



등록특허 10-2688282



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년07월26일
(11) 등록번호 10-2688282
(24) 등록일자 2024년07월22일

- (51) 국제특허분류(Int. Cl.)
C12Q 1/6886 (2018.01)
- (52) CPC특허분류
C12Q 1/6886 (2022.01)
C12Q 2600/154 (2013.01)
- (21) 출원번호 10-2018-7009227
- (22) 출원일자(국제) 2016년08월31일
심사청구일자 2021년08월31일
- (85) 번역문제출일자 2018년03월30일
- (65) 공개번호 10-2018-0081042
- (43) 공개일자 2018년07월13일
- (86) 국제출원번호 PCT/US2016/049653
- (87) 국제공개번호 WO 2017/040627
국제공개일자 2017년03월09일
- (30) 우선권주장
62/212,221 2015년08월31일 미국(US)
- (56) 선행기술조사문헌
US20140274748 A1
World J Gastroenterol, 20(34): 11991-2006
(2014.09.14.)
Mini Rev Med Chem, 10(11): 1034-1038
(2010.10.31.)
Sci Transl Med, 4(156): 156ra140
(2012.10.17.)

- (73) 특허권자
메이오 파운데이션 포 메디컬 에큐에이션 앤드 리
씨치
미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200
이그제クト 사이언시스 코포레이션
미국 위스콘신 53719 메디슨 인데비 레인 5505
- (72) 발명자
알라위, 하림 티.
미국 53562 위스콘신 미들턴 원딩 웨이 901
알퀴스트, 테이비드 에이.
미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200
(뒷면에 계속)
- (74) 대리인
특허법인 무한

전체 청구항 수 : 총 2 항

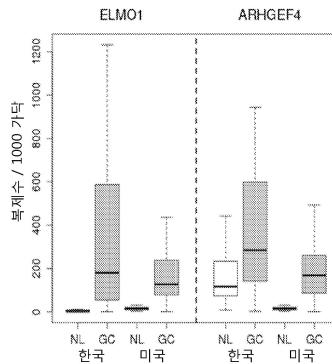
심사관 : 이준혁

(54) 발명의 명칭 위 신생물 검출 방법

(57) 요약

본원에는 신생물 검출에 관련하는 기술, 특히, 배타적이지 않지만, 위암과 같은 악성 신생물 및 전악성 신생물을 검출하는 방법, 조성물 및 관련 용도에 관한 기술이 제공된다.

대 표 도



(52) CPC특허분류

C12Q 2600/16 (2013.01)

(72) 발명자

테일러, 윌리암 알.

미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200

키시엘, 존 비.

미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200

얍, 트레이시 씨.

미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200

마호니, 더글라스 더블유.

미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200

명세서

청구범위

청구항 1

생물학적 샘플을 특성화하는 방법으로서,

(a) (aa) 생물학적 샘플에서 게놈 DNA를 비설파이트(bisulfite)로 처리하는 단계,

(bb) 선택된 2개 이상의 유전자에 대한 프라이머를 사용하여 상기 비설파이트 처리된 게놈 DNA를 증폭하는 단계, 및

(cc) 메틸화 특이적 PCR, 정량적 메틸화 특이적 PCR, 메틸화-민감성 DNA 제한 효소 분석, 정량적 비설파이트 파이로서열 결정(pyrosequencing) 또는 비설파이트 게놈 서열 결정 PCR에 의해 CpG 부위의 메틸화 수준을 결정하는 단계

를 통해 인간 개체의 생물학적 샘플에서 ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 및 ZNF569로부터 선택된 2개 이상의 유전자에 대해 CpG 부위의 메틸화 수준을 측정하는 단계;

(b) 상기 메틸화 수준을, 위암에 걸리지 않은 대조군 샘플에서 유전자의 상응하는 세트의 메틸화 수준과 비교하는 단계; 및

(c) 상기 2개 이상의 유전자에서 측정된 메틸화 수준이 상기 각각의 대조군 샘플에서 측정된 메틸화 수준보다 높은 경우, 상기 인간 개체가 위암에 걸린 것으로 결정하는 단계;를 포함하고,

이때 상기 2개 이상의 유전자는 적어도 ELM01 및 ZNF569를 포함하고,

상기 생물학적 샘플은 위 조직 또는 혈장인 것인, 방법.

청구항 2

제1항에 있어서,

상기 선택된 2개 이상의 유전자에 대한 프라이머는 다음의 프라이머 세트로부터 선택되는 것인, 방법:

ELM01에 대해, SEQ ID NOS: 21 및 22로 이루어진 프라이머 세트,

ARHGEF4에 대해, SEQ ID NOS: 23 및 24로 이루어진 프라이머 세트,

EMX1에 대해, SEQ ID NOS: 25 및 26로 이루어진 프라이머 세트,

SP9에 대해, SEQ ID NOS: 27 및 28로 이루어진 프라이머 세트,

CLEC11A에 대해, SEQ ID NOS: 29 및 30로 이루어진 프라이머 세트,

ST8SIA1에 대해, SEQ ID NOS: 13 및 14로 이루어진 프라이머 세트,

BMP3에 대해, SEQ ID NOS: 31 및 32로 이루어진 프라이머 세트,

KCNA3에 대해, SEQ ID NOS: 33 및 34로 이루어진 프라이머 세트,

DMRTA2에 대해, SEQ ID NOS: 35 및 36로 이루어진 프라이머 세트,

KCNK12에 대해, SEQ ID NOS: 37 및 38로 이루어진 프라이머 세트,

CD1D에 대해, SEQ ID NOS: 39 및 40로 이루어진 프라이머 세트,

PRKCB에 대해, SEQ ID NOS: 41 및 42로 이루어진 프라이머 세트,

CYP26C1에 대해, SEQ ID NOS: 43 및 44로 이루어진 프라이머 세트,
ZNF568에 대해, SEQ ID NOS: 45 및 46로 이루어진 프라이머 세트,
ABCB1에 대해, SEQ ID NOS: 9 및 10로 이루어진 프라이머 세트,
ELOVL2에 대해, SEQ ID NOS: 47 및 48로 이루어진 프라이머 세트,
PKIA에 대해, SEQ ID NOS: 49 및 50로 이루어진 프라이머 세트,
SFMBT2(893)에 대해, SEQ ID NOS: 5 및 6로 이루어진 프라이머 세트,
PCBP3에 대해, SEQ ID NOS: 51 및 52로 이루어진 프라이머 세트,
MATK에 대해, SEQ ID NOS: 53 및 54로 이루어진 프라이머 세트,
GRN2D에 대해, SEQ ID NOS: 55 및 56로 이루어진 프라이머 세트,
NDRG4에 대해, SEQ ID NOS: 57 및 58로 이루어진 프라이머 세트,
DLX4에 대해, SEQ ID NOS: 59 및 60로 이루어진 프라이머 세트,
PPP2R5C에 대해, SEQ ID NOS: 61 및 62로 이루어진 프라이머 세트,
FGF14에 대해, SEQ ID NOS: 63 및 64로 이루어진 프라이머 세트,
ZNF132에 대해, SEQ ID NOS: 65 및 66로 이루어진 프라이머 세트,
CHST2 (7890)에 대해, SEQ ID NOS: 67 및 68로 이루어진 프라이머 세트,
FLI1에 대해, SEQ ID NOS: 69 및 70로 이루어진 프라이머 세트,
c13orf18에 대해, SEQ ID NOS: 19 및 20로 이루어진 프라이머 세트, 및
ZNF569에 대해, SEQ ID NOS: 11 및 12로 이루어진 프라이머 세트.

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

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

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

발명의 설명

기술분야

[0001] 관련 출원 상호 참조

[0002] 본 출원은 2016년 8월 31일에 출원된 미국 가 출원 제62/212,221호에 대한 우선권 및 이익을 주장하고, 이는 본 원에 참조로 포함되어 있다.

[0003] 기술분야

[0004] 본원에 신생물을 검출하는 기술, 특히 배타적이지 않지만, 위암과 같은 전악성 및 악성 신생물을 검출하기 위한 방법, 조성물 및 관련 용도에 관한 것이다.

배경 기술

- [0005] 위암은 세계에서 암 관련 사망의 3번째로 가장 흔한 원인이고(예를 들면, 세계보건기구 Cancer: Fact Sheet No 297. WHO 참조), 서양 국가에서 대부분의 환자가 진행된 질병을 나타내고 있기 때문에 치료하는 것이 곤란하다. 미국에서, 위 악성종양은 현재 15번째로 가장 흔한 암이다(예를 들면, Surveillance, Epidemiology, and End Results Program. SEER Stat Fact Sheets: Stomach Cancer. National Cancer Institute 참조). 그러나, 대부분의 선진국에서는, 과거 반세기에 걸쳐 위암률이 매우 감소했다.
- [0006] 위암의 감소는, 부분적으로 냉동의 광범위한 사용에 기인하고, 이는 여러 유익한 효과 있었는데, 이러한 효과로는, 신선한 고기 및 야채의 소비 증가; 식품 방지제로서 사용되고 있는 소금의 섭취 감소; 및 비냉동 고기 제품의 부패에 의한 발암성 화합물에 의한 식품의 오염 감소이다. 염분 및 염분이 들어간 식품은 위 점막을 손상시켜 염증을 일으키고 DNA 합성 및 세포 증식을 증가시킬 가능성이 있다. 위암률의 감소에 기인할 가능성이 있는 그 외의 요인으로는, 위생상태의 개선 및 항생물의 사용에 의한 만성 헬리코박터 파이로리(*Helicobacter pylori*) 감염률 감소 및 일부 국가에서 스크리닝의 증가를 포함한다(예를 들면, Global Cancer Facts & Figures, 3rd ed. American Cancer Society 참조).
- [0007] 그럼에도 불구하고, 서양 국가에서 위암은 주로 대부분의 환자가 진행된 질병을 나타내기 때문에 치료하는 것이 곤란했다. 가장 호전된 상태를 나타내고 치료적 외과적 절제를 받은 환자조차 종종 재발하는 질병에 의해 사망한다.
- [0008] 따라서, 위암을 조기 검출하기 위한 개선된 방법 및 기술이 요구된다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0009] 메틸화 DNA는 대부분 종양 타입의 조직에서 바이오마커의 잠재적인 클래스로서 연구되었다. 대부분의 경우 DNA 메틸트랜스퍼라제는 유전자 발현의 후성적 제어로서 시토신-포스페이트-구아닌(CpG) 섬 부위에서 DNA에 메틸기를 부가한다. 생물학적으로 매력적인 메커니즘에서, 종양 억시약 유전자의 프로모터 영역에서 후천적인 메틸화 사상은 발현을 억제해서, 종양 형성에 기여하는 것으로 고려된다. DNA 메틸화는 RNA 또는 단백질 발현보다 화학적 및 생물적으로 안정한 진단 툴일 수 있다 (Laird (2010) "Principles and challenges of genome-wide DNA methylation analysis" *Nat Rev Genet* 11: 191-203). 또한, 산발성(sporadic) 결장암과 같은 그 외의 암에서, 메틸화 마커는 우수한 특이성을 나타내고 개개의 DNA 돌연변이보다 광범위하게 유의하고 민감하다(Zou et al (2007) "Highly methylated genes in colorectal neoplasia: implications for screening" *Cancer Epidemiol Biomarkers Prev* 16: 2686-96).
- [0010] CpG 섬의 분석은, 동물 모델 및 인간 세포주에 적용되는 경우 중요한 발견을 나타냈다. 예를 들면, Zhang 및 동료들은 동일한 CpG 섬의 상이한 부분으로부터의 앰플리콘의 메틸화의 상이한 수준을 가질 수 있다는 것을 찾았습니다 (Zhang et al. (2009) "DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution" *PLoS Genet* 5: e1000438). 또한, 메틸화 수준은 고도 메틸화된 서열과 메틸화되지 않은 서열 사이에서 2차원적으로 분포하고 DNA 메틸트랜스퍼라제 활성의 2상 스위치 유사 패턴을 지지했다 (Zhang et al. (2009) "DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution" *PLoS Genet* 5: e1000438). 생체내 뮤린 조직 및 시험관 내 세포주의 분석은, 약 0.3%의 높은 CpG 밀도 프로모터(HCP, 300 염기쌍 영역 내에서 7% 초과의 CpG 서열을 갖는 것으로 정의됨)가 메틸화되는 반면, 낮은 CpG 밀도의 영역(300 염기쌍 영역 내에서 5% 미만의 CpG 서열을 갖는 것으로 정의됨)은 동적 조직 특이적 패턴에서 빈번하게 메틸화되는 경향이 있다 (Meissner et al. (2008) "Genome-scale DNA methylation maps of pluripotent and differentiated cells" *Nature* 454: 766-70). HCP는 유비쿼터스 하우스키핑 유전자 및 고도로 조절된 발현 유전자의 프로모터를 포함한다. HCP 중에서 50% 초과의 메틸화된 부위는 Wnt 2, NDRG2, SFRP2, 및 BMP3 등의 여러 확립된 마커가 있었다 (Meissner et al. (2008) "Genome-scale DNA methylation maps of pluripotent and differentiated cells" *Nature* 454: 766-70).
- [0011] 총체적으로, 위장암은 임의의 그 외의 장기계(organ system)보다 암사망율이 높다. 그러나, 현재 결장직장암만

이 국내에서 스크리닝되고 있다. 상부 GI 암에 의한 매년 US 사망률은 결장직장암의 약 50,000에 비해 90,000을 초과한다. 위의 선암은 세계에서 2번째로 많은 암사망 원인이고 연간 약 100만명이다. 태평양 아시아, 러시아, 및 중남미의 특정 지역에서 가장 일반적이다. 대부분의 환자는 진행된 질병을 나타내기 때문에 예후는 불량하다(5년 생존률 5-15% 미만).

[0012] 위암의 기본적인 계놈 메커니즘은 복잡하고 불완전하게 이해된다. 환경 노출과 함께 *Helicobacter pylori* 감염에 의한 만성 위염이 병인에 기여한다. 염색체 및 마이크로사텔라이트(microsatellite)의 불안정성은 위암뿐만 아니라 후성적 변화에 관련된다. CpG 메틸레이터 표현형은 24%-47%의 증례로 기재되어 있다. 잘 관찰된 돌연변이는 유전자 p53, APC, β -catenin, 및 k-ras을 포함하지만, 이들은 사실상 비교적 드물고 이질적이다.

[0013] 위암은 세포 및 DNA를 소화류에 흘려보내고 궁극적으로 대변으로 배설된다. 절단된 종양이 동일한 서열을 포함하는 것으로 알려진 위암 환자의 매칭된 대변에서 돌연변이 DNA를 검출하기 위해 사전에 고감도 분석이 사용되었다. 그러나, 돌연변이 마커의 몇 개의 제한은, 근본적인 이질성과 검출이 곤란한 공정에 관련되고, 일반적으로 다수 유전자에 걸친 각각의 돌연변이 부위는 고감도를 달성하기 위해 별도로 분석되어야 한다.

[0014] DNA 메틸트랜스퍼라제에 의해 시토신-포스페이트-구아닌(CpG) 섬에서 DNA의 후성적 메틸화는, 대부분 종양 유형의 조직에서 바이오마커의 잠재적인 클래스로서 연구되었다. 생물학적으로 매력적인 메커니즘에서, 종양 억제 유전자의 프로모터 영역에서 후천적 메틸화 사상은 발현을 억제해서, 종양 형성에 기여하는 것으로 고려된다. DNA 메틸화는 RNA 또는 단백질 발현보다 화학적 및 생물학적으로 안정한 진단 툴일 수 있다. 또한, 산발성 결장암과 같은 그 외의 암에서, 이상한 메틸화 마커는 개개 DNA 돌연변이 보다도 광범위하게 유익하고 민감하고 우수한 특이성을 제공한다.

[0015] 신규한 메틸화 마커를 검색하기 위한 몇 개의 방법이 이용 가능하다. CpG 메틸화 마이크로어레이 기반 조사가 합리적이고 고-쓰로우 풋 접근방법이지만, 이러한 전략은 주로 확립된 종양 억제자 프로모터의 알려진 관심 영역에 편향되어 있다. DNA 메틸화의 계놈-와이드 분석의 또 다른 방법은 지난 10년간 개발되어 왔다. 3개의 기본적인 접근 방법이 있다. 제1 방법은 특이적 메틸화 부위를 인식하는 제한 효소에 의해 DNA를 소화한 후 정량적인 단계(예를 들면, 메틸화 특이적 PCR; MSP)에서 DNA를 증폭하기 위해 사용되는 프라이머 또는 효소 인식 부위에 제한된 메틸화 데이터를 제공하는 몇 개의 가능한 분석 기술을 수행한다. 제2 방법은 메틸 시토신 또는 그 외의 메틸화 특이적 결합 도메인에 관한 항체를 사용하는 계놈 DNA의 메틸화된 분획을 풍부하게 하고 레페런스 계놈에 대해 단편을 맵핑하기 위한 마이크로어레이 분석 또는 그 서열 결정을 수행한다. 이러한 접근방법은 단편 내에 모든 메틸화된 부위의 단일 뉴클레오티드 분해를 제공하지 않는다. 제3 방법은 모든 메틸화되지 않은 시토신을 우라실로 전환하기 위해 DNA의 비설파이트(bisulfite) 처리를 시작해서, 제한 효소 소화, 및 어댑터 리간드에 커플링 후에 모든 단편의 완전한 서열 결정을 수행한다. 제한 효소의 선택은 CpG 농후한 영역의 단편을 풍부하게 하고, 분석 중에 다수 유전자 위치에서 맵핑할 수 있도록 풍부한 서열의 수를 감소시킬 수 있다. RRBS(Reduced Representation Bisulfite Sequencing)라고 하는 후자의 방법은 위암을 연구하기 위해 아직 사용되고 있지 않다.

[0016] RRBS는 모든 CpG 섬의 80-90%의 단일 뉴클레오티드 분해능 및 중간에서부터 높은 판독 범위(read coverage)에서 종양 억제 인자 프로모터의 대부분에서 CpG 메틸화 상태 데이터를 생성했다. 암 증례-대조군 연구에서, 이들 판독의 분석에 의해 차별적으로 메틸화된 영역(DMR)을 확인한다. 퀘장암 표본의 이전의 RRBS 분석에서, 수백 개의 DMR이 발견되었지만, 대부분은 발암에 관련된 것이 아니고 대부분은 주석을 갖고 있지 않았다. 또한, 독립적인 조직 샘플에 관련된 추가의 검증 연구는 성능의 점에서 특이적이고 100% 감도를 갖는 마커-CpG를 확인했다.

[0017] 고도로 판별된 마커의 임상 적용은 큰 영향을 줄 가능성이 있다. 예를 들면, 대변 또는 혈액과 같은 원거리 배지(distant media)에서 이러한 마커의 분석에 의해 위 신생물의 검출을 위한 정확한 스크리닝 또는 진단 툴이 될 수 있다.

[0018] 본 기술의 실시형태를 개발하는 과정에서 실시된 실험은, 위암 샘플과 유사한 방법으로 생성된 123개의 신규 DNA 메틸화 마커를 열거하고 기재한다. 위암과 함께, 정상적인 위 조직, 정상적인 결장조직, 및 정상적인 백혈구 DNA가 서열 결정되었다. 마커는 가장 퉁퉁한 후보를 확인하고 최적화하기 위해 더 큰 샘플 세트에 대해 검증되었다. 추가의 실험은 위암을 검출하기 위해 10개의 최적 마커(ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 및 C13ORF18; 실시에 2 및 표 4, 5, 및 6 참조)를 확인했다. 추가의 실험은 위암을 검출하기 위한 30개의 최적 마커(ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK,

GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 또는 ZNF569; 실시예 1 및 표 1, 2 및 3 참조)을 확인했다. 추가의 실험은 위암(gastric cancer)(예를 들면, 위암(stomach cancer))을 검출하기 위한 12개의 최적 마커(ELMO1, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; 실시예 3 및 표 7 및 8 참조)를 확인했다.

[0019] 따라서, 본원에는 피험체로부터 채취한 샘플(예를 들면, 대변 샘플, 위 조직 샘플; 혈장 샘플)로부터 검출하는 경우 높은 신호 대 잡음 비 및 낮은 백그라운드 수준을 제공하는 위암 스크리닝 마커의 기술이 본원에 제공된다. 마커는 위암에 걸린 피험체의 위 조직 또는 혈장으로부터 DNA 마커의 메틸화 상태를, 대조군 피험체의 위 조직 또는 혈장으로부터 DNA 마커의 메틸화 상태와 비교함으로써 증례-대조군 연구에서 확인되었다. (예를 들면, 정상 위 조직, 정상 결장 상피세포, 정상 혈장, 및 정상 백혈구 유래 DNA, 실시예 1, 2 및 3, 및 표 1-8 참조) (예를 들면, 혈장 조직; 실시예 3 참조).

[0020] 본원에 기재된 바와 같이, 본 기술은 다수의 메틸화 DNA 마커 및 이들의 서브셋의 수(예를 들면, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12개 이상의 마커의 세트)를 제공하고, 전체적으로 위암과 잘 구별된다. 실험은 선별 필터를 후보 마커에 적용하여 높은 신호 대 잡음 비 및 낮은 백그라운드 수준을 제공하는 마커를 확인하고 위암 스크리닝 또는 진단의 목적으로 높은 특이성을 제공한다.

[0021] 일부 실시형태에서, 본 기술은 생물학적 샘플(예를 들면, 위 조직, 혈장 샘플)에서 본원에서 확인된 하나 이상의 마커의 메틸화 상태 및 그 존재의 분석에 관련된다. 이를 마커는 예를 들면, 표 1, 2, 4 및/또는 7에 제공된 본원에서 검토된 하나 이상의 차별적으로 메틸화된 영역(DMR)을 포함한다. 메틸화 상태가 본 기술의 실시형태에서 분석된다. 따라서, 본원에 제공된 본 기술은 유전자의 메틸화 상태를 측정하는 방법에서 제한되지 않는다. 예를 들면, 일부 실시형태에서 메틸화 상태는 게놈 스크리닝 방법에 의해서 측정된다. 예를 들면, 하나의 방법은 제한 랜드마크 게놈 스캐닝을 수반하고 (Kawai et al. (1994) *Mol. Cell. Biol.* 14: 7421-7427), 또 다른 예는 메틸화-민감성의 임의로 프라이밍된 PCR을 포함한다 (Gonzalgo et al. (1997) *Cancer Res.* 57: 594-599). 일부 실시형태에서, 특이적 CpG 부위에서 메틸화 패턴의 변화는 메틸화-민감성 제한 효소로 게놈 DNA를 소화하고 모니터링된 후, 관심 영역의 서던(Southern) 분석을 수행한다 (소화-서던 방법). 일부 실시형태에서, 메틸화 패턴의 변화를 분석하는 것은, PCT 증폭 전에 메틸화-민감성 제한 효소에 의해 게놈 DNA의 소화를 수반하는 PCR-기반 공정을 수반한다 (Singer-Sam et al. (1990) *Nucl. Acids Res.* 18: 687). 또한, 그 외의 기술은 메틸화 분석의 출발점으로서 DNA의 비설파이트 처리를 이용하는 것을 보고했다. 이들은 메틸화 특이적 PCR (MSP) (Herman et al. (1992) *Proc. Natl. Acad. Sci. USA* 93: 9821-9826) 및 비설파이트 전환된 DNA로부터 증폭된 PCR 생성물의 제한 효소 소화 (Sadri and Hornsby (1996) *Nucl. Acids Res.* 24: 5058-5059; and Xiong and Laird (1997) *Nucl. Acids Res.* 25: 2532-2534)를 포함한다. PCR 기술은 유전자 돌연변이 검출 (Kuppuswamy et al. (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 et al. (1992) *PCR Methods Appl.* 1: 160-163)을 위해 개발했다. 이러한 기술은 내부 프라이머를 사용하고, 이는 PCR 생성된 템플레이트에 어닐링되고, 분석될 단일 뉴클레오티드의 5' 말단을 바로 종결시킨다. 미국 특허 제7,037,650호에 기재된 "quantitative Ms-SNuPE assay"를 사용하는 방법이 일부 실시형태에서 사용되고 있다.

[0022] 메틸화 상태를 평가하면, 메틸화 상태는 종종 특정 부위를 포함하는 샘플에서 DNA의 전체 집단에 대해 특정 부위(예를 들면, 단일 뉴클레오티드에서, 특정 영역 또는 유전자좌에서, 관심의 긴 서열에서, 예를 들면, DNA의 ~100-bp, 200-bp, 500-bp, 1000-bp 서브시퀀스 이상까지)에서 메틸화되는 DNA의 개개의 가닥의 비율 또는 백분율로서 표현된다. 전통적으로, 비메틸화 핵산의 양은 캘리브레이터를 사용하여 PCR에 의해 결정된다. 그 다음, 공지된 DNA의 양은 비설파이트 처리되고, 수득된 메틸화 특이적 서열은 실시간 PCR 또는 그 외의 지수함수적 증폭, 예를 들면, QuARTS 분석을 사용하여 결정된다(예를 들면, 미국 특허 제8,361,720호; 및 미국 특허 출원 공개 제2012/0122088 및 2012/0122106호에 의해 제공되고 본원에 참조로 포함되어 있다).

[0023] 예를 들면, 일부 실시형태에서 외부 표준을 사용하여 메틸화되지 않은 표적의 표준 곡선을 생성하는 것을 포함한다. 표준 곡선은 적어도 2개의 지점으로부터 구축되고 메틸화되지 않은 DNA의 실시간 Ct 값을 공지된 정량적인 표준에 관련시킨다. 그 다음에, 메틸화 표적의 제2 표준 곡선은 적어도 2개의 지점 및 외부 표준으로부터 구축된다. 이러한 제2 표준 곡선은 메틸화 DNA의 Ct를 공지된 정량적인 표준에 관련시킨다. 다음, 시험 샘플 Ct 값은 메틸화된 및 메틸화되지 않은 집단에 대해 결정되고, DNA의 게놈 등가물은 제1의 2단계에 의해 생성된 표준곡선으로부터 산출된다. 관심 부위에서 메틸화의 백분율은 예를 들면, (메틸화된 DNAs의 수) / (메틸화된 DNAs의 수 + 메틸화되지 않은 DNAs의 수) × 100과 같은 집단에서 DNA의 전체량에 대해 메틸화된 DNA의 양으로

부터 산출된다.

[0024]

본원에는 조성물 및 방법을 수행하기 위한 키트가 제공된다. 예를 들면, 일부 실시형태에서, 하나 이상의 마커에 대해 특이적인 시약(예를 들면, 프라이머, 탐침)은 단독으로 또는 세트로 제공된다(예를 들면, 복수의 마커를 증폭하기 위한 프라이머 쌍의 세트). 검출 분석을 행하기 위한 추가의 시약(예를 들면, 효소, 버퍼, QuARTS를 수행하기 위한 양 및 음의 대조군, PCR, 서열 결정, 비설파이트, 또는 그 외의 분석)은 또한 제공될 수 있다. 일부 실시형태에서, 방법을 수행하기 위해 필요하거나, 충분하거나 유용한 하나 이상의 시약을 포함하는 키트가 제공된다. 또한, 시약을 포함하는 반응 혼합물이 제공된다. 또한, 반응 혼합물을 완료하기 위해 서로 및/또는 시험 샘플에 첨가될 수 있는 복수의 시약을 포함하는 주요한 혼합 시약(master mix reagent) 세트가 제공된다.

[0025]

일부 실시형태에서, 본원에 기재된 기술은 본원에 기재된 방법에 의해 제공된 일련의 산술연산 또는 논리연산을 실행하도록 설계된 프로그램 가능한 기기에 관련된다. 예를 들면, 기술의 일부 실시형태는 컴퓨터 소프트웨어 및/또는 컴퓨터 하드웨어(예를 들면, 이들로 실행된)와 관련된다. 일 형태에서, 본 기술은 메모리의 형태, 산술 및 논리 연산을 수행하는 소자, 및 데이터를 관독, 조작, 저장하기 위한 일련의 명령(예를 들면, 본원에 제공된 방법)을 수행하기 위한 처리 소자(예를 들면, 마이크로프로세서)를 포함하는 컴퓨터에 관련된다. 일부 실시형태에서, 마이크로프로세서는 (예를 들면, 하나 이상의 DMR의, 예를 들면, 표 1, 2, 4 및 7에 제공된 DMR 1-274의) 메틸화 상태를 결정하기 위한 시스템의 일부이고; (예를 들면, 하나 이상의 DMR의, 예를 들면, 표 1, 2, 4 및 7에 제공된 DMR 1-274의) 메틸화 상태를 비교하고; 표준 곡선을 생성하고; Ct 값을 결정하고; 예를 들면, 하나 이상의 DMR의 (예를 들면, 하나 이상의 DMR의, 예를 들면, 표 1, 2, 4 및 7에 제공된 DMR 1-274의) 메틸화 비율, 빈도 또는 백분율을 산출하고; CpG 섬을 확인하고; 분석 또는 마커의 특이성 및/또는 감도를 결정하고; ROC 곡선 또는 관련된 AUC를 산출하고; 서열을 분석하기 위한 시스템의 일부이고; 모두 본원에 기재되어 있고 당해 기술분야에 공지되어 있다.

[0026]

일부 실시형태에서, 마이크로프로세서 또는 컴퓨터는 암의 부위를 예측하기 위해 알고리즘에서 메틸화 상태 데이터를 사용한다.

[0027]

일부 실시형태에서, 소프트웨어 또는 하드웨어 성분은 다수의 분석의 결과를 수용하고 다수 분석(예를 들면, 표 1, 2, 4, 및 7에 제공된 다수의 DMR의 메틸화 상태를 결정하는 것)의 결과에 기초한 암 위험을 나타낸 사용자에게 보고하기 위해 단일 값 결과를 결정한다. 관련 실시형태는 다수 분석 결과의 수학적 조합(예를 들면, 각종 조합, 선형 조합)에 기초한 위험 인자를 산출하고, 다수 마커의 메틸화 상태를 결정한다(예를 들면, 표 1, 2, 4, 및 7에 제공된 다수의 DMR의 메틸화 상태를 결정한다). 일부 실시형태에서, DMR의 메틸화 상태는 치수를 정의하고 다차원 공간의 값을 갖고 복수의 DMR의 메틸화 상태에 의해서 정의된 좌표는 예를 들면, 암 위험에 관련된 사용자에게 보고하는 결과이다.

[0028]

일부 실시형태는 저장 매체 및 메모리 성분을 포함한다. 메모리 성분(예를 들면, 회발성 및/또는 비회발성 메모리)은 명령(예를 들면, 본원에 제공된 공정의 실시형태) 및/또는 데이터(예를 들면, 메틸화 측정, 시퀀스, 및 이와 관련된 통계적 설명과 같은 워크피스)에서 사용되는 것을 알 수 있다. 일부 실시형태는 (예를 들면, 디스플레이와 같은 출력 소자 및 키보드와 입력 소자를 포함하는)CPU, 그래픽 카드 및 사용자 인터페이스 중 하나 이상을 포함하는 시스템에 관련된다.

[0029]

본 기술과 관련된 프로그램 가능한 기기(예를 들면, 양자 컴퓨터, 화학 컴퓨터, DNA 컴퓨터, 광학 컴퓨터, 스핀 트로닉 기반 컴퓨터 등)는 기존의 현존하는 기술 및 개발중인 기술 또는 아직 개발되지 않은 기술을 포함한다.

[0030]

일부 실시형태에서, 본 기술은 데이터를 전송하기 위한 유선(예를 들면, 금속 케이블, 광 섬유) 또는 무선 전송 매체를 포함한다. 예를 들면, 일부 실시형태는 네트워크(예를 들면, 근거리 통신망(LAN), 광역통신망(WAN), 애드혹 네트워크, 인터넷 등)를 통한 데이터 전송에 관련된다. 일부 실시형태에서, 프로그램 가능한 기기는 피어(peer)와 같은 네트워크 상에 존재하고, 일부 실시형태에서 프로그램 가능한 기기는 클라이언트/서버 관계를 갖는다.

[0031]

일부 실시형태에서, 데이터는 하드디스크, 플래시메모리, 광 매체, 플로피디스크 등과 같은 컴퓨터 관독 가능한 저장 매체에 저장된다.

[0032]

일부 실시형태에서, 본원에 제공된 기술은 본원에 기재된 방법을 수행하기 위해 함께 동작하는 복수의 프로그램 가능한 디바이스와 관련된다. 예를 들면, 일부 실시형태에서, (예를 들면, 네트워크에 의해 연결된)복수의 컴퓨터는 클러스터 컴퓨터 또는 그리드 컴퓨팅 또는 이더넷, 광 섬유와 같은 종래의 네트워크 인터페이스 또는 무

선 네트워크 기술에 의해 네트워크(개인, 공용 또는 인터넷)로 연결된 (온보드 CPU, 저장, 전원공급장치, 네트워크 인터페이스 등을 갖는) 완전한 컴퓨터에 의존한 일부 다른 분포된 컴퓨터 아키텍처의 실행에서 데이터를 수집하고 처리하기 위해 병렬로 동작할 수 있다.

[0033] 예를 들면, 일부 실시형태는 컴퓨터 판독 가능한 매체를 포함하는 컴퓨터를 제공한다. 실시형태는 프로세서에 결합되는 랜덤 액세스 메모리(RAM)를 포함한다. 프로세서는 메모리에 저장된 컴퓨터-실행 가능한 프로그램 명령을 실행한다. 이러한 프로세서는 마이크로프로세서, ASIC, 스테이트 기기(state machine), 또는 그 외의 프로세서를 포함할 수 있고, 다수의 컴퓨터 프로세서, 예를 들면, Intel Corporation(Santa Clara, California 소재) 및 Motorola Corporation (Schaumburg, Illinois 소재)의 프로세서의 어느 하나일 수 있다. 이러한 프로세서는 매체, 예를 들면, 컴퓨터 판독 가능한 매체를 포함하거나 함께 매체와 통신될 수 있고, 이는 프로세서에 의해서 수행되는 경우 프로세서가 본원에 기재된 단계를 실행하도록 하는 명령을 저장한다.

[0034] 컴퓨터 판독 가능한 매체의 실시형태는 컴퓨터 판독 가능한 명령을 갖는 프로세서를 제공할 수 있는 전자, 광학, 자기 또는 그 외의 저장 또는 전송 디바이스를 포함하지만 이들로 한정되지 않는다. 적합한 매체의 그 외의 예로는 플로피 디스크, CD-ROM, DVD, 자기디스크, 메모리칩, ROM, RAM, ASIC, 구성된 프로세서, 모든 광학 매체, 모든 자기 테이프 또는 그 외의 자기 매체, 또는 컴퓨터 프로세서가 명령을 판독할 수 있는 임의의 그 외의 매체를 포함하지만, 이들로 제한되지 않는다. 또한, 다양한 형태의 컴퓨터 판독 가능한 매체는 예를 들면, 라우터, 개인 또는 공용 네트워크 또는 그 외의 전송 디바이스 또는 채널 등의 컴퓨터로 무선 및 유선으로 명령을 전송하거나 전달할 수 있다. 명령은 예를 들면, C, C++, C#, Visual Basic, Java, Python, Perl, 및 JavaScript 등의 임의의 적합한 컴퓨터 프로그래밍 언어로부터의 코드를 포함할 수 있다.

[0035] 일부 실시형태에서 컴퓨터는 네트워크에 연결된다. 컴퓨터는 또한 다수의 외부 또는 내부 디바이스, 예를 들면, 마우스, CD-ROM, DVD, 키보드, 디스플레이 또는 그 외의 입력 또는 출력 디바이스를 포함할 수 있다. 컴퓨터의 예로는 개인용 컴퓨터, 디지털 보조장치, 개인용 정보 단말기, 셀룰러폰, 모바일폰, 스마트폰, 호출기, 디지털 테블릿, 랩톱 컴퓨터, 인터넷 가전 및 그 외의 프로세서 기반 디바이스이다. 일반적으로, 본원에 제공된 기술의 형태에 관련된 컴퓨터는 본원에 제공된 기술을 포함하는 하나 이상의 프로그램을 지지할 수 있는 Microsoft Windows, Linux, UNIX, Mac OS X 등의 임의의 운영 체제에서 작동하는 임의의 유형의 프로세서 기반 플랫폼일 수 있다. 일부 실시형태는 그 외의 어플리케이션 프로그램(예를 들면, 어플리케이션)을 실행하는 개인용 컴퓨터를 포함한다. 이러한 어플리케이션은 메모리 내에 포함될 수 있고, 예를 들면, 워드 프로세싱 어플리케이션, 스프레드시트 어플리케이션, 이메일 어플리케이션, 인스턴트 메신저 어플리케이션, 프리젠테이션 어플리케이션, 인터넷 브라우저 어플리케이션, 캘린더/오거나이저 어플리케이션 및 클라이언트 디바이스에 의해 실행 가능한 임의의 그 외의 어플리케이션을 포함할 수 있다.

[0036] 본 기술에 관련된 본원에 기재된 모든 이러한 성분, 컴퓨터, 및 시스템은 논리적 또는 가상적일 수 있다.

[0037] 따라서, 본원에는 피험체로부터 수득된 샘플에서 신생물을 스크리닝하는 방법에 관련된 기술이 제공되고, 이러한 방법은, 피험체(예를 들면, 위 조직)로부터 수득된 샘플(예를 들면, 혈장 샘플)에서 마커의 메틸화 상태를 분석하는 단계 및 마커의 메틸화 상태가 신생을 갖지 않는 피험체에서 분석된 마커의 메틸화 상태와 다른 경우 피험체가 신생물을 갖는 것을 확인하는 단계를 포함하고, 여기서 상기 마커는 표 1, 2, 4 및/또는 7에서 제공된 DMR 1-274로 이루어진 군으로부터 선택된 차별적으로 메틸화된 영역(DMR)에서 염기를 포함한다. 본 기술은 위 암을 확인하고 식별하는 것에 관련되고, 예를 들면, 일부 실시형태에서 상기 신생물을 위 신생물(예를 들면, 위 암)이다. 일부 실시형태는 복수의 마커를 분석하는 단계를 포함하는 방법을 제공하고, 예를 들면, 2 내지 11 내지 100 또는 200 또는 274 개의 마커를 분석하는 단계를 포함한다.

[0038] 본 기술은 분석된 메틸화 상태로 제한되지 않는다. 일부 실시형태에서 샘플에서 마커의 메틸화 상태를 분석하는 단계는 하나의 염기의 메틸화 상태를 결정하는 단계를 포함한다. 일부 실시형태에서, 샘플에서 마커의 메틸화 상태를 분석하는 단계는 복수의 염기에서 메틸화의 정도를 결정하는 단계를 포함한다. 또한 일부 실시형태에서 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 증가된 메틸화를 포함한다. 일부 실시형태에서, 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 감소한 메틸화를 포함한다. 일부 실시형태에서 마커의 메틸화 상태는 마커의 정상 메틸화 상태에 비해 마커의 상이한 패턴의 메틸화를 포함한다.

[0039] 또한, 일부 실시형태에서 마커는 100 이하의 염기의 영역이고, 마커는 500 이하의 염기의 영역이고, 마커는 1000 이하의 염기의 영역이고, 마커는 5000 이하의 염기의 영역이거나, 일부 실시형태에서 마커는 하나의 염기이다. 일부 실시형태에서 마커는 높은 CpG 밀도 프로모터 내에 있다.

- [0040] 본 기술은 샘플 유형에 의해 제한되지 않는다. 예를 들면, 일부 실시형태에서 샘플은 대변 샘플, 조직 샘플(예를 들면, 위 조직 샘플), 혈액 샘플(예를 들면, 혈장, 혈청, 전혈), 배설 또는 소변 샘플이다.
- [0041] 또한, 본 기술은 메틸화 상태를 결정하기 위해 사용된 방법으로 제한되지 않는다. 일부 실시형태에서 분석 단계는 메틸화 특이적 폴리머라제 연쇄 반응, 핵산 서열 결정, 질량 분석, 메틸화 특이적 뉴클레아제, 질량 기반 분리 또는 표적 포획을 사용하는 단계를 포함한다. 일부 실시형태에서, 분석은 메틸화 특이적 올리고뉴클레오티드의 사용을 포함한다. 일부 실시형태에서, 본 기술은 메틸화 상태, 예를 들면, 합성에 의한 서열 결정, 실시간(예를 들면, 단분자) 서열 결정, 비즈 애밀젼 서열 결정, 나노포어 서열 결정 등을 결정하기 위해 대규모 병렬 서열 결정(예를 들면, 차세대 서열 결정)을 사용한다.
- [0042] 본 기술은 DMR을 검출하기 위한 시약을 제공하고, 예를 들면, 일부 실시형태에서 SEQ ID NO: 1-109에 의해 제공된 서열을 포함하는 일련의 올리고뉴클레오티드가 제공된다. 일부 실시형태에서 DMR에서 염기를 갖는 염색체 영역에 상보적인 서열을 포함하는 올리고뉴클레오티드 예를 들면, DMR의 메틸화 상태에 민감한 올리고뉴클레오티드가 제공된다.
- [0043] 본 기술은 마커의 다양한 패널을 제공하고, 예를 들면, 일부 실시형태에서 마커는 ARHGEF4, ELMO1, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 또는 C13ORF18인 주석을 갖고, 상기 마커를 포함하는 염색체 영역을 포함한다 (표 4 참조). 또한, 실시형태는 DMR No. 49, 43, 196, 66, 1, 237, 249, 250, 251 또는 252인 표 4로부터 DMR을 분석하는 방법을 제공한다. 일부 실시형태는 마커의 메틸화 상태를 결정하는 단계를 제공하고, 염색체 영역은 ELMO1, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 또는 ZNF569인 주석을 갖고 상기 마커를 포함한다. (표 2 참조). 또한, 실시형태는 DMR No. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 또는 237인, 표 2로부터 DMR을 분석하는 방법을 제공한다. 일부 실시형태에서, 수득된 샘플은 혈장 샘플이고, 마커는 ELMO1, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1인 주석을 갖고 상기 마커를 포함하는 염색체 영역을 포함한다. 또한, 이러한 실시형태에서 샘플은 혈장 샘플이고 DMR No. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 또는 237인 표 1, 2, 4 및 7로부터의 DMR을 분석하는 방법이다. 일부 실시형태에서 상기 방법이 표 1, 2, 4 또는 7의 행에 제공된 한 쌍의 마커와 같은 2개의 마커의 메틸화 상태를 결정하는 단계를 포함한다.
- [0044] 키트 실시형태가 제공되고, 예를 들면, 키트는 비설파이트 시약; 및 (표 1, 2, 4 및 7로부터) DMR 1-274로 이루어진 군으로부터 선택된 DMR로부터 서열을 포함하고 암(예를 들면, 위암)에 걸리지 않은 피험체에 관련된 메틸화 상태를 갖는 대조군 핵산을 포함한다. 일부 실시형태에서, 키트는 비설파이트 시약 및 본원에 기재된 올리고뉴클레오티드를 포함한다. 일부 실시형태에서, 키트는 비설파이트 시약 및 (표 1, 2, 4 및 7로부터) DMR 1-274로 이루어진 군으로부터 선택된 DMR로부터의 서열을 포함하고 암(예를 들면, 위암)에 걸리지 않은 피험체에 관련된 메틸화 상태를 갖는 대조군 핵산을 포함한다. 일부 키트는 피험체로부터의 샘플(예를 들면, 대변 샘플, 위 조직 샘플, 혈장 샘플)을 수득하기 위한 샘플 수집기; 샘플로부터 핵산을 분리하기 위한 시약; 비설파이트 시약; 및 본원에 기재된 올리고뉴클레오티드를 포함한다.
- [0045] 본 기술은 조성물(예를 들면, 반응 혼합물)의 실시형태에 관한 것이다. 일부 실시형태에서 DMR을 포함하는 핵산 및 비설파이트 시약을 포함하는 조성물을 제공한다. 일부 실시형태는 DMR을 포함하는 핵산 및 본원에 기재된 올리고뉴클레오티드를 포함하는 조성물이 제공된다. 일부 실시형태는 DMR을 포함하는 핵산 및 메틸화 민감성 제한 효소를 포함하는 조성물이 제공된다. 일부 실시형태는 DMR을 포함하는 핵산 및 폴리머라제를 포함하는 조성물이 제공된다.
- [0046] 추가의 관련된 방법 실시형태는 피험체로부터 수득된 샘플(예를 들면, 위 조직 샘플, 혈장 샘플, 대변 샘플)에서 신생물을 스크리닝하기 위한 것이 제공되고, 예를 들면, (표 1, 2, 4 또는 7로부터) 하나 이상의 DMR 1-274인 DMR 내에 염기를 포함하는 샘플에서 마커의 메틸화 상태를 결정하고; 피험체 샘플로부터의 마커의 메틸화 상태를, 암에 걸리지 않은 피험체로부터의 정상 대조군 샘플로부터 마커의 메틸화 상태와 비교하고; 상기 피험체 샘플 및 정상 대조군 샘플의 메틸화 상태의 차이의 신뢰구간 및/또는 p 값을 결정하는 것을 포함한다. 일부 실시형태에서, 상기 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 실시형태는 비설파이트 반응된 핵산

을 생성하기 위해 DMR을 포함하는 핵산을 비설파이트 시약과 반응시키는 단계; 상기 비설파이트 반응된 핵산의 뉴클레오티드 서열을 제공하기 위해 상기 비설파이트 반응된 핵산을 서열 결정하는 단계; 상기 2개 서열의 차이를 확인하기 위해 상기 비설파이트 반응된 핵산의 뉴클레오티드 서열을, 위암에 걸리지 않은 피험체로부터 DMR을 포함하는 핵산의 뉴클레오티드 서열과 비교하는 단계; 및 차이가 존재하는 경우 상기 피험체가 신생물을 갖는 것으로 확인하는 단계를 포함한다.

[0047] 피험체로부터 수득된 샘플에서 신생물을 스크리닝하기 위한 시스템은 본 기술에 의해서 제공된다. 시스템의 예시의 실시형태는 예를 들면, 피험체로부터 수득된 샘플(예를 들면, 위 조직 샘플, 혈장 샘플, 대변 샘플)에서 신생물을 스크리닝하기 위한 시스템을 포함하고, 상기 시스템은 샘플의 메틸화 상태를 결정하도록 구성된 분석성분, 상기 샘플의 메틸화 상태를, 대조군 샘플 또는 데이터베이스에 기록된 레퍼런스 샘플 메틸화 상태와 비교하도록 구성된 소프트웨어 성분, 및 암-관련된 메틸화 상태를 사용자에게 경고하도록 구성된 경고 성분을 포함한다. 일부 실시형태에서 경고는 다수의 분석에 의한 결과를 수용하는 소프트웨어 성분에 의해 결정되고(예를 들면, 다수의 마커, 예를 들면, 예를 들면, 표 1, 2, 또는 7에서 제공된 DMR의 메틸화 상태를 결정하고), 값 또는 결과를 계산하여 다수 결과에 기초해서 보고한다. 일부 실시형태는 값 또는 결과 계산에 사용하기 위해 본원에 제공된 각각의 DMR과 관련된 각종 파라미터의 데이터베이스 및/또는 사용자(예를 들면, 의사, 간호사, 임상의, 등)에게 보고하기 위한 경고를 제공한다. 일부 실시형태에서 다수 분석에 의한 모든 결과가 보고되고 일부 실시형태에서 하나 이상의 결과가 사용되어 피험체에서 암 위험을 나타내는 다수 분석에 의한 하나 이상의 결과의 복합(composite)에 기초한 스코어, 값 또는 결과를 제공한다.

[0048] 시스템의 일부 실시형태에서, 샘플은 DMR을 포함하는 핵산을 포함한다. 일부 실시형태에서, 시스템은 핵산을 분리하기 위한 성분, 샘플을 수집하기 위한 성분, 예를 들면, 대변 샘플을 수집하기 위한 성분을 더 포함한다. 일부 실시형태에서, 샘플은 DMR을 포함하는 핵산 서열을 포함한다. 일부 실시형태에서 암에 걸리지 않는 피험체로부터의 핵산 서열을 포함한다. 또한, 예를 들면, 일련의 핵산으로서, 각각의 핵산이 DMR을 포함하는 서열을 갖는 핵산이 제공된다. 일부 실시형태에서, 일련의 핵산에서 각각의 핵산은 암에 걸리지 않은 피험체로부터의 서열을 갖는다. 관련 시스템 실시형태는 기재된 일련의 핵산 및 일련의 핵산과 관련된 핵산 서열의 데이터베이스를 포함한다. 일부 실시형태는 비설파이트 시약을 더 포함한다. 일부 실시형태는 핵산 시퀀서를 더 포함한다.

[0049] 일부 실시형태에서, 인간 환자로부터 샘플(예를 들면, 위 조직 샘플, 혈장 샘플, 대변 샘플)을 특성화 하는 방법이 제공된다. 예를 들면, 일부 실시형태에서 이러한 실시형태는 인간 환자의 샘플로부터 DNA를 수득하는 단계; 표 1, 2, 4 또는 7로부터 DMR 1-274으로 이루어진 군으로부터 선택된 차별적으로 메틸화된 영역(DMR) 내에 염기를 포함하는 DNA 메틸화 마커의 메틸화 상태를 분석하는 단계; 및 하나 이상의 DNA 메틸화 마커의 분석화된 메틸화 상태를, 위 신생물을 갖지 않는 인간 환자의 하나 이상의 DNA 메틸화 마커의 메틸화 수준 레퍼런스와 비교하는 단계를 포함한다.

[0050] 이러한 방법은 인간 환자로부터의 샘플의 특이적 형태로 제한되지 않는다. 일부 실시형태에서, 샘플은 위 조직 샘플이다. 일부 실시형태에서, 샘플은 혈장 샘플이다. 일부 실시형태에서 샘플은 대변 샘플, 조직 샘플, 위 조직 샘플, 혈액 샘플, 또는 소변 샘플이다.

[0051] 일부 실시형태에서, 이러한 방법은 복수의 DNA 메틸화 마커를 분석하는 단계를 포함한다. 일부 실시형태에서, 이러한 방법은 2 내지 11 개의 DNA 메틸화 마커를 분석하는 단계를 포함한다. 일부 실시형태에서 이러한 방법은 12 내지 107 개의 DNA 메틸화 마커를 포함한다. 일부 실시형태에서, 이러한 방법은 하나의 염기의 메틸화 상태를 결정하는 단계를 포함하는 샘플에서 하나 이상의 DNA 메틸화 마커의 메틸화 상태를 분석하는 단계를 포함한다. 일부 실시형태에서, 이러한 방법은 복수의 염기의 메틸화 정도를 결정하는 단계를 포함하는 샘플에서 하나 이상의 DNA 메틸화 마커의 메틸화 상태를 분석하는 단계를 포함한다. 일부 실시형태에서 이러한 방법은 정방향 가닥의 메틸화 상태를 분석하거나 역방향 가닥의 메틸화 상태를 분석하는 단계를 포함한다.

[0052] 일부 실시형태에서, DNA 메틸화 마커는 100 개 이하의 염기의 영역이다. 일부 실시형태에서, DNA 메틸화 마커는 500 개 이하의 염기의 영역이다. 일부 실시형태에서, DNA 메틸화 마커는 1000 개 이하의 염기의 영역이다. 일부 실시형태에서, DNA 메틸화 마커는 5000 개 이하의 염기의 영역이다. 일부 실시형태에서, DNA 메틸화 마커는 1 개의 염기이다. 일부 실시형태에서, 상기 DNA 메틸화 마커는 높은 CpG 밀도 프로모터 내에 있다.

[0053] 일부 실시형태에서, 분석 단계는 메틸화 특이적 폴리미라제 연쇄 반응, 핵산 서열 결정, 질량 분석, 메틸화 특이적 뉴클레아제, 질량 기반 분리 또는 표적 포획을 사용하는 단계를 포함한다.

- [0054] 일부 실시형태에서, 상기 분석 단계는 메틸화 특이적 올리고뉴클레오티드의 사용을 포함한다. 일부 실시형태에서, 메틸화 특이적 올리고뉴클레오티드는 SEQ ID NO: 1-109로 이루어진 군으로부터 선택된다.
- [0055] 일부 실시형태에서, ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 및 C13ORF18로 이루어진 군으로부터 선택된 주석을 갖는 염색체 영역은 상기 DNA 메틸화 마커를 포함한다. 일부 실시형태에서, 상기 DMR은 표 4로부터의 것이고 DMR No. 49, 43, 196, 66, 1, 237, 249, 250, 251 및 252로 이루어진 군으로부터 선택된다.
- [0056] 일부 실시형태에서, ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 및 ZNF569로 이루어진 군으로부터 선택된 주석을 갖는 상기 염색체 영역은 상기 DNA 메틸화 마커를 포함한다.
- [0057] 일부 실시형태에서, 상기 DMR은 표 2로부터의 것이고 DMR No. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 및 237로 이루어진 군으로부터 선택된다.
- [0058] 일부 실시형태에서, 상기 샘플은 혈장 샘플이고, 상기 마커는 ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP2R5C, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1인 주석을 갖고, 상기 마커를 포함하는 염색체 영역을 포함한다.
- [0059] 일부 실시형태에서, DMR은 표 1, 2, 4, 및 7로부터의 것이고 DMR No. 253, 237, 252, 261, 251, 196, 250, 265, 256, 249, 및 274로 이루어진 군으로부터 선택된다.
- [0060] 일부 실시형태에서, 이러한 방법은 2개의 DNA 메틸화 마커의 메틸화 상태를 결정하는 단계를 포함한다. 일부 실시형태에서, 이러한 방법은 표 1, 2, 4 및 7의 행에 제공된 한 쌍의 DNA 메틸화 마커의 메틸화 상태를 결정하는 단계를 포함한다.
- [0061] 특정 실시형태에서, 본 기술은 인간 환자로부터 수득된 샘플을 특성화하는 방법을 제공한다. 일부 실시형태에서, 이러한 방법은 표 1, 2, 4 및 7로부터 DMR 1-274로 이루어진 군으로부터 선택된 DMR 내에 염기를 포함하는 샘플에서 DNA 메틸화 마커의 메틸화 상태를 결정하는 단계; 환자 샘플로부터 DNA 메틸화 마커의 메틸화 상태를 위암에 걸리지 않는 인간 환자로부터의 정상 대조군 샘플로부터 DNA 메틸화 마커의 메틸화 상태를 비교하는 단계; 및 상기 인간 환자 및 정상 대조군 샘플의 메틸화 상태 차이의 신뢰구간 및/또는 p 값을 결정하는 단계;를 포함한다. 일부 실시형태에서, 신뢰 구간은 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.
- [0062] 특정 실시형태에서, 본 기술은 인간 피험체로부터 수득된 샘플(예를 들면, 위 조직 샘플, 혈장 샘플, 대변 샘플)을 특성화하는 방법이 제공되고, 상기 방법은, 비설파이트 반응된 핵산을 생성하기 위해 DMR을 포함하는 핵산을 비설파이트 시약과 반응시키는 단계; 상기 비설파이트 반응된 핵산의 뉴클레오티드 서열을 제공하기 위해 상기 비설파이트 반응된 핵산을 서열 결정하는 단계; 상기 2개 서열의 차이를 확인하기 위해 상기 비설파이트 반응된 핵산의 뉴클레오티드 서열을, 위암에 걸리지 않은 피험체로부터 DMR을 포함하는 핵산의 뉴클레오티드 서열과 비교하는 단계를 포함한다.
- [0063] 특정 실시형태에서, 본 기술은 인간 피험체로부터 수득된 샘플(예를 들면, 위 조직 샘플, 혈장 샘플, 대변 샘플)을 특성화하는 시스템이 제공되고, 상기 시스템은 샘플의 메틸화 상태를 결정하도록 구성된 분석 성분, 상기 샘플의 메틸화 상태를, 대조군 샘플 또는 데이터베이스에 기록된 래퍼런스 샘플 메틸화 상태와 비교하도록 구성된 소프트웨어 성분, 및 메틸화 상태의 조합에 기초한 단일 값을 결정하고 위암-관련된 메틸화 상태를 사용자에게 경고하도록 구성된 경고 성분을 포함한다. 일부 실시형태에서, 상기 샘플은 DMR을 포함하는 핵산을 포함한다.
- [0064] 일부 실시형태에서, 이러한 시스템은 핵산을 분리하기 위한 성분을 더 포함한다. 일부 실시형태에서, 이러한 시스템은 샘플을 수집하기 위한 성분을 더 포함한다.
- [0065] 일부 실시형태에서, 상기 샘플은 대변 샘플, 조직 샘플, 위 조직 샘플, 혈액 샘플, 또는 소변 샘플이다.
- [0066] 일부 실시형태에서, 상기 데이터베이스는 DMR을 포함하는 핵산 서열을 포함한다. 일부 실시형태에서, 상기 데이터베이스는 위암에 걸리지 않은 피험체로부터 핵산 서열을 포함한다.

- [0067] 추가의 실시형태는 본원에 함유된 교시를 기초한 관련 기술에서 당업자에게 명백하게 될 것이다.
- ### 도면의 간단한 설명
- [0068] 본 기술의 이들 및 그 외의 특징, 형태, 및 이점은 다음의 도면을 참조하여 잘 이해될 것이다.
- 도 1에는 ELM01이 실시예 2에 기재된 위암에 대해 높은 식별성을 나타내는 것을 도시한다.
- 도 2에는 QuARTs (정량적 대립유전자 특이적 실시간 표적 및 신호 증폭) 분석에 의해서 메틸화 DNA 서명의 검출에서 사용된 FRET 카세트의 올리고뉴클레오티드 서열을 제공한다. 각각의 FRET 서열은 3개의 별도의 분석에 함께 다중화할 수 있는 형광체(fluorophore) 및 쿠엔처(quencher)를 포함한다.
- 도 3은 개개의 마커 곡선과 비교해서 3개의 마커 패널(ELM01, ZNF569, 및 c13orf18)의 혈장 중의 성능을 강조하는 수용자 동작 특성 곡선을 제공한다. 100% 특이성에서, 패널은 86%의 감도로 위암을 검출했다.
- 도 4a는 정상 점막으로부터 단계 4 위암으로의 진행을 나타내는 단계별 혈장 중의 메틸화된 ELM01의 로그 스케일 절대 가닥 카운트를 제공한다. 정량적인 마커 수준은 GC 단계에서 바로 증가했다.
- 도 4b는 3 마커 패널(ELM01, ZNF569, 및 c13orf18)의 단계별로(100% 특이성에서) 혈장 중의 위암 감도를 입증하는 막대 차트를 제공한다.
- 도 5a, 5b 및 5c는 90% 특이성 (A), 95% 특이성 (B), 및 100% 특이성 (C)에서 매트릭스 포맷에서 74 개의 혈장 샘플 중의 12 개의 위암 마커 (ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; ELM01은 2회 실시하고, 두 번째는 바이플렉스 포맷임)의 성능을 제공한다. 마커는 수직방향으로 도시되고 샘플은 수평방향으로 도시된다. 샘플은 좌측에 정상 샘플이 배열되고 우측에 암 샘플이 배열된다. 양성 히트(positive hit)는 밝은 회색으로 나타내고 짙은 회색 부분에는 존재하지 않는다. 여기서, 상부 성능 마커 - ELM01 -은 먼저 90% 특이성으로 도시되고 나머지는 패널은 100% 특이성으로 도시된다. 이러한 플롯에서 마커가 상보적 방법으로 평가될 수 있다.
- 도 6은 100% 특이성에서 74개 혈장 샘플 중의 3개의 위암 마커(최종 패널)의 성능을 제공한다. 마커는 수직방향으로 도시되고 샘플은 수평방향으로 도시된다. 샘플은 좌측에 정상 샘플이 배열되고 우측에 암 샘플이 배열된다. 암은 단계별로 배열된다. 양성 히트는 밝은 회색으로 표시되고 짙은 회색 부분에는 존재하지 않는다 (ELM01, ZNF569, 및 c13orf18).
- 도 7a~m은 혈장 중의 12개 위암 마커(ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; ELM01은 2회 실시되고 두 번째는 바이플렉스 포맷임)의 박스 플롯(선형 스케일)을 제공한다. 대조군 샘플(N=38)은 좌측에 도시되고 위암 증례(N=36)는 우측에 도시된다. 비히를 측은 β-액틴 가닥에 대해 정규화된 % 메틸화이다.
- 도면은 반드시 일정한 척도로 도시되는 것이 아니고 도면 중의 대상물이 반드시 서로 관련해서 일정한 척도로 도시되어 있는 것은 아닌 것이 이해되어야 한다. 도면은 본 명세서에 개시된 장치, 시스템, 조성물 및 방법의 다양한 실시형태를 명확하게 이해하기 위한 것이다. 가능한 한 동일한 참조번호는 도면 전체에서 동일하거나 유사한 부분을 가리키기 위해 사용된다. 또한, 도면은 본 교시의 범위를 결코 한정하는 것은 아닌 것이 이해되어야 한다.
- ### 발명을 실시하기 위한 구체적인 내용
- [0069] 본원에는 종양을 검출하는 것에 관련된 기술, 특히, 배타적이지 않지만, 전악성 신생물 및 악성 신생물, 예를 들면, 위암을 검출하기 위한 방법, 조성물 및 관련 용도에 관련된 기술이 제공되어 있다. 이러한 본 기술은 본원에 기재되어 있기 때문에, 사용된 섹션 제목은 단지 조직적인 목적인 것으로 어떤 식으로든 주제를 제한하는 것으로 해석되어서는 안된다.
- [0070] 다양한 실시형태의 상세한 설명에서, 설명목적으로, 다수의 특정 세부사항은 개시된 실시형태의 완전한 이해를 제공하기 위해서 기재되어 있다. 그러나 당업자에게는, 다양한 실시형태가 이러한 특정한 세부사항에 관계없이 실시될 수 있는 것이 명확할 것이다. 그 외의 예에서, 구조 및 장치는 블록도 형태로 표시된다. 또한, 당업자는 방법이 제시되고 수행되는 특정 배열이 예시적인 것이고 이러한 배열이 변화될 수 있고 본원에 개시된 다양한 실시형태의 사상 및 범위 내에 있다고 생각할 수 있음을 쉽게 이해할 수 있다.
- [0071] 특히, 특히 출원, 기사, 서적, 논문 및 인터넷 웹페이지를 포함하지만 이들로 제한되지 않는 이 출원에 인용된

모든 문현 및 유사한 물질은, 임의의 목적을 위해 전체 참조로 포함되어 있다. 달리 정의되어 있지 않는 한, 사용된 모든 기술 및 과학적 용어는 본원에 기재된 다양한 실시형태가 속하는 당업자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 포함된 참고 문현 내의 용어의 정의는 본 교시에서 제공된 정의와 상이한 것으로 나타나면, 본 교시에 제공된 정의가 지배적이다.

[0072] 정의

[0073] 본 교시의 이해를 용이하게 하기 위해, 많은 용어 및 어구가 하기에 정의된다. 추가의 정의는 상세한 설명에 기재되어 있다.

[0074] 명세서 및 청구범위에서, 다음의 용어는 달리 명확하게 나타내지 않는다면 본원에서 명시적으로 관련되는 의미를 취한다. 본원에서 사용되는 "일 실시형태"라는 어구는 반드시 동일한 실시형태를 의미하는 것은 아니지만, 동일한 실시형태일 수도 있다. 또한, 본원에서 사용되는 "또 다른 실시형태"라는 어구는 반드시 상이한 실시형태를 의미하는 것은 아니지만, 상이한 실시형태일 수도 있다. 따라서, 하기 기재된 바와 같이, 본 발명의 다양한 실시형태는 본 발명의 사상 또는 범위를 벗어나지 않는 한 용이하게 결합될 수 있다.

[0075] 또한, 본원에서 사용된, "또는"이라는 용어는 포괄적인 "또는" 연산자이고 달리 명확하게 나타내지 않는 한 "및 /또는"의 용어와 동일하다. "에 기초한"이라는 용어는 배타적이지 않고 달리 명확하게 나타내지 않는 한 기재되지 않은 추가의 인자에 기초될 수 있다. 또한, 본 명세서 내에, "관사(a, an, 및 the)"의 의미는 복수의 것을 포함한다. "내(in)"의 의미는 "내(in)" 및 "상(on)"을 포함한다.

[0076] 본원에 사용된 바와 같이, "핵산" 또는 "핵산 분자"는 일반적으로 임의의 리보핵산 또는 데옥시리보핵산을 의미하고, 이는 개질되지 않거나 개질된 DNA 또는 RNA일 수 있다. "핵산"은, 제한없이 단일 및 이중 가닥 핵산을 포함한다. 본원에 사용된 바와 같이, "핵산"이라는 용어는 또한 하나 이상의 개질된 염기를 포함하는 상기 기재된 DNA를 포함한다. 따라서, 안정성 또는 그 외의 이유로 개질된 주쇄를 갖는 DNA가 "핵산"이다. 본원에 사용되는 "핵산"이라는 용어는 핵산의 화학적, 효소적 또는 대사적으로 개질된 형태 뿐만아니라 예를 들면, 단일 및 복합 세포를 포함하는 바이러스 및 세포가 특징적인 DNA의 화학적 형태를 포함한다.

[0077] "올리고뉴클레오티드" 또는 "폴리뉴클레오티드" 또는 "뉴클레오티드" 또는 "핵산"은 2개 이상의 데옥시리보뉴클레오티드 또는 리보뉴클레오티드를 갖는 분자를 의미하고, 바람직하게는 3개 초과, 일반적으로 10초과이다. 정확한 크기는 주요 인자에 의존하고, 올리고뉴클레오티드의 최종적인 기능 또는 용도에 의존한다. 올리고뉴클레오티드는 임의의 방식으로 생성될 수 있는데, 예를 들면, 화학적 합성, DNA 복제, 역전사 또는 이들의 조합을 들 수 있다. DNA의 일반적인 데옥시리보뉴클레오티드는 티민, 아데닌, 시토신, 및 구아닌이다. RNA의 일반적인 리보뉴클레오티드는 우라실, 아데닌, 시토신 및 구아닌이다.

[0078] 본원에 사용되는 핵산의 "유전자좌" 또는 "영역"은 핵산의 서브영역, 예를 들면, 염색체 상의 유전자, 단일 뉴클레오티드, CpG 섬을 의미한다.

[0079] "상보적인" 및 "상보성"이라는 용어는 염기쌍화 규칙에 관련되는 뉴클레오티드(예를 들면, 1개의 뉴클레오티드) 또는 폴리뉴클레오티드(예를 들면, 뉴클레오티드의 서열)을 의미한다. 예를 들면, 서열 5'-A-G-T-3'은 서열 3'-T-C-A-5'에 대해 상보적이다. 상보성은 "부분적"일 수 있고, 핵산 염기의 일부만 염기쌍화 규칙에 따라 매칭된다. 또는, 핵산 사이의 "완전" 또는 "전체" 상보성이 있을 수 있다. 핵산 가닥 사이의 상보성 정도는 핵산 가닥 사이의 하이브리드화의 효율 및 강도에 영향을 미친다. 이는 핵산 사이의 결합에 의존하는 검출 방법 및 증폭 반응에서 특히 중요하다.

[0080] "유전자"라는 용어는 RNA 또는 폴리펩티드 또는 전구체의 생성을 위해 필요한 코딩 서열을 포함하는 핵산(예를 들면, DNA 또는 RNA) 서열을 의미한다. 기능성 폴리펩티드는 폴리펩티드의 목적의 활성 또는 기능성 특성(예를 들면, 효소적 활성, 리간드 결합, 신호 전달 등)이 유지되는 것인 한 코딩 서열의 임의의 부분 또는 전체 길이 코딩 서열에 의해서 코딩될 수 있다. 유전에 대해 사용되는 경우 "부분"이라는 용어는 유전자의 단편을 의미한다. 단편은 몇 개 뉴클레오티드로부터, 1개의 뉴클레오티드를 뺀 전체 유전자 서열까지의 크기 범위일 수 있다. 따라서, "유전자의 적어도 일부를 포함하는 뉴클레오티드"는 유전자의 단편 또는 전체 유전자를 포함할 수 있다.

[0081] "유전자"라는 용어는 또한 구조적 유전자의 코딩 영역을 포함하고, 5' 및 3' 말단에 대한 코딩 영역에 인접해서 위치된 서열, 예를 들면, 어느 말단에 대한 약 1kb의 거리에 대한 것을 포함하고, 유전자는 전체 길이 mRNA의 길이에 대응한다(예를 들면, 코딩, 조절, 구조 및 그 외의 서열을 포함한다). 코딩 영역의 5'에 위치하고 mRNA에 존재하는 서열은 5' 비번역 또는 번역되지 않은 서열로서 지칭된다. 코딩 영역의 3' 또는 하류에 위치되고

mRNA 상에 존재하는 서열은 3' 비 번역 또는 3'-번역되지 않는 서열로 지칭된다. "유전자"라는 용어는 cDNA와 유전자의 계놈 형태를 포함한다. 일부 생물체(예를 들면, 진핵생물)에서, 유전자의 계놈 형태 또는 클론은 "인트론" 또는 "개재 영역" 또는 "개재 서열"이라고 하는 비-코딩서열로 중단된 코딩 영역을 포함한다. 인트론은 핵 RNA(hnRNA)로 전사된 유전자의 세그먼트이고; 인트론은 인핸서와 같은 조절 요소를 포함해도 좋다. 인트론은 핵 또는 1차 전사물로부터 제거되거나 "연결되고(spliced out)"; 따라서 인트론은 메신저 RNA(mRNA) 전사물에 존재하지 않는다. mRNA는 신생물 폴리펩티드에서 아미노산의 서열 또는 순서를 특정하기 위해서 번역 중에 기능한다.

[0082]

인트론을 함유하는 것 이외에, 유전자의 계놈 형태는 또한 RNA 전사물에 존재하는 서열의 5' 및 3' 말단에 위치한 서열을 포함할 수 있다. 이를 서열은 "인접하는" 서열 또는 영역으로 지칭된다(이들 인접하는 서열은 mRNA 전사물 상에 존재하는 비번역 서열의 5' 또는 3'에 위치한다). 5' 인접 영역은 유전자의 전사물을 제어하거나 영향을 미치는 프로모터 및 인핸서와 같은 조절 서열을 포함할 수 있다. 3' 인접 영역은 전사의 종결, 전사 후 절단, 및 폴리아데닐화에 관한 서열을 포함할 수 있다.

[0083]

"야생형"이라는 용어는 유전자와 관련되는 경우 천연 발생원으로부터 분리된 유전자의 특성을 갖는 유전자를 의미한다. "야생형"이라는 용어는 유전자 생성물과 관련되는 경우 천연 발생원으로부터 분리된 유전자 생성물의 특성을 갖는 유전자 생성물을 의미한다. "천연 발생"이라는 용어는 대상물에 적용하는 경우 천연에서 대상물이 발견될 수 있는 사실을 의미한다. 예를 들면, 천연에서 발생원으로부터 분리될 수 있고 실험실에서 사람 손에 의해서 의도적으로 개질되지 않은 생물체(바이러스 포함)에 존재하는 폴리펩티드 또는 폴리뉴클레오티드 서열이 천연 발생이다. 야생형 유전자는 종종 집단에서 가장 빈번하게 관찰되고 유전자의 "정상" 또는 "야생형" 형태를 임의의 지정된 유전자 또는 대립유전자이다. 반면, "개질된" 또는 "돌연변이체"라는 용어는 유전자 또는 유전자 생성물과 관련되는 경우 각각 야생형 유전자 또는 유전자 생성물에 비해 서열의 개질 및/또는 기능성 특성의 개질(예를 들면, 개질된 특성)을 나타내는 유전자 또는 유전자 생성물을 의미한다. 천연 발생 돌연변이체가 분리될 수 있고; 이들이 야생형 유전자 또는 유전자 생성물에 비해 특성이 변경되는 사실에 의해서 확인되는 것을 유의해야 한다.

[0084]

"대립유전자"라는 용어는 유전자의 변형을 의미하고; 변형은 변이체 및 돌연변이체, 다형성 유전자좌 및 단일 뉴클레오티드 다형성 유전자좌, 프레임 시프트 및 스플라이스 돌연변이를 포함하지만 이들로 제한되지 않는다. 대립유전자는 집단에서 천연 발생할 수 있거나 집단의 임의의 특정 개체의 수명 중에 발생할 수 있다.

[0085]

"변이체" 및 "돌연변이체"라는 용어는 뉴클레오티드 서열과 관련되어 사용되는 경우 하나 이상의 뉴클레오티드가 상이한 핵산 서열을 의미하고, 일반적으로 관련된 뉴클레오티드 산 서열이다. "변이"는 2개의 상이한 뉴클레오티드 서열의 차이이고; 일반적으로 하나의 서열이 기준 서열이다

[0086]

"증폭"은 템플레이트 특이성을 포함한 핵산 복제의 특정한 경우이다. 비-특이적 템플레이트 복제(예를 들면, 복제는 템플레이트 의존된 것이지만 특이적 템플레이트에 의존하는 것은 아니다)와 대조한다. 템플레이트 특이성은 본원에서 복제(예를 들면, 적절한 폴리뉴클레오티드 서열의 합성) 및 뉴클레오티드(리보 또는 테옥시리보) 특이성의 충실성(fidelity)과 구별된다. 템플레이트 특이성은 빈번하게 "표적"특이성의 접에서 기재된다. 표적 서열은, 그 외의 핵산으로부터 선별되도록 하는 의미에서 "표적"이다. 증폭 기술은 주로 이러한 선별을 위해서 설계되었다.

[0087]

핵산의 증폭은 일반적으로 증폭 생성물 또는 임플리콘(amplicon)이 일반적으로 검출 가능한 일반적으로 소량의 폴리뉴클레오티드(예를 들면, 단일 폴리뉴클레오티드 분자, 10 내지 100 개 카피의 폴리뉴클레오티드 분자이고, 이는 정확하게 이와 동일하거나 동일하지 않을 수 있다)로부터 다수개 카피의 폴리뉴클레오티드 또는 폴리뉴클레오티드의 일부의 생성을 의미한다. 폴리뉴클레오티드 증폭은 다양한 화학적 및 효소적 공정을 포함한다. 폴리머라제 연쇄 반응(PCR) 또는 리가제 연쇄 반응(LCR; 미국 특허 제5,494,810 참조; 본원에 참조로 포함되어 있다) 중에 하나 또는 몇 개 카피의 표적 또는 템플레이트 DNA 분자로부터 생성된 다수 DNA 카피는 증폭된 형태이다. 증폭의 추가의 형태는 대립유전자-특이적 PCR(예를 들면, 미국특허 제5,639,611 참조, 본원에 참조로 포함되어 있다), 어셈블리 PCR(예를 들면, 미국 특허 제5,965,408호 참조; 본원에 참조로 포함되어 있다), 헬리카제-의존 PCR(예를 들면, 미국특허 제7,662,594; 본원에 참조로 포함되어 있다), Hot-start PCR (예를 들면, 미국 특허 제5,773,258 및 5,338,671 참조; 본원에 참조로 포함되어 있다), 서열간-특이적 PCR, 인버스 PCR (예를 들면, Triglia, et al. (1988) Nucleic Acids Res., 16:8186 참조; 본원에 참조로 포함되어 있다), 결찰 매개(ligation-mediated) PCR (예를 들면, Guilfoyle, R. et al. (1997) Nucleic Acids Research, 25:1854-1858 참조; 미국 특허 제5,508,169; 본원에 참조로 포함되어 있다), 메틸화 특이적 PCR (예를 들면, Herman,

et al., (1996) PNAS 93(13) 9821-9826 참조; 본원에 참조로 포함되어 있다), 미니프라이머(miniprimer) PCR, 다중 결찰 의존 탐침 증폭(multiplex ligation-dependent probe amplification) (예를 들면, Schouten, et al., (2002) Nucleic Acids Research 30(12): e57 참조; 본원에 참조로 포함되어 있다), 다중 PCR (예를 들면, Chamberlain, et al., (1988) Nucleic Acids Research 16(23) 11141-11156; Ballabio, et al., (1990) Human Genetics 84(6) 571-573; Hayden, et al., (2008) BMC Genetics 9:80 참조; 본원에 참조로 포함되어 있다), 중첩된(nested) PCR, 오버랩 신장(overlap-extension) PCR (예를 들면, Higuchi, et al., (1988) Nucleic Acids Research 16(15) 7351-7367 참조; 본원에 참조로 포함되어 있다), 실시간 PCR (예를 들면, Higuchi, et al., (1992) Biotechnology 10:413-417; Higuchi, et al., (1993) Biotechnology 11:1026-1030 참조; 본원에 참조로 포함되어 있다), 역전사 PCR (예를 들면, Bustin, S.A. (2000) J. Molecular Endocrinology 25:169-193 참조; 본원에 참조로 포함되어 있다), 고상 PCR, 열 비대칭 인터레이싱된 (thermal asymmetric interlaced) PCR, 및 터치다운(Touchdown) PCR (예를 들면, Don, et al., Nucleic Acids Research (1991) 19(14) 4008; Roux, K. (1994) Biotechniques 16(5) 812-814; Hecker, et al., (1996) Biotechniques 20(3) 478-485 참조; 본원에 참조로 포함되어 있다), 폴리뉴클레오티드 증폭은 또한 디지털 PCR을 사용하여 수행될 수 있다 (예를 들면, Kalinina, et al., Nucleic Acids Research. 25: 1999-2004, (1997); Vogelstein and Kinzler, Proc Natl Acad Sci USA. 96; 9236-41, (1999); 국제 특허 공개 제 WO05023091A2; 미국 특허 출원 공개 제 20070202525호 참조; 본원에 참조로 포함되어 있다),

[0088]

"폴리머라제 연쇄반응"("PCR")은 K.B. Mullis 미국 특허 제4,683,195, 4,683,202, 및 4,965,188호의 방법에 관한 것으로서, 게놈 DNA의 혼합물을 복제 또는 정제 없이 표적 서열의 세그먼트의 농도를 증가시키는 방법을 기재한다. 이러한 표적 서열의 증폭 공정은 과량의 2개의 올리고뉴클레오티드 프라이머를, 목적의 표적 서열을 포함하는 DNA 혼합물에 도입하는 것 이어서, DNA 폴리머라제의 존재하에서 정확한 서열의 열 사이클링을 수행한다. 2개의 프라이머는 이중가닥 표적 서열의 각각의 가닥에 대해 상보적이다. 증폭을 행하기 위해서, 혼합물을 변성시키고, 이어서 프라이머를 표적 분자 내에서 이들의 상보적 서열에 어닐링한다. 어닐링 후, 상보적 가닥의 새로운 쌍을 형성하도록 프라이머를 폴리머라제로 신장한다. 변성, 프라이머 어닐링 및 폴리머라제 신장의 단계를 수회 반복하여(즉, 변성, 어닐링, 및 신장이 1개의 "사이클"을 구성하고; 수개의 "사이클"이 있을 수 있다) 목적의 표적 서열의 높은 농도의 증폭 단편을 수득할 수 있다. 목적의 표적 서열의 증폭된 세그먼트의 길이는 서로 프라이머의 상대적 위치에 의해서 결정되고, 따라서 이러한 길이는 제어가능한 변수이다. 공정의 반복 형태에 의해, "폴리머라제 연쇄 반응"("PCR")으로 지칭된다. 표적 서열의 목적의 증폭 세그먼트가 혼합물에서 (농도의 점에서) 현저한 서열이기 때문에, 이들은 "PCR 증폭된"으로 말하고 "PCR 생성물" 또는 "증폭"이다.

[0089]

템플레이트 특이성은 효소를 선택해서 대부분 증폭 기술에서 달성된다. 증폭 효소는 이들이 사용되는 조건에서 핵산의 이종 혼합물에서 핵산의 특이적 서열만을 처리하는 효소이다. 예를 들면, Q-베타 레플리카제(replicase)의 경우, MDV-1 RNA는 레플리카제를 위한 특이적 템플레이트이다 (Kacian et al., Proc. Natl. Acad. Sci. USA, 69:3038 [1972]). 그 외의 핵산은 이러한 증폭 효소에 의해 복제되지 않을 것이다. 마찬가지로, T7 RNA 폴리머라제의 경우, 이 증폭 효소는 자체의 프로모터의 염격한 특이성을 갖는다 (Chamberlin et al., Nature, 228:227 [1970]). T4 DNA 리가제(ligase)의 경우, 효소는 2개의 올리고뉴클레오티드 또는 폴리뉴클레오티드를 결찰하지 않고, 여기서 결찰 접합에서 올리고뉴클레오티드 또는 폴리뉴클레오티드 기관과 템플레이트 사이의 미스매칭이 있다(Wu and Wallace (1989) Genomics 4:560). 최종적으로, 고온에서 기능하는 능력에 의해서 열안정성 템플레이트 의존 DNA 폴리머라제(예를 들면, Taq and Pfu DNA 폴리머라제)는 프라이머에 의해 결합되고 한정된 서열에 대해 높은 특이성을 갖는 것으로 판명되고, 고온에서는, 표적 서열에 의해 프라이머 하이브리드화되고 비-표적 서열에 의해서는 하이브리드화되지 않는 열역학 조건이 수득된다(H. A. Erlich (ed.), PCR Technology, Stockton Press [1989]).

[0090]

본원에서 사용된 "핵산 검출 분석"이라는 용어는 관심 핵산의 뉴클레오티드 조성물을 결정하는 임의의 방법을 의미한다. 핵산 검출 분석은 DNA 서열 방법, 탐침 하이브리드화 방법, 구조 특이적 절단 분석(예를 들면, the INVADER assay, Hologic, Inc., 예를 들면, 미국특허 제5,846,717, 5,985,557, 5,994,069, 6,001,567, 6,090,543, 및 6,872,816; Lyamichev et al., Nat. Biotech., 17:292 (1999), Hall et al., PNAS, USA, 97:8272 (2000), 및 US 2009/0253142에 기재되고, 각각 본원에 참조로 포함되어 있다); 효소 미스매치 절단 방법 (예를 들면, Variagenics, 미국특허 제 6,110,684, 5,958,692, 5,851,770, 본원에 참조로 포함되어 있다); 폴리머라제 연쇄 반응; 브랜치 하이브리드화(branched hybridization) 방법 (예를 들면, Chiron, 미국특허 제 5,849,481, 5,710,264, 5,124,246, 및 5,624,802, 본원에 참조로 포함되어 있다); 롤링 서클 복제 (예를 들면, 미국특허 제 6,210,884, 6,183,960 및 6,235,502, 본원에 참조로 포함되어 있다); NASBA (예를 들면, 미국특허

제 5,409,818, 본원에 참조로 포함되어 있다); 분자 비컨 기술(molecular beacon technology) (예를 들면, 미국특허 제 6,150,097, 본원에 참조로 포함되어 있다); E-센서 기술 (Motorola, U.S. 미국특허 제 6,248,229, 6,221,583, 6,013,170, 및 6,063,573, 본원에 참조로 포함되어 있다); 사이클링 탐침 기술 (예를 들면, 미국특허 제 5,403,711, 5,011,769, 및 5,660,988, 본원에 참조로 포함되어 있다); Dade Behring 신호 증폭 방법 (예를 들면, 미국특허 제 6,121,001, 6,110,677, 5,914,230, 5,882,867, 및 5,792,614, 본원에 참조로 포함되어 있다); 라가제 연쇄 반응 (예를 들면, Barnay Proc. Natl. Acad. Sci USA 88, 189-93 (1991)); 및 샌드위치 하이브리드화(sandwich hybridization) 방법 (예를 들면, 미국특허 제 5,288,609, 본원에 참조로 포함되어 있다)을 포함하지만 이들로 한정되지 않는다.

[0091] "증폭 가능한 핵산"은 임의의 증폭 방법에 의해서 증폭될 수 있는 핵산을 의미한다. "증폭 가능한 핵산"은 보통 "샘플 템플레이트"를 포함하는 것으로 고려된다.

[0092] "샘플 템플레이트"라는 용어는 (하기 정의된)"표적"의 존재하에서 분석된 샘플로부터 유래된 핵산을 의미한다. 반면 "백그라운드 템플레이트"는 샘플에서 존재하거나 존재하지 않을 수 있는 샘플 템플레이트 이외의 핵산과 관련하여 사용된다. 백그라운드 템플레이트는 대부분의 경우 부주의에 의한 것이다. 이는 캐리오버(carryover)의 결과일 수 있거나 샘플로부터 정제되도록 하는 핵산 오염물의 존재에 의한 것일 수 있다. 예를 들면, 검출되는 것 이외의 생물 유래의 핵산은 시험 샘플에서 백그라운드로서 존재할 수 있다.

[0093] "프라이머"라는 용어는 올리고뉴클레오티드를 의미하고, 정제된 제한 소화에서와 같이 천연 발생하거나 또는 합성적으로 생산되는지 관계없이, (예를 들면, 적합한 온도 및 pH에서 DNA 폴리머라제와 같은 뉴클레오티드 및 유도체의 존재하에서)핵산 가닥에 상보적인 프라이머 신장 생성물의 합성이 유도되는 조건하에 있는 경우 합성의 개시점으로서 작용할 수 있다. 프라이머는 바람직하게 증폭시에서 최대 효율을 위해 바람직하게 단일 가닥이지만 이중 가닥이어도 좋다. 이중 가닥인 경우, 프라이머가 먼저 처리되어 신장 생성물의 제조에 사용하기 전에 그 가닥을 분리한다. 프라이머는 유도체의 존재하에서 신장 생성물의 합성을 개시하기 위해 충분히 길어야 한다. 프라이머의 정확한 길이는 온도, 프라이머 공급원 및 방법의 사용을 포함하는 다수의 인자에 의존한다.

[0094] "탐침"라는 용어는 정제된 제한 소화에서와 같이 천연 발생하거나 또는 합성적, 재조합적으로 또는 PCR 증폭으로 생산되는지 관계없이, 또 다른 관심 올리고뉴클레오티드에 하이브리드화 할 수 있는 올리고뉴클레오티드(예를 들면, 뉴클레오티드의 서열)를 의미한다. 탐침은 단일 가닥 또는 이중가닥일 수 있다. 탐침(예를 들면, "포획 탐침")은 특정한 유전자 서열의 검출, 확인, 및 분리에서 유용하다. 본 발명에서 사용되는 임의 탐침은 일부 실시형태에서 임의 "리포터 분자"로 표식되어 효소(예를 들면, ELISA 및 효소 기반 조직화학적 분석), 형광, 방사성 및 발광계를 포함하지만, 이들로 제한되지 않는 임의 검출 시스템에서 검출될 수 있다.

[0095] 본원에서 사용되는 "메틸화"는 시토신의 C5 또는 N4 위치, 아데닌의 N6 위치에서 시토신 메틸화 또는 그 외의 형태의 핵산 메틸화를 의미한다. 시험관내 증폭된 DNA는 보통 메틸화되지 않는데, 일반적인 시험관내 DNA 증폭 방법이 증폭 템플레이트의 메틸화 패턴을 유지하지 않기 때문이다. 그러나, "메틸화되지 않은 DNA" 또는 "메틸화된 DNA"는 각각 본래의 템플레이트가 메틸화되지 않은 증폭된 DNA 또는 메틸화된 증폭된 DNA를 의미할 수 있다.

[0096] 따라서, 본원에서 사용된 "메틸화된 뉴클레오티드" 또는 "메틸화된 뉴클레오티드 염기"는 뉴클레오티드 염기 상의 메틸 부위를 의미하고, 여기서 메틸 부위는 인식된 일반적인 뉴클레오티드 염기가 존재하지 않는다. 예를 들면, 시토신은 피리미딘 고리 상에 메틸 부위를 포함하지 않지만, 5-메틸시토신은 피리미딘 고리의 5 위치에 메틸 부위를 포함한다. 따라서, 메틸화된 뉴클레오티드에는 시토신이 존재하지 않고 5-메틸시토신은 메틸화된 뉴클레오티드이다. 또 다른 예에서, 티민은 피리미딘 고리의 5 위치에 메틸 부위를 포함하지만, 본원의 목적을 위해, 티민이 DNA의 일반적인 뉴클레오티드 염기이기 때문에 티민이 DNA에 존재하는 경우 메틸화된 뉴클레오티드로 고려되지 않는다.

[0097] 본원에 사용되는 "메틸화된 핵산 분자"는 하나 이상의 메틸화된 뉴클레오티드를 포함하는 핵산 분자를 의미한다.

[0098] 본원에서 사용되는 핵산 분자의 "메틸화 상태", "메틸화 프로파일" 및 "메틸화 상황"은 핵산 분자의 하나 이상의 메틸화 뉴클레오티드 염기의 존재 또는 부재를 의미한다. 예를 들면, 메틸화된 시토신을 함유하는 핵산 분자는 메틸화된 것으로 고려된다(예를 들면, 핵산 분자의 메틸화 상태는 메틸화되어 있다). 임의의 메틸화된 뉴클레오티드를 포함하지 않는 핵산 분자는 메틸화되지 않는 것으로 고려된다.

[0099] 특별한 핵산 서열(예를 들면, 본원에 기재된 유전자 마커 또는 DNA 영역)의 메틸화 상태는 서열에서 모든 염기

의 메틸화 상태를 나타낼 수 있거나, 서열 내에서 염기의 서브셋(예를 들면, 하나 이상의 시토신의)의 메틸화 상태를 나타내거나, 메틸화가 발생한 서열 내에서 정확한 위치 정보를 제공하거나 하지 않는 서열 내에서 영역의 메틸화 밀도에 대한 정보를 나타낼 수 있다.

[0100] 핵산 분자내 뉴클레오티드 유전자좌의 메틸화 상태는 핵산 분자의 특정한 유전자좌에서 메틸화 뉴클레오티드의 존재 또는 부재를 의미한다. 예를 들면, 핵산 분자의 제7 뉴클레오티드에서 시토신의 메틸화 상태는, 핵산 분자의 제7 뉴클레오티드에 존재하는 뉴클레오티드가 5-메틸시토신인 경우 메틸화된다. 마찬가지로 핵산 분자에서 제7 뉴클레오티드에서 시토신의 메틸화 상태는 핵산 분자의 제7 뉴클레오티드에 존재하는 뉴클레오티드가 시토신(5-메틸시토신이 아님)인 경우 메틸화되어 있지 않다.

[0101] 메틸화 상태는 "메틸화가(methylation value)"(예를 들면, 메틸화 빈도, 비율, 비, 백분율 등을 나타낸다)에 의해서 나타내거나 표시될 수 있다. 메틸화가는 예를 들면, 메틸화 의존 제한 효소로 제한 소화 후에 존재하는 완전한 핵산의 양을 정량하거나 비설파이트 반응 후 증폭 프로파일을 비교하거나 비설파이트 처리된 및 처리되지 않은 핵산의 서열을 비교함으로써 생성될 수 있다. 따라서, 예를 들면, 메틸화가는 메틸화 상태를 나타내고 따라서 유전자좌의 다수개 카페에 대해 메틸화 상황의 정량적인 지표로서 사용될 수 있다. 이는 특히 문턱값 또는 기준값에 대해 샘플 내에서 서열의 메틸화 상황을 비교하는 것이 요구되는 경우 사용된다.

[0102] 본원에서 사용되는 "메틸화 빈도" 또는 "메틸화 백분율(%)"은 분자 또는 유전자좌가 메틸화되지 않은 경우의 수에 대해 분자 또는 유전자좌가 메틸화된 경우의 수를 의미한다.

[0103] 이와 같이 메틸화 상태는 핵산의 메틸화 상태(예를 들면, 계놈 서열)를 기재한다. 또한, 메틸화 상태는 메틸화와 관련되는 특정한 계놈 유전자좌에서 핵산 세그먼트의 특성을 의미한다. 이러한 특성은, 이들로 제한되지 않지만, 이 DNA 서열 내에 임의의 시토신(C) 잔기가 메틸화되는 것에 관계없이, 메틸화 C 잔기의 위치, 핵산의 임의 특정한 영역에 대해 메틸화 C의 비율 또는 백분율, 및 예를 들면, 대립유전자의 기원의 차이로 인해 메틸화에서 대립유전자 차이를 포함한다. "메틸화 상태", "메틸화 프로파일" 및 "메틸화 상황"은 또한 생물 샘플에서 핵산의 임의의 특정 영역에 대한 메틸화 C 또는 메틸화되지 않은 C의 상대 농도, 절대 농도 및 패턴을 의미한다. 예를 들면, 핵산 서열 내의 시토신(C) 잔기가 메틸화된 경우, 이는 "과메틸화된" 또는 "증가된 메틸화"를 갖는 것으로 지칭될 수 있고, 반면 DNA 서열 내의 시토신(C) 잔기가 메틸화되지 않은 경우, "저메틸화된" 또는 "감소된 메틸화"를 갖는 것으로 지칭될 수 있다. 마찬가지로, 핵산 서열 내의 시토신(C) 잔기가 (예를 들면, 상이한 영역 또는 상이한 개체로부터) 또 다른 핵산 서열에 비해 메틸화되는 경우 서열은 그 외의 핵산 서열에 비해 과메틸화되거나 증가한 메틸화를 갖는 것으로 고려된다. 또한, DNA 서열 내의 시토신(C) 잔기가 (예를 들면, 상이한 영역 또는 상이한 개체 등으로부터) 또 다른 핵산 서열에 비해 메틸화되지 않은 경우, 서열은 그 외의 핵산 서열에 비해 저메틸화되거나 감소한 메틸화를 갖는 것으로 고려된다. 또한, 본원에 사용되는 "메틸화 패턴"은 핵산 영역에 비해 메틸화된 및 메틸화되지 않은 뉴클레오티드의 집합 부위를 의미한다. 2개의 핵산은 동일하거나 유사한 메틸화 빈도 또는 메틸화 백분율을 가질 수 있지만 메틸화된 및 메틸화되지 않은 뉴클레오티드의 수가 상기 영역에 대해 동일하거나 유사하지만 메틸화 및 메틸화되지 않은 뉴클레오티드의 위치가 상이한 경우 상이한 메틸화 패턴을 가질 수 있다. 서열은 이들이 메틸화 정도(예를 들면, 하나가 그 외의 것에 대해 증가한 또는 감소한 메틸화를 갖는다), 빈도, 또는 패턴이 상이한 경우 "차별적으로 메틸화된" 또는 "메틸화의 차이" 또는 "상이한 메틸화 상태"를 갖는 것으로 불린다. "차별적 메틸화"라는 용어는 암 음성 샘플에서 핵산 메틸화의 수준 또는 패턴에 비해 암 양성 샘플에서 핵산 메틸화의 수준 또는 패턴의 차이를 의미한다. 이는 수술후 암 재발한 환자 대 재발하지 않은 환자 사이의 수준 또는 패턴의 차이를 의미할 수 있다. DNA의 차별적 메틸화 및 특이적 수준 또는 패턴은 예를 들면, 정확한 컷오프 또는 예측 특성이 규정되면 예후 및 예측 바이오마커이다.

[0104] 메틸화 상태 빈도는 개체의 집단 또는 단일 개체로부터의 샘플을 기재하기 위해 사용될 수 있다. 예를 들면, 메틸화 상태 빈도 50%를 갖는 뉴클레오티드 유전자좌는 50%의 경우가 메틸화되고 50% 경우가 메틸화되지 않는다. 핵산의 집합 또는 개체의 집단에서 메틸화되는 정도를 기재하기 위해서 이러한 빈도가 사용될 수 있다. 따라서, 핵산 분자의 제1 집단 또는 풀(pool)에서의 메틸화가 핵산 분자의 제2 집단 또는 풀에서의 메틸화와 상이한 경우, 제1 집단 또는 풀의 메틸화 상태 빈도가 제2 집단 또는 풀의 메틸화 상태 빈도와 상이하다. 이러한 빈도는 예를 들면, 뉴클레오티드 유전자좌 또는 핵산 영역이 단일 개체에서 메틸화되는 정도를 기재하기 위해 사용될 수 있다. 예를 들면, 이러한 빈도는 조직 샘플로부터의 샘플의 그룹이 뉴클레오티드 유전자좌 또는 핵산 영역에서 메틸화되거나 메틸화되지 않은 정도를 기재하기 위해 사용될 수 있다.

[0105] 본원에 사용되는 "뉴클레오티드 유전자좌"가 핵산 분자에서 뉴클레오티드의 위치를 의미한다. 메틸화 뉴클레오

티드의 뉴클레오티드 유전자좌는 핵산 분자에서 메틸화된 뉴클레오티드의 위치를 의미한다.

[0106] 일반적으로, 인간 DNA의 메틸화는 인접한 구아닌 및 시토신을 포함하는 디뉴클레오티드 서열 상에서 발생하고, 시토신은 구아닌의 5'에 위치한다(또한 CpG 디뉴클레오티드 서열에서 종결한다). CpG 디뉴클레오티드 내의 대부분의 시토신은 인간 게놈에서 메틸화되지만, 일부는 CpG 섬이라고 알려진 특이적 CpG 디뉴클레오티드 풍부 게놈 영역에서 메틸화되지 않는다 (예를 들면, Antequera et al. (1990) *Cel1* 62: 503-514 참조).

[0107] 본원에서 사용되는 "CpG 섬"은 전체 게놈 DNA에 대해 CpG 디뉴클레오티드의 증가한 수를 함유하는 게놈 DNA의 G:C-풍부 영역을 의미한다. CpG 섬은 길이가 적어도 100, 200개 이상의 염기쌍일 수 있고, 영역의 G:C 함량은 적어도 50%이고, 예상된 빈도에 대한 관찰된 CpG 빈도의 비는 0.6이고, 일부 예에서, CpG 섬은 길이가 적어도 500 개의 염기쌍일 수 있고, 영역의 G:C 함량은 적어도 55%이고, 예상된 빈도에 대한 관찰된 CpG 빈도의 비는 0.65이다. 예상된 빈도에 대한 관찰된 CpG 빈도의 비는 Gardiner-Garden et al (1987) *J. Mol. Biol.* 196: 261-281에 제공된 방법에 따라 산출될 수 있다. 예를 들면, 예상된 빈도에 대한 관찰된 CpG 빈도는 식 $R = (A \times B) / (C \times D)$ 에 따라 산출될 수 있고, R은 예상된 빈도에 대한 관찰된 CpG 빈도의 비이고, A는 분석된 서열에서 CpG 디뉴클레오티드의 수이고, B는 분석된 서열에서 뉴클레오티드의 총수이고, C는 분석된 서열에서 C 뉴클레오티드의 총수이고, D는 분석된 서열에서 G 뉴클레오티드의 총수이다. 메틸화 상태는 예를 들면, 프로모터 영역에서 CpG 섬에서 일반적으로 결정된다. 인간 게놈에서 그 외의 서열은 CpA 및 CpT 등의 DNA 메틸화하기 쉽다는 것이 이해될 것이다 (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 참조).

[0108] 본원에서 사용된, 핵산 분자, 또는 메틸화 특이적 시약의 메틸화 상태에 따라 핵산 분자의 뉴클레오티드를 개질하는 시약은, 핵산 분자의 메틸화 상태를 반영하도록 핵산 분자의 뉴클레오티드 서열을 변화할 수 있는 화합물 또는 조성물을 의미한다. 이러한 시약로 핵산 분자를 처리하는 방법은 상기 시약과 핵산 분자를 접촉하는 것을 포함하고, 이러한 방법은 뉴클레오티드 서열의 목적 변화를 달성하기 위해 필요에 따라 추가의 단계와 결합된다. 이러한 핵산 분자의 뉴클레오티드 서열의 변화는 각각의 메틸화 뉴클레오티드가 상이한 뉴클레오티드로 개질되는 핵산 분자를 형성할 수 있다. 이러한 핵산 뉴클레오티드 서열의 변화는 각각의 메틸화 되지 않은 뉴클레오티드가 상이한 뉴클레오티드로 개질되는 핵산 분자를 형성할 수 있다. 이러한 핵산 뉴클레오티드 서열의 변화는 메틸화되지 않은 선택된 각각의 뉴클레오티드(예를 들면, 각각의 메틸화되지 않은 시토신)가 상이한 뉴클레오티드로 개질되는 핵산 분자를 형성할 수 있다. 핵산 뉴클레오티드 서열을 변화시키기 위한 이러한 시약의 사용에 의해 메틸화된 뉴클레오티드인 각각의 뉴클레오티드 (예를 들면, 각각 메틸화된 시토신)가 상이한 뉴클레오티드로 개질되는 핵산분자를 형성할 수 있다. 사용되는 선택되는 뉴클레오티드를 개질하는 시약의 사용에 의해, 핵산 분자에서 일반적으로 발생하는 4개의 뉴클레오티드(DNA에 대해 C, G, T, 및 A 및 RNA에 대해 C, G, U, 및 A) 중 1개의 뉴클레오티드를 개질하는 시약을 의미하고, 상기 시약은 그 외의 3개의 뉴클레오티드를 개질하지 않고 1개의 뉴클레오티드를 개질한다. 하나의 예시의 실시형태에서, 이러한 시약은 상이한 뉴클레오티드를 생성하기 위해 선택된 메틸화되지 않은 뉴클레오티드를 개질한다. 또 다른 예의 실시형태에서, 이러한 시약은 메틸화되지 않은 시토신 뉴클레오티드를 결정할 수 있다. 예시의 시약은 비설파이트이다.

[0109] 본원에서 사용되는 "비설파이트 시약"은 일부 실시형태에서 예를 들면, CpG 디뉴클레오티드 서열에서 메틸화된 시티딘 및 메틸화되지 않은 시티딘을 구별하기 위해 비설파이트, 디설파이트, 수소 설파이트 또는 이들의 조합을 포함하는 시약을 의미한다.

[0110] "메틸화 분석"은 핵산의 서열에서 하나 이상의 CpG 디뉴클레오티드 서열의 메틸화 상태를 결정하기 위한 임의의 분석을 의미한다.

[0111] "MS AP-PCR" (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction)은 대부분 CpG 디뉴클레오티드를 함유하는 경향이 있는 영역에 촛점을 맞추기 위해 CG-풍부 프라이머를 사용하여 게놈의 글로벌 스캔을 허용하는 당해분야에서 승인된 기술을 의미하고 Gonzalgo et al. (1997) *Cancer Research* 57: 594-599에 기재되어 있다.

[0112] "MethyLight™"은 Eads et al. (1999) *Cancer Res.* 59: 2302-2306에 기재된 당해분야에서 승인된 형광 기반 실시간 PCR 기술을 의미한다.

[0113] "HeavyMethyl™"은 핵산 샘플의 증폭 프라이머 사이 CpG 위치를 커버하는 또는 증폭 프라이머에 의해 커버된 메

틸화 특이적 블로킹 탐침(또한 본원에서 블로커라고 지칭)는 핵산 샘플의 메틸화-특이적 선별 증폭을 가능하게 한다.

- [0114] "'HeavyMethyl™ MethyLight™'은 MethyLight™ 분석의 변형인 HeavyMethyl™ MethyLight™ 분석을 의미하고, MethyLight™ 분석은 증폭 프라이머 사이의 CpG 위치를 커버하는 메틸화 특이적 블로킹 탐침과 결합된다.
- [0115] "Ms-SNuPE" (Methylation-sensitive Single Nucleotide Primer Extension)는 Gonzalgo & Jones (1997) *Nucleic Acids Res.* 25: 2529-2531에 기재된 당해 분야에서 승인된 분석을 의미한다.
- [0116] "MSP" (Methylation-specific PCR)은 Herman et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 9821-9826, 및 미국 특허 제 5,786,146호에 의해서 기재된 당해분야에서 승인된 메틸화 분석을 의미한다.
- [0117] "COBRA" (Combined Bisulfite Restriction Analysis)은 Xiong & Laird (1997) *Nucleic Acids Res.* 25: 2532-2534에 의해서 기재된 당해분야에서 승인된 메틸화 분석을 의미한다.
- [0118] "MCA" (Methylated CpG Island Amplification)은 Toyota et al. (1999) *Cancer Res.* 59: 2307-12, 및 WO 00/26401A1에 의해서 기재된 당해분야에서 승인된 메틸화 분석을 의미한다.
- [0119] 본원에 사용되는 "선별 뉴클레오티드"는 핵산 분자에서 일반적으로 발생하는 4개의 뉴클레오티드(DNA에 대해 C, G, T, 및 A 및 RNA에 대해 C, G, U, 및 A) 중 1개의 뉴클레오티드를 의미하고, 일반적으로 발생하는 뉴클레오티드의 메틸화된 유도체를 포함할 수 있고(예를 들면, C는 선별 뉴클레오티드인 경우, 메틸화된 및 메틸화되지 않은 C는 선별 뉴클레오티드의 의미에 포함된다), 반면 메틸화된 선별 뉴클레오티드는 구체적으로 일반적으로 발생하는 메틸화된 뉴클레오티드를 의미하고, 메틸화되지 않은 선별 뉴클레오티드는 구체적으로 일반적으로 메틸화되지 않은 발생하는 뉴클레오티드를 의미한다.
- [0120] "메틸화-특이적 제한 효소" 또는 "메틸화-민감성 제한 효소"는 인식 부위의 메틸화 상태에 의존하는 핵산을 선별적으로 소화하는 효소를 의미한다. 인식 부위가 메틸화되지 않거나 반메틸화되면 특이적으로 절단된 제한 효소의 경우, 인식 부위가 메틸화되면 절단이 일어나지 않거나 상당히 감소된 효율로 일어난다. 인식 부위가 메틸화되면 특이적으로 절단된 제한 효소의 경우, 인식 부위가 메틸화되지 않으면 절단이 일어나지 않거나 상당히 감소된 효율로 일어난다. 메틸화-특이적 제한 효과가 바람직하고, 인식 서열(예를 들면, CGCG 또는 CCCGGG 등의 인식 서열)은 CG 디뉴클레오티드를 포함한다. 일부 실시형태에 대해 디뉴클레오티드에서 시토신은 탄소원자 C5에서 메틸화되면 절단되지 않은 제한 효소가 바람직하다.
- [0121] 본원에 사용되는 "상이한 뉴클레오티드"는 선별 뉴클레오티드와 화학적으로 상이한 뉴클레오티드이고, 일반적으로 상이한 뉴클레오티드는 선별 뉴클레오티드와 상이한 Watson-Crick 염기쌍 특성을 갖고, 따라서 선별 뉴클레오티드와 상보적인 일반적으로 발생하는 뉴클레오티드는 상이한 뉴클레오티드와 상보적인 일반적으로 발생하는 뉴클레오티드와 동일하지 않다. 예를 들면, C가 선별된 뉴클레오티드인 경우, U 또는 T가 상이한 뉴클레오티드이고, G에 대한 C의 상보성 및 A에 대한 U 또는 T의 상보성에 의해 예시된다. 본원에서 사용되는, 선별된 뉴클레오티드에 상보적이거나 상이한 뉴클레오티드에 상보적인 뉴클레오티드는 매우 엄격한 조건하에서 4개의 일반적으로 발생하는 뉴클레오티드 중 3개와 상보적인 뉴클레오티드 염기쌍을 형성하는 것보다 높은 친화성을 갖는 상이한 뉴클레오티드 또는 선별 뉴클레오티드와 염기쌍을 형성하는 뉴클레오티드를 의미한다. 예시의 상보성은 DNA(예를 들면, A-T 및 C-G) 및 RNA(예를 들면, A-U 및 C-G)에서 Watson-Crick 염기쌍이다. 따라서, 예를 들면, 매우 엄격한 조건하에서 G, A, 또는 T에 대한 G 염기쌍보다 C에 대한 높은 친화성을 갖는 G 염기쌍은, C가 선별된 뉴클레오티드인 경우 G는 선별된 뉴클레오티드에 대한 상보적인 뉴클레오티드이다.
- [0122] 본원에서 사용되는 소정의 마커의 "감도"는 신생물 샘플과 비신생물 샘플을 구별하는 문턱값을 초과한 DNA 메틸화가를 보고하는 샘플의 백분율을 의미한다. 일부 실시형태에서, 양성은 문턱값을 초과한 DNA 메틸화가를 보고한 신생물을 조직학적으로 확인된 것으로 정의하고(예를 들면, 그 범위는 질병과 관련된다), 거짓 음성은 문턱값 미만의 DNA 메틸화가를 보고한 조직학적으로 확인된 신생물로서 정의된다(예를 들면, 그 범위는 질병이 없는 것과 관련된다). 따라서, 감도의 값은 공지된 질병 샘플로부터 수득된 소정의 마커의 DNA 메틸화 측정이 질병 관련된 측정값 범위 내에 있을 확률을 반영한다. 본원에 정의된 바와 같이, 산출된 감도 값의 임상 관련성은 소정의 마커가 임상 질환의 피험체에 적용되는 경우 임상 질환의 존재를 검출할 확률의 추정을 나타낸다.
- [0123] 본원에 사용되는 소정의 마커의 "특이성"은 신생물 샘플과 비신생물 샘플을 구별하는 문턱값을 미만의 DNA 메틸화가를 보고하는 비신생물 샘플의 백분율을 의미한다. 일부 실시형태에서, 음성은 문턱값을 미만의 DNA 메틸화가를 보고한 조직학적으로 확인된 비신생물 샘플로 정의하고(예를 들면, 그 범위는 질병이 없는 것과 관련된다), 거짓 양성은 문턱값 초과의 DNA 메틸화가를 보고한 조직학적으로 확인된 비신생물 샘플로서 정의된다(예를

들면, 그 범위는 질병과 관련된다). 따라서, 특이성의 값은 공지된 비신생물 샘플로부터 수득된 소정의 마커의 DNA 메틸화 측정이 비-질병 관련된 측정값 범위 내에 있을 확률을 반영한다. 본원에 정의된 바와 같이, 산출된 특이성 값의 임상 관련성은 소정의 마커가 임상 질환에 걸리지 않는 환자에 적용되는 경우 임상 질환의 부재를 검출할 확률의 추정을 나타낸다.

[0124] 본원에 사용되는 "AUC"은 "곡선하 면적"의 약칭이다. 특히, 이는 수용자 동작 특성(ROC) 곡선 하 면적을 의미한다. ROC 곡선은 진단시험의 상이한 가능성 컷트 포인트의 거짓 양성률에 대한 진양성률의 플롯이다. 선택된 컷 포인트에 따른 감도와 특이성 사이의 트레이드-오프 관계를 나타낸다(특이성 감소에 의해서 임의의 감도 증가가 달성된다). ROC 곡선하 면적 (AUC)은 진단 시험의 정확성의 척도이다(면적이 클수록 우수하고, 최적값은 1이고; 랜덤 시험은 0.5의 면적을 갖는 대각선 상에 걸쳐 있는 ROC 곡선을 갖는다 (참고로 J. P. Egan. (1975) *Signal Detection Theory and ROC Analysis*, Academic Press, New York).

[0125] 본원에 사용되는 "신생물"은 "비정상 조직의 덩어리를 의미하고, 그 성장이 정상 조직의 것보다 크고 조화되지 않는 것이다. 예를 들면, Willis RA, "The Spread of Tumors in the Human Body", London, Butterworth & Co, 1952 참조.

[0126] 본원에서 사용되는 "선종"은 선 유래의 양성 종양을 의미한다. 이들의 성장은 양성이지만, 시간 경과에 따라 이들은 악성 종양으로 진행될 수 있다.

[0127] "전암성" 또는 "전-신생물" 및 그 균등물은 악성 형질 전환을 겪는 임의의 세포 증식 장애를 의미한다.

[0128] 신생물, 선종, 암 등의 "부위"는 신생물, 선종 암이 위치한 피험체 신체 내에서 조직, 기관, 세포 유형 해부학적 영역, 신체 부위이다.

[0129] 본원에서 사용되는 "진단적" 시험 적용은 피험체의 질병 또는 질환의 검출 또는 확인을 포함하고, 피험체가 소정의 질병 또는 질환을 수축할 가능성을 결정하고, 질병 또는 질환에 걸린 피험체가 치료에 반응할 가능성을 결정하고, 질병 또는 질환에 걸린 피험체의 예후(또는 진행 또는 회귀 가능성)를 결정하고, 질병 또는 질환에 걸친 피험체에 대한 치료 효과를 결정한다. 예를 들면, 진단은 신생물을 수축하는 피험체의 존재 또는 가능성 또는 이러한 피험체가 화합물(예를 들면, 약학적, 예를 들면, 약물) 또는 그 외의 치료에 바람직하게 반응할 가능성을 검출하기 위해서 사용될 수 있다.

[0130] 본원에서 사용되는 "마커"는 예를 들면, 메틸화 상태에 기초해서 정상 세포로부터 암 세포를 구별함으로써 암을 진단할 수 있는 물질 (예를 들면, 핵산 또는 핵산의 영역)을 의미한다.

[0131] "분리된 올리고뉴클레오티드"에서와 같은 핵산에 대해 사용되는 "분리된"은 천연 공급원에서 통상적으로 관련되어 있는 적어도 하나의 오염물 핵산으로부터 확인되고 분리되는 핵산 서열을 의미한다. 분리된 핵산은 자연에서 발견되는 것과 상이한 형태 또는 설정에 존재한다. 반면, 비-분리된 핵산, 예를 들면, DNA 및 RNA은 천연에서 존재하는 상태로 발견된다. 비-분리된 핵산의 예로는 이웃 유전자에 근접한 숙주 세포 염색체에서 발견되는 소정의 DNA 서열 (예를 들면, 유전자); RNA 서열, 예를 들면, 단백질의 다수의 단백질을 인코딩하는 그 외의 수 많은 mRNA와의 혼합물로서 세포에서 발견되는 특이적 단백질을 인코딩하는 특이적 mRNA 서열을 포함한다. 그러나, 특정한 단백질을 인코딩하는 분리된 핵산은, 일례로 단백질을 통상적으로 발현하는 세포 종의 핵산을 포함하고, 핵산은 천연 세포와 상이한 염색체 위치에 있거나 천연에서 발견되는 것과 상이한 핵산 서열에 의해서 인접하고 있다. 분리된 핵산 또는 올리고뉴클레오티드는 단일 가닥 또는 이중 가닥 형태로 존재할 수 있다. 분리된 핵산 또는 올리고뉴클레오티드는 단백질을 발현하기 위해 이용되는 경우, 올리고뉴클레오티드는 최소 센스 또는 코팅 가닥을 포함하지만(즉, 올리고뉴클레오티드는 단일 가닥일 수 있다), 센스 및 안티센스 가닥을 포함할 수 있다(즉, 올리고뉴클레오티드는 이중 가닥일 수 있다). 분리된 핵산은, 천연 또는 일반적인 환경으로부터 분리된 후 다른 핵산 또는 문자와 결합될 수 있다. 예를 들면, 분리된 핵산은 예를 들면, 이중 발현을 위해 배치된 숙주세포에 존재할 수 있다.

[0132] "정제된"은 천연 환경으로부터 제거되거나 단리되거나 분리되는 핵산 또는 아미노산 서열인 분자를 의미한다. "분리된 핵산 서열"은 따라서 정제된 핵산서열일 수 있다. "실질적으로 정제된" 분자는 자연적으로 회합된 다른 성분으로부터 적어도 60% 유리, 바람직하게 적어도 75% 유리, 더 바람직하게 적어도 90% 유리된다. 본원에 사용되는 "정제된" 또는 "정제하기 위해"라는 용어는 샘플로부터 오염물을 제거하는 것을 의미한다. 단백질을 오염시키는 것을 제거하면, 샘플에서 관심있는 폴리펩티드 또는 핵산의 백분율을 증가시킨다. 또 다른 예로는 재조합 폴리펩티드는 식물, 세균, 효모 또는 포유류 숙주세포에서 발현되고 폴리펩티드는 숙주 세포 단백질을 제거하여 정제되고, 이 샘플 중의 재조합 폴리펩티드의 백분율이 증가한다.

- [0133] 소정의 폴리뉴클레오티드 시퀀서 또는 폴리펩티드를 "포함하는 조성물"은 광범위하게 소정의 폴리뉴클레오티드 서열 또는 폴리펩티드를 포함하는 임의의 조성물을 의미한다. 조성물은 염(예를 들면, NaCl), 세제(예를 들면, SDS), 및 그 외의 성분(예를 들면, Denhardt's 용액, 분유, 연어 정자 DNA, 등)을 함유하는 수용액을 포함한다.
- [0134] "샘플"이 가장 넓은 의미로 사용된다. 어떤 의미에서는 동물 세포 또는 조직을 의미할 수 있다. 또 다른 의미에서는 생물학적 및 환경적 샘플뿐 아니라 임의의 공급원으로부터 수득된 표본 또는 배양물을 포함하는 것을 의미하다. 생물학적 샘플은 식물 또는 동물(인간을 포함)로부터 수득될 수 있고, 유체, 고체, 조직 및 기체를 포함한다. 일부 실시형태에서, 샘플은 혈장 샘플이다. 일부 실시형태에서, 샘플은 위 조직 샘플이다. 일부 실시형태에서, 샘플은 대변 샘플이다. 환경 샘플은 표면 물질, 토양, 물 및 산업 샘플 등의 환경 물질을 포함한다. 이러한 샘플은 본 발명에 적용 가능한 샘플 유형을 제한하는 것으로 해석되어서는 안된다.
- [0135] 본원에서 사용되는 일부 문맥에서 사용되는 "원격 샘플"은 샘플의 세포, 조직, 또는 장기 공급원이 아닌 부위로부터 간접적으로 수집된 샘플을 의미한다. 예를 들면, 췌장으로부터 유래된 샘플 물질이 대변 샘플에서 분석되는 경우(예를 들면, 췌장에서 직접 취한 샘플이 아님), 샘플은 원격 샘플이다.
- [0136] 본원에서 사용되는 "환자" 또는 "피험체"는 본 기술에 의해서 제공되는 다양한 시험을 받는 생물체를 의미한다. "피험체"는 동물을 포함하고, 바람직하게 인간 등의 포유류이다. 바람직한 실시형태에서, 피험체는 영장류이다. 또 다른 바람직한 실시형태에서, 피험체는 인간이다.
- [0137] 본원에서 사용되는 "키트"는 물질을 전달하기 위한 임의의 전달 시스템을 의미한다. 반응 분석의 문맥에서, 이러한 전달 시스템은 하나의 위치로부터 또 다른 위치로 반응 시약(예를 들면, 적절한 용기에서 올리고뉴클레오티드, 효소, 등) 및/또는 지지 물질(예를 들면, 베퍼, 분석 등을 실시하기 위한 지시사항)의 저장, 전송, 또는 전달을 허용하는 시스템을 포함한다. 예를 들면, 키트는 관련된 반응 시약 및/또는 지지 물질을 함유하는 하나 이상의 인클로저(예를 들면, 박스)를 포함한다. 본원에서 사용되는 "단편화 키트"는 각각 전체 키트 성분의 하위 요소를 포함하는 2개 이상의 분리 용기를 포함하는 전달 시스템을 의미한다. 용기는 의도된 수령인에게 함께 또는 개별적으로 전달될 수 있다. 예를 들면, 제1 용기는 분석에 사용하기 위한 효소를 포함하지만, 제2 용기는 올리고뉴클레오티드를 포함한다. "단편화 키트"는 Federal Food, Drug, and Cosmetic Act의 520(e)항에 의해 규제되는 분석물 특이적 시약(ASR)을 함유하는 키트를 포함하는 것으로 의도되지만, 이들로 제한되지 않는다. 실제, 각각 전체 키트 구성성분의 하위요소를 포함하는 2개 이상의 개별 용기를 포함하는 임의의 전달 시스템은 "단편화 키트"에 포함된다. 반면, "결합된 키트"는 단일 용기(예를 들면, 각각의 목적 성분을 수용하는 단일 박스에서)에서 반응 분석의 모든 성분을 함유하는 전달 시스템을 의미한다. "키트"는 단편화되고 결합된 키트를 포함한다.
- [0138] 기술의 실시형태
- [0139] 본원에는 신생물을 검출하는 것이 제공되고, 특히 배타적이지 않지만 위암과 같은 전악성 및 악성 신생물을 검출하기 위한 방법, 조성물 및 관련 용도가 제공된다. 마커는 위암(gastric cancer)(예를 들면, 위암(stomach cancer))에 걸린 피험체의 종양으로부터 DNA 마커의 메틸화 상태를, 대조 피험체(예를 들면, 실시예 1 및 2)로부터 동일한 DNA 마커의 메틸화 상태와 비교함으로써 증례-대조군 연구에서 확인되었다(예를 들면, 실시예 1 및 2 참조). 추가의 실험은 위암을 검출하기 위한 10개의 최적의 마커(ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 및 C130RF18; 실시예 2 및 표 4 참조)를 확인했다. 추가의 실험은 위암을 검출하기 위해 30개의 최적의 마커를 확인했다(실시예 1 및 표 2 참조). 추가의 실험은 혈장 샘플에서 위암(예를 들면, 위암)을 검출하기 위한 12개의 최적의 마커(ELMO1, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; 실시예 3 및 표 1, 2, 4 및 7 참조)를 확인했다.
- [0140] 마커 및/또는 마커의 패널(예를 들면, ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 또는 C130RF18로부터 선택된 주석을 갖고 상기 마커를 포함하는 염색체 영역(표 4 참조))은 위암(예를 들면, 위암)에 걸린 피험체의 예를 들면 위 조직으로부터 DNA 마커의 메틸화 상태를 대조군 피험체로부터 동일한 DNA 마커의 메틸화 상태를 비교하여 증례-대조군 연구에서 확인되었다(실시예 2 참조).
- [0141] 마커 및/또는 마커의 패널(예를 들면, ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 또는 ZNF569로부터 선택된 주석을 갖고 상기 마커를 포함하는 염색체 영역 (표 2 참조))은 위암(예를 들면, 위암)에 걸린 피험체의 예를 들면 위 조직으로부

터 DNA 마커의 메틸화 상태를 대조군 피험체로부터 동일한 DNA 마커의 메틸화 상태를 비교하여 증례-대조군 연구에서 확인되었다(실시예 1 참조).

[0142] 마커 및/또는 마커의 패널(예를 들면, ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1로부터 선택된 주석을 갖고, 상기 마커를 포함하는 염색체 영역 (표 1, 2, 4 및 7 참조))은 위암(예를 들면, 위암)에 걸린 피험체의 예를 들면 혈장으로부터 DNA 마커의 메틸화 상태를 대조군 피험체로부터 동일한 DNA 마커의 메틸화 상태를 비교하여 증례-대조군 연구에서 확인되었다(실시예 3 참조).

[0143] 또한, 본 기술은 다양한 마커의 패널을 제공하고, 예를 들면, 일부 실시형태에서 마커는 ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 또는 C13ORF18인 주석을 갖고 상기 마커를 포함하는 염색체 영역을 포함한다(표 4 참조). 또한, 실시형태는 DMR No. 49, 43, 196, 66, 1, 237, 249, 250, 251 또는 252인 표 4로부터 DMR을 분석하는 방법을 제공한다. 일부 실시형태는 마커의 메틸화 상태를 결정하는 것을 제공하고, 염색체 영역은 ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 또는 ZNF569인 주석을 갖고 상기 마커를 포함한다(표 2 참조). 또한, 실시형태는 DMR No. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 또는 237인 표 2로부터 DMR을 분석하는 방법을 제공한다. 수득된 샘플이 혈장 샘플인 일부 실시형태에서, 마커는 ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1인 주석을 갖고 상기 마커를 포함하는 염색체 영역을 포함한다. 또한, 샘플이 혈장 샘플인 이러한 실시형태에서 DMR No. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 또는 237인 표 1, 2, 4 및 7로부터 DMR을 분석하는 방법이 있다. 일부 실시형태에서, 상기 방법은 2개의 마커의 메틸화 상태, 예를 들면, 표 1, 2, 4 및 7의 행에 제공되는 한 쌍의 마커를 결정하는 것을 포함한다.

[0144] 본원에서의 개시 내용은 특정한 예시의 실시형태를 나타내지만, 이들의 실시형태는 일례로 나타내고 제한되지 않는 것으로 이해되어야 한다.

[0145] 특정한 형태에서, 본 개시내용은 위암(예를 들면, 위암)과 같은 암을 확인, 결정, 및/또는 분류하는 조성물 및 방법을 제공한다. 상기 방법은 피험체로부터 분리된 생물학적 샘플(예를 들면, 대변 샘플, 위 조직 샘플, 혈장 샘플)에서 적어도 하나의 메틸화 마커의 메틸화 상태를 결정하는 것을 포함하고, 마커의 메틸화 상태 변화는 위암(예를 들면, 위암)의 존재, 분류 또는 부위를 나타낸다. 특정한 실시형태는 질병의 전암 단계 중에 조기 검출을 포함하는 신생물 세포 증식 장애(예를 들면, 암)의 진단(예를 들면, 스크리닝)에 사용되는 차별적으로 메틸화된 영역(DMR, 예를 들면, 1-274, 예를 들면, 표 1, 2, 4, 및 7)을 포함하는 마커에 관련된다.

[0146] 또한, 마커는 양성 세포 증식 장애로부터 신생물의 구별에 사용된다. 특정한 형태에서, 본 기술은 신생물 세포 증식 장애가 양성 세포 증식 장애와 구별되는 방법을 개시한다.

[0147] 본원에 제공되고 표 1, 2, 4 또는 7에 열거된 DMR(예를 들면, DMR, 예를 들면, DMR 1-274)을 포함하는 마커의 염기, 마커의 영역 또는 적어도 하나의 마커의 메틸화 분석이 분석되고, 본 기술은 또한 암, 특히 위암의 검출을 위한 유용성으로 DMR을 포함하는 마커의 염기, 마커의 영역, 적어도 하나의 마커를 포함하는 마커의 패널을 제공한다.

[0148] 본 기술의 일부 실시형태는 적어도 하나의 마커, 마커의 영역 또는 DMR을 포함하는 마커의 염기의 CpG 메틸화 상황의 분석에 기초한다.

[0149] 일부 실시형태에서, 본 기술은 DMR을 포함하는 적어도 하나의 마커 내에 CpG 디뉴클레오티드 서열의 메틸화 상황을 결정하기 위해 하나 이상의 메틸화 분석과 함께 비설파이트 기술을 사용하는 것을 제공한다 (예를 들면, DMR 1-274, 표 1, 2, 4 및 7 참조). 계놈 CpG 디뉴클레오티드는 메틸화되거나 메틸화되지 않을 수 있다(또한 업- 및 다운-메틸화된으로서 알려져 있다). 그러나, 본 발명의 방법은 원격 샘플의 백그라운드 내에 이종 상태의 생물학적 샘플(예를 들면, 혈액, 장기 유출물, 또는 대변) 예를 들면, 저농도 종양 세포, 또는 이로부터 수득된 생물학적 샘플의 분석에 적합하다. 따라서, 이러한 샘플 내에 CpG 위치의 메틸화 상태를 분석하는 경우 특정한 CpG 위치에서 메틸화의 수준(예를 들면, 백분율, 분율, 비율, 비, 또는 정도)을 결정하기 위해 정량적인 분석이 사용될 수 있다.

[0150] 본 기술에 따라서, DMR을 포함하는 마커 내의 CpG 디뉴클레오티드 서열의 메틸화 상황의 결정은 위암과 같은 암

이 진단 및 특성화에서의 유용성을 갖는다.

[0151] 마커의 조합

[0152] 일부 실시형태에서, 본 기술은 표 1의 DMR(예를 들면, DMR Nos. 1-248) 또는 표4의 DMR(예를 들면, DMR Nos., 49, 43, 196, 66, 1, 237, 249, 250, 251 또는 252) 또는 표 2의 DMR(DMR Nos. 253, 251, 254, 255, 256, 249, 257, 258, 259, 260, 261, 262, 250, 263, 1, 264, 265, 196, 266, 118, 267, 268, 269, 270, 271, 272, 46, 273, 252, 또는 237) 또는 표 7의 DMR(DMR No. 274)을 포함하는 마커 또는 DMR을 포함하는 추가의 마커의 조합의 메틸화 상태를 분석하는 것에 관한 것이다. 일부 실시형태에서, 2개 이상의 마커의 메틸화 상태를 분석하면, 피험체 내의 신생물(예를 들면, 위암)을 확인하기 위한 스크린 또는 진단의 특이성 및/또는 감도를 증가시킨다. 일부 실시형태에서, 마커 또는 마커의 조합은 신생물의 유형 및/또는 위치 사이에서 구별된다.

[0153] 다양한 암은 예를 들면, 예측의 특이성 및 감도에 관한 통계기술에 의해 확인되는 다양한 마커의 조합에 의해서 예측된다. 본 기술은 일부 암의 예측 조합 및 유효성 예측 조합을 확인하는 방법을 제공한다.

[0154] 메틸화 상태를 분석하는 방법

[0155] 5-메틸시토신의 존재하에서 핵산의 분석 방법에서 가장 빈번하게 사용되는 방법은 DNA에서 5-메틸시토신의 검출을 이해 Frommer, et al에 의해 기재된 방법 또는 그 변형에 기초한다 (Frommer et al. (1992) *Proc. Natl. Acad. Sci. USA* 89: 1827-31 는 모든 목적을 위해 그 전체 참조로 본원에 포함되어 있다). 5-메틸시토신을 맵핑하는 비설파이트 방법은 5-메틸시토신이 아닌 시토신이 수소 설파이트 이온(비설파이트로 알려짐)과 반응하는 관찰에 기초한다. 반응은 보통 다음의 단계에 따라 수행된다: 먼저 시토신이 수소 설파이트와 반응하여 설폰화 시토신을 형성한다. 그 다음, 설폰화 반응 중간체의 자발적인 탈아민화에 의해서 설폰화 우라실이 형성된다. 마지막으로, 설폰화 우라실은 알칼리 조건하에서 탈설폰화되어 우라실을 형성한다. 우라실이 아데닌과 염기쌍을 형성하는 반면(따라서, 티민과 같이 작용한다) 구아닌과 5-메틸시토신 염기쌍을 형성(시토신과 같이 작용)하기 때문에, 검출 가능하다. 이는 예를 들면, 비설파이트 개념 서열 결정(Grigg G, & Clark S, *Bioessays* (1994) 16: 431-36; Grigg G, *DNA Seq.* (1996) 6: 189-98) 또는 또는 미국 특허 제5,786,146에 개시된 메틸화-특이적 PCR(MSP)에 의해 메틸화되지 않은 시토신으로부터 메틸화된 시토신을 구별할 수 있다

[0156] 일부 종래의 기술은 아가로스 매트릭스에서 분석될 DNA를 둘러싸고 DNA의 확산 및 재생을 방지하고(비설파이트 만 단일 가닥 DNA와 반응하고), 침전 및 정제단계를 빠른 투석으로 대체하는 것을 포함하는 방법에 관한 것이다 (Olek A, et al. (1996) "A modified and improved method for bisulfite based cytosine methylation analysis" *Nucleic Acids Res.* 24: 5064-6). 따라서 메틸화 상태의 개별 세포를 분석하고 상기 방법의 유용성 및 감도를 설명할 수 있다. 5-메틸시토신을 검출하는 종래의 방법의 고찰은 Rein, T., et al. (1998) *Nucleic Acids Res.* 26: 2255에 의해서 제공된다.

[0157] 비설파이트 기술은 비설파이트 처리 후 일반적으로 공지된 핵산의 특이적 짧은 단편을 증폭한 다음 서열 결정 (Olek & Walter (1997) *Nat. Genet.* 17: 275-6) 또는 프라이머 신장 반응(Gonzalgo & Jones (1997) *Nucleic Acids Res.* 25: 2529-31; WO 95/00669; 미국 특허 제 6,251,594호)에 의해 생성물을 분석하여 개별 시토신 위치를 분석하는 것을 포함한다. 일부 방법은 효소 소화를 사용한다 (Xiong & Laird (1997) *Nucleic Acids Res.* 25: 2532-4). 하이브리드화에 의한 검출이 해당 기술분야에 기재되어 있다 (Olek et al., WO 99/28498). 또한, 개별 유전자에 대해 메틸화 검출을 위해 비설파이트 기술을 사용하는 것이 기재되어 있다 (Grigg & Clark (1994) *Bioessays* 16: 431-6,; Zeschnigk et al. (1997) *Hum Mol Genet.* 6: 387-95; Feil et al. (1994) *Nucleic Acids Res.* 22: 695; Martin et al. (1995) *Gene* 157: 261-4; WO 9746705; WO 9515373).

[0158] 다양한 메틸화 분석 절차는 해당 기술분야에서 공지되어 있고 본 기술에 따라 비설파이트 처리와 함께 사용될 수 있다. 이러한 분석은 핵산 서열 내에 하나 또는 복수의 CpG 디뉴클레오티드(예를 들면, CpG 섬)의 메틸화 상태의 결정을 가능하게 한다. 이러한 분석은, 그 외의 기술 중에서 비설파이트처리된 핵산의 서열결정, PCR(서열 특이적 증폭), 서던 블롯 분석 및 메틸화-민감성 제한 효소의 이용을 포함한다.

[0159] 예를 들면, 개념 서열 결정은 비설파이트 처리를 사용하여 5-메틸시토신 분포 및 메틸화 패턴의 분석을 위해 단순하게 했다(Frommer et al. (1992) *Proc. Natl. Acad. Sci. USA* 89: 1827-1831). 또한, 비설파이트 전환된 DNA로부터 증폭된 PCR 제품의 제한 효소 소화는 예를 들면, Sadri & Hornsby (1997) *Nucleic Acids Res.* 24: 5058-5059에 기재되거나 COBRA (Combined Bisulfite Restriction Analysis)으로 알려진 방법에서 구현화된 (Xiong & Laird (1997) *Nucleic Acids Res.* 25: 2532-2534) 메틸화 상태를 분석하는 데에 사용한다.

[0160] COBRA™ 분석은 소량의 개념 DNA에서 특이적 유전자좌에서 DNA 메틸화를 수준을 결정하는 데 유용한 정량적인

메틸화 분석이다 (Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997). 간략하게, 소디움 비설파이트 처리된 DNA의 PCR 생성물에서 메틸화 의존 서열 차이를 나타내기 위해 제한 효소 소화가 사용된다. 메틸화 의존 서열 차이는 먼저 Frommer et al.에 의해 기재된 절차에 따라 표준 비설파이트 처리에 의해서 계놈 DNA에 도입된다 (Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992). 비설파이트 전환된 DNA의 PCR 증폭은 관심 CpG 섬에 특이적인 프라이머를 사용하여 수행된 후 제한 엔도뉴클레아제 소화, 겔 전기영동 및 특이적 표지된 하이브리드화 탐침을 사용하는 검출이 행해진다. 본래의 DNA 샘플에서의 메틸화 수준은 DNA 메틸화 수준의 넓은 스펙트럼에 대해 선형적으로 정량적인 방식으로 소화된 및 소화되지 않은 PCR 생성물의 상대적으로 양으로 표시된다. 또한, 본 기술은 미세절개된 파라핀-포매 조직 샘플로부터 수득된 DNA에 확실하게 적용할 수 있다.

[0161] COBRA™ 분석용 전형적 시약(예를 들면, 전형적인 COBRA™-기반 키트에서 볼 수 있다)은 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등); 제한 효소 및 적절한 버퍼; 유전자 하이브리드화 올리고뉴클레오티드; 대조군 하이브리드화 올리고뉴클레오티드; 올리고뉴클레오티드 탐침의 키나제 표지 키트; 및 표지된 뉴클레오티드를 포함하지만 이들로 제한되지 않는다. 또한, 비설파이트 전환 시약은, DNA 변성 버퍼; 설폰화 버퍼; DNA 회수 시약 또는 키트(예를 들면, 침전, 한외여과, 친화성 컬럼); 디설폰화 버퍼; 및 DNA 회수 성분을 포함할 수 있다.

[0162] 바람직하게 "MethyLight™" (형광 기반 실시간 PCR 기술) (Eads et al., Cancer Res. 59:2302-2306, 1999), Ms-SNuPE™ (메틸화 민감성 단일 뉴클레오티드 프라이머 신장) 반응 (Gonzalgo & Jones, Nucleic Acids Res. 25:2529-2531, 1997), 메틸화 특이적 PCR ("MSP"; Herman et al., Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; 미국 특허 제 5,786,146호), 및 메틸화 CpG 섬 증폭 ("MCA"; Toyota et al., Cancer Res. 59:2307-12, 1999) 등의 분석은 단독으로 또는 이들 방법 중 하나 이상과 조합해서 사용된다.

[0163] "HeavyMethyl™" 분석은 비설파이트 처리된 DNA의 메틸화 특이적 증폭에 기초한 메틸화 차이를 분석하는 정량적인 방법이다. 증폭 프라이머 사이에서 CpG를 커버하는 또는 이에 의해 커버되는 메틸화 특이적 블로킹 탐침("블로커")는 혼란 샘플의 메틸화 특이적 민감한 증폭을 가능하게 한다.

[0164] "HeavyMethyl™ MethyLight™" 분석은 HeavyMethyl™ MethyLight™ 분석을 의미하고, 이는 MethyLight™의 변형이고, MethyLight™ 분석은 증폭 프라이머 사이의 CpG 위치를 커버하는 메틸화 특이적 블로킹 탐침과 결합된다. HeavyMethyl™ 분석은 메틸화 특이적 증폭 프라이머와 함께 사용될 수 있다.

[0165] HeavyMethyl™ 분석용 전형적인 시약(예를 들면, 전형적인 MethyLight™ 기반 키트에서 볼 수 있다)은, 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등, 또는 비설파이트 처리된 DNA 서열, 또는 CpG 섬); 블로킹 올리고뉴클레오티드 및 최적화된 PCR 버퍼 및 테옥시뉴클레오티드; 및 Taq 폴리머라제를 포함하지만 이들로 제한되지 않는다.

[0166] MSP(메틸화 특이적 PCR)은 메틸화-민감성 제한 효소의 사용에 관계없이 CpG 섬 내에 CpG 부위의 사실상 임의의 그룹의 메틸화 상황을 분석할 수 있다 (Herman et al. Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; 미국 특허 제No. 5,786,146호). 간략하게 DNA는 소디움 비설파이트에 의해 개질되고, 메틸화된 시토신이 아닌 메틸화되지 않은 시토신을 우라실로 전환하고 생성물을 메틸화 DNA 대 메틸화되지 않은 DNA에 특이적인 프라이머에 의해 증폭된다. MSP는 단지 소량의 DNA만을 필요로 하고, 이는 소정의 CpG 섬 유전자좌의 0.1% 메틸화된 대립 유전자에 민감하고, 파라핀-포매된 샘플로부터 추출된 DNA 상에서 수행될 수 있다. MSP 분석용 전형적인 시약(예를 들면, 전형적인 MSP-기반 키트에서 볼 수 있다)은 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등); 최적화된 PCR 버퍼 및 테옥시뉴클레오티드 및 특이적 탐침을 포함하지만 이들로 제한되지 않는다.

[0167] MethyLight™ 분석은 PCR 단계 후 추가의 조작을 필요로 하지 않는 형광 기반 실시간 PCR(예를 들면, TaqMan®)을 이용하는 고 쓰로우풋 정량적인 메틸화 분석이다 (Eads et al., Cancer Res. 59:2302-2306, 1999). 간략하게 MethyLight™ 공정은 계놈 DNA의 혼합 샘플로 시작해서 소디움 비설파이트 반응에서 표준 절차에 따라 메틸화 의존 서열 차이의 혼합 폴로 전환된다(비설파이트 공정은 메틸화되지 않은 시토신 잔기를 우라실로 전환된다). 형광 기반 PCR은 공지된 CpG 디뉴클레오티드를 오버랩하는 PCR 프라이머와의 "편향된(biased)" 반응에서 수행된다. 서열 구별은 형광 검출 공정의 수준 및 증폭 공정의 수준에서 발생한다.

[0168] MethyLight™ 분석은 혼란, 예를 들면 계놈 DNA 샘플에서 메틸화 패턴의 정량적인 시험으로서 사용되고, 서열 구별은 탐침 하이브리드화의 수준에서 발생한다. 정량적인 버전에서, PCR 반응은 특정한 추정 메틸화 부위를 오버랩하는 형광 탐침의 존재하에서 특정한 메틸화 증폭을 제공한다. 입력 DNA 양에 대한 편향되지 않은 제어

는 프라이머 또는 탐침이 임의의 CpG 디뉴클레오티드 위에 놓이지 않는 반응에 의해서 제공된다. 또한, 계놈 메틸화의 정량적인 시험은 공지의 메틸화 부위를 커버하지 않는 대조군 올리고뉴클레오티드(예를 들면, HeavyMethyl™ 및 MSP 기술의 형광 기반 버전) 또는 잠재적인 메틸화 부위를 커버하는 올리고뉴클레오티드로 편향된 PCR 풀을 탐침함으로써 달성된다.

[0169] MethyLight™ 공정은 임의의 적합한 탐침에 의해서 사용된다(예를 들면, "TaqMan®" 탐침, Lightcycler® 탐침, 등). 예를 들면, 일부 적용에서 이중 가닥 계놈 DNA가 소디움 비설파이트로 처리되고 예를 들면 MSP 프라이머 및/또는 HeavyMethyl 블로커 올리고뉴클레오티드를 갖는 다수개의 TaqMan® 탐침 및 하나의 TaqMan® 탐침을 사용하여 2세트의 PCR 반응 중 하나를 행한다. TaqMan® 탐침은 형광 "리포터" 및 "쿠엔처" 분자로 이중 표지되고 비교적 높은 GC 함량 영역에 특이적이도록 설계되고 정방향 프라이머 또는 역방향 프라이머보다 PCR 사이클에서 약 10°C 높은 온도에서 용융된다. 이는 PCR 어닐링/신장 단계 동안 TaqMan® 탐침이 충분히 하이브리드화될 수 있다. Taq 폴리머라제는 PCR 중에 새로운 가닥을 효소적으로 합성하기 때문에, 최종적으로 어닐링된 TaqMan® 탐침에 도달한다. Taq 폴리머라제 5'에서 3'로의 엔도뉴클레아제 활성은 이를 소화함으로써 TaqMan® 탐침을 치환하고 실시간 형광 검출 시스템을 사용하여 현재 쿠엔칭되지 않은 신호의 정량적인 검출을 위해 형광 리포터 분자를 방출한다.

[0170] MethyLight™ 분석용 전형적인 시약(예를 들면, 전형적인 MethyLight™ 기반 키트에서 볼 수 있다)은, 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등); TaqMan® 또는 Lightcycler® 탐침; 최적화된 PCR 버퍼 및 데옥시뉴클레오티드 및 Taq 폴리머라제를 포함하지만 이들로 제한되지 않는다.

[0171] QM™(정량적인 메틸화) 분석은 계놈 DNA 샘플에서 메틸화 패턴의 또 다른 정량적인 시험이고, 서열 구별은 탐침 하이브리드화의 수준에서 발생한다. 정량적인 버전에서, PCR 반응은 특정한 추정 메틸화 부위를 오버랩하는 형광 탐침의 존재하에서 편향되지 않은 증폭을 제공한다. 입력 DNA 양에 대한 편향되 않은 제어는 프라이머 또는 탐침이 임의의 CpG 디뉴클레오티드 위에 놓이지 않는 반응에 의해서 제공된다. 또한, 계놈 메틸화의 정량적인 시험은 공지의 메틸화 부위를 커버하지 않는 대조군 올리고뉴클레오티드(예를 들면, HeavyMethyl™ 및 MSP 기술의 형광 기반 버전) 및 잠재적인 메틸화 부위를 커버하는 올리고뉴클레오티드로 편향된 PCR 풀을 탐침함으로써 달성된다.

[0172] QM™ 공정은 임의의 증폭 공정에서 임의의 적합한 탐침으로 사용되고, 예를 들면, "TaqMan®" 탐침, Lightcycler® 탐침, 등이다. 예를 들면, 이중 가닥 계놈 DNA가 소디움 비설파이트로 처리되고 TaqMan® 탐침 및 편향되지 않은 프라이머를 수행한다. TaqMan® 탐침은 형광 "리포터" 및 "쿠엔처" 분자로 이중 표지되고 비교적 높은 GC 함량 영역에 특이적이도록 설계되고 정방향 프라이머 또는 역방향 프라이머보다 PCR 사이클에서 약 10°C 높은 온도에서 용융된다. 이는 PCR 어닐링/신장 단계 동안 TaqMan® 탐침이 충분히 하이브리드화될 수 있다. Taq 폴리머라제는 PCR 중에 새로운 가닥을 효소적으로 합성하기 때문에, 최종적으로 어닐링된 TaqMan® 탐침에 도달한다. Taq 폴리머라제 5'에서 3'로의 엔도뉴클레아제 활성은 이를 소화함으로써 TaqMan® 탐침을 치환하고 실시간 형광 검출 시스템을 사용하여 현재 쿠엔칭되지 않은 신호의 정량적인 검출을 위해 형광 리포터 분자를 방출한다. QM™ 분석용 전형적인 시약(예를 들면, 전형적인 QM™ 기반 키트에서 볼 수 있다)은 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등); TaqMan® 또는 Lightcycler® 탐침; 최적화된 PCR 버퍼 및 데옥시뉴클레오티드 및 Taq 폴리머라제를 포함하지만 이들로 제한되지 않는다.

[0173] Ms-SNuPE™ 기술은 DNA의 비설파이트 처리에 기초하여 특이적 CPG 부위에서 메틸화 차이를 분석한 후 단일 뉴클레오티드 프라이머 신장을 행하기 위한 정량적인 방법이다 (Gonzalgo & Jones, Nucleic Acids Res. 25:2529-2531, 1997). 간략하게, 계놈 DNA는 소디움 비설파이트와 반응하여 메틸화되지 않은 시토신을 우라실로 전환하는 동안 5-메틸시토신은 변화되지 않는다. 목적의 표적 서열의 증폭은 비설파이트 전환된 DNA에 특이적인 PCR 프라이머를 사용하여 행해지고, 수득된 생성물이 관심의 CpG 부위에서 메틸화된 분석의 템플레이트로서 분리되고 사용된다. 소량의 DNA(예를 들면, 미세절개된 병리학 절편)가 분석되고 CpG 부위에서 메틸화 상황을 결정하기 위해 제한 효소의 이용을 회피한다.

[0174] Ms-SNuPE™ 분석용 전형적인 시약(예를 들면, 전형적인 Ms-SNuPE™ 기반 키트에서 볼 수 있다)은 특이적 유전자좌의 PCR 프로모터(예를 들면, 특이적 유전자, 마커, DMR, 유전자의 영역, 마커의 영역, 비설파이트 처리된 DNA 서열, CpG 섬 등); 최적화된 PCR 버퍼 및 데옥시뉴클레오티드; 젤 추출 키트; 양성 대조 프라이머; 특정 유전자좌의 Ms-SNuPE™ 프라이머; (Ms-SNuPE 반응용) 반응 버퍼; 및 표지된 뉴클레오티드를 포함하지만 이들로 제한되

지 않는다. 또한, 비설파이트 전환 시약은 DNA 변성 베퍼; 설폰화 베퍼; DNA 회수 시약 또는 키트(예를 들면, 침전, 한외여과, 친화성 컬럼); 탈설폰화 베퍼; 및 DNA 회수 성분을 포함할 수 있다.

[0175] RRBS(Reduced Representation Bisulfite Sequencing)는 모든 메틸화되지 않은 시토신을 우라실로 전환하기 위해 핵산의 비설파이트 처리를 시작하고, 제한 효소 소화로(예를 들면, MspI와 같은 CG 서열을 포함하는 부위를 인식하는 효소에 의해) 또는 어댑터 리간드에 커플링 후에 모든 단편의 완전한 서열 결정으로 이어진다. 제한 효소의 선택은 CpG 농후한 영역의 단편을 풍부하게 하고, 분석 중에 다수 유전자 위치에서 맵핑할 수 있도록 풍부한 서열의 수를 감소시킨다. 따라서, RRBS는 서열 결정을 위한 제한 단편의 서브셋을 선택(예를 들면, 분취 젤 전기영동을 사용하여 크기 선택)에 의해 핵산 샘플의 복잡성을 감소시킨다. 전체 게놈 비설파이트 서열 결정과 달리, 제한 효소 소화에 의해서 생성된 모든 단편은 적어도 하나의 CpG 디뉴클레오티드의 DNA 메틸화 정보를 함유한다. 따라서, RRBS는 이러한 영역에서 제한 효소 절단 부위의 높은 빈도로 프로모터, CpG 섬 및 그 외의 게놈 특성에 대해 샘플을 풍부하게 하고, 하나 이상의 게놈 유전자좌의 메틸화 상태를 평가하기 위한 분석을 제공한다.

[0176] RRBS의 전형적인 프로토콜은, MspI와 같은 제한 효소에 의해 핵산 샘플을 소화시키고, 오버헹 및 A-테일링에 충전, 결찰 어댑터, 비설파이트 전환 및 PCR의 단계를 포함한다. 예를 들면, et al. (2005) "Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution" *Nat Methods* 7: 133-6; Meissner et al. (2005) "Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis" *Nucleic Acids Res.* 33: 5868-77 참조

[0177] 일부 실시형태에서, 정량적인 대립유전자 특이적 실시간 표적 및 신호 증폭(QuARTS) 분석을 사용하여 메틸화 상태를 평가한다. 3개의 반응은 연속적으로 제1반응에서 증폭(반응 1) 및 표적 탐침 절단(반응2); 제2 반응에서 FRET 절단 및 형광 신호 생성(반응 3)을 포함하는 각 QuARTS 분석에서 발생한다. 표적 핵산은 특이적 프라이머로 증폭되는 경우, 플랩(flap) 서열의 특이적 검출 탐침은 앰플리콘에 느슨하게 결합한다. 표적 결합 부위에서 특이적 침습성 올리고뉴클레오티드의 존재는 검출 탐침과 플랩 서열 사이의 절단에 의해서 클리바제(cleavase)가 플랩 서열을 방출하게 한다. 플랩 서열은 상응하는 FRET 카세트의 헤어핀 부재 부분에 상보적이다. 따라서, 플랩 서열은 FRET 카세트 상에 침습성 올리고뉴클레오티드로서 기능하고 FRET 카세트 형광체와 쿠엔처 사이의 절단을 일으키고, 형광 신호를 생성한다. 절단 반응은 표적당 다수 개의 탐침을 절단하고 플랩당 다수의 형광체를 방출하고, 지수함수 신호 증폭을 제공한다. QuARTS는 상이한 염료에 의해서 FRET 카세트를 사용하여 단일 반응 웰에서 다수의 표적을 검출할 수 있다. 예를 들면, Zou et al. (2010) "Sensitive quantification of methylated markers with a novel methylation specific technology" *Clin Chem* 56: A199; 미국 특허 출원 제12/946,737, 12/946,745, 12/946,752, 및 61/548,639호 참조.

[0178] "비설파이트 시약"은 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드 서열 사이를 구별하기 위해 본원에 개시된 유용한 비설파이트, 디설파이트, 수소 설파이트, 또는 이들의 조합을 포함하는 시약을 의미한다. 상기 처리 방법은 당해 기술분야에서 공지되어 있다(예를 들면, PCT/EP2004/011715, 전체 참조로 포함되어 있다). 비설파이트 처리는 n-알킬렌글리콜 또는 디에틸렌 글리콜 디메틸에테르(DME)로 제한되지 않지만 이와 같은 변성용매의 존재 또는 디옥산 또는 디옥산 유도체의 존재하에서 수행되는 것이 바람직하다. 일부 실시형태에서, 변성 용매는 1% 내지 35% (v/v) 농도로 사용된다. 일부 실시형태에서, 비설파이트 반응은 크로만 유도체, 예를 들면, 6-하이드록시-2,5,7,8-테트라메틸크로만 2-카르복실산 또는 트리하이드록시벤조산 및 그 유도체, 예를 들면, 갈산(galllic acid)이지만 이들로 제한되지 않는 스캐빈저의 존재하에서 수행된다 (PCT/EP2004/011715 참조, 전체 참조로 포함되어 있다). 비설파이트 전환은 바람직하게 30°C 내지 70°C의 반응온도에서 수행되고, 온도는 반응 중에 약 85°C로 짧은 시간 동안 증가된다(예를 들면, PCT/EP2004/011715, 전체 참조로 포함되어 있다). 비설파이트 처리된 DNA는 바람직하게 정량화 전에 정제된다. 이는 초음파, 예를 들면, Microcon™ columns (Millipore™ 제품)에 의해 당해 기술분야에서 공지된 임의의 수단에 의해서 수행될 수 있다. 이러한 정제는 개질된 제조업자의 프로토콜에 따라 수행된다 (예를 들면, PCT/EP2004/011715, 본원에 참조로 포함된다).

[0179] 일부 실시형태에서, 표적화된 DNA의 단편은 본 발명(예를 들면, 표 3 및/또는 5 참조)에 따라 프라이머 올리고뉴클레오티드 및 증폭 효소의 세트를 사용하여 증폭된다. 몇 개의 DNA 세그먼트의 증폭은 하나 및 동일한 반응 용기에서 동시에 수행될 수 있다. 전형적으로, 증폭은 폴리머라제 연쇄 반응(PCR)을 사용하여 수행된다. 증폭은 길이가 일반적으로 100 내지 2000 개 염기쌍이다.

[0180] 방법의 또 다른 실시형태에서, DMR(예를 들면, 표 1, 2, 4, 및 7에 제공되는 DMR 1-274)을 포함하는 마커 내에 또는 근방에 CpG 위치의 메틸화 상태는 메틸화 특이적 프라이머 올리고뉴클레오티드를 사용하여 검출될 수

있다. 이러한 기술(MSP)는 미국 특허 제 6,265,171호 (Herman)에 기재되어 있다. 비설파이트 처리된 DNA의 증폭을 위한 메틸화 상황 특이적 프라이머의 사용은 메틸화된 핵산 및 메틸화되지 않은 핵산 사이의 구별이 가능하다. MSP 프라이머 쌍은 비설파이트 처리된 CpG 디뉴클레오티드로 하이브리드화하는 적어도 하나의 프라이머를 함유한다. 따라서, 상기 프라이머의 서열은 적어도 하나의 CpG 디뉴클레오티드를 포함한다. 메틸화되지 않은 DNA에 특이적인 MSP 프라이머는 CpG에서 C 위치의 위치에서 "T"를 함유한다.

[0181] 증폭에 의해서 수득된 단편은 검출 가능한 표지를 직접 또는 간접적으로 수행할 수 있다. 일부 실시형태에서, 표지는 형광 표지, 방사성 핵종, 또는 질량분석에서 검출될 수 있는 전형적인 질량을 갖는 검출 가능한 문자 단편이다. 상기 표지가 질량 표지라면, 일부 실시형태는 표지된 앰플리콘의 단일 양성 또는 음성 순 전하를 갖고 질량 분석에서 우수한 검출성이 얻어지는 것을 제공한다. 검출은 예를 들면, 매트릭스 보조 레이저 탈착/이온화 질량분석기(MALDI) 또는 전자 분무 질량 분석(ESI)를 사용하여 수행되고 가시화될 수 있다.

[0182] 이들 분석 기술에 적합한 DNA를 분리하는 방법은 당해 기술분야에서 공지되어 있다. 특히, 일부 실시형태는 미국 특허 출원 제13/470,251호 ("Isolation of Nucleic Acids")에 기재된 핵산의 분리를 포함하고, 본원에 참조로 포함되어 있다.

[0183] 방법

[0184] 일부 실시형태에서, 기술, 방법은 다음의 단계를 포함하는 것이 제공된다.

[0185] 1) 피험체로부터 수득된 핵산(예를 들면, 대변 샘플 또는 위 조직 또는 혈장 샘플 등의 체액으로부터 분리된 게놈 DNA)을, DMR을 포함하는 적어도 하나의 마커(예를 들면, 표 1, 2, 4, 및 7에 제공된 DMR 1-274) 내에서 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약과 접촉하는 단계

[0186] 2) 신생물 또는 증식 장애를 검출하는 단계(예를 들면, 80% 이상의 감도 및 80% 이상의 특이성을 갖도록 제공)

[0187] 일부 실시형태에서, 기술, 방법은 다음의 단계를 포함하는 것이 제공된다.

[0188] 1) 피험체로부터 수득된 핵산(예를 들면, 대변 샘플 또는 위 조직 등의 체액으로부터 분리된 게놈 DNA)을, ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 및 C130RF18로 이루어지는 군으로부터 선택된 주석을 갖는 염색체 영역으로부터 선택된 적어도 하나의 마커 내에서 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약과 접촉하는 단계

[0189] 2) 위암을 검출하는 단계(예를 들면, 80% 이상의 감도 및 80% 이상의 특이성을 갖도록 제공)

[0190] 일부 실시형태에서, 기술, 방법은 다음의 단계를 포함하는 것이 제공된다.

[0191] 1) 피험체로부터 수득된 핵산(예를 들면, 대변 샘플 또는 위 조직 등의 체액으로부터 분리된 게놈 DNA)을, ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 또는 ZNF569로 이루어지는 군으로부터 선택된 주석을 갖는 염색체 영역으로부터 선택된 적어도 하나의 마커 내에서 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약과 접촉하는 단계

[0192] 2) 위암을 검출하는 단계(예를 들면, 80% 이상의 감도 및 80% 이상의 특이성을 갖도록 제공)

[0193] 일부 실시형태에서, 기술, 방법은 다음의 단계를 포함하는 것이 제공된다.

[0194] 1) 피험체로부터 수득된 핵산(예를 들면, 혈장 샘플로부터 분리된 게놈 DNA)을, ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1로 이루어지는 군으로부터 선택된 주석을 갖는 염색체 영역으로부터 선택된 적어도 하나의 마커 내에서 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약과 접촉하는 단계

[0195] 2) 위암을 검출하는 단계(예를 들면, 80% 이상의 감도 및 80% 이상의 특이성을 갖도록 제공)

[0196] 바람직하게, 감도는 약 70% 내지 약 100%, 약 80% 내지 약 90%, 약 80% 내지 약 85%이다. 바람직하게, 특이성은 약 70% 내지 약 100%, 약 80% 내지 약 90%, 또는 약 80% 내지 약 85%이다.

[0197] 게놈 DNA는 시판 키트의 이용을 포함하는 임의의 수단에 의해서 분리될 수 있고, 관심 DNA는 세포막에 의해서 캡슐화되고, 생물학적 샘플은 효소적, 화학적 또는 기계적 수단에 의해서 파괴 및 용해되어야 한다. DNA 용액은 예를 들면, 프로테아나제 K로 소화함으로써 단백질 및 그 외의 오염물을 제거할 수 있다. 게놈 DNA는 용액

으로부터 회수된다. 이는 염석, 유기 추출 또는 고상 지지체에 DNA의 결합을 포함하는 다양한 방법에 의해서 수행될 수 있다. 방법의 선택은 시간, 비용 및 필요한 DNA의 양을 포함하는 몇 개의 인자에 의해 영향을 받는다. 신생물 또는 전-신생물을 포함하는 모든 임상적 샘플 유형은 본 방법에서 사용하기에 적합하고, 예를 들면, 세포주, 조직 슬라이드, 생검, 파라핀 포매 조직, 체액, 대변, 위 조직, 결장 유출물, 소변, 혈장, 혈청, 전혈, 분리된 혈액 세포, 혈액으로부터 분리된 세포 및 이들의 조합이다.

[0198] 본 기술은 샘플을 제조하고 시험에 핵산을 제공하기 위해 사용되는 방법에서 제한되지 않는다. 예를 들면, 일부 실시형태에서, DNA는 미국 특허 출원 제 61/485386호에 기재되어 있는 직접적인 유전자 포획, 또는 관련 방법을 사용하여 대변 샘플, 또는 혈액 또는 혈장 샘플로부터 분리한다.

[0199] 게놈 DNA 샘플은 DMR(예를 들면, 표 1, 2, 4 및 7에 의해 제공된 DMR 1-274)을 포함하는 적어도 하나의 마커 내에 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약으로 처리한다.

[0200] 일부 실시형태에서, 시약은 5'-위치에서 메틸화되지 않은 시토신 염기를 우라실, 티민 또는 하이브리드화 거동의 점에서 시토신과 다른 또 다른 염기로 전환된다. 그러나, 일부 실시형태에서 시약은 메틸화-민감성 제한 효소일 수 있다.

[0201] 일부 실시형태에서, 게놈 DNA 샘플은 5' 위치에서 메틸화되지 않은 시토신 염기가 우라실, 티민, 또는 하이브리드화 거동의 점에서 시토신과 다른 그 외의 염기로 전환되도록 처리된다. 일부 실시형태에서, 이러한 처리는 비설페이트(수소 설파이트, 디설파이트)를 수행한 후 알칼리 가수분해를 수행한다.

[0202] 처리된 핵산은 표적 유전자 서열(DMR, 예를 들면, 표 1, 2, 4, 및 7에 의해 제공되는 DMR 1-274로부터 선택된 적어도 하나의 DMR을 포함하는 마커로부터 적어도 하나의 유전자, 게놈 서열 또는 뉴클레오티드)의 메틸화 상태를 결정하기 위해 분석된다. 분석 방법은 본원에 열거된 방법 예를 들면, 본원에 기재된, QuARTS 및 MSP 등의 해당 기술분야에서 공지된 것으로부터 선택될 수 있다.

[0203] 이상 메틸화(Aberrant methylation), 구체적으로 DMR(예를 들면, 표 1, 2, 4 및 7에 의해 제공된 DMR1-274)를 포함하는 마커의 파메틸화는 위암에 관련되고, 일부 실시형태에서 종양 부위를 예측한다.

[0204] 본 기술은 위암과 관련되는 임의 샘플의 분석에 관련된다. 예를 들면, 일부 실시형태에서 샘플은 환자로부터 수득된 조직 및/또는 생물학적 유체를 포함한다. 일부 실시형태에서, 샘플은 분비물을 포함한다. 일부 실시형태에서, 샘플은 혈액, 혈청, 혈장, 위분비물, 췌장액(pancreatic juice), 위장 생검 샘플, 위장 생검의 미세절개된 세포, 위장 내강으로 떨어진(sloughed) 위장관 세포 및/또는 대변으로부터 회수된 위장관 세포를 포함한다. 일부 실시형태에서, 피험체는 인간이다. 이들 샘플은 상부 위장관, 하부 위장관으로부터 발생할 수 있고, 또는 상부 위장관 및 하부 위장관으로부터의 세포, 조직 및/또는 분비물을 포함할 수 있다. 샘플은 간, 담관, 췌장, 위, 결장, 직장, 식도, 소장, 충수, 십이지장, 용종, 담낭, 항문, 및/또는 복막으로부터의 세포, 분비물, 또는 조직을 포함할 수 있다. 일부 실시형태에서, 샘플은 세포 유체, 복수, 소변, 대변, 췌액(pancreatic fluid), 내시경 검사 중에 수득되는 유체, 혈액, 점액 또는 타액을 포함한다. 일부 실시형태에서 샘플은 대변 샘플이다.

[0205] 이러한 샘플은 당해 기술분야에서 공지된 임의 수단에 의해서 수득될 수 있고, 예를 들면, 당업자에게 명백할 것이다. 예를 들면, 소변 및 대변 샘플은 쉽게 수득될 수 있지만, 혈액, 복수, 혈청 또는 췌액 샘플은 예를 들면, 바늘 및 시린지를 사용하여 비경구로 수득될 수 있다. 세포가 없거나 실질적으로 세포가 없는 샘플은 샘플에 당업자에게 공지된 다양한 기술을 수행하여 수득되고 이는 원심분리 및 여과를 포함하지만 이들로 제한되지 않는다. 일반적으로 샘플을 수득하기 위해 침습적 기술이 사용되지 않는 것이 바람직하지만, 조직 균질물, 조직 철편 및 생검 표본과 같은 샘플을 수득하는 것이 여전히 바람직할 수 있다.

[0206] 일부 실시형태에서, 기술은 환자(예를 들면, 위암(예를 들면, 위암)에 걸린 환자, 조기 단계 위암에 걸린 환자, 또는 위암을 발현할 수 있는 환자)를 치료하기 위한 방법에 관한 것으로, 상기 방법은 본원에 제공된 하나 이상의 DMR의 메틸화 상태를 결정하는 단계 및 메틸화 상태를 결정한 결과에 기초해서 환자에 치료제를 투여하는 단계를 포함한다. 치료는 약학적 화합물, 백신, 수술 수행, 환자 촬영, 또 다른 시험 수행이 제공될 수 있다. 바람직하게, 상기 사용은 임상 스크리닝 방법, 예후 평가 방법, 치료결과를 모니터링하는 방법, 특정 치료제 치료에 반응할 가능성이 높은 환자를 확인하기 위한 방법, 환자 또는 피험체의 촬영 방법, 및 약물 스크리닝 및 개발 방법에서 행해진다.

[0207] 기술의 일부 실시형태에서, 피험체에서 위암(예를 들면, 위암)을 진단하는 방법이 제공된다. 본원에서 사용되

는 "진단하는" 및 "진단" 용어는 피험체가 소정의 질병 또는 질환에 걸리거나 향후 소정의 질병 또는 질환을 발현하는 것을 당업자가 추정하고 결정하는 방법을 의미한다. 당업자는 종종 하나 이상의 진단 지표 예를 들면, 바이오마커(예를 들면, 본원에 기재된 DMR)에 기초해서 진단하고, 메틸화 상태는 질환의 존재, 중증도 또는 부재를 나타낸다.

[0208] 진단과 함께, 임상적 암 예후는 가장 효과적인 치료법을 계획하기 위해 종양 재발의 가능성 및 암의 공격성을 결정하는 것에 관련된다. 추가의 정확한 예후가 이루어지고 암을 발현할 잠재적이 위험이 평가될 수 있으면, 적절한 치료, 경우에 따라 덜 심한 환자의 치료법이 선택될 수 있다. 암 바이오마커의 평가(예를 들면, 메틸화 상태를 결정)는 예후가 좋고 및/또는 암 발병 위험이 낮은 피험체를 분리하는 데 유용하고 추가의 집중적 치료에 유리한 암이 재발하거나 암을 발현할 가능성이 있는 피험체로부터 치료 또는 제한된 치료가 필요하지 않는다.

[0209] 따라서, 본원에서 사용되는 "진단하는 것"은 암을 발현할 위험을 결정하거나 예후를 결정하는 것을 포함하고, 이는 임상적 결과를 예측하고(약물 치료 유무에 관계없이), 적절한 치료를 선택하고(또는 치료가 효과적인 것에 관계없이), 또는 본원에 기재된 진단 바이오마커(예를 들면, DMR)의 측정에 기초해서 현재의 치료를 모니터링하고 잠재적으로 변화하는 것을 제공할 수 있다. 또한, 개시된 주제의 일부 실시형태에서 시간 경과에 따라 바이오마커의 다수의 결정은 진단 및/또는 예후를 촉진할 수 있다. 바이오마커의 일시적인 변화를 사용하여 임상적 결과를 예측하고, 위암의 진행을 모니터링하고, 및/또는 암에 대한 관련된 적절한 치료 효능을 모니터링하기 위해서 사용될 수 있다. 이러한 실시형태에서 효과적인 치료 과정 중에 시간 경과에 따라 예를 들면, 본원에 기재된 하나 이상의 바이오 마커(DMR)(및 모니터링되는 경우 잠재적으로 하나 이상의 추가의 바이오마커)의 메틸화 상태의 변화를 볼 수 있는 것으로 예상된다

[0210] 본원에 개시된 주제는 또한 일부 실시형태에서 피험체 내의 암의 치료 또는 예후를 개시하거나 지속하는 것을 결정하는 방법을 더 제공한다. 일부 실시형태에서, 방법은 피험체로부터 시간 경과에 따라 일련의 생물학적 샘플을 제공하고; 각각의 생물학적 샘플에서 본원에 기재된 적어도 하나의 바이오마커의 메틸화 상태를 결정하기 위해 일련의 생물학적 샘플을 분석하고; 각각의 생물학적 샘플에서 하나 이상의 바이오마커의 메틸화 상태의 임의의 측정 가능한 변화를 비교하는 것을 포함한다. 시간 경과에 따라 바이오마커의 메틸화 상태의 임의의 변화는 암을 발병할 위험을 예측하고, 임상적 결과를 예측하고, 암의 치료 또는 예후를 개시하거나 지속하는 것을 결정하고, 현재의 치료가 암을 효과적으로 치료하는지를 결정하기 위해 사용될 수 있다. 예를 들면, 제1 시점은 치료의 개시 전에 선택될 수 있고 제2 시점은 치료 개시 후 어느 시점에 선택될 수 있다. 메틸화 상태는 상이한 시점으로부터 취한 각각의 샘플에서 측정될 수 있고, 정성적 및 정량적 차이를 나타낸다. 상이한 샘플로부터 바이오마커 수준의 메틸화 상태의 변화는 위암 위험, 예후, 치료 효능 결정 및/또는 피험체 내의 암의 진행과 관련될 수 있다.

[0211] 바람직한 실시형태에서, 본 발명의 방법 및 조성물은 조기 단계에 예를 들면, 질병의 증상이 나타나기 전에 질병의 치료 또는 진단을 위한 것이다. 일부 실시형태에서, 본 발명의 방법 및 조성물은 임상적 단계에 질병의 치료 또는 진단을 위한 것이다.

[0212] 언급한 바와 같이, 일부 실시형태에서, 하나 이상의 진단 또는 예후 바이오마커의 다수 결정이 행해질 수 있고, 진단 또는 예후를 결정하기 위해 이러한 마커의 일시적 변화가 사용될 수 있다. 예를 들면, 진단 마커는 제1시기에 및 다시 제2 시기에 결정될 수 있다. 이러한 실시형태에서, 제1시기로부터 제2시기로 마커의 증가는 암의 특정한 유형 또는 중증도의 진단 또는 소정의 예후일 수 있다. 마찬가지로, 제1시기로부터 제2시기로 마커의 감소는 암의 특정한 유형 또는 중증도 또는 소정의 예후를 나타낼 수 있다. 또한, 하나 이상의 마커의 변화 정도는 암의 중증도 및 향후 악영향에 관련될 수 있다. 당업자는, 특정한 실시형태에서 다수 시점에서 동일한 바이오마커의 비교 측정이 행해질 수 있고 하나의 시점에 소정의 바이오마커 및 제2시점에 제2바이오마커를 측정할 수 있고 이들 마커의 비교는 진단 정보를 제공할 수 있다.

[0213] 본원에서 사용되는 "진단을 결정하는 것" 어구는 피험체 내의 질환의 결과 또는 과정을 예측할 수 있다. "예후" 용어는 100% 정확성의 결과 또는 과정을 예측하는 능력을 의미하는 것은 아니고 소정의 과정 또는 결과가 바이오마커(예를 들면, DMR)의 메틸화에 기초해서 발생할 다소의 가능성이 있는 것을 의미한다. 실제, 당업자는 "예후" 용어가 특정한 과정 또는 결과가 발생할 확률이 증가하고, 과정 또는 결과가 소정의 질환을 나타내는 피험체에서 발생할 가능성이, 질환을 나타내지 않는 개체의 것에 비해 큰 것을 나타낸다. 예를 들면, 질환을 나타내지 않는 개체(예를 들면, 하나 이상의 DMR의 정상적인 메틸화 상태를 갖는 것)에서, 소정의 결과(예를 들면, 위암에 걸릴) 가능성이 매우 낮을 수 있다.

- [0214] 일부 실시형태에서, 통계적 분석은 예후 지표와 불리한 결과에 대한 경향을 관련시킨다. 예를 들면, 일부 실시 형태에서, 암에 걸리지 않은 환자로부터 수득된 일반적인 대조군 샘플과 상이한 메틸화 상태는, 피험체가 통계 유의도 수준에 의해서 결정된 대조군 샘플에서 메틸화 상태와 유사한 수준을 갖는 피험체보다 암에 걸릴 가능성이 높은 것을 나타낼 수 있다. 또한, 베이스라인(예를 들면, "정상") 수준로부터 메틸화 상태의 변화는 피험체 예후를 반영할 수 있고, 메틸화 상태의 변화 정도는 유해한 사상의 중증도에 관련될 수 있다. 통계적 유의도는 종종 2개 이상의 집단을 비교함으로써 결정되고 신뢰구간 및/또는 p값을 결정된다. 예를 들면, Dowdy and Wearden, *Statistics for Research*, John Wiley & Sons, New York, 1983, 참조, 본원에 참조로 포함됨. 본 주제의 예시의 신뢰 구간은 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.9% 및 99.99%이고, 예시의 p값은 0.1, 0.05, 0.025, 0.02, 0.01, 0.005, 0.001, 및 0.0001이다.
- [0215] 그 외의 실시형태에서, 본원에 기재된 예후 또는 진단 바이오마커(예를 들면, DMR)의 메틸화 상태의 문턱값 변화 정도가 설립될 수 있고, 생물학적 샘플에서 바이오마커의 메틸화 상태의 변화 정도는 단순히 메틸화 상태에서의 문턱값 변화 정도와 비교한다. 본원에 제공된 바이오마커의 메틸화 상태에서 바람직한 문턱값 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 실시형태에서 "노모그램"이 설립되고, 진단 또는 예후 지표(바이오마커 또는 바이오마커의 조합)의 메틸화 상태는 소정의 결과에 관련된 배치에 직접 관련된다. 당업자는 이러한 노모그램의 사용에 의해 측정에서 불확실성이 마커 농도의 불확성과 동일한 것을 이해하고, 개개의 샘플 측정은 집단 평균이 아니라 참조로 된다.
- [0216] 일부 실시형태에서, 대조군 샘플은 생물학적 샘플로부터 수득된 결과를 대조군 샘플로부터 수득된 결과와 비교할 수 있도록 생물학적 샘플과 동시에 분석된다. 또한, 표준 곡선이 제공되고, 생물학적 샘플에 대한 분석 결과를 비교할 수 있는 것으로 고려될 수 있다. 이러한 표준 곡선은 형광 표지가 사용되는 경우 분석 유닛의 기능, 예를 들면, 형광 신호 강도로서 바이오마커의 메틸화 상태를 나타낸다. 다수 도너로부터 취한 샘플을 사용하면, 표준 곡선이, 위암에 걸린 도너로부터 또는 화생(metaplasia)을 갖는 도너로부터 취한 조직에서 하나 이상의 바이오마커의 "위험" 표지뿐 아니라 정상조직에서 하나 이상의 바이오마커의 대조군 메틸화 상태를 제공할 수 있다. 방법의 특정한 실시형태에서, 피험체는 피험체로부터 수득된 생물학적 샘플에서 본원에 제공된 하나 이상의 DMR의 이상 메틸화 상태를 확인할 때에 화생을 갖는 것으로 확인된다. 방법의 그 외의 실시형태에서, 피험체로부터 수득된 생물학적 샘플에서 이러한 하나 이상의 바이오마커의 이상의 메틸화 상태의 검출은, 피험체가 암에 걸린 것으로 확인된다.
- [0217] 마커의 분석은 하나의 시험 샘플 내에서 추가의 마커와 별도로 또는 동시에 수행될 수 있다. 예를 들면, 몇 개의 마커는 다수 샘플의 효율적인 처리 및 잠재적으로 제공하는 큰 진단 및/또는 예후 정확성을 제공하기 위한 시험에 결합될 수 있다. 또한, 당업자는 동일한 샘플 피험체로부터 다수의 샘플(예를 들면, 연속적인 시점)을 시험하는 값을 인식한다. 일련의 샘플의 이러한 시험은 시간 경과에 따라 마커 메틸화 상태의 변화를 확인할 수 있다. 메틸화 상태 변화 및 메틸화 상태 변화 부재는 질병 상태에 대한 유용한 정보를 제공할 수 있고, 이는 사상의 개시로부터 대략적인 시간, 구제가능한 조직의 존재 및 양, 약물 치료의 타당성, 다양한 치료의 유효성 및 장래의 사상의 위험을 포함하는 피험체의 결과의 확인을 포함하지만, 이들로 제한되지 않는다.
- [0218] 바이오마커의 분석은 다양한 물리적 포맷으로 수행될 수 있다. 예를 들면, 마이크로타이터 플레이트 또는 자동화는 다수 시험 샘플의 처리를 용이하게 하기 위해 사용될 수 있다. 또한, 단일 샘플 포맷은 외래 운송이나 응급실 설정과 같은 적절한 방식으로 즉각적인 처리 및 진단을 용이하게 하기 위한 단일 샘플 포맷을 개발할 수 있다.
- [0219] 일부 실시형태에서 대조군 메틸화 상태와 비교하여 샘플에서 적어도 하나의 바이오마커의 메틸화 상태의 측정 가능한 차이가 있는 경우, 피험체는 위암에 걸린 것으로 진단된다. 반면, 생물학적 샘플에서 메틸화 상태의 변화가 확인되지 않는 경우, 피험체는 위암에 걸리지 않는 것 또는 위암에 걸릴 위험이 없는 것으로 확인될 수 있거나 암에 걸릴 위험이 낮은 것으로 확인될 수 있다. 이점에서, 암에 걸리거나 암에 걸릴 위험이 있는 피험체는 암 또는 그 위험이 실질적으로 없거나 낮은 피험체와 구별될 수 있다. 위암을 발현할 위험이 있는 피험체는 내시경 감시 등의 집중 및/또는 정기적인 스크리닝 일정에 놓일 수 있다. 한편, 암 위험이 낮거나 실질적으로 없는 피험체는, 향후 스크리닝, 예를 들면, 본 기술에 따라 행해진 스크리닝에서, 위암의 위험이 피험체에서 나타날 때까지 내시경을 수행하는 것을 피할 수 있다.
- [0220] 상술한 바와 같이, 본 기술의 방법의 실시형태에 따라, 하나 이상의 바이오마커의 메틸화 상태 변화를 검출하는 것은 정성적 결정일 수 있거나 정량적 결정일 수 있다. 따라서, 위암(예를 들면, 위암)에 걸리거나 발현 위험이 있는 피험체를 진단하는 단계는, 특정한 문턱값 측정이 행해지고, 예를 들면, 생물학적 샘플에서 하나 이상

의 바이오마커의 메틸화 상태가 소정의 대조군 메틸화 상태로부터 변화하는 것을 나타낸다. 방법의 일부 실시 형태에서, 대조군 메틸화 상태는 바이오마커의 검출 가능한 메틸화 상태이다. 대조군 샘플이 생물학적 샘플과 동시에 시험되는 방법의 그 외의 실시형태에서, 소정의 메틸화 상태는 대조군 샘플에서 메틸화 상태이다. 방법의 그 외의 실시형태에서, 소정의 메틸화 상태는 표준 곡선에 기초하고 및/또는 표준 곡선에 의해서 확인된다. 방법의 그 외의 실시형태에서, 소정의 메틸화 상태는 구체적으로 상태 또는 상태의 범위이다. 따라서, 소정의 메틸화 상태는 실시되는 방법 및 소망의 특이성 등의 실시형태에 부분적으로 기초해서 당업자에게 명백하게 되는 허용 가능한 한도 내에서 선택될 수 있다.

[0221] 또한, 진단 방법에 관련해서, 바람직한 피험체는 척추동물 피험체이다. 바람직한 척추동물은 온혈동물이고, 바람직한 온혈 척추동물은 포유류이다. 바람직한 포유류는 가장 바람직하게 인간이다. 본원에 사용된 바와 같이, "피험체"는 인간 및 동물 피험체를 포함한다. 따라서, 수의학적 치료 용도가 본원에서 제공된다. 따라서, 본 기술은 인간과 같은 포유동물뿐 아니라 시베리아 호랑이와 같이 멸종위기로 인해 중요한 포유동물; 인간 소비를 위해 농장에서 사육되는 동물과 같은 경제적으로 중요한 포유동물; 및/또는 애완동물 또는 동물원에서 유지된 동물과 같은 인간에게 사회적으로 중요한 동물의 진단을 위해 제공한다. 이러한 동물의 예는 고양이 및 개 등의 육식동물; 피그, 호그(hog) 및 야생 맷돼지 등의 돼지; 소, 옥스, 양, 기린, 사슴, 염소, 들소, 및 낙타 등의 반추동물 및/또는 유제류; 및 말을 포함하지만 이들로 제한되지 않는다. 따라서, 가축의 진단 및 치료가 제공되고, 가축 돼지, 반추동물, 유제류, 말(경마 말 포함) 등을 포함하지만 이들로 제한되지 않는다. 현재 개시된 주제는 피험체에서 위암(예를 들면, 위암)을 진단하기 위한 시스템을 더 포함한다. 시스템은 예를 들면, 생물학적 샘플이 수집된 피험체에서 위암을 진단하거나 위암 위험을 스크리닝하기 위해 사용될 수 있는 시판 키트로서 제공될 수 있다. 본 기술에 따라 제공된 예시의 시스템은 표 1, 2, 4 및 7에서 제공된 DMR의 메틸화 상태를 평가하는 것을 포함한다.

[0222] 실시예

[0223] 실시예 1 - 위암의 검출을 위한 마커를 확인

[0224] 본 기술의 실시형태를 전개하는 과정에서 수행된 실험은, 위암에 대한 248 개 DMR(예를 들면, DMR Nos. 1-248; 표 1)에 상응하는 123개 DNA 메틸화 마커를 확인했다. 이러한 DNA 메틸화 마커는 중례-대조군 조직 샘플 세트의 대량 병렬 서열 결정에 결합된 CpG 섬 풍부에 의해서 생성된 데이터로부터 확인되었다. 대조군은 정상 위 조직, 정상 결장 상피세포 및 정상 백혈구 유래 DNA를 포함했다. 본 기술은 RRBS(reduced representation bisulfite sequencing)을 이용했다. 3단계는 모든 비공식적인 부위를 제외하고 커버리지 컷오프의 점에서 데이터를 해석하고, 논리적 회귀를 사용하여 진단 하위 그룹 사이의 메틸화백분율을 대조하고, 인접한 메틸화 부위의 정의된 그룹화 및 수용자 동작 특성 분석을 기반으로 하는 인실리코 CpG 섬을 형성한다. 4단계는 메틸화 데이터%를 여과하고(개별 CpG 및 인실리코 클러스터) 신호 대 잡음 비를 최대화하는 마커를 선택하고, 백그라운드 메틸화를 최소화하여 종양 이질성을 설명하고 ROC 성능을 강조한다. 또한, 이러한 분석적 접근방법은 다운스트림 마커 분석-메틸화 특이적 PCR, 작은 단편 깊은 서열 결정의 용이하고 최적의 설계를 위해 정의된 DNA 길이에 대해 Hotspot CpGs의 확인을 보장한다.

[0225] 각각의 샘플은 대략 2~3백만개의 높은 품질의 CpGs를 생성하고, 분석 및 여과 후 1000개 미만으로 수득되고 개별 부위가 매우 구별된다. 이들은 차별화된 메틸화의 123 개의 국제화된 영역으로 클러스터링되고, 일부는 30-40개의 염기를 신장하고 일부는 킬로베이스(kilobase)를 초과한다(표 1 참조). 12개의 DMR이 확인되고 문헌에서는 주석이 없고 어떤 언급도 없다. 이러한 DMR은 MAX로 명명되고, 염색체 위치 및 좌표가 나타났다(표 1 참조). 모든 DMRs은 0.8 이상의 AUC를 갖는다(일부는 AUC 1.0인 대조군으로부터 중례와 완벽하게 구별되는 것을 입증했다). 0.8의 AUC 문턱값을 충족하는 것 이외에, 확인된 위암 마커는 정상 위 또는 결장 점막에서보다 종양에서 20배 초과의 메틸화 밀도를 나타내고 비신생물 위 또는 결장 점막에서 1.0 %미만의 메틸화를 나타내어야 한다.

[0226]

[표 1]

| DMR No. | 유전자 | 염색체 | 섬 범위 | 가닥 |
|------------|----------|-----|---------------------|----|
| 1 | ABCB1 | 7 | 87229775-87229856 | - |
| 2 | ADAM19 | 5 | 157002684-157002766 | - |
| 3 | ADAM23 | 2 | 207307964-207308190 | + |
| 4 | ADAM23 | 2 | 207308771-207308890 | + |
| 5 | ADCY1 | 7 | 45613877-45613977 | + |
| 6 | ADRA1B | 5 | 159343458-159343934 | + |
| 7 | AFF3 | 2 | 100721738-100721967 | - |
| 8 | AGBL4 | 1 | 49242089-49242514 | - |
| 9 | ALOX5 | 10 | 45914474-45914484 | + |
| 10 | ANK1 | 8 | 41754327-41754528 | - |
| 11 | ANKRD34B | 5 | 79866019-79866185 | - |
| 12 | ANO8 | 19 | 17439445-17439845 | - |
| 13 | APOBEC3D | 22 | 39417451-39417576 | + |
| 14 | ARHGEF7 | 13 | 111767862-111767954 | + |
| 15 | BACH2 | 6 | 91004976-91005235 | - |
| 16 | BARX1 | 9 | 96715112-96715262 | - |
| 17 | BARX1 | 9 | 96715374-96715603 | - |
| 18 | BARX1 | 9 | 96721501-96721818 | - |
| 19 | BCL2 | 18 | 60985561-60985600 | - |
| 20 | BCL2 | 18 | 60985788-60985828 | - |
| 21 | BHLHE22 | 8 | 65493937-65494105 | + |
| 22 | BMP3 | 4 | 81952135-81952303 | + |
| 23 | BMP3 | 4 | 81952407-81952528 | + |
| 24 | BMP6 | 6 | 7727566-7727907 | + |
| 25 | BMP8A | 1 | 39980533-39980619 | + |
| 26 | C12orf39 | 12 | 21680721-21680828 | + |
| 27 | C12orf42 | 12 | 103889269-103889591 | - |
| 28 | C17orf46 | 17 | 43339242-43339498 | - |
| 29 | C1orf70 | 1 | 1475560-1475650 | - |
| 30 | C1orf70 | 1 | 1476065-1476127 | - |
| 31 | C1orf95 | 1 | 226737123-226737246 | + |
| 32 | C1QL3 | 10 | 16563608-16563892 | - |
| 33 | C9orf50 | 9 | 132382706-132382909 | - |
| 34 | CABIN1 | 22 | 24552655-24552724 | + |
| 35 | CACNA1C | 12 | 2162358-2162804 | + |
| 36 | CACNA1H | 16 | 1202428-1202560 | + |
| 37 | CBFA2T3 | 16 | 89007265-89007441 | - |
| 38 | CBLN1 | 16 | 49315385-49315725 | - |
| 39 | CBLN1 | 16 | 49316162-49316268 | - |
| 40 | CCDC106 | 19 | 56159280-56159589 | + |
| 41 | CCDC42B | 12 | 113592232-113592553 | + |
| 42 | CCND2 | 12 | 4378158-4378413 | + |
| 43 | CD1D | 1 | 158150624-158151205 | + |
| 44 | CDHR1 | 10 | 85954203-85954476 | + |
| 45 | CELF2 | 10 | 11059947-11060085 | + |
| 46 | CHST2 | 3 | 142838645-142839023 | + |
| 47 | CHST2 | 3 | 142839223-142839495 | + |
| 48 | CHST2 | 3 | 142839761-142840276 | + |
| 49 | CLEC11A | 19 | 51228264-51228394 | + |
| 50 | CLEC11A | 19 | 51228512-51228732 | + |
| 51 | CLSTN2 | 3 | 139654045-139654299 | + |
| 52 | CNR1 | 6 | 88876414-88876452 | - |
| 53 | COMP | 19 | 18901578-18901783 | - |

[0227]

| | | | | |
|-----|---------|----|---------------------|---|
| 54 | CPE | 4 | 166300108-166300239 | + |
| 55 | CPNE8 | 12 | 39299231-39299461 | - |
| 56 | DAB2IP | 9 | 124461296-124461420 | + |
| 57 | DBX1 | 11 | 20177747-20178529 | - |
| 58 | DOCK10 | 2 | 225906664-225906922 | - |
| 59 | DOCK10 | 2 | 225907226-225907322 | - |
| 60 | DOCK10 | 2 | 225907515-225907632 | - |
| 61 | DOK1 | 2 | 74782380-74782580 | + |
| 62 | DYNLRB2 | 16 | 80574542-80574687 | + |
| 63 | DYNLRB2 | 16 | 80574797-80574927 | + |
| 64 | EFNB3 | 17 | 7608691-7608815 | + |
| 65 | EIF4E3 | 3 | 71803220-71804279 | - |
| 66 | ELMO1 | 7 | 37487867-37488498 | - |
| 67 | ELOVL2 | 6 | 11044395-11044704 | - |
| 68 | ELOVL5 | 6 | 53213597-53214212 | - |
| 69 | EPB41L3 | 18 | 5542876-5542973 | - |
| 70 | EYA4 | 6 | 133562791-133562878 | + |
| 71 | EYA4 | 6 | 133563343-133563397 | + |
| 72 | FAIM2 | 12 | 50297863-50297988 | - |
| 73 | FAM78A | 9 | 134152466-134152692 | - |
| 74 | FBN2 | 5 | 127874257-127874290 | - |
| 75 | FER1L4 | 20 | 34189084-34189185 | - |
| 76 | FER1L4 | 20 | 34189488-34189566 | - |
| 77 | FGF2 | 4 | 123748398-123748881 | + |
| 78 | FLT3 | 13 | 28674537-28674770 | - |
| 79 | FMN2 | 1 | 240255171-240255264 | + |
| 80 | FMNL1 | 17 | 43298726-43298819 | + |
| 81 | FOXD4L1 | 2 | 114256519-114257141 | + |
| 82 | FRMD4A | 10 | 13933638-13934256 | - |
| 83 | GJC1 | 17 | 42907705-42907827 | - |
| 84 | GLB1L3 | 11 | 134146132-134146380 | + |
| 85 | GNG7 | 19 | 2702502-2702568 | - |
| 86 | GPR135 | 14 | 59930725-59931142 | - |
| 87 | GPX7 | 1 | 53067933-53068182 | + |
| 88 | GRASP | 12 | 52400819-52401166 | + |
| 89 | GRM8 | 7 | 126891909-126892517 | - |
| 90 | GSG1L | 16 | 28074472-28074761 | - |
| 91 | GSG1L | 16 | 28075145-28075290 | - |
| 92 | GUCY1A2 | 11 | 106888403-106888538 | - |
| 93 | HECW2 | 2 | 197457425-197458011 | - |
| 94 | HOPX | 4 | 57521377-57521658 | - |
| 95 | HOPX | 4 | 57521822-57522182 | - |
| 96 | HOXD8 | 2 | 176994098-176994972 | + |
| 97 | IKZF1 | 7 | 50343838-50344453 | + |
| 98 | INSM1 | 20 | 20348122-20348415 | + |
| 99 | IRF4 | 6 | 391980-392182 | + |
| 100 | ITGA4 | 2 | 182321830-182321983 | + |
| 101 | KCNK12 | 2 | 47797187-47797452 | - |
| 102 | KCNMA1 | 10 | 79397844-79397945 | - |
| 103 | KCNQ3 | 8 | 133492336-133493033 | - |

| | | | | |
|-----|-------------------------------|----|---------------------|---|
| 104 | KDM2B | 12 | 122016688-122017150 | - |
| 105 | KIF5C | 2 | 149633046-149633131 | + |
| 106 | KLHL29 | 2 | 23609989-23610119 | + |
| 107 | LBH | 2 | 30455594-30455655 | + |
| 108 | LINGO3 | 19 | 2290273-2290393 | - |
| 109 | LIPJ | 10 | 90342642-90342873 | + |
| 110 | LOC285548 | 4 | 13549245-13549556 | - |
| 111 | LOC348840 | 3 | 197807208-197807313 | - |
| 112 | LOC440461 | 17 | 66195303-66195773 | + |
| 113 | LOC645323 | 5 | 87970260-87970894 | - |
| 114 | LRRC10B | 11 | 61276493-61276910 | + |
| 115 | LRRFIP1 | 2 | 238600049-238600078 | + |
| 116 | LY86-AS1 | 6 | 6546450-6546598 | - |
| 117 | MARK1 | 1 | 220702225-220702345 | + |
| 118 | MATK | 19 | 3785828-3786371 | - |
| 119 | MAX.chr1.110626388-110626787 | 1 | 110626388-110626787 | - |
| 120 | MAX.chr1.244012787-244012875 | 1 | 244012787-244012875 | - |
| 121 | MAX.chr10.22624234-22624571 | 10 | 22624234-22624571 | - |
| 122 | MAX.chr11.120435350-120435981 | 11 | 120435350-120435981 | - |
| 123 | MAX.chr11.123301058-123301255 | 11 | 123301058-123301255 | - |
| 124 | MAX.chr11.123301366-123301506 | 11 | 123301366-123301506 | - |
| 125 | MAX.chr12.133484966-133485857 | 12 | 133484966-133485857 | - |
| 126 | MAX.chr12.4273880-4274012 | 12 | 4273880-4274012 | - |
| 127 | MAX.chr14.37116653-37117410 | 14 | 37116653-37117410 | - |
| 128 | MAX.chr15.40616377-40616419 | 15 | 40616377-40616419 | - |
| 129 | MAX.chr15.89942727-89942918 | 15 | 89942727-89942918 | - |
| 130 | MAX.chr18.5630121-5630725 | 18 | 5630121-5630725 | - |
| 131 | MAX.chr19.22034799-22034887 | 19 | 22034799-22034887 | - |
| 132 | MAX.chr2.118981853-118981946 | 2 | 118981853-118981946 | - |
| 133 | MAX.chr2.119067940-119068174 | 2 | 119067940-119068174 | - |
| 134 | MAX.chr2.177503038-177503122 | 2 | 177503038-177503122 | - |
| 135 | MAX.chr20.1784209-1784461 | 20 | 1784209-1784461 | - |
| 136 | MAX.chr3.128336893-128337084 | 3 | 128336893-128337084 | - |
| 137 | MAX.chr4.113626838-113626942 | 4 | 113626838-113626942 | - |
| 138 | MAX.chr4.113627402-113627507 | 4 | 113627402-113627507 | - |
| 139 | MAX.chr5.42951300-42951769 | 5 | 42951300-42951769 | - |
| 140 | MAX.chr5.42992128-42992212 | 5 | 42992128-42992212 | - |
| 141 | MAX.chr5.42992322-42992768 | 5 | 42992322-42992768 | - |
| 142 | MAX.chr5.42995102-42995181 | 5 | 42995102-42995181 | - |
| 143 | MAX.chr7.121950350-121951056 | 7 | 121950350-121951056 | - |
| 144 | MAX.chr7.25896389-25896501 | 7 | 25896389-25896501 | - |

| | | | | |
|-----|----------------------------|----|---------------------|---|
| 145 | MAX.chr9.113945-114435 | 9 | 113945-114435 | - |
| 146 | MAX.chr9.79627082-79627175 | 9 | 79627082-79627175 | - |
| 147 | MAX.chr9.99983730-99984118 | 9 | 99983730-99984118 | - |
| 148 | MEF2C | 5 | 88185490-88185576 | - |
| 149 | MIR129-2 | 11 | 43601090-43601396 | + |
| 150 | MMP23B | 1 | 1567269-1567478 | + |
| 151 | MPZ | 1 | 161275561-161275653 | - |
| 152 | MPZ | 1 | 161275766-161275996 | - |
| 153 | MZB1 | 5 | 138728509-138729109 | - |
| 154 | NDRG4 | 16 | 58497251-58497370 | + |
| 155 | NDRG4 | 16 | 58497979-58498250 | + |
| 156 | NEU1 | 6 | 31830234-31830331 | - |
| 157 | NEU1 | 6 | 31830456-31830522 | - |
| 158 | NEU1 | 6 | 31830682-31830917 | - |
| 159 | NEUROG2 | 4 | 113437639-113437953 | - |
| 160 | NFATC4 | 14 | 24837540-24838009 | + |
| 161 | NID2 | 14 | 52535758-52536437 | - |
| 162 | NRXN2 | 11 | 64480878-64481067 | - |
| 163 | OPLAH | 8 | 145106742-145106921 | - |
| 164 | OSR2 | 8 | 99952116-99952427 | + |
| 165 | OSR2 | 8 | 99952591-99953023 | + |
| 166 | PAX5 | 9 | 37027013-37027670 | - |
| 167 | PDE8B | 5 | 76506245-76506460 | + |
| 168 | PEBP4 | 8 | 22722910-22722994 | - |
| 169 | PHF21B | 22 | 45405722-45405819 | - |
| 170 | PLEKHO1 | 1 | 150122907-150123157 | + |
| 171 | PLXNC1 | 12 | 94543138-94543232 | + |
| 172 | PNMAL2 | 19 | 46996516-46996841 | - |
| 173 | POMC | 2 | 25391168-25391616 | - |
| 174 | POU3F1 | 1 | 38510915-38511213 | - |
| 175 | POU4F1 | 13 | 79176470-79176567 | - |
| 176 | POU4F1 | 13 | 79176954-79177042 | - |
| 177 | POU4F1 | 13 | 79177364-79177872 | - |
| 178 | PPP2R5C | 14 | 102247537-102247929 | + |
| 179 | PRDM13 | 6 | 100051231-100051668 | + |
| 180 | PRKCB | 16 | 23847288-23847339 | + |
| 181 | PRKCB | 16 | 23847575-23847699 | + |
| 182 | PRKCB | 16 | 23847825-23848020 | + |
| 183 | PRR16 | 5 | 119799921-119800390 | + |
| 184 | PRR5L | 11 | 36399221-36399473 | + |
| 185 | PYGL | 14 | 51410872-51411219 | - |
| 186 | RASA3 | 13 | 114897059-114897135 | - |
| 187 | RASSF2 | 20 | 4803267-4803703 | - |
| 188 | RBPMS2 | 15 | 65067593-65067619 | - |
| 189 | RNF165 | 18 | 43913775-43914353 | + |
| 190 | RPRM | 2 | 154334604-154335314 | - |
| 191 | RPSAP58 | 19 | 23941384-23941670 | + |
| 192 | RSPO3 | 6 | 127440413-127441057 | + |
| 193 | RUNX3 | 1 | 25256650-25257322 | - |
| 194 | SEPTIN9 | 17 | 75368800-75369056 | + |

[0230]

| | | | | |
|-----|---------|----|---------------------|---|
| 195 | SEPTIN9 | 17 | 75369224-75369327 | + |
| 196 | SFMBT2 | 10 | 7450406-7450831 | - |
| 197 | SFMBT2 | 10 | 7451097-7451185 | - |
| 198 | SFMBT2 | 10 | 7452358-7452452 | - |
| 199 | SFMBT2 | 10 | 7452746-7452779 | - |
| 200 | SFMBT2 | 10 | 7453455-7453511 | - |
| 201 | SLC35F1 | 6 | 118228394-118228979 | + |
| 202 | SLCO4C1 | 5 | 101631546-101631741 | - |
| 203 | SLCO4C1 | 5 | 101632152-101632237 | - |
| 204 | SND1 | 7 | 127671993-127672310 | + |
| 205 | SNX32 | 11 | 65601248-65601514 | + |
| 206 | SP9 | 2 | 175201823-175202389 | + |
| 207 | ST8SIA4 | 5 | 100236601-100236784 | - |
| 208 | ST8SIA6 | 10 | 17496306-17496417 | - |
| 209 | ST8SIA6 | 10 | 17496572-17496711 | - |
| 210 | TBC1D30 | 12 | 65218335-65218778 | + |
| 211 | TBC1D30 | 12 | 65220152-65220299 | + |
| 212 | TCF24 | 8 | 67874670-67875083 | - |
| 213 | TIAM1 | 21 | 32930248-32930322 | - |
| 214 | TLX2 | 2 | 74742033-74742264 | + |
| 215 | TM6SF1 | 15 | 83776207-83776706 | + |
| 216 | TRIM9 | 14 | 51560661-51561240 | - |
| 217 | TRPC3 | 4 | 122872067-122872455 | - |
| 218 | TRPC3 | 4 | 122872626-122872724 | - |
| 219 | TSC22D4 | 7 | 100075304-100075445 | - |
| 220 | TTC28 | 22 | 29075259-29075493 | - |
| 221 | UNC5CL | 6 | 40995803-40996093 | - |
| 222 | VCAN | 5 | 82768837-82769031 | + |
| 223 | VEGFC | 4 | 177713711-177713910 | - |
| 224 | WNT5A | 3 | 55522631-55522773 | - |
| 225 | WT1 | 11 | 32457189-32457408 | - |
| 226 | ZBTB16 | 11 | 113930478-113930929 | + |
| 227 | ZDBF2 | 2 | 207139497-207139592 | + |
| 228 | ZNF132 | 19 | 58951402-58951530 | - |
| 229 | ZNF134 | 19 | 58125597-58125764 | + |
| 230 | ZNF256 | 19 | 58459002-58459219 | - |
| 231 | ZNF331 | 19 | 54024023-54024436 | + |
| 232 | ZNF347 | 19 | 53662191-53662618 | - |
| 233 | ZNF354C | 5 | 178487123-178487607 | + |
| 234 | ZNF43 | 19 | 22018452-22018639 | - |
| 235 | ZNF43 | 19 | 22018852-22019004 | - |
| 236 | ZNF461 | 19 | 37157822-37158261 | - |
| 237 | ZNF569 | 19 | 37957760-37958046 | - |
| 238 | ZNF570 | 19 | 37960066-37960542 | + |
| 239 | ZNF610 | 19 | 52839503-52840013 | + |
| 240 | ZNF625 | 19 | 12267378-12267677 | - |
| 241 | ZNF665 | 19 | 53696101-53696195 | - |
| 242 | ZNF665 | 19 | 53696497-53696668 | - |
| 243 | ZNF682 | 19 | 20149796-20149923 | - |
| 244 | ZNF71 | 19 | 57106756-57106967 | + |
| 245 | ZNF788 | 19 | 12203466-12203641 | + |
| 246 | ZNF813 | 19 | 53970869-53971093 | + |
| 247 | ZNF813 | 19 | 53971239-53971374 | + |
| 248 | ZNF83 | 19 | 53193211-53193303 | - |

[0231]

AUC 및 배율 변화 비율 (암 대 정상 위 + 정상 결장)에 기초해서 123개의 개시된 마커 중 22개를 선택하고, 다른 GI 암 부위(결장, 췌장, 담관, 식도)로부터 73개의 발견 마커를 갖는 큰 조직 검증 연구에 사용되었다. 2개의 암, 정상, 및 적용 가능한 경우 선종에 대해 이러한 상이한 코호트의 샘플은 15개 내지 50개 샘플 사이의 범위에 있다. 검증 플랫폼은 메틸화 대조군 세트에 최적화된 프라이머를 갖는 정량적인 메틸화 특이적 PCR이었다. 모든 분석은 Roche 480 LightCyclers에 대해 수행하였다. 결과는 ROC 특성, 배율 변화 및 상보성에 대해 분석하였다. (암 및 선종에 대한)감도에 의해 랭크된 상부 30개가 표 2에 제공된다. 표 3은 표 2에 표시된 마커의 프라이머 정보를 제공한다. 특히, joint Mayo Korean 연구에서 추가의 조직 검증을 위해 10개가 선택되었다(실시에 2 참조)

표 2

| DMR No. | 유전자 | 감도 (100% spec) | 염색체 | 섬 범위 |
|---------|--------------|----------------|-----|---------------------|
| 253 | ELMO1 | 95% | 7 | 37487755-37488477 |
| 251 | ARHGEF4 | 91% | 2 | 131797843-131797938 |
| 254 | EMX1 | 88% | 2 | 73147710-73147772 |
| 255 | SP9 | 86% | 2 | 175202025-175202158 |
| 256 | CLEC11A | 86% | 19 | 51228217-51228732 |
| 249 | ST8SIA1 | 86% | 12 | 22487528-22487620 |
| 257 | BMP3 | 86% | 4 | 81952348-81952402 |
| 258 | KCNA3 | 84% | 1 | 111217654-111217816 |
| 259 | DMRTA2 | 81% | 1 | 50884349-50884499 |
| 260 | KCNK12 | 81% | 2 | 47797260-47797371 |
| 261 | CD1D | 79% | 1 | 158150797-158151205 |
| 262 | PRKCB | 79% | 16 | 23847825-23848025 |
| 250 | CYP26C1 | 79% | 10 | 94822416-94822607 |
| 263 | ZNF568 | 79% | 19 | 37407197-37407365 |
| 1 | ABCB1 | 79% | 7 | 87229775-87229856 |
| 264 | ELOVL2 | 77% | 6 | 11044395-11044834 |
| 265 | PKIA | 77% | 8 | 79428485-79428684 |
| 196 | SFMBT2 (893) | 74% | 10 | 7450406-7450831 |
| 266 | PCBP3 | 72% | 21 | 47063793-47064177 |
| 118 | MATK | 72% | 19 | 3785828-3786371 |
| 267 | GRIN2D | 70% | 19 | 48918144-48918350 |
| 268 | NDRG4 | 70% | 16 | 58497395-58497451 |
| 269 | DLX4 | 70% | 17 | 48042426-48042820 |
| 270 | PPP2R5C | 67% | 14 | 102247525-102247929 |
| 271 | FGF14 | 67% | 13 | 103046885-103047269 |
| 272 | ZNF132 | 67% | 19 | 58951402-58951775 |
| 46 | CHST2 (7890) | 65% | 3 | 142838645-142839023 |
| 273 | FLI1 | 63% | 11 | 128563956-128564209 |
| 252 | c13orf18 | 63% | 13 | 46960767-46961669 |
| 237 | ZNF569 | 70% (92% spec) | 19 | 37957760-37958046 |

표 3

| DMR No. | 유전자 | 정방향 MSP 프라이머 (5'-3') | 역방향 MSP 프라이머 (5'-3') |
|---------|-----------------|--|---|
| 253 | ELMO1 | TTA TAT TTT TCG TTT TTA GTA ATT TCG CGT TAG C (SEQ ID NO: 21) | GAA AAC CCG CCG AAA CAT TTC GA (SEQ ID NO: 22) |
| 251 | ARHGEF4 | TGT TTT CGC GGT CGT TAT ATA TTA CGT CGT (SEQ ID NO: 23) | GAA CTA TCC CCG AAC TCC GAC TCG A (SEQ ID NO: 24) |
| 254 | EMX1 | CGG GTT TTA GCG ATG TTT ATT TTA GTT TCC T (SEQ ID NO: 25) | CCT TTT CGT TCG TAT AAA ATT TCG TT (SEQ ID NO: 26) |
| 255 | SP8 | CGG GGG TAA AAA GGG TAG CGA TAG TGA TAC (SEQ ID NO: 27) | CGA AAA ATT CGA AAC GAA ACG TC (SEQ ID NO: 28) |
| 256 | CLEC11A | AGT TTG GCG TAG TCG GTA GAT C (SEQ ID NO: 29) | GCG CGC AAA TAC CGA ATA AAC G (SEQ ID NO: 30) |
| 249 | ST8SIA1 | GAC GTT TGT CGT CGG GTT CGT TC (SEQ ID NO: 13) | AAA AAC CCT CCG CTA CCA CTT CGC (SEQ ID NO: 14) |
| 257 | BMP3 | GTTTAATTCGTTTCGTCGTC (SEQ ID NO: 31) | CGCTACGAAACACTCCGA (SEQ ID NO: 32) |
| 258 | KCNA3 | TTT TGT TGG GTG GGT ATT TTC CGC (SEQ ID NO: 33) | ACA ACC CTC AAA ACC CCT CGA TAT C (SEQ ID NO: 34) |
| 259 | DMRTA2 | TTT ATT TAT GGA TTA CGT TTT TAG CGA (SEQ ID NO: 35) | GAC GAC CGA ACG ATC ACG CA (SEQ ID NO: 36) |
| 260 | KCNK12 | GCG TCG TTA GTA CGA CGA AGC (SEQ ID NO: 37) | GCA CCT CAA CGA AAA CAC CGA (SEQ ID NO: 38) |
| 261 | CD1D | GCG CGT AGC GGC GTT TCC (SEQ ID NO: 39) | CCC ATA TCG CCC GAC GTA A (SEQ ID NO: 40) |
| 262 | PRKCB | GTC GTT TTT GGG CGT TTT AGA GGC (SEQ ID NO: 41) | TAATCC CAA ACC AAC CGC CG (SEQ ID NO: 42) |
| 250 | CYP26C1 | GGT TTT TTG GTT ATT TCG GAA TCG T (SEQ ID NO: 43) | TAT AAA AAC CGC CGT AAT CAA CGC T (SEQ ID NO: 44) |
| 263 | ZNF568 | TTG AGA TGT TGG GTG AAG GCG ATT C (SEQ ID NO: 45) | CGC TAA CGC GAA AAA ATA ATT CGA CG (SEQ ID NO: 46) |
| 1 | ABCB1 | GAT TTT GTT CGT CGT TAG TGC (SEQ ID NO: 9) | TCT CTA AAC CGC CGA ACG A (SEQ ID NO: 10) |
| 264 | ELOVL2 | CGG TTT TAT TTA TTA TGA TTC GTA CGG G (SEQ ID NO: 47) | CGA CTA CCC TAA ACA ACG CAT CGC (SEQ ID NO: 48) |
| 265 | PKIA | CGG GGA TGA TTT TAT GTA GTC GGA GTT TCG C (SEQ ID NO: 49) | CCC GCC GAA TAC TCG ATC AAC TCG (SEQ ID NO: 50) |
| 196 | SFMBT2 (893) | AAA ACG TTT TTA GGT ATT TGG TCG T (SEQ ID NO: 5) | ATA AAA AAA CCT CCT CCG AAC CGC C (SEQ ID NO: 6) |
| 266 | PCBP3 | GGT CGC GTC GTT TTC GAT C (SEQ ID NO: 51) | GCC GCA AAC GCC GAC GA (SEQ ID NO: 52) |
| 118 | MATK | TGT ATA TTT CGA GGC GGT TTC GG (SEQ ID NO: 53) | CGC CCC CAA AAT AAA AAA ACG AA (SEQ ID NO: 54) |
| 267 | GRIN2D | TTT TTC GTT CGT TTT TTT ATT TTT TCG A (SEQ ID NO: 55) | AAC GCA TAC CAT CGA CTT CAA CTA CGA C (SEQ ID NO: 56) |
| 268 | NDRG4 | CGGGTTTCGTTCTCGTTTTTCG (SEQ ID NO: 57) | CCGGCTTCTACCGCGACTA (SEQ ID NO: 58) |
| 269 | DLX4 | GTA TTT TTA ATA TTT GGT GAG TGC G (SEQ ID NO: 59) | TAC TAA AAC GTA CGA TAA ACA TCG T (SEQ ID NO: 60) |
| 270 | PPP2R5C | TCG ATT TTA TTT TTG TTG TCG TTG TAG ATT CGC (SEQ ID NO: 61) | GAA AAA ACT AAA AAA CGA CAA AAA AAC CGG ACG (SEQ ID NO: 62) |
| 271 | FGF14 | GGA GGG AGT TTA ACG GGT TAG GTA C (SEQ ID NO: 63) | CGT AAA CAA AAA ACA ACG ACG ACG (SEQ ID NO: 64) |
| 272 | ZNF132 | ATT GCG TCG TTG TTT AGG TAA CGT A (SEQ ID NO: 65) | CGA AAA TAC CTA TCT CCT CGA CT (SEQ ID NO: 66) |
| 46 | CHST2 (780) | GGA ACG AGT GAT AGT CGG ATA GTT CGT C (SEQ ID NO: 67) | CGC CGG AAA ACG ACC CCG (SEQ ID NO: 68) |
| 273 | FLI1 | GGG AGT GAG GGT AGG GCG TTC (SEQ ID NO: 69) | CTC GCA ACC CCT TCG AAT TAA CCC G (SEQ ID NO: 70) |
| 252 | c13orf18 | TTT AGG GAA GTA AAG CGT CGT TTT C (SEQ ID NO: 19) | AAC GAC GTC TCG ATA CCT ACG A (SEQ ID NO: 20) |
| 237 | ZNF569 | TGT GGA ATC GGG GTT TGT GTT CGC (SEQ ID NO: 11) | CCC ACC CAA CAC AAA AAA TCC GAC G (SEQ ID NO: 12) |

[0234]

[0235] 실시예 2 - 신규한 메틸화 DNA 마커에 의한 위암 검출: 미국 및 한국으로부터 환자 코호트에서 조직 검증

[0236]

본 기술에 대한 실시형태를 전개하는 과정에서 수행된 실험은 전체 메틸롬(methylome) 서열 결정에 의해 신규한 메틸화된 DNA 마커를 확인한 후 미국과 한국의 환자 코호트로부터 조직에서 우수한 후보를 검증했다(표 4, 5, 및 6 참조)

[0237]

RRBS(representation bisulfite sequencing)을 사용하여 전체 메틸롬 발견 시도 후, 17개의 메틸화 DNA 마커 후보는 미국 및 한국에서 환자로부터 위 조직 세트에서 맹검 검증을 위해 선택되었다. 미세설계된 조직으로부터의 DNA는 메틸화-특이적 폴리미라제 연쇄 반응에 의해 분석되고; 마커 수준은 베타-액틴(전체 인간 DNA)에 의해 표준화되었다. GC 중례는 병리적으로 확인된 완전한 선암에 걸린 35명의 미국 환자 및 50명의 한국 환자를 포함했다. 대조군은 65명의 건강한 미국 환자에서 조직적으로 정상 위 상피조직을 포함시켰는데 이는, 중례 및 동일한 50 명의 한국인 환자와 인구 통계학적으로 일치했지만 1차 GC로부터 떨어진 장소에서 수득되었다. 각 마커에 대해 ROC 곡선(AUC) 항의 면적은 공정 로지스틱 회귀로부터 계산되었다.

[0238]

전반적으로 상위 10 개 마커는 ARHGEF4, ELM01, ABCB1, CLEC11A, ST8SIA1, SFMBT2, CD1D, CYP26C1, ZNF569, 및 C13ORF18를 포함한다 (표 4 참조). 95 %의 특이성에서, 이들 마커들의 패널은 미국 코호트의 100 % 및 한국 코호트의 94 %에서 GC를 검출 하였다. 미국 코호트에서 각각의 AUC는 0.92, 0.91, 0.89, 0.88, 0.85, 0.82, 0.82, 0.81, 0.75, 및 0.66 이었고 한국 코호트에서 0.73, 0.90, 0.78, 0.75, 0.80, 0.94, 0.87, 0.79, 0.75, 및 0.79로 나타났다 (표 6 참조). ELM01과 같은 일부 마커는 각 환자 코호트에서 GC와 유사하게 높은 차별을 보였고, ARHGEF4와 같은 다른 마커는 환자 코호트에서 정상 대조군 (NL)에 대한 높은 백그라운드 때문에 더 가변적인 차별을 나타냈다 (도 1 참조). 표 5는 표 4에 표시된 10 개의 마커에 대한 정방향 및 역방향 프라이머 정보를 제공한다.

표 4

| DMR No. | 유전자 | 염색체 | 섬 범위 |
|---------|--------------|-----|---------------------|
| 49 | CLEC11A | 19 | 51228264-51228394 |
| 43 | CD1D | 1 | 158150624-158151205 |
| 196 | SFMBT2 (893) | 10 | 7450406-7450831 |
| 66 | ELMO1 | 7 | 37487867-37488498 |
| 1 | ABCB1 | 7 | 87229775-87229856 |
| 237 | ZNF569 | 19 | 37957760-37958046 |
| 249 | ST8SIA1 | 12 | 22487528-22487620 |
| 250 | CYP26C1 | 10 | 94822416-94822607 |
| 251 | ARHGEF4 | 2 | 131797843-131797938 |
| 252 | c13orf18 | 13 | 46960767-46961669 |

[0239]

표 5

| DMR No. | 유전자 | 정방향 프라이머 (5'-3') | 역방향 프라이머 (3'-5') |
|---------|--------------|---|--|
| 49 | CLEC11A | AGT TTG GCG TAG TCG GTA GAT C (SEQ ID NO: 1) | GCG CGC AAA TAC CGA ATA AAC G (SEQ ID NO: 2) |
| 43 | CD1D | GCG CGT AGC GGC GTT TC (SEQ ID NO: 3) | CCC ATA TCG CCC GAC GTA A (SEQ ID NO: 4) |
| 196 | SFMBT2 (893) | AAA ACG TTT TTA GGT ATT TGG TCG T (SEQ ID NO: 5) | ATA AAA AAA CCT CCT CCG AAC CCG C (SEQ ID NO: 6) |
| 66 | ELMO1 | GGA GGG AGT TTA ACG GGT TAG GTA C (SEQ ID NO: 7) | CGT AAA CAA AAA ACA ACG ACG ACG (SEQ ID NO: 8) |
| 1 | ABCB1 | GAT TTT GTT CGT CGT TAG TGC (SEQ ID NO: 9) | TCT CTA AAC CCG CGA ACG A (SEQ ID NO: 10) |
| 237 | ZNF569 | TGT GGA ATC GGG GTT TGT GTT CGC (SEQ ID NO: 11) | CCC ACC CAA CAC AAA AAA TCC GAC G (SEQ ID NO: 12) |
| 249 | ST8SIA1 | GAC GTT TGT CGT CGG GTT CGT TC (SEQ ID NO: 13) | AAA AAC CCT CCG CTA CCA CTT CGC (SEQ ID NO: 14) |
| 250 | CYP26C1 | GGT TTT TTG GTT ATT TCG GAA TCG T (SEQ ID NO: 15) | TAT AAA AAC GCG CGT AAT CAA CGC T (SEQ ID NO: 16) |
| 251 | ARHGEF4 | TGT TTT CGC GGT CGT TAT ATA TTA CGT CGT (SEQ ID NO: 17) | GAA CTA TCC CCG AAC TCC GAC TCG A (SEQ ID NO: 18) |
| 252 | c13orf18 | TTT AGG GAA GTA AAG CGT CGT TTT C (SEQ ID NO: 19) | AAC GAC GTC TCG ATA CCT ACG A (SEQ ID NO: 20) |

[0240]

표 6

| DMR No. | 유전자 | 곡선하 면적 US 코호트 | 곡선하 면적 한국 코호트 |
|---------|--------------|---------------|---------------|
| 49 | CLEC11A | 0.88 | 0.75 |
| 43 | CD1D | 0.82 | 0.87 |
| 196 | SFMBT2 (893) | 0.82 | 0.94 |
| 66 | ELMO1 | 0.91 | 0.9 |
| 1 | ABCB1 | 0.89 | 0.78 |
| 237 | ZNF569 | 0.75 | 0.75 |
| 249 | ST8SIA1 | 0.85 | 0.8 |
| 250 | CYP26C1 | 0.81 | 0.79 |
| 251 | ARHGEF4 | 0.92 | 0.73 |
| 252 | c13orf18 | 0.66 | 0.79 |

[0241]

[0242] 실시예 3- 혈장에서의 신규한 메틸화 DNA 마커의 분석에 의한 위암 검출

[0243] 이 실시예는 위 선암 (GC) 검출 접근법으로서 혈장으로부터 분석하기 위해 실시예 1 및 2의 선별된 메틸화 DNA 마커 (MDM) 후보의 사용을 조사하였다. 실제, 이 실시예는 혈장 내에서 분석된 신규한 MDM 패널이 정상 대조군의 GC 증례와 정확하게 구별함을 입증한다. 마커 수준은 대조군에서는 무시할 수 있지만 경우에 따라 증가하고 GC 단계에 따라 점진적으로 증가하는 것으로 나타났다.

[0244] 단일 센터에서 증례 진입 기준을 충족시키는 아카이브 혈장 샘플이 연구에 포함되었다. 증례는 병리학적으로 확인된 완전 선암을 가진 37 명의 환자를 포함하고, 모든 단계, 조직학적 유형 및 위 부위에 걸쳐 나타났다. 증례 혈장 샘플은 절제술, 화학 요법 또는 방사선 치료 전에 수집되었다. 대조군으로 38 세 및 성별이 일치하는 건강한 지원자들의 아카이브 혈장 샘플을 선정했다. 2mL 혈장 및 비설 파이트 처리로부터 DNA를 추출하였다. 최적화되지 않은 정량적 대립 유전자 특이적 실시간 표적 및 신호 증폭 방법을 이용하여 선택된 MDM (ELMO1, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1) (표 7은 LRCC4에 대한 DMR 정보를 제공) 을 분석하였다. 이러한 마커는 임상 감도와 특이성, 위암과 정상 위 조직을 비교할 때 메틸화 신호의 배수 변화, 및 혈액 세포 유래 DNA의 매우 낮은 백그라운드 메틸화에서 랭크된 조직에서 높은 성능을 보였다.

[0245] 혈장에서 가장 높은 분석 성능을 촉진하기 위해 이전에 메틸화 특이 적 PCR 분석법을 사용하여 조직에서 절의된 마커를 원래의 DMR 서열을 다중화 QuART (정량적 대립유전자 특정 실시간 표적 및 신호 증폭) 분석법으로 재설계했다. QuARTs 분석은 PCR 프라이머 (표 8 참조), 검출 탐침 (표 8 참조) 및 침습성 올리고 (Integrated DNA Technologies), GoTaq DNA 폴리머라제 (Promega), Cleavase II (Hologic) 및 FAM, HEX 및 Quasar 670 염료 (Biosearch Technologies)가 포함 된 형광 공명 에너지 전달 리포터 카세트 (FRETs)를 포함한다.

[0246] 표 2는 QuART (정량적 대립유전자 특이적 실시간 표적 및 신호 증폭) 분석에 의한 메틸화 DNA 서명의 검출에 사용되는 FRET 카세트의 올리고 뉴클레오타이드 서열을 제공한다. 각각의 FRET 서열은 3 개의 별도의 분석법으로 함께 다중화될 수 있는 형광체 및 쿠엔처를 포함한다.

[0247] 표준은 플라스미드 대조군에서 성장하고 절단된 표적 서열로부터 유도되었다. 연속적으로 희석된 표준의 절대 카피는 포아송 (Poisson) 모델링으로부터 수득되었다.

[0248] DNA는 아카이브 혈장 샘플 2mL에서 정제한 후 사내 독자적인 방법 및 자동화 기기를 사용하여 비설파이트로 전환했다. 연구 중의 표적 회복 변동성을 보정하기 위해 모든 샘플에 공정 제어 유전자를 추가했다.

[0249] 샘플을 12 개의 프라이머 세트를 사용하여 예비 증폭시키고, 희석하고, QuART 증폭시켰다. 후자의 단계를 위한 플랫폼은 LightCycler 480 (Roche)이었다. 각각의 마커로부터 가닥 카운트를 다중 반응에 포함된 β -액틴 대조군 분석으로부터 가닥으로 정규화 하였다.

[0250] 결과를 논리적으로 분석하고 상보성을 평가하기 위해 매트릭스 형태로 플로팅하였다.

[0251] 표 3은 개별 마커 곡선과 비교하여 3개의 마커 패널 (ELMO1, ZNF569 및 c13orf18)의 혈장 성능을 강조하는 수용

자 동작 특성 곡선을 제공한다. 100 % 특이성에서 패널은 위암을 86 %의 감도로 검출했다.

[0252] 도 4a는 정상 점막에서 단계 4 위암으로의 진행을 나타내는 단계별로 혈장 내 메틸화된 ELM01의 로그 스케일 절대 가닥 카운트를 제공한다. 정량 마커 수준은 GC 단계에서 바로 증가했다.

[0253] 도 4b는 3개의 마커 패널 (ELM01, ZNF569, 및 c13orf18)의 단계별로 (100 % 특이성에서) 혈장 내 위암 감도를 보여주는 막대 차트를 제공한다.

[0254] 실제로, ELM01은 가장 차별적인 마커이었고 단독으로 94 %의 AUC를 보였다 (95 % CI 89-99 %). 100 % 특이성에서, 3 개의 MDM 패널 (ELM01, ZNF569 및 C13orf18)은 86 % (71-95 %)의 감도로 GC를 검출하였다. GC 단계 (도 4a)에서, 100 % 특이성에서 패널에 의한 감도는 단계 1 (n = 8), 2 (n = 13), 3 (n = 4) 및 4(n=11)에 대해 50 %, 92 %, 100 % 및 100 %이었다 (경향에 대해 p = 0.01). 정량적인 MDM 수준은 GC 단계에서 바로 증가했다. 메틸화된 ELM01 (도 4b)의 분포에서 볼 수 있듯이, 대조군의 중간 혈장 가닥 카운트 / ml는 0.4 였고, 단계 1, 2, 3 및 4 의 GC 중례의 가닥 카운트는 각각 16, 111, 101 및 213이었다 (경향에 대해 p = 0.01). 나이도 성별도 마커 수준에 영향을주지 않는다.

[0255] 도 5a, b 및 c는 90 % 특이성 (A), 95 % 특이성 (B) 및 100 % 특이성 (C)에서 매트릭스 형식으로 74 개의 혈장 샘플에서 12 개의 위암 마커 (ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; ELM01 은 두 번째는, 바이플렉스 형식임)을 제공한다. 마커는 수평방향으로 도시되고 샘플은 수직방향으로 도시된다. 샘플은 좌측에 정상샘플이 배열되고 우측이 암이 배열되어 있다. 양성 히트는 밝은 회색이고 짙은 회색부분에는 존재하지 않는다. 여기에서 상부 수행 마커 -ELM01-은 먼저 90 % 특이성에서 처음 나열되고 나머지 패널은 100 % 특이성으로 표시된다. 이 플롯에서 마커가 상보적 방법으로 평가될 수 있다.

[0256] 도 6은 100 % 특이성에서 74 개의 혈장 샘플 중 3 개의 위암 마커 (최종 패널)의 수행을 제공한다. 마커는 수직방향으로 도시되고 샘플 수평방향으로 도시된다. 샘플은 좌측에 정상샘플이 도시되고 우측에 암이 배열되어 있다. 암은 단계별로 배열된다. 양성 히트는 밝은 회색이고 짙은 회색부분에는 존재하지 않는다(ELM01, ZNF569 및 c13orf18).

[0257] 도 7a~m은 혈장 중 12 개의 위암 마커(ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 및 ST8SIA1; ELM01 는 두 번째는 바이플렉스 형식임)의 박스 플롯 (선형 스케일)을 제공한다. 대조 샘플 (N = 38)은 좌측에 도시되고 우측에는 위암 중례 (N = 36)가 도시된다. 수직축은 β - 액틴 가닥에 대해 정규화된 % 메틸화이다.

표 7

| DMR No. | 유전자 | 염색체 | 섬 범위 |
|---------|-------|-----|---------------------|
| 274 | LRCC4 | 7 | 127671993-127672310 |

[0258]

표 8

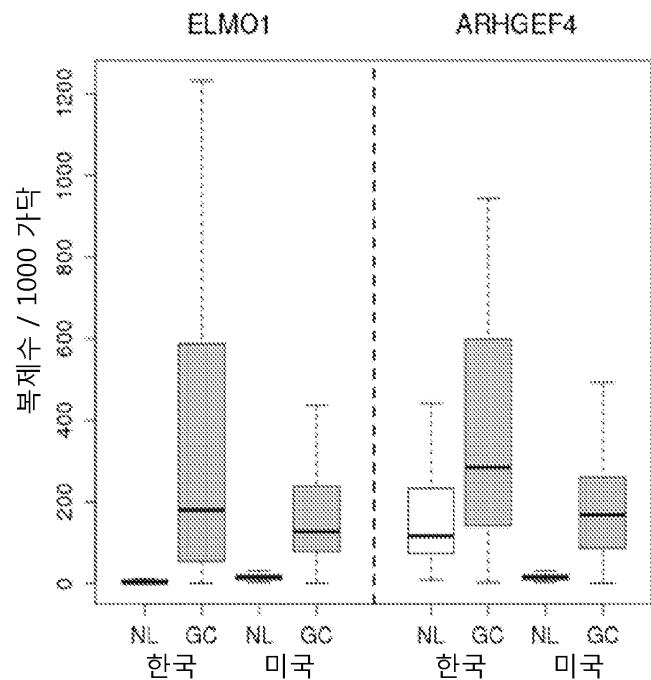
| 마커 | 올리고 타입 | 올리고 명칭 | 서열 |
|--------------|--------|------------------------|--|
| PKIA | 프라이머 | F 프라이머 PKIA | GCGAGTTGATCGAGTATTCGG (SEQ ID NO: 71) |
| | 프라이머 | R 프라이머 PKIA | CTAATCAACGCAAATCCCCGC (SEQ ID NO: 72) |
| | 프로브 | PKIA_Pb_A5 | CCACGGACGCGCACGCCGCCA/3C6/ (SEQ ID NO: 73) |
| C13ORF18 | 프라이머 | C13ORF18_FP_1 | CGTAGTTTTAGAAGTTTTAGGGA AGT (SEQ ID NO: 74) |
| | 프라이머 | C13ORF18_RP_1 | CGACGTCGATACCTACGA (SEQ ID NO: 75) |
| | 프로브 | C13ORF18_프로브_1 | CCACGGACGACGAAACGACG /3C6/ (SEQ ID NO: 76) |
| CLEC11A | 프라이머 | CLEC_FP_1 | GCGGGAGTTGGCGTAG (SEQ ID NO: 77) |
| | 프라이머 | CLEC_RP_1 | CGCGCAAATACCGAATAAACG (SEQ ID NO: 78) |
| | 프로브 | CLEC_프로브_1 | CCACGGACGTCGGTAGATCG/3C6/ (SEQ ID NO: 79) |
| ELMO | 프라이머 | F 프라이머 ELMO1 | GTTTCCGGTTTCGGGTTTG (SEQ ID NO: 80) |
| | 프라이머 | R 프라이머 ELMO1 | CGCTAAATAACGACGCCCG (SEQ ID NO: 81) |
| | 프로브 | 프로브 A5 ELMO1_v1a | CCACGGACGCGTTTATTCTG /3C6/ (SEQ ID NO: 82) |
| SFMBT2 (896) | 프라이머 | F SFMBT2_896 ver2 | GCGTTAGGTTGGTCGGAG (SEQ ID NO: 83) |
| | 프라이머 | R SFMBT2_896 ver2 | ACGCACTAACCTACGAAC (SEQ ID NO: 84) |
| | 프로브 | 프로브 A5 SFMBT2_896 ver2 | CCACGGACGCCGAAAAACTAC/3C6/ (SEQ ID NO: 85) |
| ZNF569 | 프라이머 | ZNF569 FP_V2 | AGAGTTCGGCGTTAGAGTTAG (SEQ ID NO: 86) |
| | 프라이머 | ZNF569 RP_V2 | TCGAAACCTATATCCCGCGAAC (SEQ ID NO: 87) |
| | 프로브 | ZNF569 Pb A5_V2 | CCACGGACGGCGGTTTCGA/3C6/ (SEQ ID NO: 88) |
| BTACT | 프라이머 | ACTB BT FP1 | TTTGTTTTTGATTTAGGTGTTAAGA (SEQ ID NO: 89) |
| | 프라이머 | ACTB BT RP1 | CACCAACCTCATAACCTTATC (SEQ ID NO: 90) |
| | 프로브 | ACTB BT PB A3 | GACGCAGGAGATACTGTTGTGG /3C6/ (SEQ ID NO: 91) |
| PPP2R5C | 프라이머 | F PPP2R5C | TTCGATTTTATTTGTGTCGTGTAGA (SEQ ID NO: 92) |
| | 프라이머 | PPP2R5C RP_V2 | ACGACAAAAAAACCCGACGA (SEQ ID NO: 93) |
| | 프로브 | PPP2RC_Pb_A1 | CGCCGAGGGATTCGCGTGGGT/3C6/ (SEQ ID NO: 94) |
| ARHGEF4 | 프라이머 | ARHGEF4 FP | CGTTCGCGTTATTTATTCGGCG (SEQ ID NO: 95) |
| | 프라이머 | ARHGEF4 RP | GCTCCTAATTCTCATCACGTCGT (SEQ ID NO: 96) |
| | 프로브 | ARHGEF4 Pb A1 | CGCCGAGGGCGCGTTGC/3C6/ (SEQ ID NO: 97) |
| CD1D | 프라이머 | CD1D_FP_2a | GAAGAGTCCGTAGGTTAGAGG (SEQ ID NO: 98) |
| | 프라이머 | CD1D_RP_2a | CATATCGCCGACGTAAAACC (SEQ ID NO: 99) |
| | 프로브 | CD1D_프로브_2a_A1 | CGCCGAGGGCTCGCGAACGC/3C6/ (SEQ ID NO: 100) |
| ST8SIA1 | 프라이머 | ST8SIA1_FP_ESO | AAATATGTGGTTCGTGGCGTT (SEQ ID NO: 101) |
| | 프라이머 | ST8SIA1_RP_ESO | ACGCAACACCGCGAAAAAC (SEQ ID NO: 102) |
| | 프로브 | ST8SIA1_Pb_A1_ESO | CGCCGAGGGCGCGTTGC/3C6/ (SEQ ID NO: 103) |
| CYP26C1 | 프라이머 | CYP26C1 FP | CGTTCGCGTTATTTATTCGGCG (SEQ ID NO: 104) |
| | 프라이머 | CYP26C1 RP | GCTCCTAATTCTCATCACGTCGT (SEQ ID NO: 105) |
| | 프로브 | CYP26C1 Pb A1 | CGCCGAGGGCGCGTTGC/3C6/ (SEQ ID NO: 106) |
| LRRC4 | 프라이머 | LRRC4 FP | GGCTTAATTCCCGAGGTA (SEQ ID NO: 107) |
| | 프라이머 | LRRC4 RP | ACAATACTCTTATATATTAACGCCGCCTC (SEQ ID NO: 108) |
| | 프로브 | LRRC4 Pb A1 | CGCCGAGGGAGGCGACGGAGG/3C6/ (SEQ ID NO: 109) |

[0259]

[0260] 상기 명세서에 언급된 모든 간행물 및 특허는 모든 목적을 위해 그 전체가 본원에 참고로 포함된다. 본 기술의 범위 및 사상을 벗어나지 않고 기술된 조성물, 방법 및 용도의 다양한 변경 및 변형이 당업자에게 명백할 것이다. 본 기술은 특정한 예시적인 실시형태와 관련하여 기술되었지만, 청구된 본 발명이 그러한 특정 실시형태에 부적절하게 제한되어서는 안 된다는 것을 이해해야 한다. 실제로, 약리학, 생화학, 의학 또는 관련 분야의 당업자에게 자명한 본 발명을 수행하기 위한 기술된 방식의 다양한 변경은 하기 청구 범위의 범주 내에 있는 것으로 의도된다.

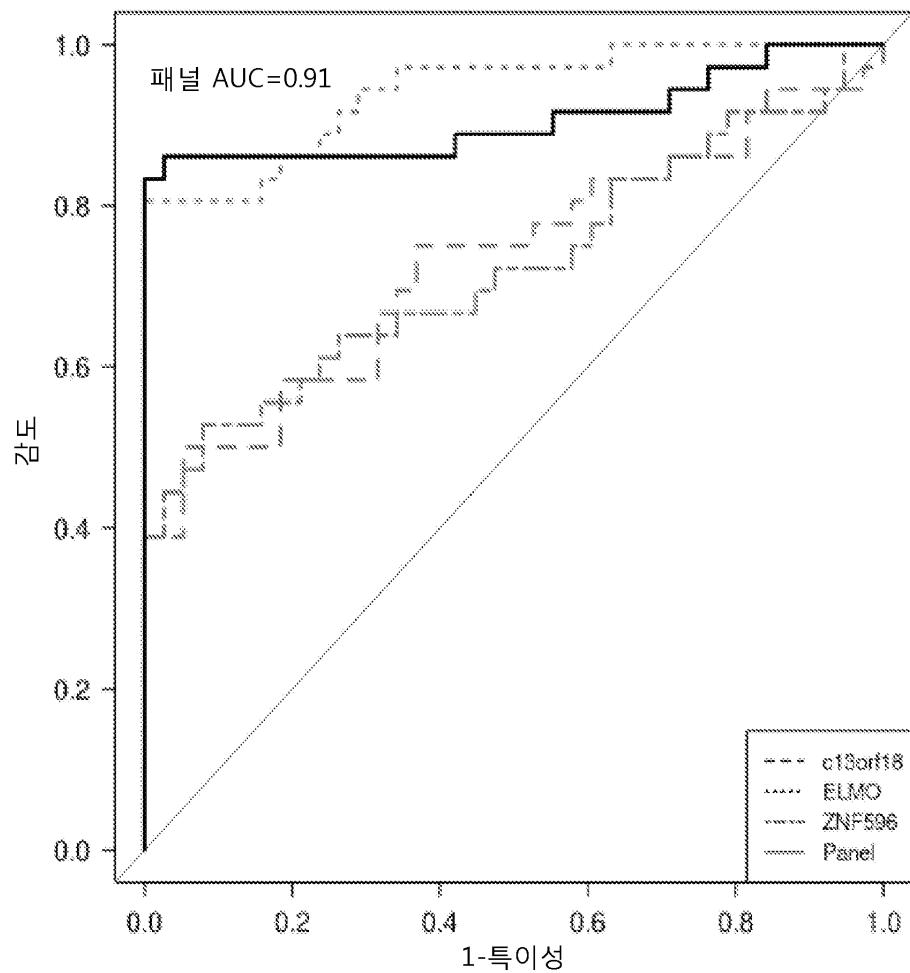
도면

도면1

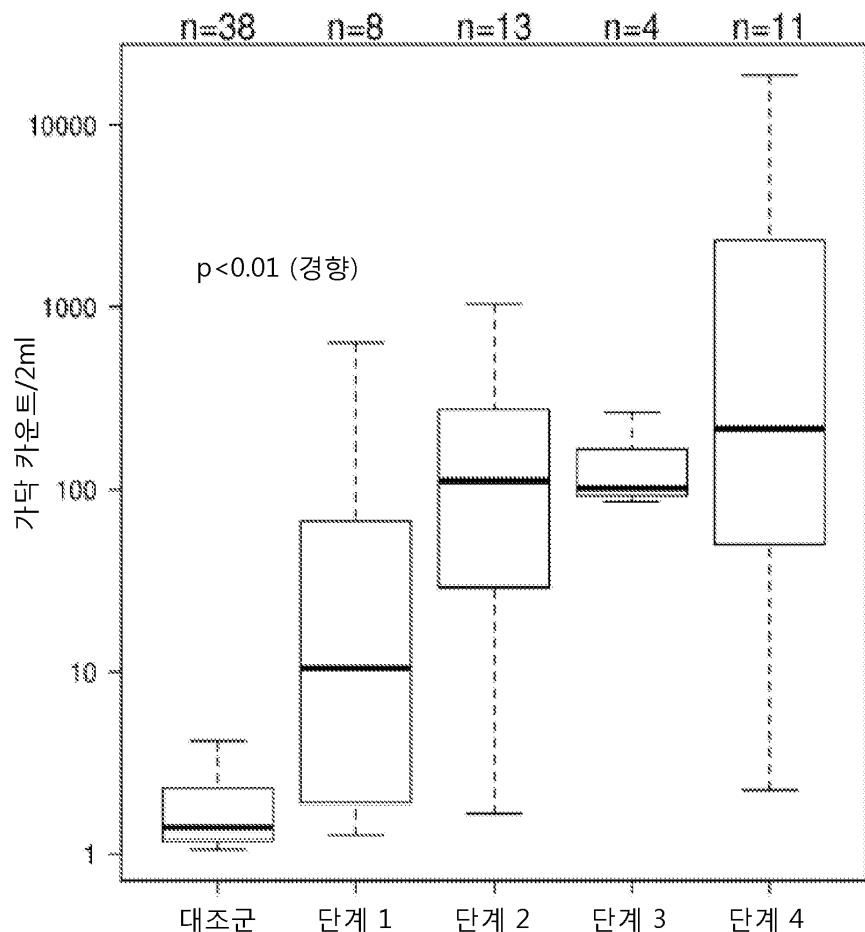


도면2

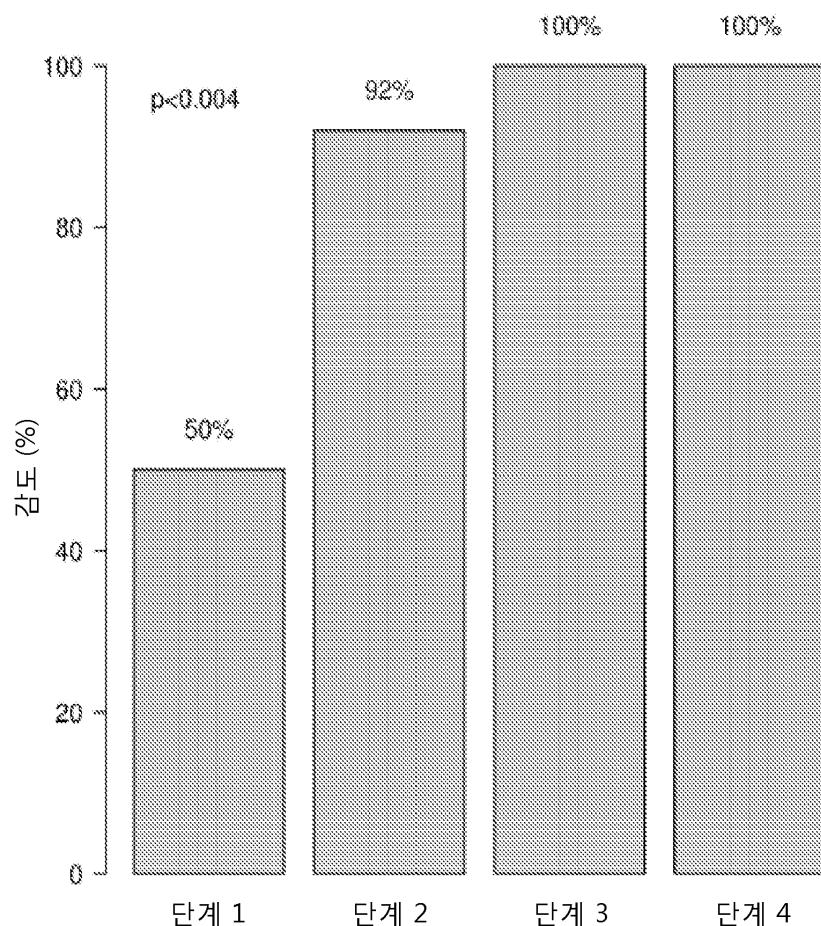
도면3



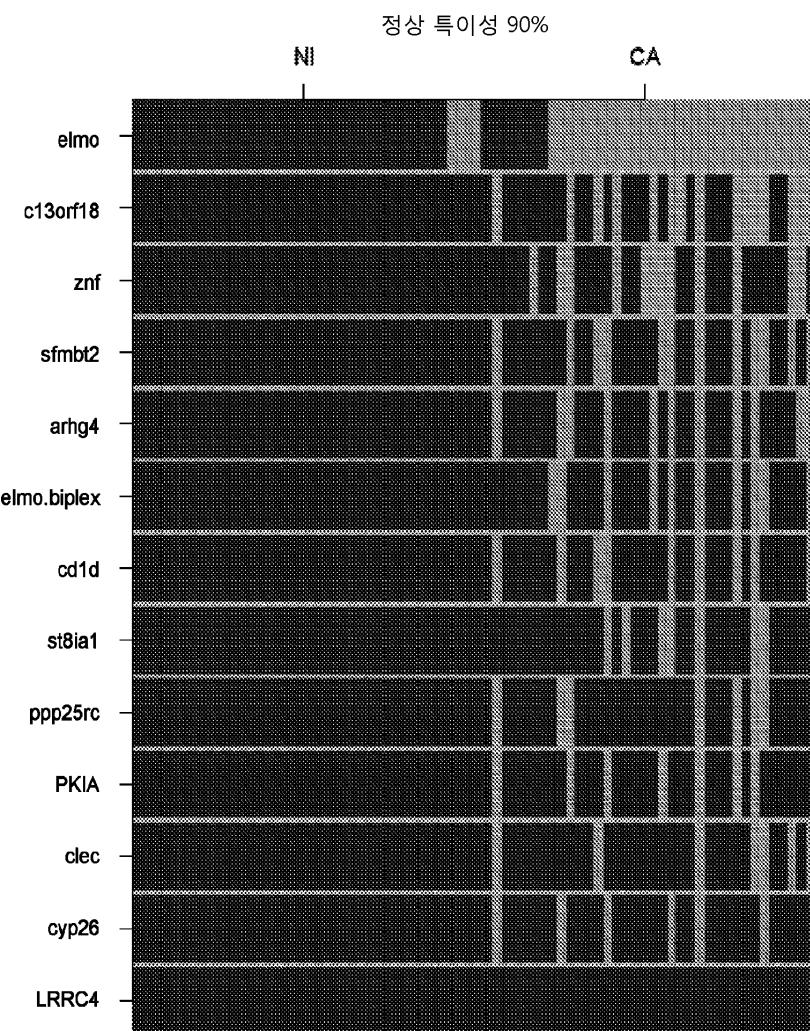
도면4a



도면4b



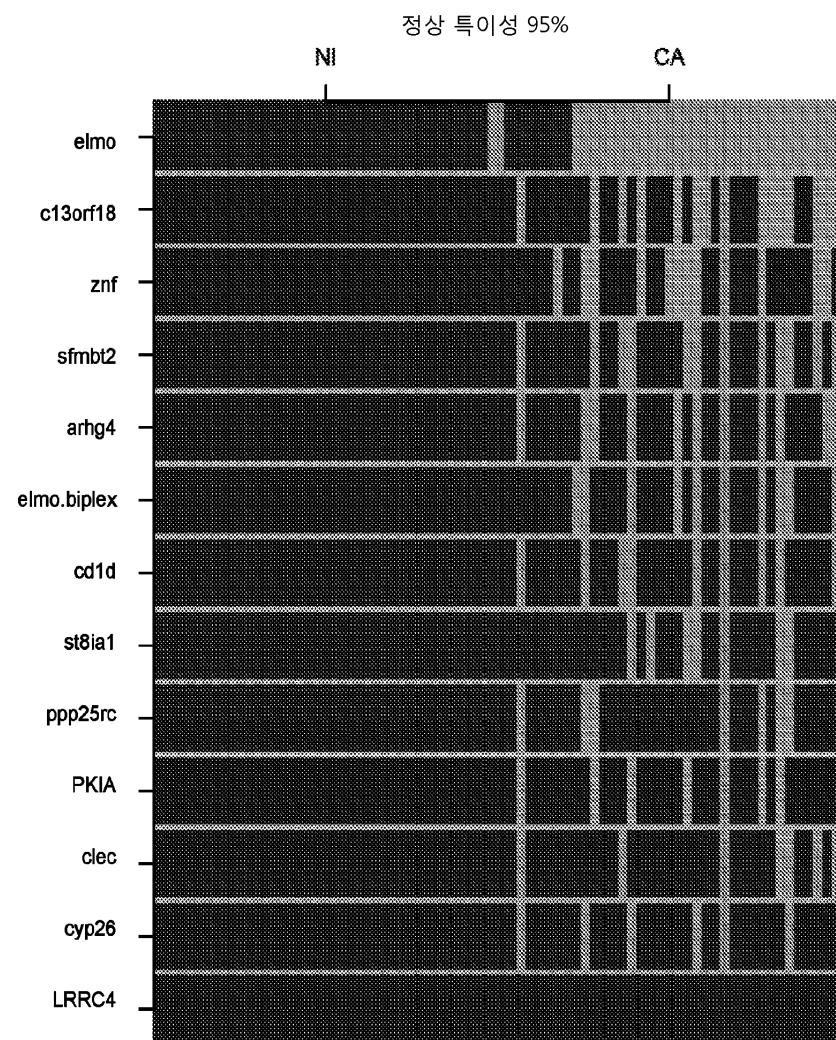
도면5a



피험체

| 대조군 | | | 증례 | | |
|-------|---|------|----|----|---|
| 거짓 | 참 | 거짓 | 증례 | 거짓 | 참 |
| 거짓 34 | 0 | 거짓 5 | 2 | | |
| 참 0 | 4 | 참 0 | 29 | | |

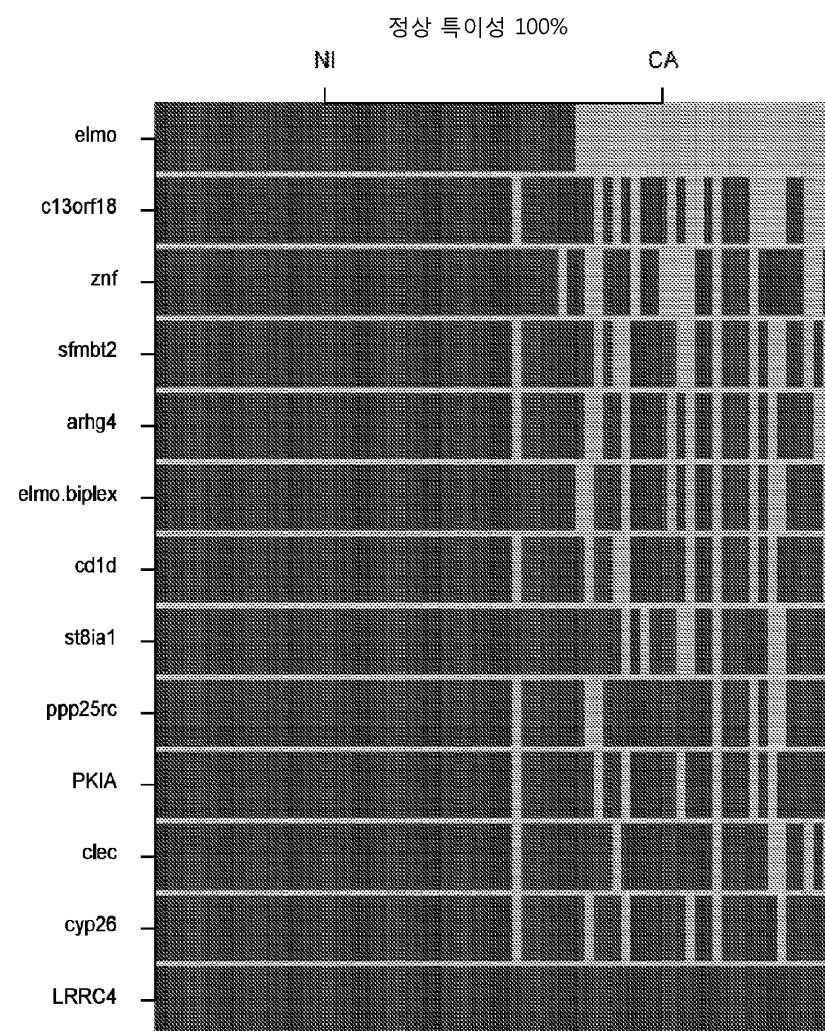
도면5b



피험체

| 대조군 | | | 증례 | | |
|-------|---|--|------|----|--|
| 거짓 | 참 | | 거짓 | 참 | |
| 거짓 36 | 0 | | 거짓 5 | 2 | |
| 참 0 | 2 | | 참 0 | 29 | |

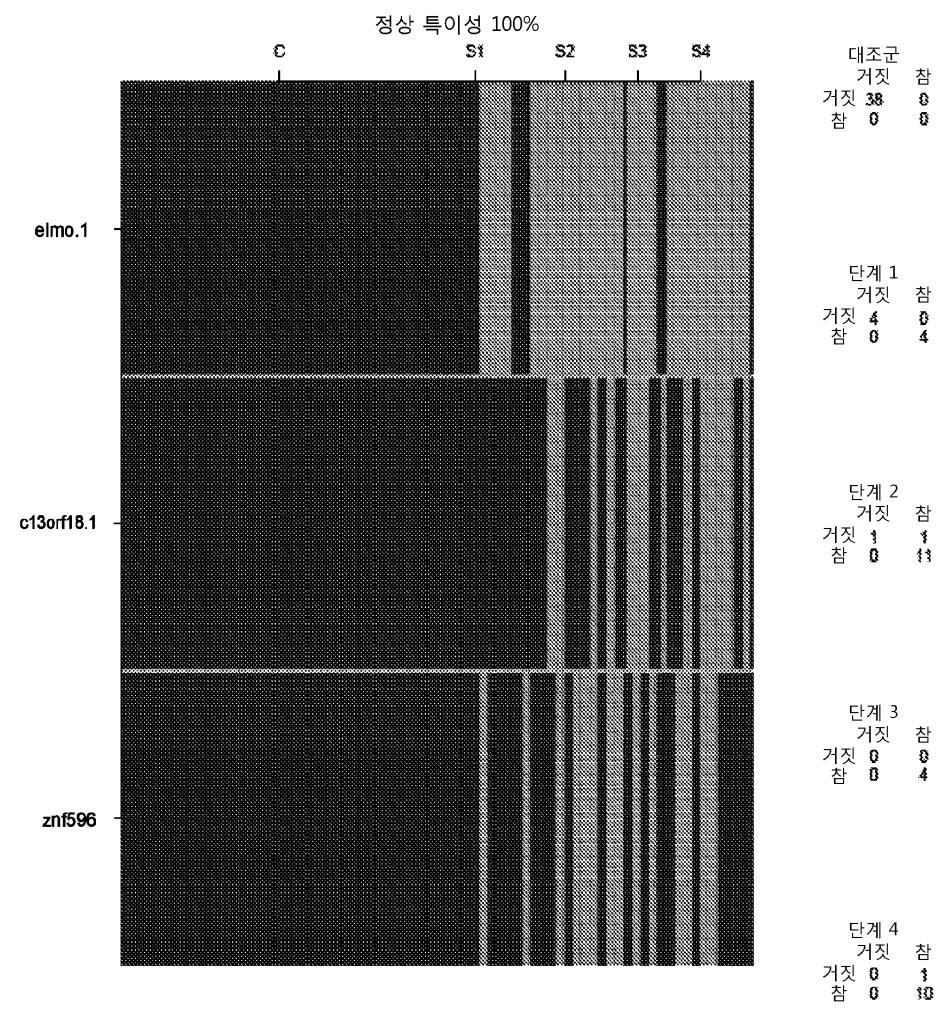
도면5c



피험체

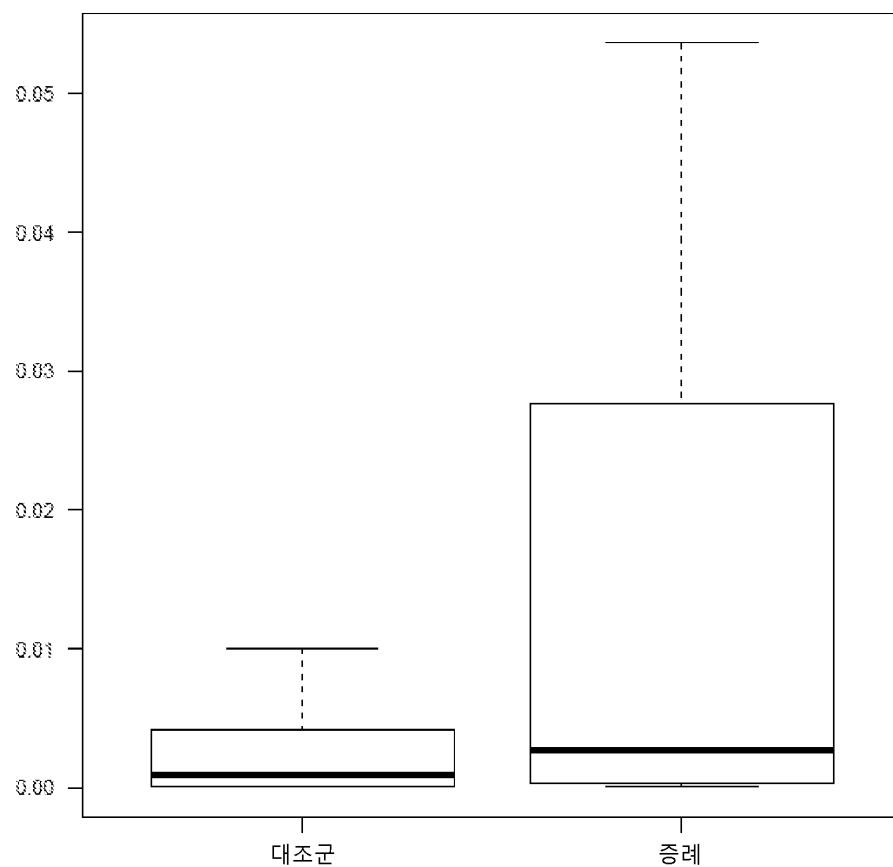
| 대조군 | | 증례 | |
|-------|---|------|----|
| 거짓 | 참 | 거짓 | 참 |
| 거짓 38 | 0 | 거짓 6 | 2 |
| 참 0 | 0 | 참 0 | 28 |

도면6



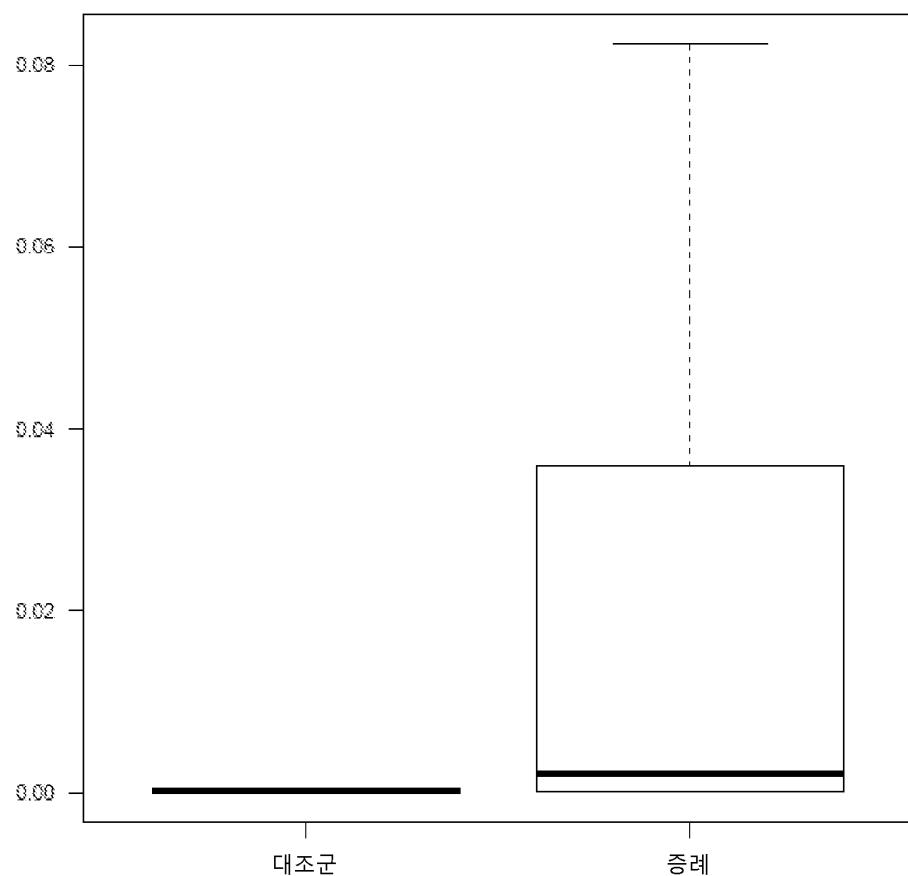
도면7a

arhg4



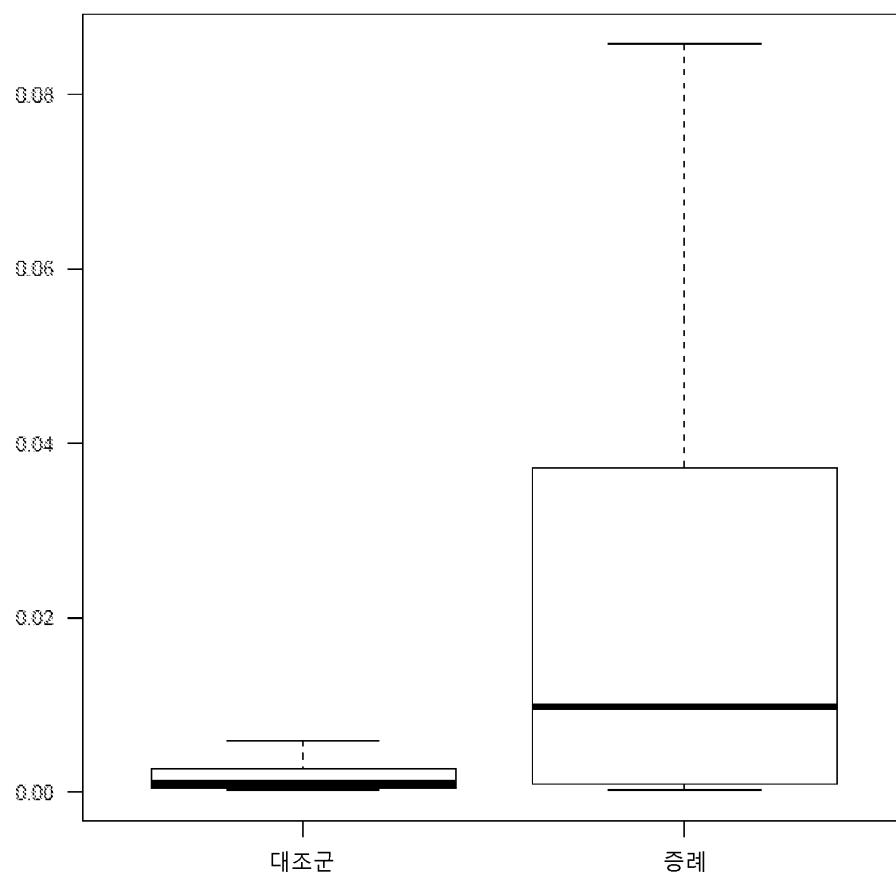
도면7b

c13orf18

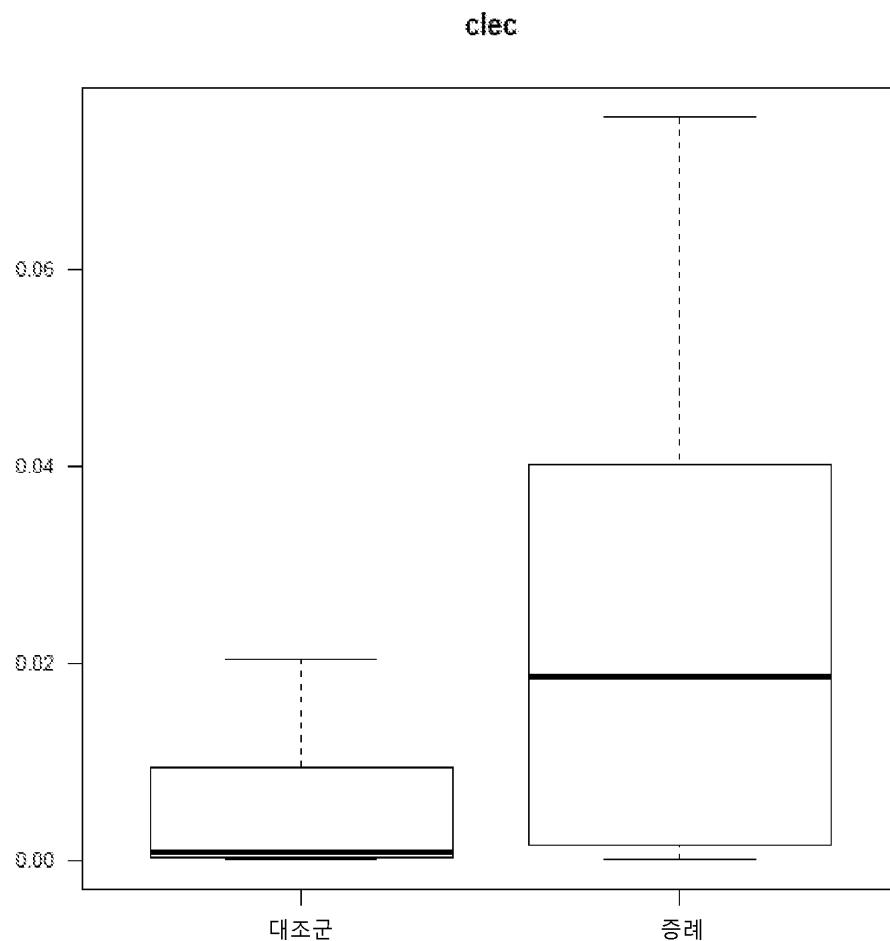


도면7c

cd1d

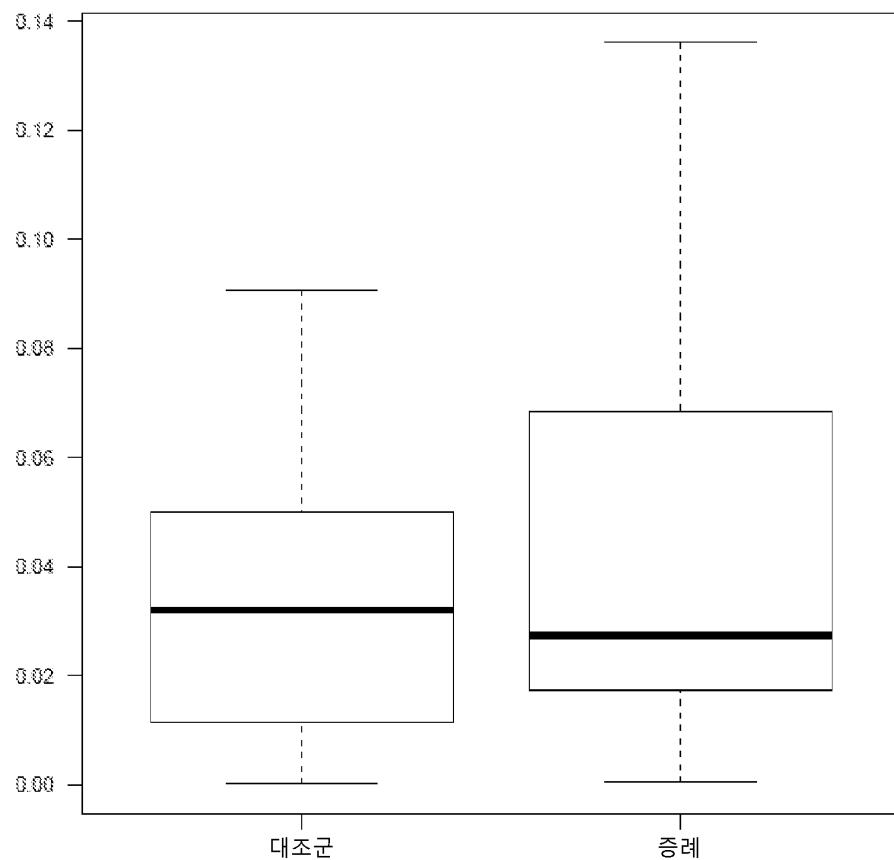


도면7d

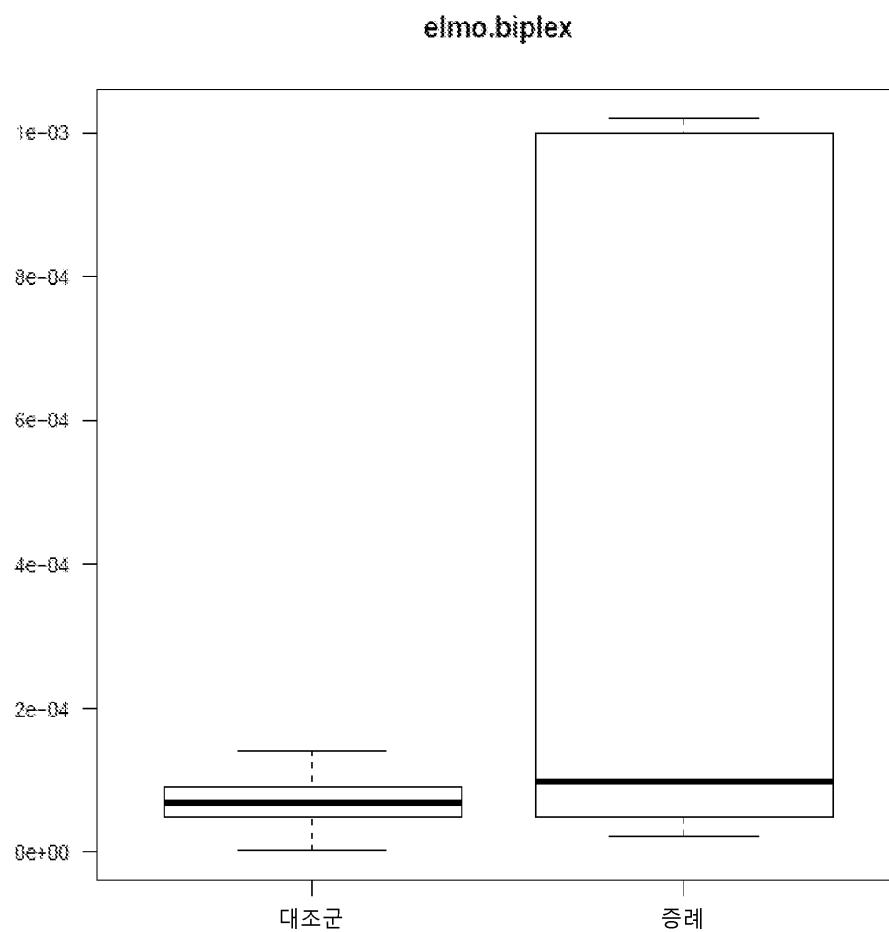


도면7e

cyp26

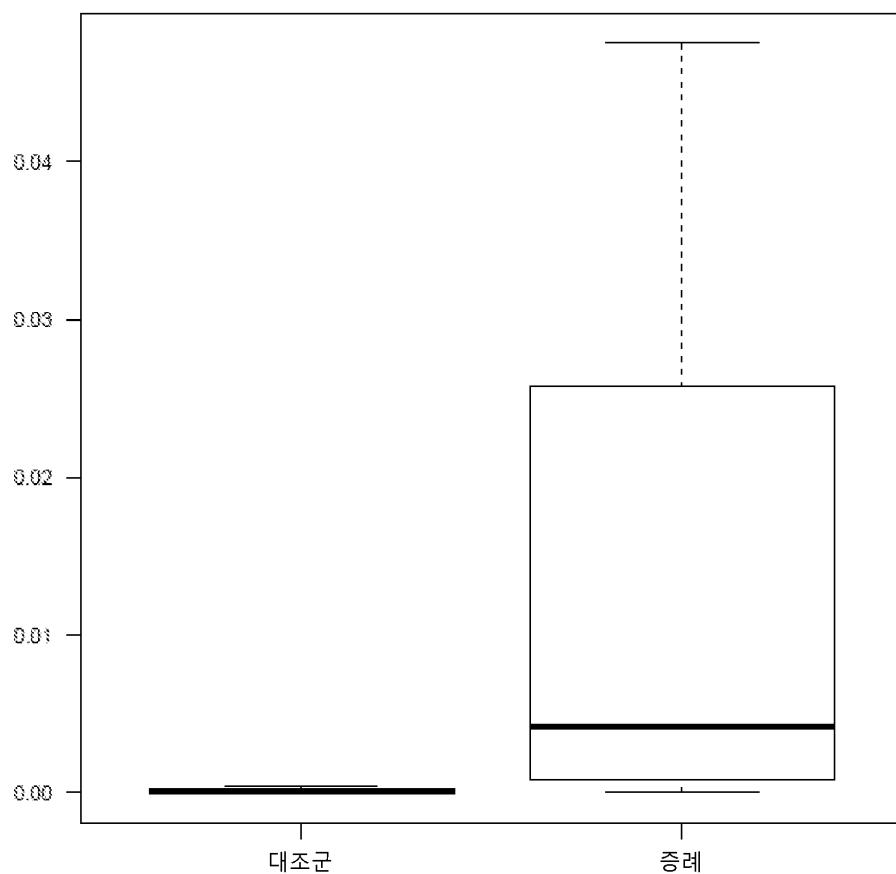


도면7f



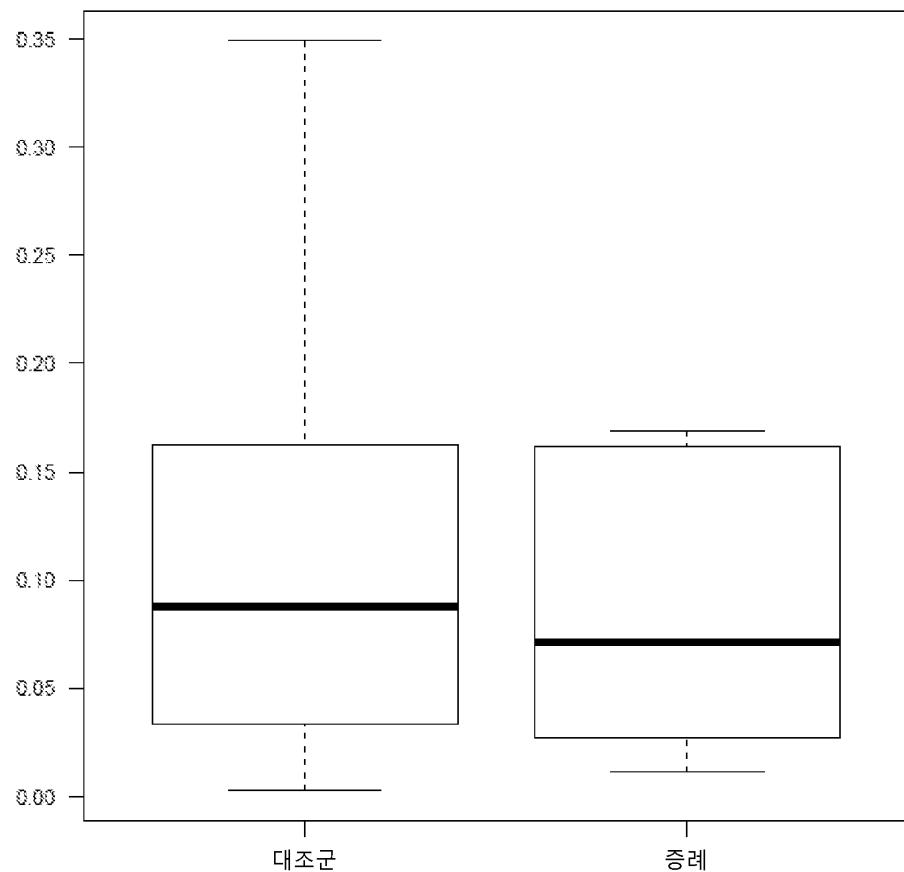
도면7g

elmo

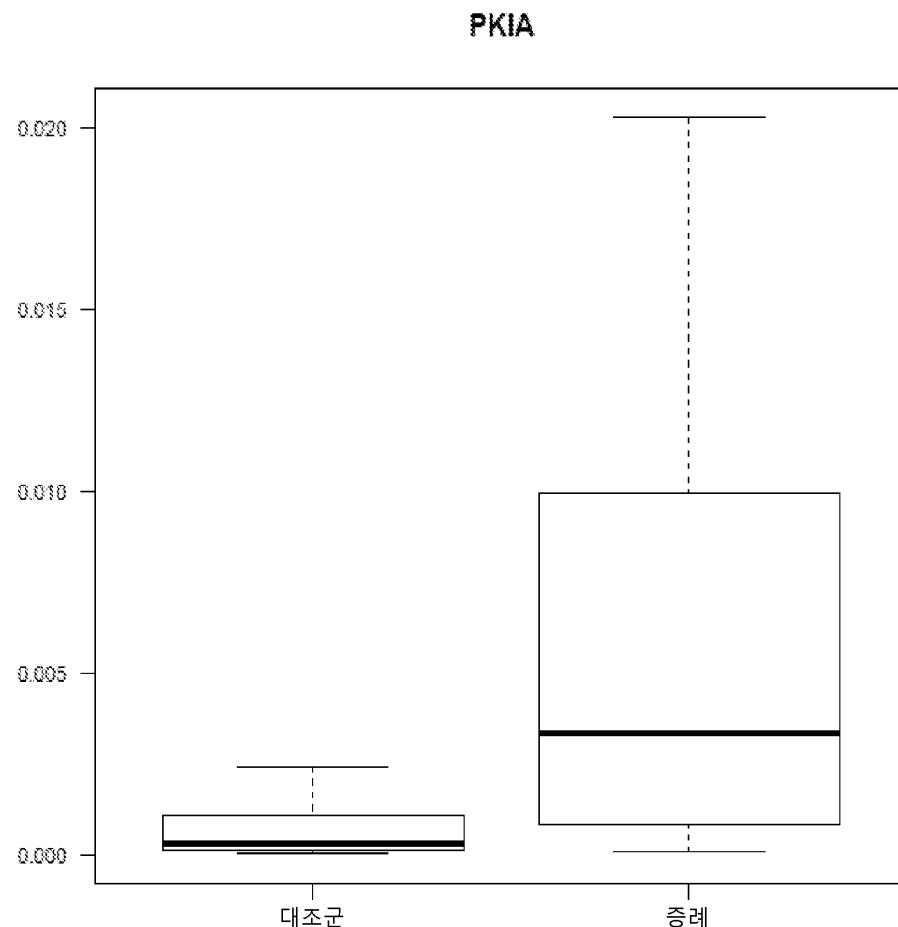


도면7h

LRRC4

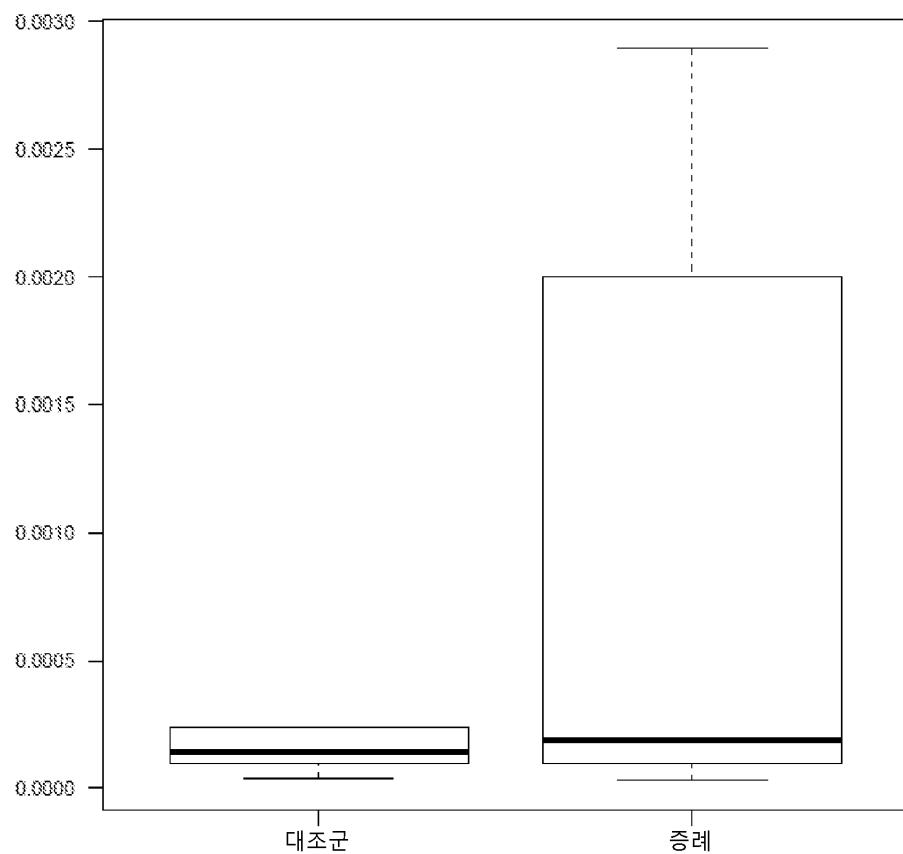


도면7i



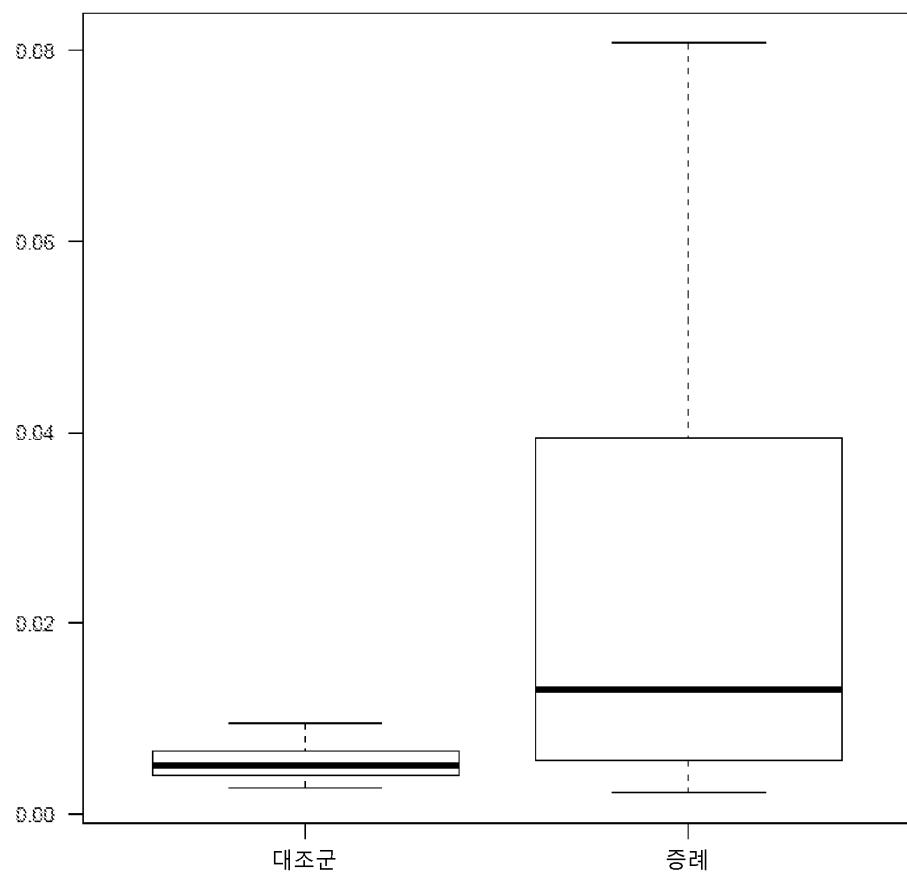
도면7j

ppp25rc



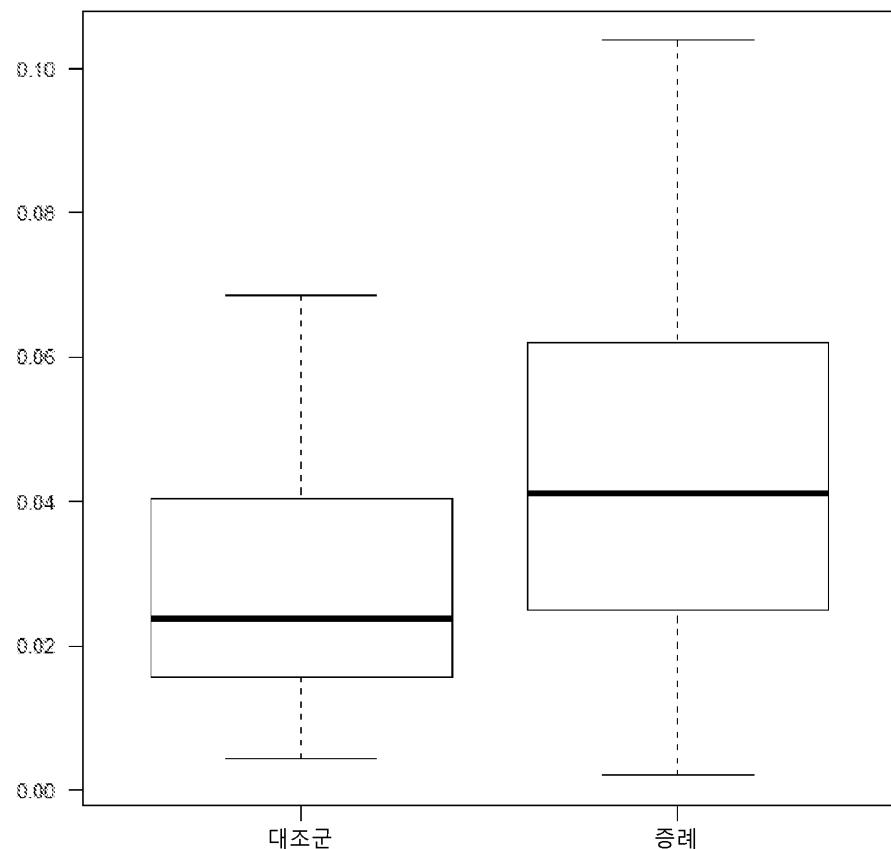
도면7k

sfmbt2

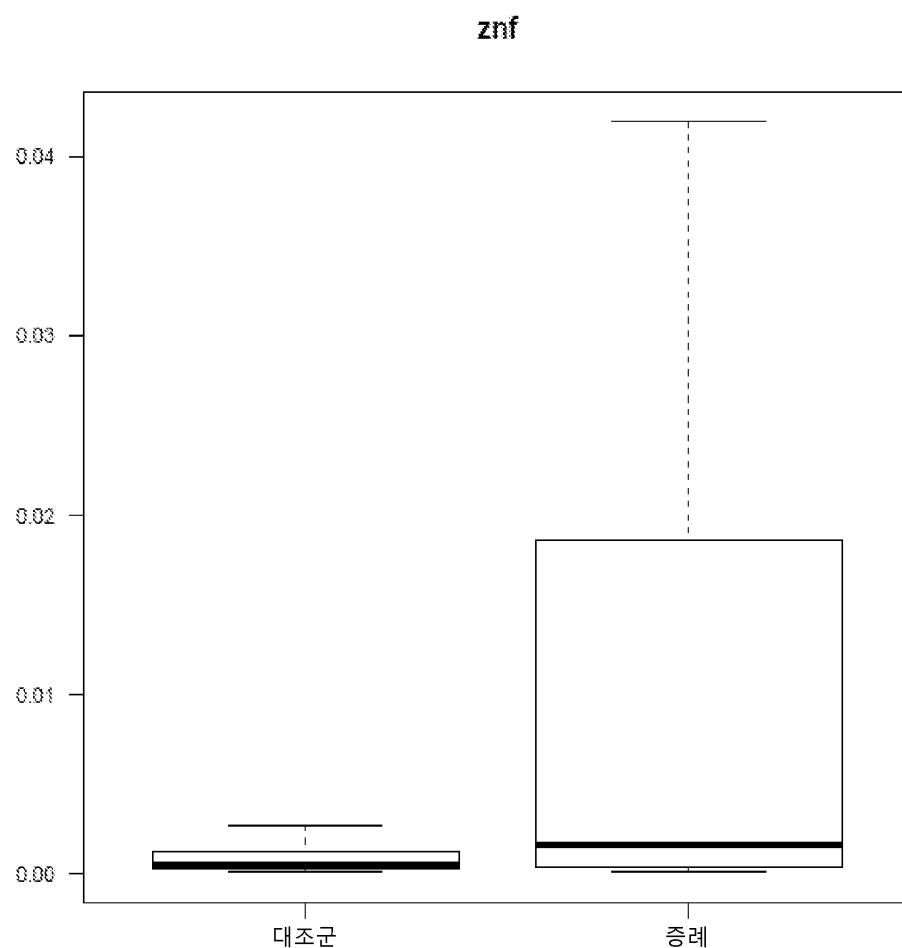


도면71

st8ia1



도면7m



서 열 목 록

SEQUENCE LISTING

<110> MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH

EXACT SCIENCES CORPORATION

<120> DETECTING GASTRIC NEOPLASM

<130> EXCTM-34505/WO-1/ORD

<150> US 62/212,221

<151> 2015-08-31

<160> 119

<170> PatentIn version 3.5

<210> 1

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

| | |
|---------------------------|----|
| <400> 1 | |
| agtttggcgt agtcggtaga tc | 22 |
| <210> 2 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| | |
| <220><223> synthetic | |
| <400> 2 | |
| gcgcgcataat accgaataaa cg | 22 |
| <210> 3 | |
| <211> 17 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| | |
| <220><223> synthetic | |
| <400> 3 | |
| gcgcgttagcg gcgttgc | 17 |
| <210> 4 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| | |
| <220><223> synthetic | |
| <400> 4 | |
| cccatatcgcc cgacgtaa | 19 |
| <210> 5 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| | |
| <220><223> synthetic | |
| <400> 5 | |
| aaaacgttt tagtatttg gtcgt | 25 |
| <210> 6 | |
| <211> 25 | |
| <212> DNA | |

| | |
|-----------------------------|----|
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 6 | |
| ataaaaaaac ctccctcgaa cccgc | 25 |
| <210> 7 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 7 | |
| ggagggagtt taacgggtta ggtac | 25 |
| <210> 8 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 8 | |
| cgtaaacaaa aaacaacgac gacg | 24 |
| <210> 9 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 9 | |
| gattttgttc gtcgttagtg c | 21 |
| <210> 10 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 10 | |
| tctctaaacc cgcaacga | 19 |
| <210> 11 | |
| <211> 24 | |

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 11

tgtggaatcg gggtttgtgt tcgc

24

<210> 12

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 12

cccacccaaac acaaaaaatc cgacg

25

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 13

gacgttggc gtcgggttcg ttc

23

<210> 14

<211> 24

<212> DNA

<213>

Artificial Sequence

<220><223> synthetic

<400> 14

aaaaaccctc cgctaccact tcgc

24

<210> 15

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 15

ggttttgg ttatcgga atcg

25

| | |
|---------------------------------|----|
| <210> 16 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 16 | |
| tataaaaacg cgcgtaatca acgct | 25 |
| <210> 17 | |
| <211> 30 | |
| | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 17 | |
| tgtttcgcg gtcgttatat attacgtcgt | 30 |
| <210> 18 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 18 | |
| gaactatccc cgaactccga ctcga | 25 |
| <210> 19 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 19 | |
| tttagggaag taaagcgtcg tttc | 25 |
| | |
| <210> 20 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |

| | |
|-------------------------------------|----|
| <400> 20 | |
| aacgacgtct cgataacctac ga | 22 |
| <210> 21 | |
| <211> 34 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 21 | |
| ttatatttt cgttttagt aatttcgcgt tagc | 34 |
| <210> 22 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 22 | |
| gaaaaccgcg cgaaacattt cga | 23 |
| <210> 23 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 23 | |
| tgtttcgcg gtcgttatat attacgtcgt | 30 |
| <210> 24 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 24 | |
| gaactatccc cgaactccga ctcgaa | 25 |
| <210> 25 | |
| <211> 31 | |
| <212> DNA | |
| <213> Artificial Sequence | |

| | |
|----------------------------------|----|
| <220><223> synthetic | |
| <400> 25 | |
| cgggttttag cgatgttat tttagttcg t | 31 |
| | |
| <210> 26 | |
| <211> 26 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 26 | |
| cctttcggt cgataaaat ttcggt | 26 |
| <210> 27 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 27 | |
| cggggtaaa aaggtagcgtatgtatac | 30 |
| <210> 28 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 28 | |
| cgaaaaattc gaaacgaaac gtc | 23 |
| | |
| <210> 29 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 29 | |
| agtttggcgt agtcggtaga tc | 22 |
| <210> 30 | |
| <211> 22 | |

| | |
|----------------------------|----|
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 30 | |
| gcgcgcaaat accgaataaa cg | 22 |
| <210> 31 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 31 | |
| gtttaattt cggttcgtc gtc | 23 |
| <210> 32 | |
| <211> 18 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 32 | |
| cgctacgaaa cactccga | 18 |
| <210> 33 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 33 | |
| ttttgttggg tgggtatttt cgcg | 24 |
| <210> 34 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 34 | |
| acaaccctca aaaccctcg atatc | 25 |

| | |
|------------------------------|----|
| <210> 35 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 35 | |
| tttatttatg gattacgtt ttagcga | 27 |
| <210> 36 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 36 | |
| gacgaccgaa cgatcacgca | 20 |
| <210> 37 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 37 | |
| gcgtcgtag tagtacgaag c | 21 |
| <210> 38 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 38 | |
| gcacctcaac gaaaacacccg a | 21 |
| <210> 39 | |
| <211> 17 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 39 | |

| | |
|-----------------------------|----|
| gcgcgttagcg gcgttgc | 17 |
| <210> 40 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 40 | |
| cccatatcgcccgacgtaa | 19 |
| | |
| <210> 41 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 41 | |
| gtcggtttttggcggttttagaggc | 24 |
| <210> 42 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 42 | |
| taatccaaa cgaaccgccc | 20 |
| <210> 43 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 43 | |
| ggttttttgg ttatccgga atcgat | 25 |
| | |
| <210> 44 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |

| | |
|-------------------------------|----|
| <220><223> synthetic | |
| <400> 44 | |
| tataaaaaacg cgcgtaatca acgct | 25 |
| <210> 45 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 45 | |
| ttgagatgtt gggtaaggc gattc | 25 |
| <210> 46 | |
| <211> 26 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 46 | |
| cgctaacgcg aaaaaataat tcgacg | 26 |
| <210> 47 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 47 | |
| cggtttatt tattatgatt cgtagcgg | 28 |
| <210> 48 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 48 | |
| cgcactaccct aaacaacgca tcgc | 24 |
| <210> 49 | |
| <211> 24 | |
| <212> DNA | |

| | | |
|----------------------------|--|----|
| <213> Artificial Sequence | | |
| <220><223> synthetic | | |
| <400> 49 | | |
| cgactaccct aaacaacgca tcgc | | 24 |
| <210> 50 | | |
| <211> 24 | | |
| <212> DNA | | |
| <213> Artificial Sequence | | |
| <220><223> synthetic | | |
| <400> 50 | | |
| ccggccgaat actcgatcaa ctcg | | 24 |
| <210> 51 | | |
| <211> 19 | | |
| <212> DNA | | |
| <213> Artificial Sequence | | |
| <220><223> synthetic | | |
| <400> 51 | | |
| ggtcgcgtcg tttcgatc | | 19 |
| <210> 52 | | |
| <211> 17 | | |
| <212> DNA | | |
| <213> Artificial Sequence | | |
| <220><223> synthetic | | |
| <400> 52 | | |
| gccgcaaacg ccgacga | | 17 |
| <210> 53 | | |
| <211> 23 | | |
| <212> DNA | | |
| <213> Artificial Sequence | | |
| <220><223> synthetic | | |
| <400> 53 | | |
| tgtatatttc gaggcggtt cggt | | 23 |
| <210> 54 | | |

| | |
|--------------------------------|----|
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 54 | |
| cgcccccaaa ataaaaaaac gaa | 23 |
| <210> 55 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 55 | |
| ttttcgttc gtttttat ttttcga | 28 |
| <210> 56 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 56 | |
| aacgcatacc atcgacttca actacgac | 28 |
| <210> 57 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 57 | |
| cggtttcgt tcgttttc g | 21 |
| <210> 58 | |
| <211> 18 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 58 | |

| | |
|--------------------------------------|----|
| ccgccttcta cgcgacta | 18 |
| | |
| <210> 59 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 59 | |
| gtattttaa tatttgtga gtgcg | 25 |
| | |
| <210> 60 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 60 | |
| tactaaaacg tacgataaac atcg | 25 |
| | |
| <210> 61 | |
| <211> 33 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 61 | |
| tcgatttat ttttgtgac gttgttagatt cgc | 33 |
| | |
| <210> 62 | |
| <211> 33 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 62 | |
| aaaaaaacta aaaaacgaca aaaaaacccg acg | 33 |
| | |
| <210> 63 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |

<220><223> synthetic
<400> 63
ggagggagtt taacgggtta ggtac 25
<210> 64
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> synthetic
<400> 64
cgtaaacaaa aaacaacgac gacg 24

<210> 65
<211> 25
<212> DNA
<213> Artificial Sequence
<220><223> synthetic
<400> 65
attgcgtcgt tgtttaggta acgta 25
<210> 66
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> synthetic
<400> 66
cgaaaatacc tatctccctcg act 23
<210> 67
<211> 28
<212> DNA
<213> Artificial Sequence
<220><223> synthetic
<400> 67
ggaacgagtg atagtcggat agttcgtc 28

<210> 68

| | |
|-----------------------------|----|
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 68 | |
| cgcccgaaaa cgaccgg | 18 |
| <210> 69 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 69 | |
| gggagtgagg gtagggcggtt c | 21 |
| <210> 70 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 70 | |
| ctcgcaaccc cttcgaatta acccg | 25 |
| <210> 71 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 71 | |
| gcgagtttatcg cgagtattcg g | 21 |
| <210> 72 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 72 | |
| ctaatcaacg caaatcccg c | 21 |
| <210> 73 | |

| | |
|---------------------------------|----|
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 73 | |
| ccacggacgc gcacggcc ac | 22 |
| | |
| <210> 74 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 74 | |
| cgtatgttt tagaaggttt ttagggaaat | 30 |
| <210> 75 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 75 | |
| cgcacgtctcg atacctacga | 20 |
| <210> 76 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 76 | |
| ccacggacga cgaaaacgac gc | 22 |
| | |
| <210> 77 | |
| <211> 17 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 77 | |

| | |
|---------------------------|----|
| gcgggagttt ggcgtag | 17 |
| <210> 78 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 78 | |
| cgcgaaata ccgaataaac g | 21 |
| <210> 79 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 79 | |
| ccacggacgg tcggtagatc gc | 22 |
| <210> 80 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 80 | |
| gtttcgggt ttcggtttt g | 21 |
| <210> 81 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 81 | |
| cgctaaaaat acgacgcccc | 20 |
| <210> 82 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |

| | |
|---------------------------|----|
| <400> 82 | |
| ccacggacgg cgtttattc gc | 22 |
| | |
| <210> 83 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 83 | |
| gcgtttaggt tggtcggag | 19 |
| <210> 84 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 84 | |
| acgcactcaa cctacgaac | 19 |
| <210> 85 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 85 | |
| ccacggacgc cgaaaaacta cc | 22 |
| | |
| <210> 86 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 86 | |
| agagttcgac gtttagatgg ag | 22 |
| <210> 87 | |
| <211> 21 | |
| <212> DNA | |

| | |
|-----------------------------|----|
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 87 | |
| tcgaaaccta tatccgcgaa c | 21 |
| <210> 88 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 88 | |
| ccacggacgg cggttttcg ac | 22 |
| | |
| <210> 89 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 89 | |
| tttgttttt tgattagtg tttaaga | 27 |
| <210> 90 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 90 | |
| caccaacctc ataaccttat c | 21 |
| <210> 91 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 91 | |
| gacgcggaga tagtgttg gc | 22 |
| | |
| <210> 92 | |

| | |
|----------------------------|----|
| <211> 29 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 92 | |
| ttcgattta ttttgttgcgtttaga | 29 |
| <210> 93 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 93 | |
| acgacaaaaa aacccgacga | 20 |
| <210> 94 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 94 | |
| cggccaggat tcgcgtgggt c | 21 |
| <210> 95 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 95 | |
| cgttcgcgtt atttatttcg gcg | 23 |
| <210> 96 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 96 | |
| gctcctaatt ctcataacg tcgt | 24 |

| | |
|---------------------------|----|
| <210> 97 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 97 | |
| cgccgaggc ggcgtttgc c | 21 |
| | |
| <210> 98 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 98 | |
| gaagagtgcg tagtttagag g | 21 |
| <210> 99 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 99 | |
| catatcgccc gacgtaaaaa cc | 22 |
| <210> 100 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 100 | |
| cgccgaggct cgcgaaacgc c | 21 |
| | |
| <210> 101 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |

| | |
|---------------------------|----|
| <400> 101 | |
| aaatatgtgg ttcgtggcgt t | 21 |
| <210> 102 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 102 | |
| acgcaacaac gcgaaaaac | 19 |
| <210> 103 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 103 | |
| cgcgaggcgc acgaaaaacg c | 21 |
| <210> 104 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 104 | |
| cgttcgcgtt atttatttcg gcg | 23 |
| <210> 105 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 105 | |
| gctcttaatt ctcataacg tcgt | 24 |
| <210> 106 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |

| | |
|--------------------------------|----|
| <220><223> synthetic | |
| <400> 106 | |
| cgccgaggc ggcgtttgc c | 21 |
| | |
| <210> 107 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 107 | |
| gcgttaattt cgcgaggta | 19 |
| <210> 108 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 108 | |
| acaatactct tataattaa cgccgctc | 28 |
| <210> 109 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 109 | |
| cgcggaggag gcgacggagg c | 21 |
| | |
| <210> 110 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 110 | |
| agccggttt ccggctgaga ctccgcgtc | 29 |
| <210> 111 | |
| <211> 32 | |

| | |
|------------------------------------|----|
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 111 | |
| tctagccgtt ttccggctg agactccgctc | 32 |
| <210> 112 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 112 | |
| agccggttt ccggctgaga cgtccgtgg | 29 |
| <210> 113 | |
| <211> 32 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 113 | |
| tctagccgtt ttccggctg agacgtccgt gg | 32 |
| <210> 114 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 114 | |
| agccggttt ccggctgaga cctcggcg | 28 |
| <210> 115 | |
| <211> 31 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 115 | |
| tctagccgtt ttccggctg agacctcggcg | 31 |

| | |
|---------------------------|---|
| <210> 116 | |
| <211> 8 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 116 | |
| cgccgagg | 8 |
| <210> 117 | |
| <211> 9 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 117 | |
| gacgcggag | 9 |
| <210> 118 | |
| <211> 9 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 118 | |
| ccacggacg | 9 |
| <210> 119 | |
| <211> 8 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> synthetic | |
| <400> 119 | |
| gcgcgtcc | 8 |