발명은 집단에서 특정 서브그룹을 검출하고 치료하는 것을 포함한다. 본 발명자들은 주어진 집단에서 염색체 상호 작용이 서브세트(예를 들어, 적어도 2개의 서브세트)마다 상이하다는 것을 발견했다. 이러한 차이점을 식별하면 의사가 상기 공정에서 설명한대로 환자를 인구의 한 서브세트의 일부로 분류할 수 있다. 따라서 본 발명은 후생유전학적 염색체 상호 작용에 기초하여 환자에게 의약을 맞춤화하는 공정을 의사에게 제공한다.
- [0039] 한 실시 양태에서, 서브그룹(예를 들어, 예후에 관한)은, 예를 들어 Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond B. 및 Nakanishi, A. (1999) ALSFRS-R : 호흡 기능 평가를 통합한 수정된 ALS 기능 등급 척도. BDNF ALS 연구 그룹(단계 III), Journal of the Neurological Sciences. 169(1-2), 13-21.,에 기술된 바와 같이 ALS 기능 등급 척도에 의해 정의된다.
- [0040] 한 실시 양태에서, 서브그룹(예를 들어, 예후 관련)은, 예를 들어 Talakad, N.S., Pradhan, C., Nalini, A., Thennarasu, K. 및 Raju T.R. (2009) 근 위축성 측삭 경화증의 폐 기능 평가, *Indian J Chest Dis Allied Sci.* 51(2): 87-91.,에 기술된 바와 같이 강제 활력 용량 (FVC)으로 정의된다.
- [0041] 연결된 핵산 생성
- [0042] 본 발명의 특정 구체적인 예는 연결된 핵산, 특히 연결된 DNA를 사용하는 것이다. 이들은 염색체 상호 작용으로 함께 모인 두 영역으로부터의 서열을 포함하므로 상호 작용에 대한 정보를 제공한다. 본원 발명에 기재된 EpiSwitch™ 방법은 염색체 상호 작용을 검출하기 위해 이러한 연결된 핵산의 생성을 사용하는 것이 기재되어있다.
- [0043] 따라서, 본 발명의 공정은 하기 단계(이들 단계를 포함하는 방법 포함)에 의해 연결된 핵산 (예를 들어, DNA)을 생성하는 단계를 포함 할 수 있다. :
- [0044] (i) 염색체 유전자좌에 존재하는 후생유전학적 염색체 상호 작용의 가교(cross-linking), 바람직하게는, 시험관 내;
- [0045] (ii) 선택적으로, 상기 염색체 유전자좌로부터 가교된(cross-linked) DNA를 분리하는 단계;
- [0046] (iii) 상기 가교된 DNA를, 예를 들어, 적어도 한 번 이상 절단하는 효소 (특히, 상기 염색체 유전자좌 내에서 적어도 한 번 절단하는 효소)에 의한 제한 소화에 의해, 절단하는 단계;
- [0047] (iv) 상기 가교된 절단된 DNA 말단을 연결시키는 단계 (특히 DNA 루프를 형성하기 위해)와;
- [0048] (v) 선택적으로, 특정 염색체 상호 작용의 존재를 식별하기 위해, 특히 PCR (중합 효소 연쇄 반응)과 같은 기술을 사용하여, 상기 연결된 DNA 및 / 또는 상기 DNA 루프의 존재를 식별하는 단계를 포함한다.
- [0049] 이들 단계는, 개체가 ALS 또는 헌팅턴병 서브그룹의 일부인지를 결정하기 위해, 본원 발명에 언급된 임의의 구체적인 예에 대한 염색체 상호 작용을 검출하기 위해 수행 될 수 있다. 상기 단계는 또한 본원 발명에 언급된 제 1 및 / 또는 제 2 핵산 세트를 생성하기 위해 수행 될 수 있다.
- [0050] 중합 효소 연쇄 반응(PCR)은 연결된 핵산을 검출 또는 식별하는데 사용될 수 있으며, 예를 들어 생성된 PCR 생성물의 크기는 존재하는 특정 염색체 상호 작용을 나타낼 수 있으므로 유전자좌 상태의 식별에 사용될 수 있다. 바람직한 구체 예에서, 표 2 또는 7에 나타낸 바와 같이 적어도 1 또는 2, 3, 4, 5, 6, 7 또는 8개 이상의 프라이머 또는 프라이머 쌍이 PCR 반응에 사용된다. 다른 바람직한 구체적인 예에서, 임의의 다른 표에 나타낸 바와

같이 적어도 1, 2, 3, 4, 5, 6, 7 또는 8개 이상의 프라이머 또는 프라이머 쌍이 PCR 반응에 사용된다. 당업자는 관심 염색체 유전자좌 내에서 DNA를 절단하는 데 사용될 수 있는 다양한 제한 효소를 인지 할 수 있다. 사용된 특정 효소는 연구된 유전자좌 및 그 안에 위치한 DNA의 서열에 의존 할 것임이 명백할 것이다. 본 발명에 기술된 바와 같이 DNA를 절단하는데 사용될 수 있는 제한 효소의 비 제한적인 예는 TaqI이다.

[0051] *EpiSwitch™ 기술과 같은 실시 예*

[0052] EpiSwitch™ 기술은 또한 표현형에 특이적인 후생유전학적 염색체 형태 시그너처의 검출에서 마이크로어레이 (microarray) EpiSwitch™ 마커 데이터를 사용하는 것과 관련이 있다. 본원 발명에 기재된 방식으로 연결된 핵산을 사용하는 EpiSwitch™와 같은 실시 양태는 몇 가지 이점을 갖는다. 이들은, 예를 들어 본 발명의 제 1 핵산 세트로부터의 핵산 서열이 제 2 핵산 세트와 혼성화되거나 혼성화되지 않기 때문에, 낮은 수준의 확률론적 잡음을 갖는다. 이것은 후생 유전적 수준에서 복잡한 메커니즘을 측정하는 비교적 간단한 방법을 허용하는 이진법 (binary) 결과를 제공한다. EpiSwitch™ 기술은 또한 빠른 처리 시간과 저렴한 비용을 가지고 있다. 일 구체 예에서, 공정 시간은 3 시간 내지 6 시간이다.

[0053] *시료 및 시료 처리*

[0054] 본 발명의 공정은 일반적으로 시료에서 수행 될 것이다. 시료는 일반적으로 개체의 DNA를 포함할 것이다. 일반적으로 세포를 포함할 것이다. 일 구체 예에서, 시료는 최소 침습적 수단에 의해 취득되며, 예를 들어 혈액 샘플 일 수 있다. DNA는 표준 제한 효소로 추출 및 절단 될 수 있다. 이를 통해 어떤 염색체 형태가 유지되고 EpiSwitch™ 플랫폼에서 탐지될지 미리 결정할 수 있다. 수평 전달을 포함하는 조직과 혈액 사이의 염색체 상호 작용의 동기화로 인해, 혈액 시료는 질병과 관련된 조직과 같은 조직에서 염색체 상호 작용을 검출 하는데 사용될 수 있다. 암과 같은 특정 조건의 경우, 돌연변이로 인한 유전 적 잡음은 관련 조직의 염색체 상호 작용 '신호'에 영향을 줄 수 있으므로, 혈액을 사용하는 것이 유리하다.

[0055] *본 발명의 핵산의 특성*

[0056] 본 발명은 본 발명의 공정에서 사용되거나 생성되는 것으로 본원 발명에 기술되어 있는 연결된 핵산과 같은 특정 핵산에 관한 것이다. 이들은 본원 발명에 언급된 제 1 및 제 2 핵산과 동일하거나 임의의 특성을 가질 수 있다. 본 발명의 핵산은 전형적으로 염색체 상호 작용에서 함께 모인 염색체의 두 영역 중 하나의 서열을 각각 포함하는 두 부분을 포함한다. 전형적으로 각각의 부분은 적어도 8, 10, 15, 20, 30 또는 40개 이상의 뉴클레오티드 길이, 예를 들어 10 내지 40개의 뉴클레오티드 길이이다. 바람직한 핵산은 임의의 표에 언급된 임의의 유전자로부터의 서열을 포함한다. 전형적으로 바람직한 핵산은 표 1 또는 5에 언급된 특정 프로브 서열이거나, 단편 및 / 또는, 이러한 서열과 상동체를 포함한다. 다른 바람직한 핵산은 표 10, 11 또는 12에 언급된 특정 프로브 서열이거나, 또는 단편 및 / 또는 이러한 서열의 상동체를 포함한다. 바람직하게는, 핵산은 DNA이다. 특정 서열이 제공되는 경우, 본 발명은 특정 실시 양태에서 요구되는 상보적 서열을 사용할 수 있는 것으로 이해된다.

[0057] 표 2 또는 7에 나타난 프라이머는 또한 본 발명에서 언급 된 바와 같이 사용될 수 있다. 일 구체 예에서, 프라이머는 표 2 또는 7에 나타난 서열이거나, 단편 및 / 또는 표 2 또는 7에 나타난 임의의 서열의 상동체를 포함하는 서열이 사용될 수 있다.

[0058] *제 2 핵산 세트 - '인덱스' 서열*

[0059] 제 2 핵산 세트의 서열은 일련의 인덱스 서열의 기능을 가지며, 본질적으로 서브그룹 특이적 서열을 식별하기에 적합한 핵산 서열의 세트이다. 그들은 '백그라운드' 염색체 상호 작용을 나타낼 수 있으며 어떤 방식으로 선택되거나 선택되지 않을 수 있다. 그들은 일반적으로 가능한 모든 염색체 상호 작용의 부분 집합이다.

[0060] 제 2 핵산 세트는 임의의 적합한 공정에 의해 유도될 수 있다. 이들은 계산적으로 유도될 수 있거나 개체의 염색체 상호 작용에 기초될 수 있다. 이들은 일반적으로 제 1 핵산 세트보다 더 큰 집단을 나타낸다. 일 특정 구체 양태에서, 제 2 핵산 세트는 특정 세트의 유전자에서 가능한 모든 후생유전학적 염색체 상호 작용을 나타낸다. 또 다른 특정 구체 양태에서, 제 2 핵산 세트는 본원 발명에 기재된 집단에 존재하는 모든 가능한 후생유전학적 염색체 상호 작용의 큰 비율을 나타낸다. 일 특정 구체 예에서, 제 2 핵산 세트는 적어도 20, 50, 100 또는 500개 이상의 유전자, 예를 들어 20 내지 100 또는 50 내지 500 개의 유전자에서, 후생유전학적 염색체 상호 작용의 적어도 50% 이상 또는 적어도 80% 이상을 나타낸다.

- [0061] 제 2 핵산 세트는 전형적으로 집단에서 질환 상태 / 표현형을 변형, 조절 또는 임의의 방식으로 매개하는 적어도 100개 이상의 가능한 후생유전학적 염색체 상호 작용을 나타낸다. 제 2 핵산 세트는 종의 질병 상태(전형적으로 진단 또는 예후와 관련된)에 영향을 미치는 염색체 상호 작용을 나타낼 수 있다. 제 2 핵산 세트는 전형적으로 ALS 서브그룹과 관련이 있거나 관련이 없는 후생유전학적 상호 작용을 나타내는 서열을 포함한다.
- [0062] 일 특정 구체 예에서, 제 2 핵산 세트는 집단에서 자연적으로 발생하는 서열로부터 적어도 부분적으로 유래하며, 전형적으로 인 실리코(*in silico*) 공정에 의해 수득된다. 상기 핵산은 자연 발생 핵산에 존재하는 핵산의 상응하는 부분과 비교하여 단일 또는 다중 돌연변이를 추가로 포함 할 수 있다. 돌연변이는 하나 또는 하나 이상의 뉴클레오티드 염기쌍의 결실, 치환 및 / 또는 첨가를 포함한다. 일 특정 구체 예에서, 제 2 핵산 세트는 자연적으로 발생하는 종에 존재하는 핵산의 상응하는 부분에 대해 적어도 70% 이상의 서열 상동성을 갖는 상동체 및 / 또는 오르톨로그(orthologue)를 나타내는 서열을 포함 할 수 있다. 다른 특정 구체 예에서, 자연적으로 발생하는 종에 존재하는 핵산의 상응하는 부분에 대해 적어도 80% 이상의 서열 상동성 또는 적어도 90% 이상의 서열 상동성이 제공된다.
- [0063] 제 2 핵산 세트의 특성
- [0064] 일 특정 구체 예에서, 제 2 핵산 세트는 적어도 100개 이상의 상이한 핵산 서열, 바람직하게는, 적어도 1000, 2000 또는 5000개 이상의 상이한 핵산 서열, 최대 100,000, 1,000,000 또는 10,000,000개의 상이한 핵산 서열이 존재한다. 전형적인 수는 100 내지 1,000,000, 예컨대 1,000 내지 100,000개의 상이한 핵산 서열 일 것이다. 전부 또는 적어도 90% 이상 또는 적어도 50% 이상 또는 이들은 상이한 염색체 상호 작용에 상응 할 것이다
- [0065] 일 특정 구체 예에서, 제 2 핵산 세트는 적어도 20개 이상의 상이한 유전자좌 또는 유전자, 바람직하게는 적어도 40개 이상의 상이한 유전자좌 또는 유전자, 보다 바람직하게는 100개 이상, 500개 이상, 1000개 이상 또는 5000개 이상의 상이한 유전자좌 또는 유전자, 예컨대 100개 내지 10,000 개의 상이한 유전자좌 또는 유전자 내의 염색체 상호 작용을 나타낸다. 제 2 핵산 세트의 길이는 서브그룹에 특이적인 염색체 상호 작용의 식별을 허용하기 위해 제 1 핵산 세트와 왓슨 크릭 염기 쌍에 따라 특이적으로 혼성화하는데 적합하다. 전형적으로, 제 2 핵산 세트는 염색체 상호 작용에서 함께 모인 2개의 염색체 영역 서열에 상응하는 2개의 부분을 포함할 것이다. 제 2 핵산 세트는 전형적으로 적어도 10개 이상, 바람직하게는 20개, 바람직하게는 여전히 30개 염기(뉴클레오티드) 길이의 핵산 서열을 포함한다. 다른 구체 예에서, 핵산 서열은 최대 500개, 바람직하게는 최대 100개, 바람직하게는 최대 50개의 염기쌍 길이 일 수 있다. 바람직한 구체 예에서, 제 2 핵산 세트는 17개 내지 25개의 염기쌍의 핵산 서열을 포함한다. 한 구체 예에서, 제 2 핵산 세트 서열의 적어도 100, 80% 또는 50% 이상은 상기 전술 한 바와 같은 길이를 갖는다. 바람직하게는, 상이한 핵산은 임의의 중첩 서열을 갖지 않으며, 예를 들어 핵산의 적어도 100%, 90%, 80% 또는 50% 이상은 적어도 5개 이상의 인접 뉴클레오티드에 걸쳐 동일한 서열을 갖지 않는다.
- [0066] 제 2 핵산 세트가 하나의 '인덱스'로서 작용하는 경우, 동일한 세트의 제 2 핵산 세트는 상이한 특성에 대한 서브그룹을 나타내는 상이한 제 1 핵산 세트와 함께 사용될 수 있으며, 즉 제 2 핵산 세트는 상이한 특성과 관련된 염색체 상호 작용을 식별하기 위해 사용될 수 있는 '보편적인' 핵산의 수집을 나타낼 것이다.
- [0067] 제 1 핵산 세트
- [0068] 제 1 핵산 세트는 일반적으로 ALS 또는 헌팅턴병의 진단 또는 예후와 관련된 서브그룹에서 나온다. 제 1 핵산은 본원 발명에 언급된 제 2 핵산 세트의 임의의 특징 및 특성을 가질 수 있다. 제 1 핵산 세트는 일반적으로 본원 발명에 기술된 바와 같이 처리 및 가공단계, 특히 EpiSwitch™ 가교 및 절단 단계를 겪는 개체로부터의 시료로부터 유래된다. 전형적으로, 제 1 핵산 세트는 개체로부터 채취한 시료에 존재하는 염색체 상호 작용의 전부 또는 적어도 80% 이상 또는 50% 이상을 나타낸다.
- [0069] 전형적으로, 제 1 핵산 세트는 제 2 핵산 세트에 의해 대표되는 염색체 상호 작용과 비교하여, 제 2 핵산 세트에 의해 대표되는 유전자좌 또는 유전자를 가로 지르는 더 작은 염색체 상호 작용 집단을 나타내며, 즉 제 2 핵산 세트는 정의된 유전자좌 또는 유전자 세트에서 상호 작용의 배경 또는 인덱스 세트를 나타낸다.
- [0070] 핵산 라이브러리
- [0071] 본원 발명에 언급된 임의의 유형의 핵산 집단은 '제 1' 또는 '제 2' 핵산과 같이, 해당 유형의 적어도 200개 이상, 적어도 500개 이상, 적어도 1000개 이상, 적어도 5000개 이상 또는 적어도 10000개 이상의 상이한 핵산을 포함하는 라이브러리의 형태로 존재할 수 있으며, 그러한 라이브러리는 어레이에 결합되는 형태 일 수 있

다.

- [0072] *혼성화*
- [0073] 본 발명은 제 1 핵산 세트 및 제 2 핵산 세트로부터의 전체적으로 또는 부분적으로 상보적인 핵산 서열이 혼성화 되도록 하는 수단을 필요로 한다. 일 구체 예에서, 모든 제 1 핵산 세트는 단일 검정(assay), 즉 단일 혼성화 단계에서 모든 제 2 핵산 세트와 접촉된다. 그러나, 임의의 적합한 검정법(assay)이 사용될 수 있다.
- [0074] *표지된 핵산 및 혼성화 패턴*
- [0075] 본원 발명에 언급된 핵산은 바람직하게는, 성공적인 혼성화의 검출을 돕는 형광 단(형광 분자) 또는 방사성 표지와 같은 독립적 표지를 사용하여 표지될 수 있다. 자외선 아래에서 특정 라벨이 탐지될 수 있다. 예를 들어 본원 발명에 기재된 어레이에서의, 혼성화 패턴은 두 서브그룹 사이의 후생유전학적 염색체 상호 작용의 차이를 나타내므로, 후생유전학적 염색체 상호 작용을 비교하는 공정과 후생유전학적 염색체 상호 작용이 본 발명의 집단 내 서브그룹에 특이적인지 결정하는 과정을 제공한다.
- [0076] 용어 "혼성화 패턴"은 제 1 및 제 2 핵산 세트 사이의 혼성화의 존재 및 부재, 즉 제 1 세트의 특정 핵산이 제 2 세트의 특정 핵산에 혼성화 되어 임의의 특정 검정 또는 기술, 또는 '패턴'이 검출될 수 있는 표면 또는 어레이를 가질 필요성에 제한되지 않는 것을 광범위하게 포함한다.
- [0077] *특정 특성을 가진 서브그룹 선택*
- [0078] 본 발명은 개체내 ALS 또는 헌팅틴병과 관련된 특성의 존재 또는 부재를 검출하기 위하여 염색체 상호 작용, 전형적으로 5 내지 20개 또는 5 내지 500개와 같은 상호 작용, 바람직하게는 20 내지 300개 또는 50 내지 100개의 상호 작용의 존재 또는 부재를 검출하는 것을 포함하는 방법을 제공한다. 바람직하게는, 염색체 상호 작용은 본원 발명에 언급된 임의의 유전자에서의 것들이다. 일 구체 예에서, 유형화(typed)된 염색체 상호 작용은 표 1 또는 5의 핵산으로 나타낸 것들이다. 바람직하게는, 유형화(typed)된 염색체 상호 작용은 표 10, 11 또는 12에서 핵산으로 나타낸 것들이다. 어떤 서브그룹(ALS 또는 대조군)이 각 프로브에 의해 감지되는지를 보여주는 표에서 열은 '루프 감지'로 지칭된다. 알 수 있는 바와 같이 본 발명의 공정은 테스트의 일부로서 ALS 서브그룹 및 / 또는 대조군 서브그룹(non-ALS)을 검출할 수 있다.
- [0079] *테스트 받는 개체*
- [0080] 테스트 받는 개체의 종의 예는 본원 발명에 언급되어 있다. 또한, 본 발명의 공정에서 테스트되는 개체는 어떤 방식으로 선택되었을 수 있다. 예를 들어, 개체는 ALS 또는 헌팅틴병에 감수성이 있다.
- [0081] *선호하는 유전자 영역, 유전자 좌, 유전자 및 염색체 상호 작용*
- [0082] 본 발명의 모든 측면에서 바람직한 유전자 영역, 유전자좌, 유전자 및 염색체 상호 작용은 표, 예를 들어 표 1 및 5에 언급되어 있다. 전형적으로 본 발명의 방법에서 염색체 상호 작용은 표 1 또는 5에 열거된 관련 유전자 중 적어도 1, 2, 3, 4, 5, 6, 7 또는 8개 이상에서 검출된다. 바람직하게는, 표 1 또는 5의 프로브 서열에 의해 표시되는 관련 염색체 상호 작용의 적어도 1, 2, 3, 4, 5, 6, 7 또는 8개 이상의 존재 또는 부재가 검출된다. 염색체 상호 작용은 본원 발명에 언급된 임의의 유전자의 상류 또는 하류는, 예를 들어 암호화(coding)서열로부터, 예를 들어 50kb 상류 또는 20kb 하류 일 수 있다.
- [0083] 바람직하게는, 표 1에서 적어도 5, 8, 10, 15개 이상 또는 모든 염색체 상호 작용이 유형화(typed)된다.
- [0084] 전형적으로, 표 5에서의 적어도 5, 7, 8개 이상 또는 모든 염색체 상호 작용이 유형화(typed)된다.
- [0085] 바람직하게는, 표 10에서 적어도 4, 6개 이상 또는 모든 염색체 상호 작용이 유형화(typed)된다.
- [0086] 일반적으로 표 11에서 적어도 1, 2 개 이상 또는 모든 염색체 상호 작용이 유형화(typed)된다.
- [0087] 바람직하게는, 표 12에서 적어도 4, 6개 이상 또는 모든 염색체 상호 작용이 유형화(typed)된다.
- [0088] 염색체 상호 작용은 질병의 존재를 결정하기 위해 유형화(typed)될 수 있다. 염색체 상호 작용은 개체가 빠르게 또는 느리게 진행되는지를 결정하기 위해 유형화(typed)될 수 있다.
- [0089] 표에 개시된 특정 프로브 및 프라이머 서열(또는 단편과 상동체를 포함하는 유도체)은 본원 발명에 개시된 임의의 타이핑 방법에 사용될 수 있다. 진단 또는 예후 방법의 용도로 제공된다.
- [0090] 일 구체 예에서, 유전자좌(유전자 및 / 또는 염색체 상호 작용이 검출되는 장소 포함)는 CTCF 결합 부위를 포함

할 수 있다. 이것은 전사 억제자 CTCF에 결합 할 수 있는 임의의 서열이다. 이 서열은 유전자좌에서 1, 2 또는 3 개의 카피(copy)로 존재할 수 있는 CCCTC 서열로 구성되거나 포함 할 수 있다. CTCF 결합 부위 서열은 CCGCGNGGNGGCAG 서열을 포함 할 수 있다(IUPAC 표기법으로). CTCF 결합 부위는 염색체 상호 작용의 적어도 100, 500, 1000 또는 4000개 이상의 염기 내에 있거나 표 1 또는 5에 나타난 임의의 염색체 영역 내에 있을 수 있다.

[0091] 일 구체 예에서, 검출된 염색체 상호 작용은 표 1 또는 5에 나타난 임의의 유전자 영역에 존재한다. 공정에서 연결된 핵산이 검출되는 경우, 표 1 또는 5의 임의의 프로브 서열에 나타난 서열이 감지 될 수 있다.

[0092] 따라서, 전형적으로 프로브의 두 영역으로부터의(즉, 염색체 상호 작용의 두 사이트로부터의) 서열이 검출 될 수 있다. 바람직한 구체적 예에서, 프로브는 임의의 표에 도시된 프로브와 동일하거나 상보적인 서열을 포함하거나 이로 구성된 공정에서 사용된다. 일부 구체적 예에서, 표에 나타난 임의의 프로브 서열과 상동인 서열을 포함하는 프로브가 사용된다.

[0093] *본원 발명에 제공된 표*

[0094] 표 1 및 5는 ALS와 관련된 염색체 상호 작용을 나타내는 프로브 (EpiSwitch™ 마커) 데이터 및 유전자 데이터를 보여준다. 프로브 서열은 염색체 상호 작용에서 함께 모인 유전자 영역의 두 부위로부터 생성된 연결된 생성물을 검출하는데 사용될 수 있는 서열, 즉 프로브는 연결된 생성물의 서열에 상보적인 서열을 포함 할 것이다. 처음 두 세트의 시작-끝 위치는 프로브 위치를 보여주고, 두 번째 두 세트의 시작-끝 위치는 관련 4kb 영역을 보여준다. 프로브 데이터 테이블에는 다음과 같은 정보가 제공된다.

[0095] HyperG_Stats: 초기하학적 보강 매개 변수를 기반으로 유전자좌에서 중요한 EpiSwitch™ 마커의 수를 찾을 확률에 대한 p- 값

[0096] 총 프로브 수(Probe Count Total) : 유전자좌에서 테스트된 총 EpiSwitch™ 형태의 총 수

[0097] Probe Count Sig : 유전자좌에서 통계적으로 유의한 것으로 밝혀진 EpiSwitch™ 형태의 수

[0098] FDR HyperG : 다중 테스트 (False Discovery Rate) 수정된 초기하학적 p-값

[0099] Percent Sig : 유전자좌에서 테스트 한 마커 수에 비해 중요한 EpiSwitch™ 마커 백분율

[0100] logFC : 후생유전학적 비율 (FC)의 로그베이스 2

[0101] AveExpr : 모든 어레이 및 채널에서 프로브에 대한 평균 log2-표현

[0102] T : 완화된 t-통계

[0103] p-값 : 원시 p-값

[0104] 조정 p-값(adj. p-value) : 조정 된 p-값 또는 q-값

[0105] B : B-통계량 (lods 또는 B)은 해당 유전자가 차별적으로 발현되는 로그 홀수다.

[0106] FC : 비 로그 폴드 변경

[0107] FC_1 : 0을 중심으로 한 비 로그 폴드 변경

[0108] LS : 이진 값 즉 FC_1 값과 관련이 있다. FC_1 값이 -1.1 미만이면 -1로 설정되고 FC_1 값이 1.1 이상이면 1로 설정된다. 이 값 사이에서 값은 0이다.

[0109] 표 1 및 5는 관련 염색체 상호 작용이 발생하는 것으로 밝혀진 유전자를 보여준다. 다른 표는 유사한 데이터를 보여준다. 유전자좌 표의 p-값은 HyperG_Stats와 동일합니다 (초기하학적 보강 매개 변수를 기반으로 유전자좌에서 중요한 EpiSwitch™ 마커를 찾을 확률에 대한 p-값).

[0110] 프로브는 Taq1 사이트에서 30bp 떨어져 있도록 설계되어 있다. PCR의 경우, PCR 프라이머는 또한 연결된 생성물을 검출하도록 설계되었지만 Taq1 부위로부터의 위치는 다양하다.

[0111] 프로브 위치 :

[0112] 시작 1 - 단편 1에서 TaqI 사이트의 30 염기 상류

[0113] 끝 1 - 단편 1의 TaqI 제한 사이트

- [0114] 시작 2 - 단편 2의 TaqI 제한 사이트
- [0115] 끝 2 - 단편 2에서 TaqI 부위의 30 염기 하류
- [0116] 4kb 서열 위치 :
- [0117] 시작 1 - 단편 1에서 TaqI 사이트의 4000 염기 상류
- [0118] 끝 1 - 단편 1의 TaqI 제한 사이트
- [0119] 시작 2 - 단편 2의 TaqI 제한 사이트
- [0120] 끝 2 - 단편 2에서 TaqI 사이트의 4000 염기 하류
- [0121] 전체 올가미 또는 탄성-망 정규화 (Lambda를 0.5 (탄성-망)로 설정)에 맞추기 위한 절차와 관련된 GLMNET 값.
- [0122] 표 1 내지 4는 ALS의 진단에 관한 것이고, 표 5 내지 9는 ALS의 예후에 관한 것이고, 일 구체 예에서 진단과 관련된 검출은 표 1 및 4에 기초하여 수행되고 예후에 관한 검출은 표 5 내지 9에 기초하여 수행 된다.
- [0123] 표 1 - 1 또는 -1이 표시된 LS 열은 ALS 사례가 존재함을 의미하고 -1은 ALS 사례가 부재함을 의미합니다.
- [0124] 표 5 - 1 또는 -1이 표시된 LS 열은 마커가 빠른 진행자에 존재하고 느린 진행자에는 부재함을 의미하고, -1은 마커가 느린 진행자에 존재하지만 빠른 진행자에는 부재함을 의미한다.
- [0125] 표 10 및 11은 ALS 예후와 관련이 있다. 표 10에는 이전 표에 표시된 마커가 포함되어 있다. 표 11의 '루프 감지' 열은 마커가 빠른 진행자에 존재하고 느린 진행자가 부재함을 의미한다.
- [0126] 마커는 EpiSwitch 3C 방법의 제품에 대한 관련 프로브를 참조하여 표에서 고유하게 식별된다. 가수 분해 프로브의 경우 이들은 Taq 사이트 (TCGA)의 상단에 있으며 EpiSwitch 상호 작용의 두 게놈 사이트를 모두 포함한다. 60개 염기 어레이 프로브 (서열 태그 각 측면의 30개 염기)와 동일한 접합을 측정하지만 양쪽에서 맞춤형 서열 길이에 걸쳐 있다. 표 19는 이에 대한 예를 제공한다.
- [0127] *시료 제조 및 염색체 상호 작용 검출을 위한 바람직한 구체 예*
- [0128] 시료를 제조하고 염색체 형태를 검출하는 방법이 본원 발명에 기재되어 있다. 예를 들어 이 부분에 기재된 대로 이러한 방법의 최적화된(비 양식화된) 버전을 사용될 수 있다.
- [0129] 전형적으로 시료에는 최소 2×10^5 개 이상의 세포가 포함된다. 시료는 최대 5×10^5 개 세포를 포함 할 수 있다. 한 구체 예에서, 시료는 2×10^5 내지 5.5×10^5 개 세포를 포함 할 것이다
- [0130] 염색체 유전자좌에 존재하는 후생유전학적 염색체 상호 작용의 가교 단계는 본원 발명에 기재되어 있다. 이것은 세포 용해가 일어나기 전에 수행 될 수 있다. 세포 용해는 3 내지 7 분, 예컨대 4 내지 6 또는 약 5 분 동안 수행 될 수 있다. 일부 구체 예에서, 세포 용해는 적어도 5 분 이상 및 10 분 미만 동안 수행된다.
- [0131] 제한 효소로 DNA를 소화시키는 단계는 본원 발명에 기술되어 있다. 전형적으로, DNA 제한은 약 10 내지 30분 동안, 예컨대 약 20분 동안 약 55°C 내지 약 70°C , 예를 들어 약 65°C 에서 수행된다.
- [0132] 바람직하게는 평균 단편 크기가 최대 4000 염기쌍인 연결된 DNA의 단편을 생성하는 빈번한 절단기 제한 효소가 사용된다. 선택적으로, 제한 효소는 평균 단편 크기가 약 200 내지 300개의 염기쌍인, 예컨대 약 256개의 염기쌍을 갖는 연결된 DNA 단편이 생성한다, 일 구체 예에서, 전형적인 단편 크기는 200 내지 4,000개의 염기쌍, 예컨대 400 내지 2,000 또는 500 내지 1,000개의 염기쌍이다.
- [0133] EpiSwitch 방법의 한 구체 예에서, DNA 제한 단계와 DNA 연결 단계 사이에는 DNA 침전 단계가 수행되지 않는다.
- [0134] DNA 연결 단계는 본원 발명에 기재되어 있다. 전형적으로 DNA 연결은 약 5 내지 30분 동안, 예컨대 10분 동안 수행된다.
- [0135] 시료 내의 단백질은 예를 들어 단백질 분해 효소, 선택적으로 단백질 분해 효소 K를 사용하여 효소적으로 분해 될 수 있다. 단백질은 약 30분 내지 1 시간동안, 예를 들어 약 45분 동안 효소적으로 분해 될 수 있다. 한 구체 예에서, 단백질의 분해 후, 예를 들어 단백질 분해 효소 K 분해, 가교 역전 또는 페놀 DNA 추출 단계는 없다.
- [0136] 한 구체 예에서, PCR 검출은 바람직하게는, 연결된 핵산의 존재 / 부재에 대한 이진법 판독으로, 연결된 핵산의

단일 카피를 검출 할 수 있다.

- [0137] *본 발명의 공정 및 용도*
- [0138] 본 발명의 공정은 다른 방식으로 설명 될 수 있다. 연결된 핵산의 제조 방법은 다음 단계를 포함하여 설명 될 수 있다: (i) 염색체 상호 작용으로 함께 모인 염색체 영역의 시험 관내 가교 단계; (ii) 상기 가교된 DNA를 절단하거나 제한 분해 절단시키는 단계; 및 (iii) 상기 가교된 절단된 DNA 말단을 연결된 핵산을 형성하여 연결하는 단계로서, 연결된 핵산의 검출은 유전자좌에서 염색체 상태를 결정하는데 사용될 수 있으며, 바람직하게는:
- [0139] -유전자좌는 표 1 또는 5에 언급된 유전자좌, 영역 또는 유전자 중 하나 일 수 있으며,
- [0140] -및 / 또는 염색체 상호 작용은 본원에 언급되거나 표 1 또는 5에 개시된 임의의 프로브에 상응하는 임의의 염색체 상호 작용 일 수 있고 / 있거나
- [0141] -연결된 생성물은 (i) 표 1 또는 5에 개시된 임의의 프로브 서열과 동일하거나 상동인 서열; 또는 (ii) (ii)에 상보적인 서열을 포함할 수 있다.
- [0142] 본 발명의 공정은 염색체 상호 작용이 게놈의 정의된 후생유전학적 활성 (질병 관련) 영역 내에 존재하는지 여부를 결정하는 것을 포함하는, 집단에서 상이한 서브그룹을 나타내는 염색체 상태를 검출하는 공정으로서 기술 될 수 있으며, 여기서 바람직하게는:
- [0143] -서브그룹은 ALS의 존재 유무 또는 ALS와 관련된 특성 (예 : 예후 또는 진행)으로 정의되고 / 되거나,
- [0144] -염색체 상태는 표 1 또는 5에 언급된 유전자좌, 영역 또는 유전자에 있을 수 있고 / 있거나;
- [0145] -염색체 상호 작용은 표 1 또는 5에 언급된 것 중 어느 하나 일 수 있거나 해당 표에 개시된 임의의 프로브에 상응할 수 있다.
- [0146] 본 발명은 표 1 또는 5에 언급된 임의의 유전자좌, 유전자 또는 영역에서 염색체 상호 작용을 검출하는 것을 포함한다. 본 발명은 염색체 상호 작용을 검출하기 위해 본원 발명에 언급된 핵산 및 프로브의 사용, 예를 들어 적어도 1, 2, 4, 6 또는 8개 이상의 사용 또는 적어도 1, 2, 4, 6 또는 8개의 상이한 유전자좌 또는 유전자에서 염색체 상호 작용을 검출하기 위한 이러한 핵산 또는 프로브를 포함한다. 본 발명은 표 2 또는 7에 열거된 프라이머 또는 프라이머 쌍 중 임의의 것을 사용하거나 본원 발명에 기재된 바와 같은 이들 프라이머의 변이체(프라이머 서열을 포함하거나 프라이머 서열의 단편 및 / 또는 상 동체를 포함하는 서열)를 사용하여 염색체 상호 작용을 검출하는 것을 포함한다.
- [0147] 염색체 상호 작용이 정의된 유전자, 영역 또는 위치 ‘내에서’ 발생하는지 여부를 분석할 때, 상호 작용에 함께 있는 염색체의 두 부분 모두가 정의된 유전자, 영역 또는 위치 내에 있고 뿐만 아니라 일부 구체 예에서 염색체의 한 부분만이 정의된 유전자, 영역 또는 위치 내에 있다 .
- [0148] *새로운 치료법을 식별하기 위한 본 발명의 방법의 용도*
- [0149] 염색체 상호 작용에 대한 지식은 컨디션(conditions)에 대한 새로운 치료법을 식별하는데 사용될 수 있다. 본 발명은 ALS를 위한 새로운 치료제(예후와 관련된 치료 포함)를 식별 또는 설계하기 위해 본원 발명에 정의된 염색체 상호 작용의 방법과 용도를 제공한다.
- [0150] *상동성*
- [0151] 폴리뉴클레오티드 / 핵산(예를 들어, DNA) 서열의 상동체가 본원 발명에서 언급된다. 이러한 상동체는 전형적으로 적어도 70% 이상, 바람직하게는 80% 이상, 85% 이상, 90% 이상, 95% 이상, 97% 이상, 98% 이상 또는 99% 이상의 상동성을 갖는다. 예를 들어, 적어도 10, 15, 20, 30, 100개 이상 또는 그 이상의 인접 뉴클레오티드의 영역, 또는 염색체 상호 작용에 관여하는 염색체의 영역으로부터 유래한 핵산의 일부를 가로 지른다. 상동성은 뉴클레오티드 동일성(때때로 "단단한 상동성"으로 지칭 됨)에 기초하여 계산 될 수 있다.
- [0152] 따라서, 특정 구체적 실시 양태에서, 폴리뉴클레오티드 / 핵산(예를 들어, DNA) 서열의 상동체는 서열 동일성 백분율을 참조하여 본원 발명에 언급된다. 전형적으로 이러한 상동체는 적어도 70% 이상, 바람직하게는 80% 이상, 85% 이상, 90% 이상, 95% 이상, 97% 이상, 98% 이상 또는 99% 이상의 서열 동일성을 갖는다. 예를 들어, 적어도 10, 15, 20, 30, 100개 이상 또는 그 이상의 인접 뉴클레오티드의 영역에 걸쳐거나, 염색체 상호 작용에 관련된 염색체의 영역으로부터 유래한 핵산의 일부를 가로 지른다.
- [0153] 예를 들어 UWGCG 패키지는 상동성 및 / 또는 % 서열 동일성을 계산하는데 사용할 수 있는 BESTFIT 프로그램을

제공한다(예 : 기본 설정에 사용) (Devereux *et al* (1984) Nucleic Acids Research 12, p387-395). PILEUP 및 BLAST 알고리즘은, 예를 들어 Altschul SF (1993) J MoI Evol 36 : 290-300; Altschul, S, F *et al* (1990) J MoI Biol 215 : 403-10.에 기재된 대로, 상동성 및 / 또는 % 서열 동일성 및 / 또는 라인업 서열(예 : 동등하거나 상응하는 서열을 식별하는 등(예 : 기본 설정에 사용))를 계산하는데 사용될 수 있다.

[0154] BLAST 분석을 수행하기 위한 소프트웨어는 국립 생명 공학 정보 센터(National Center for Biotechnology Information)를 통해 공개적으로 제공된다. 이 알고리즘은 데이터베이스 서열에서 동일한 길이의 단어와 정렬될 때 일부 양의 임계 값 스코어 T와 일치하거나 만족하는 쿼리 서열에서 길이 W의 짧은 단어를 식별함으로써 우선 높은 스코어링 서열 쌍(HSP)을 식별하는 것을 포함한다. T는 이웃 단어 점수 임계 값(상기 Altschul *et al.*)으로 지칭된다. 이러한 초기 이웃 단어 적중은 이를 포함하는 HSP를 찾기 위한 검색을 시작하기 위해 종자 역할을 한다. 누적 정렬 점수가 증가 될 수 있는 한, 단어 적중은 각 서열을 따라 양방향으로 확장된다. 각 방향으로의 단어 적중에 대한 확장은 다음과 같은 경우 중지된다. 누적 정렬 점수가 최대 달성 값에서 수량 X 만큼 떨어지고; 하나 이상의 음의 스코어링 잔기 정렬의 축적으로 인해 누적 스코어가 0 이하로 되거나; 서열의 끝에 도달하는 경우에 중지된다. BLAST 알고리즘 파라미터 W5 T 와 X는 정렬의 감도와 속도를 결정한다. BLAST 프로그램은 11개의 단어 길이 (W), 50의 BLOSUM62 스코어링 매트릭스(Henikoff and Henikoff (1992) Proc. Natl. Acad. Sci. USA 89 : 10915-10919 참조) 정렬 (B), 10개의 기대치(E), M=5, N=4, 및 두 가닥의 비교를 기본값으로 사용한다.

[0155] BLAST 알고리즘은, 예를 들어, Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90 : 5873-5787 참조와 같이, 두 서열 간의 유사성에 대한 통계적 분석을 수행한다. BLAST 알고리즘에 의해 제공되는 유사성의 한 척도는 가장 작은 합 확률 (P (N))이며, 이는 두 폴리뉴클레오티드 서열 사이의 일치가 우연히 발생할 확률의 표시를 제공한다. 예를 들어, 제 1 서열과 제 2 서열의 비교에서 가장 작은 합 확률이 약 1 미만, 바람직하게는 약 0.1 미만, 보다 바람직하게는 약 0.01 미만, 가장 바람직하게는 약 0.001 미만인 경우, 서열은 다른 서열과 유사한 것으로 간주된다.

[0156] 상동성 서열은 전형적으로 1, 2, 3, 4개 또는 그 이상의 염기, 예컨대 10, 15 또는 20개 미만의 염기(뉴클레오티드의 치환, 결실 또는 삽입일 수 있음) 차이가 있다. 이들 변화는 상동성 및 / 또는 % 서열 동일성을 계산하는 것과 관련하여 상기 언급된 임의의 영역에 걸쳐 측정될 수 있다.

[0157] '프라이머 쌍'의 상동성은, 예를 들어, 다시 단일 서열로 간주되는 다른 프라이머 쌍과 비교할 목적을 위해 두 개의 서열을 단일 서열(두 개의 서열이 서로 결합 된 것처럼)로 간주함으로써 계산될 수 있다.

[0158] *어레이(Array)*

[0159] 제 2 핵산 세트는 어레이에 결합 될 수 있고, 일 구체 예에서 적어도 15,000, 45,000, 100,000 또는 250,000개 이상의 상이한 제 2 핵산이 어레이에 결합되며, 바람직하게는, 적어도 300, 900, 2000 또는 5000개 이상의 유전자좌를 나타낸다. 일 구체 예에서, 제 2 핵산의 하나, 또는 그 이상, 또는 상이한 모든 집단은 어레이의 하나 이상의 별개의 영역에 결합되며, 사실상 어레이에서 반복되어 오류 검출이 가능하다. 어레이는 Agilent SurePrint G3 Custom CGH 마이크로 어레이 플랫폼을 기반으로 할 수 있다. 어레이에 대한 제 1 핵산의 결합의 검출은 이중 색상 시스템에 의해 수행 될 수 있다.

[0160] *치료제*

[0161] 치료제는 본원 발명에서 언급된다. 본 발명은 특정 개체, 예를 들어 본 발명의 공정에 의해 식별된 개체에서 ALS를 예방 또는 치료하는데 사용하기 위한 제제를 제공한다. 이는 치료 유효량의 작용 제제를 필요로 하는 개체에게 투여하는 것을 포함 할 수 있다. 본 발명은 특정 개체에서 ALS를 예방 또는 치료하기 위한 약제의 제조에 있어서 제제의 용도를 제공한다.

[0162] ALS에 바람직한 치료제는 다음과 같다 :

[0163] Riluzole(Rilutek) : 이 약은 전형적으로 알약으로 섭취되며 뇌의 메신저 (글루타메이트) 수치를 줄여 질병의 진행을 줄입니다. 글루타메이트(glutamate)는 ALS 환자에서 더 높은 수준으로 존재한다.

[0164] Edaravone(Radicava) : 이 약은 ALS와 관련된 일일 수행력 저하를 줄입니다. 약은 전형적으로 한 달에 한 번, 10-14일 동안 연속적으로 정맥 주사를 통해 환자에게 제공된다.

[0165] Arimoclomol : 이 약은 운동 뉴런에서 열 충격 반응 유도제로 작용하며 뉴런 장애 및 세포 사멸을 방어한다.

- [0166] Talampanel : 이 약은 근력 저하 및 증상 진행률을 감소시킨다.
- [0167] 베타-락탐 항생제(Beta-lactam antibiotics) : 페니실린 및 세팔로스포린(cephalosporin)과 같은 항체는 근육 안정성을 유지하고, GLT1, 신경교 글루타메이트 수송체 1의 수준을 상향 조절하여 수명을 증가시킨다.
- [0168] 브로모크립틴(Bromocriptine) : 이 약은 스트레스에 의해 유발되는 산화성 세포 사멸을 억제하는 자유 라디칼 제거제이다.
- [0169] Pramipexole 및 Dexamipexole : 이 약은 도파민(dopamine) 작용제 역할을 하며 자유 라디칼 제거 기능이 있다. 이 약은 미토콘드리아 기능 장애와 관련이 있다. 텍스트라미펙솔(Dexamipexole)은 프라미펙솔(Pramipexole)의 광학 거울상 이성질체이다.
- [0170] 줄기 세포 요법 : 줄기 세포 성장은 뉴런 질환의 진행을 감소시키거나 운동 뉴런을 대체한다. 줄기 세포는 척추 운동 뉴런을 만들고, 축삭을 확장하며 근육과의 시냅스를 받고 생성할 가능성이 있다. 성체 줄기 세포로부터 유래된 중간엽 줄기 세포(MSC)는 질병 진행을 지연시키기 위해 영양 인자, 항 염증성 사이토카인 및 면역 조절 케모카인(chemokines)을 방출한다.
- [0171] 면역 요법 : ICV 경로를 통한 D3H5 항체 주입과 같은 항체 요법은 ALS에 대한 유전자 변형 마우스 모델의 수명을 연장하고 더 긴 시간 동안 체중을 유지한다.
- [0172] 다음은 헌팅틴병에 대한 치료요법 목록이다. 이들은 운동 및 정신 장애의 증상을 줄이는데 도움이 될 수 있다.
- [0173] 운동 장애에 대한 약물
- [0174] ● 테트라베나진(Tetrabenazine) (Xenazine)
- [0175] ● 할로페리돌(Haldol) (Haldol), 클로르프로마진(chlorpromazine), 리스페리돈(risperidone)(Risperdal) 및 퀘티아핀(quetiapine)(Seroquel)과 같은 항 정신병 약
- [0176] ● 다른 약물로는 아만타딘(amantadine), 레벤티라세탐(levetiracetam)(Keppra, 기타) 및 클로나제팜(clonazepam)(Klonopin)이 있다.
- [0177] 정신 장애에 대한 약물
- [0178] ● 항우울제는 시탈로프람(citalopram) (Celexa), 에스시탈로프람(escitalopram) (Lexapro), 플루옥세틴(fluoxetine) (Prozac, Sarafem) 및 세르트랄린(sertraline) (Zoloft)을 포함한다.
- [0179] ● 항정신병 약에는 퀘티아핀(quetiapine) (Seroquel), 리스페리돈(risperidon) (Risperdal) 및 올란자핀(olanzapine) (Zyprexa)이 포함된다.
- [0180] ● 기분 안정화 약으로는 발프로에이트(valproate)(Depacon), 카르바마제핀 (carbamazepine)(Carbatrol, Eptol, Tegretol) 및 라모트리진(lamotrigine) (Lamictal)이 포함된다.
- [0181] 헌팅틴병의 조기 진단은 질병 경과에 따른 증상 치료를 관리하는 데 도움이 된다.
- [0182] 제제의 제형(ALS 또는 헌팅틴병의 경우)은 제제의 특성에 따라 달라진다. 제제는 제제 및 약제학적으로 허용되는 담체 또는 희석제를 함유하는 약제학적 조성물의 형태로 제공 될 것이다. 적합한 담체 및 희석제는 등장성 식염수 용액, 예를 들어 포스페이트(phosphate) 완충 식염수를 포함한다. 전형적인 경구 투여 조성물은 정제, 캡슐, 액체 용액 및 액체 현탁액을 포함한다. 상기 제제는 비경구, 정맥 내, 근육 내, 피하, 경피 또는 경구 투여용으로 제형화 될 수 있다.
- [0183] 제제의 용량은 다양한 파라미터, 특히 사용되는 물질에 따라 결정될 수 있으며; 치료될 개체의 연령, 체중 및 상태; 투여 경로; 필요한 요법에 따라 결정될 수 있다. 의사는 임의의 특정 제제에 필요한 투여 경로 및 투여량을 결정할 수 있을 것이다. 그러나 적합한 투여량은 0.1 내지 체중 당 100mg / kg, 예를 들어 체중 당 1 내지 40mg / kg, 예를 들어 매일 1 내지 3회 복용될 수 있다.
- [0184] *본원 발명에 언급된 물질의 형태*
- [0185] 본원 발명에 언급된 임의의 물질, 예컨대 핵산 또는 치료제는, 정제되거나 분리된 형태 일 수 있다. 이들은 자연에서 발견된 것과 다른 형태 일 수 있으며, 예를 들어 자연에서 발생하지 않는 다른 물질과의 조합으로 존재

할 수 있다. 핵산 (본원 발명에 정의된 서열의 일부 포함)은, 예를 들어 상기 상동성 부분에 기재된 바와 같이 서열에서 적어도 1, 2, 3, 4개 이상의 뉴클레오티드 변경을 갖는, 자연에서 발견된 것과 상이한 서열을 가질 수 있다. 핵산은 5' 또는 3' 말단에 이중 서열을 가질 수 있다. 핵산은 자연에서 발견된 것과 화학적으로 상이할 수 있으며, 예를 들어 어떤 방식으로 변형될 수 있지만, 바람직하게는 여전히 Watson-Crick 염기쌍을 만들 수 있다. 적절한 경우, 핵산은 이중 가닥 또는 단일 가닥 형태로 제공될 것이다. 본 발명은 본원 발명에 언급된 모든 특정 핵산 서열을 단일 또는 이중 가닥 형태로 제공하므로, 개시된 임의의 서열에 상보적인 가닥을 포함한다.

[0186] 본 발명은 또한 ALS와 관련된 염색체 상호 작용의 검출(예를 들어 ALS의 진단 또는 예후와 관련된)을 포함하여 본 발명의 임의의 공정을 수행하기 위한 키트를 제공한다. 이러한 키트는 본 발명의 공정에 의해 생성된 연결된 핵산을 검출 할 수 있는 제제와 같은 관련 염색체 상호 작용을 검출 할 수 있는 특이 적 결합제를 포함 할 수 있다. 키트에 존재하는 바람직한 제제는 예를 들어, 본원 발명에 기술된 바와 같이, PCR 반응에서 연결된 핵산을 증폭시킬 수 있는 연결된 핵산 또는 프라이머 쌍에 혼성화할 수 있는 프로브를 포함한다.

[0187] 본 발명은 또한 관련 염색체 상호 작용을 검출 할 수 있는 장치를 제공한다. 장치는 바람직하게는 본원 발명에 기술된 임의의 이러한 제제, 프로브 또는 프라이머 쌍과 같은, 염색체 상호 작용을 검출 할 수 있는 임의의 특정 결합제, 프로브 또는 프라이머 쌍을 포함한다.

[0188] 탐지 방법

[0189] 한 구체 예에서, 염색체 상호 작용과 관련된 연결된 서열의 정량적 검출은 PCR 반응 동안 활성화시 검출 될 수 있는 프로브를 사용하여 수행되며, 상기 연결된 서열은 후생유전학적 염색체 상호 작용으로 함께 모인 2 개의 염색체 영역으로부터의 서열을 포함하고, 상기 방법은 PCR 반응 동안 연결된 서열을 프로브와 접촉시키는 단계 및 프로브의 활성화 정도를 검출하는 단계를 포함하고, 상기 프로브는 연결 부위에 결합하는 단계를 포함한다. 상기 방법은 전형적으로 이중 표지 된 형광 가수 분해 프로브를 사용하여 특정 상호 작용이 MIQE 준수 방식으로 검출 될 수 있게 한다.

[0190] 프로브에는 일반적으로 비활성 및 활성 상태를 갖는 감지 가능한 표지가 붙어있어 활성화될 때만 감지된다. 활성화 정도는 PCR 반응에 존재하는 주형 (결합 생성물)의 정도와 관련될 것이다. PCR의 전부 또는 일부 동안, 예를 들어 적어도 50% 또는 80% 이상의 PCR 주기 동안 검출이 수행 될 수 있다.

[0191] 프로브는 올리고뉴클레오티드의 한 말단에 공유적으로 부착된 형광단 및 뉴클레오티드의 다른 말단에 부착된 소광제를 포함하여, 형광단의 형광이 소광제에 의해 소광 될 수 있다. 일 구체 예에서, 형광단은 올리고뉴클레오티드의 5' 말단에 부착되고, 소광제는 올리고뉴클레오티드의 3' 말단에 공유적으로 부착된다. 본 발명의 방법에 사용될 수 있는 형광단은 FAM, TET, JOE, Yakima Yellow, HEX, Cyanine3, ATTO550, TAMRA, ROX, Texas Red, Cyanine3.5, LC610, LC640, ATTO647N, Cyanine5, Cyanine5.5 및 ATTO 680을 포함한다. 적절한 형광단과 함께 사용될 수 있는 소광제는 TAM, BHQ1, DAB, Eclip, BHQ2 및 BBQ650을 포함하고, 선택적으로 상기 형광단은 HEX, Texas Red 및 FAM으로부터 선택된다. 형광단 및 소광제의 바람직한 조합은 BHQ1을 갖는 FAM 및 BHQ2를 갖는 Texas Red를 포함한다.

[0192] qPCR 분석에서 프로브 사용

[0193] 본 발명의 가수 분해 프로브는 전형적으로 농도 매칭 음성 대조군으로 최적화 된 온도 구배이다. 바람직하게는, 단일 단계 PCR 반응이 최적화된다. 보다 바람직하게는 표준 곡선이 계산된다. 연결된 서열의 접합 부분을 가로질러 결합하는 특정 프로브를 사용하는 이점은 증폭된 PCR 접근법을 사용하지 않고 연결된 서열에 대한 특이성이 달성 될 수 있다는 것이다. 본 명세서에 기술된 방법은 적은 카피 수 표적의 정확하고 정확한 정량을 가능하게 한다. 표적 연결된 서열은 온도 구배 최적화 전에 정제, 예를 들어 겔 정제 될 수 있다. 표적 연결된 서열은 서열화 될 수 있다. 바람직하게는, PCR 반응은 약 10ng, 또는 5 내지 15ng, 또는 10 내지 20ng, 또는 10 내지 50ng, 또는 10 내지 200ng 주형 DNA를 사용하여 수행된다. 정방향 및 역방향 프라이머는 하나의 프라이머가 연결된 DNA 서열로 표현된 염색체 영역 중 하나의 서열에 결합하고, 다른 프라이머는 연결된 DNA 서열로 표현된 다른 염색체 영역에 결합하도록 설계 되는데 예를 들어, 서열에 상보적으로 설계된다.

[0194] 연결된 DNA 표적의 선택

[0195] 본 발명은 연결된 서열에 결합하고 증폭하는 능력에 기초하여 프라이머를 선택하고, 특히 결합할 표적 서열의 프로브 서열에 기초한 특성을 선택하는 것을 포함하는데, 특히, 표적서열의 곡률의 결합을 포함하는, 본원 발명

에 정의된 PCR 방법에 사용하기 위한 프라이머 및 프로브를 선택하는 것을 포함한다.

- [0196] 프로브는 전형적으로 제한 부위에 걸친 병치 된 제한 단편 인 결합 된 서열에 결합하도록 설계 / 선택된다. 본 발명의 일 구체 예에서, 특정 염색체 상호 작용과 관련된 가능한 결합 서열의 예상 곡률은 예를 들어 본원에서 참조 된 특정 알고리즘을 사용하여 계산된다. 곡률은 예를 들어 나선형 회 전당 각도로 표현할 수 있습니다. 나선형 회 전당 10.5 ° . 결합 된 서열은 결합 된 서열이 나선형 턴당 5 ° 이상, 전형적으로 나선형 턴당 10 ° , 15 ° 또는 20 ° , 예를 들어 나선형 턴당 5 ° 내지 20 ° 의 곡률 성향 피크 스코어를 갖는 표적화를 위해 선택된다. . 바람직하게는 헬리컬 턴당 곡률 성향 스코어는 결합 부위의 상류 및 / 또는 하류에 20, 400 개의 염기와 같은 적어도 20, 50, 100, 200 또는 400 개의 염기에 대해 계산된다. 따라서 한 실시 양태에서, 결합 된 생성물에서의 표적 서열은 임의의 이러한 수준의 곡률을 갖는다. 최저 열역학적 구조 자유 에너지에 기초하여 표적 서열을 선택할 수도 있다.
- [0197] 바람직한 ALS 실시 예
- [0198] 단락 1. 염색체 상태와 관련된 염색체 상호 작용이 게놈의 정의된 영역 내에 존재하는지 여부를 결정하는 단계를 포함하는, 집단에서의 질병 서브그룹을 나타내는 염색체 상태를 검출하는 공정으로서,
- [0199] 상기 질병 서브그룹은 근위축성측삭경화증(ALS) 서브그룹이고;
- [0200] - 선택적으로, 상기 염색체 상호 작용은, 임의의 염색체 상호 작용이 집단의 ALS 서브그룹에 상응하는 염색체 상태와 관련이 있는지를 결정하는 방법에 의해 식별되고,
- [0201] 상기 결정 방법은 상이한 상태의 염색체를 갖는 서브그룹으로부터의 제1 핵산 세트와 제2 인덱스(index) 핵산 세트와 접촉시키는 단계 및, 상보적 서열이 혼성화하는 것을 허용하는 단계를 포함하며, 상기 제1 및 제2 핵산 세트에서의 핵산은 염색체 상호 작용에서 함께 모이는 염색체 영역 둘 다를로부터의 서열을 포함하는 연결(ligated)된 생성물을 나타내고, 제1 핵산세트와 제2 핵산 세트 사이의 혼성화 패턴은 ALS 서브그룹에 특이적인 염색체 상호 작용의 결정을 허용하며;
- [0202] - 상기 염색체 상호 작용은
- [0203] (i) 표 1 또는 5에 열거된 영역 또는 유전자 중 어느 하나에 존재하고/하거나;
- [0204] (ii) 표 1 또는 5에 나타난 임의의 프로브로 표시되는 염색체 상호 작용 중 어느 하나에 상응하고/하거나,
- [0205] (iii) (i) 또는 (ii)를 포함하거나 측면(flanks)에 위치하는 4,000개의 염기 영역에 존재하는, 집단에서의 질병 서브그룹을 나타내는 염색체 상태를 검출하는 공정.
- [0206] 단락 2. 제1항에 있어서, ALS를 진단하거나 ALS에 대한 예후를 결정하기 위해 수행되는 것을 특징으로 하는 공정.
- [0207] 단락 3. 제1항 또는 제2항에 있어서, 유형화(typed)되는 염색체 상호 작용의 특정 조합은
- [0208] (i) 표 1 또는 5에서 프로브로 표시되는 모든 염색체 상호 작용을 포함하거나;
- [0209] (ii) 표 1 또는 5에서 프로브로 표시되는 염색체 상호 작용 중 적어도 4, 5, 6 또는 7개 이상을 포함하거나;
- [0210] (iii) 표 1 또는 5에 열거 된 영역 또는 유전자 중 적어도 4, 5, 6 또는 7개 이상에 함께 존재하거나;
- [0211] (iv) 표 1 또는 5에서 프로브로 표시되는 염색체 상호 작용을 포함하거나 또는 측면에 위치하는 4,000개의 염기 영역에 존재하는 적어도 4, 5, 6 또는 7개 이상의 염색체 상호 작용을 포함하는 것을 특징으로 하는 공정.
- [0212] 단락 4. 제1항 내지 제3항 중 어느 한 항에 있어서, 염색체 상호 작용은 :
- [0213] -개체의 샘플 안에서 유형화(typed)되고/되거나,
- [0214] -염색체 상호 작용 부위에서 DNA 루프의 존재 또는 부재를 검출함으로써 유형화(typed)되고/되거나,
- [0215] -염색체 형태로 결합된 염색체의 원위 영역의 존재 또는 부재를 검출하여 유형화(typed)되고/되거나,
- [0216] -상기 유형화(typing) 동안 생성되고, 염색체 상호 작용에서 함께 모인 염색체의 영역에 각각 상응하는 2개의 영역을 포함하는 서열을 가지는 연결(ligated)된 핵산의 존재를 검출함으로써 염색체 상호 작용이 유형화(typed)되며,
- [0217] 상기 연결(ligated)된 핵산의 검출은, 바람직하게는, (i) 표 1 또는 5에 언급된 임의의 특정 프로브 서열과 적

어도 70% 이상의 동일성을 갖는 프로브, 및/또는 (ii) 표 2 또는 7의 임의의 프라이머 쌍과 적어도 70% 이상의 동일성을 갖는 프라이머 쌍을 사용함으로써 이루어지는 것을 특징으로 하는 공정.

- [0218] 단락 5. 제1항 내지 제4항 중 어느 한 항에 따른 공정에 있어서,
- [0219] -상기 제2 핵산 세트는 상기 제1 핵산 세트보다 더 큰 개체 그룹으로부터 유래되고/되거나;
- [0220] -상기 제1 핵산 세트는 적어도 8 개체 이상으로부터 유래되고/되거나;
- [0221] -상기 제1 핵산 세트는 제1 서브그룹으로부터 적어도 4 개체 이상과 바람직하게는, 제1 서브그룹과 겹치지 않는 제2 서브그룹으로부터 적어도 4 개체 이상으로부터 유래하고/하거나;
- [0222] -치료를 위한 개체를 선택하기 위해 수행되는 것을 특징으로 하는 공정.
- [0223] 단락 6. 제1항 내지 제5항 중 어느 한 항에 따른 공정에 있어서,
- [0224] -상기 제2 핵산 세트는 선택되지 않은 그룹을 나타내고/내거나;
- [0225] -상기 제2 핵산 세트는 정의된 위치에서 어레이에 결합되고/되거나;
- [0226] -상기 제2 핵산 세트는 적어도 100개 이상의 상이한 유전자에서 염색체 상호 작용을 나타내고/내거나;
- [0227] -상기 제2 핵산 세트는 적어도 1,000개 이상의 상이한 염색체 상호 작용을 나타내는 적어도 1,000개 이상의 상이한 핵산을 포함하고/하거나;
- [0228] -상기 제1 핵산 세트 및 제2 핵산 세트는 길이가 10 내지 100 개의 뉴클레오티드 염기를 갖는 적어도 100개 이상의 핵산을 포함하는 것을 특징으로 하는 공정.
- [0229] 단락 7. 제1항 내지 제6항 중 어느 한 항에 따른 공정에 있어서, 제1 핵산 세트는
- [0230] (i) 염색체 상호 작용으로 함께 모인 염색체 영역의 가교(cross-linking) 단계;
- [0231] (ii) 상기 가교(cross-linking) 된 영역을 선택적으로, 효소에 의한 제한 분해 절단에 의해, 절단하는 단계; 및
- [0232] (iii) 제1 핵산 세트(특히 연결된 DNA를 포함)를 형성하기 위해 상기 가교(cross-linking) 되고 절단된 DNA 말단을 연결하는 단계를 포함하는 공정에서 수득 가능한 것을 특징으로 하는 공정.
- [0233] 단락 8. 제1항 내지 제7항 중 어느 한 항에 있어서, 적어도 5 내지 9 개 이상의 상이한 염색체 상호 작용이, 바람직하게는, 5 내지 9 개의 상이한 영역 또는 유전자 안에서, 유형화(typed)되는 공정.
- [0234] 단락 9. 제1항 내지 제8항 중 어느 한 항에 따른 공정에 있어서, 상기 개념의 정의된 영역은 :
- [0235] (i) 단일 뉴클레오티드 다형성 (SNP)을 포함하고/하거나;
- [0236] (ii) 마이크로 RNA (miRNA)를 발현하고/하거나;
- [0237] (iii) 넌 코딩(non-coding) RNA (ncRNA)를 발현하고/하거나;
- [0238] (iv) 적어도 10개 이상의 인접 아미노산 잔기를 암호화하는 핵산 서열을 발현하고/하거나;
- [0239] (v) 조절 요소를 발현하고/하거나;
- [0240] (vi) CTCF 결합 부위를 포함하는 것을 특징으로 하는 공정.
- [0241] 단락 10. 염색체 상호 작용의 변화를 유발하여 치료 효과를 유발할 수 있는 제제(agent)를 선택하여 ALS를 치료하기 위한 치료제를 식별 또는 설계하는 방법으로써,
- [0242] - 상기 염색체 상호 작용은 표 1 또는 5에서 임의의 프로브로 표시되고/되거나;
- [0243] - 상기 염색체 상호 작용이 표 1 또는 5에 열거 된 임의의 영역 또는 유전자에 존재하고/하거나;
- [0244] - 상기 염색체 상호 작용은 표 10 또는 11에 나타난 상호 작용 중 하나이며,
- [0245] 선택적으로,
- [0246] - 상기 염색체 상호 작용은 어느 염색체 상호 작용이 제1항에 정의된 염색체 상태와 관련이 있는지를 결정하는 방법에 의해 식별되고/되거나,

- [0247] -염색체 상호 작용의 상기 변화는 (i) 표 1 또는 5에 언급된 임의의 프로브 서열과 적어도 70% 이상의 동일성을 갖는 프로브, 및/또는 (ii) 표 1 또는 5의 임의의 프라이머 쌍과 적어도 70% 이상의 동일성을 갖는 프라이머 쌍에 의해 모니터되고/되거나;
- [0248] - 후보 제제가 세포와 접촉되고, 후보 제제가 ALS를 치료할 수 있는지 여부를 결정하기 위해 세포에서의 염색체 상호 작용이 모니터링 되는, ALS를 치료하기 위한 치료제를 식별 또는 설계하는 방법.
- [0249] 단락 11. ALS 치료제를 식별 또는 설계하기 위한
- [0250] (vi)
- [0251] - 염색체 상호 작용이 표 1 또는 5의 프로브로 표시되고/되거나,
- [0252] - 염색체 상호 작용이 표 1 또는 5에 언급된 임의의 영역 또는 유전자에 존재하고/하거나,
- [0253] - 염색체 상호 작용이 표 10 또는 11에 나타난 상호 작용 중 어느 하나인 염색체 상호 작용의 검출의 용도; 또는
- [0254] (v) 표 1 또는 5에 언급된 임의의 프로브 서열과 적어도 70% 이상의 동일성을 갖는 프로브의 용도; 또는
- [0255] (vi) 표 2 또는 7에서 식별된 임의의 프라이머 쌍과 적어도 70% 이상의 동일성을 갖는 프라이머 쌍의 용도.
- [0256] 단락 12. 제11항에 있어서, 치료제를 식별하기 위한 용도는 후보제제가 치료제인지의 여부를 결정하기 위하여 후보 약제를 투여하는 단계, 및 염색체 상호 작용의 상기 검출을 사용하는 단계를 포함하고,
- [0257] 염색체 상태의 변화 여부를 검출하기 위해 상기 프로브 또는 상기 프라이머 쌍을 사용하며,
- [0258] 여기서 용도는 선택적으로, 시험관내, 바람직하게는, 세포에서 수행되는 것을 특징으로 하는 용도.
- [0259] 단락 13. 제1항 내지 제9항 중 어느 한 항에 따른 공정(process)에 의해 치료제가 필요한 것으로 식별된 개체에 서 ALS를 예방 또는 치료하는 방법(method)에 사용하기 위한 ALS 치료제.
- [0260] 단락 14. 제1항 내지 제13항 중 어느 한 항의 공정(process), 방법(method) 또는 용도(use)에 있어서, 상기 유형화(typing) 또는 검출은 PCR 반응 동안 연결된 생성물을 증폭시킬 수 있는 프라이머 및 연결 부위에 결합하는 프로브를 사용하는 정량적 PCR(qPCR)에 의해 연결된 생성물의 특정 검출을 포함하고,
- [0261] 상기 프로브는 염색체 상호 작용에서 함께 모인 각각의 염색체 영역으로부터의 서열에 상보적인 서열을 포함하고,
- [0262] 바람직하게는, 상기 프로브는 :
- [0263] 연결된 생성물에 특이적으로 결합하는 올리고 뉴클레오티드, 및/또는
- [0264] 올리고 뉴클레오티드의 5 '말단에 공유 결합된 형광 단(fluorophore), 및/또는
- [0265] 올리고 뉴클레오티드의 3 '말단에 공유 결합된 소광제를 포함하고,
- [0266] 선택적으로,
- [0267] 상기 형광단은 HEX, 텍사스 레드 및 FAM으로부터 선택되고/되거나;
- [0268] 상기 프로브는 길이 10 내지 40 개 뉴클레오티드 염기, 바람직하게는, 20 내지 30 개 뉴클레오티드 염기의 핵산 서열을 포함하는 것을 특징으로 하는 공정(process), 방법(method) 또는 용도(use).
- [0269] 특정 실시 예
- [0270] 한 구체 예에서, 염색체 내 상호 작용 만이 유형화(typed) / 검출되고, 염색체 외 상호 작용 (다른 염색체 사이)은 유형화(typed) / 검출되지 않는다.
- [0271] 간행물
- [0272] 본원 발명에 언급된 모든 간행물의 내용은 본 명세서에 참고로 포함되며 본 발명과 관련된 특징을 추가로 정의 하는데 사용될 수 있다.
- [0273] 특유 실시 예

[0274] EpiSwitch™ 플랫폼 기술은 유전자좌에서 정상 조건과 비정상 조건 사이의 조절 변화에 대한 후생유전학적 조절 시그니처(signatures)를 감지한다. EpiSwitch™ 플랫폼은 염색체 형태 시그니처로도 알려진 인간 염색체의 조절 고차 구조와 관련된 기본 후생유전학적 유전자 조절 수준을 식별하고 모니터링 한다. 염색체 시그니처는 일련의 유전자 탈 조절에서 뚜렷한 1 차 단계이다. 이들은 DNA 메틸화 및 RNA 프로파일링과 같은 후생유전학적 및 유전자 발현 바이오마커를 이용하는 바이오 마커 플랫폼에 비해 고유한 장점을 갖는 고차 바이오마커이다.

[0275] *EpiSwitch™* 어레이 분석

[0276] 맞춤형 EpiSwitch™ 어레이 스크리닝 플랫폼은 4개의 밀도, 15K, 45K, 100K 및 250K의 고유한 염색체 형태로 제공되며, 각 키메라(chimeric) 조각은 어레이에서 4번 반복되고, 유효 밀도는 각각 60K, 180K, 400K 및 1 백만이다.

[0277] 맞춤형 설계 *EpiSwitch™* 어레이

[0278] 15K EpiSwitch™ 어레이는 EpiSwitch™ 바이오마커 발견 기술로 조사된 약 300개의 유전자좌를 포함하여 전체 게놈을 선별할 수 있다. EpiSwitch™ 어레이는 Agilent SurePrint G3 Custom CGH 마이크로어레이 플랫폼을 기반으로 하며, 이 기술은 4 개의 밀도, 60K, 180K, 400K 및 1 백만 프로브를 제공한다. 각 EpiSwitch™ 프로브가 4배로 제공되므로 어레이 당 밀도가 15K, 45K, 100K 및 250K로 감소하여, 재현성을 통계적으로 평가할 수 있다. 유전자좌마다 조사된 잠재적 EpiSwitch™ 마커의 평균 수는 50이므로, 조사할 수 있는 유전자좌의 수는 300, 900, 2000 및 5000이다.

[0279] *EpiSwitch™* 맞춤형 어레이 파이프라인

[0280] EpiSwitch™ 어레이는 EpiSwitch™ 라이브러리 생성 후 Cy5로 표지되고 Cy3로 표지된 시료 (대조군)의 다른 하나와 비교 / 분석되는 한 세트의 시료가 포함된 이중 칼라 시스템이다. 어레이는 Agilent SureScan 스캐너를 사용하여 스캔되고 결과 기능은 Agilent Feature Extraction 소프트웨어를 사용하여 추출된다. 그런 다음 R의 EpiSwitch™ 어레이 공정 스크립트를 사용하여 데이터를 처리한다. R : Limma* 내 Biomductor의 표준 이중 칼라 패키지를 사용하여 어레이가 처리된다. 어레이의 표준화는 Limma*의 normalizedWithinArrays 함수를 사용하여 수행되며, 이는 Agilent 양성 대조군 및 EpiSwitch™ 양성 대조군의 칩(chip) 위에 수행된다. 데이터는 Agilent Flag 호출을 기반으로 필터링되고, Agilent 대조군 프로브는 제거되며 기술 복제 프로브는 평균화되어 Limma*를 사용하여 분석 할 수 있다. 프로브는 비교되는 두 시나리오 간의 차이를 기반으로 모델링 된 다음 False Discovery Rate를 사용하여 수정된다. ≤ -1.1 또는 ≥ 1.1 이고 $p \leq 0.1$ FDR p-값을 통과하는 변동 계수 (CV) $\leq 30\%$ 인 프로브는 추가 스크리닝에 사용된다. 프로브 세트를 더 줄이기 위해 R의 FactorMineR 패키지를 사용하여 다중 요인 분석이 수행된다.

[0281] * 참고 : LIMMA는 마이크로어레이 실험에서 차등 발현을 평가하기 위한 선형 모델 및 경험적 베이(Bayes) 공정이다. Limma는 마이크로어레이 또는 RNA-Seq에서 발생하는 유전자 발현 데이터 분석을위한 R 패키지이다

[0282] 프로브 풀(pool)은 처음에 최종 선택을 위해 조정 된 p-값, FC 및 CV <30 % (임의의 차단점) 매개 변수를 기반으로 선택된다. 추가 분석 및 최종 목록은 처음 두 매개 변수 (조정된 p-값; FC) 만 기준으로 작성된다.

발명의 효과

[0283] 바이오마커를 통해 질병 특성을 식별할 수 있다. 본 발명에 일반적으로 사용되는 바이오마커는 RNA 발현 패턴 및 단백질 마커를 포함한다.

도면의 간단한 설명

[0284] 도 1은 염색체 상호 작용을 탐지 할 수 있는 방법을 보여준다.

도 2는 헌팅턴병 연구에서 조사된 게놈 영역의 시각적 개요를 보여준다. HTT 유전자좌에 걸친 염색체 4의 ~ 225 kb 영역을 조사 하였다. 앵커(Anchor) 포인트(트랙 4의 "앵커")는 HTT 엑손 1의 CAG 반복 트랙에 걸친 ~ 42kb 영역으로 정의된다(도면 상단의 보라색 화살표). EpiSwitch 사이트(트랙 3), HD(트랙 5) 또는 기타 질병(트랙 6)과 관련된 SNPs와 검침을 기반으로 한 5 개의 구역(트랙 4의 구역 1-5)가 정의되었고, HC와 HD 사이의 메틸화 및 아세틸(H3K4me3, H3K36me3 및 H3K27Ac) 차이(트랙 7 ~ 12)가 관찰되었다.

도 3은 7 개의 개별 HD-Sym 샘플 중 6 개에서 3 가지 조건부 상호 작용(I5, I6 및 I7) 중 적어도 하나 이상의 존재가 관찰되었음을 보여준다. rs362331 SNP를 포함하는 영역에 걸친 상호 작용 인 I5는 가장 많은 수

의 샘플에서 관찰되었다(6/7).

도 4는 헌팅턴병의 증상 진행을 유발하는 후생 유전적 변화에 대한 모델을 제공한다. HD 진행과 관련된 염색체 형태 변화의 개요를 보여준다. 환자들이 전증후 단계에서 증상성 질환으로 진행함에 따라, HTT 유전자 좌위의 게놈 구조에서 불연속적이고, 측정 가능하고 차별적인 변화가 관찰된다.

도 5는 식별된 모든 헌팅턴병 상호 작용의 요약을 보여준다.

발명을 실시하기 위한 구체적인 내용

[0285] 본 발명은 아래의 비 제한적인 실시 예에 의해 설명된다.

[0286] <실시 예 1>

[0287] **통계 파이프라인**

[0288] EpiSwitch™ 스크리닝 어레이는 EpiSwitch™ PCR 플랫폼으로의 변환을 위한 고가의 EpiSwitch™ 마커를 선택하기 위해 R의 EpiSwitch™ 분석 패키지를 사용하여 처리된다.

[0289] 1 단계

[0290] 프로브는 변형된 선형 회귀 모델의 생산물인, 수정된 p-값 (False Discovery Rate, FDR)을 기반으로 선택된다. p-값 <= 0.1 미만의 프로브가 선택되고 후생유전학적 비율(ER)에 의해 추가로 감소되며, 추가 분석을 위해 선택되기 위해서 프로브 ER은 <=-1.1 또는 => 1.1이어야 한다. 마지막 필터는 변동 계수 (CV)이며 프로브는 <= 0.3 미만이어야 한다.

[0291] 2 단계

[0292] 통계 목록에서 상위 40개의 마커는 PCR 번역의 마커로 선택되기 위해 ER을 기반으로 선택됩니다. 음의 ER 로드 (load)가 가장 높은 상위 20개의 마커와 양의 ER 로드(load)가 가장 높은 상위 20 개의 마커가 목록을 구성한다.

[0293] 3 단계

[0294] 단계 1의 결과 마커인, 통계적으로 유의한 프로브는 초기하학 보강(HE)을 사용하여 보강 분석의 기초를 형성한다. 이 분석은 중요한 프로브 목록에서 마커를 줄일 수 있으며, 2 단계의 마커와 함께 EpiSwitch™ PCR 플랫폼으로 번역된 프로브 목록을 형성합니다.

[0295] 통계적 프로브는 HE에 의해 처리되어 어느 유전자 위치가 통계적으로 유의한 프로브를 강화시키는지 결정하고, 어느 유전자 위치가 후생유전학적 차이의 허브인가를 나타낸다.

[0296] 보정된 p-값을 기반으로 가장 중요한 보강된 유전자좌가 프로브 목록 생성을 위해 선택된다. p-값이 0.3 또는 0.2 미만인 유전자 위치가 선택된다. 2 단계의 마커를 사용하여 이러한 유전자 위치에 매핑되는 통계 프로브는 EpiSwitch™ PCR 번역에 대한 높은 값의 마커를 형성한다.

[0297] **어레이 설계 및 공정**

[0298] 어레이 설계

[0299] 1. 유전자좌는 SII 소프트웨어(현재 v3.2)를 사용하여 다음과 같이 처리된다.

- [0300] a. 특정 유전자 좌에서 게놈의 서열을 빼낸다(50kb 상류 및 20kb 하류를 갖는 유전자 서열).
- [0301] b. 이 지역 내의 서열이 CC에 포함될 확률을 정의한다.
- [0302] c. 특정 RE를 사용하여 서열을 잘라 낸다.
- [0303] d. 특정 방향에서 상호 작용할 수 있는 제한 단편을 결정한다.
- [0304] e. 서로 다른 CC가 상호 작용할 가능성을 순위 매긴다.

[0305] 2. 어레이 크기 및 사용 가능한 프로브 위치 수 결정한다(x).

[0306] 3. x / 4 상호 작용을 꺼낸다.

- [0307] 4. 각 상호 작용에 대해 파트 1의 제한 사이트에서 30bp까지, 파트 2의 제한 사이트에서 30bp까지의 서열을 정의한다. 해당 영역이 반복되지 않는지 확인하고, 해당 영역을 제외하고 목록에서 다음 상호 작용을 수행한다. 양 30bp를 결합하여 프로브를 정의한다.
- [0308] 5. $x / 4$ 프로브 및 정의된 대조군 프로브 목록을 생성하고 4번 복제하여 어레이에 생성될 목록을 생성한다.
- [0309] 6. 맞춤형 CGH 어레이를 위해 프로브 목록을 Agilent Sure 설계 웹 사이트에 업로드 한다.
- [0310] 7. 프로브 그룹을 사용하여 Agilent 맞춤형 CGH 어레이를 설계한다.
- [0311] 어레이 공정
- [0312] 1. 주형 제작을 위해 EpiSwitch™ Standard Operating Procedure(SOP)를 사용하여 시료를 처리한다.
- [0313] 2. 어레이 공정 실험실에서 에탄올 침전으로 세척한다.
- [0314] 3. Agilent SureTag complete DNA 표지 키트에 의해 시료를 처리한다. - 혈액, 세포 또는 조직에 대한 게놈 DNA 분석 효소적 표지를 위한 Agilent Oligonucleotide Array-based CGH .
- [0315] 4. Agilent 기능 추출 소프트웨어를 사용한 Agilent C 스캐너를 사용하여 스캔한다.
- [0316] **EpiSwitch™ 기술 개요**
- [0317] EpiSwitch™ 플랫폼은 비정상적이고 반응적인 유전자 발현과 관련된 주요 질병의 선별, 조기 발견, 동반자 진단, 모니터링 및 예후 분석에 매우 효과적인 수단을 제공한다. 이 접근법의 주요 장점은 비 침습적이고 빠르며 불안정한 단백질 / RNA 분자보다는 염색체 시그니처의 일부로 매우 안정적인 DNA 기반 표적에 의존한다는 것이다.
- [0318] EpiSwitch™ 바이오마커 시그니처는 복잡한 질병 표현형의 계층화에서 높은 견고성, 감도 및 특이성을 보여준다. 이 기술은 매우 유익한 종류의 후생유전학적 바이오마커로서 후생유전학, 염색체 형태 시그니처의 모니터링 및 평가에 있어 최신 혁신을 활용한다. 학계 환경에 배치된 현재의 연구 방법론은 CCS를 검출하기 위해 세포 물질의 생화학적 공정에 3일에서 7일이 필요하다. 이러한 공정은 감도와 재현성이 제한적이고; 더욱이 실제 단계에서 EpiSwitch™ 분석 패키지가 제공하는 표적 통찰력의 이점이 없다.
- [0319] ***in silico* 마커 식별의 EpiSwitch™ 어레이**
- [0320] 게놈 전체의 CCS 부위는 모든 관련 계층화 유도 바이오마커의 식별을 위해 코호트 시험의 임상 시료에서 EpiSwitch™ 어레이에 의해 직접 평가된다. EpiSwitch™ 어레이 플랫폼은 처리량이 많고, 많은 유전자좌를 신속하게 스크리닝 할 수 있기 때문에 마커 식별에 사용된다. 사용된 어레이는 Agilent 맞춤형 CGH 어레이였고, 이를 통해 *in silico* 소프트웨어를 통해 식별된 마커를 조사 할 수 있다.
- [0321] **EpiSwitch™ PCR**
- [0322] EpiSwitch™ 어레이로 식별된 잠재적 마커는 EpiSwitch™ PCR 또는 DNA 서열분석기(즉, Roche 454, Nanopore MinION 등)에 의해 검증된다. 통계적으로 유의하고 최상의 재현성을 나타내는 최고의 PCR 마커는 최종 EpiSwitch™ 시그니처 세트에 추가적 축소가 되기 위해 선택되며, 시료의 독립적인 코호트에서 검증된다. EpiSwitch™ PCR은 확립된 표준화된 운영 절차 프로토콜에 따라 숙련된 기술자에 의해 수행될 수 있다. 모든 프로토콜 및 시약 제조는 ISO 13485 및 9001 인증에 따라 수행되어 작업 품질과 프로토콜 전송 기능을 보장한다. EpiSwitch™ PCR 및 EpiSwitch™ 어레이 바이오마커 플랫폼은 전체 혈액 및 세포주 분석과 호환된다. 이 테스트는 소량의 혈액을 사용하여 매우 적은 수의 카피 수에서 이상을 감지 할 정도로 민감하다.
- [0323] **ALS 코호트 분석**
- [0324] 본 발명자들은 ALS에서 동반 진단 방법으로서 사용될 바이오마커를 식별하기 위한 기초로서 후생유전학적 염색체 상호 작용을 사용하였다. EpiSwitch™ 바이오마커 발견 플랫폼은 본 발명자들에 의해 ALS에 관련된 표현형 변화를 유도하는 것과 같은 후생유전학적 조절 시그니처 변화를 검출하기 위해 개발되었다. EpiSwitch™ 바이오마커 발견 플랫폼은 환경 신호를 후생유전학적와 전사적 기계에 통합하는데 있어 초기 규제 프로세스를 정의하는 CCS를 식별한다. 따라서 CCS는 일련의 유전자 조절에서 주요 단계이다. EpiSwitch™ 바이오마커 발견 플랫폼에 의해 분리된 CCS는 다음과 같이 잘 기록된 여러 장점을 가지고: 심각한 생화학 및 생리학적 안정성; 이진법적(binary) 특성 및 관독; 진핵의 일련의 유전자 조절에서 그들의 주요 위치를 가진다.

- [0325] 유전자 발현의 분류 후생유전학적 제어에서 불안정의 원인을 검출하는 능력은 ALS의 조기 진단 및 환자 예후의 확립을 가능하게 한다. 건강하거나 질병이 있는 환자 혈액 시료에서의 세포 조절 계놈 구조물의 비교 질문은 두 ALS 관련 후생유전학적 시그니처: 첫째 진단 잠재력과 둘째 예후 예측을 게시한다. 이 전향적 연구를 위해 옥스퍼드 대학의 Nuffield Department of Clinical Neurosciences(NDCN)의 Oxford Motor Neuron Disorders Clinic에서 임상 연구 그룹에 의해 시료가 수집되었고, 염색체 형태 시그니처 모니터링을 위해 Oxford BioDynamics가 개발한 고 처리량 플랫폼인 EpiSwitch™을 사용하여 분석되었다. 이 연구는 ALS-FRS-R, FVC(Forced Vital Capacity)의 임상 주석과 후생유전학적 시그니처에 대한 분석과 함께 ALS 진행 서브타입을 할당하기 위한 다른 임상 관찰을 비교한다.
- [0326] Oxford Motor Neuron Disorders Clinic에 제출된 총 100명의 환자가 해당 연구에 등록했으며 3개월과 6개월에 재방문하도록 요청되었다. 대조군(n=100)을 수집되었다. 각 방문 동안, 참가자들은 ALS-FRS-R 및 FVC 시험을 받고 혈액 샘플을 제공했다. ALS 질환 관련 진단 시그니처 또는 예후 질환 관련 시그니처(3개월 및 6개월)를 식별하기 위해 시료가 분석되었다. 임상 평가 결과가 0, 3 및 6개월에 EpiSwitch™ 분석과 비교되었다. ALS-FRS-R 점수의 매월 0.5-포인트 감소의 컷오프가 ALS 환자를 진행 서브타입으로 클러스터링 하는데 사용되었다.
- [0327] ALS-FRS-R 점수의 감소율에 기초하여, 3개월 시료 및 6개월 시료로부터의 예비 결과는 80%의 민감성 및 특이성을 갖는 진행성 ALS 환자가 더 빨리(> 0.5) 더 느리게(<0.5) 진행되는 선택된 후생유전학적 시그니처를 시연하였다. 현재까지 결과는 예후 시그니처가 시간이 지남에 따라 ALS의 서브타입을 선택하는데 강력하다는 것을 나타낸다. 표에 나타낸 결과는 독립적인 시료 코호트(n=50)를 갖는 75 ALS 및 75 대조군 시료(n=150)의 진단 및 예후 분석에 관한 것이다.
- [0328] 추가 결과 (표 11 포함)는 기준선, 3개월 및 6개월 생존에서 ALSFRS-R 점수 및 FVC (강제 생명력)의 표준 관행에 의해 관찰되고 평가된 ALS 환자 100명을 기초로 한다. 기준선 및 3개월에서의 측정에 기초하여, 이들 환자는 느리거나 빠른 ALS (표준 분류)로 분류되었다. 기준선에서만 3개의 마커 판독 값을 사용하여, 또한 모든 환자를 느리거나 빠른 것으로 분류했다 (EpiSwitch 분류). 이 환자들은 6개월 생존율을 기준으로 평가되고, 느리고 빠른 그룹의 생존율이 비교되었다. 표준 분류는 느린 그룹에서 상당한 수의 치명적인 결과를 가졌으며 통계적으로 느린 것과 빠른 사이의 p 값에 의한 생존 차이는 p=0.052으로서 유의하지 않았다. EpiSwitch 분류에 따르면, 6개월에서 느리고 빠른 분리는 p=0.0097로 매우 중요했으며, 대부분의 사망은 빠른 그룹에서 발생했다.
- [0329] <실시 예 2>
- [0330] HTT 유전자 좌위의 계놈 구조 차이는 헌팅턴병의 증상 및 전증후군 사례에 기초한다.
- [0331] 진행성 퇴행성 신경 질환인 헌팅턴병(Huntington's disease, HD)은 성인 뇌에서 광범위한 뉴런의 퇴행을 유발하여 궁극적으로 사망에 이르는 유전적 장애이다. HD의 근본 원인은 "헌팅턴 유전자" (HTT)에서 확장된 트리뉴클레오티드 시토신-아데닌-구아닌 (CAG) 반복이며, 이는 세포 내 응집체를 형성하고 신경 사멸을 유발하는 경향이 증된 돌연변이 헌팅턴 단백질을 초래한다. 중요하게도, CAG 반복 횟수는 질병 발병 및 중증도와 상관 관계가 있고, 39회를 초과하는 CAG 반복 환자는 그들의 삶의 어느 시점에서 HD가 발병 할 것이며, 질병의 발병은 개인 별 수십 년씩 달라질 수 있으며, 이러한 전증후 단계에 대해서는 거의 알려져 있지 않다. 유전자 구조 이해의 일환으로 DNA 메틸화, 히스톤 변형, 넌코딩 RNA 및 염색체 형태 시그니처에 대한 후생유전학적 조절 마커의 검출을 위한 후생유전학적 접근법의 발전과 새로운 기술의 출현으로, 발명자는 HTT 유전자 주위의 계놈 구조 및 질병 발현과의 관계의 전신 조절 변화를 분석하는 데 관심이 있었다. 여기서, 발명자는 영향을 받지 않는 대조군과 비교하여 전증후와 증상 HD 환자 사이의 HTT 유전자좌에서 감지적, 전신적, 비 침습적 후생유전학적 변이를 조사했다.
- [0332] 방법 : HD 및 건강한 대조군 환자의 혈액 시료를 사용하여, HTT 유전자의 바로 근처에서 염색질 구조를 평가하는 염색체 형태의 검출을 위해 검증된 고해상도 산업 플랫폼 인 EpiSwitch™를 사용했다. 발명자는 HTT 유전자 좌의 225 kb에 걸친 20 상호 작용 사이트에서 조건적, 안정적인 염색질 형태의 부재 또는 존재를 평가하고, 건강한 대조군 그룹, 확인된 증상이 있는 HD 환자 (CAG, n> 39)와 HD의 임상 증상을 나타내지 않은 유전적으로 검증된 CAG 연장을 가진 환자 간의 결과 염색체 형태를 비교하였다.
- [0333] 결과 : 말초 혈액 시료에서 테스트할 때 환자 그룹에서 일관되고 안정적인 염색체 형태가 관찰되었다. 발명자는 두 가지 구성적 상호 작용 (모든 환자군과 대조군에서 발생)과 HD에는 존재하지만 건강한 대조군에는 존재하지 않는 7개의 조건부 상호 작용을 발견했다. 가장 중요하게는, 증상 전증후가 없는 사례에는 없지만, 명백한 임상 증상 (증상 사례)을 나타내는 HD 환자에서만 존재하는 3가지 조건부 상호 작용을 관찰했다. 증상이 있는 HD 코

호트 환자의 85%(7명 중 6명)는 증상이 있는 HD와 관련된 특정 염색체 형태 중 적어도 하나 이상을 보여주었다.

[0334] 결론 : 우리의 결과는 HTT 유전자 좌위의 조절 염색질 구조가 HD 환자에서 체계적으로 변경되었다는 첫 번째 증거다. 더우기, HTT 유전자좌에서의 염색질 구조는 증상 및 전증후군 HD 환자에서 임상 단계 사이의 조건부 차이를 또한 보여준다. HD에서 질병 진행을 평가하기 위한 분자 도구를 갖는 높은 임상적 유용성을 고려할 때, 이러한 결과는 전신 염색체 형태 시그니처 (CCS)의 비 침습적 평가가 HD 환자의 예후 평가에 귀중한 추가물이 될 수 있음을 강력하게 시사한다.

[0335] **서론**

[0336] 헌팅틴병 (HD)은 기저핵에서 뉴런의 상실과 세포의 통제되지 않은 움직임, 정서적 문제 및 인지 상실에 의해 세포적으로 특징되는 퇴행성 신경 질환이다. HD는 상염색체 우세 유전적 장애이며 유병률은 지역과 민족에 따라 다양하지만, 미국과 유럽에서만 5만명 이상에게 영향을 미치는 것으로 생각된다. 근본적인 유전적 원인은 1983년 Massachusetts General Hospital에서 James Gusella에 의해 유전적 마커로 발견된 헌팅틴 유전자(HTT)의 트리뉴클레오티드 CAG 확장으로서, 독성 폴리-글루타민(polyQ) 트랙을 갖는 돌연변이 헌팅틴 단백질 (mHTT)을 생성시킨다. 그러나 수십 년에 걸친 연구와 임상 시험에도 불구하고, 성공적인 치료법은 아직 개발되지 않았다. HD의 "전형적인" 시작은 40-50세 사이이지만, 최대 15%의 사례가 매우 늦게 시작하여 60세 이후까지 임상 증상을 나타내지 않는다. 최근 조사 연구의 메타 분석에서 후기 발병 HD (LoHD) 사례의 경우, 90% 이상의 환자가 CAG 반복 길이가 ≤ 44 였다. HD에서 더 흥미로운 관찰 중 하나는 polyQ 반복 트랙의 길이와 질병의 시작 및 심각성 사이에 잘 알려진 상관관계가 있지만, 개별 환자들 사이에는 상당한 변동성이 있다는 것이다. 예를 들어, 중간 범위의 반복 길이를 가진 환자 (여기서는 40에서 50 사이로 정의된)에서 질병 발병은 개별 환자마다 60년까지 달라질 수 있다. 이것은 질병의 발병에 영향을 미치는 polyQ 트랙의 운반자인 많은 환자들이 수십 년 동안 "전증후" 상태에서 살 수 있음을 의미한다. 임상 증상의 발병을 제어하는 것은 현재 알려져 있지 않으며 HD 환자의 예후 평가를 복잡하게 한다.

[0337] 역사적으로 단순성 질환으로 간주되었지만, HD의 기본 병리에 대한 광범위한 연구는 질병의 발병 및 진행으로 이어지는 메커니즘이 원래 생각보다 복잡하다는 것을 시사한다. 유전자 발현, 단백질학, 대사체학, 네트워크 분석, 유전체학 및 단일 뉴클레오티드 다형성 (SNP) 프로파일링을 포함하여 HD에서 질병 진행의 기본 분자 변화를 관찰하기 위해 많은 다른 기술이 사용되었다. HD가 후생유전학적 조절 불량을 특징으로 하는 질병의 패러다임으로 간주됨에 따라, 최근에는 후생유전학적 접근법이 병리학 관련 변화를 평가하기 위한 유망한 새로운 도구로 등장했다. HD에서 대부분의 후생유전학적 연구는 HD와 관련된 특정 유전자좌에서의 게놈의 넓은 히스톤 변형 (아세틸화, 메틸화) 또는 히스톤 변형을 조사하는 데 중점을 두었다. 이러한 접근 방식은 질병에 대한 흥미로운 통찰력을 제공하지만, 종종 상충되는 결과를 낳고 마우스 모델과 인간 질병간에 불일치를 나타냈다. 따라서, 히스톤 변형 관독을 이용한 HD에서의 후생유전학적 비조절에 대한 일치적 그림은 아직 실현되지 않았다. 그러나, 후생유전학적 조절과 관련된 모든 분자 메커니즘이 HD와 관련하여 평가된 것은 아니다. 후생유전학적 조절의 중요한 측면은 3차원(3D) 유전체 구조의 수준에 있다.

[0338] 게놈의 3D 구성은 외부 환경 신호 및 입력의 이중 효과를 반영하며 염색체 형태의 평가로 또는 여러 형태가 염색체 형태 시그니처 (CCS)로 동시에 측정 될 때 경험적으로 측정될 수 있다. CCS는 주어진 세포 집단의 후생유전학적 지형을 관독하는 분자 바코드로 생각할 수 있다. HD의 개발에서 mHTT의 중심적인 역할을 고려할 때, 발명자는 HTT 유전자 좌위의 게놈 구조의 조절 차이가 병에 걸린 개인과 건강하고 영향을 받지 않은 대조군 사이에 존재할 수 있다는 가설을 세웠다. 발명자는 CCS 모니터링을 위해 확립된 독점적 산업 플랫폼 인 EpiSwitch를 사용하여 전증후군 및 증상이 있는 HD 환자와 건강에 영향을 받지 않는 개인 간의 염색질 구조 차이를 평가했다. EpiSwitch 관독 값은 CCS의 고해상도, 신뢰성 및 높은 처리량 감지 기능을 제공하는 동시에 품질 관리에 대한 업계 표준을 충족시킨다.

[0339] **방법**

[0340] **시료 수집**

[0341] 모든 시료는 National BioService, LLC로부터 입수하였다. 총 20개의 샘플이 연구에 사용되었는데; 10개의 건강한 대조군 시료(HC) 샘플(CAG 반복, $n < 35$) 및 10개의 HD 시료(CAG 반복, $n > 39$)다. HD 샘플의 경우, 7명은 증상 환자(HD-Sym)고, 3명은 HD 진단을 받았지만 아직 임상 증상을 보이지 않은 전증후군 환자(HD-Pre)다. 한 명의 HD 환자는 테트라베나진(tetrabenazine)을 복용하고 있었고 한 명의 환자는 세르트랄린(sertraline)을 복용하고 있었다. 모든 샘플은 인간 면역 결핍 바이러스, B형 간염 바이러스, C형 간염 바이러스 및 매독에 대해 음

성이었다 (표 14).

[0342] **연구 설계**

[0343] 발명자는 건강한 대조군(낮은 CAG), 전증후군 HD 환자(높은 CAG, 질병 징후 없음) 및 증상적 HD 환자(높은 CAG, 질병 징후) 사이에 다른 염색체 형태를 식별하고자 했다. 발명자는 분석을 위해 HTT 유전자 좌위를 둘러싸는 ~225 kb 영역(chr4 : hg38에서 주석이 달린 3,033,588에서 3,258,170까지)에 중점을 두었다. 앵커 포인트("앵커(Anchor)")로 HTT의 엑손 1(chr4 : 3,054,162 ~ 3,095,930)의 CAG 반복 확장 트랙을 사용하여, 앵커를 둘러싼 5개의 게놈 영역을 정의하여 시료 그룹간에 다양한 염색체 형태를 확인했다 (도면 2 및 표 15). 이 영역은 잠재적 EpiSwitch 앵커링(anchoring) 사이트의 존재; 알려진 질병 관련 SNP(HD 및 기타 질병)의 존재; 그리고 GWASdv2 데이터베이스 (<http://jjwanglab.org/gwasdb>)에서 발견된 HD에서 알려진 히스톤 변형 사이트(H3K4me3, H3K36me3 및 H3K27ac)의 강화(도면 2)를 기반으로 선택되었다.

[0344] **염색체 형태 동일성**

[0345] NCBI 검색 : 이전에 보고된 HD 후생유전학적 데이터에 대한 GEO 데이터베이스는 2018년 2월에 수행되었다. 사후 전두엽 피질 뇌 시료(침대 형식)의 12개(6개의 HD 및 6개의 대조군 시료)에서 H3K4me3에 대한 ChIP-seq 데이터 피크가 획득되었다(GSE68952). 또한 HD iPSC 유래 신경 세포주 및 대조 세포주에서 H3K27ac 및 H3K36me3에 대한 ChIP-seq 데이터의 Bigwig 트랙이 다운로드 되었다(GSE95342). 데이터 트랙은 EpiSwitch 및 참조 서열 주석과 함께 Integrative Genome Viewer (IGV) [41]에 로드(load)되었다. 5개의 관심 영역을 식별하기 위해 HTT 유전자 좌위에서 시각적 및 프로그래밍적 (BEDtools) 비교가 수행되었다.

[0346] 적절한 소프트웨어가 CAG 반복에 근접한 앵커 영역에서 발생하는 하나의 "말단"과 관심있는 5개의 영역 중 어느 하나와 높은 확률의 염색질 접힘 상호 작용을 식별하는데 사용되었다. 총 61개의 상호 작용이 이 기준과 일치했으며, 실제적인 이유로, 앵커 사이트와 모든 관심 영역 간의 상호 작용을 다루기 위해 20개의 상호 작용이 선택되었다. 염색체 형태 포착 분석을 수행할 때 상호 작용에 의해 야기되는 예상되는 DNA 서열을 증폭시키는 올리고뉴클레오티드 쌍을 설계하기 위해 자동화된 프라이머 설계 어플리케이션이 사용되었다.

[0347] **3C와 PCR**

[0348] 염색체 형태 포착 및 PCR에 의한 검출을 수행되었다. 각각의 환자 시료 50 μ l의 혈액 시료로부터 온전한 *?***? 염색체 형태를 갖는 염색질을 제조사의 지시(Oxford BioDynamics Plc)에 따라 EpiSwitch 분석을 사용하여 추출하였다. 모든 시료에 대한 품질 관리는 3C 분석을 위한 역사적인 내부 대조군인 MMP1 유전자좌에서 염색질 루프의 검출을 사용하여 수행되었다. 각 시료 서브그룹에 대해 일반화된 모집단 시료를 제공하기 위해 각 샘플 유형에 대한 풀링된 3C 라이브러리가 생성되었다. 샘플 유형간에 상이한 PCR 생성물 검출 패턴과의 상호 작용을 식별하기 위해 CFX-96 (Bio-Rad) 기계의 SYBR green으로 실시간 PCR이 수행되었다. 올리고뉴클레오티드를 대조군 주형에서 시험하여 각 프라이머 세트가 올바르게 작동하는지 테스트하였다. PCR 검출을 위한 로열 포렌식 프로토콜 (Royal Forensic Protocol)에 따라, 개별 HD 환자에 대한 후속 데이터를 위해 각 샘플에 대해 최종 증점 PCR이 시료당 3회 수행되었다. 이 절차를 통해 제한된 사본수의 주형을 보다 정확하게 탐지 할 수 있었다. LabChip DNA 1K Version2 키트 (Perkin Elmer)를 사용하여 Perkin Elmer의 LabChip® GX에서 모든 PCR 증폭 생성물이 모니터 되었고 형광 염료를 사용하는 제조업체의 프로토콜에 따라 DNA 칩에 내부 DNA 마커가 로드되었다. 형광은 기기 소프트웨어를 사용하여 겔 화상에서 모의 밴드로 변환된 레이저 및 심전도 판독에 의해 검출되었다. 기기의 검출 임계값은 제조사에 의해 30 형광 단위 이상으로 설정되었다.

[0349] **통계 분석**

[0350] 데이터 분석은 R (통계 컴퓨팅을 위한 언어 및 환경)에서 수행되었다. 여기에는 t-테스트 및 R-제곱 분석을 위한 통계 및 dplyr 패키지와 상자 도면 및 회귀 도면을 위한 ggplot2 패키지가 포함되었다.

[0351] **결과**

[0352] **환자 임상 특성**

[0353] HC 및 HD 시료는 연령(HC의 경우 평균 36.9세, HD의 경우 35.3세)이고 성별은 일치(남성 1/2 및 여성 1/2)했고, 증상을 가진 대부분의 HD 사례(70%)였다 (표 16). 모든 시료는 비히스패닉 또는 라틴계 백인이었다. 평균 CAG 반복 길이는 HC의 경우 25.7, HD의 경우 44.2였다(표 16). HC 및 HD 시료는 연령에서 통계적 차이를 나타내지 않았다. HC 환자는 HD-Sym과 연령의 통계적 차이가 없었다. HD-Pre 환자는 HD-Sym 환자(평균 연령=39.6)보다 더 젊었다(평균 연령=25.3) (p=0.02). 통계적으로 유의미한 것은 아니지만, 질환 지속 기간과 CAG 반복 크기 사이

에는 온건한 부정적인 관계가 있었으며 진단시 연령과 CAG 반복 크기 사이에는 온건한 긍정적인 관계가 있었다.

[0354] HC($p=1.08E^{-7}$)에 비해 HD 환자에서 CAG 반복 길이가 통계적으로 유의하게 증가하였다. HC와 HD-Pre($p=3.43E^{-6}$)와 HD-Sym($p=9.50 E^{-8}$) 사이에서 CAG 반복 길이가 통계적으로 유의하게 증가했다. HD-Pre와 HD-Sym($p=0.09$) 사이의 CAG 반복 길이에 통계적 차이가 없었다.

[0355] HD 시료의 진단시 평균 연령은 35.3세였으며 평균 질병 기간은 3.8년이었으며 10명 중 7명은 과민성, 무도병 또는 두 가지 증상이 모두 보고되었다(표 16).

[0356] **HC, HD-Pre 및 HD-Sym의 염색체 형태**

[0357] 평가된 20개의 상호 작용 중 (표 17), 발명자는 9개의 유의한 상호 작용을 식별했다. 발명자는 HD에 존재하지만 건강한 대조군이 아닌 두 가지 구성적 상호 작용과 7개의 조건부 상호 작용을 식별했다. 7개의 조건부 상호 작용 중 3개는 HD-Sym에만 있었고, HD-Pre에는 없었다.

[0358] **구성적 형태**

[0359] 모든 샘플은 MMP1 상호 작용에 대한 내부 QC 분석을 통과했다. 2개의 구성 적(모든 샘플에서 식별) 염색질 루프가 식별되었다. 양 루프 모두 첫 번째 루프는 28kb, 두 번째 루프는 34kb인 앵커(Anchor)와 영역 2 사이에 있었다. 모든 환자에서 발생하는 2 가지 구성적 상호 작용 (HD-Sym, HD-Pre 및 HC)이 이 연구에서 관찰되었다. 양 상호 작용(CC1와 CC2)은 CC1이 28kb, CC2가 34kb인 앵커와 Zone 2 사이에 있었다.

[0360] **조건적 형태**

[0361] 발명자는 이 연구에서 평가된 상이한 환자 서브그룹을 구별할 수 있는 7개의 조건부 염색체 형태를 식별했다. 특별히, 본 발명자들은 HC에 존재하지만 모든 HD 시료에는 없는 2개의 염색체 상호 작용을 식별하였다. 첫 번째 상호 작용(I1)은 앵커와 영역 4에 걸쳐 있고 77kb를 커버했고, 두 번째 상호 작용(I2)은 앵커와 영역 3에 걸쳐 있고 140kb를 커버했다. 발명자는 또한 HC와 HD-Pre에 존재했지만 HD-Sym 시료에는 없는 두 가지 염색체 상호 작용을 확인했다. 양 상호 작용(I3 및 I4) 모두 앵커와 영역 4에 걸쳐 있고 각각 92kb와 104kb를 커버했다. 마지막으로, 발명자는 HD-Sym 시료에는 존재했지만 HD-Pre 및 HC 시료에는 없는 3가지 염색체 상호 작용을 식별했다. 첫 번째 상호 작용 (I5)은 122kb를 커버하며 앵커와 영역 3에 걸쳐 있다. 흥미롭게도, 이 상호 작용에는 HD를 개발하기 위한 소인의 인자로 알려진 SNP(rs362331)가 포함되었다. 두 번째 및 세 번째 상호 작용(I6 및 I7)은 앵커와 영역 1에 걸쳐 있고 각각 185kb 및 174kb를 커버했다. 마지막으로, 발명자는 개별 HD 샘플에서 모든 조건부 상호 작용의 유무를 테스트했다. HD-Sym 시료에서, 발명자는 7개 중 6개 시료에서 하나 이상의 조건부 마커(I5, I6 및 I7)가 있음을 발견했다 (도면 3). 이 연구에서 평가된 모든 상호 작용에 대한 요약은 도면 4와 도면 5에 나타난다. 표 18은 우열차이 비를 나타낸다. 표 13은 서브그룹간에 차이가 없는 염색체 상호 작용을 보여준다.

[0362] **토론**

[0363] **문제 설명 및 결과 요약**

[0364] 본 발명은 ALS 및 헌팅틴병과 관련된 염색체 영역 및 상호 작용을 분석하는 공정으로서, CAG 반복 횟수가 39회를 초과하는 개인이 HD를 앓는 것으로 잘 알려져 있지만, 질병의 임상적 발병은 개별 환자마다 다양하며, 질병이 임상적으로 나타날 때 영향을 미치는 인자는 덜 특성화되어 있다. HD 환자와 건강하고 영향을 받지 않는 대조군에서 HTT 유전자좌의 후생유전학적 관점을 판단하기 위해 염색질 구조를 평가하기 위한 산업 플랫폼 EpiSwitch를 사용했다. CCS로 합쳐졌을 때 HD를 영향을 받지 않은 대조군과 구별할 수 있고, 더 중요한 것은 전 증후군과 증상이 있는 HD 환자를 구별할 수 있는 7가지 상호 작용을 식별했는 것이다. 증상 HD에 특이적인 이들 상호 작용 중 하나는 소인성 질환 할로 그룹(halogroup)과 연관된 것으로 나타난 SNP(rs362331)를 포함한다. 종합하면, 이들 결과는 CCS를 평가하는 간단한 비 침습적 혈액-기반 시험이 HD에서 질환 진행을 평가하기 위한 대리 바이오마커로서 작용할 수 있음을 보여준다.

[0365] **생물학적 관련성**

[0366] 폴리-Q 반복 트랙 확장 및 mHTT의 생성이 HD의 근본 원인인 것으로 알려져 있지만, 임상 증상의 발달로 이어지는 분자적 사건은 덜 특성화되어 있다. 여러 연구에서 HTT 유전자좌 내의 SNP를 질병 발병의 잠재적인 기여자로 보았다. HD 환자에 대한 최근의 한 SNP 유전자형 연구에서 HTT 유전자의 엑손 50에 있는 rs362331 C/T SNP를 포함하여 환자의 적어도 30% 이상에서 ~ 41개의 SNP 이종접합성이 식별되었다. 아마도 생물학적으로 더 관련이 있

는 것은 안티센스 올리고뉴클레오타이드, siRNA 또는 miRNA를 사용하여 rs362331 SNP의 대립 유전자가 선택적으로 녹다운될 때, mHTT 단백질 수준의 급격한 감소가 체외 및 생체 내에서 달성됨으로서 이 SNP와 그 주변이 게놈 환경은 mHTT 수준을 조절하는 데 중요한 역할을 하는 것이다. 이 연구에서 발명자는 HD-Sym 환자에게만 존재하고 rs362331 SNP와 겹쳐진 HD-Asy 및 HC에는 부재한 염색체 형태(I5)를 관찰했다. 이것은 poly-Q 트랙을 증가시킨 환자에서 신경 독성 mHTT의 생성 및 rs362331 SNP의 존재에 의한 HD의 초기 발달에 대한 유전적 소인이 고차 염색질 구조의 수준에서 조절 될 수 있음을 시사한다. HD의 또 다른 뛰어난 질문은 부모 중 어느 것도 진단을 받지 않은 경우 질병이 어떻게 유전되는지 이다. 두 가지 주요한 가설은 1) 운반체 부모가 질병의 발병 전에 다른 요인으로부터 사망했을 수 있고 2) "불안정한" CAG 반복 트랙이 각 세대마다 확장된다는 것이다. 세 번째 가능성은 중간 범위(35-50) 반복에서, 개인이 질병의 징후 없이 보인자가 될 수 있지만, 그들의 자손이 정의되지 않은 메커니즘을 통해 유전적 결함을 보상 할 수 없고 질병이 발현될 것이라는 것이다. 이 연구에서 평가된 HD 환자는 모두 이 중간 범위에서 CAG 반복을 가졌으며, 게놈 구조의 차이를 통해 질병 발생의 잠재적 보상 메커니즘이 매개될 가능성을 높은 것이다.

[0367] **임상적 관련성**

[0368] HD는 질병, HTT 유전자 서열분석 및 CAG 반복수의 측정을 결정하기 위한 간단한 테스트가 있다는 점에서 독특하다. 임상 치료 및 임상 시험의 경우, Unified Huntington 's Disease Rating Scale(UHDRS), Shoulson-Fahn Scale 및 Mini-Mental State Examination(MMSE)과 같은 질병 심각도를 측정하기 위한 여러 테스트도 있다. 이러한 평가는 질병 심각도에 대한 대리자로 HD 환자 신체적 및 정신적 복지의 여러 요소를 측정하지만, 모두 주관적이며 대부분 HD에 국한되지 않는다. 빠진 것은 질병 진행을 모니터링하는 구체적인 분자 도구이다.

[0369] 전임상 및 임상 개발의 상이한 단계에서 HD를 치료하기 위한 현재 22개의 치료제가 있으며, 그 중 절반은 임상 2상 또는 3상에 있다. 추가로 검증되면, 본원 발명에 보고된 CCS는 임상 시험에서 해당 의약품의 치료 효능을 평가하기 위한 대리 결과 바이오 마커로 사용될 수 있다. 임상 시험에서 특정 요법에 대한 증상 환자의 반응을 모니터링하는 것 외에, 본원발명에 설명된 접근법의 또 다른 장점은 전증후 환자에 대해 얻을 수 있는 정보에 있다. 대부분의 HD 환자의 경우 전증후 기간은 수십 년 지속될 수 있다. 여기에서 확인된 7개의 상호 작용 중 5개(i3-i7)는 전증후 환자와 증상이 없는 환자를 명확하게 분리하며, 추가로 검증되면 전증후 운반자에서 HD 증상의 발병에 대한 “조기 경고” 지표 시험으로 작용할 수 있다.

[0370] 이 연구는 HD의 발현에 특이적이고 알려진 질환 할로타입(halotype)과 상관 된 염색질 구조에서 검출 가능한 조 건부 차이의 첫 번째 증거를 제공한다. 이 연구의 주요 강점은 게놈 구조의 조절 역할을 이해하는 최신 개발을 기반으로 하는 독특한 접근 방식에 있다. 임상, 영상 및 분자 측정을 기반으로 질병 진행 바이오마커를 개발하는 것을 목표로 HD에 여러 역사적 연구가 있었지만, 우리가 아는 한, 임상적으로 접근 가능한 생체액에서 고차 염색질 구조의 평가가 HD로 적용되는 처음인 것이다.

[0371] [표 1a]

Probe	GeneLocus	Probe_Count_ Total	Probe_Count_ Sig
C6orf58_6_127480771_127483471_127600017_127604343_FR	C6orf58	10	5
NEFH_22_29442588_29445314_29482081_29484217_RR	NEFH	16	10
IL1A_2_112765786_112772711_112810765_112813086_RR	IL1A	23	15
FARP1_13_98271575_98282700_98346930_98348486_FR	FARP1	26	26
PRKCA_17_66441276_66447067_66475597_66481312_RF	PRKCA	30	16
RNU6-1264P_17_6162286_6163870_6195952_6199184_FR	RNU6-1264P	31	16
PON2_7_95405100_95420940_95465337_95474032_FR	PON2	40	19
CAPN9_1_230738572_230739927_230752057_230757333_RR	CAPN9	50	23
ATXN7L1_7_105654123_105657510_105741521_105750599_FR	ATXN7L1	130	61
CNTNAP2_7_146728706_146734820_146785878_146792823_RF	CNTNAP2	144	60
CTNNA3_10_66299269_66302507_66496211_66513003_FF	CTNNA3	151	46
ZFPM2_8_105632010_105638904_105814873_105824107_FR	ZFPM2	186	68
ALDH1A2_15_58325151_58334051_58485549_58488054_FR	ALDH1A2	185	70
ALDH1A2_15_58325151_58334051_58538695_58540885_FF	ALDH1A2	185	70
MAGI2_7_79009346_79018304_79275810_79284623_RF	MAGI2	186	80
FER1L6_8_123963222_123969450_124085753_124093275_FR	FER1L6	98	13
UBQLN2_X_56536168_56538402_56570114_56575112_FR	UBQLN2	10	7

[0372]

[0373] [표 1b]

HyperG_Stats	FDR_HyperG	Percent_Sig	logFC	AveExpr	t
0.192504768	0.761792372	50	-0.206940905	-0.206940905	-7.657670128
0.012576275	0.174966022	62.5	0.155885782	0.155885782	12.07017614
0.001213793	0.03415703	65.22	0.130559454	0.130559454	3.67899153
0.000000000000171	0.000000000079	100	0.254822215	0.254822215	14.73928749
0.013706774	0.176284348	53.33	0.24251793	0.24251793	7.451573283
0.019964247	0.210078331	51.61	-0.152642021	-0.152642021	-9.03898498
0.032185744	0.276690843	47.5	-0.178366996	-0.178366996	-5.268810519
0.029727242	0.269876723	46	-0.224588757	-0.224588757	-8.238208751
0.000354202	0.018221718	46.92	0.189559955	0.189559955	6.597539924
0.011284771	0.170730832	41.67	-0.440614155156995	-0.440614155156995	-7.4296031698139
0.713857099	1	30.46	-0.436584531378388	-0.436584531378387	-14.2285764776972
0.122432344	0.622894915	36.56	-0.425339105073479	-0.425339105073479	-10.493629902935
0.063739285	0.421589843	37.84	0.34586199024656	0.34586199024656	13.4772416929847
0.063739285	0.421589843	37.84	0.339861526563037	0.339861526563037	7.22282802700005
0.00139007	0.035755702	43.01	0.34944598213495	0.34944598213495	10.5772811324938
0.999996196	1	13.27	0.350542987884941	0.350542987884941	5.66944350076527
0.016370015	0.18948292	70	0.353892612578078	0.353892612578078	2.69374295077173

[0374]

[0375] [표 1c]

P.Value	adj.P.Val	B	FC	FC_1	LS	Loop detected
0.000000338	8.58E-06	6.769090112	0.866372344	-1.154238137	-1	Control
0.000000000262	7.50E-08	13.89074006	1.114105444	1.114105444	1	ALS
0.001613314	0.006000711	-1.70129901	1.094718133	1.094718133	1	ALS
0.00000000000858	1.48E-08	17.19553875	1.193188696	1.193188696	1	ALS
0.0000005	1.15E-05	6.37489208	1.183055643	1.183055643	1	ALS
0.0000000281	1.46E-06	9.26023757	0.899601504	-1.111603299	-1	Control
0.000045	0.000343598	1.852676493	0.883702704	-1.131602286	-1	Control
0.000000115	3.91E-06	7.847763841	0.85583895	-1.168444134	-1	Control
0.0000027	3.98E-05	4.678642967	1.140415818	1.140415818	1	ALS
5.21452107638594E-07	1.18E-05	6.332521031	0.736820877	-1.357181957	-1	Control
1.58303956750015E-11	1.70E-08	16.61028384	0.738881785	-1.353396471	-1	Control
2.63975825723201E-09	2.70E-07	11.61504783	0.744663678	-1.342888111	-1	Control
4.03123106207005E-11	2.85E-08	15.71064142	1.270910107	1.270910107	1	ALS
7.77879601584934E-07	1.61E-05	5.930445637	1.265635109	1.265635109	1	ALS
2.32055637317047E-09	2.48E-07	11.7426843	1.27407127	1.27407127	1	ALS
0.000188398197223922	0.000175073	2.726108027	1.275040425	1.275040425	1	ALS
0.0144551346098748	0.035577015	-3.813817915	1.278004231	1.278004231	1	ALS

[0376]

[0377] [표 1d]

Probe sequence 60 mer	Probe Location	
	Chr	Start1
TCACCACACATCACCCCTTGCTCCTCGAGTCTTGGTGACCACAACAGGGTGCCACC	6	127483440
GAGGTGGGTGAATCATGAGGTCAAGGGTTCGACAATAGTTGAGAATCTCCAACCACCTGG	22	29442590
GGCCTTATAGTCAGCTGATCAGGTGAAATCGATTGGTCCCTTAGGATCAGCTACCAATTGC	2	112765788
GAGGCAGGCGGATCACAAAGTCAAAGATCGATAACTTCAATAATAGTTACAGATGCARA	13	98282669
AGCACCATATCTGGGATGTAGCTATTGCTCGAGATTGCAGTGAGCTGTGATCACACCTCT	17	66441278
TCTTCCCTCTTTTTAAACCACCAATTCATCGACCCACACATCCTGTGCCACTCTACTGC	17	6163839
TAACCATTAIGCATCACTAACATAGCATTTCGATATGATATGCTCAGTTTAGTTAGGGAAA	7	95420909
GGCTCAGGAAGAGAACTATTGTCTCTTTCGACACGCACATGCAGGACACTCACACGTAG	1	230738574
GTTGGGTGGATCCCTTGAGCTCAGGAATTCGAAGAATGATTTTTTCAGCCCGTGTGGAAGG	7	105657479
ATCAAAAGAAAATAGATACTTGTCTTACTCGAGTTGAATAAAATCCTCAGCTTCTGTCC	7	146728708
AAAAGAACTGTGAAAAGTTGTCACATTTTCGATTAAATCCAAAAGGTTCTTCTATGAGGC	10	66302476
TTAAAAGTATAGTAGTTGGCATTAACTTCGACCTTTTTCTGTTTCAGTAACCAACCCAG	8	105638873
CATCAACTAATAGTTAAACATTATAATATCGACTGAAGACCTTCATACTGTAAGATTCA	15	58334020
CATCAACTAATAGTTAAACATTATAATATCGAGTCTGCAGTGAGCTGAGATCACACTGCC	15	58334020
TTATTCCTTTCCAATAGTTAAAATATTTCGAAACTTTAAGAATCAATATAAAAATTC	7	79009348
CATAATATAAATTAATAATGACACTATCGATTATGTCAGTGTTCCTTGGTTGGTGT	8	123969419
CAGAGCACTAAGATAGACTTCTAAGGTTTCGAGGCATATAGCTCCAGCTGTATTGAGGTA	X	56538371

[0378]

[0379] [표 1e]

Probe Location			4 kb Sequence Location				
End1	Start2	End2	Chr	Start1	End1	Start2	End2
127483469	127600019	127600048	6	127479470	127483469	127600019	127604018
29442619	29482083	29482112	22	29442590	29446589	29482083	29486082
112765817	112810767	112810796	2	112765788	112769787	112810767	112814766
98282698	98346932	98346961	13	98278699	98282698	98346932	98350931
66441307	66481281	66481310	17	66441278	66445277	66477311	66481310
6163868	6195954	6195983	17	6159869	6163868	6195954	6199953
95420938	95465339	95465368	7	95416939	95420938	95465339	95469338
230738603	230752059	230752088	1	230738574	230742573	230752059	230756058
105657508	105741523	105741552	7	105653509	105657508	105741523	105745522
146728737	146792792	146792821	7	146728708	146732707	146788822	146792821
66302505	66512972	66513001	10	66298506	66302505	66509002	66513001
105638902	105814875	105814904	8	105634903	105638902	105814875	105818874
58334049	58485551	58485580	15	58330050	58334049	58485551	58489550
58334049	58540854	58540883	15	58330050	58334049	58536884	58540883
79009377	79284592	79284621	7	79009348	79013347	79280622	79284621
123969448	124085755	124085784	8	123965449	123969448	124085755	124089754
56538400	56570116	56570145	X	56534401	56538400	56570116	56574115

[0380]

[0381] [표 2a]

probe	PCR-Primer1_ID	PCR_Primer1
C6orf58_6_127480771_127483471_127600017_127604343_FR	OBD112-021	AAGCACTTCATTCTCCCCTCACC
NEFH_22_29442588_29445314_29482081_29484217_RR	OBD112-057	ACTGAGCAATGATGGCAACAAC
IL1A_2_112765786_112772711_112810765_112813086_RR	OBD112-077	CTAGGCCTGCGTTTCTCCGT
FARPI_13_98271575_98282700_98346930_98348486_FR	OBD112-093	AGTTCTCTCTCTAAGAACTCAAGGA
PRKCA_17_66441276_66447067_66475597_66481312_RF	OBD112-101	CCCCAGGCACTCACACCTTA
RNU6-1264P_17_6162286_6163870_6195952_6199184_FR	OBD112-105	GGGCACTCAACACCCCTTTTGT
PON2_7_95405100_95420940_95465337_95474032_FR	OBD112-125	GATGGGAATCAAGGCAAGGG
CAPN9_1_230738572_230739927_230752057_230757333_RR	OBD112-145	GCATTAGCCAGCAAGCATACCT
ATXN7L1_7_105654123_105657510_105741521_105750599_FR	OBD112-177	CACCGCCTGATGCAGGCTCT
CNTNAP2_7_146728706_146734820_146785878_146792823_RF	OBD112-217	GGCACTGTTGGTCTGAAGCAC
CTNNA3_10_66299269_66302507_66496211_66513003_FF	OBD112-229	GGTCTAGATGTCAGTCTTTCC
ZFPM2_8_105632010_105638904_105814873_105824107_FR	OBD112-245	GACTATAAATCTCTCCTGTGCAGC
ALDH1A2_15_58325151_58334051_58485549_58488054_FR	OBD112-317	CCTTACCCGACACCAGGTAGC
ALDH1A2_15_58325151_58334051_58538695_58540885_FF	OBD112-309	CCCGACACCAGGTAGCATTG
MAGI2_7_79009346_79018304_79275810_79284623_RF	OBD112-329	GGGACCCCACTAGACCAC
FER1L6_8_123963222_123969450_124085753_124093275_FR	OBD112-333	ATTCTCTCCCTGGTAAATCCTGGT
UBQLN2_X_56536168_56538402_56570114_56575112_FR	OBD112-337	CGCCAGCTCAGCAGCAATAA

[0382]

[0383] [표 2b]

PCR-Primer2_ID	PCR_Primer2	GLMNET
OBD112-023	ATACTCCCATCCCCTAGGCC	0.799327478
OBD112-059	CCCTACGACTGGCAAACCCA	-0.323299991
OBD112-079	CCCTGGCATTACATCACCGA	0.422829791
OBD112-095	GTCAAGCAACTGTGCTGGGG	0.471919696
OBD112-103	GGATCCACGATCTCCCTCCAC	-0.379552353
OBD112-107	AGGTCAGGATGGTACCGTTG	0.259971774
OBD112-127	CTGGGATGATTCTCTGGACTCT	-1.132792925
OBD112-147	GGTGGCCTGGGTTAGATGC	0.980123681
OBD112-179	CAGCTGGCCGATCCATCACC	-0.201292203
OBD112-219	GAGAACGACGACCTGGCACT	-0.613932219
OBD112-231	TCATCCTATCCTCTCCTAGC	-0.36233865
OBD112-247	ACTGTAGGCCAACCAGAAG	0.541529302
OBD112-319	CTGGTCCAGTGCAGCGTGT	-0.562433398
OBD112-311	GTGACTCTGCAGCACTGTT	0.486039187
OBD112-331	AGGCTCAGCAGGTTTCTGCC	-0.759217304
OBD112-335	CTGGGCAGGTCATCCAGACAG	-1.154344187
OBD112-339	TGCTAGGCCGAGTAATCATC	0.146973789

[0384]

[0385] [표 3]

Marker	GLMNET
OBD112-021/023	0.799327478
OBD112-057/059	-0.323299991
OBD112-077/079	0.422829791
OBD112-093/095	0.471919696
OBD112-101/103	-0.379552353
OBD112-105/107	0.259971774
OBD112-125/127	-1.132792925
OBD112-145/147	0.980123681
OBD112-177/179	-0.201292203
OBD112-217/219	-0.613932219
OBD112-229/231	-0.36233865
OBD112-245/247	0.541529302
OBD112-317/319	-0.562433398
OBD112-309/311	0.486039187
OBD112-329/331	-0.759217304
OBD112-333/335	-1.154344187
OBD112-337/339	0.146973789

[0386]

[0387] [표 4a]

Genes	Full Name
C6orf58	Chromosome 6 Open Reading Frame 58
NEFH	Neurofilament Heavy
IL1A	Interleukin 1 Alpha
FARP1	FERM, ARH/RhoGEF And Pleckstrin Domain Protein 1
PRKCA	Protein Kinase C Alpha
RNU6-1264P	RNA, U6 Small Nuclear 1264, Pseudogene
PON2	Paraoxonase 2
CAPN9	Calpain 9
ATXN7L1	Ataxin 7 Like 1
CNTNAP2	Contactin Associated Protein-Like 2
CTNNA3	Catenin Alpha 3
ZFPM2	Zinc Finger Protein, FOG Family Member 2
ALDH1A2	Aldehyde Dehydrogenase 1 Family Member A2
MAGI2	Membrane Associated Guanylate Kinase, WW And PDZ Domain Containing 2
FER1L6	Fer-1 Like Family Member 6
UBQLN2	Ubiquilin 2

[0388]

[0389] [표 4b]

Activity
Protein Coding gene
maintenance of neuronal caliber, intracellular transport to axons and dendrites, neuronal damage and susceptibility to amyotrophic lateral sclerosis (ALS)
various immune responses, inflammatory processes, hematopoiesis and cell injury, activates apoptosis, correlated with rheumatoid arthritis and Alzheimer's disease
dendritic growth in neurons, link between cytoskeleton to the cell membrane
diverse cellular signalling pathways, receptors for phorbol esters, cellular processes such as cell adhesion, cell transformation, cell cycle checkpoint, and cell volume control
Pseudogene, connected with the snRNA class
cellular antioxidant, human tissues and membrane-bound, defence responses to pathogenic bacteria, protecting cells from oxidative stress
neurodegenerative processes and digestive tract, associated with gastric cancer
Protein Coding gene
regulates interactions between neurons and glia, associated with speech and language development, cell adhesion in nervous system
cell-cell adhesion in muscle cells, related to arrhythmogenic right ventricular dysplasia
regulates hematopoiesis and cardiogenesis in mammals, developing and adult tissues, posterior organ development, prevents spina bifida
dentatorubral and pallidolysian atrophy
Protein Coding gene
in vivo protein degradation

[0390]

[0391] [표 5a] Array_data_for_3_month_Samples

probe	GeneLocus	Probe_Count_Total	Probe_Count_Sig
ZFPM2_8_104964347_104973135_105209447_105218050_FF	ZFPM2	187	6
VEGFA_6_43733863_43737741_43777557_43780533_FR	VEGFA	42	1
ERBB4_2_212287088_212294815_212317329_212325591_FR	ERBB4	193	10
ZFPM2_8_105209447_105218050_105302264_105310138_FF	ZFPM2	187	6
PASD1_X_151600201_151608969_151687203_151692271_FR	PASD1	41	14
GRM7_3_7259138_7267165_7394377_7401227_RF	GRM7	174	7
GRIK2_6_102047545_102055341_102078831_102089392_FR	GRIK2	188	5
SORCS2_4_7219417_7226457_7241158_7248673_FF	SORCS2	84	6
LINGO2_9_28314155_28333777_28371887_28374860_FR	LINGO2	194	3

[0392]

[0393] [표 5b]

HyperG_Stats	FDR_HyperG	Percent_Sig	logFC	AveExpr	t	P.Value	adj.P.Val
0.067854625	0.99237389	3.21	0.272052749	7.772607282	3.665534368	0.001024476	0.999953189
0.475978534	1	2.38	-0.142175177	14.62826333	-2.054497284	0.049386319	0.999953189
0.000818098	0.03828697	5.18	0.383778852	7.597586859	3.176083618	0.003622402	0.999953189
0.067854625	0.99237389	3.21	0.229570464	9.577318093	2.801385891	0.00913346	0.999953189
7.57E-16	3.54E-13	34.15	0.341166309	9.148011765	2.62030276	0.01404442	0.999953189
0.01784403	0.37959119	4.02	0.306594489	9.301524803	2.334933566	0.026957376	0.999953189
0.16129386	1	2.66	0.286167252	9.345860096	2.824849044	0.008630812	0.999953189
0.001827031	0.0712542	7.14	-0.222042969	10.03535881	-2.171963051	0.038499943	0.999953189
0.570146657	1	1.55	0.40947654	6.9839738	2.811551759	0.008912387	0.999953189

[0394]

[0395] [표 5c]

B	FC	FC_1	LS	Loop detected
-1.464031868	1.207524741	1.207524741	1	Fs
-3.802493578	0.906151904	-1.103567731	-1	Sw
-2.223740309	1.304754929	1.304754929	1	Fs
-2.785028975	1.172485811	1.172485811	1	Fs
-3.046315984	1.266780276	1.266780276	1	Fs
-3.440610909	1.236784796	1.236784796	1	Fs
-2.750635647	1.219396451	1.219396451	1	Fs
-3.654353119	0.857350501	-1.166384109	-1	Sw
-2.770141569	1.328203808	1.328203808	1	Fs

[0396]

[0397] [표 5d]

Probe sequence	Probe Location		
	Chr	Start1	End1
TATATTTAAAAATACATACTGGTATACATCGATCTCATGACTTTGCTATTATGCATAGTG	8	104973104	104973133
CCCCAGCCAGCAACCTGGCTCACCTGATCGAGTACATCTTCAAGCCATCCTGTGTGCC	6	43737710	43737739
TATAAATAATACAGCTCTATTTGCCTACTCGATTAAGAATCATATTATCCTTAATTC	2	212294784	212294813
CACTATGCATAATAGCAAAGTCATGAGATCGAAAATGTTTGTCAAGCAGTAGGTTTTGGG	8	105218019	105218048
GATTTTAGAATCTTAAACAAGGCTGCAATCGAGTTAGCTGCTGCAGAAAGAGAGAAAA	X	151608938	151608967
GGTGACTAAAAATGAGATTGCATTTTCTTCGACATTTGGCCAGCATGCCAAACTGGT	3	7259140	7259169
TAACCTCTCTTTCTTAGGTTCTCCATATCGATAGAAAATTGTCTGCAGCCCTTAATGCC	6	102055310	102055339
AGCTCACGGTCAAGTCCGTTCCGTTTGTCTCGAATAAAGAACAAGGACCTTAAAAATAGA	4	7226426	7226455
AAAGTCATACAACCTACTATGTAAGATATTCGAATACCTGTTAGAATAGGTGAAGTTTAT	9	28333746	28333775

[0398]

[0399] [표 5e]

Probe Location		4 kb Sequence Location				
Start2	End2	Chr	Start1	End1	Start2	End2
105218019	105218048	8	104969134	104973133	105214049	105218048
43777559	43777588	6	43733740	43737739	43777559	43781558
212317331	212317360	2	212290814	212294813	212317331	212321330
105310107	105310136	8	105214049	105218048	105306137	105310136
151687205	151687234	X	151604968	151608967	151687205	151691204
7401196	7401225	3	7259140	7263139	7397226	7401225
102078833	102078862	6	102051340	102055339	102078833	102082832
7248642	7248671	4	7222456	7226455	7244672	7248671
28371889	28371918	9	28329776	28333775	28371889	28375888

[0400]

[0401] [표 6a] Array_data_for_6_month_Samples

probe	GeneLocus	Probe_Count_Total	Probe_Count_Sig	HyperG_Stats
ZFPM2_8_104964347_104973135_105209447_105218050_FF	ZFPM2	187	8	0.007868339
VEGFA_6_43733863_43737741_43777557_43780533_FR	VEGFA	42	1	0.472787206
ERBB4_2_212287088_212294815_212317329_212325591_FR	ERBB4	193	13	8.61E-06
ZFPM2_8_105209447_105218050_105302264_105310138_FF	ZFPM2	187	8	0.007868339
PASD1_X_151600201_151608969_151687203_151692271_FR	PASD1	41	18	1.87E-22
GRM7_3_7259138_7267165_7394377_7401227_RF	GRM7	174	6	0.049498361
GRIK2_6_102047545_102055341_102078831_102089392_FR	GRIK2	188	6	0.066833261
SORCS2_4_7219417_7226457_7241158_7248673_FF	SORCS2	84	3	0.134527864
LINGO2_9_28314155_28333777_28371887_28374860_FR	LINGO2	194	4	0.337532174

[0402]

[0403] [표 6b]

FDR_HyperG	Percent_Sig	logFC	AveExpr	t	P.Value	adj.P.Val
0.204139683	4.28	0.214103447	7.772607282	2.884747708	0.007463232	0.999935724
1	2.38	-0.142760759	14.62826333	-2.062959216	0.04851942	0.999935724
0.001005104	6.74	0.365781024	7.597586859	3.027136877	0.005258488	0.999935724
0.204139683	4.28	0.23990173	9.577318093	2.927455517	0.006723676	0.999935724
8.73E-20	43.9	0.331193838	9.148011765	2.543709927	0.016785386	0.999935724
0.825561956	3.45	0.354666043	9.301524803	2.701032396	0.011610123	0.999935724
1	3.19	0.294782642	9.345860096	2.909894328	0.007018965	0.999935724
1	3.57	-0.229277936	10.03535881	-2.242733507	0.033029594	0.999935724
1	2.06	0.440106238	6.9839738	3.021861688	0.005327738	0.999935724

[0404]

[0405] [표 6c]

B	FC	FC_1	LS	Loop detected
-2.662323695	1.159982827	1.159982827	1	Fs
-3.791998096	0.905784176	-1.104015754	-1	Sw
-2.449689974	1.288579031	1.288579031	1	Fs
-2.598930246	1.18091222	1.18091222	1	Fs
-3.154405805	1.25805399	1.25805399	1	Fs
-2.930776901	1.278689554	1.278689554	1	Fs
-2.625038885	1.226700147	1.226700147	1	Fs
-3.562656794	0.853061739	-1.172248096	-1	Sw
-2.457630227	1.35670423	1.35670423	1	Fs

[0406]

[0407] [표 6d]

Probe sequence 60 mer	Probe Location		
	Chr	Start1	End1
TATATTTAAAAATACATACTGGTATACATCGATCTCATGACITTTGCTATTATGCATAGTG	8	104973104	104973133
CCCCAGCCAGCAACCTGGCTCACCTGATCGAGTACATCTTCAAGCCATCCTGTGTGCC	6	43737710	43737739
TATAAATAATACAGCTCTATTTGCCTACTCGATTAAGAATCATATTATATCCCTAATTC	2	212294784	212294813
CACTATGCATAATAGCAAAGTCATGAGATCGAAAATGTTTGTCAAGCAGTAGGTTTGGG	8	105218019	105218048
GATTTTGAATCTCTAACCAAGGCTGCAATCGAGGTTAGCTGCTGCAGAAAAGAGAAAA	X	151608938	151608967
GGTGACTAAAATGAGATTGCATTTCTTTTCGACCATTTGGCCAGCATGCCAAACACTGGT	3	7259140	7259169
TAACTCTCCTTTCTTAGGTTCTCCATATCGATAGAAAATTGCTGCAGCCCTAATGCC	6	102055310	102055339
AGCTCAGGTCAGTGCCGTTCCGTTTCTCGAATAAAGAACAAGGACCTTAAAAAATAGA	4	7226426	7226455
AAAGTCATACTACTATGTAAGATTCGAATACCTGTTAGAATAGGTGAAGGTTTAT	9	28333746	28333775

[0408]

[0409] [표 6e]

Probe Location		4 kb Sequence Location				
Start2	End2	Chr	Start1	End1	Start2	End2
105218019	105218048	8	104969134	104973133	105214049	105218048
43777559	43777588	6	43733740	43737739	43777559	43781558
212317331	212317360	2	212290814	212294813	212317331	212321330
105310107	105310136	8	105214049	105218048	105306137	105310136
151687205	151687234	X	151604968	151608967	151687205	151691204
7401196	7401225	3	7259140	7263139	7397226	7401225
102078833	102078862	6	102051340	102055339	102078833	102082832
7248642	7248671	4	7222456	7226455	7244672	7248671
28371889	28371918	9	28329776	28333775	28371889	28375888

[0410]

[0411] [표 7a]

probe	Inner primers	
	PCR-Primer1_ID	PCR_Primer1
ZFPM2_8_104964347_104973135_105209447_105218050_FF	OBD112_773	GCAGCACACAGGGAACTCTCTT
VEGFA_6_43733863_43737741_43777557_43780533_FR	OBD112_765	GCTAGGAAGGCGCTGGGATG
ERBB4_2_212287088_212294815_212317329_212325591_FR	OBD112_477	TCCAAAATGTTTAATACTGCCTAGA
ZFPM2_8_105209447_105218050_105302264_105310138_FF	OBD112_777	CAGCCACTGTAGAGAGCAGT
PASD1_X_151600201_151608969_151687203_151692271_FR	OBD112_637	ACTTCTCCCAAGTCACTTTTTC
GRM7_3_7259138_7267165_7394377_7401227_RF	OBD112_517	TGACGAAGAAGCAATCCCTGGT
GRIK2_6_102047545_102055341_102078831_102089392_FR	OBD112_501	AATCTCTGCCCTCCTCATCTTG
SORCS2_4_7219417_7226457_7241158_7248673_FF	OBD112_725	AGGTATGCAGCCAGCCTGAG
LINGO2_9_28314155_28333777_28371887_28374860_FR	OBD112_561	CCGTGCCATATCCTCTGATTTATGC

[0412]

[0413] [표 7b]

Inner primers		
PCR-Primer2_ID	PCR_Primer2	GLMNET
OBD112_775	TTGTTGAGCCAGCAATTCCTTT	-0.18669073
OBD112_767	CTCAGTGGGCACACTCCA	0.15278372
OBD112_479	AGCCATGTGGTCTGGAATCT	-0.24015518
OBD112_779	TAACCCACAGCAGCAAGGT	0.22911033
OBD112_647	TGGCCATCTTGCTTTGCCTC	0.52468015
OBD112_519	GGACCTACCTCCACTGGGTTG	0.10503793
OBD112_503	CATGTTCCACAGCAAGGAAGTTA	-0.13060979
OBD112_727	TTTCCGTGCCAGTGTCTCTGT	-0.1217346
OBD112_563	GGCTGACCTTCAACAGATTCCGC	0.01500745

[0414]

[0415] [표 8]

Gene	Marker	GLMNET
ERBB4_2	OBD112_477.479_0.4ng	-0.24015518
ZFPM2_8	OBD112_773.775_6.5ng	-0.18669073
GRIK2_6	OBD112_501.503_0.8ng	-0.13060979
SORCS2_4	OBD112_725.727_3.25ng	-0.1217346
LINGO2_9	OBD112_561.563_13ng	0.01500745
GRM7_3	OBD112_517.519_13ng	0.10503793
VEGFA_6	OBD112_765.767_13ng	0.15278372
ZFPM2_8	OBD112_777.779_6.5ng	0.22911033
PASD1_X	OBD112_637.647_3.25ng	0.52468015

[0416]

[0417] [표 9]

Genes	Full name	Activity
ZFPM2	Zinc Finger Protein, FOG Family Member 2	transcription factor, regulates the activity of GATA family protein, regulates GATA-target genes expression
VEGFA	Vascular Endothelial Growth Factor A	stimulates proliferation and migration of vascular endothelial cells, stimulates physiological and pathological angiogenesis, modulates tumor stage and progression
ERBB4	Erb-B2 Receptor Tyrosine Kinase 4	stimulates a variety of cellular responses including mitogenesis and differentiation,
PASD1	PAS Domain Containing Repressor 1	encodes a cancer-associated antigen, induces autologous T-cell responses, immunotherapeutic target for the treatment of various hematopoietic malignancies
GRM7	Glutamate Metabotropic Receptor 7	activates ionotropic and metabotropic glutamate receptors, involved in most aspects of normal brain function, perturbed in neuropathologic conditions
GRIK2	Glutamate Ionotropic Receptor Kainate Type Subunit 2	modulates neurophysiologic processes
SORCS2	Sortilin Related VPS10 Domain Containing Receptor 2	protein coding gene, over expression in the central nervous system
LINGO2	Leucine Rich Repeat And Ig Domain Containing 2	protein coding gene

[0418]

[0419] [표 10a]

N	Array_Probe	Primer1_ID
1	DCLK1_13_35749195_35757694_36000600_36005542_RR	OBD112_453
2	GRM7_3_7259138_7267165_7394377_7401227_RF	OBD112_517
3	LINGO2_9_28314155_28333777_28371887_28374860_FR	OBD112_561
4	PASD1_X_151600201_151608969_151687203_151692271_FR	OBD112_637
5	UBQLN2_X_56536168_56538402_56570114_56575112_RR	OBD112_757
6	ZFPM2_8_104964347_104973135_105209447_105218050_FF	OBD112_773
7	ZFPM2_8_105209447_105218050_105302264_105310138_FF	OBD112_777
8	ZNF804B_7_89108281_89114839_89127203_89146680_RF	OBD112_797

[0420]

[0421] [표 10b]

Primer1_Seq	Primer2_ID	Primer2_Seq
TCTTGTACACGGTTGGTGGT	OBD112_455	TGTCACCTATGTGCTGAGTACTGG
TGACGAAGAAGCAATCCCTGGT	OBD112_519	GGACCTACCTCCACTGGGTTG
CCGTGCCATATCCTCTGATTTATGC	OBD112_563	GGCTGACCTTCAACAGATTCGC
ACTTCTTCCCAAGTCACTTTTTGC	OBD112_647	TGGCCATCTTGCTTTGCCTC
GGATATGCAGTTTTCTGGCACTAC	OBD112_759	CATGCTAGGGCCGAGTAATCATCT
GCAGCACACAGGGAACTCTCTT	OBD112_775	TTGTTGAGCCCAGCAATTCCTTT
CAGCCACTGTAGAGAGCAGT	OBD112_779	TAACCCACCAGCAGCAAGGT
AGTAGCTTCCCTGTTAGAGGCTTG	OBD112_799	AGCCAGTGACTCCACAACCTTCTT

[0422]

[0423] [표 10c]

Hydrolysis_probe_ID	Hydrolysis_probe_Seq
OBD112_88_FAM	AGGTTGCTTTCGAAGTACAGATATCACT
OBD112_57_FAM	ATTTTCTTTTCGACCATTTGGCCAGCA
OBD112_90_FAM	TTCTAACAGGTATTCGAATATCTTACATAGT
OBD112_26_FAM	AAGTAACATTCGACAAACCTCTGGTAA
OBD112_95_FAM	TATTCTCTTTGATCGAGGCATATAGCT
OBD112_80_FAM	AAAGTCATGAGATCGATGTATACCAGTAT
OBD112_32_FAM	TTGACAAACATTTTCGATCTCATGACTT
OBD112_33_FAM	TGTGGCTCAGTCGACAGAAAGTACA

[0424]

[0425] [표 11a(i)]

probe	GeneLocus	Probe_Count_Total	Probe_Count_Sig
DCLK1_13_35749195_35757694_36000600_36005542_RR	DCLK1	182	3
UBQLN2_X_56536168_56538402_56570114_56575112_RR	UBQLN2	10	7
ZNF804B_7_89108281_89114839_89127203_89146680_RF	ZNF804B	191	10

[0426]

[0427] [표 11a(ii)]

HyperG_Stats	FDR_HyperG	Percent_Sig	logFC	AveExpr	t
0.254316537	1	1.65	0.199442299	7.402197005	2.87151
0.000000000001	0.000000000386	70	0.283366605	9.419823948	2.114955
1.74E-05	0.002050325	5.24	0.228756351	9.389157739	2.52348

[0428]

[0429] [표 11a(iii)]

P.Value	adj.P.Val	B	FC	FC_1	LS	Loop detected
0.007707652	0.999953189	-2.681902147	1.14825439	1.14825439	0	Faster
0.043482517	0.999953189	-3.726899057	1.217031581	1.217031581	0	Faster
0.017587875	0.999953189	-3.182692934	1.171824364	1.171824364	0	Faster

[0430]

[0431] [표 11b(i)]

probe	PCR-Primer1_ID	PCR_Primer1
DCLK1_13_35749195_35757694_36000600_36005542_RR	OBD112_453	TCTTGACACGGTTGGTGGT
UBQLN2_X_56536168_56538402_56570114_56575112_RR	OBD112_757	GGATATGCAGTTTTCTGGCACTAC
ZNF804B_7_89108281_89114839_89127203_89146680_RF	OBD112_797	AGTAGCTTCCCTGTTAGAGGTCCTTG

[0432]

[0433] [표 11b(ii)]

PCR-Primer2_ID	PCR_Primer2	GLMNET
OBD112_455	TGTACCTATGTGCTGAGTACTGG	0.047800994
OBD112_759	CATGCTAGGGCCGAGTAATCATCT	-0.013191036
OBD112_799	AGCCAGTGACTCCACAACCTTCTT	-0.025426843

[0434]

[0435] [표 11c]

Marker	GLMNET
OBD112_453.455_3.25ng	0.047801
OBD112_757.759_13ng	-0.013191
OBD112_797.799_1.6ng	-0.025427

[0436]

[0437] [표 12a]

probe	GeneLocus	Loop detected	Probe sequence	
			60 mer	
HTT_4_3088056_3090591_3182426_3183765_FR	HTT	HD_Negative	AGATCTAGTTCACAGTAGCACCATAATATCGACAGATAGCTGACATCATCTCCCAATGT	
HTT_4_3076247_3078445_3182426_3183765_RR	HTT	HD_Negative	AGAGTACTTCTAACTCCTACTGTACACTCGACAGATAGCTGACATCATCTCCCAATGT	
HTT_4_3062575_3066389_3251623_3253432_RF	HTT	HD_Positive	TATAACCAGTGTCTCTACGAAGGCCGTTTGAAGTCTCAAACCTCACTTCTCTGTGCGC	
HTT_4_3076247_3078445_3253432_3258164_RR	HTT	HD_Positive	AGAGTACTTCTAACTCCTACTGTACACTCGAGGATGATCGCTCCGACAGCTCTCCAGC	
HTT_4_3070353_3072847_3212506_3214871_FR	HTT	HD_Negative	GGGTTTCGCCATGTTGGCCAGGCTGGTCCGAAGTGTATGCATCTGTGCTCAGCTTTGCA	
HTT_4_3088056_3090591_3212506_3214871_FF	HTT	HD_Positive	AGATCTAGTTCACAGTAGCACCATAATATCGACTGTCTCTGTTGGCCACTCTCACCT	
HTT_4_3088056_3090591_3167292_3170536_FR	HTT	HD_Negative	AGATCTAGTTCACAGTAGCACCATAATATCGAACTCTGACCTTGTGATCCACCCACCT	

[0438]

[0439] [표 12b]

Probe Location					probe	PCR-Primer1_ID	PCR_Primer1
Chr	Start1	End1	Start2	End2			
4	3090560	3090591	3182426	3182457	HTT_4_3088056_3090591_3182426_3183765_FR	RD031_037	GAATACCCAGGAATGCTTACTTGAGC
4	3076247	3076278	3182426	3182457	HTT_4_3076247_3078445_3182426_3183765_RR	RD031_061	ACCAAATGCCATCTGGGGACACATCCA
4	3062575	3062606	3253401	3253432	HTT_4_3062575_3066389_3251623_3253432_RF	RD031_185	CCCGCTAAGTCCACCCCTCTGTA
4	3076247	3076278	3253432	3253463	HTT_4_3076247_3078445_3253432_3258164_RR	RD031_237	ACCAAATGCCATCTGGGGACACATCCA
4	3072816	3072847	3212506	3212537	HTT_4_3070353_3072847_3212506_3214871_FR	RD031_241	AGCTCGCCACTGAGGTAACAAACA
4	3090560	3090591	3214840	3214871	HTT_4_3088056_3090591_3212506_3214871_FF	RD031_329	GAATACCCAGGAATGCTTACTTGAGC
4	3090560	3090591	3167292	3167323	HTT_4_3088056_3090591_3167292_3170536_FR	RD031_333	GAATACCCAGGAATGCTTACTTGAGC

[0440]

[0441] [표 12c]

PCR-Primer2_ID	PCR_Primer2
RD031_039	GATTCAGCCACCCACCTTTCACAAG
RD031_063	TGATTCAGCCACCCACCTTTCACAA
RD031_187	GAACCGCACTTACCCTCAGCAGT
RD031_239	GGACAAGCAGACACACTACCTGAACT
RD031_243	ATCCCTGAAACAGAAGGACCTCGTG
RD031_331	GAGCCGTCTCATAATAACCTCAGGGT
RD031_335	CAGTGGTTTAGGGCAAAGAGAGGGAG

[0442]

[0443] [표 13a]

probe	GeneLocus	Loop detected	Probe sequence	
			60 mer	
HTT_4_3088056_3090591_3112545_3114867_FR	HTT	No difference	AGATCTAGTTCACAGTAGCACCATAATATCGACATTAAGTTCATGAGAGCTTCTATATT	
HTT_4_3062575_3066389_3119060_3125597_RR	HTT	No difference	GGCCACAGGAGAAGTGAAGTTTGAGACTCGAAAAAGGAATAAATTGAAAAATAGAGGAA	
HTT_4_3076247_3078445_3241592_3247577_RR	HTT	No difference	AGAGTACTTCTAACTCCTACTGTACACTCGATGTTTTGGGTATTGAATGTGGTAAGTGG	
HTT_4_3080953_3084487_3253432_3258164_RR	HTT	No difference	ATCAAGACTGTATGGTACTGGCACAGGATCGAGGATGATCGCTCCGACAGCTCTCCAGC	
HTT_4_3080953_3084487_3241592_3247577_RR	HTT	No difference	ATCAAGACTGTATGGTACTGGCACAGGATCGATGTTTTGGGTATTGAATGTGGTAAGTGG	
HTT_4_3080953_3084487_3112545_3114867_RR	HTT	No difference	ATCAAGACTGTATGGTACTGGCACAGGATCGACATTAAGTTCATGAGAGCTTCTATATT	
HTT_4_3076247_3078445_3112545_3114867_RF	HTT	No difference	TGTCAGTTAAATTGCAGATATTTCCGATCGAGTACAGTAGGAGTATAGGAAGTACTCT	
HTT_4_3080953_3084487_3167292_3170536_RR	HTT	No difference	ATCAAGACTGTATGGTACTGGCACAGGATCGAACTCTGACCTGTGATCCACCCACCTC	
HTT_4_3038381_3040238_3088056_3090591_RF	HTT	No difference	AGATCTAGTTCACAGTAGCACCATAATATCGAATCTGGGTGTTTCAAAGACATCAACAAA	
HTT_4_3088056_3090591_3253432_3258164_FR	HTT	No difference	AGATCTAGTTCACAGTAGCACCATAATATCGAGGATGATCGCTCCGACAGCTCTCCAGC	
HTT_4_3062575_3066389_3167292_3170536_FR	HTT	No difference	GGGTTTCACCATGTTGGCCAGGATGGTCTGAACTCTGACCTGTGATCCACCCACCTC	
HTT_4_3088056_3090591_3251623_3253432_FF	HTT	No difference	AGATCTAGTTCACAGTAGCACCATAATATCGAAGCGCCTTCTGATGAGGACATGTTATA	
HTT_4_3088056_3090591_3241592_3247577_FF	HTT	No difference	AGATCTAGTTCACAGTAGCACCATAATATCGACTAAGGTGACCCCTCGTCTGGGGCT	

[0444]

[0445] [표 13b]

Probe Location					probe	PCR-Primer1_ID	PCR_Primer1
Chr	Start1	End1	Start2	End2			
4	3090560	3090591	3112545	3112576	HTT_4_3088056_3090591_3112545_3114867_FR	RD031_001	GAATACCCAGGAATGCTTACTTGAGC
4	3062575	3062606	3119060	3119091	HTT_4_3062575_3066389_3119060_3125597_RR	RD031_009	GGTCCAGAGAACCACCTTACCC
4	3076247	3076278	3241592	3241623	HTT_4_3076247_3078445_3241592_3247577_RR	RD031_025	ACCAAATGCCATCTGGGGACACATCCA
4	3080953	3080984	3253432	3253463	HTT_4_3080953_3084487_3253432_3258164_RR	RD031_041	TCCCTCTCAGAATCTTACTTGGCTTC
4	3080953	3080984	3241592	3241623	HTT_4_3080953_3084487_3241592_3247577_RR	RD031_057	TCCCTCTCAGAATCTTACTTGGCTTC
4	3080953	3080984	3112545	3112576	HTT_4_3080953_3084487_3112545_3114867_RR	RD031_065	TCCCTCTCAGAATCTTACTTGGCTTC
4	3076247	3076278	3114836	3114867	HTT_4_3076247_3078445_3112545_3114867_RF	RD031_097	GAACAGTTTTGGGTAGTATGCGGCT
4	3080953	3080984	3167292	3167323	HTT_4_3080953_3084487_3167292_3170536_RR	RD031_149	TCCCTCTCAGAATCTTACTTGGCTTC
4	3038381	3038412	3090560	3090591	HTT_4_3038381_3040238_3088056_3090591_RF	RD031_157	GAATACCCAGGAATGCTTACTTGAGC
4	3090560	3090591	3253432	3253463	HTT_4_3088056_3090591_3253432_3258164_FR	RD031_205	GAATACCCAGGAATGCTTACTTGAGC
4	3066389	3066389	3167292	3167323	HTT_4_3062575_3066389_3167292_3170536_FR	RD031_261	TCTGGACAAGGTATCTGGAAGGCTGA
4	3090560	3090591	3253401	3253432	HTT_4_3088056_3090591_3251623_3253432_FF	RD031_301	GAATACCCAGGAATGCTTACTTGAGC
4	3090560	3090591	3247546	3247577	HTT_4_3088056_3090591_3241592_3247577_FF	RD031_337	GAATACCCAGGAATGCTTACTTGAGC

[0446]

[0447] [표 13c]

PCR-Primer2_ID	PCR_Primer2
RD031_003	GCACTGCTCGCAATAGCCAAGA AACTA
RD031_011	CTTGCTGACCCTCCTATG GATGG
RD031_027	GCACAGAGTTCCAACATTTCTCCAC
RD031_043	GGACAAGCAGACACACTACCTGAACT
RD031_059	GCACAGAGTTCCAACATTTCTCCAC
RD031_067	GCACTGCTCGCAATAGCCAAGA AACTA
RD031_099	ACCAAATGCCATCTGGGACACATCCA
RD031_151	ACAGTGGTTTAGGGCAAAGAGAGGGA
RD031_159	GAAGAAGCCTCTGTTTGTCTGGAAA
RD031_207	GGAAGGACAAGCAGACACACTACCTG
RD031_263	GTGGTTTAGGGCAAAGAGAGG GAGAT
RD031_303	TCCCCGAGTCCTGTGATG GCAA AACT
RD031_339	ATGGCTACGGAAAGGGCATT CGGAC

[0448]

[0449] [표 14]

Subject ID (Unique ID for each subject)	Disease or Control	Type of Biopsiedmen (e.g., FFPE, c. serum, blood, DNA etc.)	Age (At time of collection)	Gender (M or F)	Date of collection (YBD)	Ethnicity	Age at Diagnosis	Method of diagnosis	Genetic Test Results for CAG repeats	Recorded Symptoms	Medication at time of collection	Any other genetic test results if available	Test results for HIV	Test results for HBV	Test results for HCV	Test results for syphilis
control 01	Control	Whole blood 9 mL	32	F	2017-11-07	Caucasian (White)	-	PCR	n1=17 n2=28	no	no	no	neg	neg	neg	neg
control 02	Control	Whole blood 9 mL	65	F	2017-11-10	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
control 03	Control	Whole blood 9 mL	41	M	2017-11-13	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
control 04	Control	Whole blood 9 mL	19	M	2017-11-10	Caucasian (White)	-	PCR	n1=17 n2=21	no	no	no	neg	neg	neg	neg
control 05	Control	Whole blood 9 mL	46	M	2017-11-10	Caucasian (White)	-	PCR	n1=15 n2=17	no	no	no	neg	neg	neg	neg
control 06	Control	Whole blood 9 mL	29	F	2017-11-13	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
control 07	Control	Whole blood 9 mL	47	F	2017-11-14	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
control 08	Control	Whole blood 9 mL	22	M	2017-11-14	Caucasian (White)	-	PCR	n1=19 n2=22	no	no	no	neg	neg	neg	neg
control 09	Control	Whole blood 9 mL	35	F	2017-11-15	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
control 11	Control	Whole blood 9 mL	33	M	2017-11-16	Caucasian (White)	-	PCR	n1=35 n2=35	no	no	no	neg	neg	neg	neg
HD 01	Huntington's	Whole blood 9 mL	32	M	2017-10-30	Caucasian (White)	27	PCR	n1=35 n2=48 irritability, chorea	irritability, chorea	no	no	neg	neg	neg	neg
HD 02	Huntington's	Whole blood 9 mL	32	F	2017-10-30	Caucasian (White)	29	PCR	n1=35 n2=42	no	no	no	neg	neg	neg	neg
HD 03	Huntington's	Whole blood 9 mL	28	F	2017-11-01	Caucasian (White)	26	PCR	n1=20 n2=46 irritability	irritability	no	no	neg	neg	neg	neg
HD 12	Huntington's	Whole blood 9 mL	50	M	2017-11-02	Caucasian (White)	46	PCR	n1=20 n2=48 irritability, chorea	irritability, chorea	no	no	neg	neg	neg	neg
HD 05	Huntington's	Whole blood 9 mL	52	M	2017-11-02	Caucasian (White)	49	PCR	n1=35 n2=45 irritability, chorea	irritability, chorea	no	no	neg	neg	neg	neg
HD 06	Huntington's	Whole blood 9 mL	43	M	2017-11-02	Caucasian (White)	35	PCR	n1=22 n2=40 irritability, chorea	irritability, chorea	Tetrabenazine	no	neg	neg	neg	neg
HD 08	Huntington's	Whole blood 9 mL	23	F	2017-11-02	Caucasian (White)	-	PCR	n1=35 n2=44	no	no	no	neg	neg	neg	neg
HD 09	Huntington's	Whole blood 9 mL	21	F	2017-11-02	Caucasian (White)	-	PCR	n1=22 n2=41	no	no	no	neg	neg	neg	neg
HD 10	Huntington's	Whole blood 9 mL	38	F	2017-11-02	Caucasian (White)	36	PCR	n1=35 n2=45 irritability	irritability	no	no	neg	neg	neg	neg
HD 11	Huntington's	Whole blood 9 mL	34	M	2017-11-15	Caucasian (White)	32	PCR	n1=24 n2=43 irritability, chorea	irritability, chorea	Sertraline	no	neg	neg	neg	neg

[0450]

[0451] [표 15]

Zone	hg19 coordinates	hg38 coordinates
Anchor point CAG repeat site	chr4:3055890-3097657	chr4:3054163-3095930
Zone 1	chr4:3232680-3259897	chr4:3230953-3258170
Zone 2	chr4:3107637-3130235	chr4:3105908-3128508
Zone 3	chr4:3212999-3218545	chr4:3211272-3216818
Zone 4	chr4:3161885-3185826	chr4:3160158-3184099
Zone 5	chr4:3033588-3050187	chr4:3031861-3048460

[0452]

[0453] [표 16]

	Healthy Controls (N=10)	Huntington's Disease (N=10)
Gender		
Male (N, (%))	5 (50)	5 (50)
Female (N, (%))	5 (50)	5 (50)
Ethnicity		
Non-Hispanic or Latino	10 (100)	10 (100)
Race		
White (N, (%))	10 (100)	10 (100)
Huntington's Type		
Symptomatic (N, (%))	N/A	7 (70)
Asymptomatic (N, (%))	N/A	3 (30)
CAG repeat length (Average, (SD))	25.7 (5.4)	44.2 (2.6)
Age at Diagnosis (Average, (SD))	N/A	34.9* (8.1)
Age at Sample Collection (Average, (SD))	36.9 (12.8)	35.3 (10.0)
Disease Duration** (Years) (Average, (SD))	N/A	3.8 (1.9)
% Reporting Irritability	N/A	70
% Reporting Chorea	N/A	50
*Age at Diagnosis was not available for 2 of the 10 HD patients		
N/A = Not Applicable		

[0454]

[0455] [표 17]

Interaction ID	Zone5	Anchor	Zone2	Zone4	Zone3	Zone1
RD031_025/27		✓				✓
RD031_041/43		✓				✓
RD031_057/59		✓				✓
RD031_061/63		✓		✓		
RD031_065/67		✓	✓			
RD031_097/99		✓	✓			
RD031_009/11		✓	✓			
RD031_149/151		✓		✓		
RD031_237/239		✓				✓
RD031_261/263		✓		✓		
RD031_157/159	✓	✓				
RD031_001/3		✓	✓			
RD031_329/331		✓			✓	
RD031_301/303		✓				✓
RD031_337/339		✓				✓
RD031_037/39		✓		✓		
RD031_185/187		✓				✓
RD031_205/207		✓				✓
RD031_241/243		✓			✓	
RD031_333/335		✓		✓		

[0456]

[0457] [표 18a]

Marker	HD01	HD03	HD05	HD06	HD10	HD11	HD12	Detection rates
RD031_237.239	0	0	0	1	1	1	1	57.14
RD031_329.331	0	1	1	1	1	1	1	85.71
RD031_185.187	0	0	1	1	0	1	0	42.86

[0458]

[0459] [표 18b]

237.239		Symptoms	No Symptoms		
	Loop Present	4	1	risk ratio	0.8
	Loop Absent	3	12	risk ratio	4
				Odds ratio	16

[0460]

[0461] [표 18c]

329.331		Symptoms	No Symptoms		
	Loop Present	6	1	risk ratio	0.85714
	Loop Absent	1	12	risk ratio	11.1429
				Odds ratio	72

[0462]

[0463] [표 18d]

185.187		Symptoms	No Symptoms		
	Loop Present	3	1	risk ratio	0.75
	Loop Absent	4	12	risk ratio	3
				Odds ratio	9

[0464]

[0465]

[표 19]

```

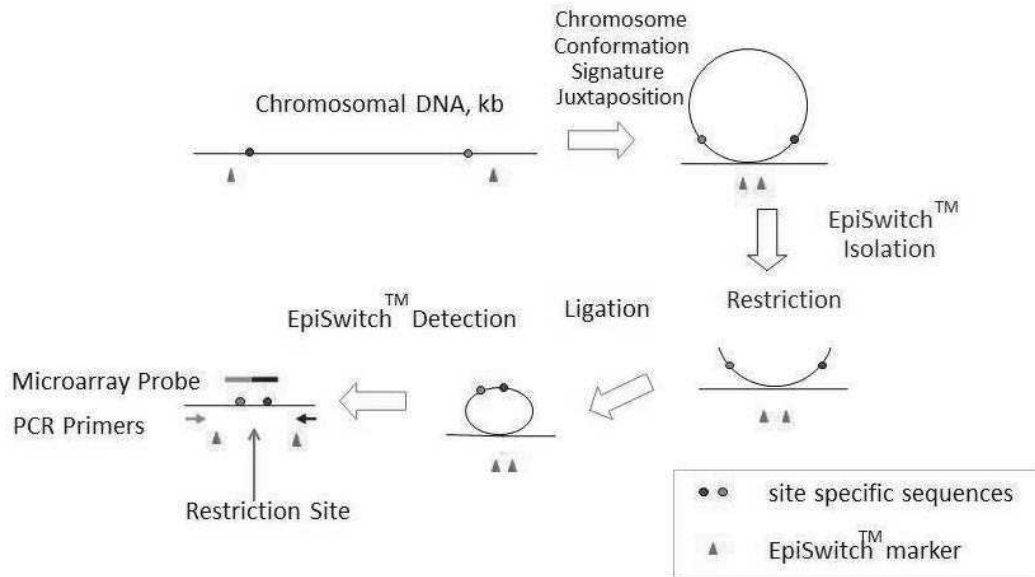
>GRM7_3_7259138_7267165_7394377_7401227_RF
TTGGGCAAATTAATCTGTATGAGTTTCAGTGTGCACATTCACCAAATGTGGATAATAATAATGG
TTAACTCCTAGGATGTTGTGAGAAATTGAAATAATATATCATCTTAGCCAGTATAATCATTTAAAT
TTTTTTTTAGTAAGCATTAGCTGGAAATCATTGGTATTTTATTTTCTTATCTACCCCTCAAACATTGA
AAGTGCTTGGAGCAGAAGTTGTATATTACATACTTTCAGAACTCTCTACCATGGGGCCTTTTGCTT
TAATTAATTTCTGCTATAAATTAATTGCTCTATTATGCCATTTTCTGATTTTCATCCAAGACACACTCTAAA
TGTATATGACGAAGAAGCAATCCCTGGT GACTAAAATGAGATTGCATTTTCTTCGACCATTTGGCC
AGCATGCCAAAACACTGGTAAATGGCAATCCCAAAGCTCACATTTGGCTGATTGATAAAATCACTACTTT
GGGCTCAGTTTGGCCATCTGTAAAGTATGGTGGTAAATGAGCAAGTATTTTATTTA TTTTAAAAA
TACTTTTATTTAGATTTTGGGATACATGTGAAGTTTGCTATATAGGTAATTCATGTCATGGGAGT
TTATTGTAAGATTATTTATCATCCAGGTACTAAGCCTAGTACACAATAGTTATTTTTTTCTGATTC
TTTTCTCCTCCCAACTT LAACCCAGTGGAGGTAGGTC CAGTGTCTGTTGTCCCCCTCTTTGTGTCCA
TGTGTTCTCATCATTTAGCCCCCCTTA TAAGTGATAAATATGTGGT

```

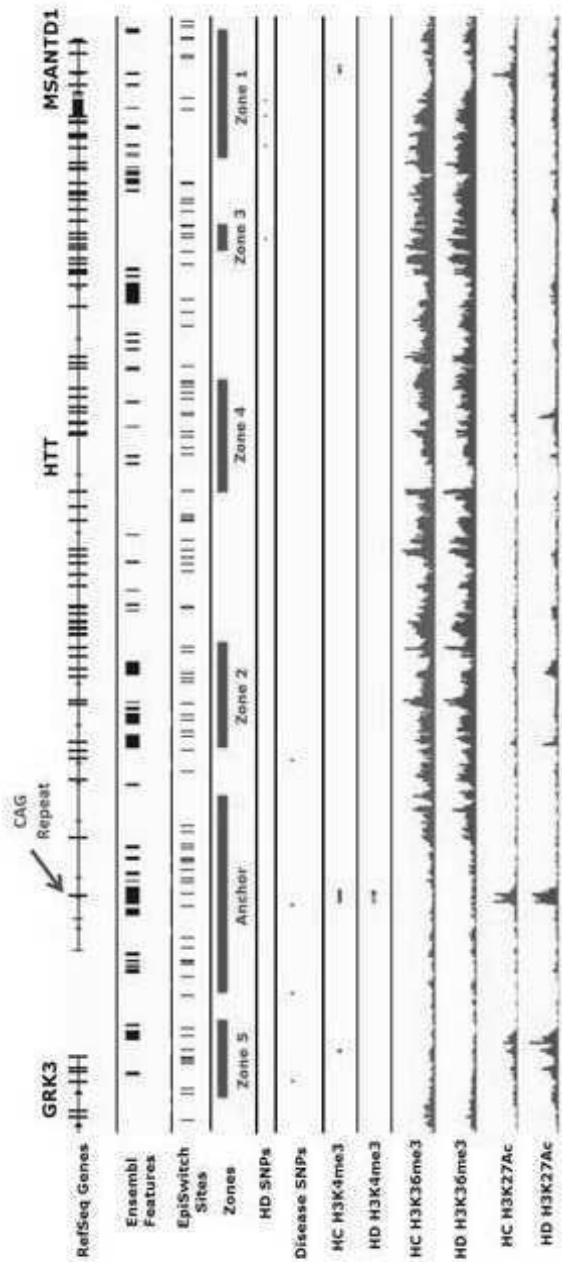
[0466]

도면

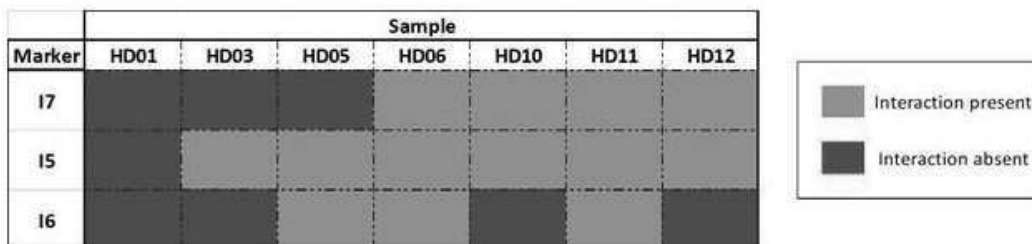
도면1



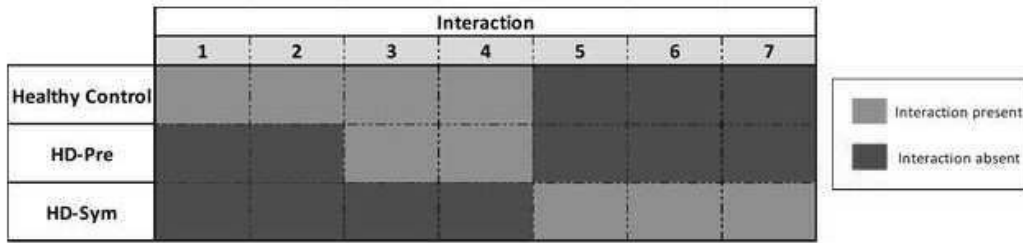
도면2



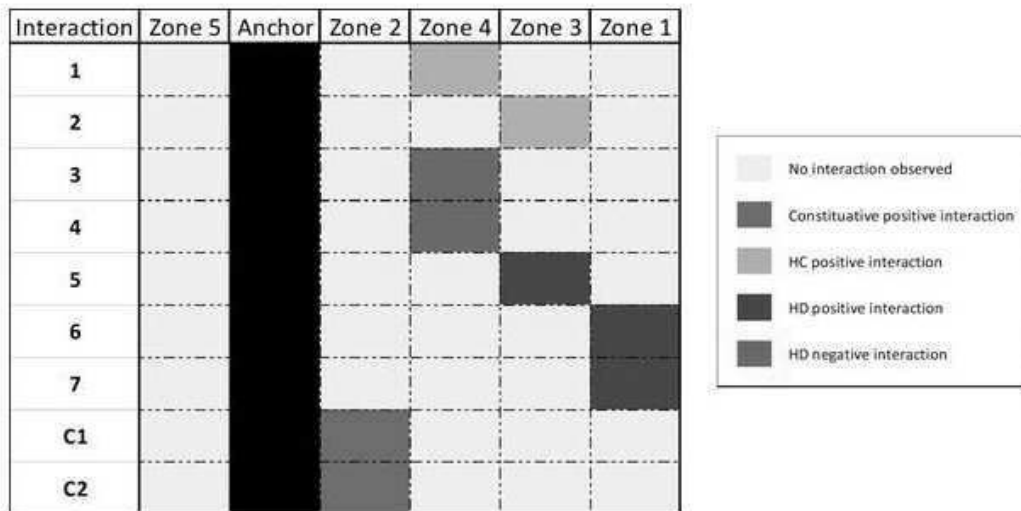
도면3



도면4



도면5



서열 목록

- <110> Oxford BioDynamics Limited
- <120> Biomarker
- <130> IP20-0009
- <150> US 62/566,756
- <151> 2017-10-02
- <150> PCT/GB 2018/052808
- <151> 2018-10-01
- <160> 164
- <170> PatentIn version 3.5
- <210> 1
- <211> 14
- <212> DNA
- <213> Homo sapiens
- <220><221> misc_feature
- <222> (6)

<223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)
 <223> n is a, c, g, or t
 <400> 1
 ccgcgnggng gcag 14

 <210> 2
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 2
 tcaccacaca tcaccacctt gctcctctc gactcttggg gaccacaaca ggggtgccacc 60
 60

 <210> 3
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 3
 gaggtgggtg aatcatgagg tcaagggtc gacaatagtt gagaatctcc aaccacctgg 60
 60

 <210> 4
 <211> 60
 <212> DNA

 <213> Homo sapiens
 <400> 4
 ggccttatag tcagctgac aggtgaaac gattggtcct taggatcagc taccattgc 60
 60

 <210> 5
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 5
 gaggcaggcg gatcacaag tcaaaagac gataacttca ataatagtta cagatgcaaa 60

60

<210> 6
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 6
 agcacatat ctgggatgta gctattgctc gagattgcag tgagctgtga tcacacctct 60

60

<210> 7
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 7
 tcttccctct ttttaaaacc accattcacc gacccacac atcctgtgcc actctactgc 60

60

<210> 8
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 8
 taaccattat gcatcactaa catagcattc gatatgatat gctcagttta gtagggaaa 60

60

<210> 9
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 9
 ggctcaggaa gagaactatt tgtctctttc gacacgcaca tgcaggacac tcacacgtag 60

60

<210> 10
 <211> 60
 <212> DNA
 <213> Homo sapiens

<400> 10
 gttgggtgga tccttgagc tcaggaattc gaagaatgat tttcagccc gtgtggaagg 60
 60

<210> 11
 <211> 60
 <212> DNA

<213> Homo sapiens
 <400> 11
 atcaaaagaa aatagatact tgtcttactic gagttgaata aaatcctcag ctttctgtcc 60
 60

<210> 12
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 12
 aaaagaaact gtgaaaagtt gtcacatttc gattaaatcc aaaaaggtct tctatgaggc 60
 60

<210> 13
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 13
 ttaaaagtat agtagttggc attaacattc gacctttttc tgtttcagta accaaccag 60
 60

<210> 14
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 14
 catcaactaa tagttaaaca ttataatc gactgaagac ctttcatact gtaagattca 60
 60

<210> 15
 <211> 60

<212> DNA
 <213> Homo sapiens
 <400> 15
 catcaactaa tagttaaaca ttataatc gagtctgcag tgagctgaga tcacactgcc 60
 60
 <210> 16
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 16
 ttattccttt ccaaatagtt aaaattattc gaaactttta agaatcaata taaaatttcc 60
 60
 <210> 17
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 17
 cataattata aattaaaaaa tgacactatc gattatgtcc agtgtttctt ggttggtgtc 60
 60
 <210> 18
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 18
 cagagcacta agatagactt ctaaggtttc gaggcataata gctccagctg tattgaggtta 60
 60
 <210> 19
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 19
 aagcattca ttctccctc acc 23
 <210> 20

<211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 20
 actgagcaat gatggcaaca ac 22

<210> 21
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 21
 ctaggcctgc gtttctcgt 20

<210> 22
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 22
 agttctctct ctaagaactc aagga 25

<210> 23
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 23
 ccccaggcac tcacacctta 20

<210> 24
 <211> 21
 <212>
 DNA
 <213> Homo sapiens
 <400> 24
 gggcactcaa caccctttg t 21

<210> 25
 <211> 21
 <212> DNA
 <213> Homo sapiens

<400> 25
 gatgggaatc aaggcaagg g 21
 <210> 26
 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 26
 gcattagcca gcaagcatac ct 22
 <210> 27
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 27
 caccgctga tgcaggtctt 20
 <210> 28
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 28
 ggcaactgttg gtctgaagca c 21
 <210> 29
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 29
 ggtctagatg tcagtctttc c 21
 <210> 30
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 30
 gactataaat ctctccttgt cagc 24
 <210> 31

<211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 31
 ccttaccga caccagtag c 21
 <210> 32
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 32
 cccgacacca ggtagcattc 20
 <210> 33
 <211> 19
 <212> DNA
 <213> Homo sapiens
 <400> 33
 gggcacccca ctagaccac 19
 <210> 34
 <211> 24
 <212>
 DNA
 <213> Homo sapiens
 <400> 34
 attctctccc tggtaaattcc tggc 24
 <210> 35
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 35
 cgccagctca gcagcaataa 20
 <210> 36
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 36

atactcccat ccctaggcc c	21
<210> 37	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 37	
ccctacgact ggcaaacca	20
<210> 38	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 38	
ccctggcatt cacatcaccg a	21
<210> 39	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 39	
gtcaagcaac tgtgtctggg g	21
<210> 40	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 40	
ggatccacga tctccctcca c	21
<210> 41	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 41	
aggtcaggat ggtaccgtt g	21
<210> 42	
<211> 24	

<212> DNA
 <213> Homo sapiens
 <400> 42
 ctgggatgat tcctctggac ttct 24
 <210> 43
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 43
 ggtgggcctg ggtagatgc 20
 <210> 44
 <211> 20
 <212>
 DNA
 <213> Homo sapiens
 <400> 44
 cagctggcgc atccatcacc 20
 <210> 45
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 45
 gagaacgacg acctggcact 20
 <210> 46
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 46
 tcatcctatc ctctcctagc 20
 <210> 47
 <211> 19
 <212> DNA
 <213> Homo sapiens
 <400> 47

actgtaggcc aaccagaag	19
<210> 48	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 48	
ctgggccagt gtcagcgtgt	20
<210> 49	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 49	
gtgactctgc acgcactgtt	20
<210> 50	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 50	
aggctcagca ggtttctgcc	20
<210> 51	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 51	
ctgggcagggt catccagaca g	21
<210> 52	
<211> 21	
<212> DNA	
<213> Homo sapiens	
<400> 52	
tgctagggcc gagtaatcat c	21
<210> 53	
<211> 60	
<212> DNA	

<213> Homo sapiens
 <400> 53
 tatatttaa aatacatact ggtatacatc gatctcatga ctttgctatt atgcatagtg 60
 60

 <210> 54
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 54
 cccagccca gcaacctggc tcacctgac gagtacatct tcaagccatc ctgtgtgccc 60
 60

 <210> 55
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 55
 tataaataat acagctctat ttgcttactc gattaaagaa tcatattata tccttaattc 60
 60

 <210> 56
 <211> 60
 <212> DNA

 <213> Homo sapiens
 <400> 56
 cactatgcat aatagcaaag tcatgagatc gaaaatgttt gtcaagcagt aggttttggg 60
 60

 <210> 57
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 57
 gattttagaa tccttaacaa ggctgcaatc gaggttagct gctgcagaaa gaagagaaaa 60
 60

 <210> 58

<211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 58

 ggtgactaaa atgagattgc attttctttc gaccatttgg ccagcatgcc aaacactggt 60
 60
 <210> 59
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 59
 taacctctcc ttcttaggt tctccatata gatagaaaat tgtctgcagc ccttaatgcc 60
 60
 <210> 60
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 60
 agctcacggt cagtgcggt cggtttgctc gaataaagaa caaggacctt aaaaaataga 60
 60
 <210> 61
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 61
 aaagtcatac aactactatg taagatattc gaatacctgt tagaataggt gaagtttat 60
 60
 <210> 62
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 62
 tatatttaaa aatacatact ggtatacatc gatctcatga ctttgctatt atgcatagtg 60

60

<210> 63

<211> 60

<212> DNA

<213> Homo sapiens

<400> 63

ccccagccca gcaacctggc tcacctgac gagtacatct tcaagccatc ctgtgtgcc

60

60

<210> 64

<211> 60

<212> DNA

<213> Homo sapiens

<400> 64

cactatgcat aatagcaaag tcatgagac gaaaatgttt gtcaagcagt aggttttggg

60

60

<210> 65

<211> 60

<212> DNA

<213> Homo sapiens

<400> 65

gattttagaa tctctaaca ggctgcaatc gaggttagct gctgcagaaa gaagagaaaa

60

60

<210> 66

<211> 60

<212> DNA

<213> Homo sapiens

<400> 66

ggtgactaaa atgagattgc attttctttc gaccatttgg ccagcatgcc aaacctggt

60

60

<210> 67

<211> 60

<212> DNA

<213> Homo sapiens

<400> 67

taacctctcc tttcttaggt tctccatc gatagaaat tgtctgcagc ccttaatgcc 60
60

<210> 68
<211> 60
<212> DNA
<213> Homo sapiens

<400> 68

agctcacggt cagtgccgtt ccgtttgctc gaataaagaa caaggacctt aaaaaataga 60
60

<210> 69
<211> 60
<212> DNA
<213> Homo sapiens

<400> 69

aaagtcatac aactactatg taagatattc gaatacctgt tagaataggt gaaggtttat 60
60

<210> 70
<211> 22
<212> DNA
<213> Homo sapiens

<400> 70

gcagcacaca gggaactctc tt 22

<210> 71
<211> 20
<212> DNA
<213> Homo sapiens

<400> 71

gctaggaagg gcctgggatg 20

<210> 72
<211> 25
<212> DNA
<213> Homo sapiens

<400> 72
tccaaaatgt ttaatactgc ctaga 25

<210> 73
<211> 20
<212> DNA
<213> Homo sapiens
<400> 73
cagccactgt agagagcagt 20

<210> 74
<211> 24
<212> DNA
<213> Homo sapiens
<400> 74
acttcttccc aagtcacttt ttgc 24

<210> 75
<211> 22
<212> DNA
<213> Homo sapiens
<400> 75
tgacgaagaa gcaatccctg gt 22

<210> 76
<211> 24
<212>
DNA
<213> Homo sapiens
<400> 76
aatctctgcc ctctctcat cttg 24

<210> 77
<211> 20
<212> DNA
<213> Homo sapiens
<400> 77
aggtatgcag ccagcctgag 20

<210> 78

<211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 78
 ccgtgccata tcctctgatt tatgc 25
 <210> 79
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 79
 ttgttgagcc cagcaattcc ttt 23
 <210> 80
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 80
 ctcagtgggc acacactcca 20
 <210> 81
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 81
 agccatgtgg tctggaatct 20
 <210> 82
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 82
 taaccaccca gcagcaaggt 20
 <210> 83
 <211> 20
 <212> DNA
 <213> Homo sapiens

<400> 83
 tggccatctt gctttgcctc 20
 <210> 84
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 84
 ggacctacct ccaactgggtt g 21
 <210> 85
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 85
 catgttccca cagcaaggaa gtta 24
 <210> 86
 <211> 20
 <212>
 DNA
 <213> Homo sapiens
 <400> 86
 tttccgtgcc agtgtcctgt 20
 <210> 87
 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 87
 ggctgacctt caacagattc gc 22
 <210> 88
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 88
 tcttgtagac ggttggtggt 20
 <210> 89
 <211> 24

<212> DNA
 <213> Homo sapiens

 <400> 89
 tgtcacctat gtgctgagta ctgg 24
 <210> 90
 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 90
 tgacgaagaa gcaatccctg gt 22
 <210> 91
 <211> 21
 <212> DNA
 <213> Homo sapiens
 <400> 91
 ggacctacct ccaactgggtt g 21
 <210> 92
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 92
 ccgtgccata tcctctgatt tatgc 25

 <210> 93
 <211> 22
 <212> DNA
 <213> Homo sapiens
 <400> 93
 ggctgacctt caacagattc gc 22
 <210> 94
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 94

acttcttccc aagtcacttt ttgc	24
<210> 95	
<211> 20	
<212> DNA	
<213> Homo sapiens	
<400> 95	
tgccatctt gctttgcctc	20
<210> 96	
<211> 25	
<212>	
DNA	
<213> Homo sapiens	
<400> 96	
ggatatgcag tttcctggc actac	25
<210> 97	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 97	
catgctaggg ccgagtaatc atct	24
<210> 98	
<211> 22	
<212> DNA	
<213> Homo sapiens	
<400> 98	
gcagcacaca ggaactctc tt	22
<210> 99	
<211> 23	
<212> DNA	
<213> Homo sapiens	
<400> 99	
ttgttgagcc cagcaattcc ttt	23
<210> 100	
<211> 20	

<212>	DNA	
<213>	Homo sapiens	
<400>	100	
	cagccactgt agagagcagt	20
<210>	101	
<211>	20	
<212>	DNA	
<213>	Homo sapiens	
<400>	101	
	taaccaccca gcagcaaggt	20
<210>	102	
<211>	25	
<212>	DNA	
<213>	Homo sapiens	
<400>	102	
	agtagcttcc ctgtagagg tcttg	25
<210>	103	
<211>	23	
<212>	DNA	
<213>	Homo sapiens	
<400>	103	
	agccagtgac tccacaactt ctt	23
<210>	104	
<211>	28	
<212>	DNA	
<213>	Homo sapiens	
<400>	104	
	aggttgcttt cgaagtacag atatcact	28
<210>	105	
<211>	26	
<212>	DNA	
<213>	Homo sapiens	
<400>	105	
	attdtttttc gaccatttgg ccagca	26

<210>	106	
<211>	31	
<212>	DNA	
<213>	Homo sapiens	
<400>	106	
	ttctaacagg tattcgaata tcttcatag t	31
<210>	107	
<211>	27	
<212>	DNA	
<213>	Homo sapiens	
<400>	107	
	aagtaacatt cgacaaacct ctggtaa	27
<210>	108	
<211>	27	
<212>	DNA	
<213>	Homo sapiens	
<400>	108	
	tattctcttt gatcgaggca tatagct	27
<210>	109	
<211>	29	
<212>	DNA	
<213>	Homo sapiens	
<400>	109	
	aaagtcatga gatcgaatga taccagtat	29
<210>	110	
<211>	28	
<212>	DNA	
<213>	Homo sapiens	
<400>	110	
	ttgacaaaca ttttcgatct catgactt	28
<210>	111	
<211>	25	
<212>	DNA	

<213> Homo sapiens
 <400> 111
 tgtggctcag tcgacagaaa gtaca 25
 <210> 112
 <211> 20
 <212> DNA
 <213> Homo sapiens
 <400> 112
 tcttgtacac ggttgggtgt 20

 <210> 113
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 113
 ggatatgcag tttcctggc actac 25
 <210> 114
 <211> 25
 <212> DNA
 <213> Homo sapiens
 <400> 114
 agtagcttcc ctgtagagg tcttg 25
 <210> 115
 <211> 24
 <212> DNA
 <213> Homo sapiens
 <400> 115
 tgtcacctat gtgctgagta ctgg 24
 <210> 116
 <211> 24

 <212> DNA
 <213> Homo sapiens
 <400> 116
 catgctaggg ccgagtaatc atct 24

<210> 117
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 117
 agccagtgac tccacaactt ctt 23
 <210> 118
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 118
 agatctagtt cacagtagca ccaatatac gacagatagc tgacatcac ctcccaatgt 60
 60
 <210> 119
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 119
 agagtacttc ctaactccta ctgtacactc gacagatagc tgacatcac ctcccaatgt 60
 60
 <210> 120
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 120
 tataaccagt gtcctacga aggcgccttc gaagtctcaa acttcacttc tctgtgcgc 60
 60
 <210> 121
 <211> 60
 <212>
 > DNA
 <213> Homo sapiens
 <400> 121
 agagtacttc ctaactccta ctgtacactc gaggatgac gctccgacag ctctccagc 60

		60
<210>	122	
<211>	60	
<212>	DNA	
<213>	Homo sapiens	
<400>	122	
	gggtttcgcc atgttgcca ggctggctc gaagttgatg catctgtgct cacgtttgca	60
		60
<210>	123	
<211>	60	
<212>	DNA	
<213>	Homo sapiens	
<400>	123	
	agatctagtt cacagtagca ccaatatac gactgtctcc tgttgccat ctctcacct	60
		60
<210>	124	
<211>	60	
<212>	DNA	
<213>	Homo sapiens	
<400>	124	
	agatctagtt cacagtagca ccaatatac gaactcctga ccttgtgatc cacccacctc	60
		60
<210>	125	
<211>	26	
<212>	DNA	
<213>	Homo sapiens	
<400>	125	
	gaataccag gaatgcttac ttgagc	26
<210>	126	
<211>	26	
<212>	DNA	
<213>	Homo sapiens	
<400>	126	

accaaagcc atctgggaca catcca	26
<210> 127	
<211> 23	
<212> DNA	
<213> Homo sapiens	
<400> 127	
cccgctaagt ccaccctct gta	23
<210> 128	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 128	
agtctgccca ctgaggtaac taacaa	26
<210> 129	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 129	
gattccagcc acccacctt cacaag	26
<210> 130	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 130	
tgattccage cacccacctt tcacaa	26
<210> 131	
<211> 23	
<212> DNA	
<213> Homo sapiens	
<400> 131	
gaaccgcact taccctcagc agt	23
<210> 132	
<211> 26	
<212> DNA	

<213> Homo sapiens

<400> 132

ggacaagcag acacactacc tgaact 26

<210> 133

<211> 26

<212> DNA

<213> Homo sapiens

<400> 133

atccccctgaa acagaaggac ctcgtg 26

<210> 134

<211> 26

<212> DNA

<213> Homo sapiens

<400> 134

gagccgtctc ataataacct cagggt 26

<210> 135

<211> 26

<212> DNA

<213> Homo sapiens

<400> 135

cagtggttta gggcaaagag agggag 26

<210> 136

<211> 60

<212> DNA

<213> Homo sapiens

<400> 136

agatctagtt cacagtagca ccaatatac gacattaagt tcatgagagt cttctatatt 60

60

<210> 137

<211> 60

<212> DNA

<213> Homo sapiens

<400> 137

gcgcacagga gaagtgaagt ttgagacttc gaaaaaggaa taaattgaaa aatagaggaa 60
60

<210> 138
<211> 60
<212>
> DNA
<213> Homo sapiens
<400> 138

agagtacttc ctaactccta ctgtacactc gatgttttgg gtattgaatg tggttaagtgg 60
60

<210> 139
<211> 60
<212> DNA
<213> Homo sapiens
<400> 139

atcaagactg tatggtactg gcacaggatc gaggatgatc gctccgacag ctctccagc 60
60

<210> 140
<211> 60
<212> DNA
<213> Homo sapiens
<400> 140

atcaagactg tatggtactg gcacaggatc gatgttttgg gtattgaatg tggttaagtgg 60
60

<210> 141
<211> 60
<212> DNA
<213> Homo sapiens
<400> 141

atcaagactg tatggtactg gcacaggatc gacattaagt tcatgagagt cttctatatt 60
60

<210> 142
<211> 60
<212> DNA

<213> Homo sapiens
 <400> 142
 tgtccagtta aattgcagat atttccgac gagtgtacag taggagttag gaagtactct 60
 60
 <210> 143
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 143
 atcaagactg tatggtactg gcacaggatc gaactcctga ccttgtgac cacccacctc 60
 60
 <210> 144
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 144
 agatctagtt cacagtagca ccaatatac gaatctgggt gtttcaaaga catcaacaaa 60
 60
 <210> 145
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 145
 agatctagtt cacagtagca ccaatatac gaggatgac gctccgacag ctctccagc 60
 60
 <210> 146
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 146
 gggtttcacc atgttgcca ggatggtctc gaactcctga ccttgtgac cacccacctc 60
 60
 <210> 147

<211> 60
 <212>
 > DNA
 <213> Homo sapiens
 <400> 147
 agatctagtt cacagtagca ccaatatac gaagcggcct tcgtaggagc actggttata 60
 60
 <210> 148
 <211> 60
 <212> DNA
 <213> Homo sapiens
 <400> 148
 agatctagtt cacagtagca ccaatatac gactaaggtc agccctccgt gctgggggct 60
 60
 <210> 149
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 149
 ggtccagaga accgcactta ccc 23
 <210> 150
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 150
 tcctctcag aatcctactt ggcttc 26
 <210> 151
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 151
 gaacagtttg tggtagtat gcggtc 26
 <210> 152
 <211> 26

<212> DNA
 <213> Homo sapiens
 <400> 152
 tctggacaag gtatctggaa ggctga 26

<210> 153
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 153
 gcactgctcg caatagccaa gaacta 26

<210> 154
 <211> 23
 <212> DNA
 <213> Homo sapiens
 <400> 154
 cttgctgacc ctctatgga tgg 23

<210> 155
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 155
 gcacagagtt ccaacatttc ctccac 26

<210> 156
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 156
 ggacaagcag acacactacc tgaact 26

<210> 157
 <211> 26
 <212> DNA
 <213> Homo sapiens
 <400> 157

gcactgctcg caatagccaa gaacta	26
<210> 158	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 158	
acagtggttt agggcaaaga gaggga	26
<210> 159	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 159	
gaagaagcct ctgtttggtc tggaaa	26
<210> 160	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 160	
ggaaggacaa gcagacacac tacctg	26
<210> 161	
<211> 26	
<212> DNA	
<213> Homo sapiens	
<400> 161	
gtggtttagg gcaaagagag ggagat	26
<210> 162	
<211> 25	
<212> DNA	
<213> Homo sapiens	
<400> 162	
tccccgagtc ctgtgatggc aaact	25
<210> 163	
<211> 25	

<212> DNA
 <213> Homo sapiens
 <400> 163
 atggctacgg aaagggcatt cggac 25
 <210> 164
 <211> 800
 <212> DNA
 <213> Homo sapiens
 <400> 164
 ttgggcaaat tacttattct gtatgagttt cagtgtgcac attcacaaa tgtggataat 60
 aataatggtt aactcctagg atgtttgtga gaattgaaat aatatatcat tcttagccca 120
 gtataatcat ttaaattttt tttttagtaa gcattagctg gaatcattgg tattgttttt 180
 attttcttat ctaccctcaa acattgaaag tgccttggag cagaagtgt atattacata 240

 ctttcagaat ctctctacca tggggccttt tgetttaatt aatttctgct ataattaatt 300
 gctctattat gccattttct gatttcatcc aagacacact ctaaatgtat atgacgaaga 360
 agcaatccct ggtgactaaa atgagattgc attttctttc gaccatttgg ccagcatgcc 420
 aaacactggt aaattggcat tccaaagctc acattggctg attgataaat tcactacttt 480
 gggctcagtt ttgcatctg taaagtgatg gtgatcaaat gagcaagtgt atttatttat 540
 tttaaaata cttttatttt agattttggg atacatgtgt aagtttgcta tataggtaaa 600
 ttcatgtcat gggagtttat tgtaaagatt atttcatcat ccaggtacta agcctagtag 660

 acaatagtta ttttttttct gattcttttc ctctcccaa cttcaacca gtggagtag 720
 gtcccagtgt ctgttgccc cctctttgtg tccatgtgtt ctcatcattt agccccact 780
 tataagtgat aatatgtggt 800