



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년03월18일

(11) 등록번호 10-2648965

(24) 등록일자 2024년03월14일

(51) 국제특허분류(Int. Cl.)  
C12Q 1/6886 (2018.01)

(52) CPC특허분류  
C12Q 1/6886 (2022.01)  
C12Q 2600/154 (2013.01)

(21) 출원번호 10-2019-7005873

(22) 출원일자(국제) 2017년09월01일  
심사청구일자 2020년08월14일

(85) 번역문제출일자 2019년02월27일

(65) 공개번호 10-2019-0045188

(43) 공개일자 2019년05월02일

(86) 국제출원번호 PCT/US2017/049915

(87) 국제공개번호 WO 2018/045322  
국제공개일자 2018년03월08일

(30) 우선권주장  
62/383,165 2016년09월02일 미국(US)

(56) 선행기술조사문헌  
Dig Dis Sci (2016) Vol.61, pp149-157, DOI  
10.1007/s10620-015-3878-3  
US20160017430 A1

(73) 특허권자  
메이오 파운데이션 포 메디칼 에듀케이션 앤드 리  
써치  
미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200  
이그젝트 사이언시즈 코포레이션  
미국 위스콘신 53719 메디슨 인테버 레인 5505

(72) 발명자  
테일러, 윌리엄 알.  
미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200 메이오 파운데이션 포 메디칼  
에듀케이션 앤드 리써치 내  
알퀴스트, 데이비드 에이.  
미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200 메이오 파운데이션 포 메디칼  
에듀케이션 앤드 리써치 내  
(뒷면에 계속)

(74) 대리인  
특허법인 무한

전체 청구항 수 : 총 10 항

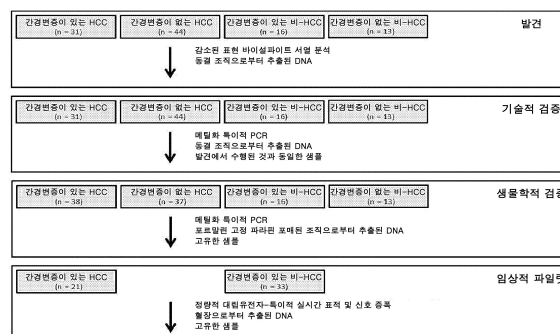
심사관 : 손영희

(54) 발명의 명칭 간세포 암종 검출

## (57) 요약

간세포 암종 스크리닝에 대한 기술과 특허, 비배타적으로, 간세포 암종의 존재를 검출하기 위한 방법, 조성물, 및 관련된 용도가 본 명세서에 제공된다.

## 대표도



(52) CPC특허분류

C12Q 2600/16 (2013.01)

C12Q 2600/166 (2013.01)

(72) 발명자

**마호니, 더글라스 더블유.**

미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200 메이오 파운데이션 포 메디칼 에  
듀케이션 앤드 리써치 내

**키셀, 존 비.**

미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200 메이오 파운데이션 포 메디칼 에  
듀케이션 앤드 리써치 내

**애플, 트레이시 씨.**

미국 55905 미네소타주 록체스터 퍼스트 스트리트  
에스.더블유. 200 메이오 파운데이션 포 메디칼 에  
듀케이션 앤드 리써치 내

**알라위, 하킴 티.**

미국 위스콘신 53719 매디슨 차머니 드라이브 441  
이그젝트 싸이언스 코퍼레이션 내

**지아코우모파울러스, 마리아**

미국 위스콘신 53719 매디슨 차머니 드라이브 441  
이그젝트 싸이언스 코퍼레이션 내

**리드가드, 그라함 피.**

미국 위스콘신 53719 매디슨 차머니 드라이브 441  
이그젝트 싸이언스 코퍼레이션 내

## 명세서

### 청구범위

#### 청구항 1

간세포 암종(HCC)을 갖거나 갖는 것으로 의심되는 대상체의 샘플에서 유전자 또는 유전자들의 메틸화 수준을 결정하는 방법으로서,

상기 방법은,

상기 대상체로부터의 샘플 중의 게놈 DNA를 바이설파이트(bisulfite)로 처리하는 단계;

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에 대한 프라이머의 세트를 사용하여 바이설파이트-처리된 게놈 DNA를 증폭시키는 단계; 및

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 적어도 하나의 CpG 부위의 메틸화 수준을 메틸화-특정 PCR, 정량적 메틸화-특이적 PCR, 메틸화 민감성 DNA 제한효소 분석, 정량적 바이설파이트 파이로서열분석, 또는 바이설파이트 게놈 서열분석 PCR에 의해 결정하는 단계

를 포함하는, 방법.

#### 청구항 2

제1항에 있어서, 상기 방법은,

상기 메틸화 수준을 HCC가 없는 대조군 샘플로부터의 상응하는 유전자 또는 유전자들의 메틸화 수준과 비교하는 단계; 및

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 측정된 메틸화 수준이 상기 대조군 샘플에서 측정된 메틸화 수준보다 더 높은 경우, 상기 HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 측정된 메틸화 수준을 대상체가 HCC를 갖는 것으로 결정하는 지표로 사용하는 단계를 더 포함하는, 방법.

#### 청구항 3

제1항에 있어서,

상기 프라이머의 세트(들)는 하기를 포함하는 것인, 방법:

DAB2IP에 대해 서열번호:4 및 5 또는 서열번호:46 및 47로 구성되는 프라이머 세트,

EMX1에 대해 서열번호:16 및 17 또는 서열번호:91 및 92로 구성되는 프라이머 세트,

HOXA1에 대해 서열번호:28 및 29 또는 서열번호:88 및 89로 구성되는 프라이머 세트, 및

TSPYL5에 대해 서열번호:1 및 2 또는 서열번호:43 및 44로 구성되는 프라이머 세트.

#### 청구항 4

제1항에 있어서,

상기 샘플은 혈장 샘플 또는 조직 샘플인 것인, 방법.

#### 청구항 5

제1항에 있어서,

상기 적어도 하나의 CpG 부위는 코딩 영역 또는 조절 영역 내에 존재하는 것인, 방법.

## 청구항 6

제1항에 있어서,

상기 적어도 하나의 CpG 부위의 메틸화 수준을 결정하는 단계는 상기 적어도 하나의 CpG 부위의 메틸화 점수를 결정하는 것 또는 상기 적어도 하나의 CpG 부위의 메틸화 빈도를 결정하는 것을 포함하는 것인, 방법.

## 청구항 7

간세포 암종(HCC)을 갖거나 갖는 것으로 의심되는 대상체의 혈장 샘플에서 유전자 또는 유전자들의 메틸화 수준을 결정하는 방법으로서,

상기 방법은,

상기 대상체로부터의 혈장 샘플 중의 게놈 DNA를 바이설파이트로 처리하는 단계;

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에 대한 하기 프라이머의 세트들을 사용하여 바이설파이트-처리된 게놈 DNA를 증폭시키는 단계:

DAB2IP에 대해 서열번호:4 및 5 또는 서열번호:46 및 47로 구성되는 프라이머 세트,

EMX1에 대해 서열번호:16 및 17 또는 서열번호:91 및 92로 구성되는 프라이머 세트,

HOXA1에 대해 서열번호:28 및 29 또는 서열번호:88 및 89로 구성되는 프라이머 세트,

TSPYL5에 대해 서열번호:1 및 2 또는 서열번호:43 및 44로 구성되는 프라이머 세트;

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 적어도 하나의 CpG 부위의 메틸화 수준을 메틸화-특정 PCR, 정량적 메틸화-특이적 PCR, 메틸화 민감성 DNA 제한효소 분석, 정량적 바이설파이트 파이로서열분석, 또는 바이설파이트 게놈 서열분석 PCR에 의해 결정하는 단계;

상기 메틸화 수준을 HCC가 없는 대조군 샘플로부터의 상응하는 유전자 또는 유전자들의 메틸화 수준과 비교하는 단계; 및

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 측정된 메틸화 수준이 상기 대조군 샘플에서 측정된 메틸화 수준보다 더 높은 경우, 상기 HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에서 측정된 메틸화 수준을 대상체가 HCC를 갖는 것으로 결정하는 지표로 사용하는 단계

를 포함하는, 방법.

## 청구항 8

제7항에 있어서,

상기 적어도 하나의 CpG 부위는 코딩 영역 또는 조절 영역 내에 존재하는 것인, 방법.

## 청구항 9

제7항에 있어서,

상기 적어도 하나의 CpG 부위의 메틸화 수준을 결정하는 단계는 상기 적어도 하나의 CpG 부위의 메틸화 점수를 결정하는 것 또는 상기 적어도 하나의 CpG 부위의 메틸화 빈도를 결정하는 것을 포함하는 것인, 방법.

#### 청구항 10

간세포 암종(HCC)을 갖거나 갖는 것으로 의심되는 대상체의 샘플에서 유전자 또는 유전자들의 메틸화 수준을 결정하기 위한 키트로서, 하기 중 둘 이상을 포함하는 키트:

바이선틸라이트 시약;

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자로부터 선택된 차별적으로 메틸화된 영역(DMR)로부터의 서열을 포함하고, 암을 갖지 않는 대상체와 관련된 메틸화 상태(methylation state)를 갖는 대조 핵산; 및

HOXA1; 및 EMX1, DAB2IP, 및 TSPYL5로부터 선택된 적어도 하나의 추가 유전자에 대한 하기 프라이머의 세트:

DAB2IP에 대해 서열번호:4 및 5, 또는 서열번호:46 및 47로 구성되는 프라이머 세트,

EMX1에 대해 서열번호:16 및 17, 또는 서열번호:91 및 92로 구성되는 프라이머 세트,

HOXA1에 대해 서열번호:28 및 29, 또는 서열번호:88 및 89로 구성되는 프라이머 세트, 및

TSPYL5에 대해 서열번호:1 및 2, 또는 서열번호:43 및 44로 구성되는 프라이머 세트.

#### 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

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

## 발명의 설명

## 기술 분야

[0001] 관련 출원들에 대한 교차참조

[0002] 본원은 미국특허의 우선권 이점을 2016년 9월 2일 출원된 미국 가특허출원 제62/383,165호로 주장하고, 이것은 전체적으로 참고로 편입된다.

[0003] 발명의 분야

[0004] 간세포 암종 스크리닝에 대한 기술과 특히, 비배타적으로, 간세포 암종의 존재를 검출하기 위한 방법, 조성물, 및 관련된 용도가 본 명세서에 제공된다.

## 배경 기술

[0005] 간세포 암종 (HCC)은 간의 주요 악성 종양이며 만성 간 질환 및 간경변증을 앓고있는 환자에서 주로 발생한다. 기원 세포(들)은 간 줄기 세포인 것으로 여겨진다 (참고 문헌, Alison MR. *Stem Cell Rev.* 2005.1(3): 253-60). 종양은 국소 팽창, 간내 전이, 및 원격 전이로 진행된다.

[0006] HCC는 현재 500,000명이 넘는 사람들이 걸려, 전세계적인 암 사망의 세 번째 주요 원인이다. HCC의 발병률은 아시아와 아프리카에서 최고인데, 여기서는 B형 간염과 C형 간염의 풍토병성 높은 유병률이 만성 간 질환의 발병과 이에 따른 HCC의 발병을 강하게 하기 쉽다.

[0007] HCC의 발효는 지난 수십 년 동안 상당히 발전했다. 과거에 HCC는 일반적으로 우상-사분면 통증, 체중 감소 및 대성기능장애 간 질환의 징후를 보였으나, 단면 이미징형성 연구 및 혈청 알파-태아단백 (AFP) 측정법을 사용하여 알려진 간경변증이 있는 환자의 일상적인 스크리닝의 결과로 이제 점진적으로 아주 초기단계에서 발견되고 있다.

[0008] HCC의 위협은 앞으로도 계속 증가할 것으로 예상된다 (참고 문헌, Llovet JM, 등, *Liver Transpl.* 2004 Feb. 10(2 Suppl 1): S115-20). 따라서, 이들 환자의 생존율을 향상시키기 위해서는 HCC의 조기 검출에 대한 필요성이 매우 크다.

## 발명의 내용

## 해결하려는 과제

## 과제의 해결 수단

[0009] 간세포 암종 (HCC)은 전세계적으로 2번째로 가장 치명적인 암이다. 초기 단계 검출로 생존율이 향상되고, 정확한 비 침습성 선별 도구가 필요하다. HCC의 가장 초기 단계의 사전-증상 검출을 위한 정확하고, 저렴하며, 안전한 선별 도구를 제공할 혁신에 대한 절박성이 있다.

[0010] 본 발명은 이러한 요구를 해결한다. 사실상, 본 발명은 HCC를 정상 대조군 (간경변증을 동반하거나 동반하지 않는 대조군)과 구별하는 신규한 메틸화된 DNA 마커를 제공한다.

[0011] 메틸화된 DNA는 대부분의 종양 유형의 조직에서 잠재적인 부류의 바이오마커로 연구되었다. 많은 사례에서, DNA 메틸전달효소는 유전자 발현의 후성유전적 제어로서 시토신-포스페이트-구아닌 (CpG) 섬 부위에서 메틸기를 DNA에 부가한다. 생물학적으로 매력적인 기전에서, 종양 억제인자 유전자의 프로모터 영역에서 획득된 메틸화 사건



은 발현을 침묵시키는 것으로 생각되며, 따라서 종양 형성에 기여한다. DNA 메틸화는 RNA 또는 단백질 발현보다 화학적으로 그리고 생물학적으로 안정한 진단 도구일 수 있다 (Laird (2010) Nat Rev Genet 11:191-203). 게다가, 산발성 결장암과 같은 다른 암에서, 메틸화 마커는 탁월한 특이성을 제공하며 개별 DNA 돌연변이보다 더 광범위하게 유익하고 민감하다 (참고 문헌, Zou et al (2007) Cancer Epidemiol Biomarkers Prev 16:2686-96).

[0012] CpG 섬의 분석은 동물 모델과 인간 세포주에 적용할 때 중요한 발견을 생성했다. 예를 들어, 코뿔과 그 동료들은 동일한 CpG 섬의 다른 부분에 있는 앰플리콘이 상이한 수준의 메틸화를 가질 수 있다는 것을 발견했다 (Zhang 등 (2009) PLoS Genet 5: e1000438). 또한, 메틸화 수준은 고도의 메틸화 서열과 비메틸화 서열 사이에 양봉형으로 분포되어 DNA 메틸전달효소 활성의 2원 스위치-유사 패턴을 추가적으로 지지한다 (Zhang 등 (2009) PLoS Genet 5: e1000438). 생체내 췌장과 조직과 시험관내 세포주의 분석은 단지 약 0.3%의 높은 CpG 밀도 프로모터 (HCP, 300 염기쌍 영역 내에서 >7% CpG 서열을 갖는 것으로 정의됨)가 메틸화된 반면에, 낮은 CpG 밀도 (300 염기쌍 영역 내에서 <5% CpG 서열을 갖는 것으로 정의됨)의 영역은 빈번하게 동적 조직-특이적 패턴에서 메틸화되는 경향이 있었다는 것을 실증했다 (Meissner 등 (2008) Nature 454:766-70). HCP는 도처에 존재하는 하우스 키팅 유전자와 고도로 조절된 발달성 유전자를 위한 프로모터를 포함한다. >50%로 메틸화된 HCP 부위 중에는 몇 개의 확립된 마커 예컨대 Wnt 2, NDRG2, SFRP2, 및 BMP3이 있다 (Meissner 등 (2008) Nature 454:766-70).

[0013] 본 발명에 대한 구현예를 전개하는 과정에서 수행된 실험은 HCC를 갖는 대상체의 혈장으로부터의 DNA 마커의 메틸화 상태를 대조군 대상체 (예를 들어, 간경변증이 있거나 또는 정상인 대상체)로부터의 동일한 DNA 마커의 메틸화 상태와 비교하였다. 이러한 실험은 HCC를 그러한 대조군 그룹과 구별하는 메틸화된 DNA 마커 후보를 확인하고 입증하였다.

[0014] 따라서, HCC 스크리닝 (예를 들어, 조사)에 대한 기술과 특히, 비배타적으로, HCC의 존재를 검출하기 위한 방법, 조성물, 및 관련된 용도가 본 명세서에 제공된다.

[0015] HCC를 검출할 수 있는 마커 및/또는 마커의 패널이 확인되었다 (다음 참고, 실시예 I, II 및 III) (ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFNB2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRNI, 및 ECE1).

[0016] 본 명세서에서 기재된 바와 같이, 본 기술은 대상체에서 HCC의 존재를 검출하기 위한 높은 식별력을 갖는 수 많은 메틸화된 DNA 마커 및 이의 서브셋 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 마커의 세트)를 제공한다. 실험은, 예를 들어 스크리닝 또는 진단 (예를 들어, HCC 스크리닝 또는 진단)의 목적으로 배지 (예를 들어, 혈장)을 분석할 때, 높은 특이성을 제공하기 위한 낮은 배경 수준과 높은 신호 대 잡음비를 제공하는 마커를 확인하기 위해 후보 마커에 선별 필터를 적용하였다.

[0017] 일부 구현예에서, 본 기술은 생물학적 샘플 (예를 들어, 혈장 샘플)에서 본 명세서에서 확인된 하나 이상의 마커 중 메틸화 상태와 그 존재를 평가하는 것에 관한 것이다. 이들 마커는 본 명세서에서 논의된 바와 같은, 예를 들어, 표 1 및 4에서 제공된 바와 같은 하나 이상의 차별적으로 메틸화된 영역 (DMR)을 포함한다. 메틸화 상태는 본 기술의 구현예에서 평가된다. 이와 같이, 본 명세서에서 제공된 기술은 유전자의 메틸화 상태가 측정되어 지는 방법에 제한되지 않는다. 예를 들어, 일부 구현예에서 메틸화 상태는 게놈 스캐닝 방법에 의해 측정된다. 예를 들어, 하나의 방법은 제한 랜드마크 게놈스캐닝을 포함하고 (참고 문헌, Kawai 등 (1994) Mol. Cell.Biol. 14:7421-7427) 그리고 또 다른 실시예는 메틸화-민감성인 임의로 프라이밍된 PCR을 포함한다 (참고 문헌, Gonzalgo 등 (1997) Cancer Res. 57:594-599). 일부 구현예에서, 특정 CpG 부위에서의 메틸화 패턴의 변화는 메틸화-민감성 제한 효소로 게놈 DNA를 절단한 후 관심 영역의 서던 분석 (절단-서던 방법)에 의해 모니터링된다. 일부 구현예에서, 메틸화 패턴의 변화를 분석하는 것은 PCR 증폭 전에 메틸화 민감성 제한 효소로 게놈 DNA의 절단을 포함하는 PCR-기재 공정을 포함한다 (참고 문헌, Singer-Sam 등 (1990) Nucl. Acids Res. 18:687). 또한, 메틸화 분석의 개시점으로 DNA의 바이선폴라이트 처리를 이용하는 다른 기술이 보고되었다. 이들은 하기를 포함한다: 메틸화-특이적 PCR (MSP) (참고 문헌, Herman 등 (1992) Proc. Natl. Acad. Sci. USA 93:9821-9826) 및 바이선폴라이트-전환된 DNA로부터 증폭된 PCR 생성물의 제한효소절단 (참고 문헌, Sadri and Hornsby (1996) Nucl. Acids Res. 24:5058-5059; and Xiong and Laird (1997) Nucl. Acids Res. 25:2532-2534). PCR 기술은 하기를 위해 개발되었다: 유전자 돌연변이의 검출 (참고 문헌, Kuppuswamy 등 (1991) Proc. Natl. Acad. Sci. USA 88:1143-1147) 및 대립유전자-특이적 발현의 정량화 (Szabo and Mann (1995) Genes Dev. 9:3097-3108; and Singer-Sam 등 (1992) PCR Methods Appl.1:160-163). 이러한 기술은 PCR-생성된 템플레이트에 어닐링하고 분석되는 단일 뉴클레오타이드의 5'를 즉시 종결시키는, 내부 프라이머를 사용한다. 하기 미국특허에서 기재된 바와 같은 “정량적 Ms-SnUPe 검정”을 사용하는 방법특허번호 제7,037,650호는 일부 구현예에서

사용된다.

[0018] 메틸화 상태를 평가함에 있어, 본 메틸화 상태는 종종 특정 부위를 포함하는 샘플 내 DNA의 총 모집단에 대해 그 특정부위에서 (예를 들어, 단일 뉴클레오타이드에서, 특정 영역 또는 유전자좌에서, 예를 들어, 최대 DNA의 ~100-bp, 200-bp, 500-bp, 1000-bp 하위서열 또는 더 긴 것과 같은, 더 긴 관심 서열에서) 메틸화되는 DNA의 개별 가닥의 분율 또는 백분율로서 표현된다. 종래에, 비메틸화된 핵산의 양은 캘리브레이터를 사용하여 PCR에 의해 결정된다. 그런 다음, 공지된 양의 DNA가 바이선폴라이트 처리되고, 수득한 메틸화-특이적인 서열은 실시간 PCR 또는 다른 지수 증폭, 예를 들어 QuARTS 검정을 사용하여 결정된다 (예를 들어, 하기 미국특허에 제공된 바와 같음특허번호 8,361,720, 8,715,937, 및 8,916,344).

[0019] 예를 들어, 일부 구현예에서 방법은 외부 표준을 사용함에 의해 비메틸화된 표적에 대한 표준 곡선을 생성하는 것을 포함한다. 본 표준곡선은 적어도 두 지점으로부터 구성되어 있으며 비메틸화된 DNA에 대한 실시간 Ct 값을 공지된 정량적 표준과 관련시킨다. 그런 다음, 메틸화된 표적에 대한 제2 표준곡선은 적어도 두 지점과 외부 표준으로부터 구성된다. 이 제2 표준곡선은 메틸화된 DNA에 대한 Ct를 공지된 정량적 표준과 관련시킨다. 다음으로, 시험 샘플 Ct 값이 메틸화된 및 비메틸화된 모집단에 대해 결정되고 DNA의 게놈 등가물은 처음 두 단계에 의해 생성된 표준 곡선으로부터 계산된다. 관심대상 부위에서 메틸화의 백분율은 모집단에서 DNA의 총량에 대해 메틸화된 DNA의 양, 예를 들어, (메틸화된 DNA의 수) / (메틸화된 DNA의 수 + 비메틸화된 DNA의 수) × 100로부터 계산된다.

[0020] 또한 본 방법을 실시하기 위한 조성물 및 키트가 본 명세서에 제공된다. 예를 들어, 일부 구현예에서, 하나 이상의 마커에 대해 특이적인 시약 (예를 들어, 프라이머, 프로브)이 단독으로 또는 세트 (예를 들어, 복수의 마커를 증폭시키기 위한 프라이머 쌍의 세트)로 제공된다. 검출 검정을 수행하기 위한 추가의 시약 (예를 들어, QuARTS, PCR, 서열분석, 바이선폴라이트, 또는 다른 검정을 수행하기 위한 효소, 완충액, 양성 및 음성 대조군)이 또한 제공될 수 있다. 일부 구현예에서, 방법을 수행하기 위해 필요하거나, 충분하거나, 또는 유용한 하나 이상의 시약을 함유하는 키트가 제공된다. 또한 시약을 함유하는 반응 혼합물이 제공된다. 반응 혼합물을 완성하기 위해 서로에 대해서 및/또는 시험 샘플에 첨가될 수 있는 복수의 시약을 함유하는 마스터 믹스 시약 세트가 추가로 제공된다.

[0021] 일부 구현예에서, 본 명세서에서 기재된 기술은 본 명세서에 기재된 방법에 의해 제공되는 바와 같은 일련의 산술 또는 논리적 연산을 수행하도록 설계된 프로그램 가능한 기계와 관련된다. 예를 들어, 본 기술의 일부 구현예는 컴퓨터 소프트웨어 및/또는 컴퓨터 하드웨어와 관련된다 (예를 들어, 구현된다). 일 양태에서, 본 기술은 데이터를 읽고, 조작하고, 저장하는 메모리의 형태, 산술 및 논리적 연산을 수행하기 위한 요소, 및 일련의 명령 (예를 들어, 본 명세서에 제공된 바와 같은 방법)을 실행하기 위한 처리 요소 (예를 들어, 마이크로프로세서)를 포함하는 컴퓨터에 관한 것이다. 일부 구현예에서, 마이크로프로세서는 (예를 들어, 하나 이상의 DMR, 예를 들어, 표 1 및 4에 제공된 바와 같은 DMR 1-400의) 메틸화 상태를 결정하는 것; (예를 들어, 하나 이상의 DMR, 예를 들어, 표 1 및 4에 제공된 바와 같은 DMR 1-400의) 메틸화 상태를 비교하는 것; 표준 곡선을 생성하는 것; Ct 값을 결정하는 것; (예를 들어, 하나 이상의 DMR, 예를 들어, 표 1 및 4에 제공된 바와 같은 DMR 1-400의) 메틸화의 분획, 빈도, 또는 백분율을 계산하는 것; CpG 섬을 확인하는 것; 검정 또는 마커의 특이성 및/또는 민감성을 결정하는 것; ROC 곡선 및 관련된 AUC를 계산하는 것; 서열분석을 위한 시스템의 일부 이고; 모두 본 명세서에서 기재된 바와 같거나 또는 당해 기술에 공지되어 있다.

[0022] 일부 구현예에서, 마이크로프로세서 또는 컴퓨터는 알고리즘에서 메틸화 상태 데이터를 사용하여 암의 부위를 예측한다.

[0023] 일부 구현예에서, 소프트웨어 또는 하드웨어 구성 요소는 다중 검정의 결과를 수신하고 다중 검정의 결과에 기초하여 암 위험을 나타내는 사용자에게 보고하기 위한 단일 값 결과를 결정한다 (예를 들어, 다중 DMR, 예를 들어 표 1 및 4에 제공된 바와 같은 DMR의 메틸화 상태를 결정함). 관련된 구현예는, 예를 들어 다중 마커 (예컨대 다중 DMR, 예를 들어 표 1 및 4에 제공된 바와 같은 DMR)의 메틸화 상태를 결정하는, 다수의 검정으로부터의 결과의 수학적 조합 (예를 들어, 가중된 조합, 선형 조합)에 기초하여 위험 인자를 계산한다. 일부 구현예에서, DMR의 메틸화 상태는 차원을 정의하고 다차원 공간에서 값을 가질 수 있고, 다중 DMR의 메틸화 상태에 의해 정의된 좌표는 예를 들어 사용자에게 보고하는 결과이다.

[0024] 일부 구현예는 저장 매체 및 메모리 구성요소를 포함한다. 메모리 구성 요소 (예를 들어, 휘발성 및/또는 비휘발성 메모리)는 명령 (예를 들어, 본 명세서에서 제공된 바와 같은 공정의 구현예) 및/또는 데이터 (예를 들어, 작업물 예컨대 메틸화 측정, 서열, 및 그와 관련된 통계적인 설명)를 저장하는데 사용된다. 일부 구현예는 하나

이상의 CPU, 그래픽 카드 및 (예를 들어, 디스플레이와 같은 출력 장치 및 키보드와 같은 입력 장치를 포함하는) 사용자 인터페이스를 또한 포함하는 시스템에 관한 것이다.

- [0025] 본 기술과 관련된 프로그래밍 가능한 기계는 전통적으로 현존하는 기술 및 개발중인 기술 또는 아직 개발되지 않은 기술 (예를 들어, 양자 컴퓨터, 화학적 컴퓨터, DNA 컴퓨터, 광학 컴퓨터, 스핀트로닉스 기반 컴퓨터 등) 을 포함한다.
- [0026] 일부 구현예에서, 본 기술은 데이터를 전송하기 위한 유선 (예를 들어, 금속 케이블, 광섬유) 또는 무선 전송 매체를 포함한다. 예를 들어, 일부 구현예는 네트워크 (예를 들어, 근거리 통신망 (LAN), 광역 통신망 (WAN), ad-hoc 네트워크, 인터넷 등)를 통한 데이터 전송에 관한 것이다. 일부 구현예에서, 프로그래밍 가능한 기계는 피어와 같은 그러한 네트워크 상에 존재하고, 일부 구현예에서는 본 프로그래밍 가능한 기계는 클라이언트/서버 관계를 갖는다.
- [0027] 일부 구현예에서, 데이터는 하드 디스크, 플래시 메모리, 광 매체, 플로피 디스크 등과 같은 컴퓨터-판독가능한 저장 매체 상에 저장된다.
- [0028] 일부 구현예에서, 본 명세서에서 제공된 기술은 본 명세서에서 기재된 바와 같은 방법을 수행하기 위해 함께 동작하는 복수의 프로그래밍 가능한 디바이스들과 관련된다. 예를 들어, 일부 구현예에서, (예를 들어, 네트워크에 의해 연결된) 복수의 컴퓨터는, 예를 들어, 클러스터 컴퓨팅 또는 그리드 컴퓨팅, 또는 전통적 네트워크 인터페이스, 예컨대 이더넷, 광섬유에 의해, 또는 무선 네트워크 기술에 의해 네트워크 (개인, 공용 또는 인터넷)에 연결된 (온보드 CPU, 스토리지, 전원 공급 장치, 네트워크 인터페이스 등을 갖는) 완전한 컴퓨터에 의존하는 일부 다른 분산 컴퓨터 구조의 실행에서 데이터를 수집하고 처리하기 위해 병렬로 동작할 수 있다.
- [0029] 예를 들어, 일부 구현예는 컴퓨터-판독가능 매체를 포함하는 컴퓨터를 제공한다. 구현예는 프로세서에 커플링된 랜덤 액세스 메모리 (RAM)를 포함한다. 프로세서는 메모리에 저장된 컴퓨터-실행가능 프로그램 명령을 실행한다. 이러한 프로세서는 마이크로프로세서, ASIC, 상태 기계 또는 다른 프로세서를 포함할 수 있으며, 캘리포니아주 산타 클라라 소재의 인텔사 및 일리노이주 샤움버그 소재의 모토로라사로부터의 프로세서와 같은 다수의 컴퓨터 프로세서 중 임의의 것일 수 있다. 이러한 프로세서는, 예를 들어, 프로세서에 의해 실행될 때, 프로세서로 하여금 본 명세서에서 기재된 단계들을 수행하게 하는 명령을 저장하는, 컴퓨터-판독가능 매체와 같은 매체를 포함하거나 또는 그 매체와 통신할 수 있다.
- [0030] 컴퓨터-판독가능 매체의 구현예는, 비제한적으로, 프로세서에 컴퓨터-판독가능 명령을 제공할 수 있는 전자, 광학, 자기 또는 다른 저장 장치 또는 전송 장치를 포함한다. 적합한 매체의 다른 예는, 비제한적으로, 플로피 디스크, CD-ROM, DVD, 자기 디스크, 메모리 칩, ROM, RAM, ASIC, 구성된 프로세서, 모든 광 매체, 모든 자기 테이프 또는 다른 자기 매체, 또는 컴퓨터 프로세서가 명령을 판독할 수 있는 임의의 다른 매체를 포함한다. 또한, 컴퓨터-판독가능 매체의 다양한 다른 형태는 라우터, 개인 또는 공용 네트워크, 유선 및 무선 양자의 다른 전송 장치 또는 채널을 포함하여, 컴퓨터로 명령을 전송하거나 전달할 수 있다. 명령어는, 예를 들어, C, C++, C#, 비주얼 베이직, 자바, 파이썬, Perl 및 자바스크립트를 비롯한 임의의 적합한 컴퓨터-프로그래밍 언어로부터의 코드를 포함할 수 있다.
- [0031] 컴퓨터는 일부 구현예에서 네트워크에 연결된다. 컴퓨터는 또한 수 많은 외부 또는 내부 디바이스 예컨대 마우스, CD-ROM, DVD, 키보드, 디스플레이, 또는 다른 입력 또는 출력 장치를 포함할 수 있다. 컴퓨터의 예는 개인용 컴퓨터, 디지털 보조장치, 개인용 디지털 보조장치, 휴대 전화, 모바일 폰, 스마트 폰, 호출기, 디지털 태블릿, 랩톱 컴퓨터, 인터넷 가전 및 기타 프로세서-기반 장치이다. 일반적으로, 본 명세서에서 제공된 기술의 양태에 관련된 컴퓨터는 본 명세서에서 제공된 기술을 포함하는 하나 이상의 프로그램을 지원할 수 있는, Microsoft Windows, Linux, UNIX, Mac OS X 등과 같이 임의의 운영 체제에서 구동하는 임의의 유형의 프로세서-기반 플랫폼 일 수 있다. 일부 구현예는 다른 애플리케이션 프로그램 (예를 들어, 애플리케이션)을 실행하는 퍼스널 컴퓨터를 포함한다. 애플리케이션은 메모리에 포함될 수 있으며, 예를 들어 워드 프로세싱 애플리케이션, 스프레드시트 애플리케이션, 이메일 애플리케이션, 인스턴트 메신저 애플리케이션, 프리젠테이션 애플리케이션, 인터넷 브라우저 애플리케이션, 캘린더/오거나이저 애플리케이션 및 클라이언트 디바이스에 의해 실행될 수 있는 임의의 다른 애플리케이션을 포함할 수 있다.
- [0032] 본 기술과 관련된 것으로 본 명세서에 기재된 모든 이러한 구성 요소, 컴퓨터 및 시스템은 논리적 또는 가상적 일 수 있다.
- [0033] 대상체로부터 획득된 샘플에서 HCC를 스크리닝하는 방법에 관한 기술이 본 명세서에서 제공되고, 상기 방법은

대상체로부터 수득된 샘플에서 마커의 메틸화 상태를 분석하는 단계; 및 마커의 메틸화 상태가 HCC를 보유하지 않은 대상체 (예를 들어, HCC를 갖지 않은 대상체) (예를 들어, HCC를 갖지 않지만, 간경변증을 갖는 대상체)에서 분석된 마커의 메틸화 상태와 상이한 경우 상기 대상체는 HCC를 갖는 것으로 동정하는 단계를 포함하고, 여기서 본 마커는 표 1 및 4에서 제공된 바와 같은 ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFNB2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRNI, 및 ECE1로부터 선택된 차별적으로 메틸화된 영역 (DMR)에 하나 이상의 염기를 포함한다.

[0034] 본 기술은 평가된 메틸화 상태에 제한되지 않는다. 일부 구현예에서 샘플 중의 마커의 메틸화 상태를 평가하는 단계는 하나의 염기의 메틸화 상태를 결정하는 단계를 포함한다. 일부 구현예에서, 샘플 중의 마커의 메틸화 상태를 평가하는 단계는 복수의 염기에서 메틸화의 정도를 결정하는 단계를 포함한다. 또한, 일부 구현예에서 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 증가된 메틸화를 포함한다. 일부 구현예에서, 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 감소된 메틸화를 포함한다. 일부 구현예에서 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 메틸화의 상이한 패턴을 포함한다.

[0035] 게다가, 일부 구현예에서 마커는 100개 이하의 염기 영역이고, 마커는 500개 이하의 염기 영역이고, 마커는 1000개 이하의 염기 영역이며, 마커는 5000개 이하의 염기 영역이거나, 또는 일부 구현예에서, 마커는 하나의 염기이다. 일부 구현예에서 마커는 높은 CpG 밀도 프로모터에 있다.

[0036] 본 기술은 샘플 유형에 제한되지 않는다. 예를 들어, 일부 구현예에서 샘플은 혈액 샘플 (예를 들어, 혈장, 혈청, 전혈), 대변 샘플, 조직 샘플 (예를 들어, 위 조직, 췌장 조직, 담관 / 간 조직, 췌장액, 및 결장직장 조직), 적출물, 또는 소변 샘플이다.

[0037] 게다가, 본 기술은 메틸화 상태를 결정하는 데 사용되는 방법에 제한되지 않는다. 일부 구현예에서 분석은 메틸화 특이적 중합효소 연쇄반응, 핵산 서열분석, 질량분광분석법, 메틸화 특이적 뉴클레아제, 질량-기반 분리, 또는 표적 포착을 사용하는 것을 포함한다. 일부 구현예에서, 분석은 메틸화 특이적 올리고뉴클레오타이드를 사용하는 것을 포함한다. 일부 구현예에서, 본 기술은 메틸화 상태를 결정하기 위한 대규모 병렬 서열분석 (예를 들어, 차세대 서열분석), 예를 들어 합성에 의한 서열분석, 실시간 (예를 들어, 단일 분자) 서열분석, 비드 에멀전 서열분석, 나노포어 서열분석 등을 사용한다.

[0038] 본 기술은 DMR을 검출하기 위한 시약을 제공하며, 예를 들어 일부 구현예에서 하기 서열번호에 제공된 서열을 포함하는 올리고뉴클레오타이드의 세트가 제공된다: 1-94 (표 2 및 5). 일부 구현예에서 DMR에서 염기를 갖는 염색체 영역에 상보적인 서열을 포함하는 올리고뉴클레오타이드, 예를 들어 DMR의 메틸화 상태에 민감한 올리고뉴클레오타이드가 제공된다.

[0039] 본 기술은 마커의 다양한 패널을 제공하고, 예를 들어, 일부 구현예에서 본 마커는 표 1 또는 3에 제공되고 마커를 포함하는 주석을 갖는 염색체 영역을 포함한다. 또한, 구현예는 DMR 번호 1-400 중 하나 이상인 표 1 및/또는 4로부터의 DMR을 분석하는 방법을 제공한다.

[0040] 키트 구현예, 예를 들어, 바이설파이트 시약; 및 DMR 1-400 (표 1 및 4로부터의 것)으로 구성된 군으로부터 선택된 DMR로부터의 서열을 포함하고 HCC를 갖지 않는 대상체 (예를 들어, HCC를 갖지 않고 간경변증을 갖지 않는 대상체) (예를 들어, HCC를 갖지 않지만 간경변증을 갖는 대상체)와 관련된 메틸화 상태를 갖는 대조 핵산을 포함하는 키트가 제공된다. 키트 구현예, 예를 들어, 바이설파이트 시약; 및 DMR 1-400 (표 1 및 4로부터의 것)으로 구성된 군으로부터 선택된 DMR로부터의 서열을 포함하고 HCC를 갖지 않는 대상체와 관련된 메틸화 상태를 갖는 대조 핵산을 포함하는 키트가 제공된다.

[0041] 일부 키트 구현예는 대상체로부터의 샘플 (예를 들어, 대변 샘플)을 얻기 위한 샘플 콜렉터; 상기 샘플로부터 핵산을 분리하기 위한 시약; 바이설파이트 시약; 및 본 명세서에서 기재된 바와 같은 올리고뉴클레오타이드를 포함한다.

[0042] 본 기술은 조성물 (예를 들어, 반응 혼합물)의 구현예에 관한 것이다. 일부 구현예에서 DMR을 포함하는 핵산 및 바이설파이트 시약을 포함하는 조성물이 제공된다. 일부 구현예는 DMR을 포함하는 핵산 및 본 명세서에서 기재된 바와 같은 올리고뉴클레오타이드를 포함하는 조성물을 제공한다. 일부 구현예는 DMR을 포함하는 핵산 및 메틸화-민감성 제한효소를 포함하는 조성물을 제공한다. 일부 구현예는 DMR을 포함하는 핵산 및 중합효소를 포함하는 조성물을 제공한다.

[0043] 대상체로부터 수득된 샘플 (예를 들어, 혈장 샘플)에서 HCC를 스크리닝하기 위한 추가의 관련된 방법 구현예가 제공되고, 예를 들어, 본 방법은 DMR 1-400 (표 1 및 4로부터의 것) 중 하나 이상인 DMR 내에 염기를 포함하는



샘플에서 마커의 메틸화 상태를 결정하는 단계; 상기 대상체 샘플로부터 마커의 메틸화 상태를 HCC를 갖지 않는 대상체로부터의 정상 대조군 샘플로부터 마커의 메틸화상태와 비교하는 단계; 및 상기 대상체 샘플과 정상 대조군 샘플의 메틸화 상태에서 차이의 신뢰 구간 및/또는 p 값을 결정하는 단계를 포함한다.

[0044] 일부 구현예에서, 신뢰구간은 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.9% 또는 99.99%이고 p 값은 0.1, 0.05, 0.025, 0.02, 0.01, 0.005, 0.001, 또는 0.0001이다. 방법의 일부 구현예는 DMR을 포함하는 핵산을 바이설펜아이트 시약과 반응시켜 바이설펜아이트-반응된 핵산을 생성하는 단계; 상기 바이설펜아이트-반응된 핵산을 서열분석하여 바이설펜아이트-반응된 핵산의 뉴클레오타이드 서열을 제공하는 단계; 바이설펜아이트-반응된 핵산의 뉴클레오타이드 서열을 암을 갖지 않는 대상체로부터의 DMR을 포함하는 핵산의 뉴클레오타이드 서열과 비교하여 두 서열의 차이를 확인하는 단계; 및 차이가 존재할 때 상기 대상체가 신생물을 갖는 것으로 동정하는 단계를 제공한다.

[0045] 대상체로부터 수득된 샘플에서 HCC를 스크리닝하기 위한 시스템이 본 기술에 의해 제공된다. 시스템의 예시적인 구현예는, 예를 들어, 대상체로부터 수득된 샘플에서 HCC를 스크리닝하기 위한 시스템을 포함하고, 상기 시스템은 샘플의 메틸화 상태를 결정하도록 구성된 분석 구성요소, 샘플의 메틸화 상태를 대조군 샘플 또는 데이터베이스 내에 기록된 참조 샘플 메틸화 상태와 비교하도록 구성된 소프트웨어 구성요소 및 HCC-관련된 메틸화 상태 (예를 들어, 무 HCC에 대한 메틸화 상태; HCC에 대한 메틸화 상태)를 사용자에게 알리도록 구성된 알림 구성요소를 포함한다. 알림은 일부 구현예에서 다중 검정 (예를 들어, 표 1 및 표 4에 제공된 바와 같은, 예를 들어 DMR인 다중 마커의 메틸화 상태를 결정함)으로부터 결과를 수신하는 소프트웨어 구성요소 및 다중 결과에 기초하여 보고하기 위한 값 또는 결과를 계산함에 의해 결정된다. 일부 구현예는 사용자 (예를 들어, 예컨대 의사, 간호사, 임상의 등)에게 보고하기 위한 값 또는 결과 및/또는 알림을 계산하는데 사용하기 위해 본 명세서에 제공된 각각의 DMR과 관련된 가중 매개변수의 데이터베이스를 제공한다. 일부 구현예에서 다중 검정으로부터의 모든 결과가 보고되고, 일부 구현예에서 하나 이상의 결과는 대상체에서의 HCC 위험을 나타내는 다중 검정으로부터의 하나 이상의 결과의 복합체에 기초하여 스코어, 값 또는 결과를 제공하기 위해 사용된다.

[0046] 시스템의 일부 구현예에서, 샘플은 DMR을 포함하는 핵산을 포함한다. 일부 구현예에서 본 시스템은 핵산을 분리하기 위한 구성요소, 혈장 샘플을 수집하기 위한 구성요소와 같은 샘플을 수집하기 위한 구성요소를 추가로 포함한다. 일부 구현예에서, 본 시스템은 DMR을 포함하는 핵산 서열을 포함한다. 일부 구현예에서 데이터베이스는 HCC를 갖지 않는 대상체로부터의 핵산서열을 포함한다. 핵산, 예를 들어, 핵산 세트가 또한 제공되고, 각각의 핵산은 DMR을 포함하는 서열을 갖는다. 일부 구현예에서 각각의 핵산이 HCC를 갖지 않는 대상체로부터의 서열을 갖는 핵산 세트. 관련된 시스템 구현예는 기재된 바와 같은 핵산 세트 및 핵산 세트와 관련된 핵산 서열의 데이터베이스를 포함한다. 일부 구현예 바이설펜아이트 시약을 추가로 포함한다. 그리고, 일부 구현예는 핵산 시퀀서를 추가로 포함한다.

[0047] HCC를 검출하는 방법인, 특정 구현예에서 대상체로부터 수득된 샘플 (예를 들어, 혈장샘플)에서 HCC를 검출하는 방법이 제공되고, 본 방법은 하기 단계를 포함한다: a) 대상체로부터 DNA를 포함하는 샘플을 수득하는 단계; b) 상기 수득된 DNA를 수득된 DNA 내 비메틸화된 시토신 잔기를 선택적으로 변형하지만 메틸화된 시토신 잔기는 변형하지 않는 시약으로 처리하여 변형된 잔기를 생성하는 단계; c) 단계 b)의 처리를 당한 DNA 내의 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계로, 여기서 하나 이상의 DNA 메틸화 마커는 다음에 의해 제공된 바와 같은 차별적으로 메틸화된 영역 (DMR)에 염기를 포함하는, 단계 :DMR 1-400 (표 1 및 4로부터의 것), d) 상기 하나 이상의 DNA 메틸화 마커의 결정된 메틸화 수준을, 대상체에 대한 하나 이상의 DNA 메틸화 마커에 대해 메틸화 수준 참조와 비교하는 단계로, 상기 대상체는 하기를 갖지 않는, 단계:HCC, 차이가 존재할 때HCC를 갖는 것으로 상기 대상체를 동정하는 단계 e)

[0048] 일부 구현예에서, 하나 이상의 DNA 메틸화 마커에서의 상승된 메틸화의 결정은 CpG 섬 및 CpG 섬 해안으로 구성된 군에서 선택된 영역 내에서 변형된 메틸화의 결정을 포함한다.

[0049] 일부 구현예에서, CpG 섬 또는 CpG 해안 내의 상승된 메틸화의 결정은 DNA 메틸화 마커의 코딩 영역 또는 조절 영역 내에서의 상승된 메틸화를 포함한다.

[0050] 일부 구현예에서, 단계 b)의 처리를 받은 DNA에서 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계는 하나 이상의 DNA 메틸화 마커의 메틸화 점수 및/또는 메틸화 빈도를 결정하는 단계를 포함한다. 일부 구현예에서, 단계 b)의 처리는 수득된 DNA의 바이설펜아이트 변형을 통해 달성된다.

[0051] 일부 구현예에서, 단계 b)의 처리를 받은 DNA에서 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계는 메틸화-특이적 PCR, 정량적 메틸화-특이적 PCR, 메틸화 민감성 DNA 제한효소 분석, 정량적 바이설펜아이트 파

이로서열분석, 및 바이셀과이트 게놈 서열분석 PCR로 구성된 군으로부터 선택된 기술에 의해 달성된다.

[0052] 추가의 구현에는 본 명세서에 함유된 교시에 기초하여 당업자에게 분명할 것이다.

### 도면의 간단한 설명

[0053] 도 1: HCC DNA 샘플을 정상 대조군 (예를 들어, 간경변증이 있거나 없는 비-HCC 개체)에서 유래된 DNA와 식별하기 위한 311개 차별적으로 메틸화된 영역 (DMR)의 확인에 관여된 4가지 단계의 묘사.

도 2: 상보적 3-마커 조합 (EMX1, LRRC4 및 BDH1)은 혈장에서 20/21 HCC 및 32/33 대조군을 확인하였고; 1개 HCC는 낮은 수준의 BDH1을 가졌고 1개 대조군은 상승된 LRRC4를 가졌다. 이 패널은 97% 특이성 (95% CI, 82% -100%)에서 HCC에 대해 95% 민감성 (95% CI, 74% -100%)이었고 0.98의 AUC를 달성했다 (도 2 참조).

도 3A-I: ACP1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, HOXA1, LRRC4, SPINT2, 및 TSPYL5에 대한 리시버 동작 특징적인 곡선 정보 하의 면적.

도 4A-CC: 생물학적 조직 검증 데이터로부터의 27개 위암 마커 (첨가된 맞춤점 분석으로 29개)의 박스 플롯 (로그 눈금). 샘플은 가장 왼쪽에 정상적인 간을 배열하고 이어서 간경변증이 없는 HCC, 간경변증이 있는 HCC, 및 간경변증 대조군 (염증성)이 따른다. 수직 축은 분수의 메틸화 ( $\beta$ -액틴 가닥으로 정규화됨)이다.

도 5: 95% 정상 특이성에서 매트릭스 형식으로 29개 대조군 (16개 간경변증, 13개 정상 간)과 75개 HCC 조직 샘플에서 27개 HCC 암 마커의 수행. 마커는 수직으로 열거되고 및 수평으로 샘플링된다. 샘플은 가장 왼쪽에 정상적인 간 (NI)을 배열하고 이어서 간경변증이 없는 HCC (HN), 간경변증이 있는 HCC (HC), 및 간경변증 대조군 (In)이 따른다. 양성 히트는 밝은 회색이고 짙은 회색에는 없다. 이 플롯은 마커를 상보적 방식으로 평가될 수 있게 한다. 주석: 2개 마커인, TBX15 및 EGR2가 qMSP 데이터에 대한 맞춤 점 방법을 사용하여 두 번 분석되었으며 여기에 포함된다.

도 6은 QuART (정량적 대립유전자-특이적 실시간 표적 및 신호 증폭) 검정에 의한 메틸화된 DNA 서명의 검출에 사용되는 FRET 카세트에 대한 올리고뉴클레오타이드 서열을 제공한다. 각각의 FRET 서열은 3개의 별개의 검정으로 함께 다중화될 수 있는 형광단 및 소광인자를 포함한다.

도 7A-D: A) 최상부 3개 마커 조합 (EMX1, BDH1, LRRC4)에 대한 컷-오프가 있는 rPart 2원 트리. 이 트리는 하기 과정에 의해 작성된다: 먼저 데이터를 두 개의 그룹으로 가장 잘 분할하는 단일 변수 (마커)가 발견된다. 본 데이터는 분리되고, 그 다음 이 과정은 각각의 하위 그룹에 별도로 적용되며 하위 그룹이 최소 크기에 도달하거나 더 이상의 개선이 이루어질 수 없을 때까지 재귀적으로 반복 적용된다. 컷-오프를 충족시키거나 충족시키지 않는 대조군 샘플 수는 분자 위치와 분모 위치에서 사례 샘플에 표시된다. 여기서, 조합은 혈장에서 21개 HCC 중 20개 및 33개 대조군 중 32개를 확인했다. B) 최상부 3개 마커 조합 (EMX1, DAB2IP, TSPYL5)에 대한 컷-오프가 있는 rPart 2원 트리. C) 최상부 3개 마커 조합 (EMX1, HOXA1, ACP1)에 대한 컷-오프가 있는 rPart 2원 트리. D) 최상부 3개 마커 조합 (EMX1, EFNB2, SPINT2)에 대한 컷-오프가 있는 rPart 2원 트리.

도 8: 메틸화 마커 EMX1을 사용하여 100% 특이성에서 단계 A-C에 의해 혈장에서 HCC 민감성을 입증하는 막대형 차트. U로 분류된 샘플은 단계에 대해 정의되지 않았다.

도 9: 이 분석에서 고려된 각각의 메틸화 마커의 상대적 중요성을 묘사하는 그래프 (ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFNB2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRN1, 및 ECE1). 전체 마커의 패널에 대한 민감성과 특이성의 교차 입증된 추정치는 각각 75%와 96 %였다.

### 발명을 실시하기 위한 구체적인 내용

[0054] 간세포 암종 스크리닝에 대한 기술과 특히, 비배타적으로, 간세포 암종의 존재를 검출하기 위한 방법, 조성물, 및 관련된 용도가 본 명세서에 제공된다.

[0055] 본 기술이 본 명세서에서 기술됨에 따라, 사용된 부문 표제는 단지 조직상의 목적을 위한 것이며 요지를 어떤 식으로든 제한하는 것으로 해석되어서는 안된다.

[0056] 다양한 구현예의 이 상세한 설명에서, 설명을 위해, 개시된 구현예의 완전한 이해를 제공하기 위해 다수의 특정 세부 사항들이 제시된다. 당해 분야의 숙련가는, 그러나 이들 다양한 구현예가 이들 특정 세부 사항을 갖거나 갖지 않고 실시될 수 있다는 것을 인정할 것이다. 다른 사례에서, 구조 및 장치는 블록 선도 형태로 도시되어 있다. 게다가, 당해 분야의 숙련가는 방법이 제시되고 수행되는 특정한 시퀀스가 예시적인 것이라라는 것과, 본

시퀀스는 변경될 수 있으며 본 명세서에서 개시된 다양한 구현예의 사상 및 범위 내에 여전히 있다는 것이 고려된다는 것을 쉽게 인정할 수 있다.

[0057] 비제한적으로, 특허, 특허 출원, 기사, 서적, 논문 및 인터넷 웹 페이지를 포함하여 본원에 인용된 모든 문헌 및 유사한 자료는 어떠한 목적으로도 전체적으로 참고로 명시적으로 편입된다. 달리 정의되지 않는 한, 본 명세서에서 사용되는 모든 기술적 및 과학적 용어는 본 명세서에 기재된 다양한 구현예가 속하는 당해 분야의 숙련가에 의해 통상적으로 이해되는 것과 동일한 의미를 갖는다. 편입된 참조에서의 용어의 정의가 본 교시에서 제공된 정의와 다른 것으로 나타나는 경우, 본 교시에 제공된 정의가 지배하여야 한다.

[0058] **정의**

[0059] 본 기술의 이해를 용이하게 하기 위해, 다수의 용어 및 어구가 이하에서 정의된다. 추가의 정의 상세한 설명 전반에 걸쳐 제시된다.

[0060] 명세서 및 청구범위 전반에 걸쳐서, 하기 용어는 문맥에 명백하게 달리 지시되어 있지 않는 한, 본 명세서에서 명백하게 관련된 의미를 갖는다. 본 명세서에서 사용된 바와 같은 어구 “일 구현예에서”는 그럴 수도 있지만, 반드시 동일한 구현예를 지칭하는 것은 아니다. 게다가, 본 명세서에서 사용된 바와 같은 어구 “또 다른 구현예에서”는 그럴 수도 있지만, 반드시 상이한 구현예를 지칭하는 것은 아니다. 따라서, 아래에 기재된 바와 같이, 다양한 본 발명의 구현예들은 본 발명의 범위 또는 사상을 벗어나지 않고 용이하게 조합될 수 있다.

[0061] 또한, 본 명세서에서 사용된 바와 같이, 용어 “또는”은 포괄적인 “또는” 연산자이며, 문맥에 달리 분명하게 지시되지 않는 한 용어 “및/또는”과 동등하다. 용어 “기반한”은 배타적인 것이 아니며 문맥상 명확하게 달리 지시하지 않는 한 기재되지 않은 추가의 인자에 기반할 수 있다. 또한, 명세서 전반에 걸쳐, “a”, “an”, 및 “the”의 의미는 복수의 참조를 포함한다. “내의”의 의미는 “내의”와 “상의”를 포함한다.

[0062] 본 명세서에서 사용된 바와 같이, “핵산” 또는 “핵산 분자”는 일반적으로 비변형되거나 또는 변형된 DNA 또는 RNA일 수 있는 임의의 리보핵산 또는 데옥시리보핵산을 지칭한다. “핵산”은, 비제한적으로, 단일- 및 이중-가닥 핵산을 포함한다. 본 명세서에서 사용된 바와 같이, 용어 “핵산”은 또한 하나 이상의 변형된 염기를 함유하는 상기에 기재된 바와 같은 DNA를 포함한다. 따라서, 안정성 또는 다른 이유로 변형된 백본을 갖는 DNA는 “핵산”이다. 본 명세서에서 사용된 바와 같이 용어 “핵산”은 그와 같이 화학적으로, 효소적으로, 또는 대사적으로 변형된 형태의 핵산뿐만 아니라, 예를 들어, 단순 및 복합체 세포를 비롯한 세포와 바이러스의 DNA 특징의 화학적 형태를 포용한다.

[0063] 용어들 “올리고뉴클레오타이드” 또는 “폴리뉴클레오타이드” 또는 “뉴클레오타이드” 또는 “핵산”은 2종 이상, 바람직하게는 3개 초과, 그리고 일반적으로 10개 초과,의 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드를 갖는 분자를 지칭한다. 정확한 크기는 올리고뉴클레오타이드의 궁극적인 기능 또는 용도에 차례로 좌우되는 많은 요소에 따라 달라질 것이다. 올리고뉴클레오타이드는 화학적 합성, DNA 복제, 역전사, 또는 이들의 조합을 포함하는 임의의 방식으로 생성될 수 있다. DNA에 대한 전형적인 데옥시리보뉴클레오타이드는 티민, 아데닌, 시토신, 및 구아닌이다. RNA에 대한 전형적인 리보뉴클레오타이드는 우라실, 아데닌, 시토신, 및 구아닌이다.

[0064] 본 명세서에서 사용된 바와 같이, 용어들 핵산의 “유전자좌” 또는 “영역”은 핵산의 부분 영역, 예를 들어 염색체, 단일 뉴클레오타이드, CpG 섬 상의 유전자 등을 지칭한다.

[0065] 용어들 “상보적” 및 “상보성”은 염기쌍화 규칙에 의해 관련된 뉴클레오타이드 (예를 들어, 1개 뉴클레오타이드) 또는 폴리뉴클레오타이드 (예를 들어, 뉴클레오타이드의 서열)를 지칭한다. 예를 들어, 서열 5'-A-G-T-3'은 서열 3'-T-C-A-5'에 상보적이다. 상보성은 핵산의 염기의 일부만이 염기쌍화 규칙에 따라 매칭되는 “부분적인” 것일 수 있다. 또는 핵산 사이에 “완전한” 또는 “전체” 상보성이 있을 수 있다. 핵산 가닥 사이의 상보성의 정도는 핵산 가닥 사이의 하이브리드화의 효율 및 강도에 영향을 미친다. 이것은 핵산 사이의 결합에 의존하는 증폭 반응 및 검출 방법에서 특히 중요하다.

[0066] 용어 “유전자”는 RNA, 또는 그것의 폴리펩타이드나 전구체의 생산에 필요한 코딩 서열을 포함하는 핵산 (예를 들어, DNA 또는 RNA) 서열을 지칭한다. 기능적 폴리펩타이드는 폴리펩타이드의 원하는 활성 또는 기능적 특성 (예를 들어, 효소적 활성, 리간드 결합, 신호 형질도입, 등)이 유지되는 한 전장 코딩 서열에 의해 또는 코딩 서열의 임의의 부분에 의해 인코딩될 수 있다. 용어 “부분”은 유전자와 관련하여 사용될 때 그 유전자의 단편을 지칭한다. 단편은 크기에서 몇 개의 뉴클레오타이드에서 전체 유전자 서열 마이너스 1개 뉴클레오타이드까지의 범위일 수 있다. 따라서, “유전자의 적어도 일부분을 포함하는 뉴클레오타이드”는 유전자의 단편 또는 전체

유전자를 포함할 수 있다.

- [0067] 용어 “유전자”는 또한 구조 유전자의 코딩 영역을 포괄하고, 그리고, 예를 들어 어느 한 말단상에서 약 1kb의 거리에 대해, 유전자가 (예를 들어, 코딩, 조절, 구조적 및 다른 서열을 포함하는) 전장 mRNA의 길이에 상응하도록 5' 및 3' 말단 둘 모두에서 코딩 영역에 인접하여 위치하는 서열을 포함한다. 코딩 영역의 5'에 위치하고 mRNA 상에 존재하는 서열은 5' 비-번역 또는 미번역된 서열로 지칭된다. 코딩 영역의 3' 또는 하류에 위치하며 mRNA 상에 존재하는 서열은 3' 비-번역 또는 3' 미번역된 서열로 지칭된다. 용어 “유전자”는 유전자의 cDNA 및 게놈 형태 둘 모두를 포함한다. 일부 유기체 (예를 들어, 진핵 생물)에서, 유전자의 게놈 형태 또는 클론은 “인트론” 또는 “개재 영역” 또는 “개입 서열”이라고 불리는 비-코딩 서열로 차단된 코딩 영역을 함유한다. 인트론은 핵 RNA (hnRNA)로 전사되는 유전자의 분절이고; 인트론은 조절인자 예컨대 증진제를 함유할 수 있다. 인트론은 핵 또는 일차 전사체로부터 제거되거나 “스플라이싱 아웃”되고; 인트론은 따라서 메신저 RNA (mRNA) 전사체에 결여되어 있다. mRNA는 발생기 폴리펩타이드 내의 아미노산의 서열 또는 순서를 특정하기 위해 번역하는 동안 기능을 한다.
- [0068] 인트론을 함유하는 것에 부가하여, 유전자의 게놈 형태는 또한 RNA 전사체 상에 존재하는 서열의 5' 및 3' 말단 둘 모두 상에 위치한 서열을 포함할 수 있다. 이들 서열은 “측접하는” 서열 또는 영역 (이들 측접 서열은 mRNA 전사체 상에 존재하는 비-번역된 서열에 대해 5' 또는 3'에 위치됨)으로 지칭된다. 5' 측접하는 영역은 유전자의 전사를 제어하거나 영향을 주는 프로모터 및 증진제와 같은 조절 서열을 함유할 수 있다. 3' 측접하는 영역은 전사의 종결, 전사 후 절단 및 폴리아데닐화를 지시하는 서열을 함유할 수 있다.
- [0069] 용어 “대립유전자”는 유전자의 변형을 지칭하고; 본 변형은 비제한적으로 변이체 및 돌연변이체, 다형성 유전자좌, 및 단일 뉴클레오타이드 다형성 유전자좌, 프레임이동, 및 스플라이싱 돌연변이를 포함한다. 대립유전자는 모집단에서 자연적으로 발생할 수 있거나 또는 모집단의 임의의 특정 개체의 수명 동안 발생할 수 있다.
- [0070] 따라서, 용어들 “변이체” 및 “돌연변이체”는 뉴클레오타이드 서열과 관련하여 사용될 때 하나 이상의 뉴클레오타이드가 일반적으로 관련된 또 다른 뉴클레오타이드산 서열과 상이한 핵산서열을 지칭한다. “변형”은 2개의 상이한 뉴클레오타이드 서열 사이의 차이이고; 전형적으로, 하나의 서열이 참조 서열이다.
- [0071] “증폭”은 템플레이트 특이성을 포함하는 핵산 복제의 특별한 경우이다. 이것은 비-특이적 템플레이트 복제 (예를 들어, 특정 템플레이트 상에 의존적이지 않지만 템플레이트-의존적인 복제)와 대조된다. 템플레이트 특이성은 복제 (예를 들어, 적절한 폴리뉴클레오타이드 서열의 합성) 및 뉴클레오타이드 (리보- 또는 디옥시리보-) 특이성의 충실도로부터 여기서 구별된다. 템플레이트 특이성은 “표적” 특이성의 관점에서 빈번하게 기재된다. 표적 서열은 이들이 다른 핵산으로부터 분류되도록 하는 의미에서 “표적”이다. 증폭 기술은 주로 이 분류를 위해 설계되었다.
- [0072] 핵산의 증폭은 전형적으로 소량의 폴리뉴클레오타이드 (예를 들어, 단일 폴리뉴클레오타이드 분자, 10 내지 100 카피의 폴리뉴클레오타이드 분자로, 정확히 동일할 수도 있고 그렇지 않을 수도 있음)로부터 출발하여 다중 카피의 폴리뉴클레오타이드, 또는 폴리뉴클레오타이드의 부분의 생산을 일반적으로 지칭하고, 여기서 증폭 생성물 또는 앰플리콘이 일반적으로 검출가능하다. 폴리뉴클레오타이드의 증폭은 다양한 화학적 및 효소과정들을 포함한다. 하나 또는 소수 카피의 표적 또는 템플레이트 DNA 분자로부터 다수의 DNA 카피의 생성은 하기의 도중에 일어난다: 증합효소연쇄반응 (PCR) 또는 리가제연쇄반응 (LCR; 참고 특허, 예를 들어, 미국특허번호 5,494,810)으로 증폭의 형태이다. 증폭의 추가의 유형은, 비제한적으로, 하기를 포함한다: 대립유전자-특정 PCR (참고 특허, 예를 들어, 미국특허번호 5,639,611), 어셈블리 PCR (참고 특허, 예를 들어, 미국특허번호 5,965,408), 헬리카제-의존적 증폭 (참고 특허, 예를 들어, 미국특허번호 7,662,594), 핫-스타트 PCR (참고 특허, 예를 들어, 미국특허번호 5,773,258 및 5,338,671), 서열간-특이적 PCR, 역 PCR (참고 문헌, 예를 들어, Triglia, et al 등 (1988) Nucleic Acids Res., 16: 8186), 결찰-매개된 PCR (참고 문헌, 예를 들어, Guilfoyle, R. 등, Nucleic Acids Research, 25: 1854-1858 (1997); 미국특허번호 5,508,169), 메틸화-특이적 PCR (참고 문헌, 예를 들어, Herman, 등, (1996) PNAS 93(13) 9821-9826), 미니프라이머 PCR, 다중 결찰-의존적 탐침 증폭(참고 문헌, 예를 들어, Schouten, 등, (2002) Nucleic Acids Research 30(12): e57), 다중 PCR (참고 문헌, 예를 들어, Chamberlain, 등, (1988) Nucleic Acids Research 16(23) 11141-11156; Ballabio, 등, (1990) Human Genetics 84(6) 571-573; Hayden, 등, (2008) BMC Genetics 9: 80), 내포된 PCR, 오버랩-연장 PCR (참고 문헌, 예를 들어, Higuchi, 등, (1988) Nucleic Acids Research 16(15) 7351-7367), 실시간 PCR (참고 문헌, 예를 들어, Higuchi, et al 등, (1992) Biotechnology 10: 413-417; Higuchi, 등, (1993) Biotechnology 11: 1026-1030), 역전사 PCR (참고 문헌, 예를 들어, Bustin, S.A. (2000) J. Molecular Endocrinology 25: 169-193),



고상 PCR, 열적비대칭 인터레이스된 PCR, 및 터치다운 PCR (참고 문헌, 예를 들어, Don, 등, Nucleic Acids Research (1991) 19(14) 4008; Roux, K.(1994) Biotechniques 16(5) 812-814; Hecker, 등, (1996) Biotechniques 20(3) 478-485). 폴리뉴클레오타이드 증폭은 또한 디지털 PCR을 사용하여 달성될 수 있다 (참고 문헌, 예를 들어, Kalinina, 등, Nucleic Acids Research.25; 1999-2004, (1997); Vogelstein and Kinzler, Proc Natl Acad Sci USA.96; 9236-41, (1999); 국제 특허공보번호 W005023091A2; 미국 특허출원공개 번호 20070202525).

[0073] 용어 “중합효소 연쇄 반응” (“PCR”)은 하기의 방법을 지칭한다: K.B.Mullis의 미국특허번호 4,683,195, 4,683,202, 및 4,965,188, 이들은 클로닝 또는 정제 없이 게놈 DNA의 혼합물에서 표적 서열의 분절 농도를 증가시키기 위한 방법을 기술한다. 표적 서열을 증폭시키기 위한 이 공정은 원하는 표적 서열을 함유하는 DNA 혼합물에 2개의 올리고뉴클레오타이드 프라이머를 아주 과도하게 도입한 후, DNA 중합 효소의 존재에서 열적 순환의 정확한 서열이 따르는 단계로 구성된다. 2개 프라이머는 이중 가닥 표적 서열의 그 각각의 가닥에 상보적이다. 증폭을 수행하기 위해, 혼합물은 변성되고 프라이머는 그 다음 표적 분자 내에서 그것의 상보적인 서열에 어닐링된다. 어닐링에 이어, 프라이머는 새로운 쌍의 상보적 가닥을 형성하도록 중합효소로 연장된다. 변성, 프라이머 어닐링, 및 중합효소 연장의 단계는 수회 반복될 수 있어 (즉, 변성, 어닐링 및 연장은 하나의 “사이클”을 구성하여; 수많은 “사이클”이 있을 수 있음) 원하는 표적 서열의 증폭된 분절의 고농도를 얻는다. 원하는 표적 서열의 증폭된 분절의 길이는 서로에 대한 프라이머의 상대 위치에 의해 결정되며, 따라서 이 길이는 조절 가능한 매개변수이다. 본 과정의 반복하는 양태에 의해, 본 방법은 “중합효소 연쇄반응” (“PCR”)으로 지칭된다. 표적 서열의 원하는 증폭된 분절이 혼합물 내에서 우세한 서열 (농도의 관점에서)이 되기 때문에, 이들은 “PCR 증폭된” 것으로 불리고 “PCR 생성물” 또는 “앰플리콘”이다.

[0074] 템플레이트 특이성은 효소의 선택에 의해 대부분의 증폭 기술에서 달성된다. 증폭 효소는 이들이 사용되는 조건 하에서 핵산의 불균질 혼합물 내 핵산의 특정 서열만을 처리할 수 있는 효소이다. 예를 들어, Q-베타 레플리카제의 경우, MDV-1 RNA는 레플리카제에 대한 특정 템플레이트이다 (Kacian 등, Proc. Natl. Acad. Sci. USA, 69: 3038 [1972]). 다른 핵산은 이 증폭 효소에 의해 복제되지 않을 것이다. 유사하게, T7 RNA 중합효소의 경우, 이 증폭효소는 그 자체의 프로모터에 대해 엄격한 특이성을 가진다 (Chamberlin et al, Nature, 228: 227 [1970]). T4 DNA 리가제의 경우, 본 효소는 결찰 접합부에서 올리고뉴클레오타이드 또는 폴리뉴클레오타이드 기질과 템플레이트 사이에 불일치가 있는 두 올리고뉴클레오타이드 또는 폴리뉴클레오타이드를 연결하지 않을 것이다 (Wu and Wallace (1989) Genomics 4: 560). 마지막으로, 열안정성 템플레이트-의존성 DNA 중합효소 (예를 들어, Taq 및 Pfu DNA 중합효소)는 고온에서 작용하는 그것의 능력으로 인해 프라이머에 의해 결합되고 따라서 한정된 서열에 대해 높은 특이성을 나타내는 것으로 밝혀졌고; 고온은 표적 서열과의 프라이머 혼성화를 선호하고 비-표적 서열과 혼성화를 선호하지 않는 열역학적 조건을 초래한다 (H.A.Erlich (ed.), PCR Technology, Stockton Press [1989]).

[0075] 본 명세서에서 사용된 바와 같은, 용어 “핵산 검출 검정”은 관심 있는 핵산의 뉴클레오타이드 조성을 결정하는 임의의 방법을 지칭한다. 핵산 검출 검정은 비제한적으로, 하기를 포함한다: DNA 서열분석 방법, 탐침 하이브리드화 방법, 구조 특이적 절단 검정 (예를 들어, INVADER 검정, Hologic, Inc.)이고, 예를 들어, 하기에 기재되어 있다: 미국특허번호 5,846,717, 5,985,557, 5,994,069, 6,001,567, 6,090,543, 및 6,872,816; Lyamichiev 등, Nat. Biotech., 17: 292 (1999), Hall 등, PNAS, USA, 97: 8272 (2000), 및 US 2009/0253142); 효소 불일치 절단 방법 (예를 들어, Variagenics의 미국특허번호 6,110,684, 5,958,692, 5,851,770); 중합효소 연쇄반응; 분지형 하이브리드화 방법 (예를 들어, Chiron의 미국특허번호 5,849,481, 5,710,264, 5,124,246, 및 5,624,802); 회전환 복제 (예를 들어, 미국특허번호 6,210,884, 6,183,960 및 6,235,502); NASBA (예를 들어, 미국특허번호 5,409,818); 분자 향로표지 기술 (예를 들어, 미국특허번호 6,150,097); E-센서 기술 (Motorola의 미국특허번호 6,248,229, 6,221,583, 6,013,170, 및 6,063,573); 사이클링 탐침 기술 (예를 들어, 미국특허번호 5,403,711, 5,011,769, 및 5,660,988); 데이드 베어링 신호증폭 방법 (예를 들어, 미국특허번호 6,121,001, 6,110,677, 5,914,230, 5,882,867, 및 5,792,614); 리가제 연쇄반응 (예를 들어, Barnay Proc. Natl. Acad. Sci USA 88, 189-93 (1991)); 및 샌드위치 하이브리드화 방법 (예를 들어, 미국특허번호 5,288,609).

[0076] 용어 “증폭가능한 핵산”은 임의의 증폭 방법에 의해 증폭될 수 있는 핵산을 지칭한다. “증폭가능한 핵산”은 일반적으로 “샘플 템플레이트”를 포함할 것으로 고려된다.

[0077] 용어 “샘플 템플레이트”는 “표적” (하기에 정의됨)의 존재에 대해 분석되는 샘플로부터 유래된 핵산을 지칭한다. 그에 반해서, “배경 템플레이트”는 샘플에 존재하거나 존재하지 않을 수 있는 샘플 템플레이트 이외의 핵

산과 관련하여 사용된다. 배경 템플레이트는 가장 흔한 우연한 것이다. 이것은 캐리오버의 결과일 수도 있고, 또는 샘플에서 정제되기를 원하는 핵산 오염 물질의 존재 때문일 수도 있다. 예를 들어, 검출되어 지는 것 이외의 유기체로부터의 핵산은 시험 샘플에 배경으로 존재할 수 있다.

[0078] 용어 “프라이머”는 정제된 제한효소분해에서와 같이 자연적으로 발생하든 또는 합성적으로 생성되든, 핵산 가닥에 상보적인 프라이머 연장 생성물의 합성이 유도되는 조건에 놓여질 때 (예를 들어, 뉴클레오타이드 및 유도제 예컨대 DNA 중합효소의 존재에서 그리고 적합한 온도 및 pH에서) 합성의 개시 지점으로 작용할 수 있는 올리고뉴클레오타이드를 지칭한다. 프라이머는 증폭에서 최대 효율을 위해 바람직하게는 단일 가닥이지만, 대안적으로 이중 가닥일 수도 있다. 이중가닥일 경우, 프라이머는 연장 생성물을 제조하기 위해 사용되기 전에 먼저 그것의 가닥을 분리하도록 처리된다. 바람직하게는, 프라이머는 올리고데옥시리보뉴클레오타이드이다. 프라이머는 유도제의 존재에서 연장 생성물의 합성을 준비할 만큼 충분히 길어야 한다. 프라이머의 정확한 길이는 온도, 프라이머의 공급원 및 방법의 사용을 포함한 많은 인자에 따라 달라질 것이다.

[0079] 용어 “탐침”은 정제된 제한효소분해에서와 같이 자연적으로 발생하든 또는 합성 적으로, 제조합으로 또는 PCR 증폭에 의해 생성되든, 관심 있는 또 다른 올리고뉴클레오타이드에 혼성화할 수 있는 올리고뉴클레오타이드 (예를 들어, 뉴클레오타이드의 서열)를 지칭한다. 탐침은 단일-가닥 또는 이중-가닥일 수 있다. 탐침은 특정 유전자 서열 (예를 들어, “캡처 탐침”)의 검출, 확인, 및 단리에 유용하다. 본 발명에서 사용되는 임의의 탐침은 일부 구현예에서 임의의 “리포터 분자”로 표지될 수 있으므로, 비제한적으로 효소 (예를 들어, ELISA, 뿐만 아니라 효소-기재 조직화화 검정), 형광, 방사성, 및 발광성 시스템을 포함한 임의의 검출 시스템에서 검출 가능한 것으로 고려된다. 본 발명이 임의의 특정 검출 시스템 또는 표지에 제한되는 것으로 의도되지 않는다.

[0080] 본 명세서에서 사용된 바와 같이, “메틸화”는 시토신의 위치 C5 또는 N4, 아데닌의 N6 위치에서의 시토신 메틸화, 또는 다른 유형의 핵산 메틸화를 지칭한다. 시험관내 증폭된 DNA는 전형적인 시험관내 DNA 증폭 방법이 증폭 템플레이트의 메틸화 패턴을 유지하지 않기 때문에 일반적으로 비메틸화된다. 그러나, “비메틸화된 DNA” 또는 “메틸화된 DNA”는 또한 그것의 최초 템플레이트가 각각 비메틸화되거나 메틸화된 증폭 DNA를 지칭할 수 있다.

[0081] 따라서, 본 명세서에서 사용된 바와 같은 “메틸화된 뉴클레오타이드” 또는 “메틸화된 뉴클레오타이드 염기”는 뉴클레오타이드 염기 상에 메틸 모이어티의 존재를 지칭하고 여기서 본 메틸 모이어티는 인식된 전형적인 뉴클레오타이드 염기 내에 존재하지 않는다. 예를 들어, 시토신은 그것의 피리미딘 고리 상에 메틸 모이어티를 함유하지 않지만, 5-메틸시토신은 그것의 피리미딘 고리의 위치 5에 메틸 모이어티를 함유한다. 따라서, 시토신은 메틸화된 뉴클레오타이드가 아니고 5-메틸시토신은 메틸화된 뉴클레오타이드이다. 또 다른 예에서, 티민은 그것의 피리미딘 고리의 위치 5에서 메틸 모이어티를 함유한다; 그러나, 본 명세서에서 목적을 위해, 티민은 DNA의 전형적인 뉴클레오타이드 염기이기 때문에, DNA에 존재할 때 티민은 메틸화된 뉴클레오타이드로 간주되지 않는다.

[0082] 본 명세서에서 사용된 바와 같이, “메틸화된 핵산 분자”는 하나 이상의 메틸화된 뉴클레오타이드를 함유하는 핵산분자를 지칭한다.

[0083] 본 명세서에서 사용된 바와 같이, 핵산 분자의 “메틸화 상태”, “메틸화 프로파일”, 및 “메틸화 상태”는 핵산 분자에서 하나 이상의 메틸화된 뉴클레오타이드 염기의 부재의 존재를 지칭한다. 예를 들어, 메틸화된 시토신을 함유하는 핵산분자는 메틸화된 것으로 간주된다 (예를 들어, 핵산 분자의 메틸화 상태는 메틸화된다). 임의의 메틸화된 뉴클레오타이드를 함유하지 않는 핵산 분자는 비메틸화된 것으로 간주된다.

[0084] 특정 핵산 서열 (예를 들어, 본 명세서에서 기재된 바와 같은 유전자 마커 또는 DNA 영역)의 메틸화 상태는 서열 내 모든 염기의 메틸화 상태를 나타낼 수 있거나 또는 서열 내의 염기의 서브셋 (예를 들어, 하나 이상의 시토신)의 메틸화 상태를 나타낼 수 있거나, 또는 메틸화가 일어나는 서열 내의 위치의 정확한 정보를 제공하거나 제공함이 없이 서열 내의 영역의 메틸화 밀도에 관한 정보를 나타낼 수 있다.

[0085] 핵산 분자 내 뉴클레오타이드 유전자좌의 메틸화 상태는 핵산 분자 내 특정 유전자좌에서 메틸화된 뉴클레오타이드의 존재 또는 부재를 지칭한다. 예를 들어, 핵산 분자 내 7번째 뉴클레오타이드에서 시토신의 메틸화 상태는 핵산 분자 내 7번째 뉴클레오타이드에 존재하는 뉴클레오타이드가 5-메틸시토신일 때 메틸화된다. 유사하게, 핵산 분자 내 7번째 뉴클레오타이드에서 시토신의 메틸화 상태는 핵산 분자 내 7번째 뉴클레오타이드에 존재하는 뉴클레오타이드가 시토신일 때 (그리고 5-메틸시토신이 아닐 때) 비메틸화된다.

[0086] 메틸화 상태는 선택적으로 “메틸화 값” (예를 들어, 메틸화 빈도, 분획, 비, 퍼센트, 등을 나타냄)에 의해 제

시되거나 또는 표시될 수 있다 메틸화 값은 예를 들어 메틸화 의존성 제한 효소로 제한 효소분해 후 존재하는 손상되지 않은 핵산의 양을 정량하거나 바이설파이트 반응 후 증폭 프로파일을 비교하거나 또는 바이설파이트-처리 및 비처리된 핵산의 서열을 비교함으로써 생성될 수 있다. 따라서, 값, 예를 들어, 메틸화 값은 메틸화 상태를 나타내고 따라서 유전자좌의 여러 카피에 걸친 메틸화 상태의 정량적 지표로서 사용될 수 있다. 이것은 샘플 내의 서열의 메틸화 상태를 역치 또는 기준 값과 비교하는 것이 바람직할 경우에 특정 용도의 것이다.

[0087] 본 명세서에서 사용된 바와 같이, “메틸화 빈도” 또는 “메틸화 퍼센트 (%)”는 분자 또는 유전자좌가 비메틸화된 경우의 수에 비해 분자 또는 유전자좌가 메틸화된 경우의 수를 지칭한다.

[0088] 이와 같이, 메틸화 상태는 핵산 (예를 들어, 유전자 서열)의 메틸화 상태를 기술한다. 또한, 메틸화 상태는 메틸화에 관련된 특정 유전자좌위에서 핵산 분절의 특징을 지칭한다. 이러한 특징은, 비제한적으로, 이 DNA 서열 내에 임의의 시토신 (C) 잔기가 메틸화되었는지에 무관하게, 메틸화된 C 잔기(들)의 위치, 핵산의 임의의 특정 영역에 걸친 메틸화된 C의 빈도 또는 백분율, 및 예를 들어, 대립 유전자의 기원에서의 차이에 기인한 메틸화에서 대립유전자 차이를 포함한다. 용어들 “메틸화 상태”, “메틸화 프로파일”, 및 “메틸화 상태”는 또한 생물학적 샘플에서 핵산의 임의의 특정 영역 전반에 걸쳐 메틸화된 C 또는 비메틸화된 C의 상대적 농도, 절대적 농도, 또는 패턴을 지칭한다. 예를 들어, 만일 핵산서열 내의 시토신 (C) 잔기(들)가 메틸화되면, 이것은 “과메틸화된” 것으로 또는 “증가된 메틸화”를 갖는 것으로 지칭될 수 있고, 반면에 만일 DNA 서열 내의 시토신 (C) 잔기(들)가 메틸화되지 않으면 이것은 “저메틸화된” 것으로 또는 “감소된 메틸화”를 갖는 것으로 지칭될 수 있다. 마찬가지로, 만일 핵산 서열 내의 시토신 (C) 잔기(들)가 또 다른 핵산 서열 (예를 들어, 상이한 영역으로부터 또는 상이한 개체, 등으로부터의 것)에 비교하여 메틸화되면 그 서열은 다른 핵산 서열에 비교하여 과메틸화되거나 또는 증가된 메틸화를 갖는 것으로 간주된다. 대안적으로, 만일 DNA 서열 내의 시토신 (C) 잔기(들)가 또 다른 핵산 서열 (예를 들어, 상이한 영역으로부터 또는 상이한 개체, 등으로부터의 것)에 비교하여 메틸화되지 않으면 그 서열은 다른 핵산 서열에 비교하여 저메틸화되거나 또는 감소된 메틸화를 갖는 것으로 간주된다. 추가로, 본 명세서에서 사용된 바와 같은 용어 “메틸화 패턴”은 핵산이 영역 상의 메틸화된 및 비메틸화된 뉴클레오타이드의 집합적인 부위를 지칭한다. 2개의 핵산은 동일하거나 유사한 메틸화 빈도 또는 메틸화 퍼센트를 가질 수 있지만, 메틸화 및 비메틸화된 뉴클레오타이드의 수가 영역 전반에 걸쳐 동일하거나 또는 유사하지만 메틸화 및 비메틸화된 뉴클레오타이드의 위치가 상이한 경우 상이한 메틸화 패턴을 갖는다. 서열은 메틸화의 정도 (예를 들어 다른 것과 비교하여 증가된 또는 감소된 메틸화를 갖는 것) 빈도, 또는 패턴에서 다를 때 본 서열은 “차별적으로 메틸화된” 또는 “메틸화에서 차이”를 갖는 것 또는 “상이한 메틸화 상태”를 갖는 것으로 지칭된다. 용어 “차별적인 메틸화”는 암 음성 샘플에서 핵산 메틸화의 수준 또는 패턴과 비교할 때 암 양성 샘플에서 핵산 메틸화의 수준 또는 패턴에서의 차이를 지칭한다. 또한 수술 후 암의 재발이 있는 환자와 재발하지 않은 환자 간의 수준 또는 패턴에서의 차이를 지칭할 수도 있다. DNA 메틸화의 차별적인 메틸화 및 특정 수준 또는 패턴은, 예를 들어 올바른 컷-오프 또는 예측의 특징이 정의되면 예후 및 예측의 바이오마커이다.

[0089] 메틸화 상태 빈도는 개체들의 모집단 또는 단일 개체로부터 샘플을 기술하는 데 사용될 수 있다. 예를 들어, 50%의 메틸화 상태 빈도를 갖는 뉴클레오타이드 유전자좌는 50%의 사례가 메틸화되고 50%의 사례는 비메틸화된 다. 그와 같은 빈도는, 예를 들어, 뉴클레오타이드 유전자좌 또는 핵산 영역이 개체들의 모집단 또는 핵산 집합에서 메틸화되는 정도를 기술하기 위해 사용될 수 있다. 따라서, 제1 모집단 또는 핵산 분자의 풀에서의 메틸화가 제2 모집단 또는 핵산 분자의 풀에서의 메틸화와 상이한 경우, 제1 모집단 또는 풀의 메틸화 상태 빈도는 제2 모집단 또는 풀의 메틸화 상태 빈도와 상이할 것이다. 그와 같은 빈도는 또한 예를 들어, 뉴클레오타이드 유전자좌 또는 핵산 영역이 단일 개체에서 메틸화되는 정도를 기술하기 위해 사용될 수 있다. 예를 들어, 그와 같은 빈도는 조직샘플로부터의 세포 군이 뉴클레오타이드 유전자좌 또는 핵산 영역에서 메틸화되거나 또는 비메틸화되는 정도를 기술하기 위해 사용될 수 있다.

[0090] 본 명세서에서 사용된 바와 같이 “뉴클레오타이드 유전자좌”는 핵산 분자에서 뉴클레오타이드의 위치를 지칭한다. 메틸화된 뉴클레오타이드의 뉴클레오타이드 유전자좌는 핵산 분자에서 메틸화된 뉴클레오타이드의 위치를 지칭한다.

[0091] 전형적으로, 인간 DNA의 메틸화는 시토신이 구아닌의 5'에 위치한 인접한 구아닌 및 시토신을 포함한 디뉴클레오타이드 서열 (또한 일명 CpG 디뉴클레오타이드 서열) 상에서 일어난다. CpG 디뉴클레오타이드 내의 대부분의 시토신은 인간 게놈에서 메틸화되지만, 그러나 일부는 CpG 섬에서 공지된 바와 같이 특정 CpG 디뉴클레오타이드 풍부 게놈 영역에서는 비메틸화된 다 (참고 문헌, 예를 들어, Antequera 등 (1990) Cell 62: 503-514).



- [0092] 본 명세서에서 사용된 바와 같이, “CpG 섬”은 총 게놈 DNA에 비하여 증가된 수의 CpG 디뉴클레오타이드를 함유하는 게놈 DNA의 G: C-풍부 영역을 지칭한다. CpG 섬은 길이가 적어도 100, 200, 또는 그 초과 염기 쌍일 수 있고, 여기서 본 영역의 G: C 함량은 적어도 50%이고 그리고 기대된 빈도에 대한 관측된 CpG 빈도의 비는 0.6이고; 일부 사례에서, CpG 섬은 길이가 적어도 500 염기 쌍일 수 있고, 여기서 본 영역의 G: C 함량은 적어도 55%이고) 그리고 기대된 빈도에 대한 관측된 CpG 빈도의 비는 0.65이다. 기대된 빈도에 대한 관측된 CpG 빈도는 하기 문헌에서 제공된 방법에 따라 계산될 수 있다: Gardiner-Garden et al (1987) J. Mol. Biol. 196: 261-281. 예를 들어, 기대된 빈도에 대한 관측된 CpG 빈도는 식  $R = (A \times B) / (C \times D)$ 에 따라 계산될 수 있고, 여기서 R은 기대된 빈도에 대한 관측된 CpG 빈도의 비이고, A는 분석된 서열에서 CpG 디뉴클레오타이드의 수이고, B는 분석된 서열에서 뉴클레오타이드의 총수이고, C는 분석된 서열에서 C 뉴클레오타이드의 총수이고, 그리고 D는 분석된 서열에서 G 뉴클레오타이드의 총수이다. 메틸화 상태는 전형적으로 CpG 섬에서, 예를 들어, 프로모터 영역에서 결정된다. 그렇지만 인간 게놈에서의 다른 서열은 CpA 및 CpT와 같은 DNA 메틸화를 받기 쉽다는 것을 인정할 것이다 (참고 문헌, 예를 들어, Ramsahoye (2000) Proc. Natl. Acad. Sci. USA 97: 5237-5242; Salmon and Kaye (1970) Biochim.Biophys.Acta.204: 340-351; Grafstrom (1985) Nucleic Acids Res. 13: 2827-2842; Nyce (1986) Nucleic Acids Res. 14: 4353-4367; Woodcock (1987) Biochem.Biophys.Res. Commun.145: 888-894).
- [0093] 본 명세서에서 사용된 바와 같이, 핵산 분자의 메틸화 상태의 기능으로서 핵산 분자의 뉴클레오타이드를 변형시키는 시약 또는 메틸화-특이적 시약은 핵산 분자의 메틸화 상태를 반영하는 방식으로 핵산 분자의 뉴클레오타이드 서열을 변화시킬 수 있는 화합물 또는 조성물 또는 다른 제제를 지칭한다. 그와 같은 시약으로 핵산 분자를 처리하는 방법은 뉴클레오타이드 서열의 원하는 변화를 달성하기 위해, 요망하는 경우, 부가적인 단계와 결합하여, 핵산 분자를 시약과 접촉시키는 단계를 포함할 수 있다. 핵산 분자의 뉴클레오타이드 서열에서 그와 같은 변화는 각각의 메틸화된 뉴클레오타이드가 상이한 뉴클레오타이드로 변형된 핵산 분자를 초래할 수 있다. 핵산 뉴클레오타이드 서열에서 그와 같은 변화는 각각의 비메틸화된 뉴클레오타이드가 상이한 뉴클레오타이드로 변형된 핵산 분자를 초래할 수 있다. 핵산 뉴클레오타이드 서열에서 그와 같은 변화는 비메틸화된 각각의 선택된 뉴클레오타이드 (예를 들어, 각각의 비메틸화된 시토신)가 상이한 뉴클레오타이드로 변형된 핵산 분자를 초래할 수 있다. 핵산 뉴클레오타이드 서열을 변화시키기 위해 그와 같은 시약의 사용은 메틸화된 뉴클레오타이드인 각각의 뉴클레오타이드 (예를 들어, 각각의 메틸화된 시토신)가 상이한 뉴클레오타이드로 변형된 핵산 분자를 초래할 수 있다. 본 명세서에서 사용된 바와 같은, 선택된 뉴클레오타이드를 변형시키는 시약의 사용은 본 시약이 다른 3개의 뉴클레오타이드를 변형시키지 없이 하나의 뉴클레오타이드를 변형시키도록, 핵산분자에서 4개의 전형적으로 발생하는 뉴클레오타이드 (DNA의 경우 C, G, T, 및 A 및 RNA의 경우 C, G, U, 및 A) 중 하나의 뉴클레오타이드를 변형시키는 시약을 지칭한다. 하나의 예시적인 구현예에서, 그와 같은 시약은 비메틸화된 선택된 뉴클레오타이드를 변형시켜 상이한 뉴클레오타이드를 생성한다. 또 다른 예시적인 구현예에서, 그와 같은 시약은 비메틸화된 시토신 뉴클레오타이드를 탈아미네이트화할 수 있다. 예시적인 시약은 바이설파이트이다.
- [0094] 본 명세서에서 사용된 바와 같이, 용어 “바이설파이트 시약”은, 예를 들어, CpG 디뉴클레오타이드 서열에서, 메틸화된 시티딘과 비메틸화된 시티딘 간을 구별하기 위해 일부 구현예에서 바이설파이트, 디설파이트, 수소 설파이트, 또는 이들의 조합을 포함하는 시약을 지칭한다.
- [0095] 용어 “메틸화 검정”은 핵산의 서열 내의 하나 이상의 CpG 디뉴클레오타이드 서열의 메틸화 상태를 결정하는 임의의 검정을 지칭한다.
- [0096] 용어 “MS AP-PCR” (메틸화-민감성인 임의로-프라이밍된 중합효소 연쇄반응)은 CG-풍부 프라이머를 사용하여 게놈의 전반적인 스캔을 가능하게 하여 CpG 디뉴클레오타이드를 포함할 공간이 가장 큰 영역에 집중할 수 있고, 그리고 하기에 기재된 기술적으로 인식된 기술을 지칭한다: Gonzalgo 등 (1997) Cancer Research 57: 594-599.
- [0097] 용어 “메틸라이트™”는 하기에 기재된 기술적으로 인식된 형광-기재 실시간 PCR 기술을 지칭한다: Eads 등 (1999) Cancer Res. 59: 2302-2306.
- [0098] 용어 “헤비메틸™”은 증폭 프라이머 사이에서 또는 증폭 프라이머로 덮인 CpG 위치를 포함하는 메틸화 특이적 차단 프로브 (또한 본 명세서에서 일명 차단제)가 핵산 샘플의 메틸화-특이적 선택적 증폭을 가능하게 하는 검정을 지칭한다.
- [0099] 용어 “헤비메틸™ 메틸라이트™” 검정은 메틸라이트™ 검정의 변형인 헤비메틸™ 메틸라이트™ 검정을 지칭하고, 여기서 본 메틸라이트™ 검정은 증폭 프라이머 사이 CpG 위치를 덮는 메틸화 특이적 차단 프로브와 조합된

다.

- [0100] 용어 “Ms-SNuPE” (메틸화-민감성 단일 뉴클레오타이드 프라이머 연장)는 하기에 의해 기재된 기술적으로 인식된 검정을 지칭한다: Gonzalgo & Jones (1997) *Nucleic Acids Res.* 25: 2529-2531.
- [0101] 용어 “MSP” (메틸화-특이적 PCR)는 하기에 의해 기재된 기술적으로 인식된 메틸화 검정을 지칭한다: Herman 등 (1996) *Proc. Natl. Acad. Sci. USA* 93: 9821-9826, 및 미국특허번호 5,786,146.
- [0102] 용어 “COBRA” (조합된 바이셀파이트 제한 분석)는 하기에 의해 기재된 기술적으로 인식된 메틸화 검정을 지칭한다: Xiong & Laird (1997) *Nucleic Acids Res.* 25: 2532-2534.
- [0103] 용어 “MCA” (메틸화된 CpG 섬 증폭)은 하기에 의해 기재된 메틸화 검정을 지칭한다: Toyota 등 (1999) *Cancer Res.* 59: 2307-12, 및 WO 00/26401A1.
- [0104] 본 명세서에서 사용된 바와 같은, “선택된 뉴클레오타이드”는 핵산분자에서 4개의 전형적으로 발생하는 뉴클레오타이드 (DNA의 경우 C, G, T, 및 A 및 RNA의 경우 C, G, U, 및 A) 중 하나의 뉴클레오타이드를 지칭하고, 그리고 전형적으로 발생하는 뉴클레오타이드의 메틸화된 유도체를 포함할 수 있고 (예를 들어, C가 선택된 뉴클레오타이드일 때, 메틸화된 및 비메틸화된 C 둘 모두가 선택된 뉴클레오타이드의 의미 내에 포함됨), 반면에 메틸화된 선택된 뉴클레오타이드는 메틸화된 전형적으로 발생하는 뉴클레오타이드를 구체적으로 지칭하고 비메틸화된 선택된 뉴클레오타이드는 비메틸화된 전형적으로 발생하는 뉴클레오타이드를 구체적으로 지칭한다.
- [0105] 용어들 “메틸화-특이적 제한효소” 또는 “메틸화-민감성 제한효소”는 그것의 인식 부위의 메틸화 상태에 의존하여 핵산을 선택적으로 단리하는 효소를 지칭한다. 인식 부위가 메틸화되지 않았거나 헤미메틸화된 경우 특이적으로 절단하는 제한 효소의 경우, 본 절단은 인식 부위가 메틸화된 경우 일어나지 않거나 현저히 감소된 효율로 일어난다. 인식 부위가 메틸화된 경우 특이적으로 절단하는 제한 효소의 경우, 본 절단은 인식 부위가 메틸화되지 않은 경우 일어나지 않거나 현저히 감소된 효율로 일어난다. 그 인식 서열이 CG 디뉴클레오타이드 (예를 들어 CGCG 또는 CCCGGG와 같은 인식 서열)를 함유하는 메틸화-특이적 제한효소가 바람직한 하다. 일부 구현예에 대해 더욱 바람직한 것은 이 디뉴클레오타이드에서의 시토신이 탄소 원자 C5에서 메틸화되는 경우 절단하지 않는 제한 효소이다.
- [0106] 본 명세서에서 사용된 바와 같이, “상이한 뉴클레오타이드”는 전형적으로 상이한 뉴클레오타이드가 선택된 뉴클레오타이드와 다른 왓슨-크릭 염기쌍짓기 특성을 가지고 그것에 의하여 선택된 뉴클레오타이드에 상보적인 전형적으로 발생하는 뉴클레오타이드가 상이한 뉴클레오타이드에 상보적인 전형적으로 발생하는 뉴클레오타이드와 동일하지 않도록 선택된 뉴클레오타이드와 화학적으로 상이한 뉴클레오타이드를 지칭한다. 예를 들어, C가 선택된 뉴클레오타이드일 때, U 또는 T는 상이한 뉴클레오타이드일 수 있으며, 이는 G에 대한 C의 상보성 및 A에 대한 U 또는 T의 상보성에 의해 예시된다. 본 명세서에서 사용된 바와 같이, 선택된 뉴클레오타이드에 대해 상보적이거나 또는 상이한 뉴클레오타이드에 상보적인 뉴클레오타이드는 4개의 전형적으로 발생하는 뉴클레오타이드 중 3개와 상보적 뉴클레오타이드의 염기쌍짓기보다 더 높은 친화성으로 고염격성 조건하에서 선택된 뉴클레오타이드 또는 상이한 뉴클레오타이드와 염기쌍을 이루는 뉴클레오타이드를 지칭한다. 상보성의 예는 DNA (예를 들어, A-T 및 C-G) 및 RNA (예를 들어, A-U 및 C-G)에서 왓슨-크릭 염기쌍짓기이다. 따라서, 예를 들어, 고염격성 조건하에서, G가 G, A, 또는 T에 대해 염기쌍을 이루기보다 G는 C에 대해 더 높은 친화성으로 염기쌍을 이루고, 따라서, C가 선택된 뉴클레오타이드인 경우, G는 선택된 뉴클레오타이드에 상보적인 뉴클레오타이드이다.
- [0107] 본 명세서에서 사용된 바와 같이, 주어진 마커의 “민감성”은 신생물성 샘플과 비-신생물성 샘플 간을 구별하는 역치 값 이상의 DNA 메틸화 값을 보고하는 샘플의 백분율을 지칭한다. 일부 구현예에서, 양성은 역치 값 (예를 들어, 질환과 관련된 범위) 이상의 DNA 메틸화 값을 보고하는 조직학-확인된 신조직형성으로 정의되고, 거짓 음성은 역치 값 (예를 들어, 무 질환과 관련된 범위) 아래의 DNA 메틸화 값을 보고하는 조직학-확인된 신조직형성으로 정의된다. 민감성의 값은 따라서 공지된 이환 샘플에서 수득된 주어진 마커에 대한 DNA 메틸화 측정이 질병-관련 측정의 범위에 있을 확률을 반영한다. 여기서 정의된 바와 같이, 계산된 민감성 값의 임상적 관련성은 주어진 마커가 그 조건을 갖는 대상체에 적용될 때 임상 조건의 존재를 감지할 개연성의 추정을 나타낸다.
- [0108] 본 명세서에서 사용된 바와 같이 주어진 마커의 “특이성”은 신생물성 샘플과 비-신생물성 샘플 간을 구별하는 역치 값 아래의 DNA 메틸화 값을 보고하는 비-신생물성 샘플의 백분율을 지칭한다. 일부 구현예에서, 음성은 역치 값 (예를 들어, 무 질환과 관련된 범위) 아래의 DNA 메틸화 값을 보고하는 조직학-확인된 비-신생물성 샘플로 정의되고, 거짓 양성은 역치 값 (예를 들어, 질환과 관련된 범위) 이상의 DNA 메틸화 값을 보고하는 조직학-확인된 비-신생물성 샘플로 정의된다. 특이성의 값은 따라서 공지된 비-신생물성 샘플에서 수득된 주어진 마커

에 대한 DNA 메틸화 측정이 무-질환 관련 측정의 범위에 있을 확률을 반영한다. 여기서 정의된 바와 같이, 계산된 특이성 값의 임상적 관련성은 주어진 마커가 그 조건을 갖는 환자에 적용될 때 임상 조건의 부재를 감지할 개연성의 추정을 나타낸다.

- [0109] 본 명세서에서 사용된 바와 같은 용어 “AUC”는 “곡선하의 영역”에 대한 약어이다. 특히 이것은 수신자 조작 특성 (ROC) 곡선 아래의 영역을 지칭한다. 본 ROC 곡선은 진단 시험의 상이한 가능한 컷 포인트에 대한 거짓 양성율에 대한 진성 양성율의 플롯이다. 이것은 선택된 컷 포인트에 따라 민감성과 특이성 사이의 절충점을 보여준다 (민감성에서의 증가는 특이성에서의 감소가 동반될 것이다). ROC 곡선하의 영역 (AUC)은 진단 시험의 정확도에 대한 척도이다 (면적이 클수록 더 양호하고; 최적은 1이고; 무작위 시험은 0.5의 영역을 갖는 대각선 상에 놓인 ROC 곡선을 갖는다; 하기 참조: J. P.Egan.(1975) Signal Detection Theory and ROC Analysis, Academic Press, New York).
- [0110] 본 명세서에서 사용된 바와 같이, 용어 “신생물”은 “정상조직의 것을 초과하고 이것과 조화를 이루지 않는 성장인, 조직의 비정상 덩어리”를 지칭한다. 참고 문헌, 예를 들어, Willis RA, “The Spread of Tumors in the Human Body”, London, Butterworth & Co, 1952.
- [0111] 본 명세서에서 사용된 바와 같이, 용어 “샘종”은 선상 기원의 양성 종양을 지칭한다. 비록 이들 성장은 양성이지만, 경시적으로 이들은 악성으로 진행될 수 있다.
- [0112] 용어 “전-암성” 또는 “전-신생물성” 및 이들의 등가물은 악성 형질전환을 겪는 임의의 세포 증식성 장애를 지칭한다.
- [0113] 신생물, 샘종, 암, 등의 “부위” 또는 “영역”은 신생물, 샘종, 암, 등이 위치한 대상체의 신체에서 조직, 장기, 세포 유형, 해부상의 영역, 신체 부위, 등이다.
- [0114] 본 명세서에서 사용된 바와 같이, “진단” 시험 적용은 대상체의 질환 상태 또는 병태의 검출 또는 확인, 대상체가 주어진 질병 또는 병태를 수축할 가능성을 결정하는 것, 질병 또는 병태를 가진 대상체가 요법에 반응할 가능성을 결정하는 것, 질환 또는 병태 (또는 그것의 유사하게 진행 또는 퇴행)을 갖는 대상체의 예후를 결정하는 것, 및 질환 또는 병태를 갖는 대상체에 대한 치료의 효과를 결정하는 것을 포함한다. 예를 들어, 진단은 신생물을 수축시키는 대상체의 존재 또는 가능성 또는 그러한 대상체가 화합물 (예를 들어, 약제, 예를 들어, 약물) 또는 다른 치료에 유리하게 반응할 가능성을 검출하기 위해 사용될 수 있다.
- [0115] 본 명세서에서 사용된 바와 같은, 용어 “마커”는 예를 들어, 그것의 메틸화 상태에 기반하여 정상 세포로부터 장애-관련된 세포 (예를 들어, 장애와 관련된 비-암성 세포) (예를 들어, 장애와 관련된 암성 세포)를 구별함에 의해 장애 (예를 들어, 비-암성 장애) (예를 들어, 암성 장애)를 진단할 수 있는 서브스טר네스 (예를 들어, 핵산 또는 핵산의 영역)를 지칭한다.
- [0116] “단리된 올리고뉴클레오타이드”에서와 같이 핵산과 관련하여 사용될 때, 용어 “단리된”은 그것이 그것의 천연 공급원과 통상적으로 관련되어 있는 적어도 하나의 오염물질 핵산으로부터 동정되고 분리된 핵산 서열을 지칭한다. 단리된 핵산은 천연에서 발견되는 것과는 다른 형태 또는 환경에 존재한다. 그에 반해서, 비-단리된 핵산, 예컨대 DNA 및 RNA는 이들이 천연에서 존재하는 상태로 발견된다. 비-단리된 핵산의 예는 하기를 포함한다: 인접하는 유전자에 근접하여 숙주 세포 염색체 상에서 발견된 주어진 DNA 서열 (예를 들어, 유전자); RNA 서열, 예컨대 다수의 단백질을 인코딩하는 수많은 다른 mRNA와의 혼합물로서 세포에서 발견된, 특정한 단백질을 인코딩하는 특정한 mRNA 서열.그러나, 특정 단백질을 인코딩하는 단리된 핵산은, 예로써 단백질을 통상적으로 발현하는 세포 내에 이러한 핵산을 포함하여, 본 핵산은 천연 세포의 것과 다른 염색체 위치에 있거나 또는 달리는 천연에서 발견된 것과 다른 핵산 서열에 의해 측정된다. 단리된 핵산 또는 올리고뉴클레오타이드는 단일-가닥 또는 이중-가닥 형태로 존재할 수 있다. 단리된 핵산 또는 올리고뉴클레오타이드가 단백질을 발현하기 위해 이용될 때, 올리고뉴클레오타이드는 센스 또는 코딩 가닥을 최소로 함유할 것이지만 (즉, 올리고뉴클레오타이드는 단일-가닥일 수 있음), 그러나 센스 및 항-센스 가닥 둘 모두를 함유할 수 있다 (즉, 올리고뉴클레오타이드는 이중-가닥일 수 있음). 단리된 핵산은 그것의 천연 또는 전형적인 환경으로부터 단리 후, 다른 핵산 또는 분자와 조합됨에 의해 될 수 있다. 예를 들어, 단리된 핵산은, 예를 들어, 이중성 발현을 위해 배치된 숙주 세포에 존재할 수 있다.
- [0117] 용어 “정제된”은 그것의 천연환경으로부터 제거되거나, 단리되거나, 또는 분리된 핵산 또는 아미노산 서열의 어느 하나의 분자를 지칭한다. “단리된 핵산 서열”은 따라서 정제된 핵산 서열일 수 있다. “실질적으로 정제된” 분자는 이들이 천연적으로 관련된 다른 성분으로부터 적어도 60% 유리, 바람직하게는 적어도 75% 유리, 그

리고 더 바람직하게는 적어도 90% 유리된다. 본 명세서에서 사용된 바와 같은, 용어들 “정제된” 또는 “정제하기 위해”는 또한 샘플로부터 오염물질의 제거를 지칭한다. 오염된 단백질의 제거는 샘플 내의 관심 있는 폴리펩타이드 또는 핵산의 백분율에서의 증가를 초래한다. 또 다른 예에서, 재조합 폴리펩타이드는 식물, 박테리아, 효모, 또는 포유동물 숙주 세포에서 발현되고 그리고 폴리펩타이드는 숙주 세포 단백질의 제거에 의해 정제되고; 재조합 폴리펩타이드의 백분율은 그렇게 함으로써 샘플에서 증가된다.

[0118] 주어진 폴리뉴클레오타이드 서열 또는 폴리펩타이드를 "포함하는 조성물"이라는 용어는 주어진 폴리뉴클레오타이드 서열 또는 폴리펩타이드를 함유하는 임의의 조성물을 광범위하게 지칭한다. 본 조성물은 염 (예를 들어, NaCl), 세제 (예를 들어, SDS), 및 다른 성분 (예를 들어, 덴하르트 용액, 건조 밀크, 연어 정자 DNA, 등)을 함유하는 수용액을 포함할 수 있다.

[0119] 용어 “샘플”은 그것의 가장 넓은 의미로 사용된다. 한 의미에서 이것은 동물 세포 또는 조직을 지칭할 수 있다. 또 다른 의미에서, 이것은 임의의 공급원뿐만 아니라 생물학적 및 환경 샘플로부터 수득된 시료 또는 배양물을 포함하는 것을 의미한다. 생물학적 샘플은 식물이나 동물 (사람을 포함함)로부터 수득될 수 있으며 유체, 고체, 조직 및 가스를 포괄한다. 환경 샘플은 환경 물질 예컨대 지표 물질, 토양, 물, 및 산업 샘플을 포함한다. 이들 예는 본 발명에 적용 가능한 샘플 유형을 제한하는 것으로 해석되어서는 안된다

[0120] 본 명세서에서 사용된 바와 같이, 일부 문맥에서 사용된 바와 같은 “원격 샘플”은 샘플의 세포, 조직 또는 장기 공급원이 아닌 부위로부터 간접적으로 수집된 샘플에 관한 것이다. 예를 들어, 채장으로부터 유래한 샘플 물질이 대변 샘플 (예를 들어, 채장에서 직접적으로 채취한 샘플로부터의 것이 아님)에서 평가되는 경우, 본 샘플은 원격 샘플이다.

[0121] 본 명세서에서 사용된 바와 같은, 용어들 “환자” 또는 “대상체”는 본 기술에 의해 제공되는 다양한 시험을 받는 유기체를 지칭한다. 용어 “대상체”는 동물, 바람직하게는 인간을 비롯한 포유동물을 포함한다. 바람직한 구현예에서, 상기 대상체는 영장류이다. 또 더욱 더 바람직한 구현예에서, 상기 대상체는 인간이다.

[0122] 본 명세서에서 사용된 바와 같이, 용어 “키트”는 물질을 전달하기 위한 임의의 전달 시스템 지칭한다. 반응 검정과 관련하여, 그러한 전달 시스템은 반응 시약 (예를 들어, 적절한 용기내의 올리고뉴클레오타이드, 효소, 등) 및/또는 지지 물질 (예를 들어, 완충액, 검정을 수행하기 위한 지시서, 등)을 한 위치에서 다른 위치로 저장, 이송 또는 전달할 수 있는 시스템을 포함한다. 예를 들어, 키트는 관련된 반응시약 및/또는 지지 물질을 함유하는 하나 이상의 봉입 (예를 들어, 박스)을 포함한다. 본 명세서에서 사용된 바와 같은, 용어 “단편화된 키트”는 각각이 전체 키트 구성요소의 하위부를 함유하는 둘 이상의 별개의 용기를 포함하는 전달 시스템을 지칭한다. 용기는 의도된 수령체에 함께 또는 별도로 전달될 수 있다. 예를 들어, 제1 용기는 검정에 사용하기 위한 효소를 함유할 수 있고, 반면 제2 용기는 올리고뉴클레오타이드를 함유한다. 용어 “단편화된 키트”는 연방 식품 의약품 및 화장품 법의 부분 520(e)에 따라 규제되는 피분산물 특정 시약 (ASR's)을 함유하는 키트를 포괄하는 것으로 의도되지만 이에 제한되지 않는다. 사실상, 각각이 전체 키트 구성요소의 하위부를 함유하는 둘 이상의 별개의 용기를 포함하는 임의의 전달 시스템은 용어 “단편화된 키트”에 포함된다. 그에 반해서, “조합된 키트”는 단일 용기에 (예를 들어, 원하는 구성요소의 각각을 수용하는 단일 박스 안에) 반응 검정의 모든 성분을 함유하는 전달 시스템을 지칭한다. 용어 “키트”는 단편화된 키트 및 조합된 키트 둘 모두를 포함한다.

## [0123] 기술의 구현예

[0124] HCC 스크리닝 (예를 들어, 감시)에 대한 기술과 특히, 비배타적으로, 대상체에서 HCC의 존재를 검출하기 위한 방법, 조성물, 및 관련된 용도가 본 명세서에 제공된다.

[0125] HCC를 검출할 수 있는 마커 및/또는 마커의 패널이 확인되었다 (예를 들어, 표 1 및 4에서 제공된 주석을 갖는 염색체 영역) (참고, 실시예 I, II, 및 III) (예를 들어, ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFNB2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRNI, 및 ECE1).

[0126] 본 명세서에서의 개시내용이 특정 설명된 구현예를 나타내지만, 이들 구현예는 제한이 아닌 예로써 제시된다는 것을 이해해야 한다.

[0127] 본 방법은 대상체로부터 단리된 생물학적 샘플에서 적어도 하나의 메틸화 마커의 메틸화 상태를 결정하는 단계를 포함하며, 여기서 상기 마커의 메틸화 상태의 변화는 HCC의 존재 또는 부류를 나타낸다. 특정 구현예는 HCC의 진단 (예를 들어, 스크리닝)을 위해 사용된 차별적으로 메틸화된 영역 (DMR, 예를 들어, DMR 1-400 (표 1 및 4로부터의 것))을 포함하는 마커에 관한 것이다.



- [0128] 적어도 하나의 마커의 메틸화 분석, 마커의 영역, 본 명세서에서 제공되고 표 1 및 3에 열거된 DMR (예를 들어, DMR 1-400)을 포함하는 또는 마커의 염기가 분석되는 구현예에 부가하여, 본 기술은 또한 적어도 하나의 마커를 포함하는 마커의 패널, 마커의 영역, 또는 대상체에서 HCC의 검출에 대해 유용성을 갖는 DMR을 포함하는 마커의 염기를 제공한다.
- [0129] 본 기술의 일부 구현예는 적어도 하나의 마커의 CpG 메틸화 상태, 마커의 영역, 또는 DMR을 포함하는 마커의 염기의 분석에 기반된다.
- [0130] 일부 구현예에서, 본 기술은 (예를 들어, 표 1 및 4 (예를 들어, DMR 1-400)에서 제공된 바와 같은) DMR을 포함하는 적어도 하나의 마커 내의 CpG 디뉴클레오타이드 서열의 메틸화 상태를 결정하기 위해 하나 이상의 메틸화 검정과 조합한 바이설파이트 기술의 사용을 제공한다. 게놈 CpG 디뉴클레오타이드는 메틸화되거나 또는 비메틸화될 수 있다 (대안적으로 각각 업- 및 다운-메틸화되는 것으로 공지됨). 그러나 본 발명의 방법은 원격 샘플 (예를 들어, 혈액, 장기 유출물, 또는 대변)의 배경 내에서, 불균질 특성, 예를 들어, 이들로부터 저농도의 종양 세포, 또는 생물학적 물질의 생물학적 샘플의 분석에 적합하다. 따라서, 그와 같은 샘플 내에서 CpG 위치의 메틸화 상태를 분석할 때, 특정 CpG 위치에서 메틸화의 수준 (예를 들어, 퍼센트, 분획, 비, 비율 또는 정도)을 결정하기 위한 정량적 검정을 사용할 수 있다.
- [0131] 본 기술에 따르면, DMR을 포함하는 마커에서 CpG 디뉴클레오타이드 서열의 메틸화 상태의 결정은 대상체에서 HCC의 진단 및 특성화 둘 모두에서 유용성을 갖는다.
- [0132] **마커의 조합**
- [0133] 일부 구현예에서, 기술은 표 1 및/또는 4로부터의 2종이상의 DMR (예를 들어, DMR 번호 1-400으로부터 2종 이상의 DMR)을 포함하는 마커의 조합의 메틸화 상태를 평가하는 것에 관한 것이다. 일부 구현예에서, 1초과 마커의 메틸화 상태를 평가하는 것은 대상체에서 HCC의 존재를 확인하기 위한 스크린 또는 진단의 특이성 및/또는 민감성을 증가시킨다.
- [0134] 다양한 암은 예를 들어 예측의 특이성 및 민감성과 관련된 통계적인 기술에 의해 확인되는 바와 같이 다양한 마커의 조합에 의해 예측된다. 본 기술은 일부 암에 대한 예측의 조합과 입증된 예측의 조합을 동정하는 방법을 제공한다.
- [0135] 예를 들어, HCC를 검출할 수 있는 마커 및/또는 마커의 패널이 확인되었다 (예를 들어, 표 1 및 4에서 제공된 주석을 갖는 염색체 영역) (참고, 실시예 I, II, 및 III) (예를 들어, ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFN2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRN1, 및 ECE1).
- [0136] **메틸화 상태를 분석하기 위한 방법**
- [0137] 5-메틸시토신의 존재에 대해 핵산을 분석하기 위해 가장 빈번하게 사용된 방법은 DNA 내의 5-메틸시토신의 검출에 대해 Frommer 등에 의해 기재된 바이설파이트 방법 또는 하기에 기초된다: (Frommer 등 (1992) Proc. Natl. Acad. Sci. USA 89: 1827-31) 또는 이의 변형. 5-메틸시토신을 맵핑하는 바이설파이트 방법은 5-메틸시토신이 아닌 시토신이 수소 설파이트 이온 (바이설파이트로도 공지됨)과 반응하는 관찰에 기반한다. 본 반응은 일반적으로 하기 단계에 따라 수행된다: 먼저, 시토신은 수소 설파이트와 반응하여 설폰화된 시토신을 형성한다. 다음으로, 설폰화된 반응 중간체의 자발적인 탈아민화는 설폰화된 우라실을 초래한다. 마지막으로, 설폰화된 우라실은 알칼리성 조건하에서 탈설폰화되어 우라실을 형성한다. 우라실이 아데닌과 염기쌍을 형성하고 (따라서 티민 같이 거동함), 반면에 5-메틸시토신은 구아닌과 염기쌍을 이루기 때문에 (따라서 시토신 같이 거동함) 검출이 가능하다. 이는 예를 들어 하기에 의해, 비-메틸화된 시토신으로부터의 메틸화된 시토신의 식별을 가능하게 한다: 바이설파이트 게놈 서열분석 (Grigg G, & Clark S, Bioessays (1994) 16: 431-36; Grigg G, DNA Seq.(1996) 6: 189-98) 또는 예를 들어, 하기에 개시된 바와 같은 메틸화-특정 PCR (MSP), 미국특허번호 5,786,146.
- [0138] 일부 전통적 기술은 아가로스 매트릭스 내에 분석되는 DNA를 둘러싸는 단계, 그렇게 함으로써 DNA의 확산 및 복원을 방지하는 단계 (단지 바이설파이트만 단일가닥 DNA와 반응함), 및 빠른 투석으로 침전 및 정제 단계를 대체하는 단계를 포함하는 방법에 관한 것이다 (Olek A, 등 (1996) "A modified and improved method for bisulfite based cytosine methylation analysis" Nucleic Acids Res. 24: 5064-6). 따라서 본 방법의 유용성 및 민감성을 설명하는, 메틸화 상태에 대해 개별 세포를 분석하는 것이 가능하다. 5-메틸시토신을 검출하기 위한 전통적 방법의 개요가 하기에 의해 제공되어 있다: Rein, T., 등 (1998) Nucleic Acids Res. 26: 2255.
- [0139] 바이설파이트 기술은 전형적으로 하기 단계를 포함하여; 바이설파이트 처리에 후속하는 공지된 핵산의 짧은, 특



정한 단편을 증폭시키는 단계, 그 다음 생성물을 하기의 어느 하나에 의해 분석하는 단계: 서열분석 (Olek & Walter (1997) Nat. Genet. 17: 275-6) 또는 프라이머 연장 반응 (Gonzalzo & Jones (1997) Nucleic Acids Res. 25: 2529-31; WO 95/00669; 미국특허번호 6,251,594), 개별 시토신 위치를 분석한다. 일부 방법은 효소적 소화를 사용하고 (Xiong & Laird (1997) Nucleic Acids Res. 25: 2532-4). 하이브리드화에 의한 검출이 또한 본 기술 분야에 기재되어 있다 (Olek 등, WO 99/28498). 추가로, 개별 유전자에 관한 메틸화 검출을 위해 바이설파이트 기술의 사용이 기재되어 있다 (Grigg & Clark (1994) Bioessays 16: 431-6.; Zeschnick 등 (1997) Hum Mol Genet. 6: 387-95; Feil 등 (1994) Nucleic Acids Res. 22: 695; Martin 등 (1995) Gene 157: 261-4; WO 9746705; WO 9515373).

[0140] 다양한 메틸화 검정 절차가 당해 기술에 공지되어 있고 그리고 본 기술에 따른 바이설파이트 처리와 조합하여 사용될 수 있다. 이들 검정은 핵산 서열 내에 하나 또는 복수의 CpG 디뉴클레오타이드 (예를 들어, CpG 섬)의 메틸화 상태의 결정을 허용한다. 다른 기술들 중에서 이러한 검정은 바이설파이트-처리된 핵산의 서열분석, PCR (서열-특정 증폭용), 서던 블랏 분석, 및 메틸화-민감성 제한 효소의 사용을 포함한다.

[0141] 예를 들어, 게놈 서열분석은 바이설파이트 처리를 사용함에 의해 메틸화 패턴 및 5-메틸시토신 분포의 분석을 위해 단순화되어 졌다 (Frommer 등 (1992) Proc. Natl. Acad. Sci. USA 89: 1827-1831). 추가로, 바이설파이트-전환된 DNA로부터 증폭된 PCR 생성물의 제한 효소 단리는, 예를 들어, 하기에 의해 기재된 바와 같이 메틸화 상태를 분석하는데 그 용도가 있다: Sadri & Hornsby (1997) Nucl. Acids Res. 24: 5058-5059 또는 COBRA (조합된 바이설파이트 제한 분석)와 같은 공지된 방법에 형체화된 것 (Xiong & Laird (1997) Nucleic Acids Res. 25: 2532-2534).

[0142] COBRA™ 분석은 소량의 게놈 DNA 내의 특정 유전자좌에서 DNA 메틸화 수준을 결정하는데 유용한 정량적 메틸화 검정이다 (Xiong & Laird, Nucleic Acids Res. 25: 2532-2534, 1997). 간단히, 제한 효소 단리는 아황산수소나 트롬-처리된 DNA의 PCR 생성물에서 메틸화-의존적 서열 차이를 밝히는데 사용된다. 메틸화-의존적 서열 차이는 Frommer 등에 의해 기재된 절차에 따라 표준 바이설파이트 처리에 의해 게놈 DNA 안으로 먼저 도입된다 (Proc. Natl. Acad. Sci. USA 89: 1827-1831, 1992). 그 다음 관심 있는 CpG 섬에 대해 특이적인 프라이머를 사용하여 바이설파이트 전환된 DNA의 PCR 증폭이 수행되고, 이어서 제한 엔도뉴클레아제 단리, 겔 전기영동, 및 특정한 표지된 하이브리드화 프로브를 사용한 검출이 수행된다. 최초 DNA 샘플 내 메틸화 수준은 넓은 범위의 DNA 메틸화 수준에 걸쳐서 선형으로 정량적인 방식으로 단리된 PCR 생성물과 비단리된 PCR 생성물의 상대적인 양으로 표시된다. 또한, 이 기술은 미세해부된 파라핀-포매된 조직 샘플로부터 수득된 DNA에 안정적으로 적용될 수 있다.

[0143] COBRA™ 분석을 위한 전형적인 시약 (예를 들어, 전형적인 COBRA™-기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 특정유전자좌 (예를 들어, 특정한 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 바이설파이트 처리된 DNA 서열, CpG 섬, 등)에 대한 PCR 프라이머; 제한 효소 및 적절한 완충액; 유전자-하이브리드화 올리고뉴클레오타이드; 대조 하이브리드화 올리고뉴클레오타이드; 올리고뉴클레오타이드 탐침을 위한 키나제 라벨링 키트; 및 표지된 뉴클레오타이드. 추가로, 바이설파이트 전환 시약은 하기를 포함할 수 있다: DNA 변성 완충액; 실존화 완충액; DNA 회수 시약 또는 키트 (예를 들어, 침전, 한외여과, 친화성 칼럼); 탈선평화 완충액; 및 DNA 회수 성분.

[0144] 바람직하게는, 검정 예컨대 “메틸라이트™” (형광-기재 실시간 PCR 기술) (Eads 등, Cancer Res. 59: 2302-2306, 1999), Ms-SNuPE™ (메틸화-민감성 단일 뉴클레오타이드 프라이머 연장) 반응 (Gonzalzo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997), 메틸화-특이적 PCR (“MSP”; Herman 등, Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996; 미국특허번호 5,786,146), 및 메틸화된 CpG 섬 증폭 (“MCA”; Toyota 등, Cancer Res. 59: 2307-12, 1999)은 단독으로 또는 이들 방법 중 하나 이상과 조합하여 사용된다.

[0145] “헤비메틸™” 검정인, 기술은 바이설파이트-처리된 DNA의 메틸화-특이적 증폭에 기반한 메틸화 차이를 평가하기 위한 정량적 방법이다. 증폭 프라이머 사이에 또는 이에 의해 커버된 CpG 위치를 덮는 메틸화-특이적 차단 프로브 (“차단제”)는 핵산 샘플의 메틸화-특이적 선택적 증폭을 가능하게 한다.

[0146] 용어 “헤비메틸™ 메틸라이트™” 검정은 메틸라이트™ 검정의 변형인 헤비메틸™ 메틸라이트™ 검정을 지칭하고, 여기서 본 메틸라이트™ 검정은 증폭 프라이머 사이 CpG 위치를 덮는 메틸화 특이적 차단 프로브와 조합된다. 헤비메틸™ 검정은 또한 메틸화 특이적 증폭 프라이머와 함께 사용될 수 있다.

[0147] 헤비메틸™ 분석을 위한 전형적인 시약 (예를 들어, 전형적인 메틸라이트™-기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 특정 유전자좌 (예를 들어, 특정한 유전자, 마커, DMR, 유전자

의 영역, 마커의 영역, 바이선폴아이트 처리된 DNA 서열, CpG 섬, 또는 바이선폴아이트 처리된 DNA 서열 또는 CpG 섬, 등)에 대한 PCR 프라이머; 올리고뉴클레오타이드를 차단하는 것; 최적화된 PCR 완충액 및 데옥시뉴클레오타이드; 및 Taq 중합효소.

[0148] MSP (메틸화-특이적 PCR)는 메틸화 민감성 제한효소의 사용과 독립하여, CpG 섬 내의 CpG 부위의 사실상 임의의 그룹의 메틸화 상태를 평가할 수 있다 (Herman 등 Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996; 미국특허 번호 5,786,146). 간단히, DNA는 비메틸화되지만, 메틸화되지 않은 시토신을 우라실로 전환시키는 아황산수소나트륨에 의해 변형되고, 생성물은 메틸화된 DNA 대 비메틸화된 DNA에 대해 특이적인 프라이머로 후속적으로 증폭된다. MSP는 단지 작은 양의 DNA 만 필요로 하고, 주어진 CpG 섬 유전자좌의 0.1% 메틸화된 대립 유전자에 민감하며, 파라핀-포매된 샘플로부터 추출된 DNA에서 수행될 수 있다. MSP 분석을 위한 전형적인 시약 (예를 들어, 전형적인 MSP -기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 메틸화된 및 비메틸화된 PCR 프라이머로 다음에 대한 것: 특정유전자좌 (예를 들어, 특정한 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 바이선폴아이트 처리된 DNA 서열, CpG 섬, 등); 최적화된 PCR 완충액 및 데옥시뉴클레오타이드, 및 특정한 프로브.

[0149] 메틸라이트™ 검정은 PCR 단계 후 추가의 조작을 요하지 않는 형광-기재 실시간 PCR (예를 들어, TaqMan®)을 이용하는 고-처리량 정량적 메틸화 검정이다 (Eads 등, Cancer Res. 59: 2302-2306, 1999). 간단히, 메틸라이트™ 과정은 아황산수소나트륨 반응에서, 표준절차에 따라 메틸화-의존적 서열 차이의 혼합된 풀로 전환된 게놈 DNA의 혼합된 샘플로 시작한다 (바이선폴아이트 과정은 비메틸화된 시토신 잔기가 우라실로 전환됨). 형광-기재 PCR은 그 다음 “편향된” 반응에서, 예를 들어, 공지된 CpG 디뉴클레오타이드를 중첩시키는 PCR 프라이머로 수행된다. 서열 식별력은 증폭 과정의 수준 및 형광 검출 과정의 수준 둘 모두에서 발생한다.

[0150] 메틸라이트™ 검정은 핵산, 예를 들어 게놈 DNA 샘플에서 메틸화 패턴에 대한 정량적 시험으로서 사용되며, 여기서 서열 식별력은 탐침 하이브리드화의 수준에서 발생한다. 정량적 버전에서, PCR 반응은 특정 추정 메틸화 부위와 중첩하는 형광성 탐침의 존재에서 메틸화 특이적 증폭을 제공한다. 프라이머 또는 탐침이 임의의 CpG 디뉴클레오타이드를 위에 놓이지 않는 반응에 의해 유입 DNA의 양에 대한 비편향된 제어가 제공된다. 대안적으로, 게놈 메틸화에 대한 정성적 시험은 공지된 메틸화 부위를 덮지 않는 대조 올리고뉴클레오타이드 (예를 들어, 헤비메틸™ 및 MSP 기술의 형광-기재 버전) 또는 잠재적인 메틸화 부위를 덮는 올리고뉴클레오타이드 중 어느 하나로 편향된 PCR 풀을 탐침 검사함에 의해 달성된다.

[0151] 메틸라이트™ 과정은 임의의 적합한 탐침 (예를 들어 “TaqMan®” 탐침, 라이트사이클러® 탐침, 등)으로 사용된다. 예를 들어, 일부 적용에서 이중-가닥 게놈 DNA는 아황산수소나트륨으로 처리되고, 그리고 예를 들어, MSP 프라이머 및/또는 헤비메틸 차단제 올리고뉴클레오타이드 및 TaqMan® 탐침과 함께, TaqMan® 탐침을 사용하여 2 세트의 PCR 반응 중 하나를 거친다. TaqMan® 탐침은 형광 “리포터” 및 “켄처” 분자로 이중-표지되고 상대적으로 높은 GC 함량 영역에 특이적으로 되도록 설계되어 이것은 정방향 또는 역방향 프라이머보다 PCR 사이클에서 약 10° C 더 높은 온도에서 용융된다. 이것은 TaqMan® 탐침이 PCR 어닐링/연장 단계 동안 완전하게 혼성화된 상태로 유지되게 한다. Taq 중합 효소가 PCR 동안 효소적으로 새로운 가닥을 합성하기 때문에, 결국 어닐링된 TaqMan® 탐침에 도달할 것이다. Taq 중합효소 5' 에서 3' 으로 엔도뉴클레아제 활성은 그런 다음 TaqMan® 탐침을 단리하여 실시간 형광 검출 시스템을 사용하여 그것의 현재 켄칭되지 않은 신호의 정량적 검출을 위해 형광 리포터 분자를 방출함으로써 대체될 것이다.

[0152] 메틸라이트™ 분석을 위한 전형적인 시약 (예를 들어, 전형적인 메틸라이트™-기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 특정한 유전자좌 (예를 들어, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 바이선폴아이트 처리된 DNA 서열, CpG 섬, 등)에 대한 PCR 프라이머; TaqMan® 또는 라이트사이클러® 탐침; 최적화된 PCR 완충액 및 데옥시뉴클레오타이드; 및 Taq 중합효소.

[0153] QM™ (정량적 메틸화) 검정은 게놈 DNA 샘플 내 메틸화 패턴에 대한 대안적인 정량적 시험이고, 여기서 서열 식별력은 탐침 하이브리드화의 수준에서 발생한다. 이 정량적 버전에서, PCR 반응은 특정 추정 메틸화 부위를 중첩하는 형광 탐침의 존재에서 비편향된 증폭을 제공한다. 프라이머 또는 탐침이 임의의 CpG 디뉴클레오타이드를 위에 놓이지 않는 반응에 의해 유입 DNA의 양에 대한 비편향된 제어가 제공된다. 대안적으로, 게놈 메틸화에 대한 정성적 시험은 공지된 메틸화 부위를 덮지 않는 대조 올리고뉴클레오타이드 (예를 들어, 헤비메틸™ 및 MSP 기술의 형광-기재 버전) 또는 잠재적인 메틸화 부위를 덮는 올리고뉴클레오타이드 중 어느 하나로 편향된 PCR 풀을 탐침 검사함에 의해 달성된다.

[0154] QM™ 과정은 증폭 과정에서 임의의 적합한 탐침, 예를 들어 “TaqMan®” 탐침, 라이트사이클러® 탐침으로 사

용될 수 있다. 예를 들어, 이중-가닥 게놈 DNA는 아황산수소나트륨으로 처리되고 비편향된 프라이머 및 TaqMan® 탐침을 거친다. TaqMan® 탐침은 형광 “리포터” 및 “켄처” 분자로 이중-표지되고 상대적으로 높은 GC 함량 영역에 특이적으로 되도록 설계되어 이것은 정방향 또는 역방향 프라이머보다 PCR 사이클에서 약 10° C 더 높은 온도에서 용융된다. 이것은 TaqMan® 탐침이 PCR 어닐링/연장 단계 동안 완전하게 혼성화된 상태로 유지되게 한다. Taq 중합 효소가 PCR 동안 효소적으로 새로운 가닥을 합성하기 때문에, 결국 어닐링된 TaqMan® 탐침에 도달할 것이다. Taq 중합효소 5' 에서 3' 으로 엔도뉴클레아제 활성은 그런 다음 TaqMan® 탐침을 단리하여 실시간 형광 검출 시스템을 사용하여 그것의 현재 켄칭되지 않은 신호의 정량적 검출을 위해 형광 리포터 분자를 방출함으로써 대체될 것이다. QM™ 분석을 위한 전형적인 시약 (예를 들어, 전형적인 QM™-기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 특정한 유전자좌 (예를 들어, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 바이설파이트 처리된 DNA 서열, CpG 섬, 등)에 대한 PCR 프라이머; TaqMan® 또는 라이트사이클러® 탐침; 최적화된 PCR 완충액 및 데옥시뉴클레오타이드; 및 Taq 중합효소.

[0155] Ms-SNuPE™ 기술은 DNA의 바이설파이트 처리와, 이어서 단일-뉴클레오타이드 프라이머 연장에 기반한 특정 CpG 부위에서 메틸화 차이를 평가하기 위한 정량적 방법이다 (Gonzalzo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997). 간단히, 게놈 DNA는 5-메틸시토신이 변함 없는 남아 있으면서 비메틸화된 시토신을 우라실로 전환시키기 위해 아황산수소나트륨과 반응된다. 원하는 표적서열의 증폭은 그런 다음 바이설파이트-전환된 DNA에 대해 특이적인 PCR 프라이머를 사용하여 수행되고, 그리고 수득한 생성물은 단리되고 관심 있는 CpG 부위에서 메틸화 분석을 위한 템플레이트로 사용된다. 소량의 DNA가 분석될 수 있고 (예를 들어, 미세해부된 병리학 부문) 및 이것은 CpG 부위에서 메틸화 상태를 결정하기 위한 제한 효소의 이용을 회피한다.

[0156] Ms-SNuPE™ 분석을 위한 전형적인 시약 (예를 들어, 전형적인 Ms-SNuPE™-기재 키트에서 발견될 수 있는 것과 같음)은, 비제한적으로, 하기를 포함할 수 있다: 특정한 유전자좌 (예를 들어, 특정한 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 바이설파이트 처리된 DNA 서열, CpG 섬, 등)에 대한 PCR 프라이머; 최적화된 PCR 완충액 및 데옥시뉴클레오타이드; 겔 추출 키트; 양성 대조군 프라이머; 특정한 유전자좌에 대한 Ms-SNuPE™ 프라이머; 반응 완충제 (Ms-SNuPE 반응용); 및 표지된 뉴클레오타이드. 추가로, 바이설파이트 전환 시약은 하기를 포함할 수 있다: DNA 변성 완충액; 설폰화 완충액; DNA 회수 시약 또는 키트 (예를 들어, 침전, 한외여과, 친화성 칼럼); 탈설폰화 완충액; 및 DNA 회수 성분.

[0157] 감소된 표현 바이설파이트 서열 분석 (RRBS)은 모든 비메틸화된 시토신을 우라실로 전환시키기 위해 핵산의 바이설파이트 처리로 시작하고, 이어서 제한 효소 단리 (예를 들어, MspI과 같은 CG 서열을 포함하는 부위를 인식하는 효소에 의함) 및 어댑터간드에 커플링 후 단편의 완전한 서열 분석이 따른다. 제한효소의 선택은 CpG 치밀한 영역에 대한 단편을 풍부하게 하여, 분석 중에 여러 유전자 위치에 매핑될 수 있는 중복적인 서열의 수를 감소시킨다. 이와 같이, RRBS는 서열분석을 위한 제한 단편의 서브셋을 (예를 들어, 예비 겔 전기 영동을 사용하여 크기 선택에 의해) 선택함으로써 핵산 샘플의 복잡성을 감소시킨다. 전체의-게놈 바이설파이트 서열분석과는 대조적으로, 제한 효소 단리에 의해 생산된 모든 단편은 적어도 하나의 CpG 디뉴클레오타이드에 대한 DNA 메틸화 정보를 함유한다. 이와 같이, RRBS는 프로모터, CpG 섬 및 기타 게놈 특징에 대한 샘플을 이들 영역에서 제한 효소 절단 부위가 풍부하게 하고 따라서 하나 이상의 유전자 좌위의 메틸화 상태를 평가하는 검정을 제공한다.

[0158] RRBS에 대한 전형적인 프로토콜은 제한효소 예컨대 MspI로 핵산 샘플을 단리하는 단계, 돌출부를 충전하고 A-테일링하는 단계, 어댑터를 접합하는 단계, 바이설파이트 전환 및 PCR의 단계를 포함한다. 참고 문헌, 예를 들어, 등 (2005) “Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution” Nat Methods 7: 133-6; Meissner 등 (2005) “Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis” Nucleic Acids Res. 33: 5868-77.

[0159] 일부 구현예에서, 정량적 대립유전자-특이적 실시간 표적 및 신호 증폭 (QuARTS) 검정이 메틸화 상태를 평가하기 위해 사용된다. 증폭 (반응 1) 및 1차 반응에서 표적 탐침 절단 (반응 2); 및 2차 반응에서 FRET 절단 및 형광 신호 생성 (반응 3)을 포함하는 3개의 반응이 각각의 QuARTS 검정에서 순차적으로 일어난다. 표적 핵산이 특정한 프라이머 증폭될 때, 플랩 서열을 갖는 특정한 검출 탐침은 앰플리콘에 느슨하게 결합한다. 상기 표적 결합 부위에 특정 침습성 올리고뉴클레오타이드의 존재는 절단화 효소가 검출 탐침과 플랩 서열 사이를 절단함으로써 플랩 서열을 방출하게 한다. 플랩 서열은 상응하는 FRET 카세트의 비 헤어핀 부분에 상보적이다. 따라서, 플랩 서열은 FRET 카세트상의 침습성 올리고뉴클레오타이드로서 작용하고 FRET 카세트 형광단과 켄처 사이에서 절단을 유효하게 하여 형광 신호를 생성한다. 절단 반응은 표적당 다중 탐침을 절단할 수 있고 따라서 플랩당



다중 형광단을 방출하여 지수 신호 증폭을 제공할 수 있다. QuARTS는 상이한 염료로 FRET 카셋트를 사용하여 단일 반응 웰에서 다수의 표적을 검출할 수 있다. 참고 문헌, 예를 들어, Zou 등 (2010) “Sensitive quantification of methylated markers with a novel methylation specific technology” Clin Chem 56: A199; 미국특허 출원일련번호 12/946,737, 12/946,745, 12/946,752, 및 61/548,639.

[0160] 용어 “바이선티프 시약”은 본 명세서에서 개시된 바와 같이 메틸화된 CpG 디뉴클레오타이드 서열과 비메틸화된 CpG 디뉴클레오타이드 서열 간을 구별하는데 유용한, 바이선티프, 디선티프, 수소선티프, 또는 이들의 조합을 포함하는 시약을 지칭한다. 상기 처리의 방법은 당해 기술에 공지되어 있다 (예를 들어, PCT/EP2004/011715). 바람직하게는, 바이선티프 처리는 변성 용매 예컨대 비제한적으로 n-알킬렌글리콜 또는 디에틸렌 글리콜 디메틸 에테르 (DME)의 존재에서, 또는 디옥산 또는 디옥산 유도체의 존재에서 수행된다. 일부 구현예에서 본 변성 용매는 1% 내지 35% (v/v) 사이의 농도로 사용된다. 일부 구현예에서, 바이선티프 반응은 포착제 예컨대 비제한적으로 크로만 유도체, 예를 들어, 6-하이드록시-2,5,7,8,-테트라메틸크로만 2-카복실산 또는 트리하이드록시벤존산 및 이들의 유도체, 예를 들어, 갈산의 존재에서 수행된다 (참고: PCT/EP2004/011715). 바이선티프 전환은 바람직하게는 30° C 내지 70° C 사이의 반응 온도에서 수행되고, 그것에 의하여 온도는 반응 동안 단시간에 85° C 이상으로 증가된다 (참고: PCT/EP2004/011715). 바이선티프 처리된 DNA는 바람직하게는 정량화 이전에 정제된다. 이것은 예를 들어, (Millipore™에 의해 제조된) Microcon™ 칼럼의 수단에 의해, 비제한적으로 환외여과와 같이, 본 기술분야에서 공지된 임의의 수단에 의해 수행될 수 있다. 정제는 변형된 제조자의 프로토콜에 따라 수행된다 (참고, 예를 들어, PCT/EP2004/011715).

[0161] 일부 구현예에서, 처리된 DNA의 단편은 본 발명에 따른 프라이머 올리고뉴클레오타이드 세트 (예를 들어, 표 2 참고) 및 증폭효소를 사용하여 증폭된다. 몇 개의 DNA 분절의 증폭은 하나 및 동일한 반응 용기에서 동시에 수행될 수 있다. 전형적으로, 증폭은 증합효소 연쇄 반응 (PCR)을 사용하여 수행된다. 앰플리콘은 전형적으로 길이에서 100 내지 2000 염기쌍이다.

[0162] 본 방법의 또 다른 구현예에서, DMR (예를 들어, DMR 1-400; 표 1 및 4)을 포함하는 마커 내 또는 근처에서 CpG 위치의 메틸화 상태는 메틸화-특이적 프라이머 올리고뉴클레오타이드의 사용에 의해 검출될 수 있다. 이 기술 (MSP)은 하기에 기재되어 있다: 미국특허번호 6,265,171 (Herman). 바이선티프 처리된 DNA의 증폭을 위해 메틸화 상태 특이적 프라이머의 사용은 메틸화된 핵산과 비메틸화된 핵산 사이의 분화를 가능하게 한다. MSP 프라이머 쌍은 바이선티프 처리된 CpG 디뉴클레오타이드에 혼성화하는 적어도 하나의 프라이머를 함유한다. 따라서, 상기 프라이머의 서열은 적어도 하나의 CpG 디뉴클레오타이드를 포함한다. 비-메틸화된 DNA에 특이적인 MSP 프라이머는 CpG 내의 C 위치의 위치에 “T”를 함유한다.

[0163] 증폭의 수단에 의해 수득된 단편은 검출가능한 표지를 직접적으로 또는 간접적으로 담지할 수 있다. 일부 구현예에서, 라벨은 질량 분광분석기에서 검출될 수 있는 전형적인 질량을 갖는 형광 라벨, 방사선 핵종 또는 탈착 가능한 분자 단편이다. 상기 라벨이 질량 라벨인 경우, 표지된 앰플리콘이 단일 양성 또는 음성의 순 전하를 가져 질량 분광분석기에서 더 나은 유전성을 가능하게 하는 일부 구현예를 제공한다. 검출은 예를 들어, 매트릭스 보조 레이저 탈착/이온화 질량 분광분석법 (MALDI)의 수단에 의하거나 또는 전자 분무 질량 분광분석법 (ESI)을 사용하여 수행되고 가시화될 수 있다.

[0164] 이들 검정 기술에 적합한 DNA를 분리하는 방법은 당해 기술에 공지되어 있다. 특히, 일부 구현예는 하기에 기재된 바와 같은 핵산의 단리를 포함한다: 미국특허 출원일련번호 13/470,251 (“Isolation of Nucleic Acids”).

# [0165] 방법

[0166] 일부 구현예 기술에서, 하기 단계를 포함하는 방법이 제공된다:

[0167] 1) 대상체로부터 수득된 핵 (예를 들어, 체액 예컨대 혈액 샘플 (예를 들어, 혈장 샘플), 대변 샘플, 또는 조직 샘플로부터 단리된, 예를 들어, 게놈 DNA)을 DMR (예를 들어, DMR 1-400 (표 1 및 4로부터의 것))을 포함하는 적어도 하나의 마커 내의 메틸화된 CpG 디뉴클레오타이드와 비-메틸화된 CpG 디뉴클레오타이드 간을 구별하는 적어도 하나의 시약 또는 일련의 시약들과 접촉시키는 단계 및

[0168] 2) HCC의 결핍을 검출하는 단계 (예를 들어, 80% 이상의 민감성 및 80% 이상의 특이성이 부여됨).

[0169] 일부 구현예 기술에서, 하기 단계를 포함하는 방법이 제공된다:

[0170] 1) 대상체로부터 수득된 핵산 (예를 들어, 체액 예컨대 혈액 샘플 (예를 들어, 혈장 샘플), 대변 샘플, 또는 조직 샘플로부터 단리된, 예를 들어, 게놈 DNA)을 DMR (예를 들어, DMR 1-400 (표 1 및 4로부터의 것))을 포함하

는 적어도 하나의 마커 내의 메틸화된 CpG 디뉴클레오타이드와 비-메틸화된 CpG 디뉴클레오타이드 간을 구별하는 적어도 하나의 시약 또는 일련의 시약들과 접촉시키는 단계 및

- [0171] 2) HCC를 분류하는 단계 (예를 들어, 80% 이상의 민감성 및 80% 이상의 특이성이 부여됨).
- [0172] 바람직하게는, 민감성은 약 70% 내지 약 100%, 또는 약 80% 내지 약 90%, 또는 약 80% 내지 약 85%이다. 바람직하게는, 특이성은 약 70% 내지 약 100%, 또는 약 80% 내지 약 90%, 또는 약 80% 내지 약 85%이다.
- [0173] 게놈 DNA는 상업적으로 입수가 가능한 키트의 사용을 포함한 임의의 수단에 의해 분리될 수 있다. 간단히, 관심 대상의 DNA가 세포막에 의해 캡슐화된 생물학적 샘플은 효소적, 화학적 또는 기계적 수단에 의해 파괴되고 용해되어야 한다. DNA 용액은 그런 다음, 예를 들어, 프로테이나제 K로의 단리에 의해 단백질을 및 다른 오염물질이 소거될 수 있다. 게놈 DNA는 그런 다음 용액으로부터 회수된다. 이것은 염색, 유기 추출 또는 고상 지지체에 대한 DNA의 결합을 포함하는 다양한 방법에 의해 수행될 수 있다. 방법의 선택은 시간, 비용 및 요구된 DNA 양을 포함한 여러 인자에 의해 영향을 받을 것이다. 예를 들어, 세포주, 조직학적 슬라이드, 생검, 파라핀-포매된 조직, 체액, 대변, 결장 유출물, 소변, 혈장, 혈청, 전혈, 분리된 혈구, 혈액으로부터 분리된 세포, 및 이들의 조합인, 신생물성 물질 또는 전-신생물성 물질을 포함하는 모든 임상 샘플 유형이 본 방법에서의 사용에 적합하다.
- [0174] 본 기술은 샘플을 준비하고 시험용 핵산을 제공하는 데 사용되는 방법에 제한되지 않는다. 예를 들어, 일부 구현예에서, DNA는, 예를 들어, 하기에 상세히 기술된 바와 같이, 직접적인 유전자 포착을 사용하여 대변 샘플 또는 혈액 또는 혈장 샘플로부터 분리된다: 미국특허 출원일련번호 61/485386 또는 그와 관련된 방법.
- [0175] 게놈 DNA 샘플은 그런 다음 DMR (예를 들어, 표 1 및 4에 의해 제공된 바와 같은, 예를 들어, DMR 1-400)을 포함하는 적어도 하나의 마커 내의 메틸화된 CpG 디뉴클레오타이드와 비-메틸화된 CpG 디뉴클레오타이드 간을 구별하는 적어도 하나의 시약, 또는 일련의 시약들로 처리된다.
- [0176] 일부 구현예에서, 본 시약은 5' -위치에서 비메틸화된 시토신 염기를 하이브리드화 행동의 관점에서 시토신에 비유사한 우라실, 티민, 또는 또 다른 염기로 전환시킨다. 그러나 일부 구현예에서, 본 시약은 메틸화 민감성 제한 효소일 수 있다.
- [0177] 일부 구현예에서, 게놈 DNA 샘플은 5' -위치에서 비메틸화된 시토신 염기가 하이브리드화 행동의 관점에서 시토신에 비유사한 우라실, 티민, 또는 또 다른 염기로 전환되는 그와 같은 방식으로 처리된다. 일부 구현예에서, 이 처리는 바이셀레이트 (수소 셀파이트, 디셀파이트)로 수행되고 이어서 알칼리성 가수분해가 따른다.
- [0178] 처리된 핵산은 그런 다음 분석되어 표적 유전자 서열 (DMR, 예를 들어, 표 1 및 4에서 제공된 바와 같은, 예를 들어, DMR 1-400으로부터 선택된 적어도 하나의 DMR을 포함하는 마커로부터의 적어도 하나의 유전자, 유전자 서열, 또는 뉴클레오타이드)의 메틸화 상태를 결정한다. 분석 방법은 본 명세서에 열거된 것들, 예를 들어, 본 명세서에서 기재된 바와 같은 QuARTS 및 MSP를 포함하여, 당해 분야에서 공지된 것들로부터 선택될 수 있다.
- [0179] 본 기술은 HCC와 관련된 임의의 샘플의 분석에 관한 것이다. 예를 들어, 일부 구현예에서 샘플은 환자로부터의 혈장 샘플을 포함한다. 일부 구현예에서, 샘플은 환자로부터 획득된 조직 및/또는 생체액을 포함한다. 일부 구현예에서, 샘플은 간 조직을 포함한다. 일부 구현예에서, 샘플은 분비물을 포함한다. 일부 구현예에서, 샘플은 혈액, 혈장 및/또는 혈청을 포함한다. 일부 구현예에서, 대상체는 인간이다. 이들 샘플은 상부 위장관, 하부 위장관에서 유래할 수 있고, 상부 위장관 및 하부 위장관 둘 모두로부터의 세포, 조직 및/또는 분비물을 포함할 수 있다. 샘플은 간, 담관, 췌장, 위, 결장, 직장, 식도, 소장, 맹장, 십이지장, 용종, 담낭, 항문, 및/또는 복막으로부터의 세포, 분비물, 또는 조직을 포함할 수 있다. 일부 구현예에서, 샘플은 세포액, 복수, 소변, 대변, 췌장 유체, 내시경술 동안 획득된 유체, 혈액, 점액, 또는 타액을 포함한다. 일부 구현예에서, 샘플은 대변 샘플이다.
- [0180] 이러한 샘플은 숙련자에게 분명한 바와 같이 본 기술분야에 공지된 임의의 수의 수단에 의해 획득될 수 있다. 예를 들어, 소변 및 대변 샘플은 쉽게 획득 가능한 반면, 혈액, 복수, 혈청, 또는 췌장 유체 샘플은 예를 들어 바늘 및 주사기를 사용함에 의해 비경구적으로 획득될 수 있다. 무세포 또는 실질적으로 무세포 샘플은, 비제한적으로, 원심분리 및 여과를 포함하여 당해 분야의 숙련자에게 공지된 다양한 기술로 샘플을 처리함에 의해 획득될 수 있다. 샘플을 얻기 위해 침습성 기술을 사용하지 않는 것이 일반적으로 바람직하지만, 조직 균질물, 조직 절편 및 생검 시료와 같은 샘플을 얻는 것은 여전히 바람직할 수 있다.
- [0181] 기술의 일부 구현예에서, 대상체에서 HCC를 진단하는 방법이 제공된다. 본 명세서에서 사용된 바와 같은 용어들

“진단하는” 및 “진단”은 숙련가가 대상체가 주어진 질환 또는 병태를 앓고 있는지 또는 미래에 주어진 질환 또는 병태를 나타낼 수 있는지 여부를 평가할 수 있고 또한 더욱이 결정할 수 있는 방법을 지칭한다. 숙련가는 종종 하나 이상의 진단 지표, 예컨대 예를 들어, 바이오마커 (예를 들어, 본 명세서에서 개시된 바와 같은 DMR), 병태의 존재, 중증도, 또는 부재를 나타내는 그 메틸화 상태에 기초하여 진단을 한다.

[0182] 진단과 함께, 임상 암 예후 (예를 들어, HCC)는 가장 효과적인 요법을 계획하기 위해 암의 공격성 및 종양 재발의 가능성을 결정하는 것과 관련된다. 만일 보다 정확한 예후가 이루어 질 수 있거나 암을 진행하는 잠재적인 위험조차도 평가될 수 있다면, 적절한 요법과 일부 사례에서는 환자에 대해 덜 심한 요법이 선택될 수 있다. 암 바이오마커의 평가 (예를 들어, 메틸화 상태를 결정하는 것)는 양호한 예후 및/또는 치료법을 요하지 않을 암을 전개하는 낮은 위험 또는 암을 전개하기가 보다 쉬운 이들로부터의 제한된 요법 또는 보다 집중적인 치료로부터 이점이 있을 암의 재발을 앓고 있는 개별 대상체들에게 유용하다.

[0183] 이와 같이, 본 명세서에서 사용된 바와 같은 “진단을 하는” 또는 “진단하는”은 암을 유발할 위험을 결정하는 것 또는 예후를 결정하는 것을 포함하며, 이는 본 명세서에서 개시된 진단 바이오마커 (예를 들어, DMR)의 측정에 기반하여, (의료적 치료가 있거나 없이) 임상 결과를 예측하거나, 적절한 치료 (또는 치료가 효과적인지 여부)를 선택하거나, 또는 현재 치료를 모니터링하고, 잠재적으로 치료를 변경하기 위해 제공될 수 있다. 또한, 현재 개시된 요지의 일부 구현예에서, 진단 및/또는 예후를 용이하게 하기 위해 경시적으로 바이오마커의 다중 결정이 이루어질 수 있다. 바이오마커에서의 일시적 변화는 임상 결과를 예측하고, HCC의 진행을 모니터링하고 및/또는 암에 대한 적절한 요법의 효능을 모니터링하는데 사용될 수 있다. 예를 들어 그와 같은 구현예에서, 효과적인 요법의 과정 도중 경시적으로 생물학적 샘플에서 본 명세서에 개시된 하나 이상의 바이오마커 (예를 들어, DMR) (그리고, 모니터링된다면, 잠재적으로 하나 이상의 추가의 바이오마커(들))의 메틸화 상태에서의 변화를 알 수 있는 것이 기대될 수 있다.

[0184] 일부 구현예에서 현재 개시된 요지는 대상체에서 HCC의 예방 또는 치료를 개시할지 또는 계속할지를 결정하는 방법을 추가로 제공한다. 일부 구현예에서, 상기 방법은 대상체로부터의 기간에 걸쳐 일련의 생물학적 샘플을 제공하는 단계; 일련의 생물학적 샘플을 분석하여 각각의 생물학적 샘플에서 본 명세서에 개시된 적어도 하나의 바이오마커의 메틸화 상태를 결정하는 단계; 및 각각의 생물학적 샘플에서 바이오마커 중 하나 이상의 메틸화 상태에서의 임의의 측정가능한 변화를 비교하는 단계를 포함한다. 일정 시간 기간에 걸쳐 바이오마커의 메틸화 상태에서의 임의의 변화는 HCC를 전개할 위험을 예측하고, 임상 결과를 예측하고, 암의 예방 또는 치료의 시작하거나 지속할지 여부와, 현재 요법이 HCC를 효과적으로 치료하는지 여부를 결정하는데 사용될 수 있다. 예를 들어, 제1 시점은 치료 개시 이전에 선택될 수 있고 제2 시점은 치료의 개시 후 어느 시점에서 선택될 수 있다. 메틸화 상태는 상이한 시점에서 취한 각각의 샘플에서 측정될 수 있으며, 정성적 및/또는 정량적 차이가 나타내어 질 수 있다. 상이한 샘플로부터의 바이오마커 수준의 메틸화 상태에서의 변화는 대상체에서의 위험 (예를 들어, HCC의 위험), 예후, 치료 효능을 결정하는 것 및/또는 장애의 진행과 상관될 수 있다.

[0185] 바람직한 구현예에서, 본 발명의 방법 및 조성물은, 예를 들어, 질환의 증상이 나타나기 전에, 초기 단계에서 질환의 치료 또는 진단을 위한 것이다. 일부 구현예에서, 본 발명의 방법 및 조성물은 임상 단계에서 질환의 치료 또는 진단을 위한 것이다.

[0186] 지적된 바와 같이, 일부 구현예에서, 하나 이상의 진단 또는 예후 바이오마커의 다중 결정이 이루어질 수 있고, 마커에서의 일시적 변화는 진단 또는 예후를 결정하는데 사용될 수 있다. 예를 들어, 진단 마커는 초기 시간에 결정될 수 있고, 두 번째 시간에 다시 결정될 수 있다. 그와 같은 구현예에서, 초기 시간으로부터 두 번째 시간으로 마커에서의 증가는 장애의 특정 유형 또는 중증도 또는 주어진 예후를 진단할 수 있다. 마찬가지로, 초기 시간으로부터 두 번째 시간으로 마커에서의 감소는 장애의 특정 유형 또는 중증도 또는 주어진 예후를 나타낼 수 있다. 게다가, 하나 이상의 마커의 변화 정도는 장애의 중증도 및 장애의 유해 사례와 관련될 수 있다. 숙련가는 특정 구현예에서 다수의 시점에서 동일한 바이오마커를 비교 측정할 수 있으면서, 하나의 시점에서 주어진 바이오마커를 그리고 두 번째 시점에서 제2 바이오마커를 측정할 수 있으며, 이들 마커의 비교는 진단 정보를 제공할 수 있다는 것을 이해할 것이다.

[0187] 본 명세서에서 사용된 바와 같이, 어구 “예후를 결정하는 것”은 숙련가가 대상체의 병태의 경과 또는 결과를 예측할 수 있는 방법을 지칭한다. 용어 “예후”는 100% 정확도로 병태의 경과 또는 결과를 예측하는 능력을 지칭하지 않으며, 또는 심지어 주어진 경과 또는 결과는 바이오마커 (예를 들어, DMR)의 메틸화 상태에 기반하여 예상대로 보다 많이 또는 보다 적게 발생하기 쉽다. 대신에, 숙련가는 용어 “예후”는 특정 경과 또는 결과가 발생할 증가된 개연성을 지칭한다; 즉, 경과 또는 결과는 병태를 나타내지 않는 개체들과 비교할 때 주어진 병

태를 나타내는 대상체에서 보다 발생하기 쉽다는 것을 이해할 것이다. 예를 들어, 병태를 나타내지 않는 개체 (예를 들어, 하나 이상의 DMR의 정상 메틸화 상태를 갖는 것)에서, 주어진 결과의 변화는 아주 낮을 수 있다.

[0188] 일부 구현예에서, 통계적인 분석은 부정적인 결과에 대한 소인과 예후 지표를 연관시킨다. 예를 들어, 일부 구현예에서, 장애를 갖지 않는 환자로부터 수득된 정상 대조군 샘플에서의 것과 상이한 메틸화 상태는 대상체가 통계적 유의도의 수준에 의해 결정된 바와 같이, 대조군 샘플에서 메틸화 상태에 보다 유사한 수준을 갖는 대상체보다 더 장애를 앓기 쉽다는 신호일 수 있다. 추가로, 기준선 (예를 들어, “정상”) 수준으로부터 메틸화 상태에서의 변화는 대상체 예후를 반영할 수 있고, 메틸화 상태에서의 변화의 정도는 유해 사례의 중증도와 관련될 수 있다. 통계적 유의도는 2종 이상의 모집단을 비교하고 신뢰 구간 및/또는 p 값을 결정함으로써 종종 결정된다 (참고 문헌, 예를 들어, Dowdy and Wearden, Statistics for Research, John Wiley & Sons, New York, 1983). 본 요지의 예시적인 신뢰구간은 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.9% 및 99.99%인 반면, 예시적인 p 값은 0.1, 0.05, 0.025, 0.02, 0.01, 0.005, 0.001, 및 0.0001이다.

[0189] 다른 구현예에서, 본 명세서에 개시된 예후 또는 진단 바이오마커 (예를 들어, DMR)의 메틸화 상태에서의 변화의 역치 정도가 확립될 수 있고, 생물학적 샘플에서 바이오마커의 메틸화 상태에서 변화의 정도는 메틸화 상태에서의 변화의 역치 정도에 간단히 비교된다. 본 명세서에 제공된 바이오마커에 대한 메틸화 상태에서 바람직한 역치 변화는 약 5%, 약 10%, 약 15%, 약 20%, 약 25%, 약 30%, 약 50%, 약 75%, 약 100%, 및 약 150%이다. 또 다른 구현예에서, 예후 또는 진단 지표 (바이오마커 또는 바이오마커의 조합)의 메틸화 상태가 주어진 결과에 대한 관련된 성향과 직접적으로 관련되는 “노도 그래프”가 확립될 수 있다. 숙련가는 개별 샘플 측정은 모집단 평균이 아니라 참조되기 때문에, 이 측정에서 불확실성이 마커 농도에서의 불확실성과 동일하다는 것을 이해하면서 두 개의 수치를 관련시키는 그러한 노도그래프의 사용에 정통한다.

[0190] 일부 구현예에서, 생물학적 샘플로부터 수득된 결과가 대조군 샘플로부터 수득된 결과에 비교될 수 있도록 대조군 샘플은 생물학적 샘플과 동시에 분석된다. 추가로, 생물학적 샘플에 대한 검정 결과가 비교될 수 있는 표준 곡선이 제공될 수 있다는 것이 고려된다. 이러한 표준 곡선은 형광 표지가 사용되는 경우 검정 단위, 예를 들어, 형광 신호 강도의 함수로서 바이오마커의 메틸화 상태를 나타낸다. 다중 공여체로부터 취한 샘플을 사용하여, 표준 곡선은 정상 조직에서 하나 이상의 바이오마커의 메틸화 상태를 제어할 뿐만 아니라 화생이 있는 공여체로부터 또는 장애 (예를 들어, HCC)가 있는 공여체로부터 취한 조직에서 하나 이상의 바이오마커의 “위험이 있는” 수준에 대해 제공될 수 있다. 본 방법의 특정 구현예에서, 대상체는 대상체로부터 수득된 생물학적 샘플에서 본 명세서에서 제공된 하나 이상의 DMR의 비정상적인 메틸화 상태를 확인시 HCC를 갖는 것으로 확인된다. 본 방법의 다른 구현예에서, 대상체로부터 수득된 생물학적 샘플에서 이러한 바이오마커 중 하나 이상의 비정상적인 메틸화 상태의 검출은 상기 대상체가 HCC를 갖는 것으로 확인된다.

[0191] 마커의 분석은 하나의 시험 샘플 내에 추가 마커와 별도로 또는 동시에 수행될 수 있다. 예를 들어, 다수의 샘플을 효율적으로 처리하고 잠재적으로 더 큰 진단 및/또는 예후 정확도를 제공하기 위해 몇 개의 마커가 하나의 시험으로 조합될 수 있다. 또한, 당해 분야의 숙련가는 동일한 대상체로부터의 다수의 샘플 (예를 들어, 연속적인 시점에서)을 시험한 값을 인식할 것이다. 연속 샘플의 이러한 시험은 경시적으로 마커 메틸화 상태의 변화를 확인하게 할 수 있다. 메틸화 상태에서 변화뿐만 아니라 메틸화 상태에서 변화의 부재는, 비제한적으로, 질환의 발현으로부터 대략적인 시간, 구제가능한 조직의 존재 및 양, 약물 요법의 적절성, 다양한 요법의 유효성, 및 미래 사건의 위험을 포함하여 대상체의 결과의 동정을 확인하는 것을 포함하는 질환 상태에 대한 유용한 정보를 제공할 수 있다.

[0192] 바이오마커의 분석은 다양한 물리적 포맷으로 수행될 수 있다. 예를 들어, 미세적정 플레이트 또는 자동화의 사용은 다수의 시험 샘플의 처리를 용이하게 하는데 사용될 수 있다. 대안적으로, 예를 들어, 외래 운송이나 응급실 환경설정과 같은 시기 적절한 방식으로 즉각적인 치료 및 진단을 용이하게 하기 위해 단일 샘플 포맷이 개발될 수 있다.

[0193] 일부 구현예에서, 대상체는 대조 메틸화 상태에 비교하여 샘플에서 적어도 하나의 바이오마커의 메틸화 상태에서 측정가능한 차이가 있는 경우 HCC를 갖는 것으로 진단된다. 반대로, 생물학적 샘플에서 메틸화 상태에서의 변화가 확인되지 않는 경우, 대상체는 HCC가 아니거나, HCC에 대한 위험에 있지 않거나 또는 저위험의 HCC를 갖는 것으로 확인될 수 있다. 이와 관련하여, HCC 또는 이의 위험을 갖는 대상체는 낮은 것부터 실질적으로 무 HCC 또는 이의 위험이 있는 대상체와 구별될 수 있다. HCC를 전개할 위험이 있는 이들 대상체는 보다 집중적인 및/또는 규칙적인 검사 일정에 배치될 수 있다.

[0194] 상기에서 언급된 바와 같이, 본 기술의 방법의 구현예에 따라, 하나 이상의 바이오마커의 메틸화 상태에서 변화



를 검출하는 것은 정성적 결정일 수 있거나 또는 이것은 정량적 결정일 수 있다. 이와 같이, HCC를 갖거나 전개할 위험이 있는 것으로 대상체를 진단하는 단계는, 예를 들어, 생물학적 샘플에서 하나 이상의 바이오마커의 메틸화 상태가 사전 결정된 대조군 메틸화 상태로부터 변하는 것과 같은 특정 역치 측정이 이루어짐을 나타낸다. 본 방법의 일부 구현예에서, 대조군 메틸화 상태는 바이오마커의 임의의 검출가능한 메틸화 상태이다. 대조군 샘플이 생물학적 샘플과 동시에 시험되는 방법의 다른 구현예에서, 사전 결정된 메틸화 상태는 대조군 샘플에서의 메틸화 상태이다. 본 방법의 다른 구현예에서, 사전 결정된 메틸화 상태는 표준 곡선에 기초되고 및/또는 이에 의해 확인된다. 본 방법의 다른 구현예에서, 사전 결정된 메틸화 상태는 구체적으로 상태 또는 상태의 범위이다. 이와 같이, 사전 결정된 메틸화 상태는 실시되는 방법 및 원하는 특이성 등의 구현예에 부분적으로 기초하여 당해 분야의 숙련자에게 명백하게 될 허용가능한 한계 내에서 선택될 수 있다.

[0195] 또한 진단 방법과 관련하여, 바람직한 대상체는 척추동물 대상체이다. 바람직한 척추동물은 온혈 동물이고; 바람직한 온혈 척추동물은 포유동물이다. 바람직한 포유동물은 가장 바람직하게는 인간이다. 본 명세서에서 사용된 바와 같은, 용어 “대상체”는 인간 및 동물 대상체 둘 모두를 포함한다. 따라서, 수의과 치료 용도가 본 명세서에 제공된다. 이와 같이, 본 기술은 포유동물 예컨대 인간뿐만 아니라 시베리아 호랑이와 같이 멸종 위기에 기한 중요성; 인간에 의해 소비되기 위해 농장에서 자란 동물과 같은 경제적 중요성의 이들 포유동물; 및/또는 애완 동물이나 동물원으로 보관된 동물과 같이 인간에 사회적으로 중요한 동물의 진단을 제공한다. 이러한 동물의 예는 비제한적으로 하기를 포함한다: 육식동물 예컨대 고양이 및 개; 피그, 호그 및 야생 멧돼지를 포함한 돼지; 반추동물 및/또는 유제류 예컨대 소, 황소, 양, 기린, 사슴, 염소, 들소, 및 낙타; 및 말. 따라서, 비제한적으로, 사육된 돼지, 반추동물, 유제류, 말 (경주 말 포함), 및 기타 동종의 것을 포함한 가축의 진단 및 치료가 또한 제공된다. 현재-개시된 요지는 대상체에서 HCC를 진단하기 위한 시스템을 추가로 포함한다. 본 시스템은 예를 들어, 생물학적 샘플이 수집된 대상체에서 그와 같은 장애의 위험을 스크리닝하기 위해 사용될 수 있는 상업용 키트로서 제공될 수 있다. 본 기술에 따라 제공된 예시적인 시스템은 표 1 및/또는 4에서 제공된 바와 같은 DMR의 메틸화 상태를 평가하는 것을 포함한다.

## [0196] 실시예

### [0197] 실시예 I.

[0198] 본 실시예는 정상 대조군 (예를 들어 간경변증이 있거나 없는 비-HCC 개체)으로부터 유래된 DNA로부터 HCC DNA 샘플을 식별하기 위한 311개 차별적으로 메틸화된 영역 (DMR)의 확인을 기술한다.

[0199] 실험은 4 단계로 수행되었다.

[0200] 첫 번째로, DNA 메틸화 마커 발견은 (간경변증이 있거나 없는) 냉동된 종양 HCC 조직 및 (간경변증이 있거나 없는) 냉동된 정상 간 조직으로부터 추출된 DNA 및 건강한 지원자로부터의 버피 코트 샘플에 대해 RRBS를 사용하여 수행하였다. 판별 차별적으로 메틸화된 영역 (DMR)을 엄격한 여과 기준에 의해 확인하였고 실시간 메틸화 특이적 정량적 PCR 검정 (qMSP) (기술 검증)을 사용하여 결과의 재현성을 보장하기 위해 동일 또는 팽창된 시료 중 어느 하나에 대해 재분석하였다.

[0201] 두 번째로, 후보 마커는 독립적인 기록 케이스 및 대조군 조직으로부터 추출된 DNA에 대한 맹검된 qMSP 검정에 의해 생물학적 검증을 위한 기준을 추가로 선정함으로써 선택하였다.

[0202] 세 번째로, 후보 마커에 대한 서열분석 결과는 pan-GI RRBS 서열분석 세트를 통해 비교하여 메틸화 특이성의 수준을 결정하였다.

[0203] 네 번째로, 대다수의 DNA가 비-간 공급원으로부터 나온 혈액-기재 환경에서 가장 잘 수행할 수 있는 작은 HCC 마커의 세트를 선택하기 위해 별개의 결정 모델을 적용하였다. 선택된 마커는 그런 다음 맹검된 독립적인 혈장 샘플에서 시험하여 임상 배치에서 HCC 검출을 평가하였다.

[0204] 도 1은 이들 4개의 단계를 요약한다.

## [0205] 결과

[0206] 비-편향된 전체의 메틸롬 서열분석을 18개 HCC 및 35개 대조군 (9개 간경변증, 26개 정상 간) 조직에서 추출된 DNA에 대해 수행하였다. 조직 검증을 통해, 최상의 DMR은 메틸화 특이적 PCR을 사용하여 75개 HCC 및 29개 대조군 (16개 간경변증, 13개 정상 간)으로부터 독립적인 조직에서 추출된 DNA에서 확인되었다. 최상부 DMR을 표적으로 하는 맹검된 정량적 대립유전자-특이적 실시간 표적 및 신호 증폭 검정을 그런 다음 21개 HCC 사례 (9개 BCLC [바르셀로나 클리닉 간암 병기 결정] 단계 A, 6개 단계 B, 6개 단계 C) 및 33개 간경변증 대조군을 포함하



는 독립적인 환자 세트로부터의 혈장 DNA에 대해 수행하였다. 재귀 분할 결정 분석은 최상의 DMR 조합을 동정하기 위해 사용하였다. 초기 서열분석은 0.75 초과인 AUC를 갖는 311개 DMR을 동정하였다. 생물학적 검증에 따라, 최상부 12 DMR (ACP1, BDH1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, EFNB2, HOXA1, LRRC4, SPINT2, TSPYL5, CCNJ\_3707, CCNJ\_3124, PFKP, SCRNI, 및 ECE1)을 선택하여 혈장 시험을 수행하였다. 혈장에서 가장 두드러진 마커인, EMX1 단독 만 0.89의 AUC를 가졌다. 상보적 3-마커 조합 (EMX1, LRRC4 및 BDH1)은 혈장에서 20/21 HCC 및 32/33 대조군을 확인하였고; 1개 HCC는 낮은 수준의 BDH1을 가졌고 1개 대조군은 상승된 LRRC4를 가졌다. 패널은 97% 특이성 (95% CI, 82%-100%)에서 HCC에 대해 95% 민감성 (95% CI, 74%-100%)이었고 0.98의 AUC를 달성하였다 (도 2 참고). 생물학적 조직 검증 단계로부터 ACP1, Chr12.133, CLEC11A, DAB2IP, DBNL, EMX1, HOXA1, LRRC4, SPINT2, 및 TSPYL5에 대한 리시버 동작 특징적인 곡선 정보 하의 영역은 도 3A-I에 제공되어 있다.

[0207] 도 4A-CC는 생물학적 조직 검증 데이터로부터 27개 위암 마커 (첨가된 맞춤점 분석으로 29개)의 박스 플롯 (로그 눈금)을 제공한다. 샘플은 가장 왼쪽에 정상적인 간을 배열하고 이어서 간경변증이 없는 HCC, 간경변증이 있는 HCC, 및 간경변증 대조군 (염증성)이 따른다. 수직 축은 분수의 메틸화 ( $\beta$ -액틴 가닥으로 정규화됨)이다.

[0208] 도 5는 95% 정상 특이성에서 매트릭스 형식으로 29개 대조군 (16개 간경변증, 13개 정상 간)과 75개 HCC 조직 샘플에서 27개 HCC 암 마커의 수행을 도시한다. 마커는 수직으로 열거되고 및 수평으로 샘플링된다. 샘플은 가장 왼쪽에 정상적인 간 (NI)을 배열하고 이어서 간경변증이 없는 HCC (HN), 간경변증이 있는 HCC (HC), 및 간경변증 대조군 (In)이 따른다. 양성 히트는 밝은 회색이고 짙은 회색에는 없다. 이 플롯은 마커를 상보적 방식으로 평가될 수 있게 한다. 주석: 2개 마커인, TBX15 및 EGR2가 qMSP 데이터에 대한 맞춤 점 방법을 사용하여 두 번 분석되었으며 여기에 포함된다.

[0209] 표 1은 정상 대조군으로부터 HCC를 구별하는 DMR에 대한 정보를 제공한다.

[0210] 표 2는 표 1로부터 선택된 DMR에 대한 프라이머 정보를 제공한다.

[0211] 표 3은 HCC 대 정상 간 사이의 비교에서 특정 DMR에 대한 AUC 및 배수-변화 정보를 제공하고, 여기서 배수-변화는 대조군의 단편적인 메틸화에 대한 사례의 단편적인 메틸화의 비이다.

표 1.

DMR 번호.	유전자 주석	염색체 번호.	DMR 개시 및 말단 위치
1	Septin9	chr17	75368812-75368887
2	Septin9	chr17	75369228-75369313
3	Septin9	chr17	75368948-75369052
4	ACP1	chr2	264204-264250
5	ACP1	chr2	264087-264151
6	ACTG1	chr17	79481554-79481612
7	AKR1B1	chr7	134143298-134143341
8	AKR1B1	chr7	134143650-134143684
9	AKR1B1	chr7	134143461-134143500
10	ALDH1A3	chr15	101419689-101419707
11	ALDH1A3	chr15	101419988-101420013
12	ANKRD33B	chr5	10564710-10564793
13	ANKRD34B	chr5	79866119-79866169
14	ARHGEF19	chr1	16533241-16533318
15	ARL4C	chr2	235404686-235404738
16	ATL1	chr14	51027344-51027417
17	B3GNT4	chr12	122688498-122688504
18	B3GNT4	chr12	122688516-122688554
19	B3GNT5	chr3	182971710-182971817
20	B3GNT5	chr3	182971523-182971580
21	BAALC	chr8	104153072-104153134
22	BACE2	chr21	42539722-42539743
23	BACE2	chr21	42539915-42539974
24	BLVRB	chr19	40973164-40973202
25	BLVRB	chr19	40973019-40973041
26	BMP6	chr6	7727949-7728013
27	BMP6	chr6	7726324-7726384
28	BVES	chr6	105584689-105584756
29	BVES	chr6	105584565-105584590
30	BVES	chr6	105584763-105584789
31	C6orf126	chr6	35744417-35744467
32	C6orf174	chr6	127837271-127837331
33	C7orf51	chr7	100091317-100091392
34	C7orf57	chr7	48075209-48075261
35	CATSPERG	chr19	38852416-38852441
36	CATSPERG	chr19	38853253-38853336
37	CCDC48	chr3	128745984-128746119
38	CCNI2	chr5	132082945-132083024
39	CCNJ	chr10	97803226-97803325
40	CCNJ	chr10	97803156-97803203
41	CCNJ	chr10	97803124-97803152
42	CCNJ	chr10	97803707-97803733
43	CDKN2A	chr9	21974872-21974890
44	CDKN2A	chr9	21974943-21975018
45	CELF2	chr10	11059845-11059861
46	CELF2	chr10	11059991-11060017

[0212]

47	CELSR3	chr3	48693898-48693974
48	CHST2	chr3	142839074-142839107
49	CHST2	chr3	142839408-142839457
50	CHST2	chr3	142838525-142838580
51	CHST2	chr3	142838038-142838164
52	CHST2	chr3	142838701-142838753
53	CHST2	chr3	142838686-142838696
54	CHST2	chr3	142839245-142839275
55	CHST2	chr3	142838916-142838930
56	CLIP4	chr2	29338189-29338335
57	CORO2B	chr15	68871554-68871567
58	CORO2B	chr15	68871572-68871581
59	CORO2B	chr15	68871623-68871639
60	CORO2B	chr15	68871190-68871224
61	CRHR2	chr7	30722095-30722147
62	DCLK2	chr4	150999533-150999622
63	DCLK2	chr4	150999775-150999832
64	DGCR14	chr22	19137109-19137145
65	DGCR14	chr22	19137067-19137086
66	DKFZp686O24166	chr11	17373279-17373346
67	DNAJB6	chr7	157129265-157129325
68	DUSP4	chr8	29207072-29207112
69	DUSP4	chr8	29207421-29207440
70	ECE1	chr1	21616801-21616877
71	EFNB2	chr13	107188064-107188108
72	EFNB2	chr13	107188196-107188212
73	EFNB2	chr13	107187748-107187766
74	EFNB2	chr13	107187768-107187785
75	EFNB2	chr13	107188435-107188478
76	EFNB2	chr13	107188935-107189019
77	EFNB2	chr13	107188790-107188904
78	EFNB2	chr13	107187702-107187716
79	EFNB2	chr13	107188503-107188558
80	EGR2	chr10	64575092-64575141
81	EHD3	chr2	31456843-31456921
82	EHD3	chr2	31457043-31457061
83	EIF4E3	chr3	71803308-71803340
84	EIF4E3	chr3	71803236-71803290
85	EMILIN2	chr18	2847067-2847090
86	EMILIN2	chr18	2847092-2847100
87	EMILIN2	chr18	2847233-2847274
88	FAHD2B	chr2	97760723-97760782
89	FAM105A	chr5	14582008-14582040
90	FAM105A	chr5	14582046-14582064
91	FAM105A	chr5	14582380-14582417
92	FAM55C	chr3	101498249-101498337
93	FBXL19	chr16	30935481-30935570
94	FCHSD1	chr5	141031071-141031122
95	FCHSD1	chr5	141031057-141031069
96	FHOD1	chr16	67281136-67281198

[0213]

97	FHOD1	chr16	67281037-67281085
98	FIBCD1	chr9	133815197-133815238
99	FIBCD1	chr9	133814914-133814942
100	FOXD1	chr5	72743623-72743647
101	FOXD1	chr5	72743444-72743467
102	FOXD1	chr5	72743469-72743495
103	FUT4	chr11	94277618-94277669
104	FYN	chr6	112194124-112194158
105	FYN	chr6	112194611-112194633
106	GALNT12	chr9	101569940-101569975
107	GALNT12	chr9	101570002-101570029
108	GNG4	chr1	235813128-235813161
109	GNG4	chr1	235813683-235813724
110	GPX7	chr1	53068129-53068171
111	GPX7	chr1	53068016-53068063
112	GPX7	chr1	53068089-53068108
113	GRID2IP	chr7	6543224-6543246
114	GRID2IP	chr7	6570807-6570820
115	GRID2IP	chr7	6543285-6543348
116	GRID2IP	chr7	6570602-6570650
117	HDGFRP3	chr15	83876476-83876501
118	HDGFRP3	chr15	83876552-83876612
119	HDGFRP3	chr15	83876614-83876641
120	HK1	chr10	71078304-71078341
121	HK1	chr10	71078661-71078739
122	HoxA13	chr7	27239880-27239914
123	HoxA13	chr7	27239850-27239876
124	HPDL	chr1	45792443-45792512
125	IKZF1	chr7	50343562-50343603
126	IKZF1	chr7	50343493-50343512
127	IKZF1	chr7	50343859-50343873
128	IKZF1	chr7	50344046-50344106
129	IKZF1	chr7	50343990-50344044
130	ITPR3	chr6	33601477-33601537
131	KCNQ3	chr8	133493278-133493345
132	KCNS2	chr8	99439489-99439524
133	KCNS2	chr8	99439451-99439464
134	KCTD12	chr13	77460548-77460628
135	KCTD12	chr13	77460089-77460122
136	KCTD12	chr13	77459810-77459868
137	KIAA1614	chr1	180881686-180881721
138	KIAA1614	chr1	180881734-180881776
139	LCNL1	chr9	139872612-139872667
140	LGALS3	chr14	55596202-55596277
141	LIMD2	chr17	61777597-61777647
142	LMO2	chr11	33891092-33891223
143	LOC440461	chr17	66195307-66195331
144	LOC440461	chr17	66195475-66195486
145	LOC440461	chr17	66195649-66195686
146	LOC441617	chr11	74953198-74953215

[0214]



147	LOC441617	chr11	74952822-74952901
148	LOC441617	chr11	74953185-74953195
149	LPAR2	chr19	19739243-19739256
150	LPAR2	chr19	19739094-19739141
151	LRRC10B	chr11	61276689-61276765
152	LRRC10B	chr11	61276161-61276167
153	LRRC34	chr3	169530174-169530220
154	LRRC34	chr3	169530317-169530418
155	LRRK1	chr15	101459830-101459847
156	LRRK1	chr15	101459790-101459821
157	LRRN1	chr3	3841285-3841390
158	MATK	chr19	3785888-3785923
159	MATK	chr19	3785942-3785986
160	MAX.chr1.1535351-1535441	chr1	1535351-1535441
161	MAX.chr1.161582155-161582236	chr1	161582155-161582236
162	MAX.chr1.227976136-227976180	chr1	227976136-227976180
163	MAX.chr1.227976189-227976206	chr1	227976189-227976206
164	MAX.chr1.41847970-41848023	chr1	41847970-41848023
165	MAX.chr11.1357749-1357833	chr11	1357749-1357833
166	MAX.chr11.45377012-45377082	chr11	45377012-45377082
167	MAX.chr11.75947388-75947438	chr11	75947388-75947438
168	MAX.chr17.1132739-1132794	chr17	1132739-1132794
169	MAX.chr19.37288291-37288390	chr19	37288291-37288390
170	MAX.chr19.37464150-37464218	chr19	37464150-37464218
171	MAX.chr19.9896775-9896833	chr19	9896775-9896833
172	MAX.chr2.144694740-144694873	chr2	144694740-144694873
173	MAX.chr2.219773668-219773754	chr2	219773668-219773754
174	MAX.chr20.62461671-62461721	chr20	62461671-62461721
175	MAX.chr3.184243258-184243328	chr3	184243258-184243328
176	MAX.chr4.56915281-56915399	chr4	56915281-56915399
177	MAX.chr5.178957711-178957760	chr5	178957711-178957760
178	MAX.chr6.130686855-130686917	chr6	130686855-130686917
179	MAX.chr6.155316859-155316913	chr6	155316859-155316913
180	MAX.chr7.139930482-139930532	chr7	139930482-139930532
181	MAX.chr9.140024004-140024057	chr9	140024004-140024057
182	MAX.chr9.99983940-99983992	chr9	99983940-99983992
183	MECOM	chr3	169380257-169380286
184	MECOM	chr3	169380412-169380427
185	MEX3B	chr15	82339734-82339758
186	MEX3B	chr15	82339894-82339967
187	MGC16703	chr22	21368623-21368674
188	MIR3132	chr2	220417386-220417427
189	MIR3132	chr2	220417233-220417279
190	MIR92B	chr1	155164705-155164747
191	MIR92B	chr1	155164749-155164784
192	MNT	chr17	2297347-2297397
193	N4BP3	chr5	177540980-177541072
194	NCRNA00085	chr19	52207589-52207731
195	NCRNA00085	chr19	52207475-52207527
196	OBSCN	chr1	228401498-228401559

[0215]

197	OBSN	chr1	228463603-228463619
198	ODF2L	chr1	86861515-86861590
199	OVOL1	chr11	65554351-65554410
200	OXT	chr20	3052753-3052809
201	PAQR8	chr6	52227453-52227521
202	PDE4D	chr5	58335632-58335685
203	PFKP	chr10	3110620-3110694
204	PFKP	chr10	3110554-3110609
205	PFKP	chr10	3111174-3111244
206	PFKP	chr10	3110913-3110978
207	PLEKHO1	chr1	150122743-150122798
208	PMAIP1	chr18	57567061-57567089
209	PMAIP1	chr18	57567094-57567115
210	PPP2R5C	chr14	102248080-102248126
211	PPP2R5C	chr14	102248158-102248201
212	PRDM13	chr6	100061630-100061691
213	PRDM5	chr4	121843466-121843518
214	PSMG2	chr18	12658195-12658238
215	PSMG2	chr18	12658247-12658264
216	RANGRF	chr17	8192319-8192372
217	RANGRF	chr17	8192381-8192390
218	RASL11B	chr4	53728280-53728354
219	RASSF2	chr20	4803295-4803348
220	RASSF2	chr20	4803983-4804072
221	RECK	chr9	36037020-36037061
222	RECK	chr9	36037093-36037184
223	RGS10	chr10	121302301-121302309
224	RGS10	chr10	121302871-121302922
225	RGS10	chr10	121302655-121302672
226	RGS20	chr8	54793664-54793736
227	RPP25	chr15	75248658-75248693
228	RPP25	chr15	75249650-75249686
229	SCRN1	chr7	30029222-30029267
230	SCRN1	chr7	30029115-30029160
231	SCRN1	chr7	30029187-30029209
232	SDK1	chr7	3340834-3340842
233	SDK1	chr7	3340543-3340635
234	SDK1	chr7	3340737-3340813
235	SDK1	chr7	3340869-3340900
236	SLC16A3	chr17	80186851-80186883
237	SLC16A3	chr17	80186820-80186835
238	SLC35F1	chr6	118228406-118228451
239	SLC35F1	chr6	118228566-118228578
240	SLC35F1	chr6	118228886-118228910
241	SLC6A20	chr3	45838034-45838084
242	SLC6A6	chr3	14443943-14444020
243	SLC7A5	chr16	87902822-87902863
244	SLC7A5	chr16	87903100-87903127
245	SPNS2	chr17	4403134-4403184
246	ST8SIA1	chr12	22487540-22487711

[0216]

247	ST8SIA6	chr10	17496576-17496610
248	ST8SIA6	chr10	17496378-17496406
249	STK32C	chr10	134121022-134121044
250	STK32C	chr10	134121172-134121183
251	STK32C	chr10	134122104-134122151
252	STK32C	chr10	134122249-134122269
253	STXBP1	chr9	130370839-130370874
254	STXBP1	chr9	130370883-130370925
255	SYCE1L	chr16	77246378-77246432
256	TAF4B	chr18	23806638-23806671
257	TAF4B	chr18	23806718-23806723
258	TCF24	chr8	67875000-67875071
259	TMEM143	chr19	48837211-48837295
260	TMEM163	chr2	135476181-135476192
261	TMEM163	chr2	135476038-135476121
262	TMEM163	chr2	135476286-135476322
263	TRIM17	chr1	228604397-228604446
264	TRPC3	chr4	122872963-122873035
265	TSC22D1	chr13	45150818-45150847
266	TSC22D1	chr13	45150763-45150815
267	UAP1L1	chr9	139972155-139972218
268	VASH1	chr14	77228066-77228133
269	VASH2	chr1	213124546-213124598
270	WDR66	chr12	122356382-122356450
271	WNT7A	chr3	13921575-13921631
272	ZC3HAV1L	chr7	138720665-138720718
273	ZEB2	chr2	145274897-145275040
274	ZNF160	chr19	53606314-53606402
275	ZNF160	chr19	53606256-53606282
276	ZNF160	chr19	53606572-53606659
277	ZNF354C	chr5	178487210-178487287
278	ZNF468	chr19	53360825-53360874
279	ZNF506	chr19	19932506-19932566
280	ZNF510	chr9	99540238-99540295
281	ZNF549	chr19	58038905-58038971
282	ZNF568	chr19	37407242-37407355
283	ZNF607	chr19	38210327-38210438
284	ZNF611	chr19	53238186-53238286
285	ZNF671	chr19	58238870-58238942
286	ZNF816	chr19	53466145-53466206
287	ZNF816	chr19	53466048-53466114
288	SPINT2	chr19	38755180-38755164
289	DBNL	chr7	44080227-44080310
290	PRKAG2	chr7	151329671-151329753
291	TDH	chr8	11205044-11205092
292	COTL1	chr16	84651245-84651314
293	PHF21B	chr22	45403120-45403172
294	PHF21B	chr22	45404843-45404858
295	PHF21B	chr22	45405759-45405766

[0217]

296	MCF2L2	chr3	183146162-183146204
297	MCF2L2	chr3	183146290-183146355
298	MCF2L2	chr3	182897154-182897207
299	AMN	chr14	103394822-103394849
300	MARK1	chr1	220701604-220701619
301	MARK1	chr1	220701582-220701599
302	MARK1	chr1	220701657-220701686
303	EMX1	chr2	73147710-73147772
304	TSPYL5	chr8	98289858-98290220
305	DAB2IP	chr9	124461305-124461420
306	CLEC11A	chr19	51228217-51228732
307	HXA1	chr7	27136145-27136425
308	ADCY1	chr7	45613877-45614572
309	DMRTA2	chr1	50884349-50884499
310	TBX15	chr1	119527066-119527655
311	AK055957	chr12	133484978-133485739

표 2.

마커	OLIGO 유형	서열
TSPYL5	프라이머	GCGCGGGAGGATTTTCG (서열번호:1)
	프라이머	CCGCCACCATAAACGACC (서열번호:2)
	탐침	CCACGGACG CGAAATCGAAAT/3C6/ (서열번호:3)
DAB2IP	프라이머	CGTTCGTTACGTCGTTTTCGT (서열번호:4)
	프라이머	GATCGACGCGACTCGAC (서열번호:5)
	탐침	CCACGGACGCTCGACGTCGCC/3C6/ (서열번호:6)
CLEC11A	프라이머	GCGGGAGTTTGGCGTAG (서열번호:7)
	프라이머	CGCGCAAATACCGAATAAACG (서열번호:8)
	탐침	CCACGGACGGTCGGTAGATCG/3C6/ (서열번호:9)
ACP1	프라이머	GCGCGTTGTTTCGTTTCG (서열번호:10)
	프라이머	CGTCACCTACCGCAAATACG (서열번호:11)
	탐침	CCACGGACGGCGGATAAGGAG/3C6/ (서열번호:12)
BDH1	프라이머	AGTACGTAAGTAGAGCGCG (서열번호:13)
	프라이머	CTAAAATTAACCTACGCCGCCGT (서열번호:14)
	탐침	CCACGGACGGGAGAACGTTTCGA/3C6/ (서열번호:15)
EMX1	프라이머	GGCGTCGCGTTTTTTTAGAGAA (서열번호:16)
	프라이머	TTCTTTTTTCGTTTCGTATAAAATTTTCGTT (서열번호:17)
	탐침	CCACGGACGATCGGGTTTTAG/3C6/ (서열번호:18)
ZF_RASSF1	프라이머	TGCGTATGGTGGGCG AG (서열번호:19)
	프라이머	CCTAATTTACACGTCAACCAATCGAA (서열번호:20)
	탐침	CCACGGACGGCGCGTGCCTTT/3C6/ (서열번호:21)
BTACT	프라이머	TTTGTTTTTTTGATTAGGTGTTTAAAG (서열번호:22)
	프라이머	CACCAACCTCATAACCTTATC (서열번호:23)

[0218]



AK055957 (chr12.133)	탐침	GACGCGGAGATAGTGTGTGG /3C6/ (서열번호:24)
	프라이머	GCGTTTTAGTTAGATAGGGCGG (서열번호:25)
	프라이머	GAAAACCCCTTCCCCGAAAC (서열번호:26)
HOXA1	탐침	CGCCGAGGCGCACGCCTAAA/3C6/ (서열번호:27)
	프라이머	AGTCGTTTTTTTAGGTAGTTTAGGCG (서열번호:28)
	프라이머	CGACCTTTACAATCGCCGC (서열번호:29)
SPINT2	탐침	CGCCGAGGGCGGTAGTTGT/3C6/ (서열번호:30)
	프라이머	GGGAGCGGTGCGCTAG (서열번호:31)
	프라이머	GCACCTAACTAAACAAAACGAACTAAAC (서열번호:32)
DBNL	탐침	CGCCGAGGCGCAAACGCAAA/3C6/ (서열번호:33)
	프라이머	AGGTGGCGCGTATTACG (서열번호:34)
	프라이머	CCTACTAAACGCGCTCAACC (서열번호:35)
EFNB2	탐침	CGCCGAGGCGCTCGATTCCC/3C6/ (서열번호:36)
	프라이머	TTCGATATTGGGTGTCGCG (서열번호:37)
	프라이머	CGCGAAAACCAAAAACGAAAC (서열번호:38)
LRRC4	탐침	CGCCGAGG GAGGCGGGGTTC/3C6/ (서열번호:39)
	프라이머	GCGTTAATTTGCGGAGGTA (서열번호:40)
	프라이머	ACAATACTCTTATATATTAACGCCGCTC (서열번호:41)
	탐침	CGCCGAGGAGGCGACGGAGG/3C6/ (서열번호:42)

표 3.

유전자 주석	곡선하 면적	배수-변화 (정상 간 대비)	배수-변화 (정상 비피 대비)
DAB2IP	0.94	55.04	403.81
TSPYL5	0.94	159.16	648.30
AK055957 (chr12.133)	0.94	166.22	483.57
EMX1	0.90	578.89	154.47
X57SPINT2	0.89	15.73	83.12
X61ACP1	0.89	378.07	17.96
TBX15.fit	0.86	10.02	N/A
X44SCRN1	0.85	5.68	12.94
TBX15	0.85	8.57	41.50
X63HDGFRP3	0.83	2.85	74.06
OPLAH	0.83	2.16	900.80
X49ST8SIA6	0.82	5405378.50	90.32
X52HDGFRP3	0.82	285.68	42.86
CLEC11A	0.82	30.73	51.09
X63DBNL	0.81	3.04	32.20
X65EGR2.fit	0.80	9.83	N/A
HOXA1	0.80	9.13	2746.00
X50VOL1	0.79	17.01	66.53
X65EGR2	0.78	23.71	12.57
X71COTL1	0.77	2407.68	10613.37
X72RASSF2	0.77	4.35	6960.58
X47TDH	0.76	8.55	13.52
X74GALNT12	0.75	16.31	42.20
X45PRKAG2	0.72	20.02	923.17
ST8SIA1	0.71	1.66	56.25
LRRC4	0.71	1.47	433.42
C13orf18	0.68	1.99	27.00
X39ECE1	0.65	3.36	163.46
X67chr4.56915281	0.64	6.39	102.50

이러한 실험은 정상 백혈구 DNA 샘플에서는 아니지만 간 (암 및 정상)에서 메틸화된 89개 간 상피성 DMR의 확인을 추가로 초래했다.

표 4는 정상 백혈구 DNA 샘플에는 아니지만 간 (암 및 정상)에서 메틸화된 간 상피성 DMR에 대한 정보를 제공한

다.

표 4.

DMR 번호	유전자 주석	DMR Chr.번호 및 개시 말단 위치
312	ABHD8	chr19:17403265-17403457
313	ABTB1	chr3:127391081-127391163
314	ADAM8	chr10:135090174-135090246
315	AGAP3	chr7:150812289-150812467
316	AGAP3	chr7:150812127-150812178
317	AMIGO3	chr3:49757070-49757161
318	ANXA2	chr15:60690852-60690949
319	APBB2	chr4:40859188-40859260
320	ATP2B4	chr1:203598701-203598782
321	B3GALT4	chr6:33245156-33245191
322	B3GALT4	chr6:33244921-33245067
323	BDH1	chr3:197282831-197282922
324	BDH1	chr3:197281722-197281827
325	BMP2	chr20:6750738-6750803
326	C17orf64	chr17:58499118-58499182
327	C6orf223	chr6:43970506-43970593
328	CELF6	chr15:72612688-72612773
329	DAP2IP	chr9:124461326-124461415
330	DLEC1	chr3:38080680-38080710
331	EPS8L2	chr11:726256-726367
332	ESPNP	chr1:17027786-17027840
333	F12	chr5:176830943-176831003
334	F12	chr5:176830740-176830823
335	FIGNL2	chr12:52214703-52214931
336	FLJ45983	chr10:8097592-8097679
337	GAL3ST2	chr2:242743108-242743318
338	GRIN2D	chr19:48901811-48901852
339	GSTO2	chr10:106028898-106028963
340	IL17C	chr16:88701236-88701393
341	IL4I1	chr19:50393418-50493491
342	IRF4	chr6:393636-393768
343	JARID2	chr6:15244974-15245009
344	KCNQ4	chr1:41249959-41250042
345	KDM2B	chr12:121904281-121904436
346	LFNG	chr7:2559582-2559607
347	LGALS3	chr14:55595740-55595831
348	LIMD2	chr17:61778042-61778108
349	LOC389333	chr5:138728888-138728935
350	LOC389333	chr5:138728233-138728342
351	LRRC4	chr7:127671917-127672169
352	LRRC4	chr7:127672388-127672445
353	LRRFIP1	chr2:238600061-238600124
354	LTBP4	chr19:41119803-41119889
355	LYL1	chr19:13215369-13215437
356	LYL1	chr19:13210345-13210498
357	MACROD1	chr11:63828351-63828427

[0223]

358	MARVELD1	chr10:99474326-99474382
359	MAX.chr10:22765154-22765214	chr10:22765154-22765214
360	MAX.chr11:518982-519057	chr11:518982-519057
361	MAX.chr14:107253126-107253203	chr14:107253126-107253203
362	MAX.chr20:1784481-1784547	chr20:1784481-1784547
363	MAX.chr6:167763903-167763975	chr6:167763903-167763975
364	MAX.chr8:142215988-142216025	chr8:142215988-142216025
365	MAX.chr8:145104176-145104352	chr8:145104176-145104352
366	MBOAT2	chr2:9144074-9144113
367	MTHFD1L	chr6:151187945-151188021
368	MTHFD2	chr2:74425838-74425898
369	MYPOP	chr19:46405010-46405058
370	NGEF	chr2:2337929980233793066
371	NR2F6	chr19:17346368-17346464
372	NR2F6	chr19:17346575-17346695
373	NTN1	chr17:9143556-9143609
374	NTRK1	chr1:156786617-156786674
375	PDZD7	chr10:102792180-102792249
376	PTK2B	chr8:27183159-27183229
377	PTPRE	chr10:129845681-129845740
378	RASSF1	chr3:50378497-50378540
379	RUNX3	chr1:25256236-25256294
380	S1PR4	chr19:3179884-3179960
381	SBNO2	chr19:1131812-1131869
382	SH3PXD2A	chr10:105453034-105453075
383	SHH	chr7:155597905-155597937
384	SLC25A36	chr3:140661249-140661281
385	SOCS3	chr17:76355477-76355512
386	SP9	chr2:175202178-175202322
387	SPDYA	chr2:29033684-29033774
388	TFR2	chr7:100230996-100231029
389	TIAM1	chr21:32930248-32930318
390	TIAM1	chr21:32931595-32931681
391	TIAM1	chr21:32931280-32931363
392	UCP2	chr11:73693845-73693912
393	VIM	chr10:17271919-17271971
394	WNT1	chr12:49375089-49375117
395	WNT11	chr11:75917494-75917626
396	ZFYVE28	chr4:2415181-2415265
397	ZMIZ1	chr10:81003086-81003157
398	ZMIZ1	chr10:81002084-81002169
399	ZMIZ1	chr10:81002889-81002992
400	ZNF703	chr8:37554906-37554971

[0224]

[0225]

연구 대상체 및 샘플

[0226]

본 연구는 Mayo Clinic Institutional Review Board (미네소타주 로체스터 소재)의 승인을 받았다. 신선한 냉동된 (FF) 조직, 혈장, 및 버피 코트 샘플은 IRB-승인 환자 바이오뱅크에서 제공받았다. 진단을 확정하고 신생물성 세포질을 평가하기 위해 종양 조직 절편을 전문가 GI 병리학자가 재검토했다. 절편을 그 다음 매크로-해부했다. QiaAmp Mini 키트 (Qiagen, 캘리포니아주 발렌시아 소재)를 사용하여 게놈 DNA를 정제하고 후속적으로 AMPure XP 키트 (Beckman Coulter, 캘리포니아주 브레이아 소재)로 재-정제했다.

[0227]

감소된 표현 바이설파이트 서열분석 라이브러리의 준비

[0228]

서열분석 라이브러리는 이전에 공개된 방법의 변형된 버전을 사용하여 준비되었다. 게놈 DNA (300 ng)를 10 U의 MspI로 밤새 단리시켰다. 이 단계 및 후속적인 단계에 사용된 모든 효소는 달리 구체화되지 않는 한 New England Biolabs (NEB)에서 제공되었다. 단편을 말단-보수하고 5 U의 Klenow 단편 (3' -5' exo-)으로 A-제단하고, 그리고 바코드 서열과 보편적으로 메틸화된 시토신을 함유하는 TruSeq 어댑터 (Illumina)에 밤새 결합했다. SYBR GREEN qPCR (라이트사이클러 480 - Roche)을 사용하여 결합 효율 및 단편 품질을 측정했다. 샘플은 바



이설파이트 처리하고 변형된 EpiTect 프로토콜 (Qiagen)을 사용하여 정제 (2회)하고 그 다음 최종 AMPure XP 세척을 수행했다. qPCR을 사용하여 라이브러리 강화를 위한 최적의 PCR 사이클을 결정하였다. 하기 조건을 강화 PCR을 위하여 사용하였다: 각각의 50uL 반응은 5uL의 10X 완충액, 1.25uL의 10 mM 각 데옥시리보뉴클레오타이드 삼인산 (dNTP), 5uL 프라이머 콕테일 (~5uM), 15uL 샘플, 1uL PfuTurbo Cx 핫스타트 및 22.75  $\mu$ L 물을 함유했고; 온도 및 시간은 각각 95C-5분; 98C-30초; 12 내지 16주기의 98C-10초, 65C-30초, 72C-30초, 72C-5분 및 4C 유지였다. 샘플은 PicoGreen 검정 (Molecular Probes)에 의해 정량화하고, 무작위화된 4-plex 라이브러리에 조합하고, 그리고 크기입증을 위해 Bioanalyzer 2100 (Agilent)으로 시험하였다. 추가의 라운드의 AMPure XP 정제/크기 선택을 실험적으로 결정된 완충액 농도에서 수행하여 어댑터 이량체 오염을 최소화하고 350 bp보다 큰 인서트를 제거하였다. phiX 대조군 표준 (Illumina) 및 어댑터-특이적 프라이머를 사용하여 qPCR에 의해 최종 라이브러리 평가를 달성하였다.

[0229] 대량으로 병렬 서열분석 및 생물정보학

[0230] 샘플은 내부 검정 대조군을 위해 예비된 추가의 레인이 있는 무작위화된 레인 배정에 따라 플로우 셀 상에 장입되었다. 서열분석은 Illumina HiSeq 2000 상의 Mayo 클리닉 의료 계층 설비에서 차세대 서열분석 코어에 의해 수행되었다. 관독은 101개 사이클에 대해 단방향이었다. 각각의 플로우 셀 레인은 정렬된 서열에 대해 30-50배 서열분석 깊이의 중앙 적용범위에 대해 충분한 100-120 백만 관독을 생성했다. 표준 일루미나 파이프라인 소프트웨어는 염기를 소환하여 fastq 형식으로 관독을 생성했다. 서열 관독치 평가 및 정리, 참조 계층에 대해 정렬, 메틸화 상태 추출 및 CpG 보고와 주석을 위해 SAAP-RRBS (감소된 표현 바이설파이트 서열분석에 대한 유선형의 분석 및 주석 파이프라인)을 사용하였다. 낮은 적용범위 ( $\leq 10$ )를 갖는 CpG들은 배제했다. 3차 분석은 비-정보제공 또는 낮은 샘플 적용범위 CpG를 제거하는 단계 및 슬라이딩 100 bp 윈도우 내의 낮은 배경과 치밀한 클러스터를 갖는 메틸화된 CpG 영역을 동정하는 단계로 구성되었다. 관독-깊이 기준은 사례와 대조군 사이의 %-메틸화에서의 10% 차이를 검출하기 위해 원하는 통계력에 기초하였다. 통계적 유의성은 관독 횟수에 기초한 DMR 당 메틸화 백분율의 로지스틱 회귀에 의해 결정되었다. 개별 대상체에 걸친 다양한 관독 깊이를 설명하기 위해, 과-분산된 로지스틱 회귀 모델을 사용하였으며, 분산 파라미터는 적합화된 모델로부터 잔기의 Pearson Chi-square 통계를 사용하여 추정되었다. DMR는 그것의 유의 수준에 따라 등급화되어, 대조군 그룹에서 %-메틸화가  $\leq 1\%$ 이었고 암에서  $\geq 10\%$ 인 경우 추가로 고려되었다. 대부분의 장기 부위에서, 이것은 수백 개의 잠재적인 후보를 초래했다. 이용된 추가의 필터는 리시버 동작 특징적인 곡선 아래의 영역, 신호 대 배경 % 메틸화 비 (배수-변화), 및 DMR 전반에 걸친 CpG의 양성 샘플 대 샘플 공동-메틸화 (및 대조군에서 이의 결핍)였다.

[0231] 기술 및 생물학적 조직 검증

[0232] 메틸화 특이적 PCR (MSP) 마커 검정은 상기 열거된 기준에 따라 결정된 바와 같은 - 간 검색 데이터 세트로부터 가장 유망한 DMR 중 30개에 대해 전개되었다. 프라이머는 소프트웨어 (Methprimer - University of California, 캘리포니아주 샌프란시스코 소재, MSP Primer - Johns Hopkins University, 메틸랜드주 볼티모어 소재) 또는 수동으로 설계되었다. 검정은 바이설파이트 전환된 (메틸화된 및 비메틸화된 게놈 DNA), 비전환된, 및 비-템플레이트 대조군 상에서 SYBR GREEN qPCR에 의해 엄격하게 시험되고 최적화되었다. 음성 대조군과 교차-반응한 검정은 재설계되거나 폐기되었다. 또한, 특정 증폭이 일어났는지 보장하기 위해 용융 곡선 분석을 수행하였다. 기술적 검증 단계에서 RRBS 발견에 사용된 것과 동일한 샘플이 qMSP에 의해 재시험되었다. 메틸화 블라인드 되도록 설계된  $\beta$ -액틴 검정은 총 DNA 복제를 나타내는 분자로 사용되었다. 데이터는 로지스틱 회귀 분석과 AUC 및 신호 대 배경 결과를 발견 값과 비교하여 분석했다. 마커의 약 절반 미만이 이하로 수행되어 제거되었다. 나머지 (N=16)는 104개의 독립적인 조직 샘플의 팽창된 세트에서 qMSP에 의해 시험되었다. 또한, 실험은 다른 GI 암 (결장, 식도, 췌장, 담관)에 대한 초기 서열분석 연구에서 확인되고 입증되었고 그리고 강한 다중-장기 암 마커인 11개 메틸화 암 마커를 포함했다. 결과 매트릭스는 AUC 및 배수 변화 비였다 (표 3). 분석된 마커에 대한 박스 플롯 및 상보성 매트릭스는 각각 도 4 및 5에 묘사되어 있다.

[0233] 전반적 장기 검증

[0234] 최상의 메틸화 마커가 간의 외부에서 어떻게 수행되는지 평가하기 위해, 실험은 HCC 샘플 및 서열분석된 초기 - 결장, 췌장, 식도, 및 위인 기타 주요 GI 암에 걸쳐서 검증 DMR에 대한 서열분석 관독을 사용하여 비교 CpG % 메틸화 매트릭스를 구성하였다. 마커의 최종 패널은 1) 생물학적 조직 검증 단계에서의 전반적인 성능 및 2) 다른 암에 걸친 마커의 부위 특이적 특징에 기초하여 혈장에서 시험하기 위해 선택되었다. 혈액 내 HCC를 가장 잘 검출하기 위해, 비-HCC DNA가 과잉인 경우 보편적이고 간 특이적인 암 신호 둘 모두를 나타내는 강력한 12개 마커 패널이 선택되었다. 마커 중 10개는 조직 검증에서 유래되었고; 기이한 간 부위 특이성을 실증한 2개의 추가



마커인 EFN2와 BDH1은 후속적인 조직 검증 없이 RRBS 데이터로부터 직접적으로 설계되고 이용되었다.

[0235] 혈장 검증

[0236] 혈장 DNA는 Exact Sciences에서 개발된 자동화 실리카 비드 방법에 의해 2mL 분획으로부터 추출되었다.

1	2ml의 Te 완충액 (1mM Tris 0.1mM EDTA)
2	0.4 ng/μl의 피쉬 DNA 희석제 내 100μl의 120 cp/ul 제브라피쉬 DNA
3	7ml의 혈장 세포용해 버퍼 (4.3M GTC 10% IGEPAL)
4	2ml의 혈장
5	55°C에서 1시간 동안 인큐베이션한다
6	200μl 결합 비드를 부가한다
7	2.8ml의 100% 이소프로판올을 부가한다
8	30°C에서 30분 동안 인큐베이션한다.
9	비드를 자화시키고 상청액을 제거한다
10	750μl 3M GuHCl 56.8% EtOH을 재현탁된 결합 비드에 부가한다
11	400 RPM에서 2분 동안 웨이킹한다.
12	비드를 결합시키고 상청액을 대기로 발산하여 폐기한다
13	1000μl로 1회 수세하고 (80%ETOH), 30°C에서 3분 동안 인큐베이션하고, 비드를 결합시키고 상청액을 대기로 발산하여 폐기한다
14	500μl로 1회 수세하고 (80%ETOH), 30°C에서 3분 동안 인큐베이션하고, 비드를 결합시키고 상청액을 대기로 발산하여 폐기한다
15	250μl로 1회 수세하고 (80%ETOH), 30°C에서 3분 동안 인큐베이션하고, 비드를 결합시키고 상청액을 대기로 발산하여 폐기한다
16	250μl로 1회 수세하고 (80%ETOH), 30°C에서 3분 동안 인큐베이션하고, 비드를 결합시키고 상청액을 대기로 발산하여 폐기한다
17	웨이킹하면서 70°C에서 15분 건조시킨다.
18	125μl의 Te 완충액 (1mM Tris 0.1mM EDTA)을 부가하고 웨이킹하면서 65°C에서 25분 동안 인큐베이션한다
19	비드를 결합시키고 DNA를 함유하는 상청액을 깨끗한 튜브로 이송한다
20	사용할 때까지 -20°C에서 저장한다.

[0237]

[0238] DNA는 그런 다음 아래 설명된 전매 방법을 사용하여 바이설파이트 전환되고 정제된다.

1	5ul 0.36% BSA
2	70ul의 샘플
3	5ul 1.6N NaOH
4	42°C에서 20' 인큐베이션한다 (변성)
5	8' 동안 냉각시킨다
6	120ul의 암모늄 바이설파이트를 추가한다
7	65°C에서 75' 인큐베이션한다 (전환) (쉐이크 3')
8	750ul의 7M GuHCl을 추가한다
9	50ul 결합 비드를 추가한다
10	쉐이킹하면서 30°C에서 30' 동안 인큐베이션한다
11	비드를 결합시킨다
12	상청액을 대기로 발산시켜 폐기한다
13	1000ul 80% ETOH을 추가한다
14	쉐이킹하면서 30°C에서 3' 동안 인큐베이션한다
15	비드를 결합시킨다
16	상청액을 대기로 발산시켜 폐기한다
17	200ul 디설포화 용액을 분배한다
18	쉐이킹하면서 30°C에서 7' 동안 인큐베이션한다
19	비드를 결합시킨다
20	상청액을 대기로 발산시켜 폐기한다
21	250ul 80% ETOH을 추가한다
22	쉐이킹하면서 30°C에서 3' 동안 인큐베이션한다
23	비드를 결합시킨다
24	상청액을 대기로 발산시켜 폐기한다
25	쉐이킹하면서 70°C에서 15' 동안 비드를 건조시킨다
26	80ul의 Te 완충액 (1mM Tris 0.1mM EDTA)을 추가한다
27	쉐이킹하면서 25분 동안 65°C 인큐베이션한다
28	비드를 결합시키고 DNA를 함유하는 상청액을 깨끗한 튜브로 이송한다
29	사용할 때까지 -20°C에서 저장한다.

[0239]

[0240] 샘플은 실시간 PCR 기기 (Roche LC480) 상에서 다음과 같이 수행되었다: QuART 형식 (참고, 미국특허번호 8,361,720), DMR 서열(표 2 참고)로부터 전개된 프라이머 및 탐침, GoTaq DNA 중합효소 (Promega), Cleavase II (Hologic), 및 FAM, HEX, 및 Quasar 670 염료 (Biosearch Technologies)를 함유하는 형광 공명 에너지 전달 리포터 카세트 (FRET)를 사용함.

[0241] 도 6은 QuART (정량적 대립유전자-특이적 실시간 표적 및 신호 증폭) 검정에 의한 메틸화된 DNA 서명의 검출에 사용되는 FRET 카세트에 대한 올리고뉴클레오타이드 서열을 제공한다. 각각의 FRET 서열은 3개의 별개의 검정으로 함께 다중화될 수 있는 형광단 및 소광인자를 포함한다.

[0242] 관심 있는 마커 서열을 함유하는 플라스미드는 Genscript로부터 획득되었고 15 ul 반응당 1 복제의 명목 농도로 1X QuART 시약에서 희석하였다. 반응 혼합물은 각각 384개 각 웰에 분주하고, 라이트사이클러 상에서 45 사이클 동안 사이클링하고, 데이터를 수집하였다. 웰들은 샘플을 함유하거나 함유하지 않는 것 어느 하나로 불렀다. 푸와송 랜덤 가변성은 1로 설정되었고 평균 성공률 값은 시행 착오에 의해 입력되어 그 값에 대한 누적 개연성을 계산하는 데 사용되었다. 누적 개연성이 신호가 있는 웰들의 퍼센트와 같을 때, 이 경우에 복제수인, 정확한 평균 성공률이 밝혀졌다. 이들 플라스미드는 희석되고 검정 표준으로 사용되었다.

[0243] QuARTs-X (참고, 미국가특허번호 62/249,097) 상기는 11 사이클의 증폭을 당한 최대 12개 표적에 대해 프라이머

로 수행된 샘플의 사전-증폭 플레이트를 먼저 형성함에 의해 수행된다. 이 생성물은 그 다음 1: 9로 희석되고 삼중 반응에서 3개의 표적만을 함유하는 후속적인 QuART 반응에 대한 템플레이트로 사용된다. 가닥 수를 계산하기 위해 사용된 표준은 사전-증폭을 거치지 않는다. 표준이 아닌 샘플을 사전-증폭함으로써 검정의 민감성이 증가된다.

[0244] 결과는 회귀적 분할 (rPart)에 의해 분석되었다. 다중 메틸화 마커를 단일 위험 스코어에 결합시키는 로지스틱 회귀분석을 사용하는 것이 표준기술이다. 그러나, 로지스틱 모델 내에서 마커들 사이의 고차원 상호작용을 발견하고 및/또는 모델링하는 것은 어렵다. 이것은 그러한 영향이 존재할 때 당해 마커 패널의 예측 능력을 제한한다. 회귀 분할 트리 (rPart)는 마커 패널의 예측 정확도를 극대화하기 위해 그와 같은 방식으로 마커 간의 고차원 상호작용을 발견할 수 있는 결정 트리 접근법이다. rPart로, 최상부 3개 마커 조합 (EMX1, BDH1, LRRC4)에 대한 HCC 혈액 샘플의 감수성 및 특이성은 각각 97% 및 95%였다. (도 7A) 3개의 대안적인 마커 조합이 모델링되었다:

[0245] 대안 #1 (EMX1, DAB2IP, TSPYL5): 특이성=100% 민감성=90% (도 7B)

[0246] 대안 #2 (EMX1, HOXA1, ACP1): 특이성=88% 민감성=100% (도 7C)

[0247] 대안 #3 (EMX1, EFNB2, SPINT2): 특이성=100% 민감성=90% (도 7D)

[0248] 혈장에서 최상의 수행성 단일 마커인, EMX1은 100% 특이성에서 77% 민감성으로 0.89의 AUC를 가졌다. EMX1에 대한 신호는 단계적으로 증가하는 더 높은 베타 액틴 정규화된 신호를 나타냈다. (도 8).

[0249] **실시예 II.**

[0250] 조직 샘플 (75개 HCC, 20개 간경변증, 및 30개 정상)은 실시간 PCR 기기 (Roche LC480) 상에서 다음과 같이 수행되었다: QuART 형식 (참고, 미국특허번호 8,361,720), DMR 서열(표 5 참고)로부터 전개된 프라이머 및 탐침, GoTaq DNA 중합효소 (Promega), Cleavase 2.0 (Hologic), 및 FAM, HEX, 및 Quasar 670 염료 (Biosearch Technologies)를 함유하는 형광 공명 에너지 전달 리포터 카세트 (FRET)를 사용함. 표 6은 100% 민감성에서 간경변증 및 정상으로부터 HCC를 구별하는 각각의 마커의 능력을 나타낸다.

표 5.

유전자	정방향 QuARTS 프라이머 (5'-3')	역방향 QuARTS 프라이머 (5'-3')	탐침
TSPYL5	TTTGTTCGGTTTTTGGCG (서열번호:43)	ACCATAAACGACCGAAATCGA (서열번호:44)	CCACGGACG GCGGGAGGA TTT /3C6/ (서열번호:45)
DAB2IP	CGTTCGTTACGTCGTTTTCTG (서열번호:46)	GATCGACGCGACTCGAC (서열번호:47)	CCACGGACG CTCGACGTC GCC/3C6/ (서열번호:48)
Chr12_1 33_V2	GCGTTTTAGTTAGATAGGGCGG (서열번호:49)	GAAAACCCCTTCCCCGAAAC (서열번호:50)	CGCCGAGG CGCACGCCTA AA/3C6/ (서열번호:51)
CLEC11 A	GCGGGAGTTTGGCGTAG (서열번호:52)	CGCGCAAATACCGAATAAACG (서열번호:53)	CCACGGACG-GTCGGTAGA TCG/3C6/ (서열번호:54)
SPINT2	GGGAGCGGTCGCGTAG (서열번호:55)	GCACCTAACTAAACAAAACGA ACTAAC (서열번호:56)	CGCCGAGG CGCAAACGCAA/3C6/ (서열번호:57)
ACP1	GCGCGTTGTTTCGTTTCG (서열번호:58)	CGTCACCTACCGCAAATACG (서열번호:59)	CCACGGACG GCGGATAAG GAG/3C6/ (서열번호:60)
BDH1	AGTACGTAAGTAGAGCGCG (서열번호:61)	CTAAAATTAACACTACGCCCGCT (서열번호:62)	CCACGGACG GAGAACGTT CGA/3C6/ (서열번호:63)
EFNB2	CGTTCGTTCTGTTATTTTTTCG A (서열번호:64)	GCCCGCGAAAACCAAAAC (서열번호:65)	CGCCGAGG CGAAACTCAC CT/3C6/ (서열번호:66)
CCNJ_37 07	GCGTTTTTTTTAGCGGGGTTA (서열번호:67)	CCGAACTAAAATTCTCCCGC (서열번호:68)	CGCCGAGG ATGAGCGTGT TA/3C6/ (서열번호:69)
CCNJ_31 24	CGTTTTTCGTTTGGGTACG (서열번호:70)	CCAACCCAAACACGCC (서열번호:71)	CCACGGACG CGCGCCGTA CGA/3C6/ (서열번호:72)
PFKP	GGAGGTTGGCGGGGAG (서열번호:73)	CATATACTATCGCCTTCGACT C (서열번호:74)	CGCCGAGG CGTAACAAAA AC/3C6/ (서열번호:75)
SCRN1	TCGTTTTAGGTGAGTCGCG (서열번호:76)	AAATAAACCGCCGAAAAAAC AAC (서열번호:77)	CCACGGACG CGCTACAAA CGC/3C6/ (서열번호:78)
ECE1	GGAGGGGTTTCGTTCTG (서열번호:79)	CTACTATCGACGCTAAAAATAA AACGAAC (서열번호:80)	CGCCGAGG CGCGACCTAA AA/3C6/ (서열번호:81)
DBNL	AGGTGGCGCGTATTACG	CCTACTAAACGCGCTCAACC	CGCCGAGG CGCTCGATTC

	(서열번호:82)	(서열번호:83)	CC/3C6/ (서열번호:84)
LRRC4	GCGTTAATTCGCGAGGTA (서열번호:85)	ACAATACTCTTATATATTAACG CCGCTC (서열번호:86)	CGCCGAGG AGGCGACGGA GG/3C6/ (서열번호:87)
HOXA1	AGTCGTTTTTTAGGTAGTTAG GCG (서열번호:88)	CGACCTTACAATCGCCGC (서열번호:89)	CGCCGAGG GGCGGTAGTT GT/3C6/ (서열번호:90)
EMX1	GGCGTCGCGTTTTTAGAGAA (서열번호:91)	TTCCTTTTCGTTCTGATAAAATT TCGT (서열번호:92)	CCACGGACG ATCGGGTTTT AG/3C6/ (서열번호:93) 또는 CGCCGAGG ATCGGGTTTTA G/3C6/ (서열번호:94)



표 6.

유전자	100% 특이성에서 간경변증 또는 정상으로부터 HCC를 식별하는 민감성
TSPYL5	79%
DAB2IP	91%
Chr12_133_V2	91%
CLEC11A	69%
SPINT2	68%
ACP1	84%
BDH1	8%
EFNB2	45%
CCNJ_3707	53%
CCNJ_3124	49%
PFKP	57%
SCRN1	51%
ECE1	49%
EMX1	67%
LRRC4	20%
HOXA1	52%

[0253]

[0254]

### 실시예 III.

[0255]

이의 주요 목표는 간세포 암종 (HCC)을 예측하는 마커의 패널을 결정하는 것이 었다. 244명 대상체 (간세포 암 95명과 대조군 149명)로부터의 혈장을 2mL로 조정하고 추출하였다. 대조군 149명은 51명의 간경변증과 98명의 정상 환자로 구성되었다.

[0256]

도 9는 이 분석에서 고려된 각각의 메틸화 마커의 상대적 중요성을 묘사한다. 전체 마커의 패널에 대한 민감성과 특이성의 교차 입증된 추정치는 각각 75%와 96 %였다.

[0257]

대조군 대 정상의 88.6% 특이성에서, 하기 마커 패널 (Chr12.133, CLEC11A, EMX1, HOXA1, CCNJ\_3707)은 HCC에 대해 85.3% 민감성을 수득했다 (표 7).

표 7.

	총 환자	예측		민감성
		음성	양성	
HCC	95	14	81	85.3%
대조군	149	132	17	

[0258]

[0259]

동일한 패널 (Chr12.133, CLEC11A, EMX1, HOXA1, CCNJ\_3707)은 정상 및 간경변증 (대조군) 환자에 대해 하기 명세의 특이성을 수득했다 (표 8).

표 8.

환자 군	총 환자	예측		특이성
		음성	양성	
간경변증	51	41	10	80.4%
정상	98	91	7	92.9%
총	149	132	17	88.6%

[0260]

[0261]

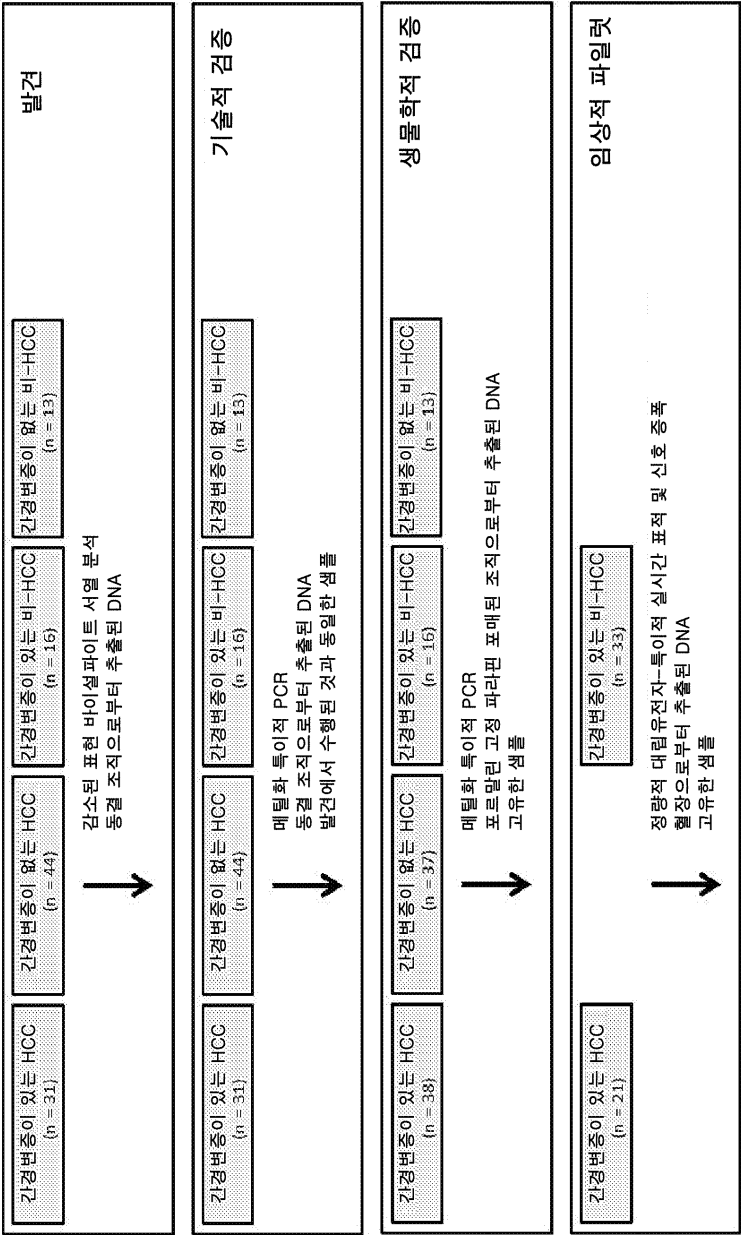
4 mL 혈장 샘플로부터 DNA를 분리하는 예시적인 절차는 아래와 같이 수행될 것이다:

- [0262] ● 2 mL 혈장 샘플에, 300  $\mu$ L의 프로테이나제 K (20mg/mL)를 첨가하고 혼합한다. 샘플이 2 mL 미만의 혈장인 경우, 10 mM 트리스-HCl, 0.1 mM EDTA 용액을 부가하여 2 mL로 조정한다
- [0263] ● 혈장에 6 mL의 혈장 세포용해 버퍼 1을 부가하고 실온에서 혼합한다
- [0264] 혈장 세포용해 버퍼는 다음과 같다:
- [0265] - 4.3M 구아니딘티오시아네이트
- [0266] - 10% IGEPAL CA-630 (옥틸페녹시폴리(에틸렌옥시)에탄올, 분지형)
- [0267] (45 mL의 4.8 M 구아니딘티오시아네이트와 조합된 5.3g의 IGEPAL CA-630)
- [0268] ● 200  $\mu$ L 자기 실리카 결합 비드를 부가하고 [16  $\mu$ g의 비드/ $\mu$ L] 다시 혼합한다.
- [0269] ● 7 mL의 세포용해 버퍼 2를 튜브에 부가한다.
- [0270] 혈장 세포용해 버퍼 2는 40% 이소프로판올과 60%의 세포용해 버퍼 1의 비로 혼합함에 의해 제조된다.
- [0271] ● 샘플을 세포 용해버퍼 2와 60분 동안 혼합한다
- [0272] ● 자석 상에 튜브를 놓고 비드가 10분 동안 수집되게 한다. 상청액을 대기로 발산하여 폐기한다.
- [0273] ● 1000  $\mu$ L 세정 완충액 (10 mM Tris HCl, 80% EtOH)을 비드에 부가하고, 그리고 웨이킹하면서 3분 동안 30° C에서 인큐베이션한다.
- [0274] ● 자석 상에 튜브를 놓고 비드가 수집되게 한다. 상청액을 대기로 발산하여 폐기한다.
- [0275] ● 500  $\mu$ L 세정 완충액을 비드에 부가하고 그리고 웨이킹하면서 3분 동안 30° C에서 인큐베이션한다.
- [0276] ● 자석 상에 튜브를 놓고 비드가 수집되게 한다. 상청액을 대기로 발산하여 폐기한다.
- [0277] ● 250  $\mu$ L 세정 완충액을 부가하고 그리고 웨이킹하면서 3분 동안 30° C에서 인큐베이션한다.
- [0278] ● 자석 상에 튜브를 놓고 비드가 수집되게 한다. 잔여 완충액을 대기로 발산하여 폐기한다.
- [0279] ● 250  $\mu$ L 세정 완충액을 부가하고 그리고 웨이킹하면서 3분 동안 30° C에서 인큐베이션한다.
- [0280] ● 자석 상에 튜브를 놓고 비드가 수집되게 한다. 잔여 완충액을 대기로 발산하여 폐기한다.
- [0281] ● 웨이킹하면서 70° C에서 15분 동안 비드를 건조시킨다.
- [0282] ● 125  $\mu$ L 용출 완충액 (10 mM Tris HCl, pH 8.0, 0.1 mM EDTA)을 비드에 부가하고 웨이킹하면서 65° C에서 25분 동안 인큐베이션한다.
- [0283] ● 자석 상에 튜브를 놓고 비드가 10분 동안 수집되게 한다.
- [0284] ● DNA를 함유하는 상청액을 대기로 발산하고 신규한 용기 또는 튜브로 이송한다.
- [0285] 바이설파이트 전환
- [0286] *I. 암모늄 수소 설파이트를 사용한 DNA의 설폰화*
- [0287] 1. 각각의 튜브에, 64  $\mu$ L DNA, 7  $\mu$ L 1 N NaOH, 및 0.2 mg/mL BSA와 0.25 mg/mL의 피쉬 DNA를 함유하는 9  $\mu$ L의 담체 용액을 조합한다.
- [0288] 2. 42° C에서 20분 동안 인큐베이션한다.
- [0289] 3. 120  $\mu$ L의 45% 암모늄 수소 설파이트를 부가하고 66° 에서 75분 동안 인큐베이션한다.
- [0290] 4. 4° C에서 10분 동안 인큐베이션한다.
- [0291] *II. 자기 비드를 사용한 탈설폰화*
- [0292] 물질
- [0293] 자기 비드 (Promega MagneSil Paramagnetic Particles, Promega 카탈로그 번호 AS1050, 16  $\mu$ g/ $\mu$ L).
- [0294] 결합 완충액: 6.5-7 M 염산 구아니딘.

- [0295] 후-전환 세정 완충액: 10 mM Tris HCl을 갖는 80% 에탄올 (pH 8.0).
- [0296] 탈설폰화 완충액: 70% 이소프로필알코올, 0.1 N NaOH이 탈설폰화 완충액에 대해 선택되었다.
- [0297] 샘플을 본질적으로 아래에 기재된 바와 같은 온도 및 혼합 속도에서 샘플을 혼합하거나 인큐베이션하기 위해 임의의 적절한 디바이스 또는 기술을 사용하여 혼합하였다. 예를 들어, 샘플의 혼합 또는 인큐베이션을 위해 써모믹서 (에펜도르프)가 사용될 수 있다. 예시적인 탈설폰화는 아래와 같다:
- [0298] 1. 1분 동안 와동하는 병으로 비드 스톱을 철저하게 혼합한다.
- [0299] 2. (예를 들어, USA Scientific으로부터의) 2.0 mL 튜브 안에 50  $\mu$ L의 비드를 분취한다.
- [0300] 3. 750  $\mu$ L의 결합 완충액을 비드에 부가한다.
- [0301] 4. 단계 I로부터의 150  $\mu$ L의 설폰화된 DNA를 부가한다.
- [0302] 5. 혼합한다 (예를 들어, 30° C에서 30분 동안 1000 RPM).
- [0303] 6. 자석 상에 튜브를 위치시키고 5분 동안 원위치에 정치한다. 튜브를 정치하면서, 상청액을 제거하고 폐기한다.
- [0304] 7. 1,000  $\mu$ L의 세정 완충액을 부가한다. 혼합한다 (예를 들어, 30° C에서 3분 동안 1000 RPM).
- [0305] 8. 자석 상에 튜브를 위치시키고 5분 동안 원위치에 정치한다. 튜브를 정치하면서, 상청액을 제거하고 폐기한다.
- [0306] 9. 250  $\mu$ L의 세정 완충액을 부가한다. 혼합한다 (예를 들어, 30° C에서 3분 동안 1000 RPM).
- [0307] 10. 자기 랙 상에 튜브 위치시키고; 1분 후 상청액을 제거하고 폐기한다.
- [0308] 11. 200  $\mu$ L의 탈설폰화 완충액을 부가한다. 혼합한다 (예를 들어, 30° C에서 5분 동안 1000 RPM).
- [0309] 12. 자기 랙 상에 튜브 위치시키고; 1분 후 상청액을 제거하고 폐기한다.
- [0310] 13. 250  $\mu$ L의 세정 완충액을 부가한다. 혼합한다 (예를 들어, 30° C에서 3분 동안 1000 RPM).
- [0311] 14. 자기 랙 상에 튜브 위치시키고; 1분 후 상청액을 제거하고 폐기한다.
- [0312] 15. 250  $\mu$ L의 세정 완충액을 튜브에 부가한다. 혼합한다 (예를 들어, 30° C에서 3분 동안 1000 RPM).
- [0313] 16. 자기 랙 상에 튜브 위치시키고; 1분 후 상청액을 제거하고 폐기한다.
- [0314] 17. 15분 동안 뚜껑 개방으로 30° C에서 모든 튜브를 인큐베이션한다.
- [0315] 18. 자기 랙으로부터 튜브를 제거하고 비드에 직접적으로 70  $\mu$ L의 용출 완충액을 부가한다.
- [0316] 19. 용출-완충액으로 비드를 인큐베이션한다 (예를 들어, 40° C에서 45분 동안 1000 RPM).
- [0317] 20. 약 1분 동안 자기 랙 상에 튜브 위치시키고; 상청액을 제거하고 저장한다.
- [0318] 전환된 DNA는 그 다음 사전-증폭 및/또는 플랩 엔도뉴클레아제 검정에 사용된다.
- [0319] 상기 명세서에 언급된 모든 공보 및 특허는 모든 목적을 위해 전체적으로 본 명세서에 참고로 편입된다. 기재된 바와 같은 기술의 범위 및 사상을 벗어남이 없이 본 기술의 기재된 조성물, 방법, 및 사용의 다양한 변형 및 변화는 당해 분야의 숙련자에게 분명할 것이다. 본 기술은 특정 예시적인 구현예와 관련하여 기재되었지만, 청구된 바와 같은 본 발명은 그러한 특정 구현예에 지나치게 제한되어서는 안된다는 것을 이해해야 한다. 사실상, 약리학, 생화학, 의학 또는 관련 분야의 숙련자에게 명백한 본 발명을 수행하기 위한 기재된 방식의 다양한 변형은 하기 청구범위의 범위 내에 있는 것으로 의도된다.

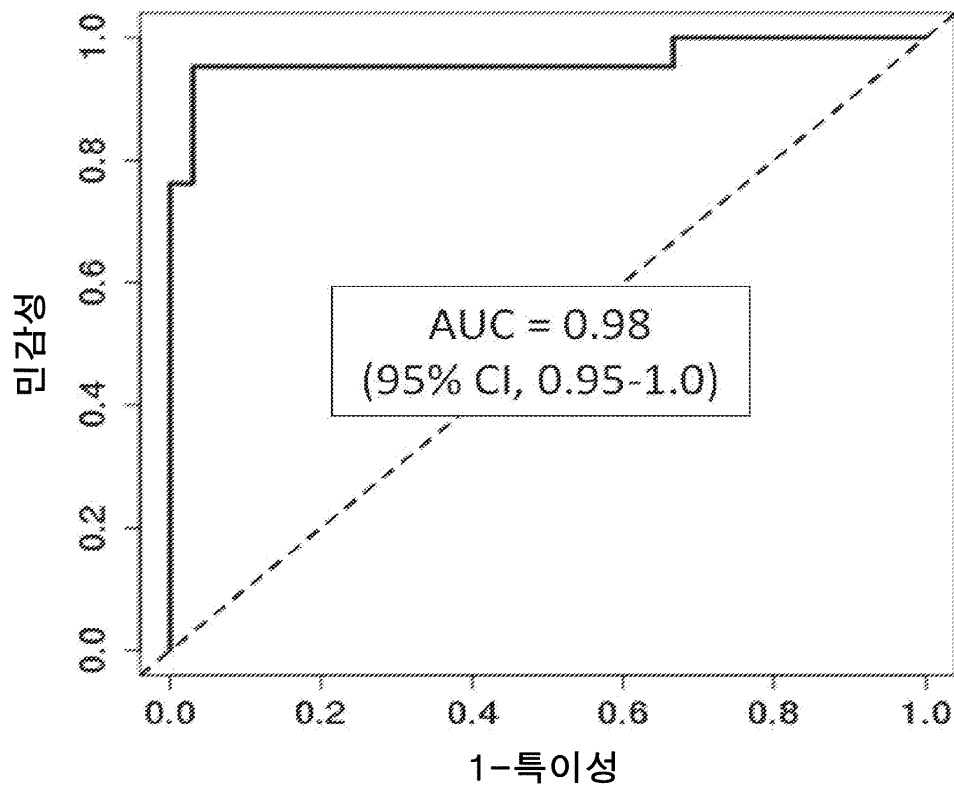
도면

도면1

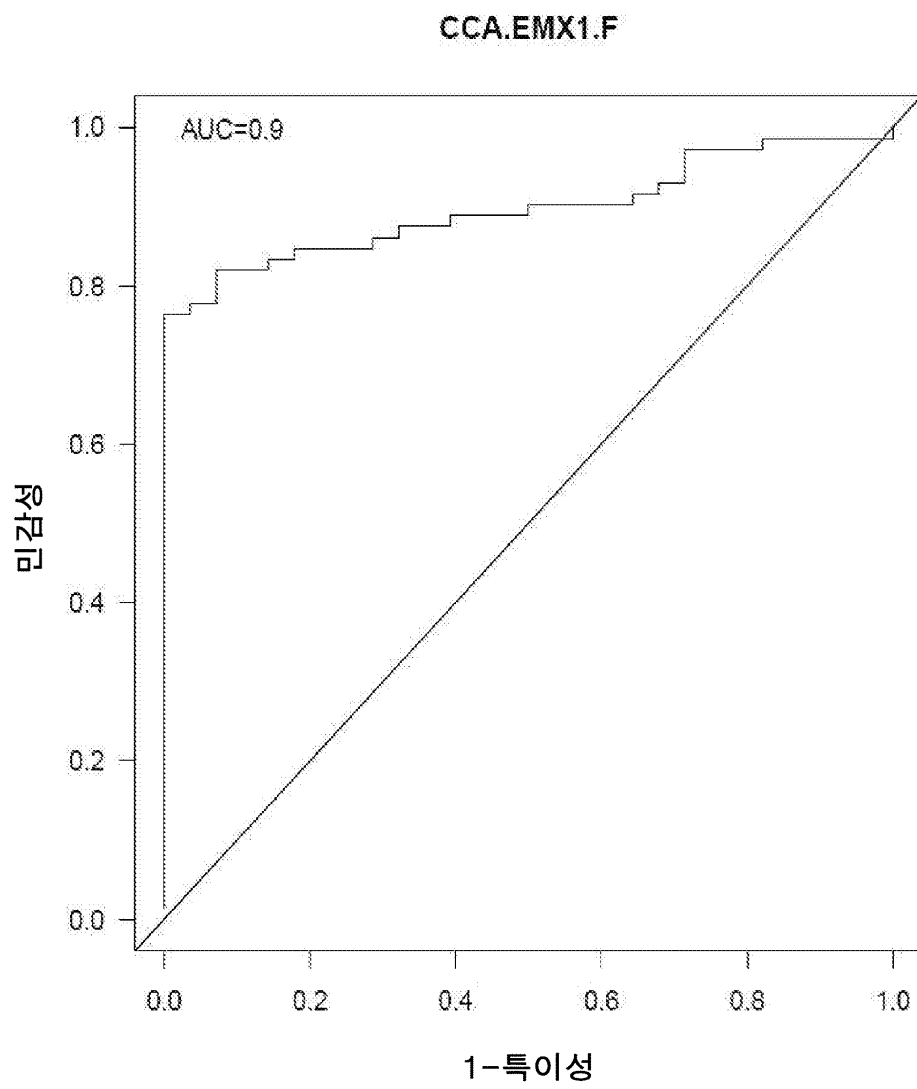




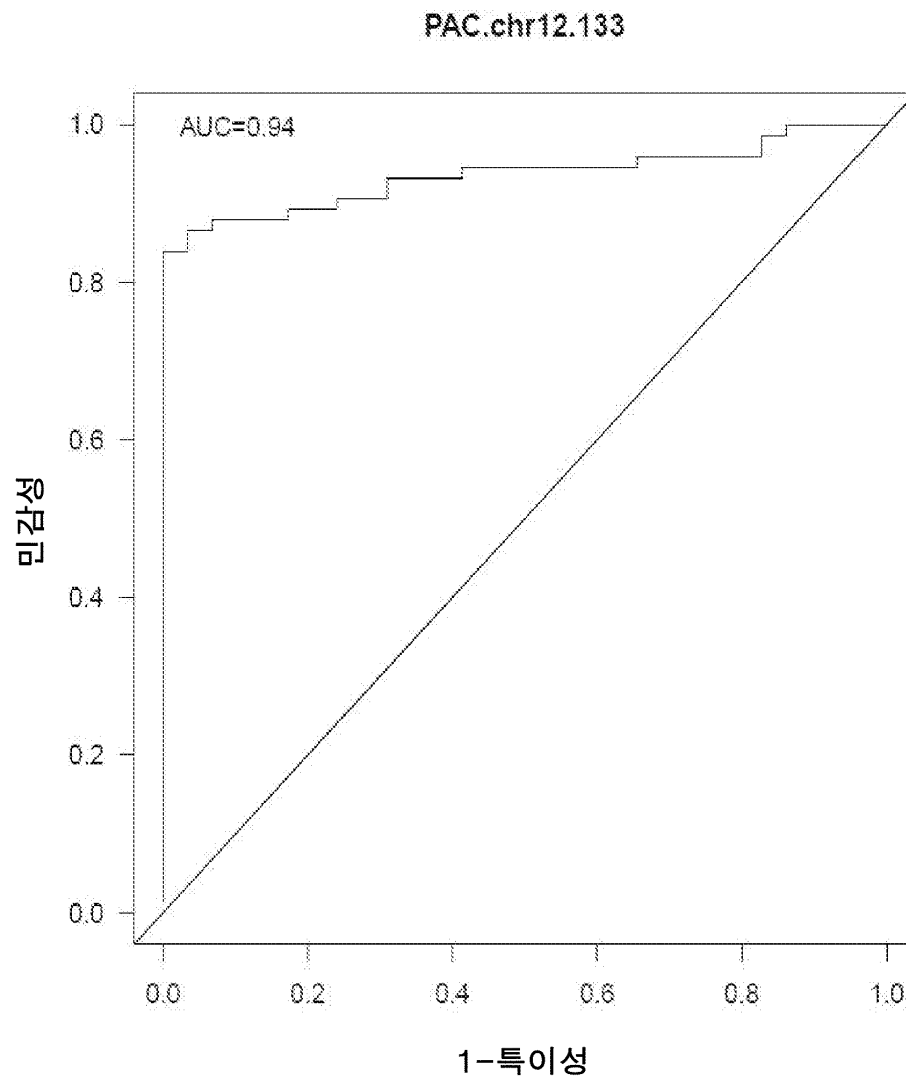
도면2



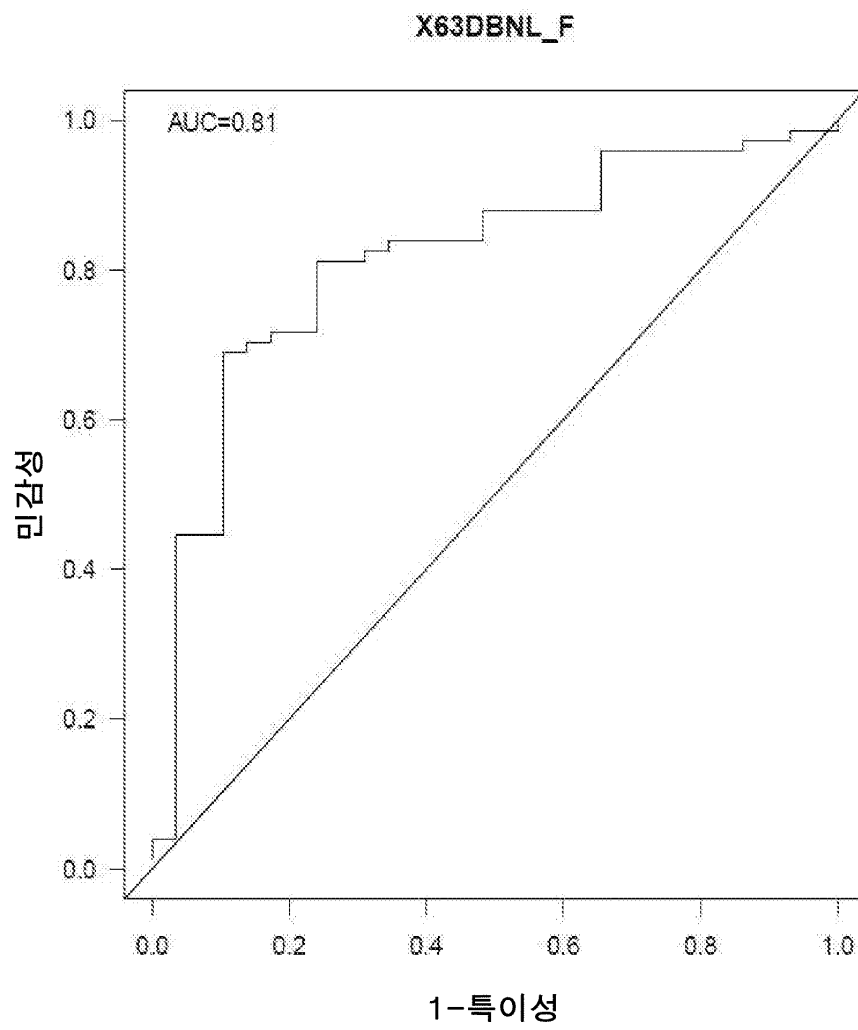
도면3a



도면3b

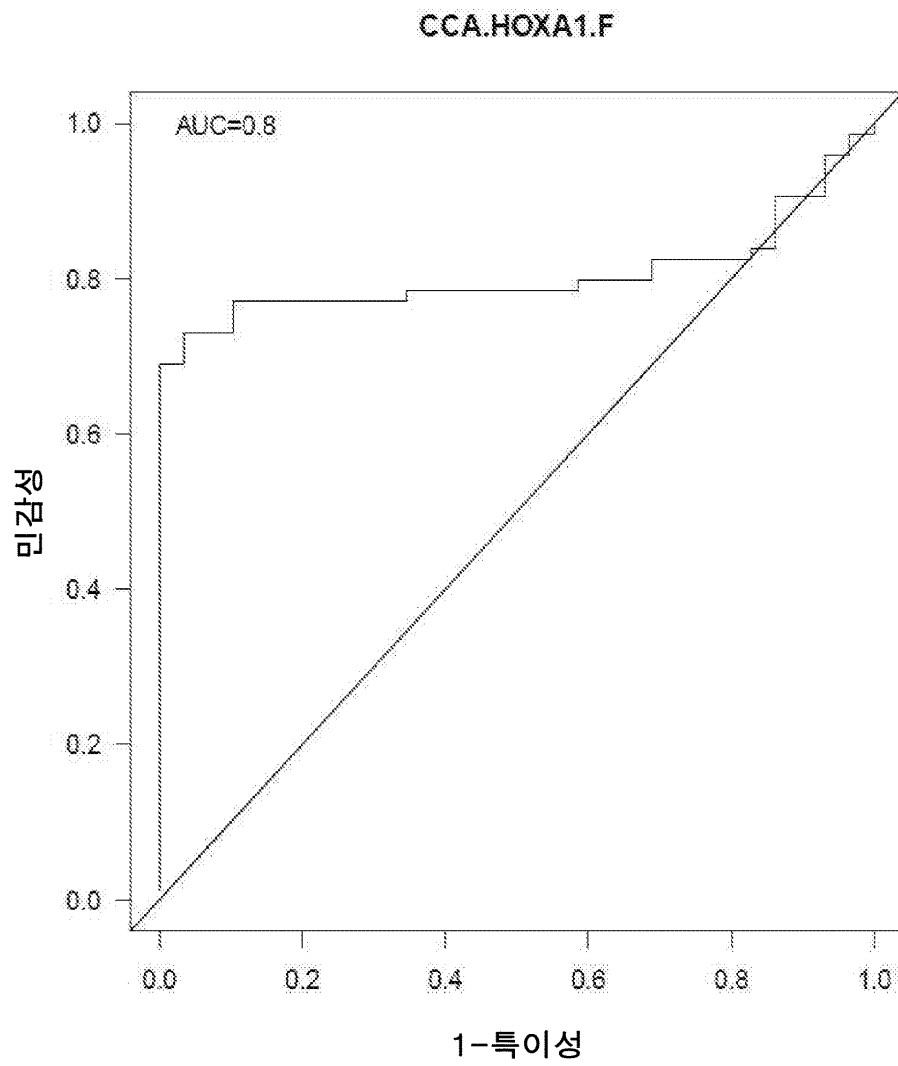


도면3c

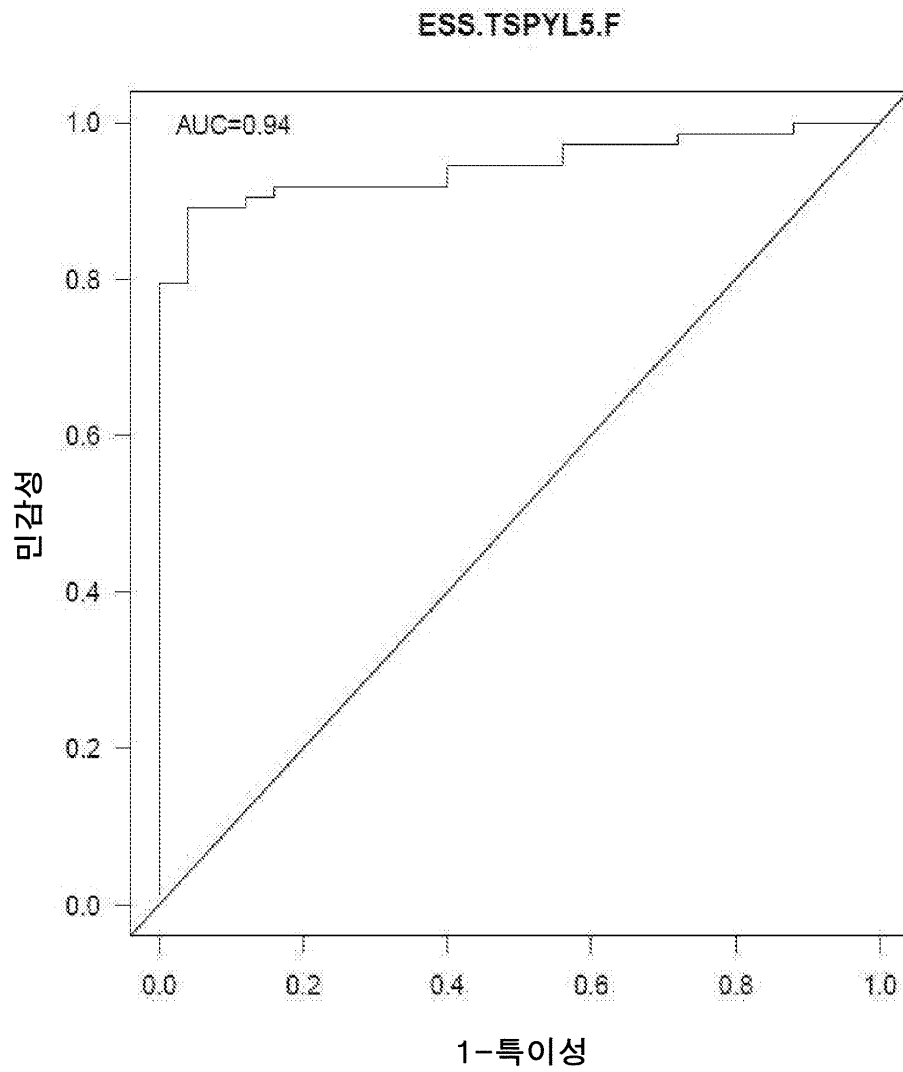




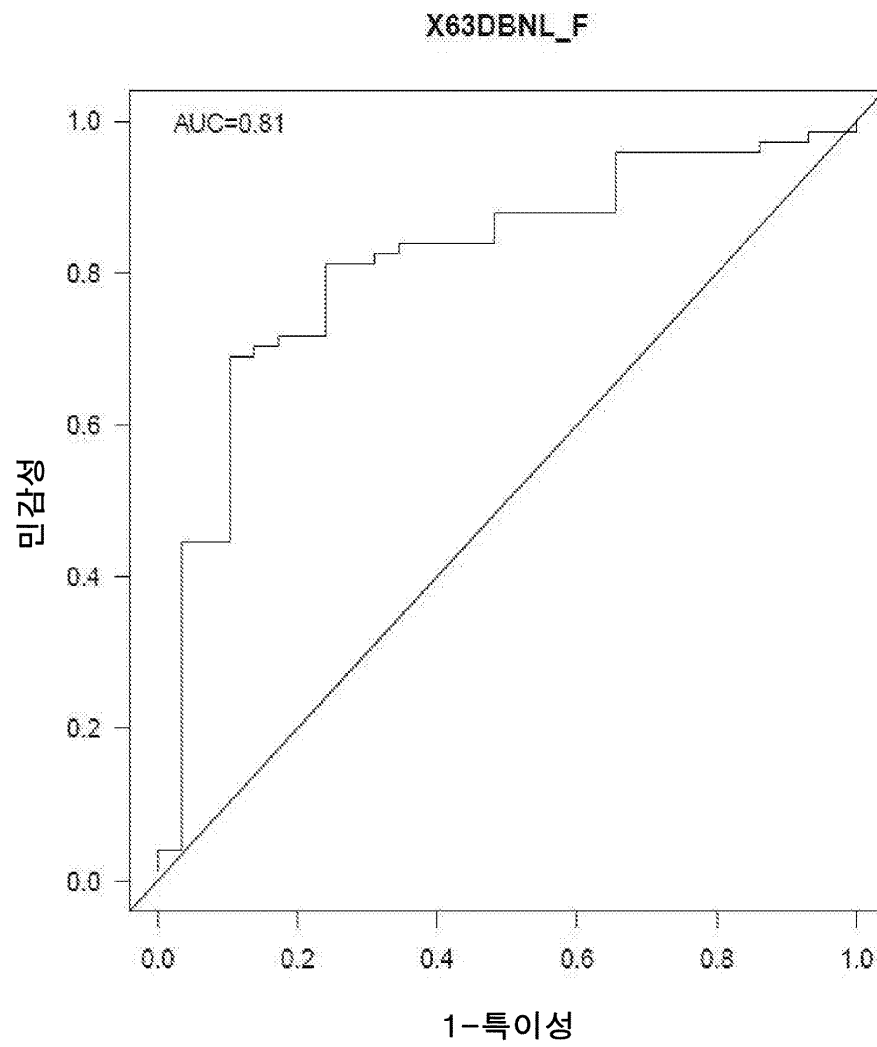
도면3d



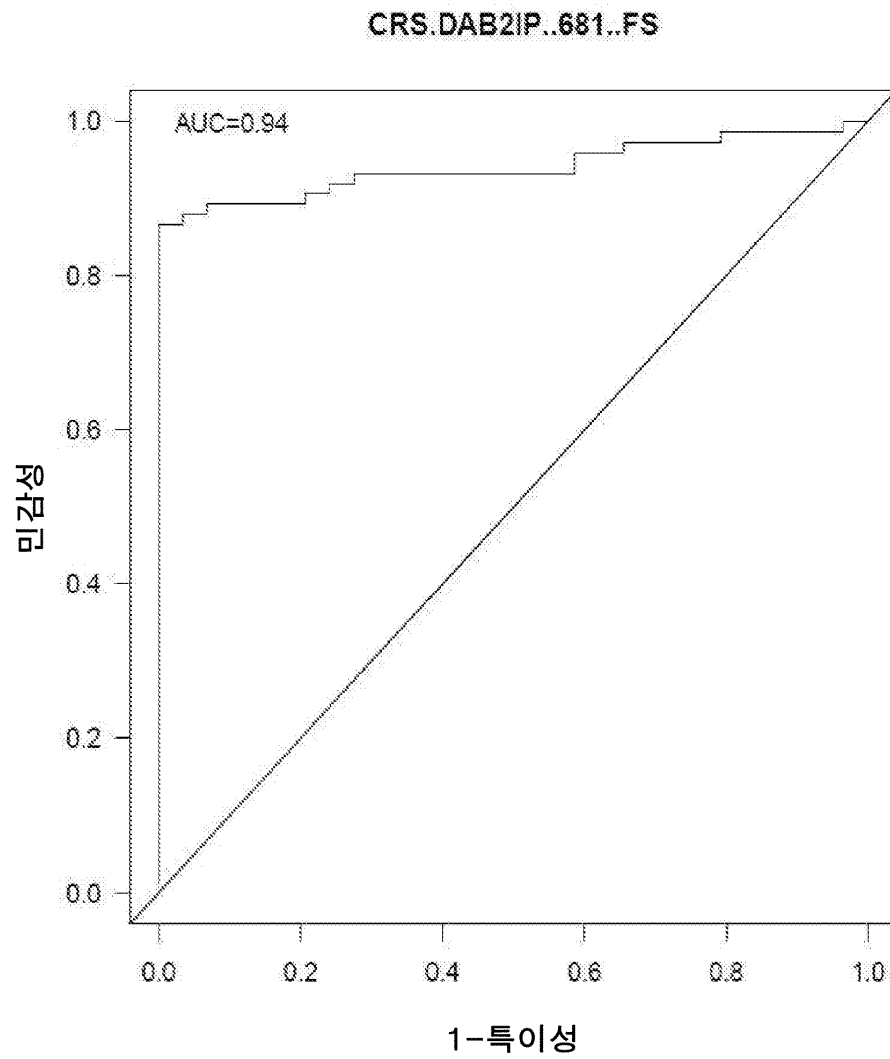
도면3e



도면3f

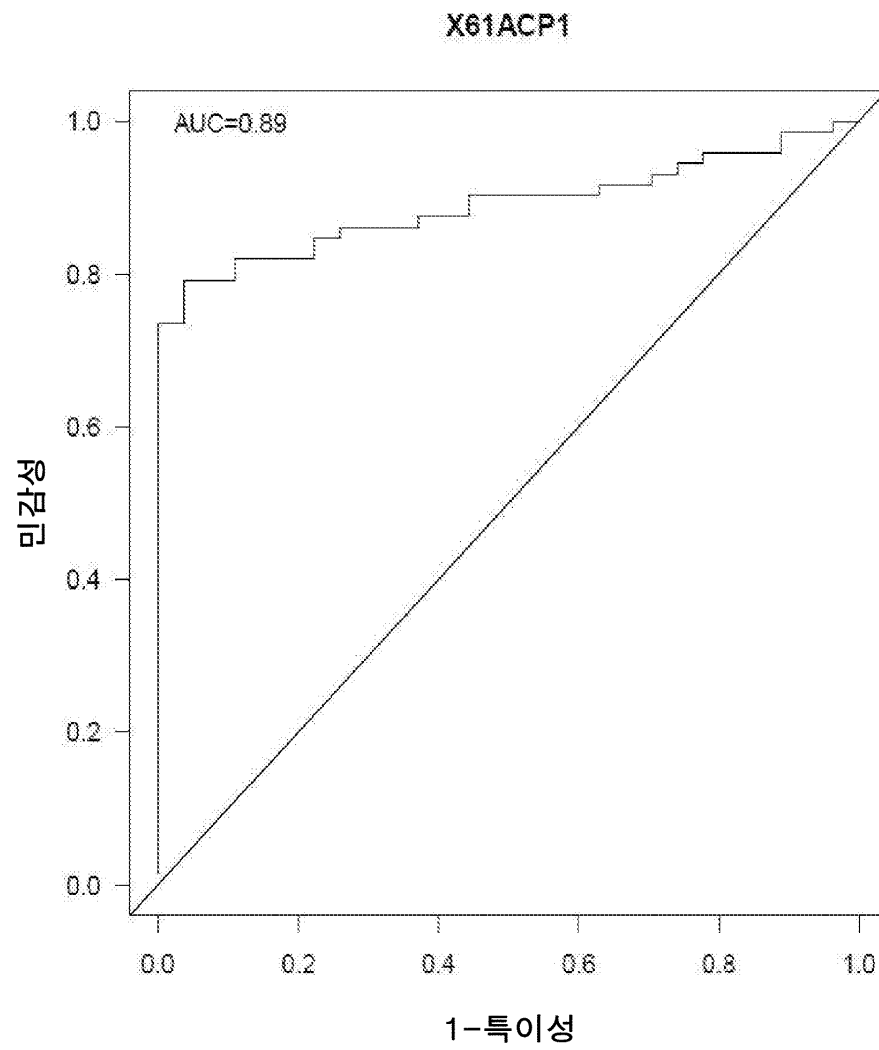


도면3ga

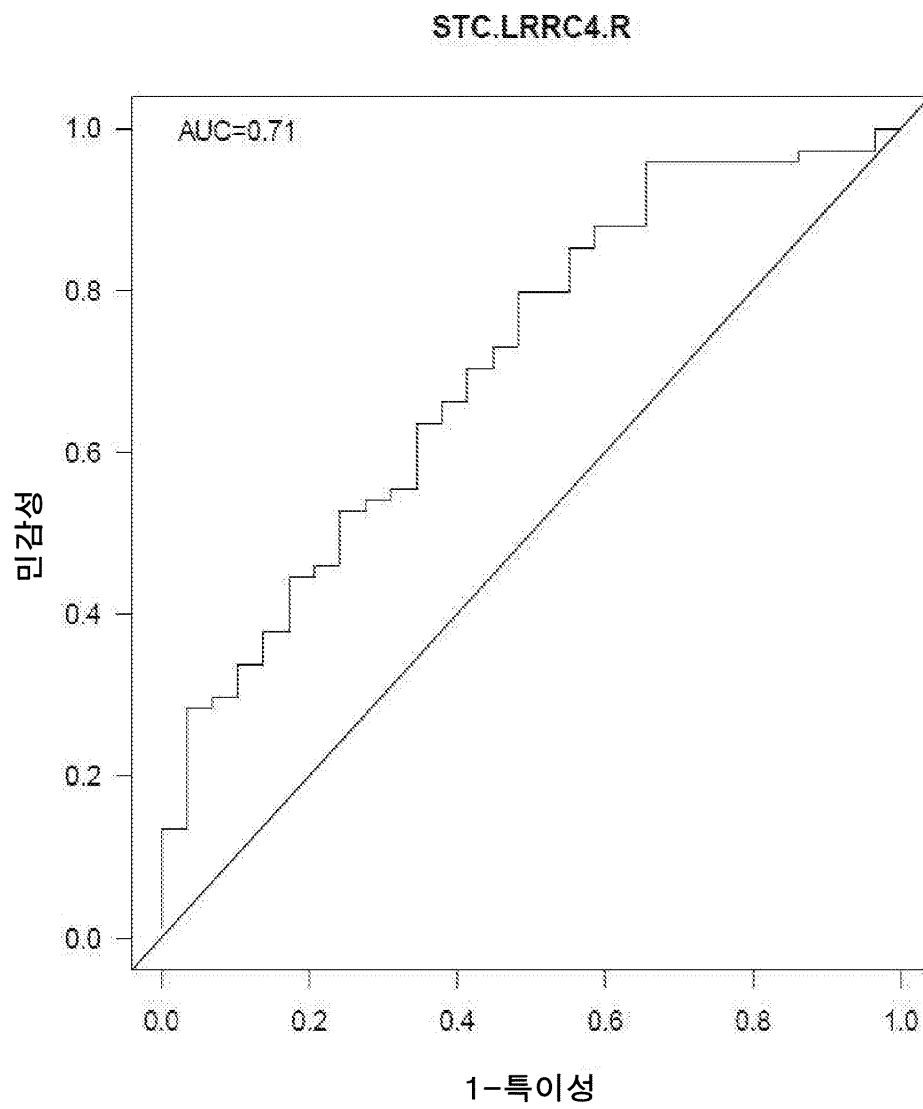




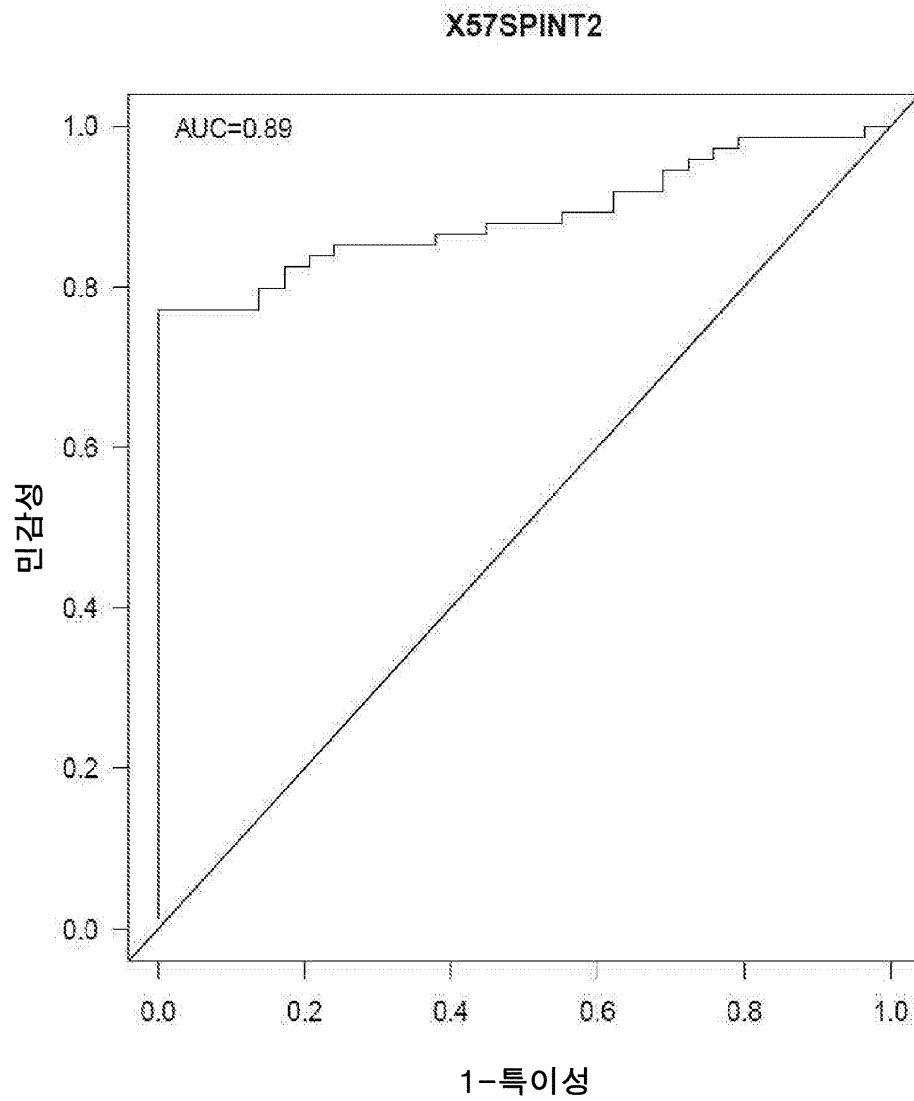
도면3gb



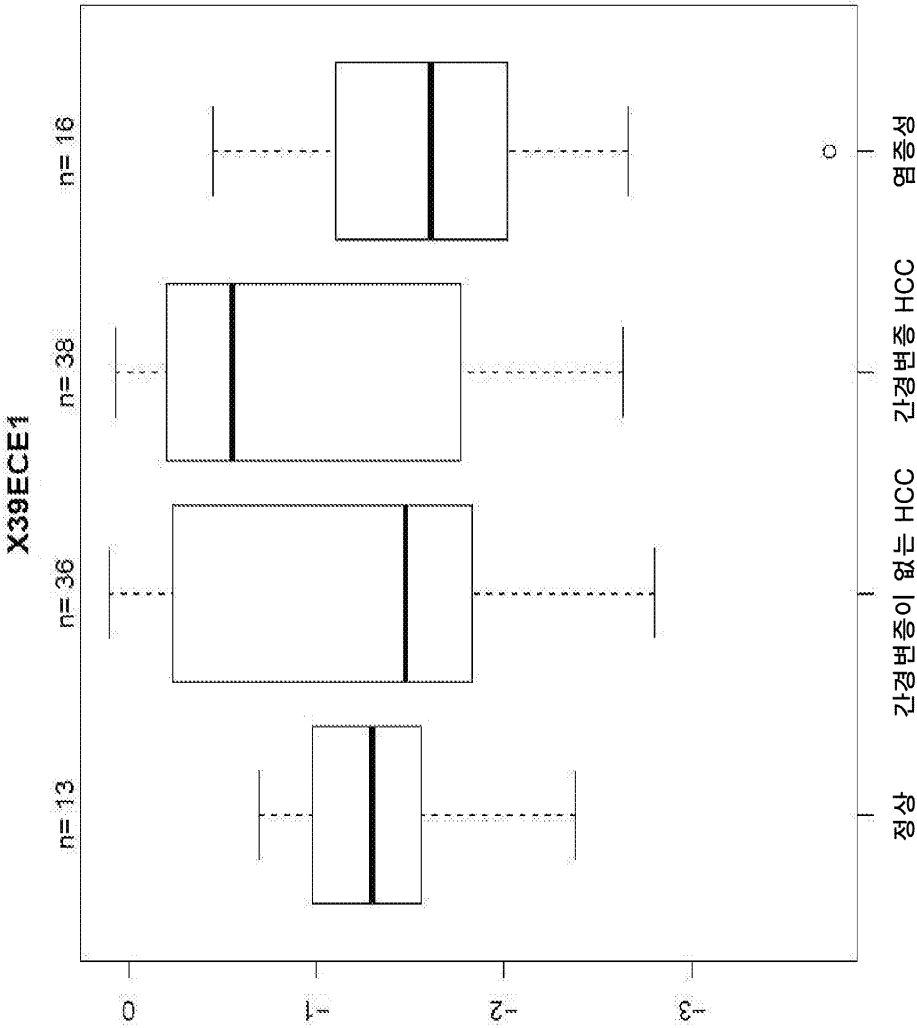
도면3h



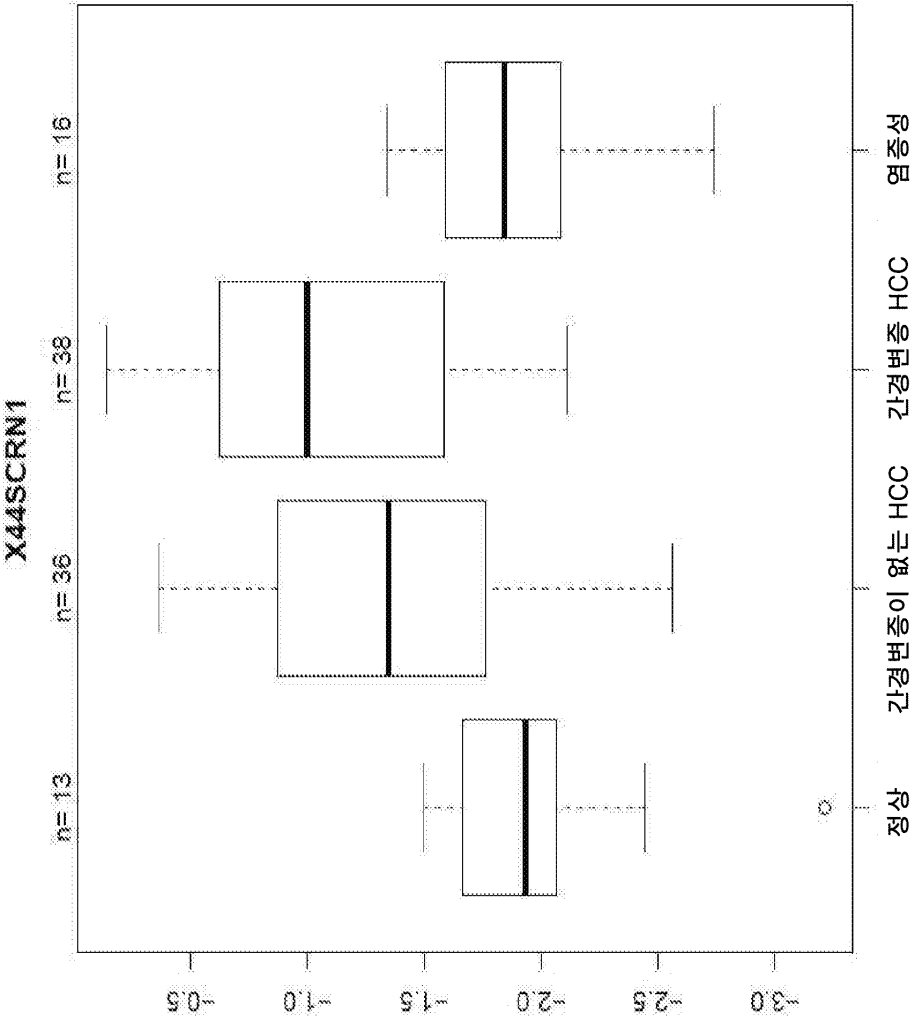
도면3i



도면4a

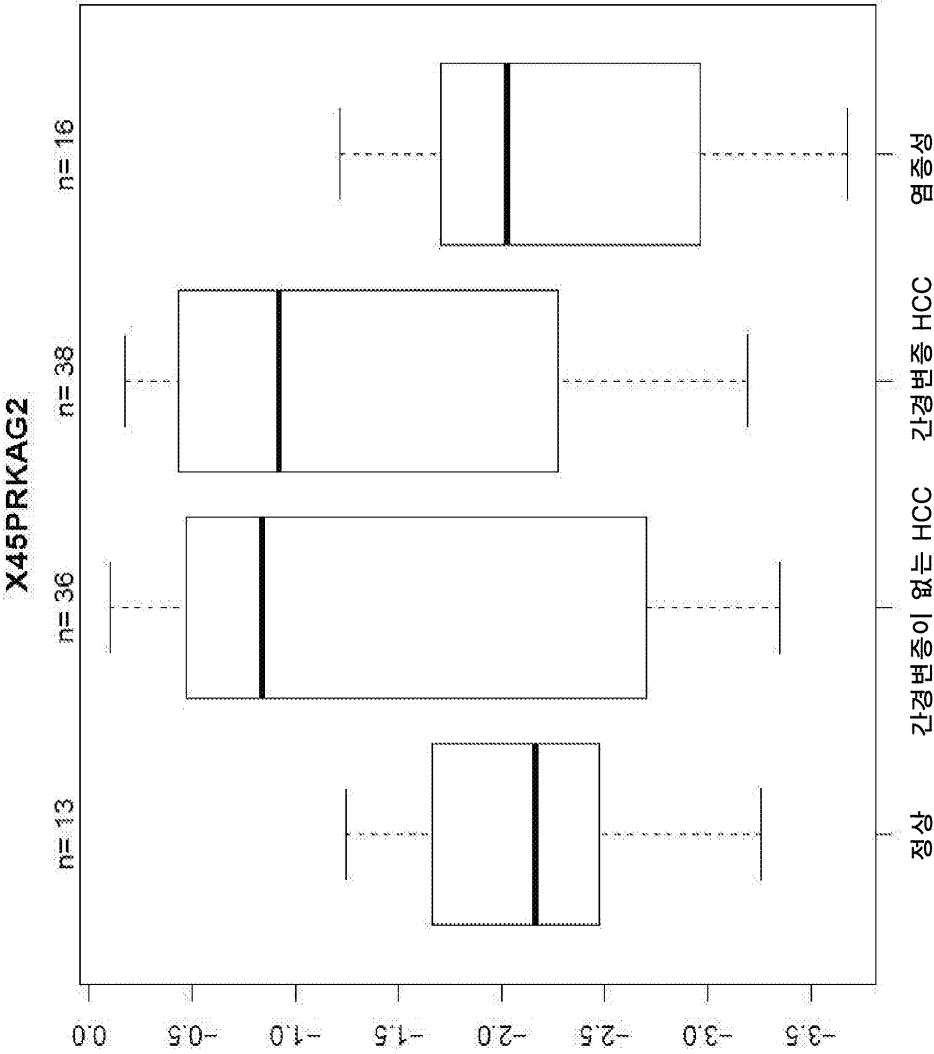


도면4b

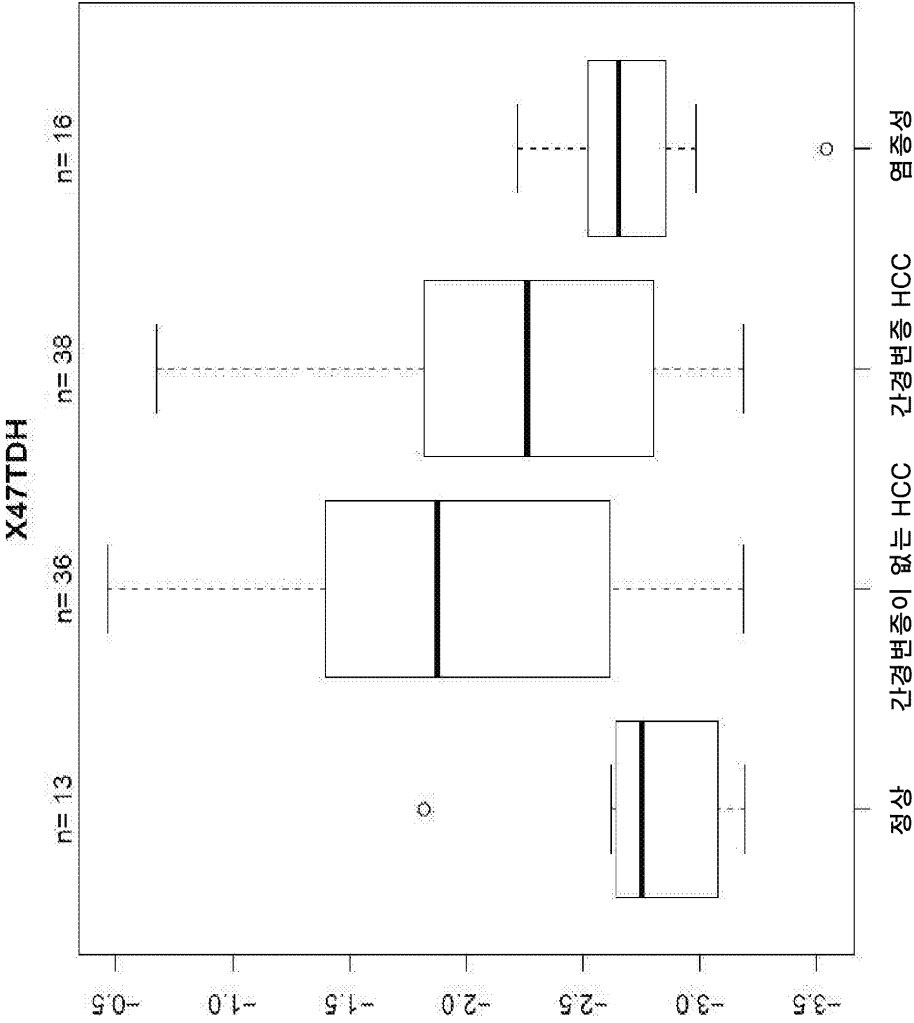




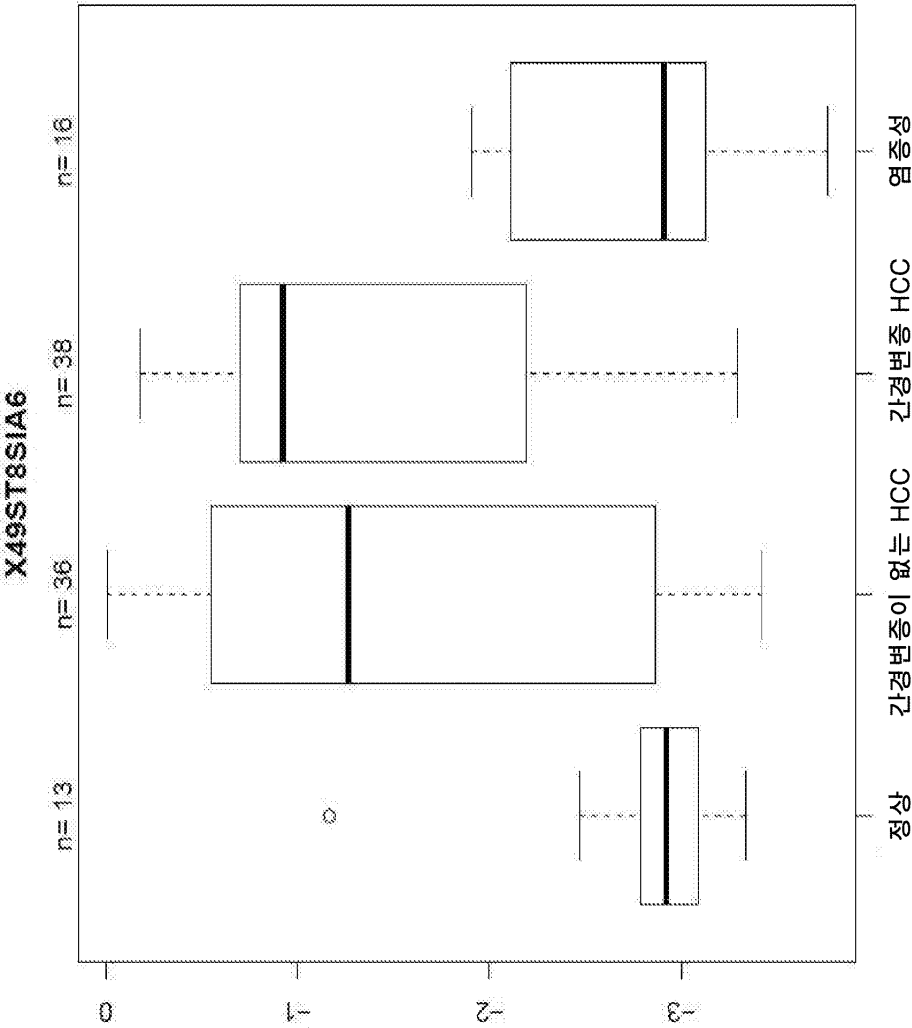
도면4c



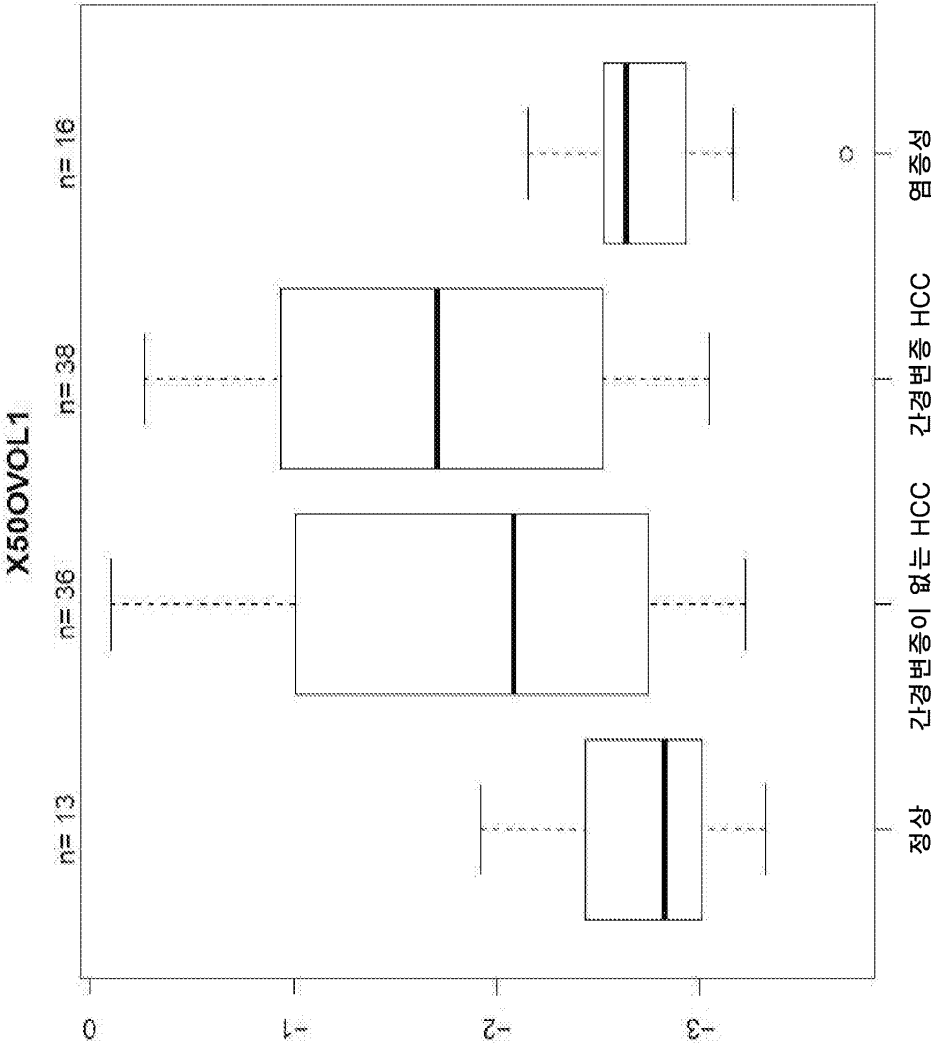
도면4d



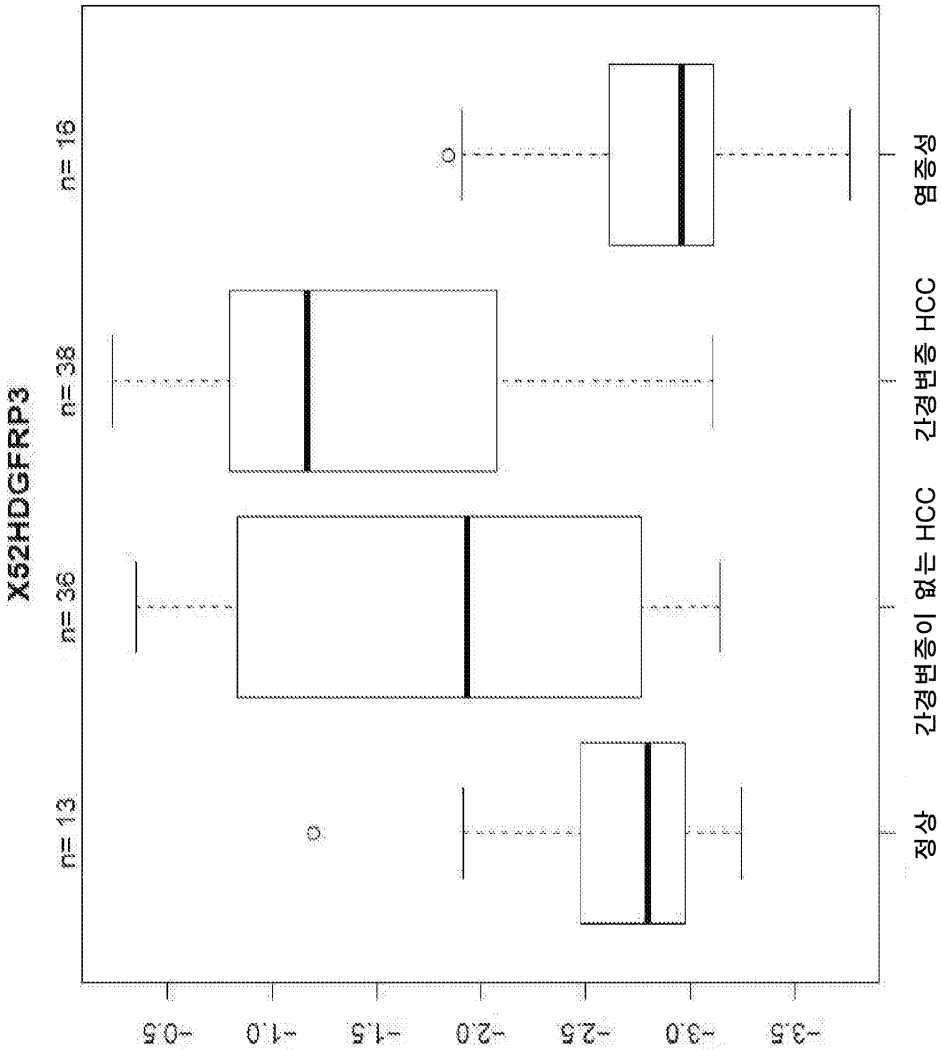
도면4e



도면4f

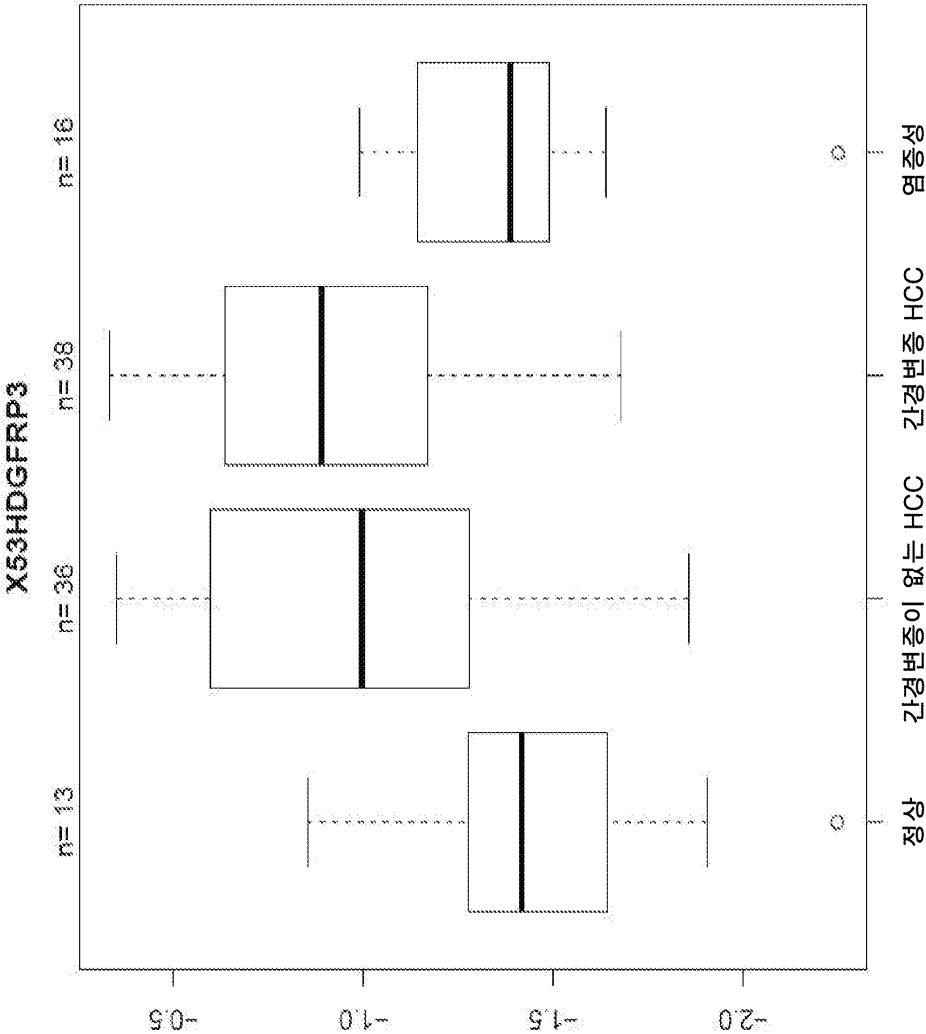


도면4g

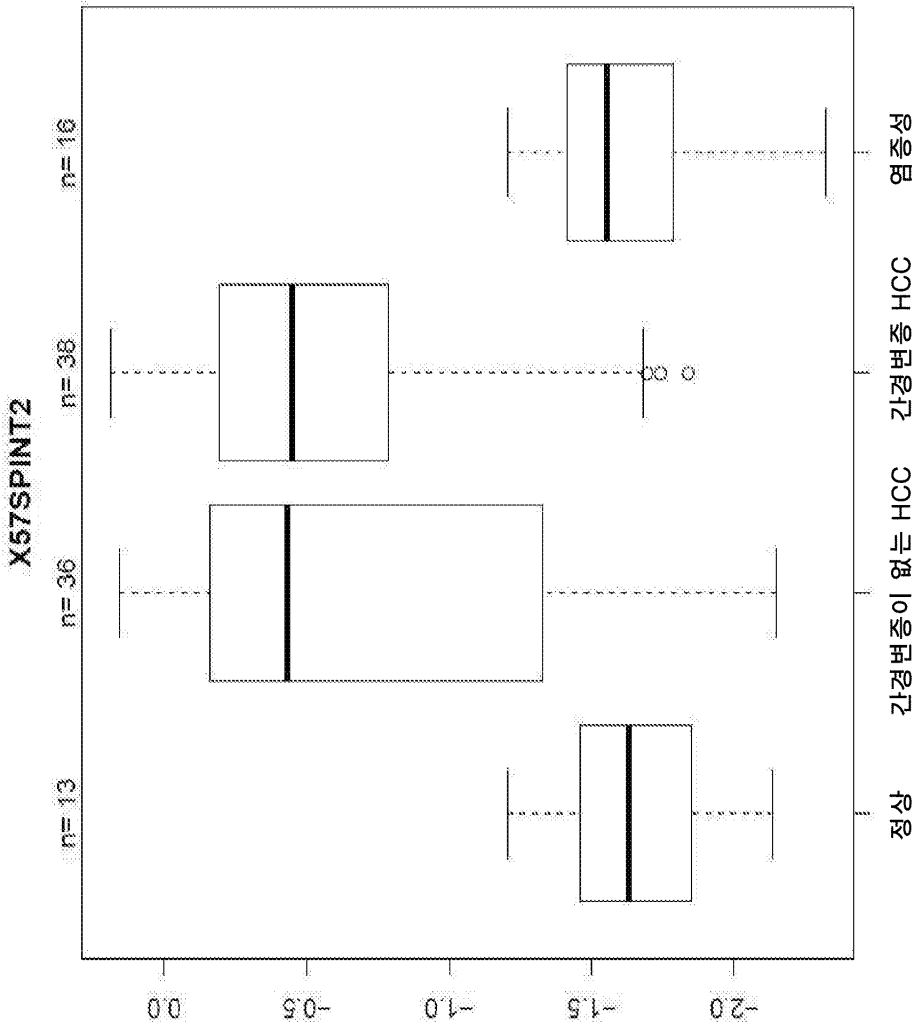




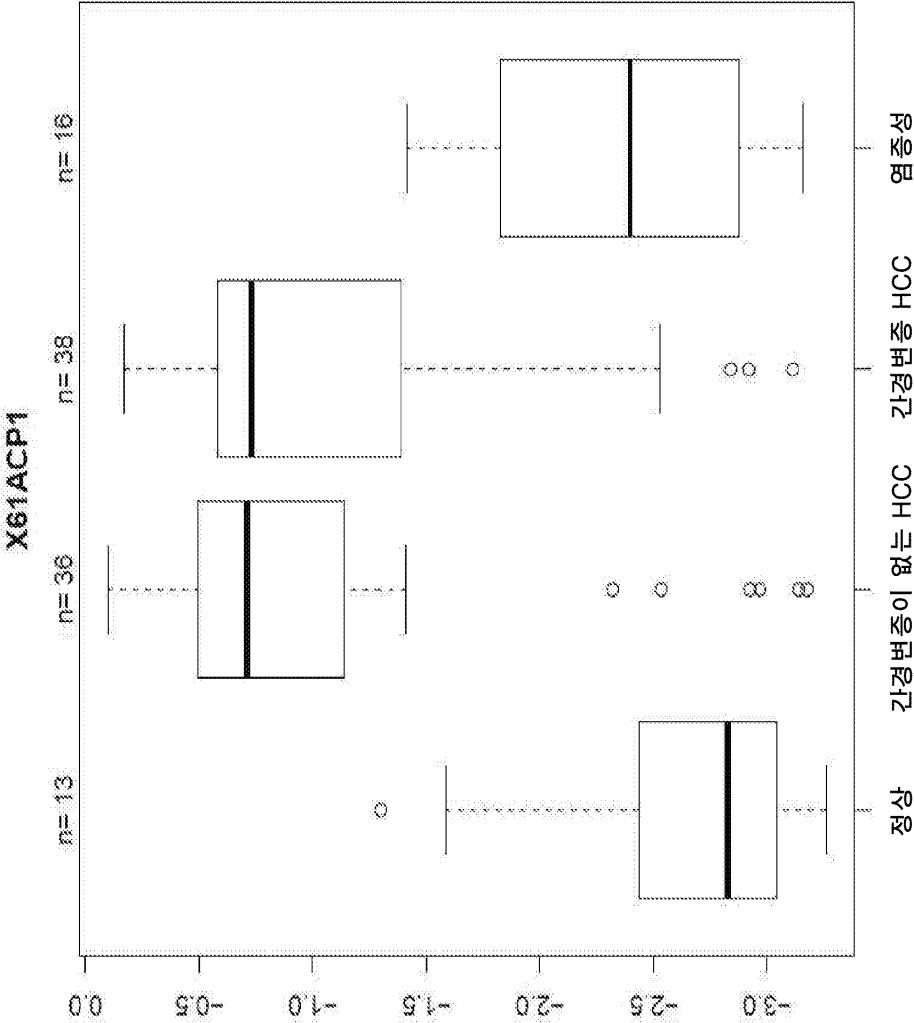
도면4h



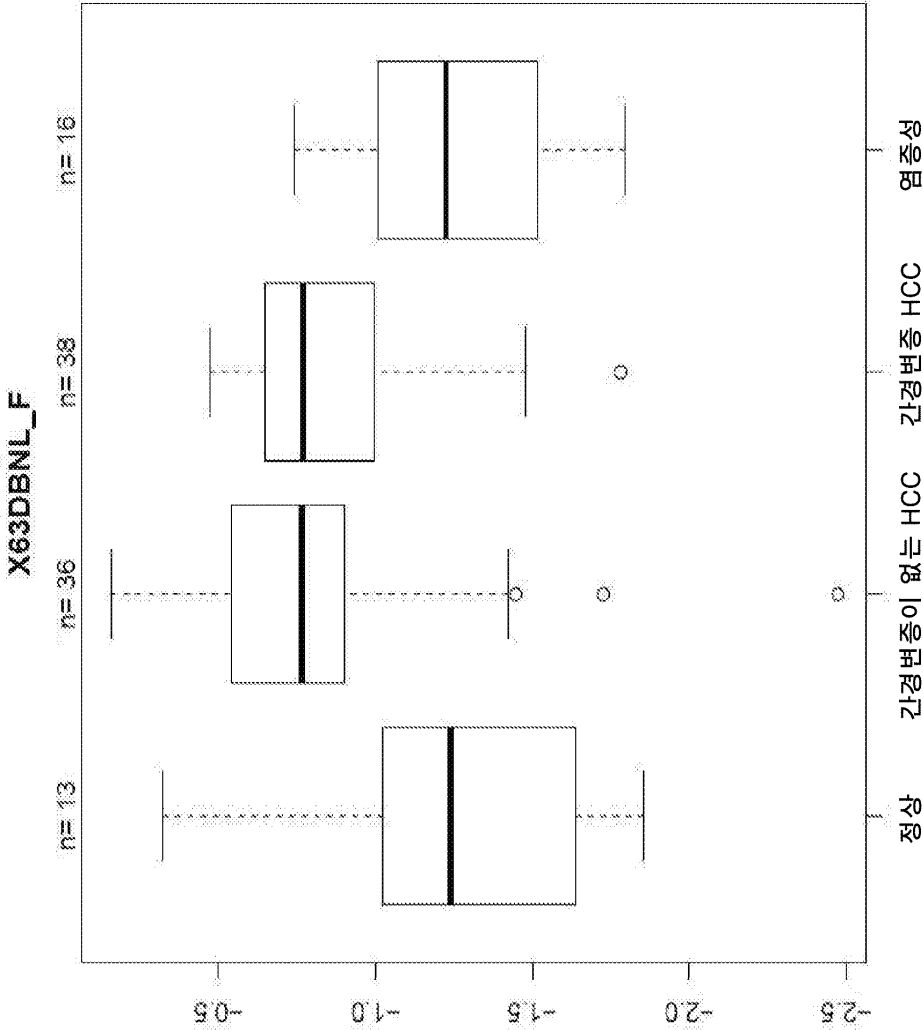
도면4i



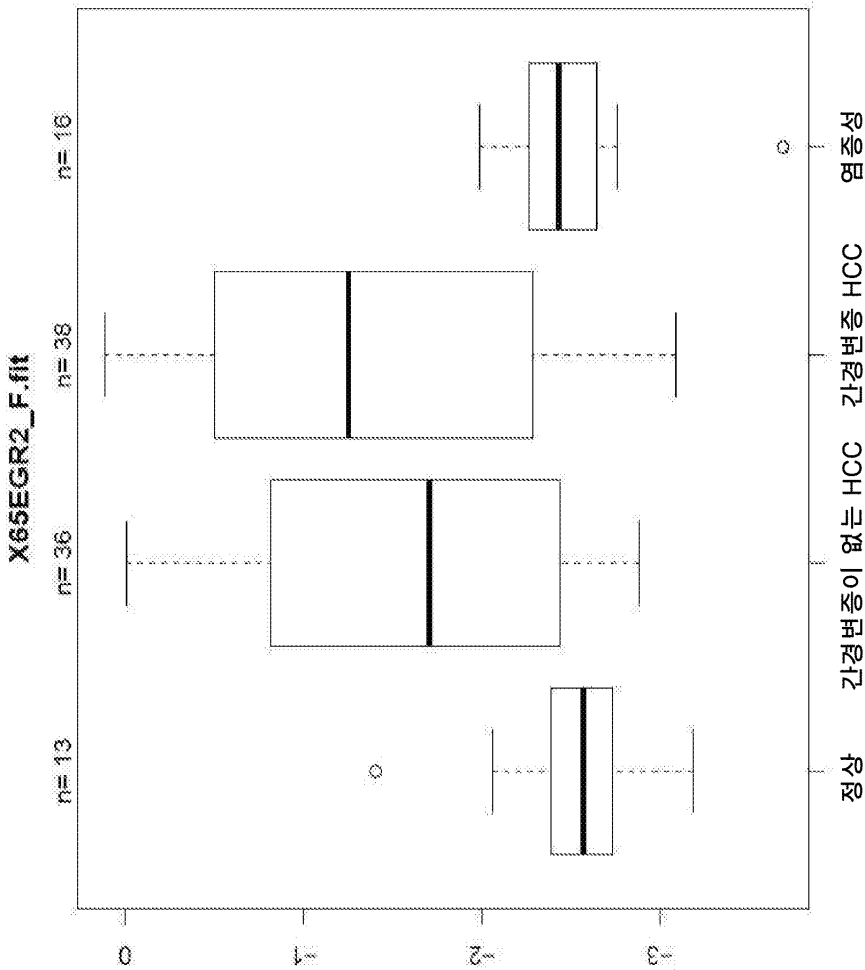
도면4j



도면4k

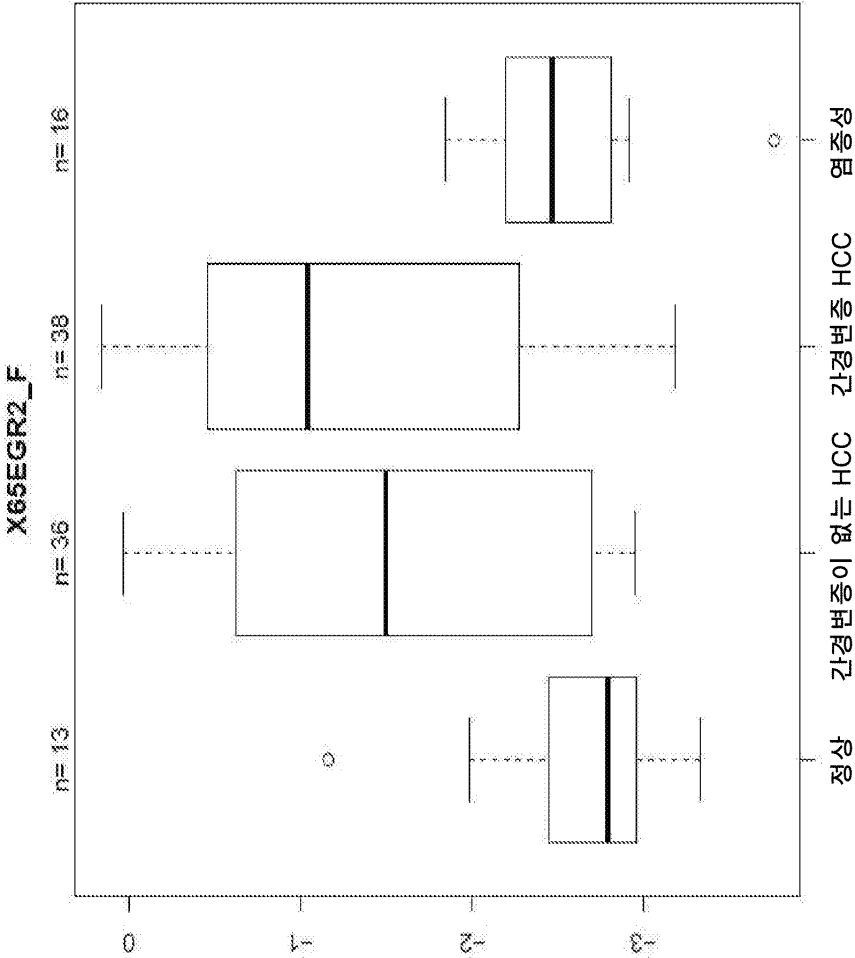


도면41

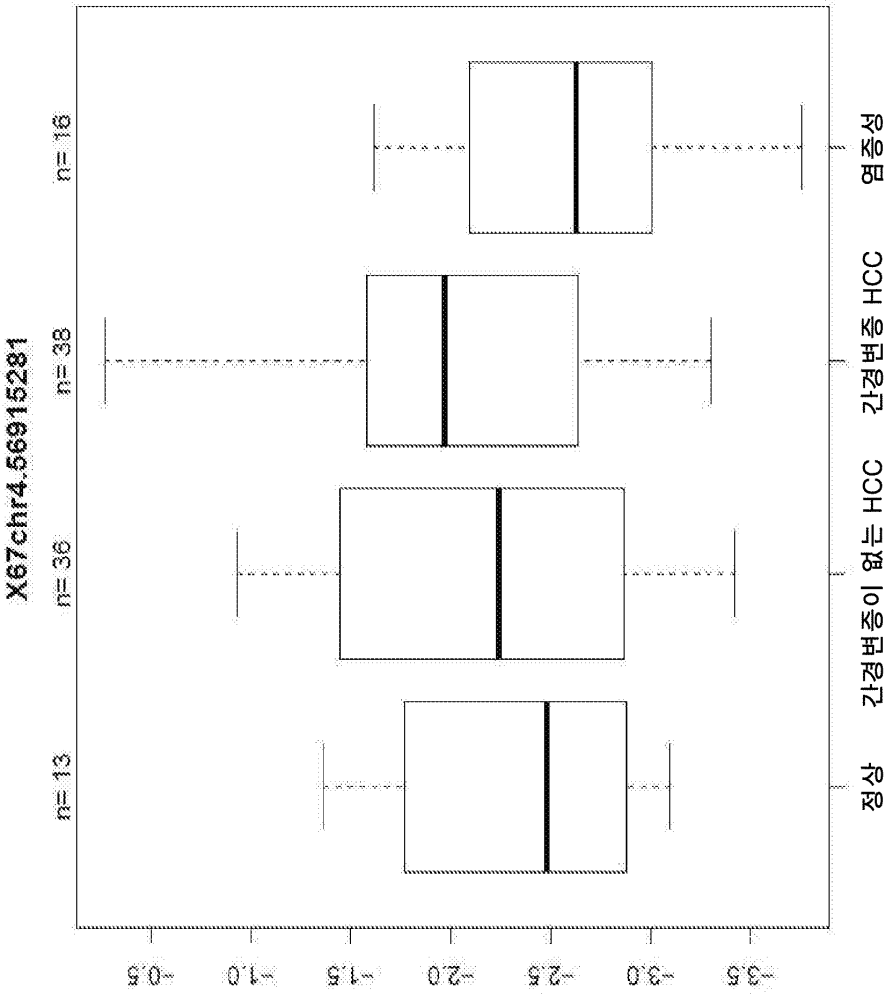




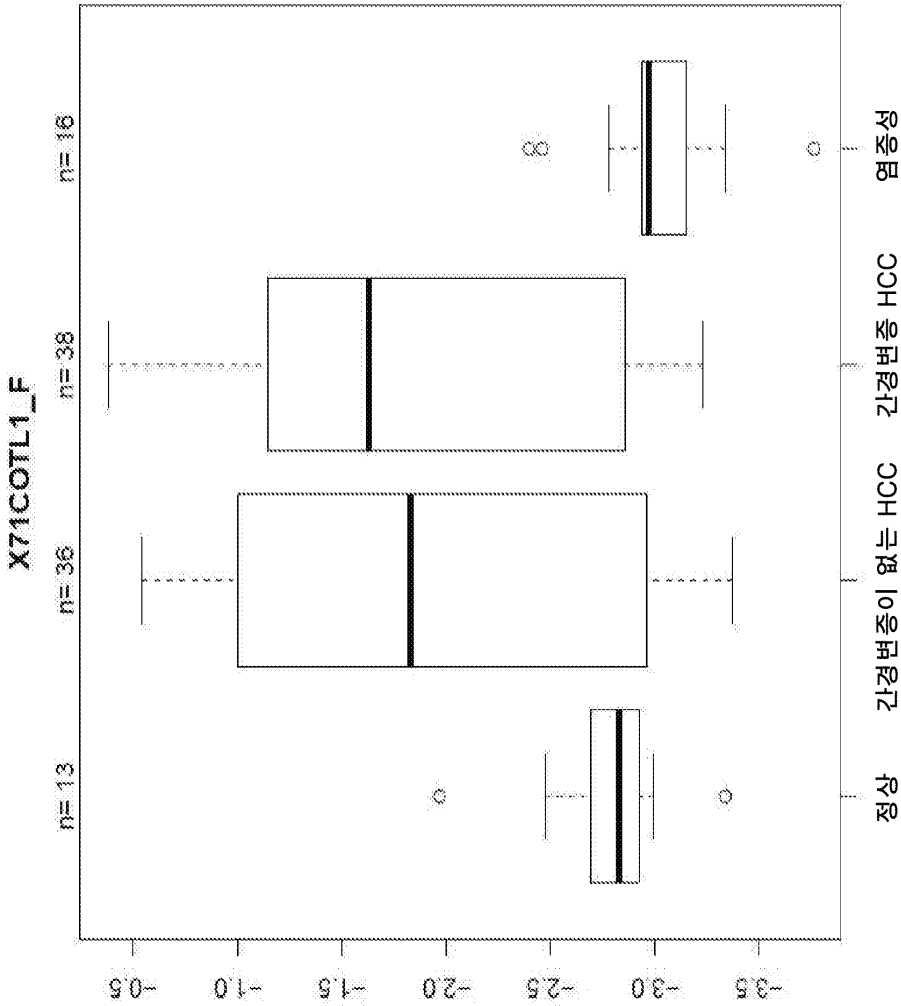
도면4m



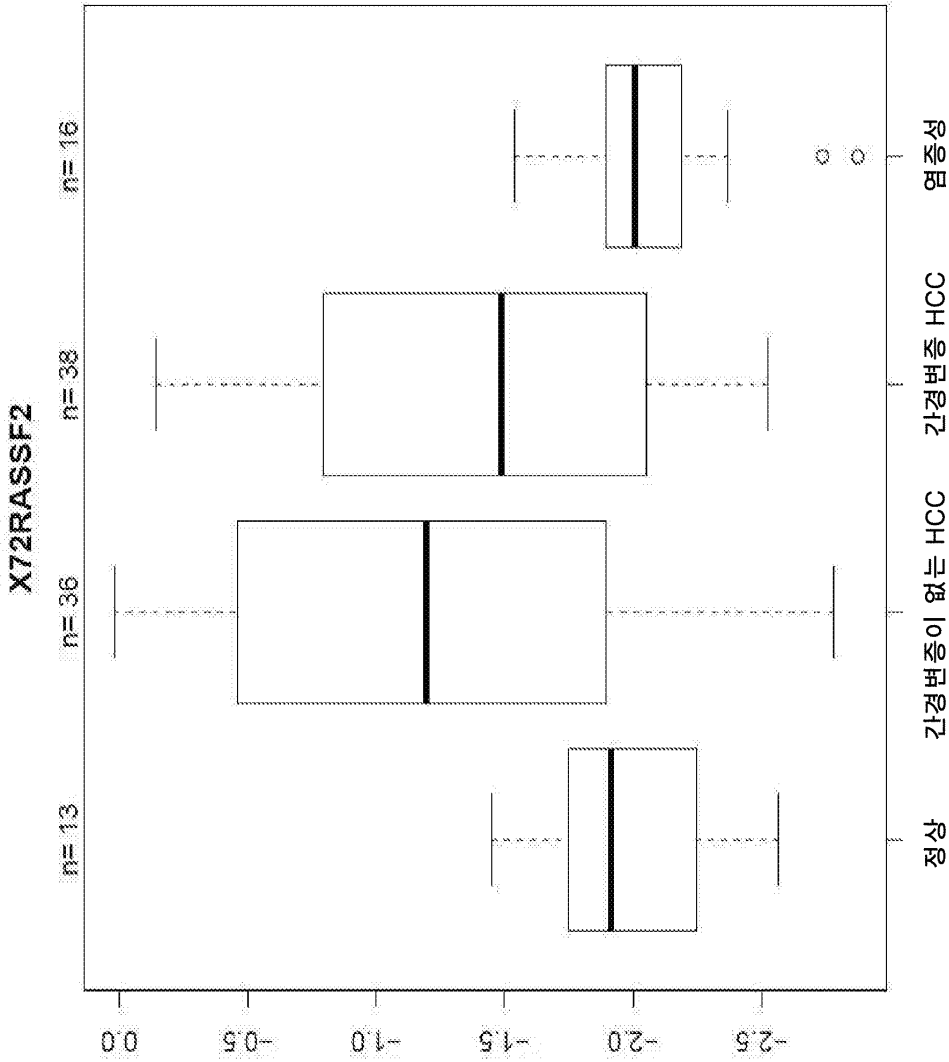
도면4n



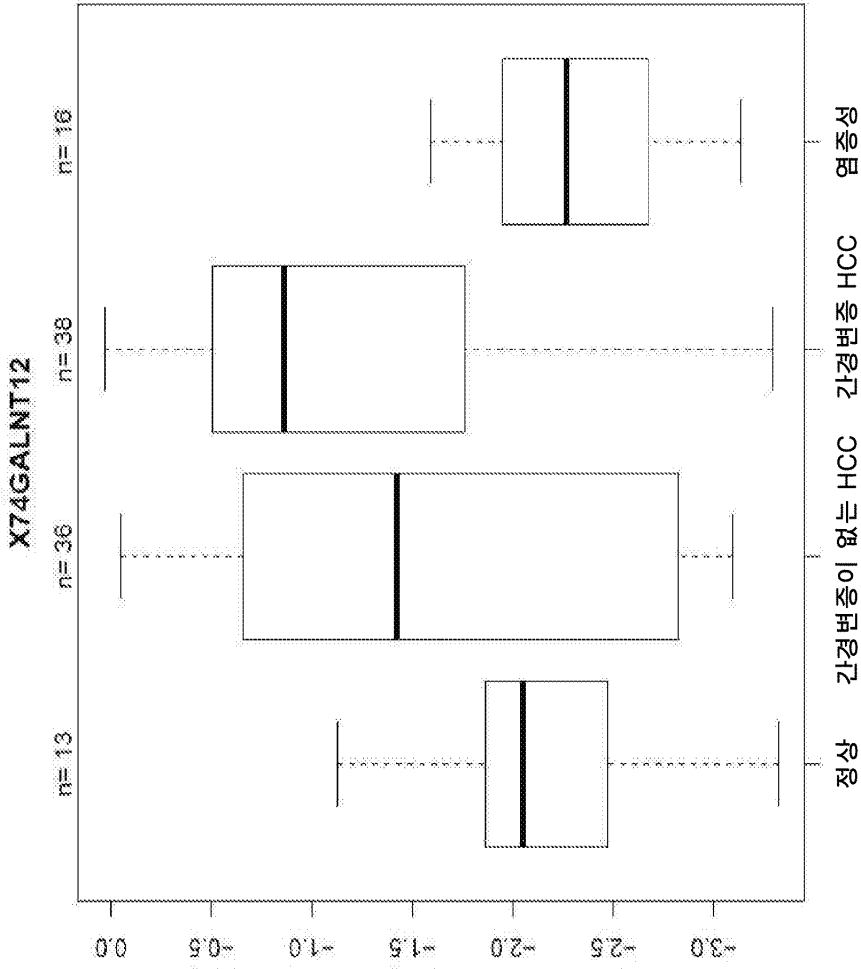
도면40



도면4p

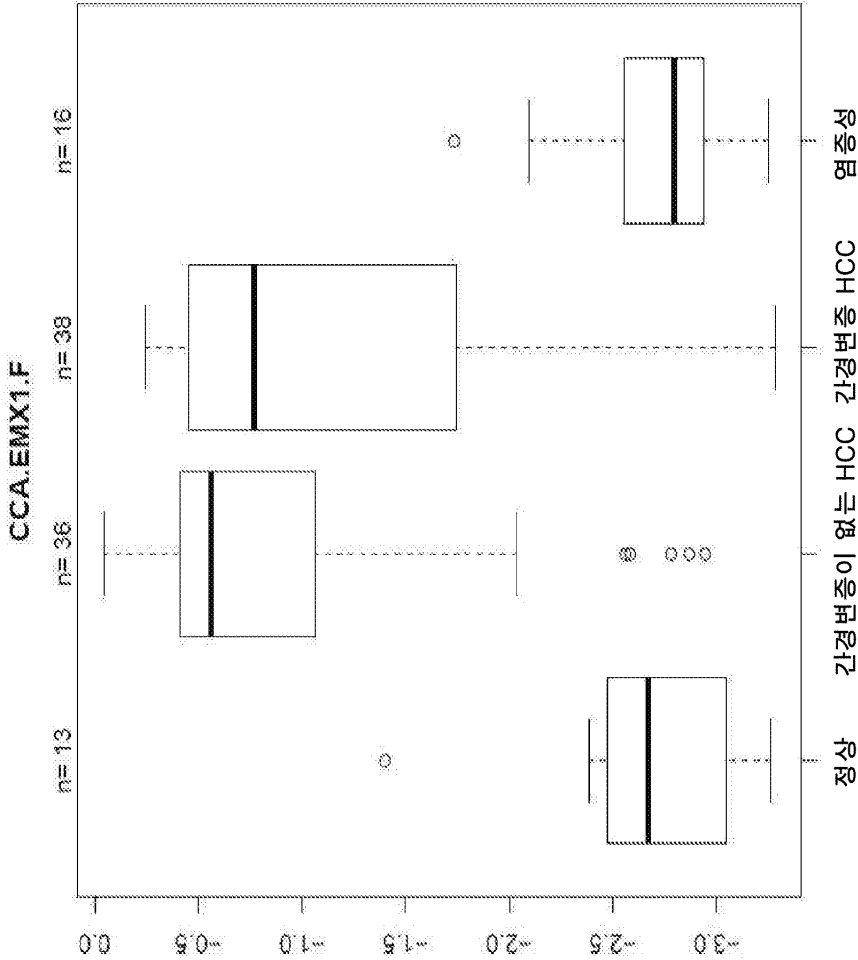


도면4q

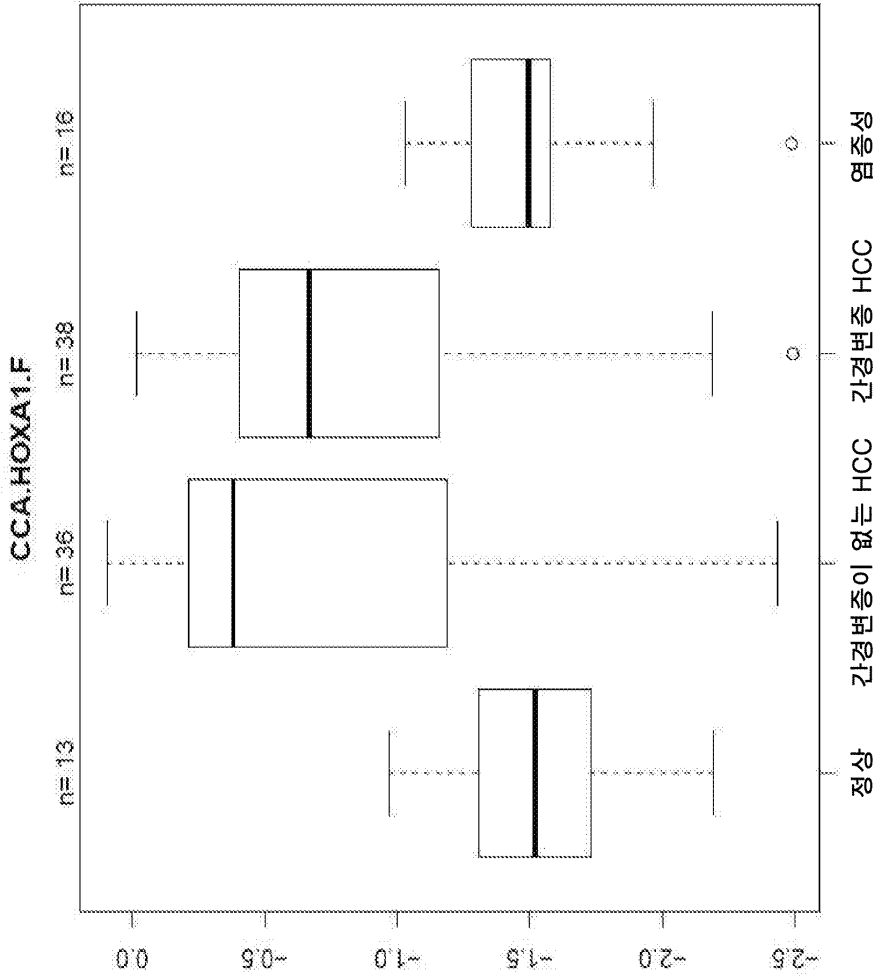




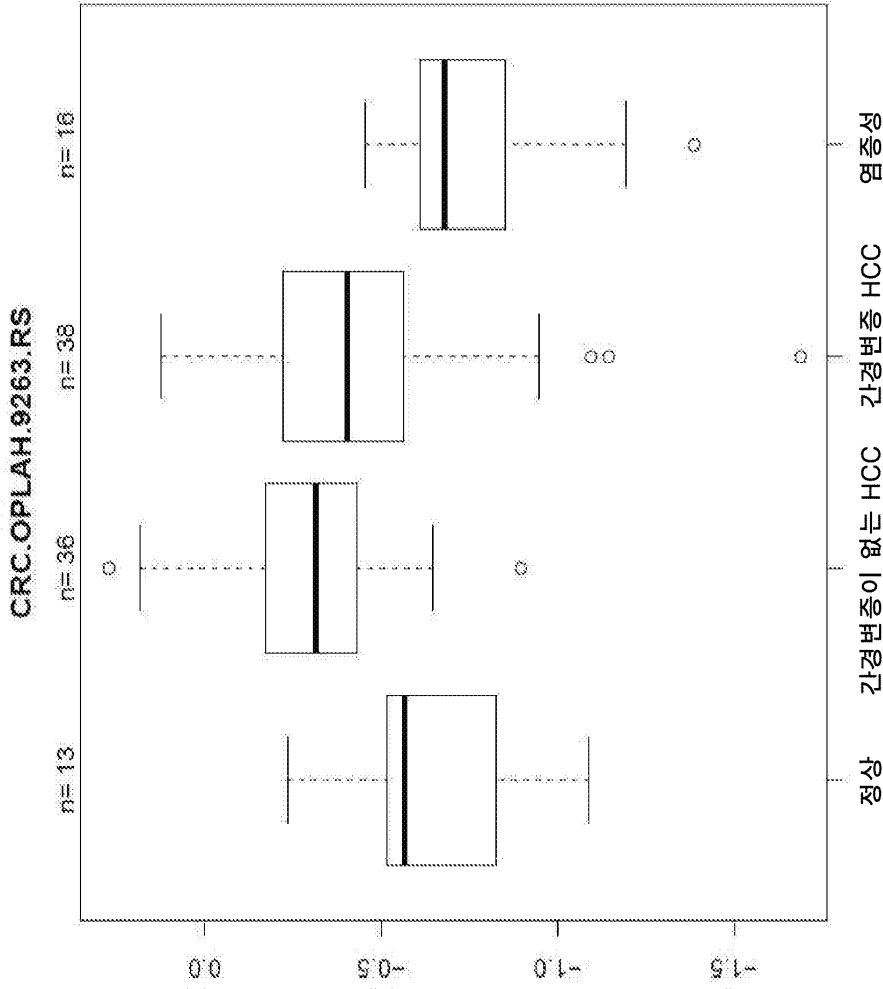
도면4r



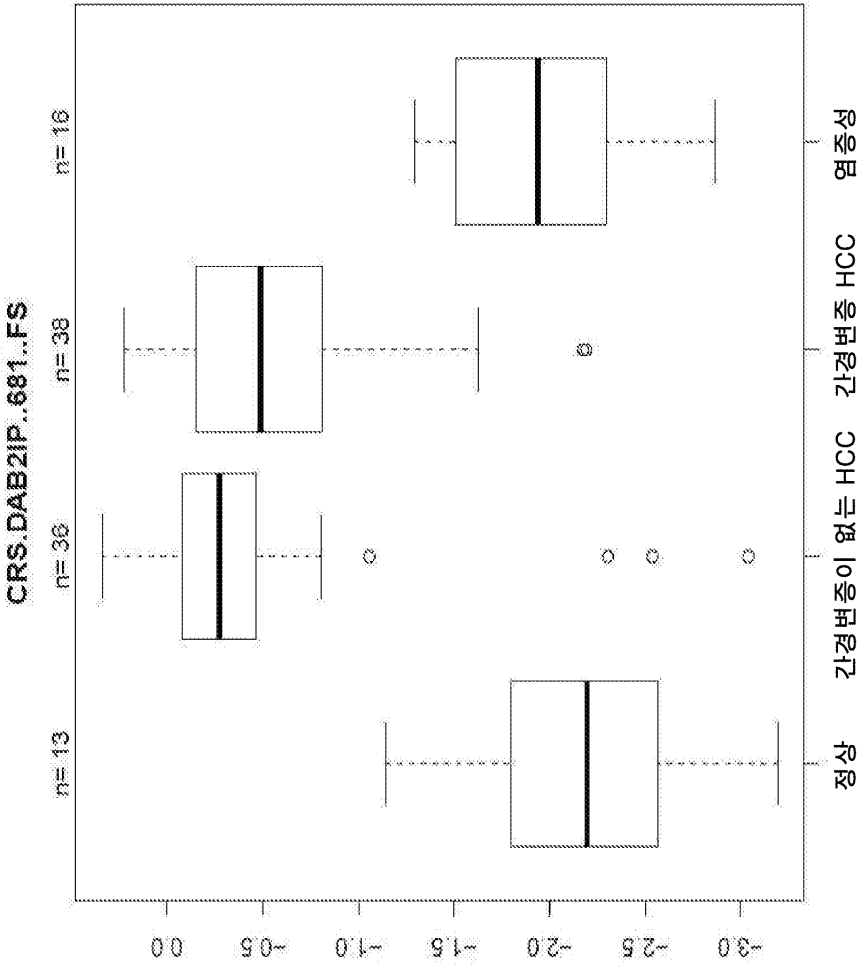
도면4s



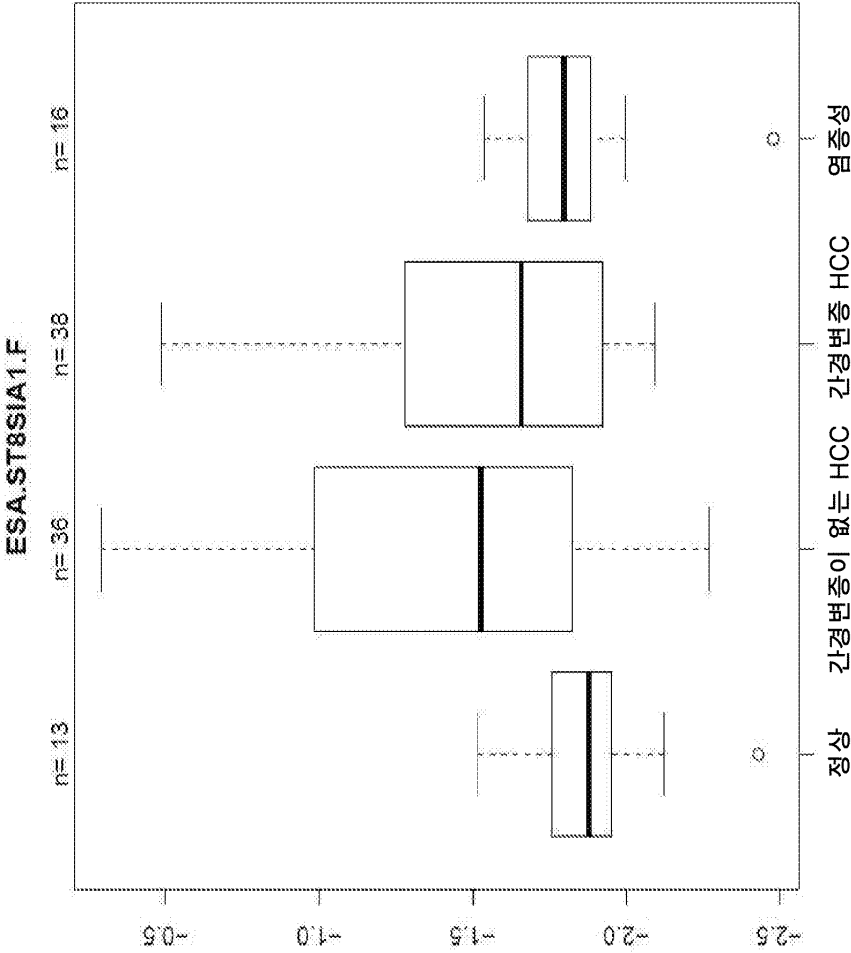
도면4t



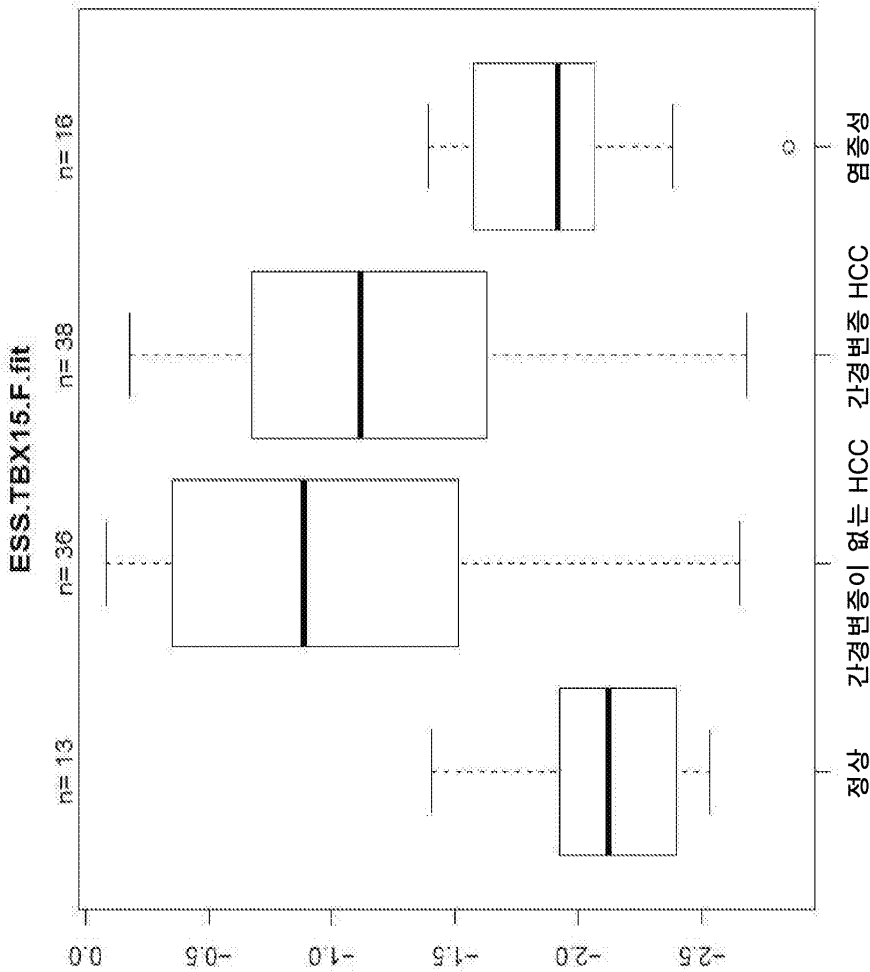
도면4u



도면4v

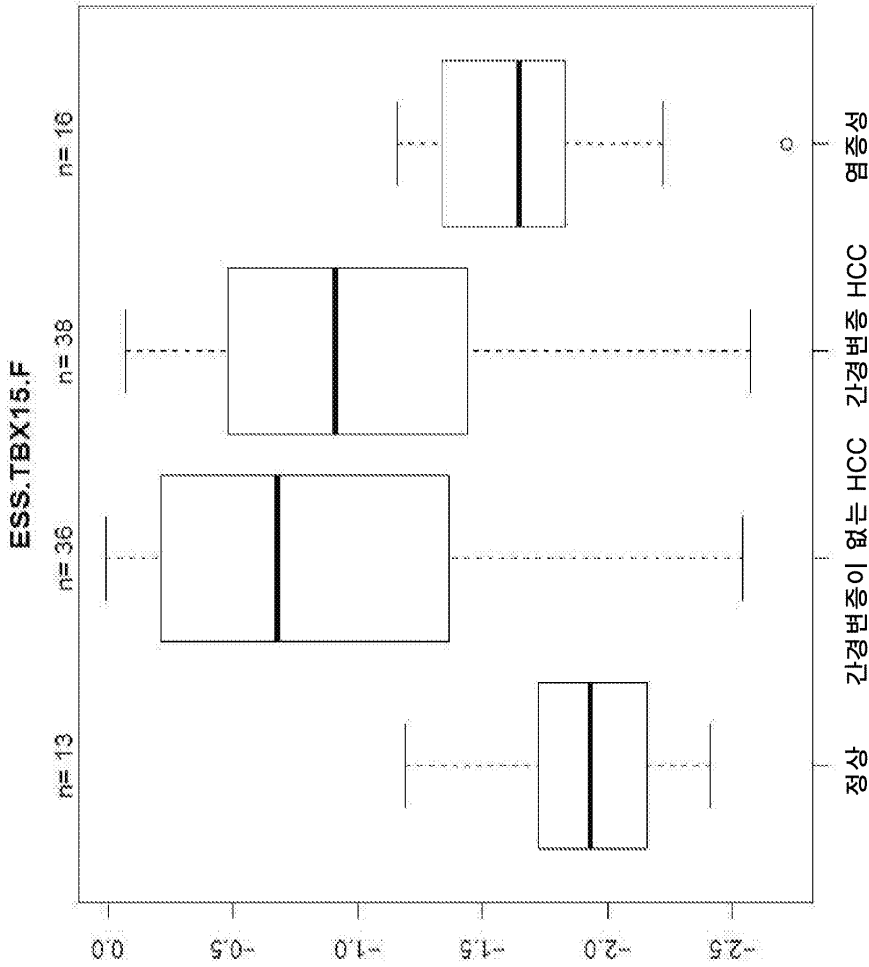


도면4w

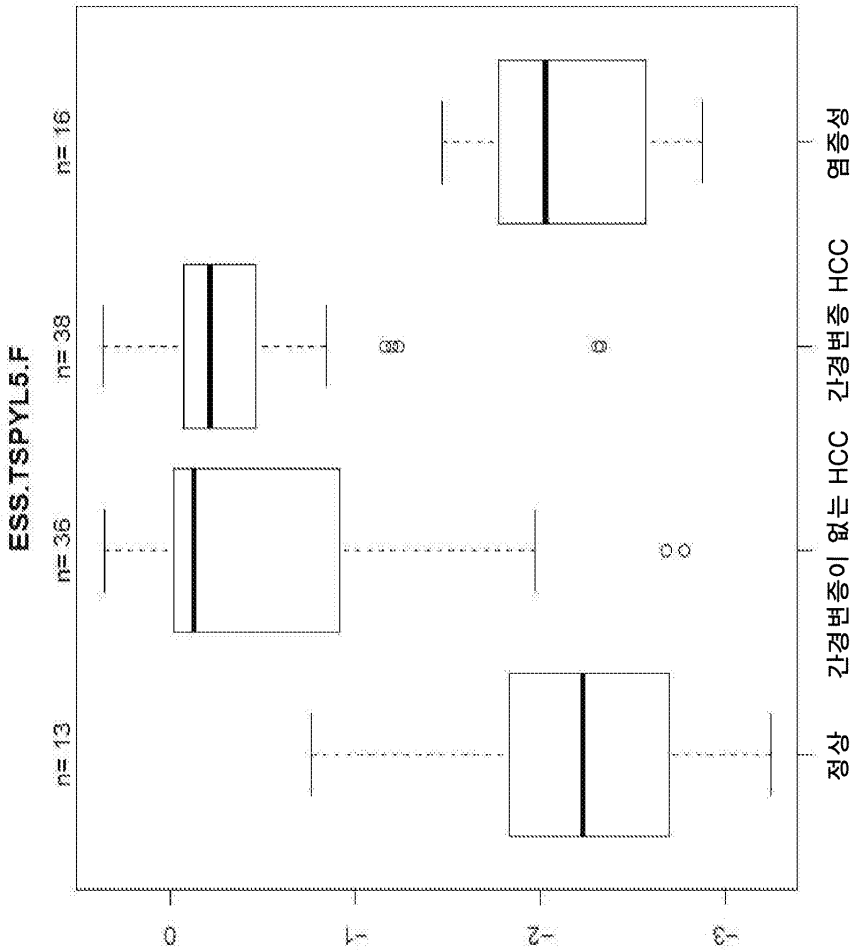




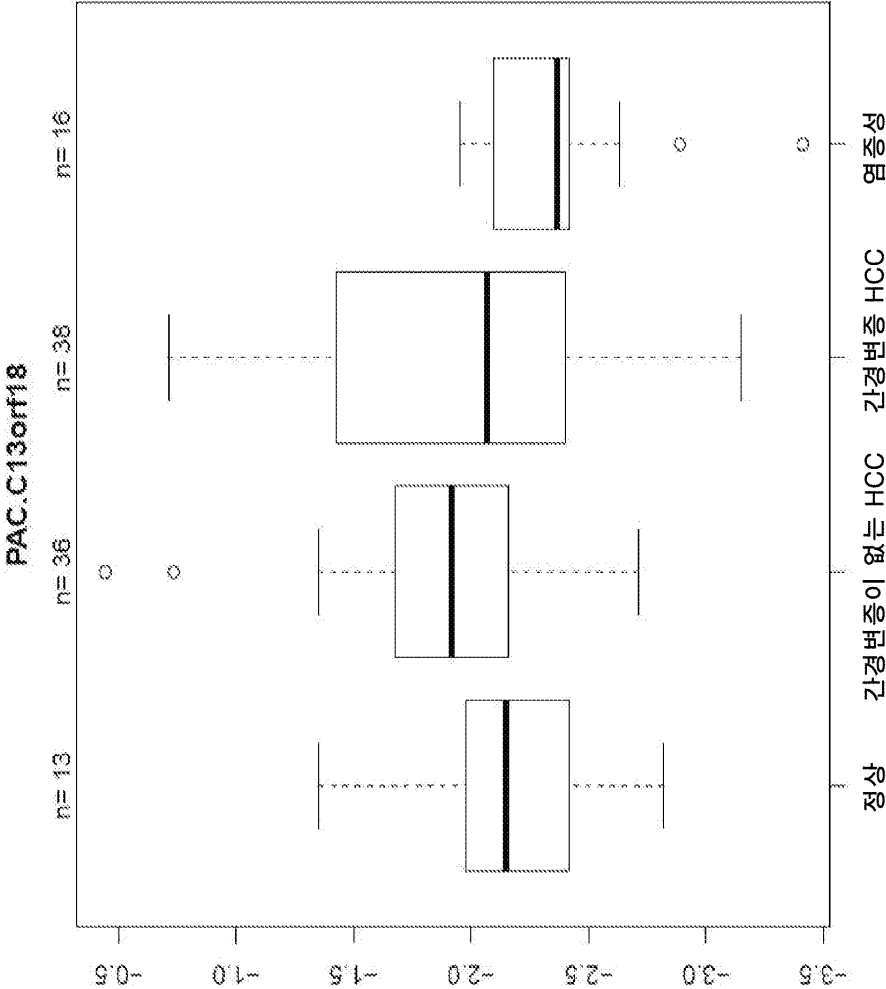
도면4x



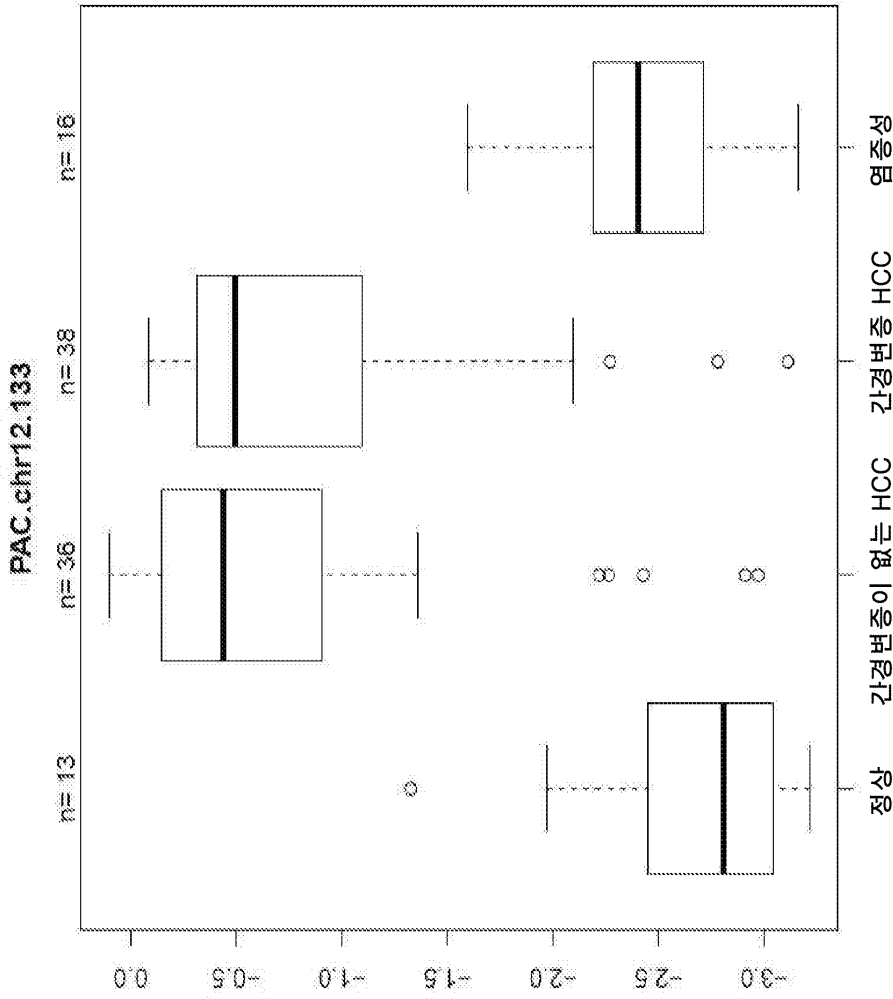
도면4y



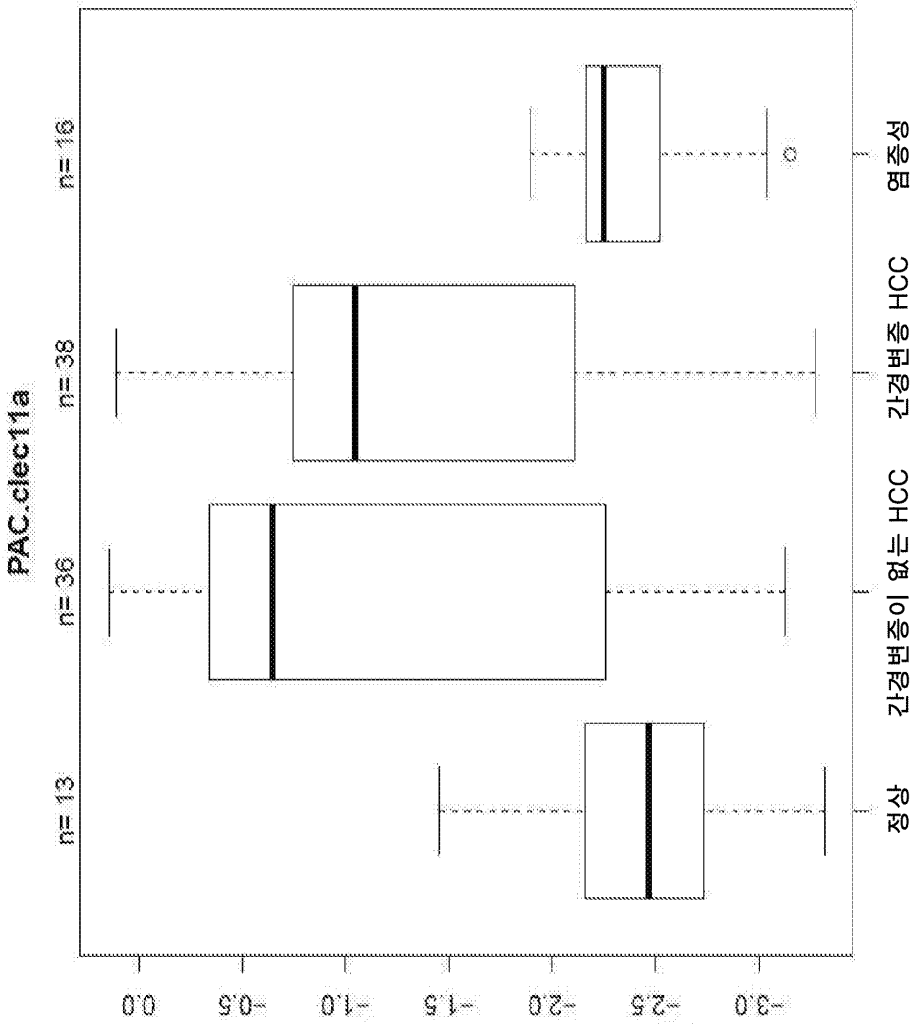
도면4z



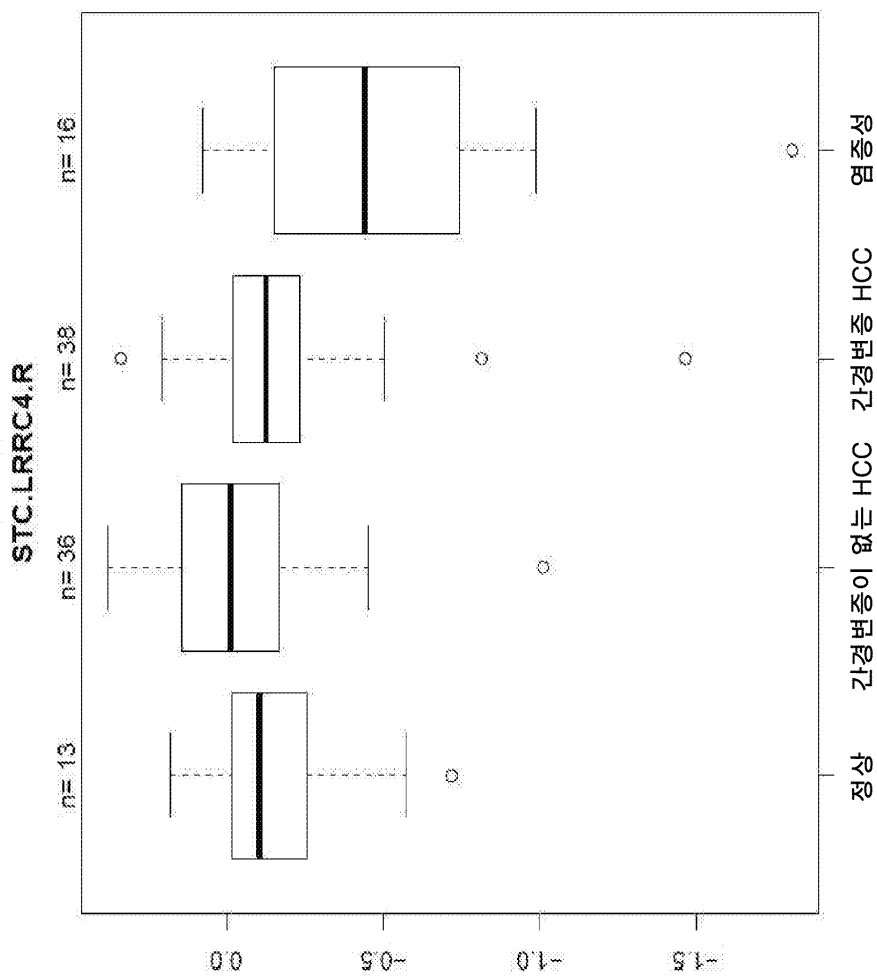
도면4aa



도면4bb

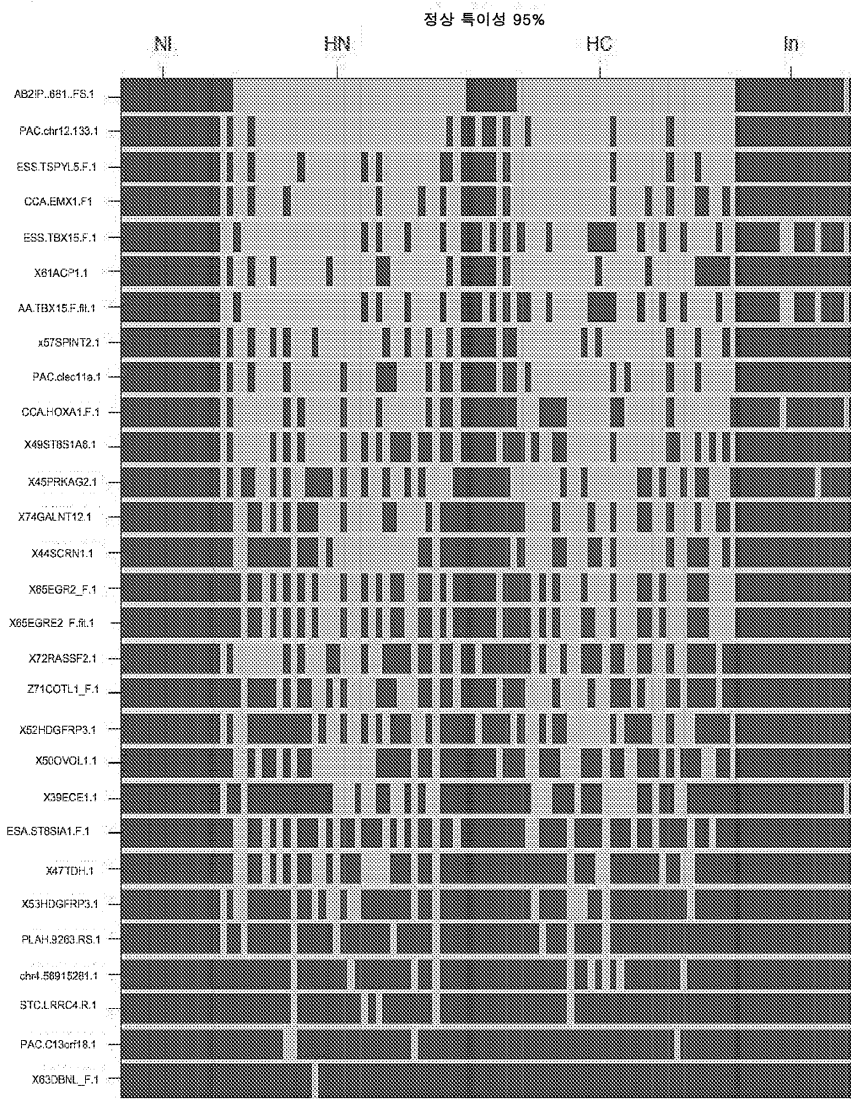


도면4cc





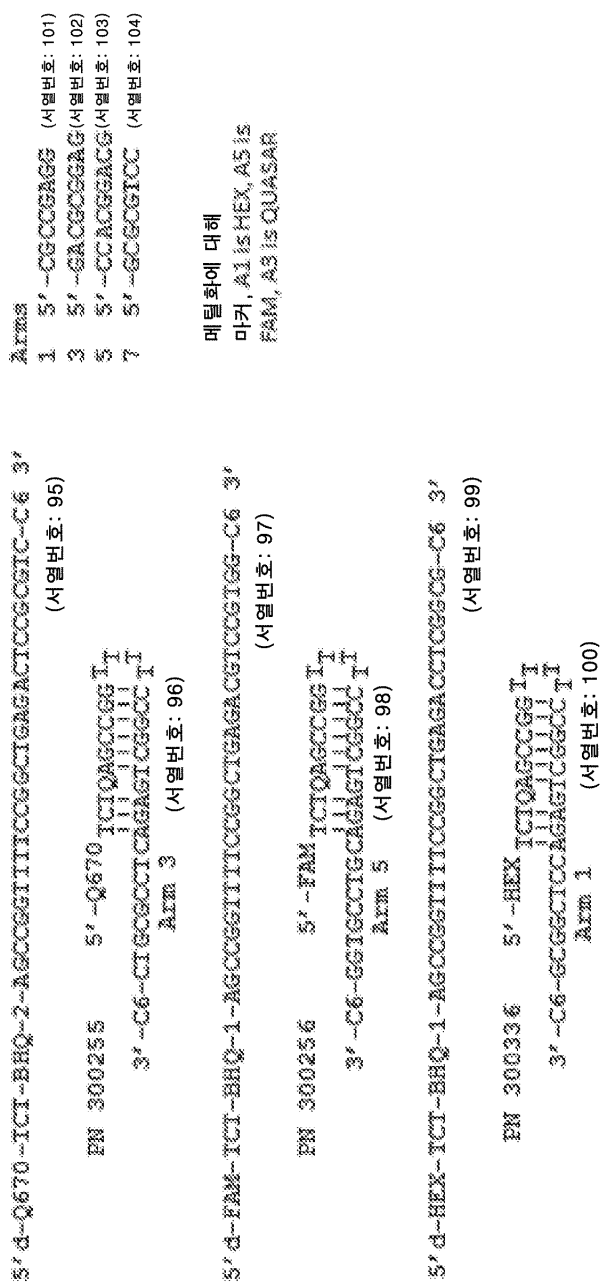
도면5



대상체

정상			간경변증이 없는 HCC				간경변증 HCC				염증성		
거짓	진성	진성	거짓	진성	거짓	진성	거짓	진성	거짓	진성	거짓	진성	진성
13	0	0	1	2	3	4	3	4	12	3	12	3	3
0	0	0	0	33	0	0	31	0	0	0	0	1	1

도면6

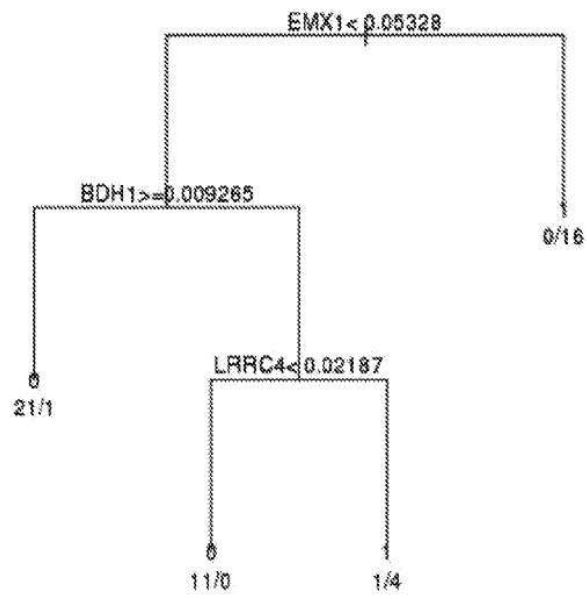


Arms

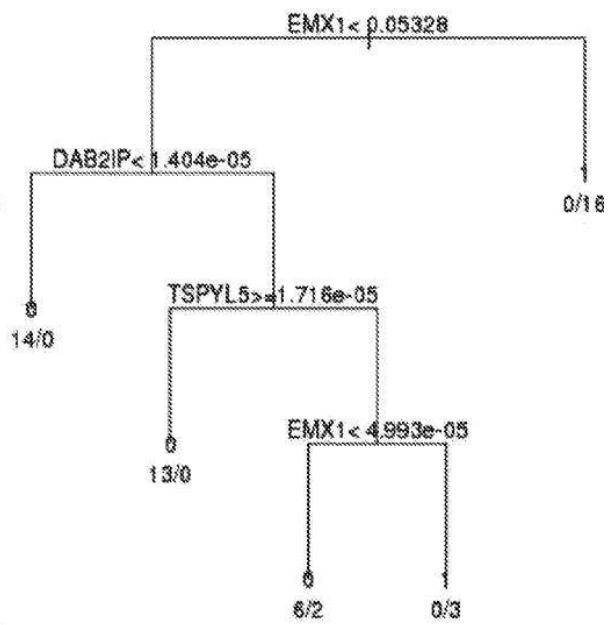
1	5'-CGCCGAGG	(서열번호: 101)
3	5'-GACGGGAG	(서열번호: 102)
5	5'-CCAGGAGC	(서열번호: 103)
7	5'-GCGGTC	(서열번호: 104)

메틸화에 대해  
마커, A1 is HEX, A5 is  
FAM, A3 is QUASAR

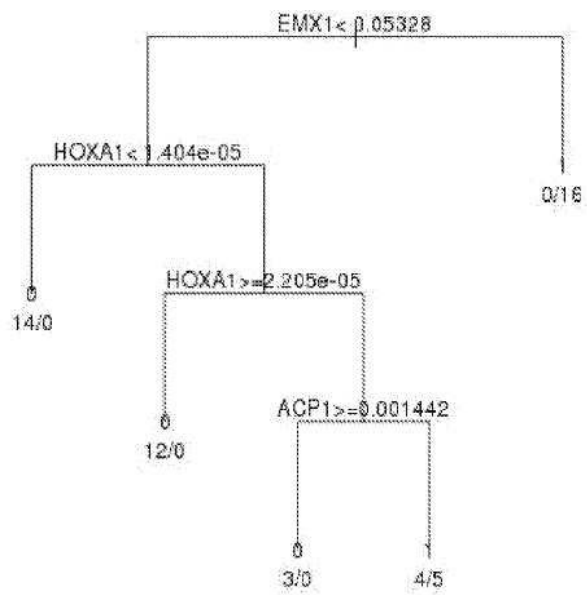
도면7a



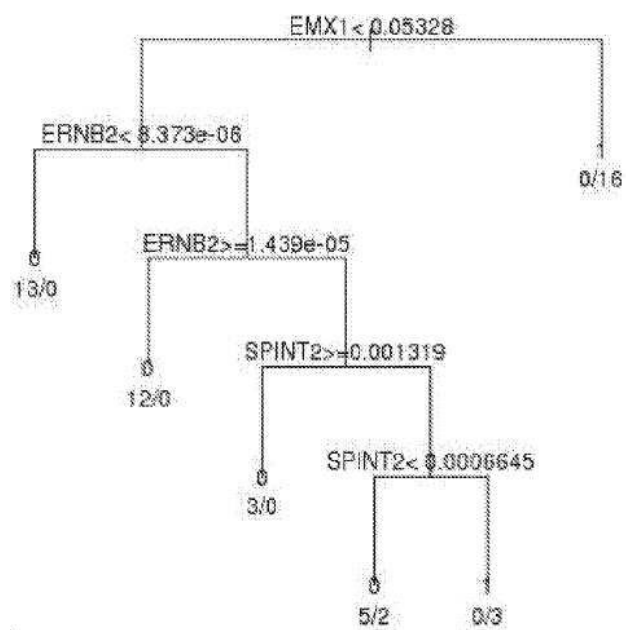
도면7b



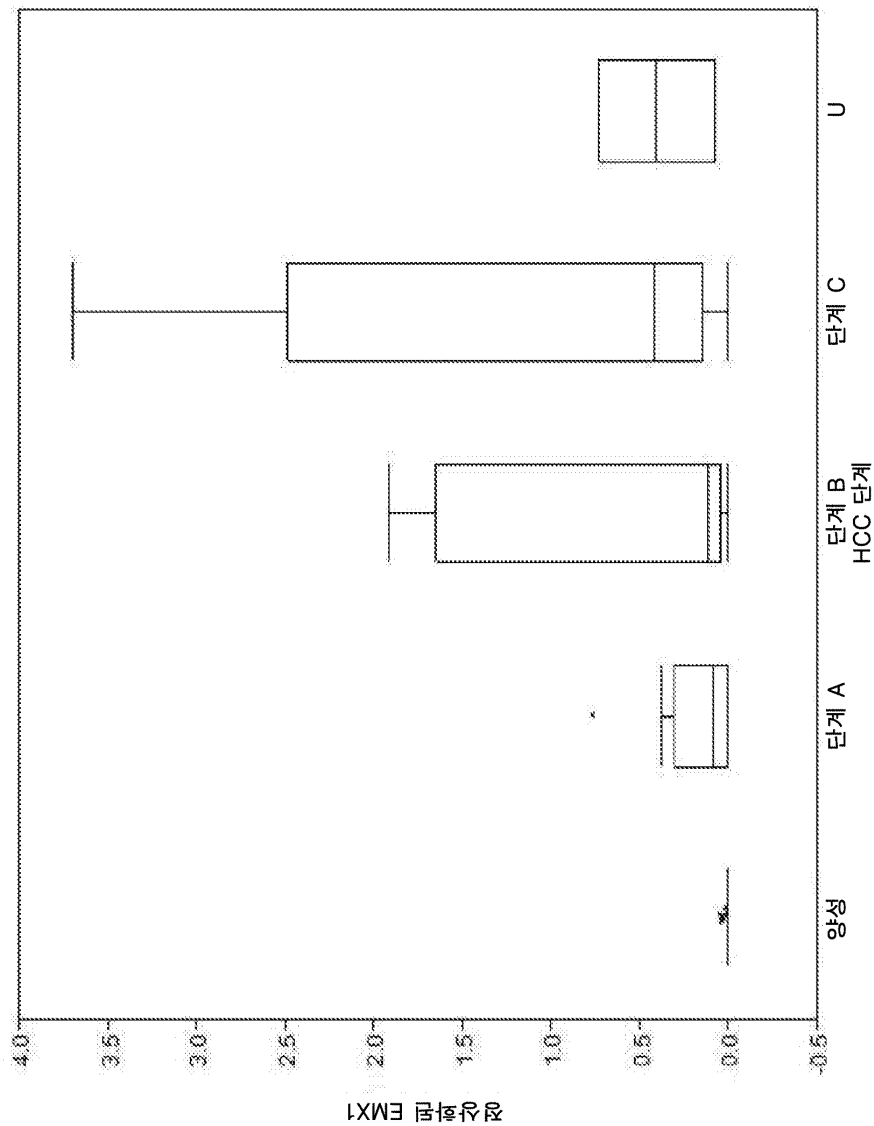
도면7c



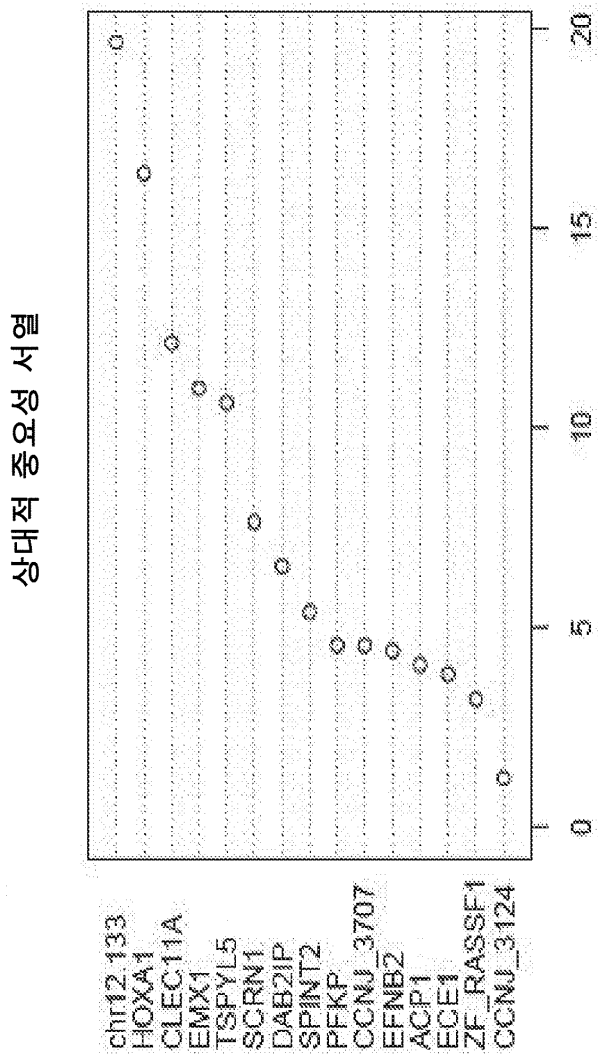
도면7d



도면8



도면9



서열 목록

- <110> EXACT SCIENCES CORPORATION  
MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH
- <120> DETECTING HEPATOCELLULAR CARCINOMA
- <130> EXCTM-35013/WO-1/ORD
- <150> US 62/383,165
- <151> 2016-09-02
- <160> 104
- <170> KoPatentIn version 3.0
- <210> 1
- <211> 17
- <212> DNA
- <213> Artificial Sequence
- <220><223> Synthetic

<400>	1	
gcgcgggagg attttcg		17
<210>	2	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	2	
ccgccaccat aaacgacc		18
<210>	3	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	3	
ccacggacgc gaaatcgaaa t		21
<210>	4	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	4	
cgttcgttac gtcgttttcg t		21
<210>	5	
<211>	17	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	5	
gatcgacgcg actcgac		17
<210>	6	
<211>	21	
<212>	DNA	



<213> Artificial Sequence

<220><223> Synthetic

<400> 6

ccacggacgc tcgacgtcgc c 21

<210> 7

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 7

gcgggagttt ggcgtag 17

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 8

cgcgcaaata ccgaataaac g 21

<210> 9

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 9

ccacggacgg tcggtagatc g 21

<210> 10

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 10

gcgcgttggt tcgtttcg 18

<210> 11

<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	11	
cgtcacctac cgcaaatacg		20
<210>	12	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	12	
ccacggacgg cggataagga g		21
<210>	13	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	13	
agtacgtaag tagagcgcg		19
<210>	14	
<211>	22	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	14	
ctaaaattaa ctacgccgcc gt		22
<210>	15	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	15	
ccacggacgg agaacgttcg a		21

<210> 16  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 16  
 ggcgtcgcgt ttttagaga a 21

<210> 17  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 17  
 ttcttttcg ttcgtataaa atttcgtt 28

<210> 18  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 18  
 ccacggacga tcgggtttta g 21

<210> 19  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 19  
 tgcgtatggt gggcgag 17

<210> 20  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic

<400>	20	
cctaatttac acgtcaacca atcgaa		26
<210>	21	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	21	
ccacggacgg cgcgtgcgtt t		21
<210>	22	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	22	
tttgtttttt tgattaggtg tttaaga		27
<210>	23	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	23	
caccaacctc ataaccttat c		21
<210>	24	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	24	
gacgcggaga tagtgtgtg g		21
<210>	25	
<211>	22	
<212>	DNA	
<213>	Artificial Sequence	

<220><223>	Synthetic	
<400>	25	
gcgttttagt tagatagggc gg		22
<210>	26	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	26	
gaaaaccct tccccgaaac		20
<210>	27	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	27	
cgccgaggcg cagcctaaa		20
<210>	28	
<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	28	
agtcgtttt ttaggtagtt taggcg		26
<210>	29	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	29	
cgaccttac aatgccgc		19
<210>	30	
<211>	20	

<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	30	
cgccgagggg cggtagttgt		20
<210>	31	
<211>	16	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	31	
gggagcggtc gcgtag		16
<210>	32	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	32	
gcacctaact aaacaaaacg aactaac		28
<210>	33	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	33	
cgccgaggcg caaacgcaaa		20
<210>	34	
<211>	17	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	34	
aggtggcgcg tattacg		17

<210>	35	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	35	
cctactaac gcgtcaacc		20
<210>	36	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	36	
cgccgaggcg ctcgattccc		20
<210>	37	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	37	
ttcgatattg ggtgtcgcg		19
<210>	38	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	38	
cgcgaaaacc aaaaacgaaa c		21
<210>	39	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	39	



cgccgaggga ggcggggttc	20
<210> 40	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 40	
gcgttaattt cgcgaggta	19
<210> 41	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 41	
acaatactct tatatattaa cgccgctc	28
<210> 42	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 42	
cgccgaggag gcgacggagg	20
<210> 43	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 43	
tttgtttcgg tttttggcg	19
<210> 44	
<211> 21	
<212> DNA	
<213> Artificial Sequence	

<220><223> Synthetic  
 <400> 44  
 accataaacg accgaaatcg a 21  
 <210> 45  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 45  
 ccacggacgg cgggaggatt t 21  
 <210> 46  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 46  
 cgttcgttac gtcgttttcg t 21  
  
 <210> 47  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 47  
 gatcgacgacg actcgac 17  
 <210> 48  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 48  
 ccacggacgc tcgacgtcgc c 21  
 <210> 49  
 <211> 22  
 <212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 49

gcgttttagt tagatagggc gg 22

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 50

gaaaaccct tccccgaaac 20

<210> 51

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 51

cgccgaggcg cagcctaaa 20

<210> 52

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 52

gcgggagttt ggcgtag 17

<210> 53

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 53

cgcgcaaata ccgaataaac g 21

<210> 54

<211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 54  
 ccacggacgg tcggtagatc g 21  
 <210> 55  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 55  
 gggagcggtc gcgtag 16  
 <210> 56  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 56  
 gcacctaact aaacaaaacg aactaaac 28  
 <210> 57  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 57  
 cgccgaggcg caaacgcaaa 20  
 <210> 58  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 58

gcgcggttggt tcgtttcg	18
<210> 59	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 59	
cgtcacctac cgcaaatacg	20
<210> 60	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 60	
ccacggacgg cggataagga g	21
<210> 61	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 61	
agtacgtaag tagagcgcg	19
<210> 62	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 62	
ctaaaattaa ctacgccgcc gt	22
<210> 63	
<211> 21	
<212> DNA	
<213> Artificial Sequence	

<220><223> Synthetic  
 <400> 63  
 ccacggacgg agaacttcg a 21  
 <210> 64  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 64  
 cgttcgttcg ttatTTTTT tcga 24  
  
 <210> 65  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 65  
 gcccgcgaaa accaaaaac 19  
 <210> 66  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 66  
 cgccgaggcg aaactcacct 20  
 <210> 67  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 67  
 gcgtTTTTT ttagcgggt ta 22  
  
 <210> 68  
 <211> 21

<212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 68  
 ccgaaactaa aattctcccg c 21  
 <210> 69  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 69  
 cgccgaggat gagcgtgtta 20  
 <210> 70  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 70  
 cggttttcgt ttgggtacg 19  
  
 <210> 71  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 71  
 ccaacccaaa ccacgcc 17  
 <210> 72  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 72  
 ccacggacgc gcgccgtacg a 21  
 <210> 73

<211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 73  
 ggaggttggc ggggag 16

<210> 74  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 74  
 catatactat cgccttccga ctc 23

<210> 75  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 75  
 cgccgaggcg taacaaaaac 20

<210> 76  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 76  
 tcgttttagg tgagtcgcg 19

<210> 77  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 77



aaataaaacc gccgaaaaa caac	24
<210> 78	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 78	
ccacggacgc gctacaaacg c	21
<210> 79	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 79	
ggaggggttt cgttcgt	17
<210> 80	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 80	
ctactatcga cgctaaaaat aaaacgaac	29
<210> 81	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	
<400> 81	
cgccgaggcg cgacctaaaa	20
<210> 82	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> Synthetic	

<400> 82  
aggtggcgcg tattacg 17

<210> 83  
<211> 20  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic  
<400> 83  
cctactaac gcgtcaacc 20

<210> 84  
<211> 20  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic  
<400> 84  
cgccgaggcg ctcgattccc 20

<210> 85  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic  
<400> 85  
gcgttaattt cgcgaggta 19

<210> 86  
<211> 28  
<212> DNA  
<213> Artificial Sequence  
<220><223> Synthetic  
<400> 86  
acaatactct tatatattaa cgccgctc 28

<210> 87  
<211> 20  
<212> DNA

<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	87	
cgccgaggag	gcgacggagg	20
<210>	88	
<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	88	
agtcgttttt	ttaggtagtt taggcg	26
<210>	89	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	89	
cgacctttac	aatcgccgc	19
<210>	90	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	90	
cgccgagggg	cggtagtgt	20
<210>	91	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	91	
ggcgtcgcgt	tttttagaga a	21
<210>	92	

<211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 92  
 ttctttttcg ttcgtataaa atttcgt 27  
 <210> 93  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 93  
 ccacggacga tcgggtttta g 21  
 <210> 94  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 94  
 cgccgaggat cgggttttag 20  
  
 <210> 95  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 95  
 agccggtttt ccggctgaga ctccgcgtc 29  
 <210> 96  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature  
 <222> (3)

<223> modified with quencher  
 <400> 96  
 tctagccggt tttccggtg agactccgcg tc 32  
 <210> 97  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> Synthetic  
 <400> 97  
 agccggtttt ccggctgaga cgtccgtgg 29  
 <210> 98  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature  
 <222> (3)  
 <223> modified with quencher  
 <400> 98  
 tctagccggtt ttccggtga gacgtccgtg g 31  
 <210> 99  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <400> 99  
 agccggtttt ccggctgaga cctcggcg 28  
  
 <210> 100  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Synthetic  
 <220><221> misc\_feature

<222>	(3)	
<223>	modified with quencher	
<400>	100	
tctagccggt tttccggctg agacctcggc g		31
<210>	101	
<211>	8	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	101	
cgccgagg		8
<210>	102	
<211>	9	
<212>	DNA	
<213>		
>	Artificial Sequence	
<220><223>	Synthetic	
<400>	102	
gacgcggag		9
<210>	103	
<211>	9	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	103	
ccacggacg		9
<210>	104	
<211>	8	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Synthetic	
<400>	104	
gcgcgtcc		8