



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년03월29일

(11) 등록번호 10-2380690

(24) 등록일자 2022년03월25일

(51) 국제특허분류(Int. Cl.)

C12Q 1/6886 (2018.01)

(52) CPC특허분류

C12Q 1/6886 (2022.01)

C12Q 2600/154 (2013.01)

(21) 출원번호 10-2018-7032924

(22) 출원일자(국제) 2017년04월13일

심사청구일자 2020년04월02일

(85) 번역문제출일자 2018년11월13일

(65) 공개번호 10-2018-0133497

(43) 공개일자 2018년12월14일

(86) 국제출원번호 PCT/US2017/027439

(87) 국제공개번호 WO 2017/180886

국제공개일자 2017년10월19일

(30) 우선권주장

62/322,610 2016년04월14일 미국(US)

(56) 선행기술조사문헌

Cancer J, 18(6): 492-501 (21012.11.)*

World J Gastroenterol, 21(40): 11387-11395

(2015.10.28.)*

US20140274748 A1

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

메이오 파운데이션 포 메디칼 에듀케이션 앤드 리
써치미국 55905 미네소타주 록체스터 퍼스트 스트리트
에스.더블유. 200

(72) 발명자

알퀴스트, 데이비드 에이.

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

테일러, 윌리엄 알.

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

(뒷면에 계속)

(74) 대리인

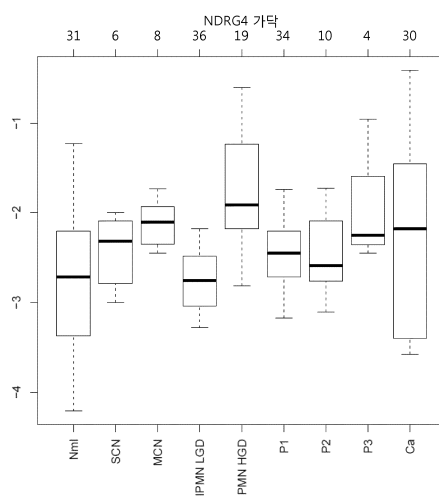
특허법인 무한

전체 청구항 수 : 총 7 항

심사관 : 이준혁

(54) 발명의 명칭 **췌장 고도 이형성증의 검출방법****(57) 요약**

본 개시내용은 췌장 고등급 이형성증의 스크리닝을 위한 기술, 특히 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 존재를 검출하는 방법, 조성물, 및 관련 용도에 관한 것이나, 이로만 제한되는 것은 아니다.

대표도 - 도1a

(72) 발명자

키셀, 존 비.

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

애플, 트레이시 씨.

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

마호니, 더글라스 더블유.

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

명세서

청구범위

청구항 1

삭제

청구항 2

삭제

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

생물학적 시료를 특성화하는 방법으로서, 상기 방법은

인간의 생물학적 시료에서 22개 이하의 유전자에 대한 CpG 부위의 메틸화 수준을,

상기 생물학적 시료 중의 유전체 DNA를 바이셀파이트로 처리하는 단계;

상기 바이셀파이트-처리된 유전체 DNA를 상기 선택된 22개 이하의 유전자 각각에 대한 프라이머 세트를 사용하여 증폭시키는 단계; 및

상기 CpG 부위의 메틸화 수준을 메틸화-특이 PCR, 정량적 메틸화-특이 PCR, 메틸화-민감성 DNA 제한 효소 분석, 정량적 바이셀파이트 파이로시퀀싱, 또는 바이셀파이트 유전체 시퀀싱 PCR에 의해 결정하는 단계;를 통해 측정하는 것을 포함하고,

상기 22개 이하의 유전자는 BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781로부터 선택되며,

상기 22개 이하의 유전자는 적어도 AK055957 및 ST8SIA1를 포함하고,

상기 22개 이하의 유전자는 적어도 하기 유전자들: DLX4, EMX1, FRMD4A, LRRC4, SP9, TBX15, 및 ZNF781

중 하나를 포함하는 것인, 생물학적 시료를 특성화하는 방법.

청구항 45

제44항에 있어서,

상기 생물학적 시료는 채장 조직 및 채장 낭액 중에서 선택된 하나 이상을 포함하는 것인, 생물학적 시료를 특성화하는 방법.

청구항 46

제44항에 있어서,

상기 생물학적 시료는 대변 시료, 혈액 시료 및 혈액 분획 시료 중 선택된 하나 이상을 포함하는 것인, 생물학적 시료를 특성화하는 방법.

청구항 47

제44항에 있어서,

상기 22개 이하의 유전자에 대한 프라이머 세트는 하기의 군에서 선택되는 것인, 생물학적 시료를 특성화하는 방법:

BMP3에 대한 SEQ ID NO: 45 및 46로 이루어진 프라이머 세트;

NDRG4에 대한 SEQ ID NO: 47 및 48로 이루어진 프라이머 세트;

ABCB1에 대한 SEQ ID NO: 1 및 2로 이루어진 프라이머 세트;

C13ORF18에 대한 SEQ ID NO: 7 및 8로 이루어진 프라이머 세트;

CD1D에 대한 SEQ ID NO: 43 및 44로 이루어진 프라이머 세트;

CLEC11A에 대한 SEQ ID NO: 9 및 10로 이루어진 프라이머 세트;

DLX4에 대한 SEQ ID NO: 11 및 12로 이루어진 프라이머 세트;

ELM01에 대한 SEQ ID NO: 13 및 14로 이루어진 프라이머 세트;

EMX1에 대한 SEQ ID NO: 15 및 16로 이루어진 프라이머 세트;

FER1L4에 대한 SEQ ID NO: 17 및 18로 이루어진 프라이머 세트;

FRMD4A에 대한 SEQ ID NO: 19 및 20로 이루어진 프라이머 세트;

GRIN2D에 대한 SEQ ID NO: 21 및 22로 이루어진 프라이머 세트;

HOXA1에 대한 SEQ ID NO: 23 및 24로 이루어진 프라이머 세트;

LRRC4에 대한 SEQ ID NO: 25 및 26로 이루어진 프라이머 세트;

PRKCB에 대한 SEQ ID NO: 29 및 30로 이루어진 프라이머 세트;

SP9에 대한 SEQ ID NO: 31 및 32로 이루어진 프라이머 세트;
 ST6GAL2에 대한 SEQ ID NO: 33 및 34로 이루어진 프라이머 세트;
 TBX15에 대한 SEQ ID NO: 37 및 38로 이루어진 프라이머 세트;
 VWC2에 대한 SEQ ID NO: 39 및 40로 이루어진 프라이머 세트;
 ZNF781에 대한 SEQ ID NO: 41 및 42로 이루어진 프라이머 세트
 AK055957에 대한 SEQ ID NO: 3 및 4로 이루어진 프라이머 세트; 및
 ST8SIA1에 대한 SEQ ID NO: 35 및 36로 이루어진 프라이머 세트.

청구항 48

제44항에 있어서,
 상기 CpG 부위는 암호화 영역 또는 조절 영역에 존재하는 것인, 생물학적 시료를 특성화하는 방법.

청구항 49

제44항에 있어서,
 상기 22개 이하의 유전자에 대한 CpG 부위의 메틸화 수준을 측정하는 단계는 상기 CpG 부위의 메틸화 점수를 결정하는 단계와 상기 CpG 부위의 메틸화 빈도를 결정하는 단계로 이루어진 군으로부터 선택된 결정을 포함하는 것인, 생물학적 시료를 특성화하는 방법.

청구항 50

생물학적 시료를 특성화하는 방법으로서, 상기 방법은
 인간의 생물학적 시료에서 3개 이상의 유전자에 대한 CpG 부위의 메틸화 수준을,
 상기 생물학적 시료 중의 유전체 DNA를 바이설파이트로 처리하는 단계;
 상기 바이설파이트-처리된 유전체 DNA를 상기 선택된 3개 이상의 유전자 각각에 대한 프라이머 세트를 사용하여 증폭시키는 단계; 및
 상기 CpG 부위의 메틸화 수준을 메틸화-특이 PCR, 정량적 메틸화-특이 PCR, 메틸화-민감성 DNA 제한 효소 분석, 정량적 바이설파이트 파이로시퀀싱, 또는 바이설파이트 유전체 시퀀싱 PCR에 의해 결정하는 단계;를 통해 측정하는 것을 포함하고,
 상기 3개 이상의 유전자는 BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781로부터 선택되며,
 상기 3개 이상의 유전자는 적어도 AK055957 및 ST8SIA1를 포함하고,
 상기 3개 이상의 유전자는 적어도 하기 유전자들: DLX4, EMX1, FRMD4A, LRRC4, SP9, TBX15, 및 ZNF781 중 하나를 포함하는 것인, 생물학적 시료를 특성화하는 방법.

발명의 설명

기술 분야

[0001]

본 개시내용은 췌장 고도 이형성증의 스크리닝을 위한 기술, 특히 췌장 고도 이형성증 [췌장관내 유두상 점액 종양 고도 이형성증(intraductal papillary mucinous neoplasm high grade dysplasia, IPMN-HGD), 췌장 상피내

종양 3형(pancreatic intraepithelial neoplasia 3, PanIN-3), 또는 췌관 선암종(pancreatic ductal adenocarcinoma, PDAC)]의 존재를 검출하는 방법, 조성물, 및 관련 용도에 관한 것이나, 이로만 제한되는 것은 아니다.

배경 기술

[0002] 의학적 진보에도 불구하고, 췌장암은 가장 치명적인 질병 중 하나이다. 2011년, 미국에서 췌장암으로 진단받은 사람들은 44,030명으로 추산되었고, 이 중 약 37,660명의 사람들이 사망했다. 췌장암은 미국 남성들과(폐암, 전립선암, 및 대장암에 이어서) 여성들(폐암, 유방암, 및 대장암에 이어서) 사이에서 4번째로 가장 흔한 암-관련 사망의 원인이다. 남성과 여성 모두에서 거의 동일한 발병률로 70대 및 80대에서 최고의 발병률이 발생한다. 폐암, 대장암, 유방암, 및 전립선암이 2003년 이래로 미국에서 감소하고 있는 반면에, 같은 시기 동안 췌장암은 증가해오고 있다. 추정된 사망자가 82,700 (남성) 및 79,100 (여성)인 반면에, 2011년 전세계적으로 선진국에서 추정된 췌장암의 새로운 사례는 84,200명 (남성) 및 80,900명 (여성)이다. 불행하게도, 대부분의 증상을 보이는 환자는 치료가 불가능하다. 췌장암 환자에 대한 예후는 합쳐진 모든 단계에 대해 6%의 5년 상대 생존율(5-year relative survival rate)으로 매우 좋지 않다. 이는 진단 시기에서 질병이 말기이기 때문이다.

[0003] 이러한 환자들의 생존율을 향상시키기 위해 췌장암의 조기 검출이 매우 필요하다.

발명의 내용

[0004] 현재 미국에서 4번째로 가장 흔한 암으로 인한 사망의 원인인 췌장암(Siegel, R., et al., CA Cancer J Clin, 2014. 64(1): p. 9-29를 참고)은 전체 5년 생존(overall 5-year survival)이 5% 미만으로, 모든 암들 중 가장 치명적이다(Wolfgang, C.L., et al., CA: a cancer journal for clinicians, 2013. 63(5): p. 318-48을 참고). 매년, 진단된 췌장암의 수는 그 질병에 대한 사망자들의 수와 거의 동일하다. 치료법을 향상시키려는 노력에도 불구하고, 이 참담한 결과는 지난 30년동안 변함없이 왔다(Wolfgang, C.L., et al., CA: a cancer journal for clinicians, 2013. 63(5): p. 318-48을 참고). 중대한 관심사인 췌장암은 증가 추세에 있다. 미국에서, 췌장암은 2030년까지 암으로 인한 사망의 두번째로 가장 흔한 원인이 될 것으로 예측된다 (Rahib, L., et al., Cancer research, 2014. 74(14): p. 4006을 참고). 세계의 다른 지역에서는, 다른 통상적인 암(예를 들어, 유방암, 결장암, 및 전립선암)의 발병률 및 사망률이 감소하는 추세인 반면에, 췌장암은 악화된 예측이 있는 유일한 암 중 하나이다(Malvezzi, M., et al., Annals of oncology : official journal of the European Society for Medical Oncology / ESMO, 2014을 참고).

[0005] 효과적이거나 광범위하게 사용되는 스크리닝 도구가 없기 때문에, 췌장암은 대개 증상 후 및 말기에서 진단된다. 현재, 사례의 85%가 전이성 질병과 존재하며 일반적으로 진단 후 6개월 미만 동안 생존한다 (Wolfgang, C.L., et al., CA: a cancer journal for clinicians, 2013. 63(5): p. 318-48을 참고). 분명한 국소적으로 진행된 질병 또는 원격 전이 없이 징후적으로 존재하고 절제술(resection)로 진행하는 소수에 대해서, 전체 생존은 5년 차에서 20% 미만이다(Wolfgang, C.L., et al., CA: a cancer journal for clinicians, 2013. 63(5): p. 318-48을 참고). 희망적으로, 초기 종양으로 발견된 사례의 소규모 하위 집단이 최선의 결과-2 cm 미만의 병변으로 30-60%의 5년 생존율 및 1 cm 미만의 병변으로 75%가 넘는 5년 생존율-를 가진다(Sohn, T., et al., Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract, 2000. 4(6): p. 567-579; Furukawa, H., et al., Cancer, 1996. 78(5): p. 986-990; Shimizu, Y., et al., Journal of gastroenterology and hepatology, 2005. 20(10): p. 1591-1594; Ishikawa, O., et al., Hepato-Gastroenterology, 1999. 46(25): p. 8-15; Tsuchiya, R., et al., Annals of Surgery, 1986. 203(1): p. 77-81을 참고). 더욱이, 증상이 없는 개인에서 우연히 발견된 초기 췌장암은 치료적 절제술(curative resection)로 높은 생존율을 갖는 것으로 보이며, 이는 소규모 일화적 시리즈 (small anecdotal series)에 기초한다 (Yeo, C.J. and J.L. Cameron, Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie, 1998. 383(2): p. 129-33; Okano, K. and Y. Suzuki, World journal of gastroenterology : WJG, 2014. 20(32): p. 11230-11240을 참고). 따라서, 초기 췌장암의 증상 전 검출(pre-symptomatic detection)을 높은 치료율로 이끌 수 있는 확실한 증거가 있으며, 이는 효과적인 스크리닝 도구에 대한 필요를 두드러지게 강조할 수 있다.

[0006] 산발형(sporadic form)의 췌장암은 모든 사례 중 90% 넘게 차지하며, 10% 미만의 사례는 근본적인 유전적 또는 후천적 소인성 질병을 포함한다. 이러한 고위험의 후천적인 질병은 특정한 회귀 유전적 질환, 만성 췌장염

(chronic pancreatitis), 우연히 발견된 낭포성 췌장 전암(cystic pancreatic precancers) (특히, 췌장관내 유두상 점액 종양(intraductal papillary mucinous neoplasms, IPMNs)), 비만, 최근 발병한 당뇨병(recent-onset diabetes mellitus)을 포함한다 (Chakraborty, S., et al., Biochimica et biophysica acta, 2011. 1815(1): p. 44-64을 참고). 다양한 내시경적인 및 영상화 방법을 이용하는 스크리닝과 감시 프로그램은 이러한 알려진 유전적이거나 후천적인 위험 인자가 있는 환자에게 일관성 없이 적용된다. 대다수의 췌장암이 명백한 위험 인자가 없는 개인에서 발생하기 때문에, 고위험 하위 집단에서 선택적 스크리닝은 전체 암 발병률 또는 사망률에서 영향이 거의 없을 것이다. 췌장암 사망률에 최대한의 영향을 갖기 위해, 효과적인 전 집단 스크리닝 도구(population-wide screening tool)가 요구된다.

[0007] 이 두려운 질병의 막심한 피해를 줄이기 위해, 효과적인 스크리닝 접근방법이 시급히 필요하다. 초기암과 진행성 전암(advanced precancer)의 증상전 검출(pre-symptomatic detection)을 위한 정확하고, 가격이 적정하며, 안전한 스크리닝 도구를 전달할 혁신에 대한 시급한 과제가 있다.

[0008] 본 발명은 이 필요성을 해결한다. 실제로, 본 발명은 고등급 전구 병변(high-grade precursor lesion)이 있는 환자군과 정상적인 췌장 또는 저등급 전구체(low-grade precursor)가 있는 대조군으로부터의 암을 구별하는 신규한 메틸화된 DNA 마커를 제공한다.

[0009] 메틸화된 DNA는 대부분의 종양 유형의 조직에서 바이오마커의 잠재적인 종류로서 연구되어왔다. 많은 경우에서, DNA 메틸 전달효소(methyltransferases)는 유전자 발현의 후성적 제어(epigenetic control)로서 메틸 그룹을 시토신-포스페이트-구아닌 섬 (cytosine-phosphate-guanine island, CpG island) 부위의 DNA에 더했다. 생물학적으로 매력적인 메커니즘에서, 종양 억제 유전자의 프로모터 영역에서 후천적 메틸화 사건은 발현을 침묵시키고, 따라서 종양형성(oncogenesis)에 기여한다고 생각된다. DNA 메틸화는 RNA 또는 단백질 발현보다 화학적 및 생물학적으로 더욱 안정한 진단 도구일 수 있다 (Laird (2010) Nat Rev Genet 11: 191-203). 더욱이, 산발성 결장암(sporadic colon cancer)과 같은 다른 암에서, 메틸화 마커는 우수한 특이도를 제공하고 개별적인 DNA 돌연변이보다 더 광범위하게 정보를 제공하며 민감하다(Zou et al (2007) Cancer Epidemiol Biomarkers Prev 16: 2686-96).

[0010] CpG 섬의 분석은 동물 모델 및 인간 세포주에 적용될 경우 중요한 발견들을 가져왔다. 예를 들어, Zhang과 동료들은 동일한 CpG 섬의 상이한 부분으로부터 증폭산물(amplicon)이 상이한 수준의 메틸화를 가질 수 있음을 발견하였다 (Zhang et al. (2009) PLoS Genet 5: e1000438). 게다가, 메틸화 수준은 고도로 메틸화된 서열과 메틸화 되지 않은 서열 사이에서 이봉 분포(bimodal distribution)를 보였는데, 이는 DNA 메틸 전달효소 활성의 바이너리 스위치 유사 패턴(binary switch-like pattern)을 더욱 지지한다(Zhang et al. (2009) PLoS Genet 5: e1000438). 체내에서 쥐과(murine) 조직 및 시험관에서 세포주의 분석은 고 CpG 밀도 프로모터 [(high CpG density promoter, HCP), 300 염기 쌍 영역내에서 7% 초과 CpG 서열을 갖는 것으로 정의됨]의 약 0.3%만 메틸화 된 것으로 입증하였고, 반면에 저 CpG 밀도[(low CpG density, LCP), 300 염기 쌍 영역내에서 5% 미만의 CpG 서열을 갖는 것으로 정의됨]의 영역은 다양한 조직 특이적 패턴에서 빈번히 메틸화되는 경향이 있었다 (Meissner et al. (2008) Nature 454: 766-70). HCP는 유비퀴틴 하우스키핑 유전자(housekeeping gene)와 고도로 조절되는 발달 유전자(developmental gene)에 대한 프로모터를 포함한다. 50% 넘게 메틸화된 HCP 부위 중에 있는 것은 Wnt 2, NDRG2, SFRP2, 및 BMP3와 같은 몇몇의 확인된 마커들이었다 (Meissner et al. (2008) Nature 454: 766-70).

[0011] 본 발명에 대한 양태를 개발하는 과정에서 수행된 실험은 췌장관내 유두상 점액 종양 고등급 이형성증 (intraductal papillary mucinous neoplasm high grade dysplasia, IPMN-HGD), 췌장 상피내 종양 3형 (pancreatic intraepithelial neoplasia 3, PanIN-3), 또는 췌관 선암종(pancreatic ductal adenocarcinoma, PDAC)을 갖는 개체의 췌장 조직으로부터 DNA 마커의 메틸화 상태를 대조 개체로부터 동일한 DNA 마커의 메틸화 상태와 비교하였다 (예, 각 조직 유형에 대한 IPMN-저등급 이형성증 (IPMN-low grade dysplasia, IPMN-LGD), PanIN-1, 및 PanIN-2를 갖는 개체). 이러한 실험은 고등급 전구체(IPMN-HGD, PanIN-3) 또는 침습성 암(PDAC)이 있는 환자군을 정상적인 췌장 또는 저등급 전구 병변(IPMN-저등급 이형성증 (IPMN-low grade dysplasia, IPMN-LGD), PanIN-1, 및 PanIN-2)이 있는 대조군으로부터 구별하는 메틸화된 DNA 마커 후보를 식별하고 검증하였다(실시예 I, II, IV, 및 V를 참고).

[0012] 따라서, 본 개시내용은 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)의 스크리닝을 위한 기술(예, 감시(surveillance)), 특히 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 존재를 검출하는 방법, 조성물, 및 관련 용도에 관한 것이나, 이로만 제한되는 것은 아니다.

- [0013] 마커 및/또는 마커의 패널(panel) (예, 표 2에 제공된 주석(annotation)을 갖는 염색체 부위)은 개체에서 체장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)을 검출할 수 있는 것으로 식별되었다 (실시예 I, II, IV, 및 V을 참고) (BMP3, NDRG4, ABCB1, AKO55957, C13ORF18, CD1D, CLEC11A, DLX4, ELMO1, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781).
- [0014] 본 명세서에 설명된 바와 같이, 본 발명의 기술은 개체에서 체장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)의 존재를 검출하는 고 구별(high discrimination)이 있는 다수의 메틸화된 DNA 마커와 이의 부분집합 (예, 2, 3, 4, 5, 6, 7, 10, 15, 19, 20, 21, 50, 75, 94 마커의 세트)을 제공한다. 예를 들어, 스크리닝 또는 진단 (예, 암 스크리닝 또는 진단)의 목적을 위한 매질(media) (예, 대장 조직, 대변 시료)을 검정할 경우, 실험은 선택 필터(selection filter)를 후보 마커에 적용하여, 높은 신호 대 잡음비(signal to noise ratio)와 낮은 백그라운드 수준(background level)을 제공하여 높은 특이도를 제공하는 마커를 식별하였다.
- [0015] 일부 양태에서, 본 발명의 기술은 생물학적 시료에서 본 명세서에서 식별된 하나 이상의 마커의 존재 및 메틸화 상태를 평가하는 것과 관련된다. 이러한 마커는 본 명세서에 기술된 바와 같은 하나 이상의 차등 메틸화 영역 (differentially methylated region, DMR)-예를 들어 표 2에 제공된 바와 같은-을 포함한다. 메틸화 상태는 본 발명의 기술의 양태에서 평가된다. 이와 같이, 본 명세서에 제공된 기술은 유전자의 메틸화 상태가 측정된 방법에서 제한되지 않는다. 예를 들어, 일부 양태에서, 메틸화 상태는 유전체 스캐닝 방법(genome scanning method)에 의해 측정된다. 예를 들어, 한 방법이 제한 랜드마크 유전체 스캐닝 (restriction landmark genomic scanning)을 수반하고 (Kawai et al. (1994) Mol. Cell. Biol. 14: 7421-7427) 다른 예시는 메틸화 민감성 임의 프라이머 PCR (methylation-sensitive arbitrarily primed PCR)을 수반한다(Gonzalzo et al. (1997) Cancer Res. 57: 594-599). 일부 양태에서, 특정 CpG 부위에서 메틸화 패턴의 변화는 메틸화 민감성 제한효소 (methylation-sensitive restriction enzyme)를 이용한 유전체 DNA의 절단에 이어서 관심 영역의 서던 분석 (Southern analysis)에 (절단-서던 방법(digestion-Southern method))의해 모니터링할 수 있다. 일부 양태에서, 메틸화 패턴의 변화를 분석하는 것은 PCR 증폭 전 유전체 DNA를 메틸화 민감성 제한 효소를 이용한 절단을 수반하는 PCR에 기초한 공정을 수반한다(Singer-Sam et al. (1990) Nucl. Acids Res. 18: 687). 더욱이, 다른 기법이 DNA의 바이설파이트 처리를 메틸화 분석을 위한 시작점(starting point)으로 이용하는 것으로 보고되고 있다. 이는 메틸화 특이적 PCR(methylation-specific PCR, MSP) (Herman et al. (1992) Proc. Natl. Acad. Sci. USA 93: 9821-9826)과 바이설파이트 전환된 DNA(bisulfite-converted DNA)로부터 증폭된 PCR 생성물의 제한 효소 절단(Sadri and Hornsby (1996) Nucl. Acids Res. 24: 5058-5059; and Xiong and Laird (1997) Nucl. Acids Res. 25: 2532-2534)을 포함한다. PCR 기법은 유전자 변이의 검출 (Kuppuswamy et al. (1991) Proc. Natl. Acad. Sci. USA 88: 1143-1147)과 대립유전자 특이적 발현(allelic-specific expression)의 정량(Szabo and Mann (1995) Genes Dev. 9: 3097-3108; and Singer-Sam et al. (1992) PCR Methods Appl. 1: 160-163)에 대해 발달해왔다. 이러한 기법은 인터널 프라이머(internal primer)를 이용하며, 이는 PCR 생성된 템플레이트 (PCR-generated template)에 어닐(anneal)하고 검정될 단일 뉴클레오티드의 5'을 즉시 종결시킨다. 미국 특허 제7,037,650호에 기술된 "정량적 Ms-SNuPE 검정(quantitative Ms-SNuPE assay)"을 사용하는 방법이 일부 양태에서 사용된다.
- [0016] 메틸화 상태를 평가한 후, 메틸화 상태는 종종 특정 부위(예, 단일 뉴클레오티드, 특정 영역 또는 유전자 좌위, 더 긴 관심서열, 예를 들어, ~100-bp, 200-bp, 500-bp, 1000-bp 까지의 DNA의 하위서열(subsequence) 또는 더 긴)에서 메틸화된 DNA의 개별적 가닥의 백분율 또는 분수로서 그 특정 부위를 포함하는 시료 내 DNA의 총 집단에 관하여 표현된다. 전통적으로, 메틸화되지 않은 핵산의 양은 캘리브레이터(calibrator)를 이용하여 PCR에 의해 결정된다. 그 다음, 알려진 양의 DNA를 바이설파이트 처리하고 수득한 메틸화 특이적 서열을 실시간 PCR(real-time PCR) 또는 다른 지수적 증폭(exponential amplification), 예를 들어, QuARTS 검정 (예를 들어, 미국특허 제8,361,720호; 미국특허 제8,916,344호; 및 미국특허 제2012/0122088호와 2012/0122106호에 의해 제공됨)을 이용하여 결정한다.
- [0017] 예를 들어, 일부 양태에서, 방법은 메틸화되지 않은 표적에 대한 표준 곡선(standard curve)을 외부 표준 (external standard)을 이용하여 생성하는 단계를 포함한다. 표준 곡선은 최소 둘 이상의 점(point)으로부터 제작되고 알려진 정량 표준(quantitative standard)에 대한 메틸화되지 않은 DNA를 위한 실시간 Ct 값에 관련된다. 그 다음, 메틸화된 표적에 대한 두번째 표준 곡선이 최소 둘 이상의 점과 외부 표준으로부터 구성된다. 이 두번째 표준 곡선은 알려진 정량 기준에 대한 메틸화된 DNA를 위한 Ct값과 관련된다. 다음에, 테스트 시료 Ct값을 메틸화된 집단 및 메틸화 되지 않은 집단에 대해 결정하고 DNA의 유전체 당량 (genomic equivalent)을 처음의 두 단계에 의해 생성된 표준 곡선으로부터 계산한다. 관심 부위에서 메틸화의 백분율을 집단에서 DNA의

총량에 관하여 메틸화된 DNA의 양으로부터 계산하며 그 예는 (메틸화된 DNA의 수) / (메틸화된 DNA의 수 + 메틸화되지 않은 DNA의 수) X 100과 같다.

[0018] 본 명세서에 또한 제공되는 것은 방법을 실시하는 조성물 및 키트이다. 예를 들어, 일부 양태에서, 하나 이상의 마커에 특이적인 시약 (예, 프라이머, 탐침(probe))은 단독으로 또는 세트로 제공된다(예, 복수의 마커들을 증폭시키기 위한 프라이머 쌍의 세트). 검출 검정을 수행하기 위한 추가의 시약은 또한 제공될 수 있다 (예를 들어, QuARTS, PCR, 시퀀싱, 바이선폴라이트, 또는 다른 검정을 수행하기 위한 효소, 완충액, 양성 및 음성 대조군). 일부 양태에서, 방법을 수행하는데 필요한, 충분한, 또는 유용한 하나 이상의 시약을 함유하는 키트가 제공된다. 또한 제공되는 것은 시약을 함유하는 반응 혼합액이다. 더 제공되는 것은 서로에게 및/또는 테스트 시료에 더해져서 반응 혼합액을 완성할 수 있는 복수의 시약을 포함하는 마스터 혼합 시약(master mix reagent) 세트이다.

[0019] 일부 양태에서, 본 명세서에 기술된 기술은 본 명세서에 기술된 방법에 의해 제공된 바와 같은 일련의 산술적 또는 논리적인 연산을 수행하도록 설계된 프로그래밍 가능한 기계와 관련된다. 예를 들어, 본 발명의 기술의 일부 양태는 컴퓨터 소프트웨어 및/또는 컴퓨터 하드웨어(예, 이에서 수행됨)와 관련된다. 일 측면에서, 본 발명의 기술은 메모리의 형태, 산술적이고 논리적인 연산을 수행하기 위한 요소, 및 일련의 명령(예, 본 명세서에 제공된 바와 같은 방법)을 실행하기 위한 프로세싱 요소 (예, 마이크로프로세서)를 포함하는 데이터를 읽고, 조작하고, 저장하기 위한 컴퓨터에 관련된다. 일부 양태에서, 마이크로프로세서는 메틸화 상태(예, 하나 이상의 DMR의 메틸화 상태, 예를 들어, 표 2와 6에 제공된 바와 같은 DMR 1-96)를 결정하고; 메틸화 상태(예, 하나 이상의 DMR의 메틸화 상태, 예를 들어, 표 2와 6에 제공된 바와 같은 DMR 1-96)를 비교하며; 표준 곡선을 생성하고; Ct 값을 결정하며; 메틸화(예, 하나 이상의 DMR의 메틸화 상태, 예를 들어, 표 2와 6에 제공된 바와 같은 DMR 1-96)의 분수, 빈도, 또는 백분율을 계산하고; CpG 섬을 식별하며; 검정 또는 마커의 특이도(specificity) 및/또는 감도(sensitivity)를 결정하고; ROC 곡선 및 관련 AUC를 계산하며; 서열 분석; 본 명세서에 기술된 또는 기술분야에 알려진 모든 것에 대한 시스템의 일부이다.

[0020] 일부 양태에서, 마이크로프로세서 또는 컴퓨터는 메틸화 상태 데이터를 암의 부위를 예측하기 위한 알고리즘에서 사용한다.

[0021] 일부 양태에서, 소프트웨어 또는 하드웨어 구성요소는 복수 검정(multiple assay)의 결과를 수신하고 복수 검정의 결과에 기초한 암 위험을 나타내는 사용자에게 보고하기 위한 하나의 값 결과를 결정한다 (예를 들어, 표 2와 6에 제공된 바와 같은, 예를 들어, 복수 DMR의 메틸화 상태를 결정하는 것). 관련된 양태는 복수 검정, 예를 들어, 복수 마커의 메틸화 상태를 결정하는 것 (예를 들어, 표 2와 6에 제공된 바와 같은 복수 DMR과 같은) 으로부터 결과의 수학적 결합(mathematical combination) (예, 가중치 결합(weighted combination), 선형 결합(linear combination))에 기초하여 위험 요인을 계산한다. 일부 양태에서, DMR의 메틸화 상태는 차원(dimension)을 정의하고 다차원적인 공간에서 값을 가질 수 있으며, 복수 DMR의 메틸화 상태에 의해 정의된 좌표가 결과, 예를 들어, 사용자에게 보고하기 위한 결과이다.

[0022] 일부 양태는 저장 매체와 메모리 구성요소를 포함한다. 메모리 구성요소 (예, 휘발성 및/또는 비휘발성 메모리)는 저장 명령 (예, 본 명세서에 제공된 프로세서의 양태) 및/또는 데이터 (예, 메틸화 측정, 서열, 및 이와 관련된 통계적 기술과 같은 작업물)에서 사용한다. 일부 양태는 하나 이상의 CPU, 그래픽 카드, 및 유저 인터페이스(예를 들어, 디스플레이와 같은 출력 장치 및 키보드와 같은 입력장치를 포함)를 또한 포함하는 시스템에 관련된다.

[0023] 본 발명의 기술과 관련된 프로그램 가능한 기계는 종래의 현존하는 기술 및 발달중이거나 아직 발달할 기술[예를 들어, 양자 컴퓨터, 화학 컴퓨터, DNA 컴퓨터, 광학 컴퓨터, 스핀트로닉스 기반 컴퓨터(spintronics based computer) 등]을 포함한다.

[0024] 일부 양태에서, 본 발명의 기술은 데이터 전송을 위한 유선 (예를 들어, 금속 케이블, 광섬유) 또는 무선 전송 매체를 포함한다. 예를 들어, 일부 양태는 네트워크[예를 들면, 근거리 통신망 (local area network, LAN), 광역통신망 (wide area network, WAN), 애드-혹 네트워크(ad-hoc network), 인터넷 등]를 통한 데이터 전송에 관련된다. 일부 양태에서, 프로그램 가능한 기계는 이러한 네트워크에 피어(peer)로서 존재하며, 일부 양태에서 프로그램 가능한 기계는 클라이언트/서버 관계를 갖는다.

[0025] 일부 양태에서, 데이터는 하드 디스크, 플래쉬 메모리, 광학 매체, 플로피 디스크 등과 같은 컴퓨터-판독 가능 매체에 저장된다.

- [0026] 일부 양태에서, 본 명세서에 제공되는 기술은 본 명세서에 기술된 바와 같은 방법을 수행하기 위해 함께 작동하는 복수의 프로그램 가능한 장치와 관련된다. 예를 들어, 일부 양태에서, 복수의 컴퓨터 (예를 들어, 네트워크에 의해 연결됨)는 예를 들어, 이더넷, 광섬유와 같은 종래의 네트워크 인터페이스, 또는 무선 네트워크 기술에 의해 네트워크(개인, 공용, 또는 인터넷)에 연결된 완전 컴퓨터(complete computer) (온보드 CPU, 저장, 전원 공급장치, 네트워크 인터페이스 등이 있는)에 의존하는 다른 분산형 컴퓨터 구조 또는 그리드 컴퓨팅(grid computing) 또는 클러스터 컴퓨팅(cluster computing)의 실행에서 병렬적으로 동작하여 데이터를 수집하고 처리할 수 있다.
- [0027] 예를 들어, 일부 양태는 컴퓨터 관독가능한 매체를 포함하는 컴퓨터를 제공한다. 상기 양태는 프로세서에 커플링된 랜덤 액세스 메모리 (random access memory, RAM)를 포함한다. 프로세서는 메모리에 저장된 컴퓨터 실행 가능한 명령을 실행한다. 이러한 프로세서는 마이크로 프로세서, ASIC, 스테이트 기계(state machine), 또는 다른 프로세서를 포함할 수 있고, California Santa Clara의 Intel Corporation 및 Illinois Schaumburg의 Motorola Corporation의 프로세서와 같은 다수의 컴퓨터 프로세서 중 임의의 것이 될 수 있다. 이러한 프로세서는 매체, 예를 들어, 컴퓨터 관독가능한 매체를 포함하거나, 이와 커뮤니케이션 될 수 있고, 이는 프로세서에 의해 실행될 경우, 프로세서를 본 발명에 기술된 단계를 수행하도록 하는 명령을 저장한다.
- [0028] 컴퓨터 관독가능 매체의 예시는 컴퓨터 관독가능한 명령이 있는 프로세서를 제공할 수 있는 전자, 광학, 자기, 또는 다른 저장 또는 전송 장치를 포함하나, 이에 한정되는 것은 아니다. 적합한 매체의 다른 예시는 플로피 디스크, CD-ROM, DVD, 자기 디스크, 메모리칩, ROM, RAM, ASIC, 구성된 프로세서, 모든 광학 매체, 모든 자기 테이프 또는 다른 자기 매체, 또는 컴퓨터 프로세서가 명령을 관독할 수 있는 임의의 다른 매체를 포함하나, 이에 한정되는 것은 아니다. 또한, 컴퓨터 관독가능한 매체의 다양한 다른 형태는 라우터, 개인 또는 공용 네트워크, 또는 다른 전송 장치 또는 채널을 포함하는 컴퓨터에 명령을 전송하거나 전달할 수 있으며, 이는 유선 및 무선 모두 일 수 있다. 명령은, 예를 들어, C, C++, C#, 비주얼 베이직(Visual Basic), 자바(Java), 파이썬(Python), 펄(Perl), 및 자바 스크립트(JavaScript)를 포함하는 임의의 적합한 컴퓨터 프로그래밍 언어의 코드를 포함할 수 있다.
- [0029] 컴퓨터는 일부양태에서 네트워크로 연결된다. 컴퓨터는 마우스, CD-ROM, DVD, 키보드, 디스플레이, 또는 다른 입력 또는 출력장치와 같은 다수의 외부의 또는 내부의 장치를 또한 포함할 수 있다. 컴퓨터의 예시는 개인 컴퓨터, 디지털 보조장치, 개인 디지털 보조장치, 셀룰러 폰(cellular phone), 모바일 폰(mobile phone), 스마트폰(smart phone), 호출기, 디지털 테블릿, 랩탑 컴퓨터, 인터넷 가전, 및 다른 프로세서에 기초한 장치이다. 일반적으로, 본 명세서에 제공되는 기술의 측면에 관련된 컴퓨터는 본 명세서에 제공된 기술을 포함하는 하나 이상의 프로그램을 지지할 수 있는 마이크로소프트 윈도우(Microsoft Window), 리눅스(Linux), 유닉스(UNIX), 맥 OS X(Mac OS X) 등을 비롯한 임의의 운영체제에서 작동하는 프로세서에 기초한 플랫폼의 임의의 유형일 수 있다. 일부 양태는 다른 응용 프로그램(예, 어플리케이션)을 실행하는 개인 컴퓨터를 포함한다. 어플리케이션은 메모리에 포함될 수 있고, 이는 예를 들어, 워드 프로세싱 어플리케이션, 스프레드시트 어플리케이션, 이메일 어플리케이션, 인스턴트 메신저 어플리케이션, 프레젠테이션 어플리케이션, 인터넷 브라우저 어플리케이션, 달력/오거나이저 어플리케이션(calendar/organizer application), 및 클라이언트 장치에 의해 실행될 수 있는 임의의 다른 어플리케이션을 포함할 수 있다.
- [0030] 본 발명의 기술에 관련된 본 명세서에 기술된 모든 이러한 구성요소, 컴퓨터, 및 시스템은 논리적이거나 가상적일 수 있다.
- [0031] 본 명세서에 제공된 것은 개체로부터 수득한 시료에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 스크리닝하는 방법에 관련된 기술로, 그 방법은 개체로부터 수득한 시료에서 마커의 메틸화 상태를 결정하는 단계; 및 마커의 메틸화 상태가 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체에서 결정된 마커의 상태와 상이한 경우, 개체를 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 식별하는 단계로서, 상기 마커는 표 2와 6에서 제공되는 바와 같은 BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELMO1, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781로부터 선택된 차등 메틸화 영역(differentially methylated region, DMR)에서 하나 이상의 염기를 포함하는 것인, 단계;를 포함한다.
- [0032] 본 발명의 기술이 평가된 메틸화 상태에 한정되는 것은 아니다. 일부 양태에서, 시료에서 마커의 메틸화 상태를 평가하는 단계는 한 염기의 메틸화 상태를 결정하는 단계를 포함한다. 일부 양태에서, 시료에서 마커의 메틸화 상태를 결정하는 단계는 복수의 염기에서 메틸화의 정도를 결정하는 단계를 포함한다. 더욱이, 일부 양태

에서, 마커의 메틸화 상태는 마커의 정상적인 메틸화 상태에 비하여 마커의 증가된 메틸화를 포함한다. 일부 양태에서, 마커의 메틸화 상태는 마커의 정상적인 메틸화 상태에 비하여 마커의 감소된 메틸화를 포함한다. 일부 양태에서, 마커의 메틸화 상태는 마커의 정상적인 메틸화 상태에 비하여 마커의 메틸화의 상이한 패턴을 포함한다.

- [0033] 더욱이, 일부 양태에서, 마커는 100 염기 이하의 영역이고, 마커는 500 염기 이하의 영역이며, 마커는 1000 염기 이하의 영역이고, 마커는 5000 염기 이하의 영역이거나, 또는, 일부 양태에서, 마커는 하나의 염기일 수 있다. 일부 양태에서, 마커는 고 CpG 밀도 프로모터(high CpG density promoter)이다.
- [0034] 본 발명의 기술은 시료의 유형에 의해 한정되지 않는다. 예를 들어, 일부 양태에서, 시료는 대변 시료, 조직 시료(예를 들면, 위 조직, 췌장 조직, 담관(bile duct) / 간 조직, 췌액(pancreatic juice), 췌장 낭액, 및 대장 조직), 혈액 시료(예를 들어, 혈장, 혈청, 전혈), 배설물, 또는 소변 시료이다.
- [0035] 더욱이, 본 발명의 기술은 메틸화 상태를 결정하는데 사용된 방법에 한정되지 않는다. 일부 양태에서, 검정 단계는 메틸화 특이적 중합효소 연쇄반응(methylation specific polymerase chain reaction), 핵산 시퀀싱(nucleic acid sequencing), 질량분석법(mass spectrometry), 메틸화 특이적 핵산분해효소(methylation specific nuclease), 질량에 기초한 분리(mass-based separation), 또는 표적 포획(target capture)을 사용하는 것을 포함한다. 일부 양태에서, 검정 단계는 메틸화 특이적 올리고뉴클레오티드의 사용을 포함한다. 일부 양태에서, 본 발명의 기술은 초병렬 시퀀싱(massively parallel sequencing)(예, 차세대 시퀀싱(next-generation sequencing))을 메틸화 상태를 결정하는데 사용하며, 그 예로는 합성을 통한 시퀀싱(sequencing-by-synthesis), 실시간(예, 단일-분자(single-molecule)) 시퀀싱, 비드 에멀전 시퀀싱(bead emulsion sequencing), 나노포어 시퀀싱(nanopore sequencing) 등이 있다.
- [0036] 본 발명의 기술은 DMR을 검출하기 위한 시약을 제공하며, 예를 들어, 일부 양태에서, 제공되는 것은 SEQ ID NO: 1-44 (표 3) 및 SEQ ID NO: 45, 46, 47 및 48 (표 7)에 의해 제공되는 서열을 포함하는 올리고뉴클레오티드의 세트이다. 일부 양태에서, 제공되는 것은 DMR 중의 염기를 갖는 염색체 부위에 상보적인 서열을 포함하는 올리고뉴클레오티드이며, 그 예로는 DMR의 메틸화 상태에 민감한 올리고뉴클레오티드가 있다.
- [0037] 본 발명의 기술은 다양한 마커의 패널을 제공하며, 예를 들어, 일부 양태에서, 마커는 표 2와 6에서 제공되고 마커를 포함하는 주석(표 2와 6을 참고)을 갖는 염색체 부위를 포함한다. 게다가, 양태는 표 2와 6의 DMR을 하나 이상의 DMR No. 1-96로 분석하는 방법을 제공한다.
- [0038] 키트 양태가 제공되며, 예를 들어, 키트는 바이설과이트 시약; 및 DMR 1-96 (표 2와 6으로부터)로 이루어진 군으로부터 선택된 DMR로부터 서열을 포함하고 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체와 관련된 메틸화 상태를 포함하는 대조 핵산;을 포함한다. 키트 양태가 제공되며, 예를 들어, 키트는 바이설과이트 시약; 및 DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 92, 95 및 96 (표 2와 6으로부터)로 이루어진 군으로부터 선택된 DMR로부터 서열을 포함하고 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체와 관련된 메틸화 상태를 갖는 대조 핵산;을 포함한다.
- [0039] 일부 키트 양태는 개체로부터 본 명세서에 기술된 바와 같은 시료(예, 대변 시료)를 획득하는 시료 수집기; 시료로부터 핵산을 분리하기 위한 시약; 바이설과이트 시약; 및 올리고뉴클레오티드;를 포함한다.
- [0040] 본 발명의 기술은 조성물(예, 반응 혼합액)의 양태와 관련된다. 일부 양태에서, 제공되는 것은 DMR을 포함하는 핵산과 바이설과이트를 포함하는 조성물이다. 일부 양태는 본 명세서에 기술된 바와 같은 DMR을 포함하는 핵산과 올리고뉴클레오티드를 포함하는 조성물을 제공한다. 일부 양태는 DMR을 포함하는 핵산과 메틸 민감성 제한 효소를 포함하는 조성물을 제공한다. 일부 양태는 DMR을 포함하는 핵산과 중합효소를 포함하는 조성물을 제공한다.
- [0041] 추가의 관련된 방법 양태는 개체로부터 수득한 시료에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 스크리닝 하기 위해 제공되며, 예를 들어, 방법은 DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 92, 95, 및 96 (표 2와 6 참고) 중 하나 이상인 DMR 중의 염기를 포함하는 시료에서 마커의 메틸화 상태를 결정하는 단계; 개체 시료로부터 마커의 메틸화 상태와 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체로부터 정상 대조 시료의 마커의 메틸화 상태를 비교하는 단계; 및 개체 시료와 정상 대조 시료의 메틸화 상태의 차이의 신뢰구간 및/또는 p값을 결정하는 단계;를 포함한다.
- [0042] 일부 양태에서, 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 일부양태는 DMR을 포함하는 핵산과 바이설과이트

시약을 반응시켜 바이설파이트-반응된 핵산을 생성하는 단계; 바이설파이트-반응된 핵산을 시퀀싱하여 바이설파이트-반응된 핵산의 뉴클레오티드 서열을 제공하는 단계; 바이설파이트-반응된 핵산의 뉴클레오티드 서열을 얻을 갖지 않는 개체로부터의 DMR을 포함하는 핵산의 뉴클레오티드 서열과 비교하여 두 서열간의 차이점을 식별하는 단계; 및 차이점이 존재하는 경우, 개체를 신생물(neoplasm)을 갖는 것으로 식별하는 단계;를 제공한다.

[0043] 개체에서 수득한 시료에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 스크리닝하는 시스템은 본 발명의 기술에 의해 제공된다. 시스템의 예시적 양태는 예를 들어, 개체에서 수득한 시료에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 스크리닝하는 시스템, 시료의 메틸화 상태를 결정하도록 구성된 분석 구성요소를 포함하는 시스템, 시료의 메틸화 상태를 데이터베이스에 기록된 참조 시료 메틸화 상태 또는 대조 시료와 비교하도록 구성된 소프트웨어, 및 사용자에게 췌장 고등급 이형성증 메틸화 상태(예, 췌장 고등급 이형성증이 아닌 메틸화 상태, 췌장 고등급 이형성증에 대한 메틸화 상태)를 알려주도록 구성된 알림 구성요소를 포함한다. 알림은 일부 양태에서 복수 검정(예, 복수 마커의 메틸화 상태를 결정하는 것, 예를 들어, 표 2와 6에서 제공된 바와 같은 예를 들어, DMR)으로부터 결과를 수신하는 소프트웨어 구성요소에 의해서 및 보고하기 위한 값 또는 결과를 복수 결과에 기초하여 계산함으로써 결정된다. 일부 양태는 값 또는 결과를 계산하는 것 및/또는 사용자(예, 내과의, 간호사, 임상의 등과 같은)에게 보고를 위한 알림 사용을 위한 본 명세서에 제공된 각 DMR과 관련된 가중치 매개변수(weighted parameter)의 데이터 베이스를 제공한다. 일부 양태에서, 복수 검정으로부터 모든 결과들은 보고되며, 일부 양태에서, 하나 이상의 결과는 개체에서 췌장 고등급 이형성증을 나타내는 복수 검정으로부터 하나 이상의 결과의 종합에 기초하여 점수, 값, 또는 결과를 제공하는데 이용된다.

[0044] 시스템의 일부 양태에서, 시료는 DMR을 포함하는 핵산을 포함한다. 일부 양태에서, 시스템은 핵산을 분리하는 구성요소, 대변 시료를 수집하는 구성요소와 같은 시료를 수집하는 구성요소를 더 포함한다. 일부 양태에서, 시스템은 DMR을 포함하는 핵산서열을 포함한다. 일부 양태에서, 데이터베이스는 고등급 췌장 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체로부터 핵산서열을 포함한다. 또한 포함되는 것은 핵산으로, 그 예로 핵산의 세트, DMR을 포함하는 서열을 갖는 각 핵산이 있다. 일부 양태에서, 각 핵산이 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체로부터 서열을 갖는 핵산의 세트. 관련된 시스템 양태는 기술된 바와 같은 핵산의 세트와 핵산의 세트와 관련된 핵산서열의 데이터베이스를 포함한다. 일부 양태는 바이설파이트 시약을 더 포함한다. 일부 양태는 핵산 시퀀서(sequencer)를 더 포함한다.

[0045] 어떤 양태에서, 개체로부터 수득한 시료에서 췌장 고등급 이형성증을 검출하는 방법이 제공되며, 방법은 a) 개체로부터 DNA를 포함하는 시료를 획득하는 단계; b) 수득한 DNA를, 수득한 DNA에서 메틸화되지 않은 시토신 잔기를 선택적으로 변형시키나, 메틸화된 시토신 잔기는 변형하지 않는 시약으로 처리하여 변형된 잔기를 생성하는 단계; c) 단계 b)의 처리를 겪은 DNA에서 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계로서, 상기 하나 이상의 DNA 메틸화 마커는 DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 92, 95 및 96(표 2와 6으로부터)에 의해 제공되는 바와 같은 차등 메틸화 영역(differentially methylated region, DMR) 중의 염기를 포함하는 것인, 단계; d) 하나 이상의 DNA 메틸화 마커의 결정된 메틸화 수준을 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 개체에 대한 하나 이상의 DNA 메틸화 마커에 대한 메틸화 수준 참조를 비교하는 단계; 및 e) 차이점이 존재할 경우, 개체를 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 식별하는 단계;를 포함한다.

[0046] 일부 양태에서, 하나 이상의 DNA 메틸화 마커에서 상승된 메틸화의 결정은 CpG 섬과 CpG 섬 쇼어(CpG island shore)로 이루어진 군으로부터 선택된 영역내의 변경된 메틸화의 결정을 포함한다.

[0047] 일부 양태에서, CpG 섬과 CpG 쇼어 내의 상승된 DNA 메틸화 마커의 결정은 DNA 메틸화 마커의 조절 영역 또는 암호화 영역내의 상승된 메틸화를 포함한다.

[0048] 일부 양태에서, 단계 b)의 처리를 겪은 DNA에서 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계는 하나 이상의 DNA 메틸화 마커의 메틸화 빈도 및/또는 메틸화 점수를 결정하는 단계를 포함한다. 일부 양태에서, 단계 b)의 처리는 수득한 DNA의 바이설파이트 변형을 통해 수행된다.

[0049] 일부 양태에서, 단계 b)의 처리를 겪은 DNA에서 하나 이상의 DNA 메틸화 마커의 메틸화 수준을 결정하는 단계는 메틸화 특이적 PCR, 정량적 메틸화 특이적 PCR, 메틸화 민감성 DNA 제한 효소 분석, 정량적 바이설파이트 파이로시퀀싱, 및 바이설파이트 유전체 시퀀싱 PCR로 이루어진 군으로부터 선택된 기법에 의해 달성된다.

[0050] 일부 양태에서, 시료는 췌장 조직 및/또는 췌액, 대장 조직을 포함한다. 일부 양태에서, 시료는 췌장 낭액을 포함한다. 일부 양태에서, 시료는 대변 시료를 포함한다. 일부 양태에서, 시료는 혈액 시료를 포함한다.

- [0051] 어떤 양태에서, 본 발명은 인간의 생물학적 시료에서 BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781로부터 선택된 둘 이상의 유전자에 대한 CpG 부위의 메틸화 수준 측정을 통해 생물학적 시료를 특성화하는 방법을 제공한다.
- [0052] 이러한 방법은 CpG 부위의 메틸화 수준을 결정하는 특정 방법에 한정되는 것은 아니다. 일부 양태에서, 이러한 방법은 생물학적 시료 중의 유전체 DNA를 바이설파이트로 처리하는 단계; 바이설파이트 처리된 유전체 DNA를 선택된 둘 이상의 유전자에 대한 프라이머 세트를 사용하여 증폭시키는 단계; CpG 부위의 메틸화 수준을 메틸화 특이적 PCR, 정량적 메틸화 특이적 PCR, 메틸화 민감성 DNA 제한 효소 분석, 정량적 바이설파이트 파이로시퀀싱, 또는 바이설파이트 유전체 시퀀싱 PCR에 의해 결정하는 단계;를 통해 CpG 부위의 메틸화 수준을 측정한다.
- [0053] 일부 양태에서, 둘 이상의 유전자에 대한 CpG 부위의 메틸화 수준을 측정하는 단계는 상기 CpG 부위의 메틸화 점수를 결정하는 단계와 상기 CpG 부위의 메틸화 빈도를 결정하는 단계로 이루어진 군으로부터 선택된 결정을 포함한다.
- [0054] 일부 양태에서, 이러한 방법은 메틸화 수준을 체장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)이 없는 대조 시료에서 대응하는 유전자 세트의 메틸화 수준에 비교하는 단계; 및 둘 이상의 유전자에서 측정된 메틸화 수준이 각 대조 시료에서 측정된 메틸화 수준보다 큰 경우, 개체는 체장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 결정하는 단계;를 더 포함한다.
- [0055] 이러한 방법은 생물학적 시료의 특정 유형에 한정되지 않는다. 일부 양태에서, 생물학적 시료는 체장 조직 및/또는 체장 낭액을 포함한다. 일부 양태에서, 생물학적 시료는 대변 시료, 혈액 시료, 및/또는 혈액 분획 시료를 포함한다.
- [0056] 일부 양태에서, CpG 부위는 암호화 영역 또는 조절 영역에 존재한다.
- [0057] 일부 양태에서, 선택된 둘 이상의 유전자에 대한 프라이머 세트는 다음과 같다:
- [0058] BMP3에 대한 SEQ ID NO: 45 및 46로 이루어진 프라이머 세트;
- [0059] NDRG4에 대한 SEQ ID NO: 47 및 48로 이루어진 프라이머 세트;
- [0060] ABCB1에 대한 SEQ ID NO: 1 및 2로 이루어진 프라이머 세트;
- [0061] AK055957에 대한 SEQ ID NO: 3 및 4로 이루어진 프라이머 세트;
- [0062] C13ORF18에 대한 SEQ ID NO: 7 및 8로 이루어진 프라이머 세트;
- [0063] CD1D에 대한 SEQ ID NO: 43 및 44로 이루어진 프라이머 세트;
- [0064] CLEC11A에 대한 SEQ ID NO: 9 및 10로 이루어진 프라이머 세트;
- [0065] DLX4에 대한 SEQ ID NO: 11 및 12로 이루어진 프라이머 세트;
- [0066] ELM01에 대한 SEQ ID NO: 13 및 14로 이루어진 프라이머 세트;
- [0067] EMX1에 대한 SEQ ID NO: 15 및 16로 이루어진 프라이머 세트;
- [0068] FER1L4에 대한 SEQ ID NO: 17 및 18로 이루어진 프라이머 세트;
- [0069] FRMD4A에 대한 SEQ ID NO: 19 및 20로 이루어진 프라이머 세트;
- [0070] GRIN2D에 대한 SEQ ID NO: 21 및 22로 이루어진 프라이머 세트;
- [0071] HOXA1에 대한 SEQ ID NO: 23 및 24로 이루어진 프라이머 세트;
- [0072] LRRC4에 대한 SEQ ID NO: 25 및 26로 이루어진 프라이머 세트;
- [0073] PRKCB에 대한 SEQ ID NO: 29 및 30로 이루어진 프라이머 세트;
- [0074] SP9에 대한 SEQ ID NO: 31 및 32로 이루어진 프라이머 세트;
- [0075] ST6GAL2에 대한 SEQ ID NO: 33 및 34로 이루어진 프라이머 세트;

- [0076] ST8SIA1에 대한 SEQ ID NO: 35 및 36로 이루어진 프라이머 세트;
- [0077] TBX15에 대한 SEQ ID NO: 37 및 38로 이루어진 프라이머 세트;
- [0078] VWC2에 대한 SEQ ID NO: 39 및 40로 이루어진 프라이머 세트; 및
- [0079] ZNF781에 대한 SEQ ID NO: 41 및 42로 이루어진 프라이머 세트;를 포함한다.
- [0080] 추가적인 양태는 본 명세서에 포함된 교시(teaching)에 기초한 관련 기술분야의 통상적 기술자에게 명백할 것이다.

도면의 간단한 설명

- [0081] 도 1a 내지 도 1y는 실시예 I에 기술된 바와 같은 채장 조직으로부터 검정된 상이한 마커의 분포 플롯(distribution plot)을 제공한다.
- 도 2는 채장 낭액에서 상위 3개의 메틸화된 DNA 마커의 분포(베타-액틴에 의해 정규화됨)를 환자군 (HGD/암)와 대조군 (LGD/이형성증 없음)에 걸쳐 보여준다 (실시예 V를 참고).
- 도 3은 2개의 상위 후보 메틸화된 DNA 마커의 채장 낭액 수준(베타-액틴에 의해 정규화됨)을 각 낭포 카테고리에 걸쳐 나타낸다(실시예 VI를 참고).

발명을 실시하기 위한 구체적인 내용

- [0082] 본 개시내용은 채장 고등급 이형성증의 스크리닝을 위한 기술, 특히 채장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 존재를 검출하는 방법, 조성물, 및 관련 용도에 관한 것이나, 이로만 제한되는 것은 아니다.
- [0083] 본 발명의 기술이 본 명세서에 기재되나, 사용된 섹션 제목은 조직적인 목적만을 위한 것으로, 어떤 방식으로나 대상을 제한하는 것으로 해석되는 것이 아니다.
- [0084] 설명의 목적을 위한 다양한 양태의 상세한 설명에서, 다수의 구체적인 세부사항들이 개시된 양태의 완전한 이해를 제공하기 위해 제시된다. 그러나, 기술분야의 통상적인 기술자는 이러한 다양한 양태들이 이러한 구체적인 세부사항들과 또는 이들 없이 실행될 수 있음을 이해할 것이다. 다른 경우에서, 구조와 장치는 블록도 형태로 나타난다. 더욱이, 기술분야의 통상적인 기술자는 방법이 제시 및 수행되는 특정한 순서가 예시적임을 용이하게 인식할 것이며, 이 순서가 달라질 수 있으며 본 명세서에 개시된 다양한 양태의 사상과 범위내에 여전히 남아 있음이 의도된다.
- [0085] 특허, 특허 출원, 기사, 서적, 논문, 및 인터넷 웹페이지를 포함하나, 이에 한정되지 않는 이 출원에 인용된 모든 문헌 및 유사한 자료는 임의의 목적을 위한 참고문헌 전체에 의해 명확히 포함된다. 다르게 정의되지 않는 한, 본 명세서에 사용된 모든 기술적 용어와 과학적 용어는 본 명세서에 기술된 다양한 양태가 속하는 기술분야의 통상적인 기술자에 의해 통상적으로 이해되는 동일한 의미를 갖는다. 참고문헌에 포함된 용어의 정의가 본 발명의 교시에 제공된 정의와 상이한 것으로 나타나면, 본 발명의 교시에 제공된 정의가 지배적일 것이다.
- [0086] **정의**
- [0087] 본 발명의 기술의 이해를 용이하게 하기 위해, 복수의 용어와 구들이 아래에 정의된다. 추가의 정의는 상세한 설명 전체에 걸쳐 제시된다.
- [0088] 명세서와 청구항 전체에 걸쳐, 다음의 용어는 문맥이 명백히 다르게 지시하지 않는 한, 본 명세서에 명확히 관련된 의미를 취한다. 본 명세서에서 사용된 바와 같은 구 "일 양태에서"는 동일한 양태를 지칭할 수 있으나, 반드시 동일한 양태를 지칭하는 것은 아니다. 더욱이, 본 명세서에 사용된 바와 같은 구 "다른 양태에서"는 상이한 양태를 지칭할 수 있으나, 반드시 상이한 양태를 지칭하는 것은 아니다. 따라서, 아래에 기술된 바와 같이, 본 발명의 다양한 양태는 본 발명의 범위 또는 정신으로부터 벗어나지 않고 용이하게 결합될 수 있다.
- [0089] 또한, 본 명세서에서 사용된 바와 같이, 용어 "또는"은 문맥이 명백히 다르게 지시하지 않는 한, 포괄적인 "또는" 연산자이고 용어 "및/또는"과 동등하다. 용어 "~에 기초한"은 문맥이 명백히 다르게 지시하지 않는 한, 배타적이지 않고 기재되지 않은 추가의 요인에 기초되는 것을 허용한다. 또한, 본 명세서 전체에 걸쳐, 단수의 표현 ("a", "an", 및 "the")의 의미는 복수의 것을 포함한다. "~에서, ~내에서, 및 ~내의(in)"의 의미는 "~에서, ~내에서, 및 ~내의(in)"와 "~상에서, ~상의, 및 ~상(on)"을 포함한다.

- [0090] 본 명세서에 사용된 바와 같이, "핵산" 또는 "핵산 분자"는 일반적으로 임의의 리보핵산(ribonucleic acid) 또는 데옥시리보핵산(deoxyribonucleic acid)을 의미하고, 이는 변형되지 않거나 변형된 DNA 또는 RNA일 수 있다. "핵산"은, 제한없이 단일 가닥 및 이중 가닥 핵산을 포함한다. 본 명세서에 사용된 바와 같이, 용어 "핵산"은 하나 이상의 변형된 염기를 포함하는 상기 기재된 바와 같은 DNA를 또한 포함한다. 따라서, 안정성 또는 그 외의 이유로 변형된 뼈대(backbone)를 갖는 DNA가 "핵산"이다. 본 명세서에 사용된 바와 같은 용어 "핵산"은 이러한 화학적, 효소적 또는 대사적으로 변형된 핵산의 형태뿐만 아니라 예를 들어, 단순 세포 및 복합 세포를 포함하는 바이러스 및 세포의 DNA 특징의 화학적 형태를 포괄한다.
- [0091] 용어 "올리고뉴클레오타이드" 또는 "폴리뉴클레오타이드" 또는 "뉴클레오타이드" 또는 "핵산"은 2개 이상의, 바람직하게는 3개 초과, 일반적으로는 10개 초과, 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드를 갖는 분자를 의미한다. 정확한 크기는 주요 인자에 의존할 것이며, 이는 결국 올리고뉴클레오타이드의 궁극적인 기능 또는 용도에 의존하는 것이다. 올리고뉴클레오타이드는 화학적 합성, DNA 복제, 역전사 또는 이들의 조합을 포함하는 임의의 방식으로 생성될 수 있다. DNA에 대한 일반적인 데옥시리보뉴클레오타이드는 티민, 아데닌, 시토신, 및 구아닌이다. RNA에 대한 일반적인 리보뉴클레오타이드는 우라실, 아데닌, 시토신 및 구아닌이다.
- [0092] 본 명세서에 사용된 바와 같이, 용어 핵산의 "유전자 좌위" 또는 "영역"은 핵산의 하위영역(subregion), 예를 들면, 염색체 상의 유전자, 단일 뉴클레오타이드, CpG 섬 등을 의미한다.
- [0093] 용어 "상보적인" 및 "상보성"은 염기쌍화 규칙(base-pairing rule)에 의해 관련되는 뉴클레오타이드(예를 들면, 1개의 뉴클레오타이드) 또는 폴리뉴클레오타이드(예를 들면, 뉴클레오타이드의 서열)를 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 매칭(matching)된다. 또는, 핵산 사이에 "완전" 또는 "전체" 상보성이 있을 수 있다. 핵산 가닥 사이의 상보성 정도(degree of complementarity)는 핵산 가닥 사이의 혼성(hybridization)의 효율 및 강도에 영향을 미친다. 이는 핵산 사이의 결합에 의존하는 검출 방법 및 증폭 반응에서 특히 중요하다.
- [0094] 용어 "유전자"는 RNA 또는 폴리펩티드 또는 이의 전구체의 생성을 위해 필요한 암호화 서열을 포함하는 핵산(예를 들면, DNA 또는 RNA) 서열을 의미한다. 기능성 폴리펩티드는 폴리펩티드의 목적하는 활성 또는 기능적 특성(예를 들면, 효소적 활성, 리간드 결합, 신호 전달 등)이 유지되는 한, 암호화 서열의 임의의 부분 또는 암호화 서열의 전체 길이에 의해서 암호화될 수 있다. 유전자에 대해 사용되는 경우 용어 "부분(portion)"은 그 유전자의 단편(fragment)을 의미한다. 단편은 크기가 몇 개의 뉴클레오타이드 내지 1개의 뉴클레오타이드를 뺀 전체 유전자 서열 사이 일 수 있다. 따라서, "유전자의 적어도 한 부분을 포함하는 뉴클레오타이드"는 유전자의 단편 또는 전체 유전자를 포함할 수 있다.
- [0095] 용어 "유전자"는 구조적 유전자의 암호화 영역을 또한 망라하고, 5' 및 3' 말단 모두에서 암호화 영역에 인접하게 위치한 서열, 예를 들면, 양 말단에서 약 1kb의 거리를 포함해서, 유전자는 전체 길이의 mRNA의 길이에 대응한다(예를 들면, 암호화, 조절, 구조적 및 그 외의 서열을 포함). 암호화 영역의 5'에 위치하고 mRNA 상에 존재하는 서열은 5' 비번역(non-translated) 또는 번역되지 않은(untranslated) 서열로서 지칭된다. 암호화 영역의 3' 또는 하류(downstream)에 위치되고 mRNA 상에 존재하는 서열은 3' 비 번역 또는 3' 번역되지 않은 서열로 지칭된다. 용어 "유전자"는 cDNA와 유전자의 유전체 형태 모두를 아우른다. 일부 유기체(예를 들면, 진핵생물)에서, 유전자의 유전체 형태 또는 클론은 "인트론" 또는 "개재 영역(intervening regions)" 또는 "개재 서열(intervening sequence)"이라고 하는 비암호화 서열로 중단된 암호화 영역을 포함한다. 인트론은 핵 RNA(hnRNA)로 전사된 유전자의 단편(segment)이고; 인트론은 인핸서와 같은 조절 인자(regulatory element)를 포함할 수 있다. 인트론은 핵 또는 1차 전사물(transcript)로부터 제거되거나 "스플라이스 아웃(spliced out)"되고; 따라서 인트론은 메신저 RNA(messenger RNA, mRNA) 전사물에 존재하지 않는다. mRNA는 초기의 폴리펩티드(nascent polypeptide)에서 아미노산의 서열 또는 순서를 구체화하기 위해서 번역 시 가능하다.
- [0096] 인트론을 함유하는 것 이외에, 유전자의 유전체 형태는 RNA 전사물에 존재하는 서열의 5' 및 3' 두 말단에 위치한 서열을 또한 포함할 수 있다. 이들 서열은 "플랭킹(flanking)" 서열 또는 영역으로 지칭된다(이들 플랭킹 서열은 mRNA 전사물 상에 존재하는 비번역 서열의 5' 또는 3'에 위치한다). 5' 플랭킹 영역은 유전자의 전사를 제어하거나 이에 영향을 미치는 프로모터 및 인핸서(enhancer)와 같은 조절 서열을 포함할 수 있다. 3' 플랭킹 영역은 전사의 종결, 전사 후 절단(posttranscriptional cleavage), 및 폴리아데닐화(polyadenylation)를 지시하는 서열을 포함할 수 있다.
- [0097] 용어 "대립유전자"는 유전자의 차이(variation)를 의미하고; 차이는 이형(variant) 및 돌연변이, 다형성 유전자 좌위(polymorphic loci), 및 단일 뉴클레오타이드 다형성 유전자 좌위, 프레임 시프트 및 스플라이스 돌연변이를

포함하나, 이에 제한되는 것은 아니다. 대립유전자는 집단에서 자연적으로 발생할 수 있거나 집단의 임의의 특정 개체의 생애 중 발생할 수 있다.

[0098] 따라서, 용어 "이형" 및 "돌연변이"는 뉴클레오타이드 서열에 관하여 사용되는 경우, 대개 관련된 다른 뉴클레오타이드 산 서열과 하나 이상의 뉴클레오타이드가 상이한 핵산 서열을 의미한다. "이형"은 2개의 상이한 뉴클레오타이드 서열간의 차이이고; 일반적으로, 하나의 서열이 대조 서열(reference sequence)이다.

[0099] "증폭"은 템플레이트 특이도를 수반하는 핵산 복제의 특정한 경우이다. 이는 비특이적 템플레이트 복제(예를 들면, 템플레이트 의존적이나, 특이적 템플레이트에 의존하는 것은 아닌 복제)와 대조되는 것이다. 템플레이트 특이도는 여기서 복제(예를 들면, 적절한 폴리뉴클레오타이드 서열의 합성) 및 뉴클레오타이드(리보 뉴클레오타이드 또는 데옥시리보 뉴클레오타이드) 특이도의 정확도(fidelity)와 구별된다. 템플레이트 특이도는 "표적"특이도에 관하여 자주 기술된다. 표적 서열은 이들이 그 외의 핵산으로부터 선별되도록 하는 의미에서 "표적"이다. 증폭 기법은 주로 이 선별을 위해서 설계되었다.

[0100] 핵산의 증폭은 증폭 생성물 또는 증폭산물(amplicon)이 일반적으로 검출 가능한 소량의 폴리뉴클레오타이드(예를 들면, 단일 폴리뉴클레오타이드 분자, 10 내지 100 카피의 폴리뉴클레오타이드 분자이고, 이는 정확하게 동일하거나 동일하지 않을 수 있다)로부터 일반적으로 시작하는 복수 카피의 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 부분의 생성을 일반적으로 의미한다. 폴리뉴클레오타이드 증폭은 다양한 화학적 및 효소적 공정을 망라한다. 중합 효소 연쇄 반응(polymerase chain reaction, PCR) 또는 리가아제 연쇄 반응(ligase chain reaction, LCR; 예를 들어, 미국 특허 제5,494,810호를 참고)시, 하나 또는 소수 카피의 표적 또는 템플레이트 DNA 분자로부터 복수의 DNA 카피의 생성은 증폭의 형태이다. 증폭의 추가적 유형은 대립유전자 특이적 PCR(예를 들면, 미국 특허 제5,639,611호를 참고), 어셈블리 PCR(예를 들면, 미국 특허 제5,965,408호를 참고), 헬리카제 의존적 증폭(예를 들면, 미국 특허 제7,662,594호를 참고), 핫-스타트 PCR [Hot-start PCR, (예를 들면, 미국 특허 제5,773,258호 및 5,338,671호를 참고)], 서열간 특이적 PCR(intersequence-specific PCR), 역 PCR[inverse PCR, (예를 들면, Triglia, et al. (1988) Nucleic Acids Res., 16:8186을 참고)], 라이게이션 매개 PCR [ligation-mediated PCR, (예를 들면, Guilfoyle, R. et al. (1997) Nucleic Acids Research, 25:1854-1858 (1997); 미국 특허 제5,508,169호를 참고)], 메틸화 특이적 PCR (예를 들면, Herman, et al., (1996) PNAS 93(13) 9821-982을 참고), 미니 프라이머 PCR(miniprimer PCR), 멀티플렉스 라이게이션 의존적 탐침 증폭[multiplex ligation-dependent probe amplification, (예를 들면, Schouten, et al., (2002) Nucleic Acids Research 30(12): e57 참고)], 멀티플렉스 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 9:80을 참고], 네스티드 PCR(nested PCR), 오버랩 신장 PCR[overlap-extension PCR, (예를 들면, Higuchi, et al., (1988) Nucleic Acids Research 16(15) 7351-7367을 참고)], 실시간 PCR (예를 들면, Higuchi, et al. (1992) Biotechnology 10:413-417; Higuchi, et al., (1993) Biotechnology 11:1026-1030을 참고), 역전사 PCR (예를 들면, Bustin, S.A. (2000) J. Molecular Endocrinology 25:169-193을 참고), 고체상 PCR(solid phase PCR), 열적 비대칭 인터레이싱된 PCR (thermal asymmetric interlaced PCR), 및 터치 다운 PCR [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) 478-485을 참고)]를 포함할 수 있으나, 이에 한정되는 것은 아니다. 폴리뉴클레오타이드 증폭은 디지털 PCR(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 제W005023091A2호; 미국 특허 출원 공개 제20070202525호를 참고)을 사용하여 또한 수행될 수 있다.

[0101] 용어 "중합효소 연쇄반응(polymerase chain reaction, PCR)"은 유전체 DNA의 혼합물에서 클로닝(cloning) 또는 정제 없이 표적 서열의 단편(segment)의 농도를 증가시키는 방법을 기재하는 K.B. Mullis 미국 특허 제 4,683,195호, 4,683,202호, 및 4,965,188호의 방법을 지칭한다. 이 표적 서열을 증폭시키는 공정은 과량의 2개의 올리고뉴클레오타이드 프라이머를 목적하는 표적 서열을 포함하는 DNA 혼합물에 도입시키고, 이어서 DNA 중합 효소가 존재하는 정밀한 일련의 열적 사이클링(thermal cycling)으로 이루어진다. 2개의 프라이머는 이중가닥 표적 서열의 각각의 가닥에 대해 상보적이다. 증폭하기 위해서, 혼합물을 변성시키고, 이어서 프라이머를 표적 분자 내에서 이들의 상보적 서열에 어닐링한다. 어닐링 후, 프라이머를 중합효소로 신장하여 상보적 가닥의 새로운 쌍을 형성하도록 한다. 변성, 프라이머 어닐링, 및 중합효소 신장의 단계를 수 회 반복하여(즉, 변성, 어닐링, 및 신장이 하나의 "사이클(cycle)"을 구성하고; 많은 "사이클"이 있을 수 있음) 목적하는 표적 서열의 높은 농도의 증폭된 단편을 획득할 수 있다. 목적하는 표적 서열의 증폭된 단편의 길이는 서로에 대한 프라이머

의 상대적 위치에 의해서 결정되고, 따라서 이 길이는 제어가능한 매개변수이다. 공정의 반복 측면 덕분에, 이 방법은 "중합효소 연쇄반응(polymerase chain reaction, PCR)"으로 지칭된다. 표적 서열의 목적하는 증폭된 단편이 혼합물에서 (농도의 관점에서) 우세한 서열이 되기 때문에, 이들은 "PCR 증폭된"이라고 하며, "PCR 생성물" 또는 "증폭산물"이다.

[0102] 템플레이트 특이도는 효소의 선택에 의한 대부분 증폭 기술에서 달성된다. 증폭 효소는 이들이 사용되는 조건 하에서 핵산의 이중 혼합물에서 핵산의 특이적 서열만을 처리할 효소이다. 예를 들면, Q-베타 복제효소 (replicase)의 경우, MDV-1 RNA는 복제효소에 대한 특이적 템플레이트이다 (Kacian et al., Proc. Natl. Acad. Sci. USA, 69:3038 [1972]). 다른 핵산은 이 증폭 효소에 의해 복제되지 않을 것이다. 마찬가지로, T7 RNA 중합효소의 경우, 이 증폭 효소는 이의 자체 프로모터에 대한 엄격한 특이도를 갖는다 (Chamberlin et al, Nature, 228:227 [1970]). T4 DNA 리가아제(ligase)의 경우, 효소는 라이게이션 접합(ligation junction)에서 올리고뉴클레오타이드 또는 폴리뉴클레오타이드 기질과 템플레이트 사이의 미스매칭이 있는 2개의 올리고뉴클레오타이드 또는 폴리뉴클레오타이드를 라이게이팅(ligating) 하지 않을 것이다 (Wu and Wallace (1989) Genomics 4:560). 최종적으로, 고온에서 기능하는 능력 덕분에 열안정성 템플레이트 의존적 DNA 중합효소[thermostable template-dependant DNA polymerases, (예를 들면, Taq 및 Pfu DNA 중합효소)]는 프라이머에 의해 결합된 소정의 서열에 대해 높은 특이도를 드러내는 것으로 판명되었고, 고온은 표적 서열과 프라이머 혼성화(primer hybridization)를 장려하고 비표적 서열과의 혼성화는 장려하지 않는 열역학적 조건을 야기한다(H. A. Erlich (ed.), PCR Technology, Stockton Press [1989]).

[0103] 본 명세서에서 사용된 바와 같이, 용어 "핵산 검출 검정"은 관심 핵산의 뉴클레오타이드 조성물을 결정하는 임의의 방법을 의미한다. 핵산 검출 검정은 DNA 시퀀싱 방법, 탐침 혼성화 방법, 구조 특이적 절단 검정[structure specific cleavage assays, (예를 들면, Hologic, Inc의 INVADER 검정)]을 포함하나 이에 한정되지 않으며, 예를 들어, 미국특허 제5,846,717호, 5,985,557호, 5,994,069호, 6,001,567호, 6,090,543호, 및 6,872,816호; Lyamichiev et al., Nat. Biotech., 17:292 (1999), Hall et al., PNAS, USA, 97:8272 (2000), 및 미국 특허 제2009/0253142호; 효소 미스매치 절단 방법 [enzyme mismatch cleavage method, (예를 들면, Variagenics, 미국 특허 제6,110,684호, 5,958,692호, 5,851,770호)]; 중합효소 연쇄 반응; 가지 혼성화 방법[branched hybridization method, (예를 들면, Chiron, 미국특허 제5,849,481호, 5,710,264호, 5,124,246호, 및 5,624,802호)]; 롤링 서클 복제 [rolling circle replication, (예를 들면, 미국특허 제6,210,884호, 6,183,960호 및 6,235,502호)]; NASBA (예를 들면, 미국특허 제5,409,818호); 분자 비컨 기술[molecular beacon technology, (예를 들면, 미국특허 제6,150,097호)]; E-센서 기술 (Motorola, U.S. 미국특허 제 6,248,229호, 6,221,583호, 6,013,170호, 및 6,063,573호); 사이클링 탐침 기술 (예를 들면, 미국특허 제 5,403,711호, 5,011,769호, 및 5,660,988호); Dade Behring 신호 증폭 방법 (예를 들면, 미국특허 제6,121,001호, 6,110,677호, 5,914,230호, 5,882,867호, 및 5,792,614호); 리가아제 연쇄 반응 (예를 들면, Barnay Proc. Natl. Acad. Sci USA 88, 189-93 (1991)); 및 샌드위치 혼성화 방법[sandwich hybridization method, (예를 들면, 미국특허 제5,288,609호)]에 기재되어 있다.

[0104] 용어 "증폭 가능한 핵산"은 임의의 증폭 방법에 의해서 증폭될 수 있는 핵산을 의미한다. "증폭 가능한 핵산"은 보통 "시료 템플레이트"를 포함할 것으로 고려된다.

[0105] 용어 "시료 템플레이트"는 "표적"(하기에 정의됨)의 존재에 대해 분석된 시료로부터 유래한 핵산을 의미한다. 반면, "백그라운드 템플레이트"는 시료에서 존재하거나 존재하지 않을 수 있는 시료 템플레이트 이외의 핵산과 관하여 사용된다. 백그라운드 템플레이트는 대부분 의도치 않은 것이다. 이는 캐리오버(carryover)의 결과일 수 있거나 시료로부터 정제되도록 하는 핵산 오염물(nucleic acid contaminant)의 존재에 의한 것일 수 있다. 예를 들면, 검출될 핵산 이외의 생물 유래의 핵산은 시험 시료에서 백그라운드로서 존재할 수 있다.

[0106] 용어 "프라이머"는 핵산 가닥에 상보적인 프라이머 신장 생성물의 합성이 유도되는 조건하에 있는 경우(예를 들면, DNA 중합효소와 같은 유도제와 뉴클레오타이드의 존재 및 적합한 온도 및 pH에서), 정제된 제한 절단처럼 자연적으로 발생하거나 또는 합성적으로 생산되는지와 관계없이 합성의 개시점으로서 작용할 수 있는 올리고뉴클레오타이드를 의미한다. 프라이머는 바람직하게는 증폭에서 최대 효율을 위한 단일 가닥이나, 대안적으로 이중 가닥일 수 있다. 이중 가닥인 경우, 신장 생성물의 제조에 사용되기 전에 프라이머가 먼저 처리되어 이의 가닥을 나눈다. 바람직하게는, 프라이머는 올리고데옥시리보뉴클레오타이드(oligodeoxyribonucleotide)이다. 프라이머는 유도제의 존재 하에 신장 생성물의 합성을 준비(prime) 하기에 충분히 길어야 한다. 프라이머의 정확한 길이는 온도, 프라이머 공급원, 및 방법의 사용을 포함하는 다수의 인자에 의존할 것이다.

- [0107] 용어 "탐침"은 정제된 제한 절단처럼 자연적으로 발생하거나 또는 합성적, 재조합적으로, 또는 PCR 증폭으로 생산되는지 관계없이, 다른 관심 올리고뉴클레오타드에 혼성화 할 수 있는 올리고뉴클레오타드(예를 들면, 뉴클레오타드의 서열)를 의미한다. 탐침은 단일 가닥 또는 이중 가닥일 수 있다. 탐침은 특정한 유전자 서열의 검출, 확인, 및 분리에서 유용하다(예를 들면, "포획 탐침"). 본 발명에서 사용되는 임의의 탐침은 일부 양태에서 임의의 "리포터 분자"로 표지되어서 효소(예를 들어, 효소에 기초한 조직화학적 검정 뿐만 아니라 ELISA), 형광, 방사성 및 발광 시스템을 포함하나, 이에 한정되지 않는 임의의 검출 시스템에서 검출될 수 있다. 본 발명이 임의의 특정한 검출 시스템 또는 표지에 제한되는 것은 의도되지 않는다.
- [0108] 본 명세서에서 사용되는 바와 같이, "메틸화"는 시토신의 C5 또는 N4 위치, 아데닌의 N6 위치에서 시토신 메틸화 또는 그 외의 유형의 핵산 메틸화를 의미한다. 시험관에서, 증폭된 DNA는 보통 메틸화되지 않는데, 그 이유는 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 또는 메틸화된 증폭된 DNA를 의미할 수 있다.
- [0109] 따라서, 본 명세서에서 사용된 바와 같이, "메틸화된 뉴클레오타드" 또는 "메틸화된 뉴클레오타드 염기"는 메틸 모이어티(moiety)가 인식된 일반적인 뉴클레오타드 염기에 존재하지 않는 뉴클레오타드 염기 상의 메틸 모이어티의 존재를 의미한다. 예를 들면, 시토신은 이의 피리미딘 고리 상에 메틸 모이어티를 포함하지 않으나, 5-메틸시토신은 이의 피리미딘 고리의 위치 5에서 메틸 모이어티를 포함한다. 그러므로, 시토신은 메틸화된 뉴클레오타드가 아니며, 5-메틸시토신은 메틸화된 뉴클레오타드이다. 다른 예에서, 티민은 이의 피리미딘 고리의 위치 5에서 메틸 모이어티를 포함하나, 본 명세서의 목적을 위해, 티민이 DNA에 존재하는 경우, 티민이 DNA의 일반적인 뉴클레오타드 염기이기 때문에, 메틸화된 뉴클레오타드로 고려되지 않는다.
- [0110] 본 명세서에 사용된 바와 같이, "메틸화된 핵산 분자"는 하나 이상의 메틸화된 뉴클레오타드를 포함하는 핵산 분자를 의미한다.
- [0111] 본 명세서에서 사용되는 바와 같이, 핵산 분자의 "메틸화 상태", "메틸화 프로필" 및 "메틸화 상황"은 핵산 분자에서 하나 이상의 메틸화된 뉴클레오타드 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 뉴클레오타드를 포함하지 않는 핵산 분자는 메틸화되지 않은 것으로 고려된다.
- [0112] 특정한 핵산 서열(예를 들면, 본 명세서에 기재된 유전자 마커 또는 DNA 영역)의 메틸화 상태는 서열에서 모든 염기의 메틸화 상태를 나타낼 수 있거나, 서열 내에서 염기의 부분집합(예를 들면, 하나 이상의 시토신의)의 메틸화 상태를 나타낼 수 있거나, 메틸화가 발생한 서열 내에서 정확한 위치 정보 제공이 있거나 없는 서열 내에서 지역적 메틸화 밀도(regional methylation density)에 대한 정보를 나타낼 수 있다.
- [0113] 핵산 분자내 뉴클레오타드 좌위의 메틸화 상태는 핵산 분자내 특정한 유전자 좌위에서 메틸화된 뉴클레오타드의 존재 또는 부재를 의미한다. 예를 들면, 핵산 분자에서 7번째 뉴클레오타드에 존재하는 뉴클레오타드가 5-메틸시토신인 경우, 핵산 분자내 7번째 뉴클레오타드에서 시토신의 메틸화 상태는 메틸화된다. 마찬가지로, 핵산 분자내 7번째 뉴클레오타드에 존재하는 뉴클레오타드가 시토신(5-메틸시토신이 아님)인 경우, 핵산 분자내 7번째 뉴클레오타드에서 시토신의 메틸화 상태는 메틸화되지 않는다.
- [0114] 메틸화 상태는 "메틸화 값(methylation value)"(예를 들면, 메틸화 빈도, 분수, 비, 백분율 등을 나타냄)에 의해서 선택적으로 나타내거나 표시될 수 있다. 메틸화 값은 예를 들면, 메틸화 의존적 제한 효소로 제한 절단 후에 존재하는 온전한 핵산의 양을 정량하거나 바이설파이트 반응 후 증폭 프로필을 비교하거나 바이설파이트 처리된 및 처리되지 않은 핵산의 서열을 비교함으로써 생성될 수 있다. 따라서, 값, 예를 들면, 메틸화 값은 메틸화 상태를 나타내고 따라서 유전자 좌위의 복수 카피 전체에 걸친 메틸화 상황의 정량적인 지표로서 사용될 수 있다. 임계 값 또는 기준 값에 대해 시료 내에서 서열의 메틸화 상황을 비교하는 것이 필요할 경우, 이는 특히 유용하다.
- [0115] 본 명세서에서 사용되는 바와 같이, "메틸화 빈도" 또는 "메틸화 백분율(%)"은 분자 또는 유전자 좌위가 메틸화되지 않은 경우의 수에 관하여 분자 또는 유전자 좌위가 메틸화된 경우의 수를 의미한다.
- [0116] 이와 같이, 메틸화 상태는 핵산(예를 들면, 유전체 서열)의 메틸화 상태를 기재한다. 또한, 메틸화 상태는 메틸화에 관련된 특정한 유전체 유전자 좌위에서 핵산 단편의 특성을 의미한다. 이 DNA 서열 내에 임의의 시토신(C) 잔기가 메틸화되는 것에 관계없이, 이러한 특징은 메틸화된 C 잔기(들)의 위치, 핵산의 임의의 특정한 영역에 전체에 걸친 메틸화된 C의 빈도 또는 백분율, 및 예를 들면, 대립유전자의 기원의 차이로 인한 메틸화에서

대립유전자 차이(allelic difference)를 포함하나, 이로 한정되는 것은 아니다. 용어 "메틸화 상태", "메틸화 프로파일", 및 "메틸화 상황"은 또한 생물학적 시료에서 핵산의 임의의 특정 영역에 전체에 걸친 메틸화된 C 또는 메틸화되지 않은 C의 상대적 농도, 절대적 농도, 또는 패턴을 의미한다. 예를 들면, 핵산 서열 내의 시토신(C) 잔기(들)가 메틸화 되는 경우, 이는 "과메틸화된(hypermethylated)" 또는 "증가된 메틸화"를 갖는 것으로 지칭될 수 있고, 반면 DNA 서열 내의 시토신(C) 잔기(들)가 메틸화되지 않은 경우, "저메틸화된(hypomethylated)" 또는 "감소된 메틸화"를 갖는 것으로 지칭될 수 있다. 마찬가지로, 핵산 서열 내의 시토신(C) 잔기(들)가 다른 핵산 서열(예를 들면, 상이한 영역 또는 상이한 개체 등으로부터)에 비해 메틸화되는 경우, 그 서열은 그 외의 핵산 서열에 비해 과메틸화되거나 증가한 메틸화를 갖는 것으로 고려된다. 대안적으로, DNA 서열 내의 시토신(C) 잔기(들)가 다른 핵산 서열(예를 들면, 상이한 영역 또는 상이한 개체 등으로부터)에 비해 메틸화되지 않은 경우, 그 서열은 그 외의 핵산 서열에 비해 저메틸화되거나 감소한 메틸화를 갖는 것으로 고려된다. 또한, 본 명세서에 사용되는 용어 "메틸화 패턴"은 핵산 영역에 비해 메틸화된 및 메틸화되지 않은 뉴클레오타이드의 집합 부위(collective site)를 의미한다. 메틸화된 및 메틸화되지 않은 뉴클레오타이드의 수가 영역 전체에 걸쳐 동일하거나 유사하나, 메틸화 및 메틸화되지 않은 뉴클레오타이드의 위치가 상이한 경우, 2개의 핵산은 동일하거나 유사한 메틸화 빈도 또는 메틸화 백분율을 가질 수 있으나, 상이한 메틸화 패턴을 가질 수도 있다. 이들이 메틸화 정도(예를 들면, 하나가 그 외의 것에 대해 증가한 또는 감소한 메틸화를 갖음), 빈도, 또는 패턴이 상이한 경우, 서열은 "차등적으로 메틸화된" 또는 "메틸화의 차이" 또는 "상이한 메틸화 상태"를 갖는 것으로 불린다. 용어 "차등적 메틸화"는 암 음성 시료에서 핵산 메틸화의 수준 또는 패턴과 비교한 암 양성 시료에서 핵산 메틸화의 수준 또는 패턴의 차이를 의미한다. 이는 수술 후 암이 재발한 환자 대 재발하지 않은 환자 사이의 수준 또는 패턴의 차이를 또한 의미할 수 있다. 예를 들어, 정확한 컷오프(cut-off) 또는 예측 특징이 정의되면, 차등적 메틸화 및 DNA 메틸화의 특정 수준 또는 패턴은 예후 바이오마커(prognostic biomarker) 및 예측 바이오마커(predictive biomarker)이다.

[0117] 메틸화 상태 빈도는 개체의 집단 또는 단일 개체로부터의 시료 기재에 사용될 수 있다. 예를 들어, 50%의 메틸화 상태 빈도를 갖는 뉴클레오타이드 좌위는 50%의 경우가 메틸화되고 50% 경우가 메틸화되지 않는다. 이러한 빈도를 사용하여 예를 들어, 핵산의 집합 또는 개개의 집단에서 뉴클레오타이드 유전좌위 또는 핵산 영역이 메틸화되는 정도를 기재할 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 상이할 것이다. 이러한 빈도를 또한 사용하여, 예를 들면, 단일 개체에서 뉴클레오타이드 좌위 또는 핵산 영역이 메틸화되는 정도를 기재할 수 있다. 예를 들면, 이러한 빈도를 사용하여 조직 시료로부터의 세포의 그룹이 뉴클레오타이드 좌위 또는 핵산 영역에서 메틸화되거나 메틸화되지 않은 정도를 기재할 수 있다.

[0118] 본 명세서에 사용된 바와 같이, "뉴클레오타이드 좌위"는 핵산 분자에서 뉴클레오타이드의 위치를 의미한다. 메틸화된 뉴클레오타이드의 뉴클레오타이드 좌위는 핵산 분자에서 메틸화된 뉴클레오타이드의 위치를 의미한다.

[0119] 일반적으로, 인간 DNA의 메틸화는 시토신이 구아닌의 5'에 위치하는(또한 CpG 디뉴클레오타이드 서열에서 종결함) 인접한 구아닌 및 시토신을 포함하는 디뉴클레오타이드 서열 상에서 발생한다. CpG 디뉴클레오타이드 내의 대부분의 시토신은 인간 유전체에서 메틸화되나, 일부는 CpG 섬이라고 알려진 특정 CpG 디뉴클레오타이드가 풍부한 유전체 영역에서 메틸화되지 않고 남아있다 (예를 들면, Antequera et al. (1990) Cell 62: 503-514를 참고).

[0120] 본 명세서에서 사용된 바와 같이, "CpG 섬"은 전체 유전체 DNA에 대해 CpG 디뉴클레오타이드의 증가된 수를 함유하는 유전체 DNA의 G:C 풍부 영역을 의미한다. 영역의 G:C 함량이 적어도 50%이고, 예상된 빈도에 대한 관찰된 CpG 빈도의 비가 0.6인 경우, CpG 섬은 길이가 적어도 100, 200개 이상의 염기쌍일 수 있고, 일부 예에서, 영역의 G:C 함량이 적어도 55%이고, 예상된 빈도에 대한 관찰된 CpG 빈도의 비가 0.65인 경우, CpG 섬은 길이가 적어도 500 개의 염기쌍일 수 있다. 예상된 빈도에 대한 관찰된 CpG 빈도는 Gardiner-Garden et al (1987) J. Mol. Biol. 196: 261-281에서 제공된 방법에 따라 계산될 수 있다. 예를 들면, 예상된 빈도에 대한 관찰된 CpG 빈도는 식 $R = (A \times B) / (C \times D)$ 에 따라 계산될 수 있고, 이 식에서 R은 예상된 빈도에 대한 관찰된 CpG 빈도의 비이고, A는 분석된 서열에서 CpG 디뉴클레오타이드의 수이며, B는 분석된 서열에서 전체 뉴클레오타이드의 수이고, C는 분석된 서열에서 전체 C 뉴클레오타이드의 수이며, D는 분석된 서열에서 전체 G 뉴클레오타이드의 수이다. 메틸화 상태는 예를 들면, 프로모터 영역의 CpG 섬에서 일반적으로 결정된다. 인간 유전체에서 그 외의 서열이 CpA 및 CpT와 같은 DNA 메틸화의 경향이 있는 것 역시 인식될 것이다. (예를 들어, Ramsahoye (2000) Proc. Natl. Acad. Sci. USA 97: 5237-5242; Salmon and Kaye (1970) Biochim. Biophys. Acta. 204: 340-351; Grafstrom (1985) Nucleic Acids Res. 13: 2827-2842; Nyce (1986) Nucleic Acids Res. 14: 4353-4367; Woodcock (1987) Biochem. Biophys. Res. Commun. 145: 888-894를 참고).

- [0121] 본 명세서에서 사용된 바와 같이, 핵산 분자의 뉴클레오티드를 핵산 분자의 메틸화 상태의 작용으로서 변형시키는 시약, 또는 메틸화 특이적 시약은 핵산 분자의 메틸화 상태를 반영하는 방식으로 핵산 분자의 뉴클레오티드 서열을 변화시킬 수 있는 화합물 또는 조성물 또는 다른 시약을 의미한다. 이러한 시약으로 핵산 분자를 처리하는 방법은 시약과 핵산 분자를 접촉시키는 단계를 포함하고, 필요한 경우, 추가의 단계와 결합되어 뉴클레오티드 서열의 목적하는 변화를 달성할 수 있다. 이러한 핵산 분자의 뉴클레오티드 서열의 변화는 각각의 메틸화된 뉴클레오티드가 상이한 뉴클레오티드로 변형되는 핵산 분자를 야기할 수 있다. 이러한 핵산 뉴클레오티드 서열의 변화는 각각의 메틸화 되지 않은 뉴클레오티드가 상이한 뉴클레오티드로 변형되는 핵산 분자를 야기할 수 있다. 이러한 핵산 뉴클레오티드 서열의 변화는 메틸화되지 않은 각각의 선택된 뉴클레오티드(예를 들면, 각각의 메틸화되지 않은 시토신)가 상이한 뉴클레오티드로 변형되는 핵산 분자를 야기할 수 있다. 핵산 뉴클레오티드 서열을 변화시키기 위한 이러한 시약의 사용은 메틸화된 뉴클레오티드인 각각의 뉴클레오티드 (예를 들면, 각각 메틸화된 시토신)가 상이한 뉴클레오티드로 변형되는 핵산분자를 야기할 수 있다. 본 명세서에 사용된 바와 같이, 선택된 뉴클레오티드를 변형시키는 시약의 사용은 핵산 분자에서 일반적으로 발생하는 4개의 뉴클레오티드(DNA에 대한 C, G, T, 및 A와 RNA에 대한 C, G, U, 및 A) 중 1개의 뉴클레오티드를 변형시키는 시약을 의미해서, 시약은 그 외의 3개의 뉴클레오티드 변형없이 1개의 뉴클레오티드를 변형시킨다. 하나의 예시적 양태에서, 이러한 시약은 메틸화되지 않은 선택된 뉴클레오티드를 변형시켜 상이한 뉴클레오티드를 생성한다. 다른 예시적 양태에서, 이러한 시약은 메틸화되지 않은 시토신 뉴클레오티드를 탈아미노(deaminated) 할 수 있다. 예시적 시약은 바이설파이트이다.
- [0122] 본 명세서에서 사용된 바와 같이, 용어 "바이설파이트 시약"은 일부 양태에서, 바이설파이트, 디설파이트, 하이드로젠 설파이트, 또는 이들의 조합을 포함하여 예를 들면, CpG 디뉴클레오티드 서열에서 메틸화된 시티딘 및 메틸화되지 않은 시티딘을 구별하는 시약을 의미한다.
- [0123] 용어 "메틸화 검정"은 핵산의 서열에서 하나 이상의 CpG 디뉴클레오티드 서열의 메틸화 상태를 결정하는 임의의 검정을 의미한다.
- [0124] 용어 "MS AP-PCR" [메틸화 민감성 임의적 프라임드 중합효소 연쇄 반응, (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction)]은 CG 풍부 프라이머를 사용하여 유전체의 포괄적인 스캔(global scan)을 허용하고 CpG 디뉴클레오티드를 가장 함유할 것 같은 영역에 초점을 맞추는 기술분야에서 알려진 기술을 의미하고, Gonzalgo et al. (1997) Cancer Research 57: 594-599에 기재되어 있다.
- [0125] 용어 "MethyLight™"은 Eads et al. (1999) Cancer Res. 59: 2302-2306에 기재된 기술분야에서 알려진 형광에 기초한 실시간 PCR 기법을 의미한다.
- [0126] 용어 "HeavyMethyl™"은 증폭 프라이머 사이의 CpG 위치를 커버하는(covering) 또는 이에 의해 커버된(covered) 메틸화 특이적 블로킹 탐침(또한 본 명세서에서 블로커(blocker)라고 지칭)이 핵산 시료의 메틸화 특이적 선택 증폭을 가능하게 하는 검정을 의미한다.
- [0127] 용어 "HeavyMethyl™ MethyLight™" 검정은 MethyLight™ 검정이 증폭 프라이머 사이의 CpG 위치를 커버하는 메틸화 특이적 블로킹 탐침과 합쳐진 MethyLight™ 검정의 변형인 HeavyMethyl™ MethyLight™ 검정을 의미한다.
- [0128] 용어 "Ms-SNuPE" [메틸화 민감성 단일 뉴클레오티드 프라이머 신장, (Methylation-sensitive Single Nucleotide Primer Extension)]는 Gonzalgo & Jones (1997) Nucleic Acids Res. 25: 2529-2531에 기재된 기술분야에서 알려진 검정을 의미한다.
- [0129] 용어 "MSP"[메틸화 특이적 PCR, (Methylation-specific PCR)]은 Herman et al. (1996) Proc. Natl. Acad. Sci. USA 93: 9821-9826, 및 미국 특허 제5,786,146호에 의해 기재된 기술분야에서 알려진 메틸화 검정을 의미한다.
- [0130] 용어 "COBRA" [합쳐진 바이설파이트 제한 분석, (Combined Bisulfite Restriction Analysis)]은 Xiong & Laird (1997) Nucleic Acids Res. 25: 2532-2534에 의해 기재된 기술분야에서 알려진 메틸화 검정을 의미한다.
- [0131] 용어 "MCA" [메틸화된 CpG섬 증폭, (Methylated CpG Island Amplification)]은 Toyota et al. (1999) Cancer Res. 59: 2307-12, 및 WO 00/26401A1호에 기재된 메틸화 검정을 의미한다.
- [0132] 본 명세서에 사용된 바와 같이, "선택된 뉴클레오티드"는 핵산 분자에서 일반적으로 발생하는 4개의 뉴클레오티드(DNA에 대한 C, G, T, 및 A와 RNA에 대한 C, G, U, 및 A) 중 1개의 뉴클레오티드를 의미하고, 일반적으로 발

생하는 뉴클레오타이드의 메틸화된 유도체를 포함할 수 있다(예를 들면, C가 선택된 뉴클레오타이드인 경우, 메틸화된 및 메틸화되지 않은 C는 선택된 뉴클레오타이드의 의미에 모두 포함됨), 반면 메틸화된 선택된 뉴클레오타이드는 메틸화된 일반적으로 발생하는 뉴클레오타이드를 구체적으로 의미하고, 메틸화되지 않은 선택된 뉴클레오타이드는 메틸화되지 않은 일반적으로 발생하는 뉴클레오타이드를 구체적으로 의미한다.

[0133] 용어 "메틸화 특이적 제한 효소" 또는 "메틸화 민감성 제한 효소"는 이의 인식 부위의 메틸화 상태에 의존하는 핵산을 선택적으로 절단하는 효소를 의미한다. 인식 부위가 메틸화되지 않거나 헤미메틸화되면(hemimethylated) 특이적으로 절단하는 제한 효소의 경우, 인식 부위가 메틸화된다면, 절단이 일어나지 않거나 현저히 감소된 효율로 일어날 것이다. 인식 부위가 메틸화되면 특이적으로 절단하는 제한 효소의 경우, 인식 부위가 메틸화되지 않으면, 절단이 일어나지 않거나 현저히 감소된 효율로 일어날 것이다. 바람직한 것은 메틸화 특이적 제한 효소이며, 인식 서열은 CG 디뉴클레오타이드를 함유한다(예를 들면, CGCG 또는 CCCGGG와 같은 인식 서열). 일부 양태에 더욱 바람직한 것은 이 디뉴클레오타이드의 시토신이 탄소원자 C5에서 메틸화될 경우, 절단하지 않는 제한 효소이다.

[0134] 본 명세서에 사용된 바와 같이, "상이한 뉴클레오타이드"는 선택된 뉴클레오타이드와 화학적으로 상이한 뉴클레오타이드를 의미하여서, 일반적으로 상이한 뉴클레오타이드는 선택된 뉴클레오타이드와 상이한 왓슨-크릭 염기쌍(Watson-Crick base-pairing) 특성을 가지므로, 선택된 뉴클레오타이드와 상보적인 일반적으로 발생하는 뉴클레오타이드는 상이한 뉴클레오타이드와 상보적인 일반적으로 발생하는 뉴클레오타이드와 동일하지 않다. 예를 들면, C가 선택된 뉴클레오타이드인 경우, U 또는 T가 상이한 뉴클레오타이드일 수 있고, 이는 G에 대한 C의 상보성 및 A에 대한 U 또는 T의 상보성에 의해 예시된다. 본 명세서에서 사용된 바와 같이, 선택된 뉴클레오타이드에 상보적이거나, 상이한 뉴클레오타이드에 상보적인 뉴클레오타이드는 매우 엄격한 조건하에서 상보적인 뉴클레오타이드의 염기쌍-4개의 일반적으로 발생하는 뉴클레오타이드 중 3개와-보다 높은 친화도를 갖는 상이한 뉴클레오타이드 또는 선택된 뉴클레오타이드와 염기쌍을 형성하는 뉴클레오타이드를 의미한다. 상보성의 예시는 DNA(예를 들면, A-T 및 C-G) 및 RNA(예를 들면, A-U 및 C-G)에서 왓슨-크릭 염기쌍이다. 따라서, 예를 들면, 매우 엄격한 조건하에서 G, A, 또는 T에 대한 G 염기쌍보다 C에 대한 높은 친화도를 갖는 G 염기쌍은 그러므로 C가 선택된 뉴클레오타이드인 경우, G는 선택된 뉴클레오타이드에 대한 상보적인 뉴클레오타이드이다.

[0135] 본 명세서에서 사용된 바와 같이, 소정의 마커의 "감도"는 신생성 시료(neoplastic sample)와 비신생성 시료(non-neoplastic sample)를 구별하는 임계 값을 초과한 DNA 메틸화 값을 보고하는 시료의 백분율을 의미한다. 일부 양태에서, 양성(+)은 임계 값을 초과한 DNA 메틸화 값(예를 들면, 질병과 관련된 범위)을 보고하는 조직학적으로 확인된 신생으로 정의되고, 위음성(false negative)은 임계 값 미만의 DNA 메틸화 값(예를 들면, 질병 없음과 관련된 범위)을 보고한 조직학적으로 확인된 신생으로서 정의된다. 따라서, 감도의 값은 알려진 질병에 걸린 시료로부터 수득한 소정의 마커에 대한 DNA 메틸화 측정이 질병과 관련된 측정값 범위 내에 있을 확률을 반영한다. 본 명세서에 정의된 바와 같이, 계산된 감도 값의 임상 관련성은 소정의 마커가 임상 질환이 있는 개체에 적용되는 경우, 그 질환의 존재를 검출할 확률의 추정을 나타낸다.

[0136] 본 명세서에 사용된 바와 같이, 소정의 마커의 "특이도"는 신생물성 시료와 비신생물성 시료를 구별하는 임계 값 미만의 DNA 메틸화 값을 보고하는 비신생물성 시료의 백분율을 의미한다. 일부 양태에서, 음성은 임계 값 미만의 DNA 메틸화 값(예를 들면, 질병 없음과 관련된 범위)을 보고한 조직학적으로 확인된 비신생물성 시료로 정의되고, 위양성(false positive)은 임계 값 초과한 DNA 메틸화 값(예를 들면, 질병과 관련된 범위)을 보고하는 조직학적으로 확인된 비신생물성 시료로서 정의된다. 따라서, 특이도의 값은 알려진 비신생물성 시료로부터 수득한 소정의 마커에 대한 DNA 메틸화 측정이 질병 없음에 관련된 측정값 범위 내에 있을 확률을 반영한다. 본 명세서에 정의된 바와 같이, 계산된 특이도 값의 임상 관련성은 소정의 마커가 임상 질환이 없는 환자에 적용되는 경우, 임상 질환의 부재를 검출할 확률의 추정을 나타낸다.

[0137] 본 명세서에 사용된 용어 "AUC"는 "곡선하 면적(area under a curve)"의 약칭이다. 특히, 이는 수신자 동작 특성(Receiver Operating Characteristic, ROC) 곡선하 면적을 의미한다. ROC 곡선은 진단시험의 상이한 가능한 절단 점(cut point)에 관한 위양성율(false positive rate)에 대한 진양성율(true positive rate)의 플롯이다. 이는 선택된 절단점에 따른 감도와 특이도 사이의 트레이드-오프(trade-off)를 나타낸다(특이도 감소에 의해 임의의 감도 증가가 달성될 것임). ROC 곡선하 면적(AUC)은 진단 시험의 정확도에 대한 척도이다(면적이 클수록 우수하고, 최적은 1이며; 랜덤 시험은 0.5의 면적을 갖는 대각선 상에 있는 ROC 곡선을 가질 것임; J. P. Egan, (1975) Signal Detection Theory and ROC Analysis, Academic Press, New York을 참고).

[0138] 본 명세서에서 사용된 바와 같이, 용어 "신생물"은 "비정상적 조직의 덩어리, 그 성장이 정상 조직의 것을 넘어

서며 통제되지 않는 것"을 의미하며, 예를 들면, Willis RA, "The Spread of Tumors in the Human Body", London, Butterworth & Co, 1952를 참고할 수 있다.

- [0139] 본 명세서에서 사용된 바와 같이, "선종(adenoma)"은 선 유래(glandular origin)의 양성 종양을 의미한다. 이들의 성장은 양성이지만, 시간 경과에 따라 이들이 진행되어 악성 종양으로 될 수 있다.
- [0140] 용어 "전암성(pre-cancerous)" 또는 "전신생물성(pre-neoplastic)" 및 이의 증가물은 악성 형질 전환을 겪는 임의의 세포 증식 질환을 의미한다.
- [0141] 신생물, 선종, 암 등의 "부위" 또는 "영역"은 신생물, 선종, 암 등이 위치한 개체의 신체 내에서 조직, 기관, 세포 유형, 해부학적 영역, 신체 부위 등이다.
- [0142] 본 명세서에서 사용된 바와 같이, "진단적" 시험 적용은 개체의 질병 상태 또는 질환의 검출 또는 식별을 포함하는데, 이는 개체가 소정의 질병 또는 질환에 걸릴 가능성을 결정하고, 질병 또는 질환이 있는 개체가 치료에 반응할 가능성을 결정하며, 질병 또는 질환이 있는 개체의 예후를 결정하고(또는 이의 진행 또는 퇴보 가능성), 질병 또는 질환이 있는 개체에 대한 치료 효과를 결정한다. 예를 들면, 진단은 신생물을 축소시키는 개체의 존재 또는 가능성 또는 이러한 개체가 화합물(예를 들면, 약학적, 예를 들면, 약물) 또는 그 외의 치료에 순조롭게 반응할 가능성의 검출에 사용될 수 있다.
- [0143] 본 명세서에서 사용되는 용어 "마커"는 예를 들면, 이의 메틸화 상태에 기초하여 정상 세포로부터 질환-관련 세포[(예를 들어, 질환과 관련된 비암성 세포) (예를 들어, 질환과 관련된 암성 세포)]를 구별함으로써 질환(예를 들어, 비암성 질환)(예를 들어, 암성 질환)]을 진단할 수 있는 물질 (예를 들면, 핵산 또는 핵산의 영역)을 의미한다.
- [0144] 용어 "분리된(isolated)"이 "분리된 올리고뉴클레오타이드"에서와 같이 핵산에 대해 사용되는 경우, 이의 자연적인 공급원에서 통상적으로 관련되어 있는 적어도 하나의 오염물 핵산으로부터 식별되고 나뉜(separated) 핵산 서열을 의미한다. 분리된 핵산은 자연에서 발견된 것으로부터 상이한 형태 또는 설정에 존재한다. 반면, DNA 및 RNA와 같은 비분리 핵산은 이들이 자연에서 존재하는 상태로 발견된다. 비분리 핵산의 예시는 이웃하는 유전자에 근접한 숙주 세포 염색체 상에서 발견되는 소정의 DNA 서열 (예를 들면, 유전자), 다양한 단백질을 암호화하는 수 많은 그 외의 mRNA와의 혼합물로서 세포에서 발견되는 특정 단백질을 암호화하는 특정 mRNA 서열과 같은 RNA 서열을 포함한다. 그러나, 핵산이 자연적인 세포의 염색체 위치와 상이한 염색체 위치에 있거나, 그렇지 않으면 자연에서 발견되는 것보다 상이한 핵산 서열에 의해 플랭킹(flanking)되는 경우, 특정한 단백질을 암호화하는 분리된 핵산은 일례로 단백질을 통상적으로 발현하는 이러한 세포내 핵산을 포함한다. 분리된 핵산 또는 올리고뉴클레오타이드는 단일 가닥 또는 이중 가닥 형태로 존재할 수 있다. 분리된 핵산 또는 올리고뉴클레오타이드가 단백질을 발현하기 위해 이용되는 경우, 올리고뉴클레오타이드는 최소 센스 가닥(sense strand) 또는 암호화 가닥을 포함할 것이나(즉, 올리고뉴클레오타이드는 단일 가닥일 수 있음), 센스 및 안티센스 가닥 둘을 함유할 수 있다(즉, 올리고뉴클레오타이드는 이중 가닥일 수 있음). 자연적 또는 일반적인 환경으로부터 분리된 후, 분리된 핵산은 다른 핵산 또는 분자와 결합될 수 있다. 예를 들어, 분리된 핵산은 예를 들면, 이중 발현(heterologous expression)을 위해 배치된 숙주세포에 존재할 수 있다.
- [0145] 용어 "정제된"은 이들의 자연 환경으로부터 제거되거나, 분리되거나, 나뉜 핵산 또는 아미노산 서열인 분자를 의미한다. "분리된 핵산 서열"은 따라서 정제된 핵산서열일 수 있다. "실질적으로 정제된" 분자는 이들이 자연적으로 회합된 다른 성분으로부터 적어도 60% 유리되고, 바람직하게는 적어도 75% 유리되며, 더욱 바람직하게는 적어도 90% 유리된다. 본 명세서에 사용된 바와 같이, 용어 "정제된" 또는 "정제하기 위해"는 시료로부터 오염물의 제거를 또한 의미한다. 단백질을 오염시키는 것의 제거는 시료에서 관심 폴리펩티드 또는 핵산의 백분율의 증가를 야기한다. 다른 예로, 재조합 폴리펩티드는 식물, 세균, 효모 또는 포유류 숙주세포에서 발견되고, 폴리펩티드는 숙주 세포 단백질 제거에 의해 정제되며, 이 시료에서 재조합 폴리펩티드의 백분율은 그로 인해 증가된다.
- [0146] 용어 소정의 폴리뉴클레오타이드 시퀀스 또는 폴리펩티드를 "포함하는 조성물"은 소정의 폴리뉴클레오타이드 서열 또는 폴리펩티드를 포함하는 임의의 조성물을 광범위하게 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 포함할 수 있다.
- [0147] 용어 "시료"가 이의 가장 넓은 의미에서 사용된다. 한 의미에서, 이는 동물 세포 또는 조직을 의미할 수 있다. 다른 의미에서, 생물학적 및 환경적 시료뿐만 아니라 임의의 공급원으로부터 수득한 표본 또는 배양물을 포함하

는 것을 의미한다. 생물학적 시료는 식물 또는 동물(인간을 포함)로부터 획득할 수 있고, 유체, 고체, 조직 및 기체를 망라한다. 환경적 시료는 물질, 토양, 물, 및 공업용 시료와 같은 환경적 물질을 포함한다. 이러한 시료는 본 발명에 적용가능한 시료 유형을 제한하는 것으로 해석되는 것은 아니다.

[0148] 본 명세서에서 사용되는 바와 같이, 일부 문맥에서 사용된 "원격 시료"는 시료의 세포, 조직, 또는 기관 공급원이 아닌 부위로부터 간접적으로 수집된 시료에 관련된다. 예를 들면, 채장으로부터 유래된 시료 물질이 대변 시료에서 평가되는 경우(예를 들면, 채장에서 직접 취한 시료가 아님), 이 시료는 원격 시료다.

[0149] 본 명세서에서 사용된 바와 같이, 용어 "환자(patient)" 또는 "피험자 또는 개체 (subject)"는 본 발명의 기술에 의해서 제공되는 다양한 시험의 대상이 되는 유기체를 의미한다. 용어 "피험자" 또는 "개체"는 동물을 포함하고, 바람직하게는 인간을 비롯한 포유류이다. 바람직한 양태에서, 피험자 또는 개체는 영장목(primates)이다. 더욱 바람직한 양태에서, 피험자 또는 개체는 인간이다.

[0150] 본 명세서에서 사용된 바와 같이, 용어 "키트"는 물질을 전달하기 위한 임의의 전달 시스템을 의미한다. 반응 검정의 문맥에서, 이러한 전달 시스템은 하나의 위치에서 다른 위치로 반응 시약(예를 들면, 적절한 용기에서 예를 들어, 올리고뉴클레오타이드, 효소 등) 및/또는 지지 물질(예를 들면, 버퍼, 검정 등을 수행하기 위한 작성된 지시사항)의 저장, 전송, 또는 전달을 허용하는 시스템을 포함한다. 예를 들면, 키트는 관련된 반응 시약 및/또는 지지 물질을 함유하는 하나 이상의 봉입물(예를 들면, 박스)을 포함한다. 본 명세서에서 사용된 바와 같이, 용어 "단편화 키트"는 전체 키트 구성의 하위구성(subportion)을 각각 포함하는 2개 이상의 별개의 용기를 포함하는 전달 시스템을 의미한다. 용기는 의도된 수령인에게 함께 또는 별도로 전달될 수 있다. 예를 들면, 제1 용기는 검정에 사용하기 위한 효소를 포함하는 반면, 제2 용기는 올리고뉴클레오타이드를 함유한다. 용어 "단편화 키트"는 미국 식품의약청(Federal Food, Drug, and Cosmetic Act)의 520(e)절 하에 규제되는 분석물 특이적 시약(Analyte specific reagents, ASR)을 함유하는 키트를 망라하는 것으로 의도되나, 이로 한정되는 것은 아니다. 실제, 전체 키트 구성의 하위구성을 각각 포함하는 2개 이상의 개별 용기를 포함하는 임의의 전달 시스템은 용어 "단편화 키트"에 포함된다. 반면, "합쳐진 키트"는 단일 용기(예를 들면, 각각의 목적하는 성분을 수용하는 단일 박스에서)에서 반응 검정의 모든 성분을 함유하는 전달 시스템을 의미한다. 용어 "키트"는 단편화되고 합쳐진 키트 모두를 포함한다.

[0151] 기술의 양태

[0152] 췌장암을 필연적으로 높은 사망률을 갖는 본질적으로 빠르게 자라는 악성 종양으로 보는 전통적인 관점에 대조적으로, 최근의 계산 모델(computational model)은 전암 발달(precancer development), 암 진화(cancer evolution), 및 전이를 통한 분자적 개시로부터 진행에 대해 평균적으로 20년이 걸릴 수 있음을 제시하며 (Yachida, S., et al., Nature, 2010. 467(7319): p. 1114-1117를 참고) 이는 대장 신생(colorectal neoplasia)의 성장률과 발달사와 흡사하다 (Stryker, S.J., et al., Gastroenterology, 1987. 93(5): p. 1009-13; Bozic, I., et al., Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(43): p. 18545-50를 참고). 전암 및 초기 T1암 단계 모두에서 긴 증상 전 체류시간(pre-symptomatic dwell time)은 스크린 검출(screen detection)에 대해 상대적으로 긴 절호의 기회를 사실상 제공할 수 있다.

[0153] 췌장에서 암 전구체 병변은 조직학 적으로 잘 기술되어 있다 (Zamboni, G., et al., Clinical gastroenterology, 2013. 27(2): p. 299-322을 참고). 이러한 병변은 현저히 통상적으로, 일반 대중의 13 - 25%에서 발생하며, 췌장암 예방으로 잠재적으로 번역될 수 있는 스크리닝 표적을 제안한다. 그러나, 이들의 검출은 현재 우연한 것이며 이들의 관리는 문제가 많다. 이들은 증상을 거의 야기하지 않고 현재 주로 우연한 발견으로 검출되고 있다. 가장 일반적인 유형인 췌장 상피내 종양(pancreatic intra-epithelial neoplasia, PanIN)은 영상화로 검출하기 어렵고 주로 조직학적인 발견이다. 영상화로 검출가능한 전암은 일반적으로 낭포성이고 대개 무해한 장액 낭포성 종양(serous cystadenomas), 진행의 중간 등급 위험(intermediate risk)을 갖는 췌장관내 유두상 점액 종양(intraductal papillary mucinous neoplasms, IPMNs), 및 높은 악성 잠재성을 갖는 점액 낭포성 신생물을 포함한다(Morris-Stiff, G., et al., HPB : the official journal of the International Hepato Pancreato Biliary Association, 2013. 15(3): p. 175-81을 참고). 대부분 전암은 진행하지 않고, 위험 예측에 대한 가능한 접근법이 불완전하다. 방사성 구조적 특징(radiographic structural feature), 세포학적 시료 채취(cytologic sampling), 또는 분자적 검사(molecular testing)를 기반으로 하는 현재 사용되는 평가 시스템(scoring system)은 조직학을 정확히 예측하는데 실패했다 (Correa-Gallego, C., et al., Pancreatology : official journal of the International Association of Pancreatology, 2010. 10(2-

3): p. 144-50을 참고). 따라서, 췌장 절제(pancreatic resection)를 촉발시키는 많은 병변은 비신생물성이 되거나 고등급 이형성증을 포함하지 않음을 증명했다 (Correa-Gallego, C., et al., Pancreatology : official journal of the International Association of Pancreatology, 2010. 10(2-3): p. 144-50을 참고).

- [0154] 췌장암 진행의 느린 자연적 발달사는 치료 가능한 췌장 신생의 검출을 위한 충분한 기회를 제공한다. 이 혁신적인 도전은 높은 임상적 감도와 절제에 가장 적합한 고등급 이형성증이 있는 전암을 정확하게 구별하는 능력이 있는 비침습적이거나 최소한으로 침습적인 도구를 개발하는 것이다.
- [0155] 본 발명에 대한 양태를 개발하는 과정에서 수행된 실험은 IPMN-HGD, PanIN-3, 또는 PDAC을 갖는 개체의 췌장 조직으로부터 DNA 마커의 메틸화 상태를 대조 개체[예를 들어, 각 조직 유형에 대해 IPMN-저등급 이형성증 (IPMN-low grade dysplasia, IPMN-LGD), PanIN-1, 및 PanIN-2를 갖는 개체]로부터 동일한 DNA 마커의 메틸화 상태를 비교하였다. 이러한 실험은 고등급 전구체 (IPMN-HGD, PanIN-3) 또는 침습성 암 (PDAC)이 있는 환자군을 정상적 췌장 또는 저등급 전구체 병변 [IPMN-저등급 이형성증 (IPMN-low grade dysplasia, IPMN-LGD), PanIN-1, 및 PanIN-2]이 있는 대조군으로부터 구별하는 메틸화된 DNA 마커 후보를 식별하고 검증하였다(실시예 I, II, IV, 및 V를 참고).
- [0156] 따라서, 본 개시내용은 췌장 고등급 이형성증의 스크리닝(예를 들어, 감시)을 위한 기술, 특히 개체에서 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)의 존재 검출에 대한 방법, 조성물 및 관련 용도에 관한 것이나, 이로만 제한되는 것은 아니다.
- [0157] 마커 및/또는 마커의 패널(예를 들어, 표2와 6에서 제공되는 주석을 갖는 염색체 부위)을 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)을 검출할 수 있는 것으로 식별하였다 (실시예 I, II, IV 및 V를 참고) (BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781).
- [0158] 본 명세서의 개시가 어떤 도시된 양태를 나타냄에도 불구하고, 이러한 양태가 한 예로서 존재하며 제한의 방법으로서 존재하는 것이 아님이 이해될 것이다.
- [0159] 방법은 개체로부터 분리된 생물학적 시료에서 적어도 하나의 메틸화 마커의 메틸화 상태를 결정하는 단계를 포함하며, 마커의 메틸화 상태의 변화가 존재를 나타내거나 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)의 종류를 나타낸다. 특정 양태는 차등 메틸화 영역 [(differentially methylated region, DMR), 예를 들어, DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 92, 95, 및 96 (표 2와 6으로부터)]을 포함하는 진단 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 진단 (예, 스크리닝)에 사용되는 마커와 관련된다.
- [0160] 본 명세서에 제공되고 표 2와 6에 나열된 DMR(예를 들어, DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 92, 95 및 96)을 포함하는 적어도 하나 이상의 마커, 마커의 영역, 또는 마커의 염기의 메틸화 분석이 분석되는 양태 외에도, 본 발명의 기술은 개체에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 검출에 대한 유용성이 있는 DMR을 포함하는 적어도 하나의 마커, 마커의 영역, 마커의 염기를 포함하는 마커의 패널을 또한 제공한다.
- [0161] 본 발명의 기술의 일부 양태는 DMR을 포함하는 적어도 하나의 마커, 마커의 영역, 또는 마커의 염기의 CpG 메틸화 상태의 분석을 기반으로 한다.
- [0162] 일부 양태에서, 본 발명의 기술은 DMR[예를 들어, 표2와 6에서 제공됨 (예를 들어, DMR 1-96)]을 포함하는 적어도 하나 이상의 마커 내에서 CpG 디뉴클레오타이드 서열의 메틸화 상태를 결정하는 하나 이상의 메틸화 검정이 있는 조합에서 바이셀파이트 기법의 사용을 제공한다. 유전체 CpG 디뉴클레오타이드는 메틸화되거나 메틸화되지 않을 수 있다 (대안적으로 각각 업-메틸화된(up-methylated) 및 다운-메틸화된(down-methylated)으로 알려짐). 그러나, 본 발명의 방법은 원격 시료(예를 들어, 혈액, 장기 유출액(organ effluent), 또는 대변)의 백그라운드 내에서 이질성(heterogeneous nature)의 생물학적 시료, 예를 들어, 낮은 농도의 종양 세포, 또는 이로부터의 생물학적 물질의 분석에 적합하다. 따라서, 이러한 시료내에서 CpG 위치의 메틸화 상태를 분석하는 경우, 특정 CpG 위치에서 메틸화의 수준(예를 들어, 백분율, 분수, 비, 비율, 또는 정도)을 결정하는 정량적 검정을 사용할 수 있다.
- [0163] 본 발명의 기술에 따르면, DMR을 포함하는 마커에서 CpG 디뉴클레오타이드의 서열의 메틸화 상태의 결정은 개체에서 췌장 고등급 이형성증 (IPMN-HGD, PanIN-3, 또는 PDAC)의 특정 및 진단 모두에서 유용성을 갖는다.

- [0164] **마커의 조합**
- [0165] 일부 양태에서, 본 발명의 기술은 표 2와 6으로부터 둘 이상의 DMR (예, DMR No. 1-96로부터 둘 이상의 DMR)을 포함하는 마커 조합의 메틸화 상태를 평가하는 것에 관련된다. 일부 양태에서, 하나 이상의 마커의 메틸화 상태를 평가하는 것은 개체에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 존재를 식별하는 스크린 또는 진단의 특이도 및/또는 감도를 증가시킨다.
- [0166] 다양한 암은 마커의 다양한 조합에 의해 예측될 수 있으며 이는 예를 들어, 예측의 특이도 및 감도에 관련된 통계적 기법에 의해 식별되었다. 본 발명의 기술은 예측적 조합을 식별하는 방법을 제공하고 일부 암에 대한 예측적 조합을 검증하였다.
- [0167] 일부 양태에서, 마커의 조합(예를 들어, DMR을 포함)은 신생물의 부위를 예측한다.
- [0168] 예를 들어, 마커 및/또는 마커의 패널은 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 검출할 수 있는 것으로 식별되었다 (예, 표 2와 6에 제공된 주석을 갖는 염색체 부위) (실시예 I, II, IV 및 V을 참고) (BMP3, NDRG4, ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, LRRC4, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781).
- [0169] **메틸화 상태를 검정하는 방법**
- [0170] 5' 메틸시토신의 존재에 대해 핵산을 분석하는 가장 빈번히 사용되는 방법은 Frommer등의 DNA에서 5-메틸시토신의 검출 [Frommer, et al. for the detection of 5-methylcytosines in DNA (Frommer et al. (1992) Proc. Natl. Acad. Sci. USA 89: 1827-31)]에 기재된 바이설파이트 방법 또는 이의 변형에 기초한다. 5-메틸시토신을 매핑(mapping)하는 바이설파이트 방법은 5-메틸시토신이 아닌 시토신이 하이드로젠 설파이트 이온과 반응하는 (또한, 바이설파이트로 알려짐) 관측에 기초한다. 반응은 대개 다음의 단계를 따라 수행된다: 첫째, 시토신을 하이드로젠 설파이트와 반응시켜 술폰화된 시토신을 생성한다. 다음, 술폰화된 반응 중간체의 자발적인 탈아미노화가 술폰화된 우라실을 야기한다. 마지막으로, 술폰화된 우라실을 염기성 조건에서 탈술폰화시켜 우라실을 생성한다. 5-메틸시토신이 구아닌과 염기쌍을 형성하는 반면(따라서, 시토신처럼 작용함), 우라실은 아데닌과 염기쌍을 형성하기 때문에(따라서, 티민처럼 작용함) 검출이 가능하다. 이는 메틸화된 시토신을 메틸화되지 않은 시토신으로부터 구별하며, 바이설파이트 유전체 시퀀싱(Grigg G, & Clark S, Bioessays (1994) 16: 431-36; Grigg G, DNA Seq. (1996) 6: 189-98) 또는 미국 특허 제5,786,146호에 개시된 메틸화 특이적 PCR(methylation-specific PCR, MSP)에 의해 가능하다.
- [0171] 일부 종래 기술은 분석될 DNA를 아가로오스 기질에 봉입하는 단계, 이로써 DNA의 확산과 원형회복(바이설파이트는 단일가닥 DNA와만 반응함)을 예방하는 단계, 침전과 정제 단계를 빠른 투석(fast dialysis)으로 대체하는 단계를 포함하는 방법에 관련된다(Olek A, et al. (1996) "A modified and improved method for bisulfite based cytosine methylation analysis" Nucleic Acids Res. 24: 5064-6). 따라서, 메틸화 상태에 대한 개별적인 세포를 분석하는 것이 가능하며, 이는 방법의 유용성과 감도를 도시한다. 5-메틸시토신을 검출하는 종래의 방법의 개요가 Rein, T., et al. (1998) Nucleic Acids Res. 26: 2255에 의해 제공된다.
- [0172] 바이설파이트 기법은 바이설파이트 처리 후에 알려진 핵산의 짧고 특정한 단편을 증폭시키는 것에 이어서 시퀀싱에 의한 산물(Olek & Walter (1997) Nat. Genet. 17: 275-6) 또는 프라이머 신장 반응(Gonzalzo & Jones (1997) Nucleic Acids Res. 25: 2529-31; WO 95/00669호; 미국 특허 제6,251,594호)을 검정하는 것을 일반적으로 수반하여 개별적인 시토신 위치를 분석한다. 일부 방법은 효소적 절단을 사용한다 (Xiong & Laird (1997) Nucleic Acids Res. 25: 2532-4). 혼성에 의한 검출은 또한 기술분야에 기재되었다 (Olek등의 WO 99/28498호). 또한, 개별적인 유전자에 대한 메틸화 검출을 위한 바이설파이트 기술의 사용이 기재되었다 (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 9746705호; WO 9515373호).
- [0173] 다양한 메틸화 검정 절차는 기술분야에 알려져 있으며, 본 발명의 기술에 따라서 바이설파이트 처리와 함께 사용될 수 있다. 이러한 검정은 핵산 서열내에서 하나 또는 복수의 CpG 디뉴클레오타이드(예를 들어, CpG 섬)의 메틸화 상태의 결정을 허용한다. 이러한 검정은 다른 기술 사이에서, 바이설파이트 처리된 핵산의 시퀀싱, PCR (서열 특이적 증폭), 서던 블랏 분석, 및 메틸화 민감성 제한 효소의 사용을 수반한다.
- [0174] 예를 들어, 유전체 시퀀싱은 바이설파이트 처리를 이용하여 메틸화 패턴 및 5-메틸시토신 분포의 분석에 대해 단순화되었다 (Frommer et al. (1992) Proc. Natl. Acad. Sci. USA 89: 1827-1831). 더욱이, 바이설파이트 전

환 DNA로부터 증폭된 PCR 생성물의 제한효소 절단은, 예를 들어, Sadri & Hornsby (1997) Nucl. Acids Res. 24: 5058-5059에 의해 기술된 또는 COBRA [합쳐진 바이설파이트 제한 분석 (Combined Bisulfite Restriction Analysis)] (Xiong & Laird (1997) Nucleic Acids Res. 25: 2532-2534)으로 알려진 방법에서 구현되는 바와 같이 메틸화 상태를 평가하는데 사용된다.

[0175] COBRA™ 분석은 소량의 유전체 DNA 내 특정 유전자 좌위에서 DNA 메틸화 수준 결정에 유용한 정량적인 메틸화 검정이다(Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997). 간단하게, 제한 효소 절단은 소듐 바이설파이트 처리된 DNA의 PCR 생성물의 메틸화 의존적인 서열의 차이점을 밝히는데 사용된다. 메틸화 의존적 서열 차이는 Frommer 등에 의해 기술된 절차에 따라 기준 바이설파이트 처리(standard bisulfite treatment)에 의해 유전체 DNA로 처음에 도입되었다 (Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992). 바이설파이트 전환된 DNA의 PCR 증폭은 그 다음에 관심 CpG 섬에 특이적인 프라이머를 사용하여 수행하고, 이어서 제한 엔도뉴클레아제 절단(restriction endonuclease digestion), 전기영동, 및 특이적인 표지된 혼성화 탐침을 사용한 검출이 수행된다. 본래의 DNA 시료에서 메틸화 수준은 DNA 메틸화 수준의 넓은 스펙트럼에 걸친 선형 정량적 방식으로 절단된 PCR 생성물과 절단되지 않은 PCR 생성물의 상대적인 양에 의해 나타난다. 게다가, 이 기법은 현미경 해부된 파라핀 포매 조직 시료 (microdissected paraffin-embedded tissue sample)로부터 수득한 DNA에 확실히 적용될 수 있다.

[0176] COBRA™ 분석을 위한 일반적인 시약 (예를 들어, 일반적으로 COBRA™에 기초한 키트에서 찾을 수 있는 바와 같음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이설파이트 처리된 DNA 서열, CpG 섬 등)에 대한 PCR 프라이머; 제한 효소 및 적합한 완충액; 유전자 혼성화 올리고뉴클레오티드; 대조 혼성화 올리고뉴클레오티드; 올리고뉴클레오티드 탐침에 대한 키나아제 표지 키트; 및 표지된 뉴클레오티드를 포함하나, 이에 한정되는 것은 아니다. 더욱이, 바이설파이트 전환 시약은 DNA 변성 완충액; 술폰화 완충액; DNA 회복 시약(DNA recovery reagent) 또는 키트 (예를 들어, 침전, 한외 여과(ultrafiltration), 친화도 컬럼); 탈술폰화 완충액; 및 DNA 회복 구성요소;를 포함할 수 있다.

[0177] 바람직하게는, "MethylLight™"[형광에 기초한 실시간 PCR 기법 (Eads et al., Cancer Res. 59:2302-2306, 1999)], Ms-SNuPE™ [메틸화 민감성 단일 뉴클레오티드 프라이머 신장 (Methylation-sensitive Single Nucleotide Primer Extension) 반응 (Gonzalzo & 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 제5,786,146호), 및 메틸화 CpG 섬 증폭 ("MCA"; Toyota et al., Cancer Res. 59:2307-12, 1999)과 같은 검정은 단독으로 또는 하나 이상의 이러한 방법들의 조합으로 사용될 수 있다.

[0178] "HeavyMethyl™" 검정 기술은 바이설파이트 처리된 DNA의 메틸화 특이적 증폭에 기초한 메틸화 차이를 평가하는 정량적인 방법이다. 증폭 프라이머 사이의 CpG 위치를 커버하는 또는 이에 의해 커버된 메틸화 특이적 블로킹 탐침("블로커(blocker)")은 핵산 시료의 메틸화 특이적 선택적 증폭을 가능하게 한다.

[0179] 용어 "HeavyMethyl™ MethylLight™" 검정은 MethylLight™ 검정의 변형인 HeavyMethyl™ MethylLight™ 검정을 의미하고, MethylLight™ 검정은 증폭 프라이머 사이의 CpG 위치를 커버하는 메틸화 특이적 블로킹 탐침과 합쳐진다. HeavyMethyl™ 검정은 메틸화 특이적 증폭 프라이머와의 조합에서 또한 사용될 수 있다.

[0180] HeavyMethyl™ 분석에 대한 일반적인 시약(예를 들어, 일반적인 MethylLight™에 기초한 키트에서 찾을 수 있음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이설파이트 처리된 DNA 서열, CpG 섬, 또는 바이설파이트 처리된 DNA 서열 또는 CpG 섬 등)에 대한 PCR 프라이머; 블로킹 올리고뉴클레오티드; 최적화된 PCR 완충액과 데옥시뉴클레오티드; 및 Taq 중합효소를 포함하나, 이에 한정되는 것은 아니다.

[0181] MSP[메틸화 특이적 PCR (methylation-specific PCR)]은 CpG 섬 내의 CpG 부위의 사실상 임의의 그룹의 메틸화 상태를 평가하는 것을 허용하며, 이는 메틸화 민감성 제한 효소의 사용과 관계없다 (Herman et al. Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; 미국 특허 제5,786,146호). 간단하게, DNA는 메틸화된 시토신이 아닌 메틸화되지 않은 시토신을 우라실로 전환하는 소듐 바이설파이트에 의해 변형되며, 생성물은 메틸화되지 않은 DNA에 비해 메틸화된 DNA에 특이적인 프라이머로 차후에 증폭될 수 있다. MSP는 소량의 DNA만을 필요로 하고, 소정의 CpG 섬 유전자 좌위의 0.1% 메틸화된 대립유전자에 민감하며, 파라핀 포매 시료로부터 추출된 DNA에서 수행될 수 있다. MSP 분석을 위한 일반적인 시약 (예를 들어, 일반적인 MSP에 기초한 키트에서 발견될 수 있음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이설파이트 처리된 DNA 서열, CpG 섬 등)에 대한 메틸화된 및 메틸화되지 않은 PCR 프라이머; 최적화된 PCR 완충액과 데옥시뉴클레오티

드; 및 특정 탐침을 포함하나, 이에 한정되는 것은 아니다.

[0182] MethyLight™ 검정은 PCR 단계 후 추가적인 조작이 필요 없는 형광에 기초한 실시간 PCR(예를 들어, TaqMan®)을 이용하는 대량 신속처리 정량적 메틸화 검정(high-throughput quantitative methylation assay)이다 (Eads et al., Cancer Res. 59:2302-2306, 1999). 간단하게, MethyLight™ 과정은 기준 절차에 따른 소듬 바이선폴과이트 반응에서 메틸화 의존적 서열 사이의 혼합 풀(mixed pool)로 전환된 유전체 DNA의 혼합된 시료로 시작한다 (바이선폴과이트 과정은 메틸화되지 않은 시토신 잔기를 우라실로 전환시킴). 형광에 기초한 PCR은 그 다음 “편향된” 반응 (“biased” reaction)에서 예를 들어, 알려진 CpG 디뉴클레오티드와 겹쳐지는 프라이머로 수행된다. 서열 구별은 증폭 과정의 수준과 형광 검출 과정의 수준 모두에서 일어난다.

[0183] MethyLight™ 검정은 핵산, 예를 들어, 유전체 DNA 시료에서 메틸화 패턴에 대한 정량적 시험으로 사용되며, 서열 구별은 탐침 혼성화의 수준에서 일어난다. 정량적 버전(quantitative version)에서, PCR 반응은 메틸화 부위로 추정되는 특정한 부위와 겹치는 형광 탐침이 존재하는 메틸화 특이적 증폭을 제공한다. 인풋 DNA(input DNA)량에 대한 비편향된(unbiased) 대조는 프라이머도 아니고 탐침도 아닌 임의의 CpG 디뉴클레오티드 위에 겹쳐지는(overlie) 반응에 의해 제공된다. 대안적으로, 유전체 메틸화에 대한 정량적 시험은 알려진 메틸화 부위를 커버하지 않는 대조 올리고뉴클레오티드(예를 들어, HeavyMethyl™ 및 MSP 기법의 형광에 기초한 버전) 또는 잠재적 메틸화 부위를 커버하는 올리고뉴클레오티드가 있는 편향된 PCR 풀을 탐침함으로써(probing) 달성될 수 있다.

[0184] MethyLight™ 과정은 임의의 적합한 탐침(예를 들어, “TaqMan®” 탐침, Lightcycler® 탐침 등)과 사용된다. 예를 들어, 일부 적용에서, 이중가닥 유전체 DNA는 소듬 바이선폴과이트로 처리되고, TaqMan® 탐침을 이용한 PCR 반응 두 세트 중 하나-예를 들어, MSP 프라이머 및/또는 헤비메틸 블로커(HeavyMethyl blocker) 올리고뉴클레오티드와 TaqMan® 탐침-의 대상이 된다. TaqMan® 탐침은 형광 “보고(reporter)” 분자 및 “소광(quencher)” 분자로 이중 표지되고 상대적으로 높은 GC 함유 영역에 대해 특이적하도록 설계되어서, PCR 사이클에서 정방향 또는 역방향 프라이머보다 약 10° C 높은 온도에서 녹는다. 이는 PCR 어닐링/신장 단계 시 TaqMan® 탐침을 완전 혼성화로 남도록 허용한다. PCR 동안에 Taq 중합효소가 새로운 가닥을 효소적으로 합성함에 따라, 이는 결국에 어닐링된 TaqMan® 탐침에 도달할 것이다. Taq 중합효소의 5' 에서 3'로 엔도뉴클레아제 활성은 그 다음 TaqMan® 탐침을 절단하여 이를 대체하고 실시간 형광 검출 시스템을 이용하여 이의 지금 소광되지 않은(unquenched) 신호의 정량적 검출에 대한 형광 보고 분자를 방출할 것이다.

[0185] MethyLight™ 분석에 대한 일반적인 시약 (예를 들어, 일반적인 MethyLight™에 기초한 키트에서 발견될 수 있음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이선폴과이트 처리된 DNA 서열, CpG 섬 등)에 대한 PCR 프라이머; TaqMan® 또는 Lightcycler® 탐침; 최적화된 PCR 완충액 및 데옥시뉴클레오티드; 및 Taq 중합효소를 포함하나, 이에 한정되는 것은 아니다.

[0186] QM™ (정량적 메틸화(quantitative methylation)) 검정은 유전체 DNA 시료에서 메틸화 패턴에 대한 대안적인 정량적 시험이며, 서열 구별이 탐침 혼성화의 수준에서 일어난다. 이 정량적 버전에서, PCR 반응은 메틸화 부위로 추정되는 특정한 부위와 겹치는 형광 탐침이 존재하는 비편향된 증폭을 제공한다. 인풋 DNA(input DNA)량에 대한 비편향된(unbiased) 대조는 프라이머도 아니고 탐침도 아닌 임의의 CpG 디뉴클레오티드 위에 겹쳐지는(overlie) 반응에 의해 제공된다. 대안적으로, 유전체 메틸화에 대한 정량적 시험은 알려진 메틸화 부위를 커버하지 않는 대조 올리고뉴클레오티드(HeavyMethyl™ 및 MSP 기법의 형광에 기초한 버전) 또는 잠재적 메틸화 부위를 커버하는 올리고뉴클레오티드가 있는 편향된 PCR 풀을 탐침함으로써 달성될 수 있다.

[0187] QM™ 과정은 임의의 적합한 탐침, 예를 들어, “TaqMan®” 탐침, Lightcycler® 탐침과 증폭 과정에서 사용된다. 예를 들어, 이중가닥 유전체 DNA는 소듬 바이선폴과이트로 처리되고, 비편향된 프라이머와 TaqMan® 탐침의 대상이 된다. TaqMan® 탐침은 형광 “보고(reporter)” 분자 및 “소광(quencher)” 분자로 이중 표지되고, 상대적으로 높은 GC 함유 영역에 대해 특이적하도록 설계되어서, PCR 사이클에서 정방향 또는 역방향 프라이머보다 약 10° C 높은 온도에서 녹는다. 이는 PCR 어닐링/신장 단계 시 TaqMan® 탐침을 완전 혼성화로 남도록 허용한다. PCR 동안에 Taq 중합효소가 새로운 가닥을 효소적으로 합성함에 따라, 이는 결국에 어닐링된 TaqMan® 탐침에 도달할 것이다. Taq 중합효소의 5' 에서 3'의 엔도뉴클레아제 활성은 그 다음 TaqMan® 탐침을 절단하여 이를 대체하고 실시간 형광 검출 시스템을 이용하여 이의 지금 소광되지 않은(unquenched) 신호의 정량적 검출에 대한 형광 보고 분자를 방출할 것이다. QM™ 분석에 대한 일반적인 시약 (예를 들어, 일반적인 QM™에 기초한 키트에서 찾을 수 있음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이선폴과이트 처리된 DNA 서열, CpG 섬 등)에 대한 PCR 프라이머; TaqMan® 또는 Lightcycler® 탐침; 최

적화된 PCR 완충액 및 데옥시뉴클레오타이드; 및 Taq 중합효소를 포함하나, 이에 한정되는 것은 아니다.

[0188] Ms-SNuPE™ 기법은 DNA의 바이설파이트 처리에 기초하여 특정 CpG 부위에서 메틸화 차이를 평가하고, 이어서 단일 뉴클레오타이드 프라이머 신장을 하는 정량적 방법이다 (Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997). 간단하게, 유전체 DNA는 소듐 바이설파이트와 반응하여 5-메틸시토신은 변화시키지 않는 반면, 메틸화되지 않은 시토신을 우라실로 전환한다. 목적하는 표적 서열의 증폭은 이어서 바이설파이트 전환 DNA에 특이적인 PCR 프라이머를 이용하여 수행하고, 수득한 생성물은 분리되어 관심 CpG 부위에서 메틸화 분석에 대한 템플레이트로 사용된다. 소량의 DNA는 분석될 수 있고(예를 들어, 현미경 해부된 병리 절편(microdissected pathology section)), 이는 CpG 부위에서 메틸화 상태를 결정하는 제한 효소의 이용을 방지한다.

[0189] Ms-SNuPE™ 분석에 대한 일반적인 시약 (예를 들어, 일반적인 Ms-SNuPE™에 기초한 키트에서 찾을 수 있음)은 특정 유전자 좌위(예를 들어, 특정 유전자, 마커, DMR, 유전자 영역, 마커 영역, 바이설파이트 처리 DNA 서열, CpG 섬 등)에 대한 PCR 프라이머; 최적화된 PCR 완충액 및 데옥시뉴클레오타이드; 겔 추출 키트; 양성 대조 프라이머; 특정 유전자 좌위에 대한 Ms-SNuPE™ 프라이머; 반응 완충액 (Ms-SNuPE반응을 위한), 및 표지된 뉴클레오타이드를 포함하나, 이에 한정되는 것은 아니다. 또한, 바이설파이트 전환 시약은 DNA 변성 완충액; 술포화 버퍼; DNA 회복 시약 또는 키트(예를 들어, 침전, 한외여과, 친화도 컬럼); 탈술포화 완충액; 및 DNA 회복 구성 요소를 포함할 수 있다.

[0190] 감소된 표현 바이설파이트 시퀀싱(Reduced Representation Bisulfite Sequencing, RRBS)은 핵산의 바이설파이트 처리와 시작하여 모든 메틸화되지 않은 시토신을 우라실로 전환하고, 이어서 제한 효소 절단(예를 들어, MspI와 같은 CG 서열을 포함하는 부위를 인식하는 효소에 의함)과 어댑터 리간드에 결합(coupling) 후, 단편의 완전 시퀀싱(complete sequencing)을 한다. 제한 효소의 선택은 CpG 밀집 영역에 대한 단편을 강화하고, 이는 분석시에 다수의 유전자 위치에 매핑할 수 있는 불필요한 서열(redundant sequence)의 수를 감소시킨다. 이와 같이, RRBS는 핵산 시료의 복잡도(complexity)를 시퀀싱을 위한 제한 단편의 부분집합을 선택함으로써(예를 들어, 예비적인 전기영동을 이용한 크기 선택(size selection)) 감소시킨다. 온전한 유전체 바이설파이트 시퀀싱과는 대조적으로, 제한 효소 절단에 의해 생성된 모든 단편이 적어도 하나의 CpG 디뉴클레오타이드에 대한 DNA 메틸화 정보를 함유할 수 있다. 이와 같이, RRBS는 이러한 영역에서 높은 빈도의 제한 효소 절단 부위가 있는 프로모터, CpG 섬, 및 다른 유전체 특성에 대한 시료를 강화하고, 따라서 검정을 제공하여 하나 이상의 유전체 좌위의 메틸화 상태를 평가한다.

[0191] RRBS를 위한 일반적인 프로토콜은 핵산 시료를 MspI와 같은 제한 효소로 절단시키는 단계, 돌출부(overhang)을 채우고 A-테일링(A-tailing)하는 단계, 어댑터를 라이게이션하는 단계, 바이설파이트 전환단계, 및 PCR 하는 단계를 포함하며, 예를 들어, 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: 5868-77를 참고할 수 있다.

[0192] 일부 양태에서, 정량적 대립유전자 특이적 실시간 표적 및 신호 증폭 (quantitative allele-specific real-time target and signal amplification, QuARTS) 검정은 메틸화 상태를 평가에 사용된다. 세계의 반응이 각 QuARTS 검정에서 연속하여 일어나며, 이는 일차 반응에서 증폭(반응 1) 및 표적 탐침 절단(cleavage) (반응 2); 이차 반응에서 FRET 절단 및 형광 신호 생성(반응 3);을 포함한다. 표적 핵산이 특정 프라이머로 증폭될 경우, 플랩 서열(flap sequence)이 있는 특정한 검출 탐침은 증폭산물에 느슨하게 결합한다. 표적 결합 부위에서 특정한 침습적 올리고뉴클레오타이드의 존재는 분할을 야기하여 검출 탐침과 플랩 시퀀스 사이를 절단함으로써 플랩 시퀀스를 방출한다. 플랩 서열은 대응하는 FRET 카세트 (cassette)의 헤어핀이 아닌 부분에 상보적이다. 따라서, 플랩 서열은 FRET 카세트 상에서 침습적 올리고뉴클레오타이드로서 기능하고 소광물질과 FRET 카세트 형광단 (FRET cassette fluorophore) 사이의 분할을 가져오며, 이는 형광 신호를 생성한다. 분할 반응은 표적당 복수 탐침을 자를 수 있고, 따라서 플랩당 복수 형광단을 방출할 수 있고, 이는 지수적 신호 증폭을 제공한다. QuARTS는 하나의 반응 웰에서 FRET 카세트를 상이한 염료와 이용하여 복수 표적을 검출할 수 있다. 예를 들어, Zou et al. (2010) "Sensitive quantification of methylated markers with a novel methylation specific technology" Clin Chem 56: A199; 미국 특허 제12/946,737호, 제12/946,745호, 제12/946,752호, 및 제 61/548,639호를 참고할 수 있다.

[0193] 용어 "바이설파이트 시약"은 바이설파이트, 디설파이트, 하이드로젠 설파이트, 또는 이들의 조합물을 포함하는 시약을 의미하며, 본 명세서에 기재된 바와 같이 메틸화된 및 메틸화되지 않은 CpG 디뉴클레오타이드 서열 구

별에 유용하다. 상기 치료의 방법은 기술분야에 알려져 있다 (예를 들어, PCT/EP2004/011715호). 바람직한 것은 바이설파이트 처리가 n-알킬렌글리콜(n-alkylenglycol) 또는 디에틸렌 글리콜 디메틸 에테르(diethylene glycol dimethyl ether, DME)을 비롯하나 이에 한정되지 않는 변성 용액의 존재에서, 또는 디옥산 또는 디옥산 유도체의 존재에서 수행되는 것이다. 일부 양태에서, 변성 용액은 1%와 35% (v/v) 사이의 농도에서 사용될 수 있다. 일부 양태에서, 바이설파이트 반응은 크로만 유도체, 예를 들어, 6-히드록시-2,5,7,8,-테트라메틸크로만 2-카복실산 또는 트리히드록시벤존산 및 이의 유도체, 예를 들어, 갈산 (PCT/EP2004/011715호를 참고)을 비롯하나 이들에 한정되지 않는 스캐빈저(scavenger)의 존재하에 수행될 수 있다. 바이설파이트 전환은 바람직하게는 30° C와 70° C 사이의 반응 온도에서 수행될 수 있고, 그에 의해 온도는 반응 시에 짧은 시간동안 85° C 넘도록 증가된다 (PCT/EP2004/011715호를 참고). 바이설파이트 처리된 DNA는 바람직하게는 정량전에 정제된다. 이는 기술분야에 알려진 임의의 수단에 의해 수행될 수 있으며, 그 예는 Microcon™ 컬럼(Millipore™에 의해 제작됨)에 의한 한외 여과를 비롯하나 이에 한정되는 것은 아니다. 정제는 변형된 제조자의 프로토콜에 따라 수행될 수 있다(예를 들어, PCT/EP2004/011715호를 참고).

[0194] 일부 양태에서, 처리된 DNA의 단편은 본 발명에 따른 프라이머 올리고뉴클레오티드의 세트(예를 들어, 표 3과 7을 참고)와 증폭 효소를 이용하여 증폭된다. 일부 DNA 단편의 증폭은 동시에 하나의 반응 용기 및 동일한 반응 용기에서 수행될 수 있다. 일반적으로, 증폭은 중합효소 연쇄반응(polymerase chain reaction, PCR)을 이용하여 수행된다. 증폭산물은 일반적으로 길이가 100 내지 2000 염기쌍이다.

[0195] 방법의 다른 양태에서, DMR을 포함하는 마커 내의 또는 마커 근처의 CpG 위치의 메틸화 상태 (예를 들어, DMR 1-96; 표 2와 6)은 메틸화 특이적 프라이머 올리고뉴클레오티드의 사용에 의해 검출될 수 있다. 이 기술(MSP)은 Herman의 미국 특허 제6,265,171호에 기재되어 있다. 바이설파이트 처리된 DNA의 증폭에 대한 메틸화 상태 특이적 프라이머의 사용은 메틸화된 및 메틸화되지 않은 핵산 사이의 구별을 허용한다. MSP 프라이머 쌍은 바이설파이트 처리된 CpG 디뉴클레오티드에 혼성화하는 적어도 하나의 프라이머를 함유한다. 그러므로, 상기 프라이머의 서열은 적어도 하나의 CpG 디뉴클레오티드를 포함한다. 비메틸화된 DNA에 특이적인 MSP 프라이머는 CpG 내의 C 위치의 위치에서 “T”를 포함한다.

[0196] 증폭에 의해 수득한 단편은 직접적으로 또는 간접적으로 검출가능한 표지를 지닌다. 일부 양태에서, 표지는 형광 표지, 방사성 핵종(radionuclide), 또는 질량 분석계에서 검출될 수 있는 일반적인 질량을 갖는 분리 가능한 분자 단편이다. 상기 표지들이 질량 표지(mass label)인 경우, 일부 양태는 단일 양성 또는 음성 순전하를 갖는 표지된 증폭산물을 제공하며, 이는 질량 분석계에서 더 좋은 검출도(delectability)를 허용한다. 검출은 예를 들어, 매트릭스 보조 레이저 탈착/이온화 (matrix assisted laser desorption/ionization, MALDI) 질량 분석계 또는 전기 분무 질량분석법 (electron spray mass spectrometry, ESI)을 이용하여 수행되거나 시각화 될 수 있다.

[0197] 이러한 검정 기술에 적합한 DNA를 분리하는 방법이 기술분야에 알려져 있다. 특히, 일부 양태는 미국 특허 제 13/470,251호 (“핵산의 분리(Isolation of Nucleic Acids)”)에 기재된 바와 같이 핵산의 분리를 포함한다.

[0198] 방법

[0199] 일부 양태에서, 다음의 단계를 포함하는 기술과 방법이 제공된다.

[0200] 1) 개체로부터 수득한 핵산(예를 들어, 대변 시료, 혈액 시료, 또는 조직 시료(예, 체장 조직)와 같은 체액으로부터 분리된, 예를 들면, 유전체 DNA)을 DMR (예, DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 및 92, 95 및 96 (표 2와 6으로부터))을 포함하는 적어도 하나의 마커 내의 메틸화된 또는 비메틸화된 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약들과 접촉시키는 단계 및

[0201] 2) 체장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC) (예를 들면, 80% 이상의 감도 및 80% 이상의 특이도와 제공됨)의 부족을 검출하는 단계.

[0202] 일부 양태에서, 다음의 단계를 포함하는 기술과 방법이 제공된다.

[0203] 1) 개체로부터 수득한 핵산(예를 들어, 대변 시료, 혈액 시료, 또는 조직 시료(예, 체장 조직)와 같은 체액으로부터 분리된, 예를 들면, 유전체 DNA)을 DMR (예, DMR 1-96 또는 DMR 1, 21, 24, 25, 26, 55, 70, 77, 81, 84, 및 92, 95 및 96 (표 2와 6으로부터))을 포함하는 적어도 하나의 마커 내의 메틸화된 또는 비메틸화된 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약들과 접촉시키는 단계 및

[0204] 2) 체장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)(예를 들면, 80% 이상의 감도 및 80% 이상의 특

이도와 제공됨)을 분류하는 단계.

- [0205] 바람직하게는, 감도는 약 70% 내지 약 100%, 약 80% 내지 약 90%, 또는 약 80% 내지 약 85%이다. 바람직하게는, 특이도는 약 70% 내지 약 100%, 약 80% 내지 약 90%, 또는 약 80% 내지 약 85%이다.
- [0206] 유전체 DNA는 시판중인 키트의 사용을 포함한 임의의 수단에 의해 분리될 수 있다. 간단하게, 관심 DNA가 세포 막에 의해 캡슐화된 생물학적 시료는 효소적, 화학적, 또는 기계적 수단에 의해 파괴 및 용해 되어야한다. DNA 용액에서 예를 들어, 단백질 분해효소 K를 이용한 절단에 의한 다른 오염물과 단백질이 이어서 제거될 수 있다. 유전체 DNA는 이어서 용액으로부터 회복될 수 있다. 이는 염색, 유기적 추출(organic extraction), 또는 고체 상 지지체에 DNA의 결합을 포함한 다양한 방법에 의해 수행될 수 있다. 방법의 선택은 시간, 비용, 및 요구되는 DNA의 양을 포함한 몇 개의 요인들에 의해 영향을 받을 것이다. 신생성 물질(neoplastic matter) 또는 전 신생성 물질(pre-neoplastic matter)을 포함하는 모든 임상적 시료 유형은 본 발명의 방법에서 사용하기에 적합하며, 그 예로는 세포주, 조직학적 슬라이드 (histological slide), 조직 검사(biopsy), 파란핀 포매 조직, 체액, 대변, 결장 유출액, 소변, 혈장, 혈청, 전혈, 분리된 혈구, 혈액으로부터 분리된 세포, 및 이들의 조합이 있다.
- [0207] 일부 양태에서, 시료는 체장 조직 및/또는 체액을 포함한다.
- [0208] 본 발명의 기술은 시료를 제조하고 검사를 위한 핵산을 제공하는데 사용된 방법에서 제한되지 않는다. 예를 들어, 일부 양태에서, DNA는 예를 들어, 관련 방법 또는 미국 특허 제61/485386호에 상술된 바와 같은 직접적 유전자 포획(direct gene capture)을 이용하여 대변 시료, 혈액, 또는 혈장 시료로부터 분리된다.
- [0209] 유전체 DNA 시료를 DMR(예를 들어, 표2 와 6에서 제공된 바와 같은 예를 들면, DMR 1-96)을 포함하는 적어도 하나의 마커 내에서 메틸화된 및 비메틸화된 CpG 디뉴클레오티드를 구별하는 적어도 하나의 시약 또는 일련의 시약으로 이어서 처리한다.
- [0210] 일부 양태에서, 시약은 5' 위치에서 메틸화되지 않은 시토신 염기를 혼성화 작용(hybridization behavior)의 관점에서 시토신과 유사하지 않은 우라실, 티민, 또는 다른 염기로 전환한다. 그러나, 일부 양태에서, 시약은 메틸화 민감성 제한 효소일 수 있다.
- [0211] 일부 양태에서, 유전체 DNA 시료는 5' 위치에서 메틸화되지 않은 시토신 염기가 혼성화 작용의 관점에서 시토신에 유사하지 않은 우라실, 티민, 또는 다른 염기로 전환되는 이러한 방식에서 처리된다. 일부 양태에서, 이 처리는 바이설파이트(하이드로젠 설파이트, 디설파이트)와 수행되며, 이어서 염기성 가수분해(alkaline hydrolysis)를 수행한다.
- [0212] 처리된 핵산을 이어서 분석하여 표적 유전자 서열(DMR- 예를 들어, 표 2와 6에서 제공되는 바와 같은 예를 들면, DMR 1-96로부터 선택된 적어도 하나의 DMR-을 포함하는 마커로부터 적어도 하나의 유전자, 유전체 서열, 또는 뉴클레오티드)의 메틸화 상태를 결정한다. 분석의 방법은 본 명세서에 기술된 QuARTS 및 MSP를 비롯한 본 명세서에 열거된 방법을 포함하는 기술분야에 알려진 방법으로부터 선택될 수 있다.
- [0213] 본 발명의 기술은 체장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)과 관련된 임의의 시료의 분석에 관련된다. 예를 들어, 일부 양태에서, 시료는 환자로부터 수득한 조직 및/또는 생물학적 유체를 포함한다. 일부 양태에서, 시료는 대장암 시료를 포함한다. 일부 양태에서, 시료는 분비물을 포함한다. 일부 양태에서, 시료는 혈액, 혈청, 혈장, 위 분비물, 체액, 위장관 조직검사 시료, 식도 조직검사 시료로부터 현미경 해부된 세포, 위장 내강(gastrointestinal lumen)안으로 벗겨진 식도 세포, 및/또는 대변으로부터 회복된 식도세포를 포함한다. 일부 양태에서, 개체는 인간이다. 이러한 시료는 상부 위장관, 하부 위장관으로부터 유래할 수 있거나, 상부 위장관 및 하부 위장관 모두로부터 세포, 조직, 및/또는 분비물을 포함할 수 있다. 시료는 간, 담관(bile duct), 체장, 위, 결장, 직장, 식도, 소장, 충수, 십이지장, 폴립(polyp), 담낭, 항문, 및/또는 복막(peritoneum)으로부터 세포, 분비물, 또는 조직을 포함할 수 있다. 일부 양태에서, 시료는 세포액, 복수(ascites), 소변, 배설물, 체장액, 내시경 시 수득한 유동액, 혈액, 점액, 또는 타액을 포함한다. 일부 양태에서, 시료는 대변 시료이다.
- [0214] 이러한 시료는 통상적인 기술자에게 분명할 것 같은 기술분야에 알려진 임의의 수의 수단에 의해 수득될 수 있다. 예를 들어, 혈액, 복수, 혈청, 또는 체액 시료가 바늘과 주사기를 이용하여 비경구적으로 획득될 수 있는 반면에, 예를 들어, 소변과 배설물 시료는 용이하게 획득할 수 있다. 세포가 없는 또는 실질적으로 세포가 없는 시료는 통상적인 기술자에게 알려진 다양한 기법에 시료를 적용하여 획득할 수 있으며, 원심분리 및 여과를 포함하나, 이에 한정되는 것은 아니다. 침습적 기법이 시료를 획득하기 위해 사용되지 않는 것이 일반적으로

바람직 하지 않음에도 불구하고, 조직 균질액, 조직 절편(tissue section), 및 조직검사 표본과 같은 시료를 획득하는 것이 여전히 바람직할 수 있다.

[0215] 본 발명의 기술의 일부 양태에서, 개체에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 진단하는 방법이 제공된다. 본 명세서에서 사용된 용어 “진단하기(diagnosing)” 및 “진단(diagnosis)”은 개체가 소정의 질병 또는 질환으로부터 고통받을 것인지 또는 미래에 소정의 질병 또는 질환을 발달시킬 수 있을 것인지에 대한 여부를 통상적 기술자가 추정 및 심지어 결정할 수 있는 방법을 의미한다. 통상의 기술자는 예를 들어, 바이오 마커 (예, 본 명세서에 개시된 DMR), 질환의 존재, 중증도, 또는 부재를 나타내는 메틸화 상태를 비롯한 하나 이상의 진단 지표에 근거하여 보통 진단을 내린다.

[0216] 진단을 따라서, 임상 암 예후(clinical cancer prognosis) (예, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)에 대한)는 암의 공격성(aggressiveness)과 암 재발의 가능성을 결정하고 가장 효과적인 치료법을 계획하는 것에 관련된다. 더욱 정확한 예후가 만들어 질 수 있거나 심지어 암을 발달시키는 잠재적 위험을 평가할 수 있을 경우, 적합한 치료법 및 일부 경우에서 환자를 위한 덜 심각한 치료법이 선택될 수 있다. 암 바이오 마커 평가 (예, 메틸화 상태 결정)는 암을 발달시킬 위험이 낮거나 및/또는 좋은 예후가 있는 치료법이 필요 없거나, 제한적 치료법이 필요할 개체를 암을 발달시키거나 암의 재발을 겪을 가능성이 높은 더욱 집중적인 치료로부터 이득을 얻을 수 있는 개체와 구별하는데 유용하다.

[0217] 이와 같이, 본 명세서에 사용된 바와 같은 “진단 내리기” 또는 “진단하기”는 암을 발달시키는 위험을 결정하거나 예후를 결정하는 것을 더 포함할 수 있고, 이는 임상적 결과를 예측하는 것(의학적 치료가 있거나 없는), 적합한 치료 선택을 선택하는 것(또는 치료가 효과적일 것인지에 대한 여부), 또는 현재의 치료를 모니터링 하거나 치료를 잠재적으로 바꾸는 것을 본 명세서에 개시된 진단 바이오마커(예, DMR)의 척도에 기초하여 제공할 수 있다. 더욱이, 본 명세서에 기재된 대상의 일부 양태에서, 시간에 따른 바이오마커의 다중결정(multiple determination)을 해서 진단 및/또는 예후를 용이하게 할 수 있다. 바이오마커에서 시간적 변화를 사용하여 임상 결과를 예측하고, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 진행을 모니터링 하며/하거나, 암에 대응하는 적절한 치료의 효능을 모니터링 할 수 있다. 이러한 양태에서, 예를 들어, 효율적인 치료 과정 동안, 시간에 따른 생물학적 시료에서 본 명세서에 기재된 하나 이상의 바이오마커(예, DMR) (및 모니터링 한다면, 잠재적으로 하나 이상의 추가적인 바이오마커(들))의 메틸화 상태의 변화를 보는 것을 기대할 수 있다.

[0218] 본 명세서에 개시된 대상은 개체에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 예방법(prophylaxis) 또는 치료를 개시하거나 지속할 것인지에 대한 여부를 결정하는 방법을 일부 양태에서 더 제공한다. 일부 양태에서, 방법은 개체로부터 시간에 따른 일련의 생물학적 시료를 제공하는 단계; 일련의 생물학적 시료를 분석하여 본 명세서에 개시된 하나 이상의 바이오마커의 메틸화 상태를 각각의 생물학적 시료에서 결정하는 단계; 및 하나 이상의 바이오마커의 메틸화 상태의 임의의 측정 가능한 변화를 각각의 생물학적 시료에서 비교하는 단계;를 포함한다. 시간에 따른 바이오마커의 메틸화 상태의 임의의 변화는 사용되어 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 발달시킬 위험을 예측하고, 임상 결과를 예측하며, 암의 예방법 또는 치료를 개시하거나 지속시킬지에 대한 여부와 현재의 치료법이 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 효과적으로 치료할 것인지에 대한 여부를 결정할 수 있다. 예를 들어, 제1 시점은 치료의 개시 전 선택될 수 있고, 제2 시점은 치료의 개시 후 어느 때에 선택될 수 있다. 메틸화 상태는 상이한 시점에서 취해진 각각의 시료에서 측정될 수 있고, 정성적 및/또는 정량적 차이점을 나타낼 수 있다. 상이한 시료로부터 바이오 마커 수준의 메틸화 상태의 변화는 개체에서 위험[예를 들어, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC) 발달의 위험]예후, 치료 효능을 결정하는 것, 및/또는 질환의 진단과 상관될 수 있다(correlated).

[0219] 바람직한 양태에서, 본 발명의 방법과 조성물은 초기에, 예를 들어, 질병의 증상이 나타나기 전에, 질병의 치료와 진단을 위한 것이다. 일부 양태에서, 본 발명의 방법과 조성물은 임상단계 (clinical stage)에서 질병의 치료 또는 진단을 위한 것이다.

[0220] 언급한 바와 같이, 일부 양태에서, 하나 이상의 진단 또는 예후 바이오마커의 다중 결정을 내릴 수 있고, 마커에서 시간적 변화가 진단 또는 예후의 결정에 사용될 수 있다. 예를 들어, 진단 마커는 초기 시기에 결정될 수 있고, 두번째 시기에 다시 결정될 수 있다. 이러한 양태에서, 초기 시기 내지 제2 시기에서 마커의 증가는 질환의 중증도 또는 특정한 유형의 진단, 또는 소정의 예후가 될 수 있다. 마찬가지로, 초기 시기 내지 두번째 시기에서 마커의 감소는 질환의 중증도 또는 특정 유형, 또는 소정의 예후를 나타낼 수 있다. 더욱이, 하나 이상의 마커의 변화도는 질병의 중증도 및 미래의 부작용에 관련될 수 있다. 통상적인 기술자는 어떤 양태에서,

비교 측정(comparative measurement)이 복수의 시점에서 동일한 바이오마커로 만들어 질 수 있는 반면에, 첫번째 시점에서 소정의 바이오마커, 두번째 시점에서 두번째 바이오마커를 또한 측정할 수 있고, 이러한 마커의 비교가 진단 정보를 제공할 수 있음을 이해할 것이다.

[0221] 본 명세서에 사용된 바와 같이, 구 “예후를 결정하는 것”은 통상적 기술자가 개체에서 질환의 결과 또는 과정을 예측할 수 있는 방법을 나타낸다. 용어 “예후”는 100%의 정확도로 질환의 결과 또는 과정을 예측할 수 있는 능력 또는 심지어 소정의 과정 또는 결과가 바이오마커(예, DMR)의 메틸화 상태에 기초하여 일어날 가능성이 예측대로 더 높거나 낮은 것을 나타내는 것은 아니다. 대신, 통상적인 기술자는 용어 “예후”가 어떤 과정 또는 결과가 일어날 증가된 확률을 나타내는 것임을 이해할 것이며, 이는 즉, 질환을 보이지 않는 개인들과 비교할 경우, 과정 또는 결과는 소정의 질환을 보이는 개체에서 일어날 가능성이 더 높다. 예를 들어, 질환을 보이지 않는 개인들에서(예, 하나 이상의 DMR의 정상적인 메틸화 상태를 갖는), 소정의 결과의 가능성은 매우 낮을 수 있다.

[0222] 일부 양태에서, 통계적 분석은 예후적 지표를 부정적인 결과에 대한 소인과 관련 짓는다. 예를 들어, 일부 양태에서, 질환을 갖지 않는 환자로부터 수득한 정상 대조 시료에서의 메틸화 상태와 상이한 메틸화 상태는 개체가 통계적 유의성(statistical significance)의 수준에 의해 결정된 대조시료에서 메틸화 상태와 더 유사한 수준이 있는 개체 보다 질환으로부터 고통받을 확률이 더 크다는 것을 시사할 수 있다. 또한, 기초선(baseline)(예를 들어, “정상”) 수준으로부터 메틸화 상태의 변화는 개체 예후를 반영할 수 있고, 메틸화 상태의 변화도는 부작용의 중증도와 관련될 수 있다. 통계적 유의성은 둘 이상의 집단을 비교하고 신뢰구간 및/또는 p값을 결정함으로써 보통 결정된다(Dowdy and Wearden, Statistics for Research, John Wiley & Sons, New York, 1983을 참고). 본 발명의 대상의 예시적인 신뢰구간은 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.9%, 및 99.99%이며, 반면 예시적인 p값은 0.1, 0.05, 0.025, 0.02, 0.01, 0.005, 0.001, 및 0.0001이다.

[0223] 다른 양태에서, 본 명세서에 개시된 예후적 또는 진단적 바이오마커(예, DMR) 메틸화 상태의 임계 변화도(threshold degree of change)가 밝혀질 수 있고, 생물학적 시료에서 바이오마커의 메틸화 상태의 변화도는 메틸화 상태의 임계 변화도와 간단히 비교될 수 있다. 본 명세서에 제공된 바이오마커에 대한 메틸화 상태의 바람직한 임계 변화는 약 5%, 약 10%, 약 15%, 약 20%, 약 25%, 약 30%, 약 50%, 약 75%, 약 100%, 및 약 150%이다. 다른 양태에서, 소정의 결과에 대한 관련된 소인에 직접적으로 관련된 예후적 또는 진단적 지표(바이오마커 또는 바이오마커의 조합)의 메틸화 상태에 의해 “노모그램”이 밝혀질 수 있다. 집단 평균이 아닌 개별적 시료 측정이 참조되기 때문에, 통상적인 기술자는 두 수치 값에 관련된 이러한 노모그램의 사용을 이 측정에서 불확실성(uncertainty)이 마커 농도(marker concentration)에서 불확실성과 동일한 것을 이해하는 것을 비롯하여 숙지하고 있다.

[0224] 일부 양태에서, 대조 시료는 생물학적 시료와 함께 분석되어서, 생물학적 시료로부터 수득한 결과를 대조 시료로부터 수득한 결과와 비교할 수 있다. 더욱이, 생물학적 시료에 대한 검정 결과가 비교될 수 있는 표준 곡선이 제공될 수 있음이 고려될 수 있다. 형광 표지가 사용되는 경우, 이러한 표준 곡선은 검정 유닛의 함수로서 바이오마커의 메틸화 상태를 나타내며, 그 예로는 형광 신호 감도가 있다. 복수의 공여자로부터 취한 시료를 사용하여, 표준곡선은 질환(예, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC))이 있는 공여자 또는 화생(metaplasia)이 있는 공여자로부터 취한 조직에서 하나 이상의 바이오마커의 “위험 상태” 수준 뿐만 아니라 정상 조직에서 하나 이상의 바이오마커의 메틸화 상태를 제어하기 위해 제공될 수 있다. 방법의 어떤 양태에서, 개체로부터 수득한 생물학적 시료에서 본 명세서에 제공된 하나 이상의 DMR의 비정상적인 메틸화 상태를 식별하면, 개체는 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 식별된다. 방법의 다른 양태에서, 개체로부터 수득한 생물학적 시료에서 이러한 하나 이상의 바이오마커의 비정상적 메틸화 상태의 검출은 개체를 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 식별한다.

[0225] 마커의 분석은 하나의 검사 시료내에서 추가의 마커로 별도로 또는 동시에 수행할 수 있다. 예를 들어, 몇몇 마커들은 복수의 시료들을 효율적으로 처리하고 더 높은 진단 및/또는 예후 정확도를 잠재적으로 제공하기 위해 하나의 시험으로 합쳐질 수 있다. 게다가, 통상적인 기술자는 동일한 개체로부터 복수 시료(예를 들면, 연속적인 시점에서)를 검사하는 것의 가치를 인식할 것이다. 이러한 일련의 시료의 시험은 시간에 따라 마커 메틸화 상태의 변화의 식별을 허용할 수 있다. 메틸화 상태 변화의 부재 뿐만 아니라 메틸화 상태 변화는 질병 상태에 대한 유용한 정보를 제공할 수 있고, 이는 사건(event)의 시작으로부터 대략적인 시간, 구조할 수 있는 조직의 존재 및 양, 약물 치료의 타당성, 다양한 치료의 효능, 및 미래 사건의 위험을 포함하는 개체의 결과의 식별을 포함하나, 이로 한정하는 것은 아니다.

- [0226] 바이오마커의 분석은 다양한 물리적 형식으로 수행될 수 있다. 예를 들면, 마이크로타이터 플레이트의 사용 또는 자동화는 사용되어 다수의 검사 시료의 처리를 용이하게 할 수 있다. 대안적으로, 단일 시료 형식을 개발하여 시기 적절한 즉각적인 치료 및 진단을 용이하게 할 수 있고, 그 예로는 보행가능한 운송(ambulatory transport)이나 응급실 설정이 있다.
- [0227] 일부 양태에서, 개체는 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖는 것으로 진단되고, 대조 메틸화 상태와 비교하는 경우, 시료에서 적어도 하나의 바이오마커의 메틸화 상태의 측정 가능한 차이가 있다. 반면, 생물학적 시료에서 메틸화 상태의 변화가 식별되지 않는 경우, 개체를 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖지 않는 것으로, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)에 대한 위험이 없는 것으로, 또는 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)의 낮은 위험을 갖는 것으로 식별할 수 있다. 이런 점에서, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC) 또는 이의 위험을 갖는 개체는 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC) 또는 이의 위험이 낮거나 실질적으로 없는 개체와 구별될 수 있다. 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 발달시킬 위험을 갖는 이러한 개체는 집중적인 및/또는 정기적인 스크리닝 일정에 배치할 수 있다.
- [0228] 위에서 언급한 바와 같이, 본 발명의 기술의 방법의 양태에 따라, 하나 이상의 바이오마커의 메틸화 상태 변화를 검출하는 것은 정성적 결정일 수 있거나 정량적 결정일 수 있다. 이와 같이, 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 갖거나 또는 발달시킬 위험에 있는 개체를 진단하는 단계는 특정 임계 측정을 만드는 것, 예를 들면, 생물학적 시료에서 하나 이상의 바이오마커의 메틸화 상태가 소정의 대조 메틸화 상태로부터 달라지는 것을 나타낸다. 방법의 일부 양태에서, 대조 메틸화 상태는 바이오마커의 임의의 검출 가능한 메틸화 상태이다. 대조 시료가 생물학적 시료와 동시에 검사되는 방법의 다른 양태에서, 소정의 메틸화 상태는 대조 시료에서 메틸화 상태이다. 방법의 다른 양태에서, 소정의 메틸화 상태는 표준 곡선에 기초하고/거나 표준 곡선에 의해서 확인된다. 방법의 다른 양태에서, 소정의 메틸화 상태는 구체적인 상태 또는 상태의 범위이다. 이와 같이, 소정의 메틸화 상태는 통상의 기술자에게 명백할 허용 가능한 한계 내에서 실시되는 방법 및 목적하는 특이도 등의 양태에 부분적으로 기초하여 선택될 수 있다.
- [0229] 또한 진단 방법에 대해서, 바람직한 개체는 척추동물 개체다. 바람직한 척추동물은 온혈동물이고, 바람직한 온혈 척추동물은 포유류이다. 바람직한 포유류는 가장 바람직하게는 인간이다. 본 명세서에 사용된 바와 같이, 용어 "개체"는 인간 및 동물 개체를 모두 포함한다. 따라서, 수의학적 치료 용도가 본 명세서에서 제공된다. 이와 같이, 본 발명의 기술은 인간과 같은 포유동물 뿐만 아니라 시베리아 호랑이와 같이 멸종위기에 처한 중요한 포유동물; 인간에 의한 소비를 위해 농장에서 사육되는 동물과 같은 경제적으로 중요한 포유동물; 및/또는 애완동물로서 또는 동물원에서 길러진 동물과 같은 인간에게 사회적으로 중요한 동물의 진단을 위해 제공한다. 이러한 동물의 예는 고양이 및 개와 같은 육식동물(carnivores); 피그(pig), 호그(hog) 및 야생 멧돼지(wild boar)를 포함하는 돼지(swine); 캐틀(cattle), 옥스(ox), 양, 기린, 사슴, 염소, 들소(bison) 및 낙타와 같은 유제류(ungulates) 및/또는 반추동물(ruminants); 및 말을 포함하나, 이로 제한하는 것은 아니다. 따라서, 또한 제공되는 것은 가축의 진단 및 치료이며, 길들인 돼지, 반추동물, 유제류, 말(경주마 포함) 등을 포함하나, 이에 제한되는 것은 아니다. 본 발명에 개시된 대상은 개체에서 췌장 고등급 이형성증(IPMN-HGD, PanIN-3, 또는 PDAC)을 진단하는 시스템을 더 포함한다. 시스템은 예를 들면, 생물학적 시료가 수집된 개체에서 이러한 질환 위험의 스크리닝에 사용되는 시판 키트로서 제공될 수 있다. 본 발명의 기술에 따라 제공된 예시적 시스템은 표 2와 6에서 제공된 DMR의 메틸화 상태를 평가하는 것을 포함한다.
- [0230] **실시예**
- [0231] [실시예 I]
- [0232] 췌장관내 유두상 점액 종양(Intraductal papillary mucinous neoplasm, IPMN) 및 췌장 상피내 종양(pancreatic intraepithelial neoplasia, PanIN)은 췌관 선암종(pancreatic ductal adenocarcinoma, PDAC)의 주요 관 전구체(ductal precursor)이다. 고등급 이형성증(high-grade dysplasia, HGD)과 초기 PDAC를 갖는 전구 병변이 외과적으로 치료되어야 한다는 여론이 있는 반면, 현재 영상화 도구와 위험 예측 모델은 불완전하며 이러한 병변을 검출하는데 보통 실패한다. 췌액, 췌장 낭액, 또는 대변과 같은 매질(media)에 적용된 구별 분자 마커(discriminant molecular marker)는 고등급 병변을 더욱 정확하게 검출할 필요가 있다.
- [0233] 실험을 수행하여 고등급 이형성증(IPMN-HGD, PanIN-3) 또는 침습성 암(PDAC)이 있는 사례군을 정상적인 췌장 또는 저등급 전구 병변(IPMN-저등급 이형성증 (IPMN-LGD), PanIN-1, 및 PanIN-2)이 있는 대조군으로부터 구별하는

메틸화된 DNA 마커 후보를 식별하고 검증하였다.

- [0234] 후보 마커의 식별은 감소된 표현 바이설파이트 시퀀싱(Reduced Representation Bisulfite Sequencing, RRBS)을 동결된 췌장 정상 및 신생 조직으로부터 DNA에 이용하여 비편향된 전체 메틸롬(whole methylome) 발견 접근법에 기초하였고; 후보 마커를 ROC 곡선 하 면적 (area under the ROC, curve AUC), 배수 변화(fold-change), 및 p 값에 기초하여 선택하였다. 검증(validation)은 메틸화 특이적 PCR을 이용하여 사례군 및 대조군으로부터 독립적인 현미경 해부된 조직과 선택된 마커의 블라인드 검정(blinded assay)에 기초하였고; 마커 수준은 췌장 상피(pancreatic epithelia)에 대한 마커인 메틸화된 LRRC4에 표준화하였다. 사례군 (n=53)은 IPMN-HGD (19), PanIN-3 (4), 및 PDAC (30)을 포함하였다. 대조군 (n=111)은 조직학적으로 정상적인 췌장 조직 (31), IPMN-LGD (36), PanIN-1 (34), 및 PanIN-2 (10)을 포함하였다. AUC를 로지스틱 회귀(logistic regression)에 의해 각 마커 별로 계산하였고, 마커 조합에 의한 구별을 분석하였다.
- [0235] RRBS 데이터로부터, 28개의 후보 마커를 MSP 검증(MSP validation)을 위해 선택하였다. 이들 중 3개는 Z-마커였고, 이는 모든 조직 시료에서 CpG 과메틸화(hypermethylation)를 보이는 영역을 정상, 형성 이상(dysplastic), 또는 암에 걸린 것인지에 대한 여부의 식별에 이용된 용어-그러나 순환하는 백혈구(leukocyte)와 같은 조혈 유래 DNA에서 메틸화를 거의 갖지 않는-이다 (예를 들어, 미국 특허 제14/966,938호를 참고). Z-마커는 2개의 중요한 기능-신생성 마커를 생물학적 시료(대개 일부 면역침투도(degree of immunologic infiltration)를 함유함)내 총 내피 신호(epithelial signal)에 정규화 하는 경우, 이들은 효과적인 분모(denominator)이고, 종양 세포 탈락(shedding), 자살, 또는 괴사가 정상 세포/DNA 또한 방출되도록 야기하는 환자군에서 이들은 그 자체로 종양 특이적 마커일 수 있다-을 한다.
- [0236] 23개의 마커가 발굴 시료(discovery sample)에서 0.8 이상의 MSP AUC 값과 0.01 이하의 p값을 가졌고, 방법에 설명된 바와 같이 3개의 Z-마커와 함께 확장된 독립적 검증 세트(expanded independent validation set)에 넣었다(실시예 III). 보편적인 메틸화 표준(Universal methylation standard)을 적절한 음성 대조군과 함께 검증에서 사용하였다. 결과는 % 메틸화(췌장 조직에 대한 마커인 LRRC4에 정규화 됨)로 표현하였다. 4개의 마커 [환자군에서 낮은 % 메틸화, 0.79 미만의 낮은 AUC, 또는 높은 백그라운드]는 제거되고, 19개의 유망한 후보(ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781)를 남겼다.
- [0237] 사례군과 대조군의 구별은 검증세트에서 밝혀졌고, 검증에서 0.79 이상의 AUC 값을 가진 마커가 표 1a에 열거된다 (19개의 유망한 후보 - ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781-에 대한 AUC 값을 제공). 표 1b는 0.85이상의 AUC 값을 가진 마커의 패널을 나타낸다. 표 1b는 0.85 이상의 AUC 값이 있는 마커의 패널을 나타낸다. 표 1b에 나타낸 패널은 85, 90, 95, 및 100%의 각 특이도에서 환자군의 89, 87, 77, 및 74%를 검출하였다.
- [0238] 도 1은 췌장 조직으로부터 검정된 상이한 마커의 분포 플롯(distribution plot)을 제공한다. 개별적 조직 카테고리 고리에 걸친 하나의 상위 마커(*TBX15*)의 분포 플롯이 대조군보다 환자군에서(IPMN-HGD, PanIN-3, 및 PDAC) 매우 높은 수준을 분명하게 도시한다(도 1). 상위 마커의 패널은 0.91의 AUC (95% CI: 85-96%)를 얻었다. 85, 90, 95, 및 100%의 각 특이도에서 환자군의 89, 87, 77, 및 74%를 검출하였다.
- [0239] 표 2는 고등급 전구체 (IPMN-HGD, PanIN-3) 또는 침습성 암 (PDAC)에 대한 고 구별 마커[highly discriminant marker, (ABCB1, AK055957, C13ORF18, CD1D, CLEC11A, DLX4, ELM01, EMX1, FER1L4, FRMD4A, GRIN2D, HOXA1, PRKCB, SP9, ST6GAL2, ST8SIA1, TBX15, VWC2, 및 ZNF781을 포함)]로 식별된 이러한 마커에 대한 DMR 시작 / 종결 위치, 유전자 주석(gene annotation), 및 염색체 번호를 포함하는 DMR 정보를 제공한다. 표 3은 표 2에 제공된 DMR에 대한 프라이머를 제공한다.

[0240] [표 1a]

마커	AUC
VWC2	0.89
TBX15	0.89
CLEC11A	0.89
PRKCB	0.89
EMX1	0.88
ELMO1	0.88
ST8SIA1	0.87
ABCB1	0.87
DLX4	0.87
SP9	0.86
FRMD4A	0.85
GRIN2D	0.85
FER1L4	0.84
CD1D	0.83
HOXA1	0.83
AK055957	0.82
C13orf18	0.82
ZNF781	0.81
ST6GAL2	0.79

[0241]

[0242] [표 1b]

[0243] 환자군 (IPMN-HGD, PanIN-3, PDAC)과 대조군 (정상, IPMN-LGD, PanIN-1, 2)사이의 고 판별(AUC>0.85)의 선택된 메틸화된 DNA 마커

마커	AUC	95% 신뢰구간	p
TBX15	0.89	0.83-0.95	<0.001
VWC2	0.89	0.83-0.95	<0.001
PRKCB	0.89	0.83-0.95	<0.001
CLEC11A	0.89	0.83-0.94	<0.001
EMX1	0.88	0.82-0.94	<0.001
ELMO1	0.88	0.81-0.94	<0.001
DLX2	0.87	0.81-0.92	<0.001
ABCB1	0.87	0.80-0.93	<0.001
ST8SIA1	0.87	0.80-0.93	<0.001
SP9	0.86	0.79-0.92	<0.001

[0244]

표 2

[0245] DMR에 대한 정보

DMR No.	유전자 주석 (Gene Annotation)	염색체 No.	DMR 시작-종결 위치
1	ABCB1	7	87229775-87229856
2	ACAN	15	89346282-89346382
3	ADAM23	2	207308771-207308844
4	AJAP1	1	4714768-4714799
5	AK055957	12	133484978-133485739
6	ANKS1B	12	99288410-99288435
7	ANKS1B	12	99288269-99288361
8	ARHGEF4	2	131797843-131797938
9	ASCL1	12	103352254-103352327
10	BEND4	4	42153551-42153626
11	BTBD11	12	107713354-107713454
12	C13orf18	13	46960767-46961669
13	C8orf42	8	495602-495648
14	CACNB2	10	18429175-18429297
15	CADM3	1	159141662-159141748
16	CADM3	1	159158449-159158481
17	CCNA1	13	37005929-37005955
18	CD1D	1	158150797-158151205
19	CELF2	10	11059508-11059637

20	CELF2	10	11059947-11060085
21	CLEC11A	19	51228217-51228732
22	COL23A1	5	178017202-178017291
23	DAB1	1	58715853-58715890
24	DLX4	17	48042426-48042820
25	ELMO1	7	37487755-37488477
26	EMX1	2	73147710-73147772
27	EYA4	6	133562302-133562380
28	FBLN1	22	45898783-45898862
29	FBLN1	22	45898874-45898984
30	FER1L4	20	34189488-34189566
31	FJX1	11	35640997-35641094
32	FJX1	11	35641325-35641485
33	FLI1	11	128564620-128564788
34	FLT1	13	29068359-29068403
35	FRMD4A	10	13933742-13933802
36	FRMD4A	10	13933653-13933799
37	GAD2	10	26505523-26505589
38	GALNT13	2	154728133-154728196
39	GDNF	5	37839926-37840050
40	GFPT2	5	179780663-179781008
41	GRIN2D	7	37487755-37488477
42	HMX3	10	124895177-124895244
43	HOXA1	7	27136145-27136425
44	ILDR2	1	166916866-166917004
45	ISLR2	15	74420532-74420572
46	KCNA1	12	5019401-5019496
47	KCNA1	12	5019741-5019848
48	KCNA1	12	5019499-5019633
49	KIAA1239	4	37246813-37246866
50	KIAA1239	4	37245707-37245735
51	KIF19	17	72322322-72322421
52	LOC100130155	8	65282172-65282334
53	LOC100192379	4	122686300-122686394
54	LOC642345	13	88324257-88324286
55	LRRC4	7	127671993-127672310
56	LY6H	8	144241567-144241620
57	MAX.chr12.133484966-133485681	12	133484966-133485681
58	MAX.chr2.137523242-137523296	2	137523242-137523296
59	MAX.chr2.468116-468180	2	468116-468180
60	MAX.chr3.69591901-69591966	3	69591901-69591966
61	MAX.chr4.8859999-8860038	4	8859999-8860038
62	MAX.chr8.72459953-72459992	8	72459953-72459992
63	NEFL	8	24813584-24813624
64	NEFL	8	24814127-24814163
65	NTRK3	15	88800287-88800464
66	OPLAH.9263	8	145106349-145106456
67	PDE3A	12	20522579-20522673
68	PDE4B	1	66258534-66258614
69	PHF21B	22	45404844-45404938
70	PRKCB	16	23847825-23848025
71	PROKR2	20	5296943-5297041
72	RASSF2	20	4803295-4803585
73	RBFOX3	17	77179670-77179722
74	RIMS2	8	104512722-104512817
75	SLC35F1	6	118228394-118228509
76	SOX11	2	5836458-5836545
77	SP9	2	175202025-175202158

78	SPSB4	3	140770490-140770550
79	ST6GAL2	2	107502965-107503050
80	ST6GALNAC5	1	77334012-77334125
81	ST8SIA1	12	22487528-22487620
82	ST8SIA6	10	17495971-17496068
83	SYT6	1	114695394-114695533
84	TBX15	1	119527066-119527655
85	TCF24	8	67874063-67874135
86	TCF4	18	53255304-53255565
87	TFPI2	7	93519404-93519768
88	TIAM1	21	32930771-32930935
89	TSHZ3	19	31839809-31840038
90	TSHZ3	19	31842197-31842291
91	VIPR2	7	158937370-158937481
92	VWC2	7	49813135-49814168
93	WBSCR17	7	70596576-70596675
94	ZNF781	19	38182950-38183127

표 3

표 2에 제공된 DMR에 대한 프라이머

DMR No.	유전자	정방향 MSP 프라이머 (5'-3')	역방향 MSP 프라이머 (5'-3')
1	ABCB1	GAT TTT GTT CGT CGT TAG TGC (SEQ ID NO: 1)	TCT CTA AAC CCG CGA ACG A (SEQ ID NO: 2)
5	AK055957	TCG GCG TAT TTT TCG TAG ACG C (SEQ ID NO: 3)	CGC AAT CTT AAA CGT ACG CTT CGA (SEQ ID NO: 4)
8	ARHGEF4	TGT TTT CGC GGT CGT TAT ATA TTA CGT CGT (SEQ ID NO: 5)	GAA CTA TCC CCG AAC TCC GAC TCG A (SEQ ID NO: 6)
12	C13orf18	TTT AGG GAA GTA AAG CGT CGT TTT C (SEQ ID NO: 7)	AAC GAC GTC TCG ATA CCT ACG A (SEQ ID NO: 8)
18	CD1D	GCG CGT AGC GGC GTT TC (SEQ ID NO: 43)	CCC ATA TCG CCC GAC GTA A (SEQ ID NO: 44)
21	CLEC11A	AGT TTG GCG TAG TCG GTA GAT C (SEQ ID NO: 9)	GCG CGC AAA TAC CGA ATA AAC G (SEQ ID NO: 10)
24	DLX4	GTA TTT TTA ATA TTT GGT GAG TGC G (SEQ ID NO: 11)	TAC TAA AAC GTA CGA TAA ACA TCG T (SEQ ID NO: 12)
25	ELMO1	TTA TAT TTT TCG TTT TTA GTA ATT TCG CGT TAG C (SEQ ID NO: 13)	GAA AAC CCG CCG AAA CAT TTC GA (SEQ ID NO: 14)
26	EMX1	CGG GTT TTA GCG ATG TTT ATT TTA GTT TCG T (SEQ ID NO: 15)	CCT TTT CGT TCG TAT AAA ATT TCG TT (SEQ ID NO: 16)
30	FER1L4	TGG AGT TTT TTT CGT CGT TGA CGC (SEQ ID NO: 17)	AAT ACC GCG TCG ACC AAA AAC G (SEQ ID NO: 18)
36	FRMD4A	TGTAAGCGGTTGGTATTTCGGAACG (SEQ ID NO: 19)	CAAAAACCTTCGCCGAACGAA (SEQ ID NO: 20)
41	GRIN2D	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)
43	HOXA1	TGG GTT ATC GGT TTT TTA AGT TCG G (SEQ ID NO: 23)	GAA TTC CTC CCA ACC AAC CCT CTA CG (SEQ ID NO: 24)
55	LRRC4	CGT GGT TGT CGT TTG TAA TAT CGC (SEQ ID NO: 25)	ATA ATC CCC GCT AAC TAA CGA CG (SEQ ID NO: 26)
66	OPLAH.9263	TGC GTA GGT GAT AGG GAG GGG TTA C (SEQ ID NO: 27)	ACA AAA CAC ATC CTA TTA ACG CGA A (SEQ ID NO: 28)
70	PRKCB	GTC GTT TTT GGG CGT TTT AGA GGC (SEQ ID NO: 29)	TAA TCC CAA ACG AAC CGC CG (SEQ ID NO: 30)
77	SP9	CGG GGG TAA AAA GGG TAG CGA TAG TGA TAC (SEQ ID NO: 31)	CGA AAA ATT CGA AAC GAA ACG TC (SEQ ID NO: 32)

79	ST6GAL2	GTTAGTTTTTCGGTCGGGATCGAGC (SEQ ID NO: 33)	CGTTCGTAACGTCGACGATTACCGTA (SEQ ID NO: 34)
81	ST8SIA1	GAC GTT TGT CGT CGG GTT CGT TC (SEQ ID NO: 35)	AAA AAC CCT CCG CTA CCA CTT CGC (SEQ ID NO: 36)
84	TBX15	CGC GTA GGT CGG ATT GAT TTT TCG T (SEQ ID NO: 37)	TTA ATT AAC GCC CGA ACA CCC GC (SEQ ID NO: 38)
92	VWC2	TTG GAG AGT TTT TCG AAT TTT TTC GC (SEQ ID NO: 39)	GAA AAC CAC CCT AAC GCC G (SEQ ID NO: 40)
94	ZNF781	TCG TTA GGA ATT AGT GGT TGG ATT TAT CGC (SEQ ID NO: 41)	CGT TCT CTT ATT TTC CGA ATA CGC GAA (SEQ ID NO: 42)

[실시예 II]

마커를 비정상적인 췌장 조직학의 정도(degrees of abnormal pancreatic histology)가 다른 8명의 환자로부터의 블라인드 낭액 시료(blinded cyst fluid sample)에서 검사하였다. MSP는 바이선티카이트 전환 낭액 DNA에서 우리 마커 중 7개를 사용하여 가동시켰다. 카피(copy)를 분모로 LRRC4를 사용한 % 메틸화에 정규화 하였다. 4개의 마커로부터 결과가 표 4에 도시된다.

이러한 마커 별로, 조직학적 중증도 (degree of histologic severity)와 메틸화 확실성(positivity) 사이에 상관관계(correlation)가 있다. 선암종 환자 시료 (P-1)는 모든 마커에 대해서 가장 높은 % 메틸화를 보였다. 환자 P-2, P-3, 및 P-6은 이러한 3개의 조직학의 저등급 악성 가능성이 있는 정렬(alignment) 전역에 걸쳐 0 이었다. MCN과 IPMN은 경도 내지 중간 등급 이형성증 (고등급 IPMN은 검사에 이용할 수 없음)에 가장 덜 민감한 PRKCB 및 가장 민감한 CD1D으로 상이하게 빛난다(light up).

표 4

	LRRC4 정규화				
	ST8SIA1	CD1D	VWC2	PRKCB	임상 조직학
P-1	23.74	112.15	84.67	65.70	낭포성 패턴(cystic pattern)이 있는 PDAC
P-2	0.00	0.00	0.00	0.00	인접한 저류 낭종(retention cyst)이 있는 팽대부 선종(ampullary adeno)
P-3	0.00	0.00	0.00	0.00	장액 낭포성 종양(serous cystadenoma)
P-4	3.49	7.06	0.86	0.04	경도(mild) 내지 중등도(moderate) 이형성증이 있는 MCN
P-5	0.00	7.52	0.00	0.00	저등급(low) 내지 중간 등급(intermediate) 이형성증이 있는 MCN
P-6	0.00	0.00	0.00	0.00	가유두상 종양(pseudopapillary tumor)
P-7	8.52	15.12	22.34	0.33	BD-IPMN 중간 등급 이형성증
P-8	0.06	1.71	0.42	0.09	국소적 중간 등급 이형성증과 함께 저등급 이형성증이 있는 BD-IPMN

[실시예 III]

이 실시예는 실시예 I과 II에 대한 재료, 방법, 및 결과를 제공한다.

발굴 조직 시료 (신선 동결(fresh frozen) 및 FFPE)를 메이요 클리닉 로체스터(Mayo Clinic Rochester)의 기관암 레지스트리(institutional cancer registries)로부터 선택하였고, 숙련된 병리학자가 검토하여 정확한 분류를 확인하였다. 이들은 췌장암, IPMN, PanIN (P1, P2, 및 P3), 정상적인 췌장, 및 정상적인 결장을 포함하였다. 임상군과 하위군(sub-group) 당 N = 10 내지 26. 18개의 정상적인 백혈구 대조군이 Mayo Biospecimens Linking Investigators와 Clinicians to GIH Cell Signalling Research Clinical Core에 의해 제공되었다. 유전체 DNA를 Qiagen DNA Mini Kit와 Micro Kit를 사용하여 제조하였다.

감소된 표현 바이선티카이트 시퀀싱(Reduced Representation Bisulfite Sequencing, RRBS): RRBS는 암세포에서 과메틸화되고 정상세포에서 저메틸화되는 것으로 알려진 영역인 프로모터와 CpG 섬에서 후생적 시그니처(epigenetic signature) 또는 메틸화 시그니처(methylation signature)에서 비편향된 보기(unbiased look)를 허용하는 절차이다. 우리가 한 많은 연구들은-암 유형과 필터링하는 방법 모두에서-최초이다. 우리가 마커 영

역을 식별하고 평가하는데 사용한 기준을 과정을 고유하게 만드는 임상적 메틸화 특이적 증폭 검정을 설계하는 최종 목표에 맞췄다. 4개의 주요 필터는 1) 소정의 CpG 밀도 기준과 베타 회귀 모델링(beta regression modeling)을 이용한 DMR 호출, 2) ROC 곡선 분석, 3) 배수 변화 (정상 조직 또는 정상 백혈구에 대한), 및 4) 영역에 대한 증명된 CpG 공메틸화(co-methylation) (환자군에서) 및 공메틸화의 부족 (대조군에서)이다.

[0255] 라이브러리 제조: 유전체 DNA (300 ng)를 CpG 함유 모티프(motif)를 인식하는 메틸화 특이적 제한효소인 MspI 10 Unit을 이용한 절단에 의해 단편으로 만들어서 시료 CpG 함량을 높이고 유전체의 불필요한 영역을 제거하였다. 절단된 단편을 5 Unit의 클레노브 절편(Klenow fragment) (3'-5' 엑소-)으로 말단 회복(end-repaired) 및 A-테일링하고, 바코드 서열을 함유하는(각 단편을 이의 시료 ID에 연결하기 위함) 메틸화된 TruSeq 어댑터 (Illumina, San Diego CA)에 밤새 라이게이션 하였다. 160-340bp 단편 (40-220 bp 인서트)의 크기 선택은 Agencourt AMPure XP SPRI 비드/버퍼 (Beckman Coulter, Brea CA)를 이용하여 수행하였다. 완충액 컷오프 (cutoffs)는 비드/버퍼의 0.7X - 1.1X 시료 부피였다. 최종 용출 부피는 22 uL (EB 버퍼 - Qiagen, Germantown MD) 이었고; qPCR을 사용하여 작은 시료 분획에서 라이게이션 효율과 단편 품질(fragment quality)을 측정하였다. 시료는 이어서 변형된 EpiTect 프로토콜 (Qiagen)을 이용하여 바이선허과이트 전환 (2회)을 겪었다. qPCR과 통상의 PCR(conventional PCR) (PfuTurbo Cx hotstart - Agilent, Santa Clara CA)에 이어서 전환된 시료 분획에서 Bioanalyzer 2100 (Agilent) 평가는 최종 라이브러리 증폭 전 최적의 PCR 사이클 횟수를 결정한다. 다음의 조건이 최종 PCR에 사용되었다: 1.) 각 50uL 반응물은 5uL의 10X 완충액, 1.25uL의 10 mM 각 데옥시리보뉴클레오타이드 트리포스페이트(deoxyribonucleotide triphosphate, dNTP), 5uL의 프라이머 혼합물 (~5uM), 15uL 템플레이트 (시료), 1uL의 PfuTurbo Cx hotstart, 및 22.75의 물을 함유한다; 온도와 시간은 각각 95° C-5분; 98° C-30초; 12-16 사이클의 98° C-10초, 65° C-30초, 72° C-30초, 72° C-5분 및 4° C 홀드였다. 시료는 무작위 기법(randomization scheme)에 기초하여 4-플렉스 라이브러리로 합쳐졌고 (등물), 최종 크기 확인(verification)에 대한 bioanalyzer와 phiX 표준 및 어댑터 특이적 프라이머를 이용하는 qPCR으로 검사하였다.

[0256] 시퀀싱 및 생명정보(Bioinformatics): 내부 검정 대조군(internal assay control) 전용의 추가의 레인(lane)이 있는 무작위 레인(lane) 배치에 따라서 시료를 플로우 셀(flow cell)에 로딩하였다. Illumina HiSeq 2000상에서 Mayo Clinic Medical Genome Facility에서 Next Generation Sequencing Core에 의해 시퀀싱을 수행하였다. 리드는 101 사이클 동안 단일 방향이었다. 각 플로우 셀 레인은 1억-1억 2천만 개의 리드를 생성하고, 이는 정렬된 서열에 대한 30-50 배수 시퀀싱 텀쓰[sequencing depth, (CpG당 리드 수)]의 중앙값 커버리지(median coverage)에 충분하다. 표준 Illumina 파이프라인 소프트웨어는 fastq 형식에서 시퀀싱된 리드 생성(read generation)과 염기를 호출했다. 감소된 표현 바이선허과이트 시퀀싱에 대한 주석 파이프라인과 스트림라인 분석(streamlined analysis)인 SAAP-RRBS를 서열 정렬과 메틸화 추출에 사용하였다. RRBS 발굴 시기(discovery phase)를 위해, 1차 관심 비교(comparison of interest)는 각각의 매핑된 CpG에서 메틸화 차이-채장 신생 환자군, 채장 대조군, 백혈구 대조군, 및 특정 적용을 위한 직장 대조군 사이의-였다. CpG 점은 0.6 보다 큰 예상된 CpG 비율에 대한 관측에 의해 생화학적으로 소정되었다. 그러나, 이 모델을 위해, CpG 분석의 타일 유닛(tiled unit) “차등 메틸화 영역(methylated region, DMR)” 을 염색체별로 CpG 부위 위치 사이의 차이에 기초하여 고안하였다. 단 하나의 CpG가 있는 점은 제외하였다. 질병군 당 총 커버리지의 텀쓰(depth of coverage)가 200 리드 이상 (평균 10 리드/개체)이고, %-메틸화의 분산이 0보다 큰 (비정보적인 CpG는 제외함) 경우에만, 개개의 CpG 부위를 차등적 분석(differential analysis)에 대해 고려하였다. 리드 텀쓰 기준(Read-depth criteria)은 목적하는 검정력(statistical power)에 기초하여 각 군의 시료 크기가 18 개체인 임의의 두 군 사이의 %-메틸화의 10% 차이를 검출하였다. 통계적 유의성(Statistical significance)을 DMR 당 메틸화 백분율의 로지스틱 회귀에 의해 리드 수(read count)에 기초하여 결정하였다. 개별적 개체에 걸친 변화하는 리드 텀쓰를 설명하기 위해, 산포 모수(dispersion parameter)를 피팅된 모델(fitted model)로부터 잔차(residual)의 피어슨 카이-제곱 통계량(Pearson Chi-square statistic)을 이용하여 추정하는 경우, 과대산포 로지스틱 회귀 모델(over-dispersed logistic regression model)을 사용하였다. 이들의 유의수준(significance level)에 따라 평가된 DMR은 조직과 백혈구 대조군을 합하여 %-메틸화가 1% 이하이나, 환자군에서는 10% 이상일 경우, 더 고려된다. 이는 94번 마커 (표 2)를 야기한다. 모든 것은 0.85보다 큰 AUC, 25보다 큰 신호 배수 변화(signal fold change) (환자군 v. 정상 채장 조직)과 5보다 큰 신호 배수 변화 (환자군 v. 저등급 및 중등도 이형성증 시료를 포함하는 모든 대조군), 및 0.001보다 작은 p값을 가졌다.

[0257] 메틸화 특이적 PCR (methylation specific PCR, MSP) 검증: 후보 마커(대개 상위 20 - 30의 부위당 RRBS DMR)을 조직에서 메틸화 특이적 PCR-처음에는 발굴 시료(discovery sample)에서 그 뒤에 큰 독립 시료 세트

(independent sample set)에서-에 의해 검증하였다.

- [0258] MSP 프라이머 설계: 28개의 상위 마커에 특이적인 프라이머를 설계하여 주문하였다 (IDT, Coralville IA). 설계는 Methprimer software (University of California, San Francisco CA) 또는 MSPPrimer (Johns Hopkins University, Baltimore, MD)를 이용하여 수행하였다. 검정을 보편적으로 메틸화된 및 메틸화되지 않은 유전체 DNA 대조군의 회색 상에서 SYBR Green으로 qPCR에 의해 검사하고 최적화하였다.
- [0259] 메틸화 특이적 PCR: 정량적 MSP 반응을 발굴 시료(discovery sample)에서 수행하여 성능 (performance)을 확인하였다. 전장 유전체 차세대 시퀀싱(genome wide next gen sequencing) 연구가 높은 오류 발견율(false discovery rate)을 겪을 수 있기 때문에, 후보 DMR을 있는 플랫폼 상에서 높은 분석 감도 및 특이도로 검사하는 것은 중대하다. 검증은 메틸화 특이적 PCR을 이용하여 사례군 및 대조군으로부터 독립적 현미경 해부된 조직과 선택된 마커의 블라인드 검정에 기초하였고; 마커 수준은 메틸화된 췌장 상피에 대한 마커 LRRC4에 표준화시켰다. 환자군 (n=53)은 IPMN-HGD (19), PanIN-3 (4), 및 PDAC (30)을 포함하였다. 대조군 (n=111)은 조직학적으로 정상적인 췌장 조직 (31), IPMN-LGD (36), PanIN-1 (34), 및 PanIN-2 (10)을 포함하였다. AUC를 마커 별로 로지스틱 회귀에 의해 계산하였고, 마커 조합에 의한 식별을 탐구하였다.
- [0260] 추가의 검증을 Boston의 Massachusetts General Hospital의 공동 연구자로부터 제공된 낭액 시료에서 수행하였다. 8개의 200uL 시료 (낭포성 패턴이 있는 췌장암 1개, 인접한 저류 낭종이 있는 췌대부 선종 1개, 장액 낭포성 종양 1개, 경도(mild) 내지 중등도(moderate) 이형성증이 있는 MCN 1개, 저등급(low) 내지 중간 등급(intermediate) 이형성증 이 있는 MCN 1개, 가유두상 종양 1개, 중간 등급 이형성증이 있는 BD-IPMN 1개, 및 국소적 중간 등급 이형성증 과 함께 저등급 이형성증이 있는 BD-IPMN 1개). DNA를 추출하였고, 바이설파이트 전환하였으며, 상위 10개의 마커 후보를 이용하여 MSP에 의해 작동시켰다 (블라인드). 정규화된 LRRC4 결과와 블라인드 되지 않은 임상정보를 동시에 교환하였다.
- [0261] [실시에 IV]
- [0262] 이 실시예는 신규한 DNA 메틸화 마커로 췌장 낭포에서 진행성 신생(advanced neoplasia)의 정확한 검출을 보여준다.
- [0263] 낭액(cyst fluid, CF) 분석이 있거나 없는 영상화에 의한 췌장 낭포에서 진행성 신생 (고등급 이형성증(high grade dysplasia, HGD) 또는 암)을 검출하기 위한 현재의 임상적 알고리즘은 진단 정확도가 부족하며, 이는 불필요한 수술을 자주 야기한다. 조직으로부터 DNA의 전체 메틸롬 시퀀싱에 의해, 실험은 저등급 이형성증 (low grade dysplasia, LGD)이 있거나 이형성증이 없는 낭포로부터 HGD 또는 암을 숨긴 낭포를 정확하게 나눌 수 있는 신규한 메틸화된 DNA 마커를 식별하였다. CF에 대한 연구에서, 이러한 마커의 패턴이 진행성 신생이 없는 낭포로부터 진행성 신생이 있는 낭포를 고도로 구별하는 것처럼 보이는 것을 발견하였다.
- [0264] 확장된 다기관(multicenter) 사례-대조 연구에서, 실험을 수행하여 (1) 낭포 환자군(HGD 또는 선암종)와 낭포 대조군(LGD 또는 이형성증 없음)에서 블라인드 CF 분석에 기초한 신규한 메틸화된 DNA 마커의 검정에 의한 췌장 낭포에서 전이성 신생의 검출 정확도를 평가하고 (2) CF에서 메틸화된 DNA 마커 분산을 CEA와 변이 KRAS의 이들을 비교하였다.
- [0265] 외과적으로 절제된 낭포로부터 보관 CF 시료(archival CF sample)를 연구하였다. CF를 내시경 초음파(endoscopic ultrasound) 또는 수술에서 낭포 천자(cyst puncture)에 의해 획득하였다. 0.2 ml CF로부터 추출된 DNA (QiaAmp Mini kit, Qiagen, Valencia CA)를 바이설파이트 전환하였고 (EZ-96 DNA Methylation kit, Zymo Research, Irvine CA), 선택된 메틸화된 DNA 마커 (BMP3, NDRG4, SP9, DLX4, ABCB1, CD1D, CLEC11A, EMX1, PRKCB, ST8SIA1, VWC2, TBX15, LRRC4, ELMO1)를 이어서 메틸화 특이적 PCR 또는 정량적 대립유전자 특이적 실시간 표적 및 신호 증폭 (quantitative allele-specific real-time target and signal amplification, QuARTs)에 의해 검정하였다. KRAS (7개의 변이)를 시판중인 방법 (MILLIPLEX® MAP Kit) (예를 들어, 미국 특허 제9,127,318호를 참고)에 의해 QuARTs와 CEA로 검정하였다.
- [0266] 134개의 췌장 낭포 [41개의 췌장관내 유두상 점액 종양(IPMN), 35개의 점액 낭포성 신생물(Mucinous Cystic Neoplasm, MCN), 23개의 장액 낭포성 종양, 13개의 낭포성 침습 선암종(cystic invasive adenocarcinoma) (3개의 IPMN, 1개의 MCN, 9개의 indeterminate precursor), 및 22개의 염증성 또는 다른 낭포]로부터, 실험을 21개의 환자군 (8개의 HGD, 13개의 선암종)와 113개의 대조(45개의 이형성증 없음, 68개의 LGD)로 분류하였다. 환자군에 대한 나이의 중앙값 (IQR)은 71세 (56-77)이고 대조군에 대한 나이의 중앙값은 61세 (46-69) ($p<0.01$); 환자군의 61%와 대조군의 31%는 남성이었다 ($p=0.034$). 상위 3개의 메틸화된 마커 (TBX15, BMP3, CLEC11A)의

CF 분포는 환자군과 대조군 사이에서 보이는 높은 구별을 도시한다 (도 2를 참고). 상위 4개의 메틸화 마커는 개별적으로 0.90보다 큰 ROC 곡선 하의 영역 (area under the ROC curve, AUC)을 달성하였다(표 5를 참고). 변이 *KRAS* 또는 CEA에 의한 것 보다 이러한 메틸화된 마커 각각에 의한 사례 검출이 실질적이고 현저하게 더 나았으며, 이는 설정된 특이도에서 민감도 또는 AUC에 기초한다(표 5를 참고). 매칭된 특이도에서, 상위 개별적인 메틸화된 마커는 변이 *KRAS* 또는 CEA가 한 것의 2배로 환자군을 검출하였다. 표 6은 BMP3와 NDRG4에 대한 DMR 정보를 제공한다. 표 7은 BMP3과 NDRG4에 대한 프라이머 정보를 제공한다.

표 5

[0267] 췌장 낭액 검사에서 환자군 (HGD/암)를 대조 (LGD, 이형성증 없음)로부터 구별하기 위한 진단 정확도: 메틸화된 DNA 마커를 변이 *KRAS* 및 CEA와 비교하였다.

낭액 마커	감도, (%)		AUC (95% CI)	P 값 vs <i>KRAS</i>
	85% 특이도	95% 특이도		
메틸화된 DNA	81	76	0.92 (0.85-0.99)	0.002
<i>TBX15</i>	81	67	0.92 (0.85-0.98)	0.002
<i>BMP3</i>	86	67	0.91 (0.85-0.97)	0.002
<i>CLEC11A</i>	86	57	0.90 (0.84-0.97)	0.004
<i>ST8SIA1</i>				
CEA	43	29	0.72 (0.60-0.84)	0.906
변이 <i>KRAS</i>	48	33	0.71 (0.57-0.85)	---

표 6

DMR No.	유전자 주석	유전체 No.	DMR 시작-종결 위치
95	BMP3	4	81952348-81952402
96	NDRG4	16	58497382-58497492

표 7

[0269] 표 6에 제공된 DMR에 대한 프라이머.

DMR No.	유전자	정방향 MSP 프라이머 (5'-3')	역방향 MSP 프라이머 (5'-3')	탐침
95	BMP3	GTTTAATTTTCGGTTTCGTCGTC (SEQ ID NO: 45)	CGCTACGAAACACTCCGA (SEQ ID NO: 46)	CGCCGAGGCGGTTTTTTCG/3C6/ (SEQ ID NO: 49)
96	NDRG4	CGGTTTTTCGTTTCGTTTTTCG (SEQ ID NO: 47)	CCGCCTTCTACGCGACTA (SEQ ID NO: 48)	CCACGGACGGTTTCGTTTATCG/3C6/ (SEQ ID NO: 50)

[0270] [실시예 V]

[0271] 이 실시예는 낭액에서 신규한 DNA 메틸화 마커의 검정에 의한 췌장 낭포에서 진행성 신생의 검출을 보여준다.

[0272] 췌장 낭포의 중요한 부분집합은 진행성 신생 (고등급 이형성증 (high-grade dysplasia, HGD) 또는 암)을 숨기고 있다. 그러나, 이 고위험군을 검출하는 현재의 진단 접근법은 정확도가 부족한 영상화 기준 또는 낭액 분석에 기초한다. 췌장 낭포가 있는 환자의 관리는 진행성 신생에 대한 정보를 주는 분자 마커에 의해 더 최적화될 수 있다. 실험은 HGD 또는 선암종을 저등급 이형성증 (low-grade dysplasia, LGD)과 이형성증이 없는 조직(non-dysplastic tissue)으로부터 구별하는 췌장 조직에서 신규한 메틸화된 DNA 마커(methylated DNA marker, MDM)의 패널을 최근 식별하고 검증하였다 (예를 들면, Majumder S, et al., Gastroenterology 2016; S120-S121을 참고). 낭액에 적용될 경우, 조직에서 이 구별이 유지될지는 아직 결정되지 않았다. 이와 같이, 실험은 진행성 신생을 함유하는 낭포가 있는 환자군과 저등급 이형성증 (low-grade dysplasia, LGD)을 함유하거나 이형성증이 없는 낭포가 있는 대조군으로부터 낭액에서 후보 MDM의 구별을 평가하고 비교하는 목표로 수행되었다.

[0273] 이 블라인드 다기관 연구에서, 보관 췌장 낭액(archival pancreatic cyst fluid)을 83개의 외과적으로 절제된 낭포로부터 식별하였다. 0.2 ml의 낭액으로부터의 DNA를 QiaAmp Mini 키트 (Qiagen, Valencia CA)를 이용하여

추출하였고 EZ-96 DNA 메틸화 키트 (Zymo Research, Irvine CA)를 이용하여 밤새 바이설파이트 전환시켰다. 14개의 선택된 MDM(BMP3, NDRG4, SP9, DLX4, ABCB1, CD1D, CLEC11A, EMX1, PRKCB, ST8SIA1, VWC2, TBX15, LRRC4, ELM01)의 검정을 이어서 메틸화 특이적 PCR 또는 정량적 대립유전자 특이적 실시간 PCR 표적 및 신호 증폭(methylation specific PCR or quantitative allele-specific real-time target and signal amplification, QuARTs)에 의해 수행하였고; 수준을 베타 액틴 (총 DNA)과 나이에 정규화시켰다. 환자군 (n=14)은 선암종 또는 HGD가 있는 낭포였고, 대조군은 (n=61) LGD가 있거나 이형성증이 없는 낭포였다. 중간 등급 이형성증 (intermediate grade dysplasia)이 있는 낭포 (n=8)를 별도로 분석하였다.

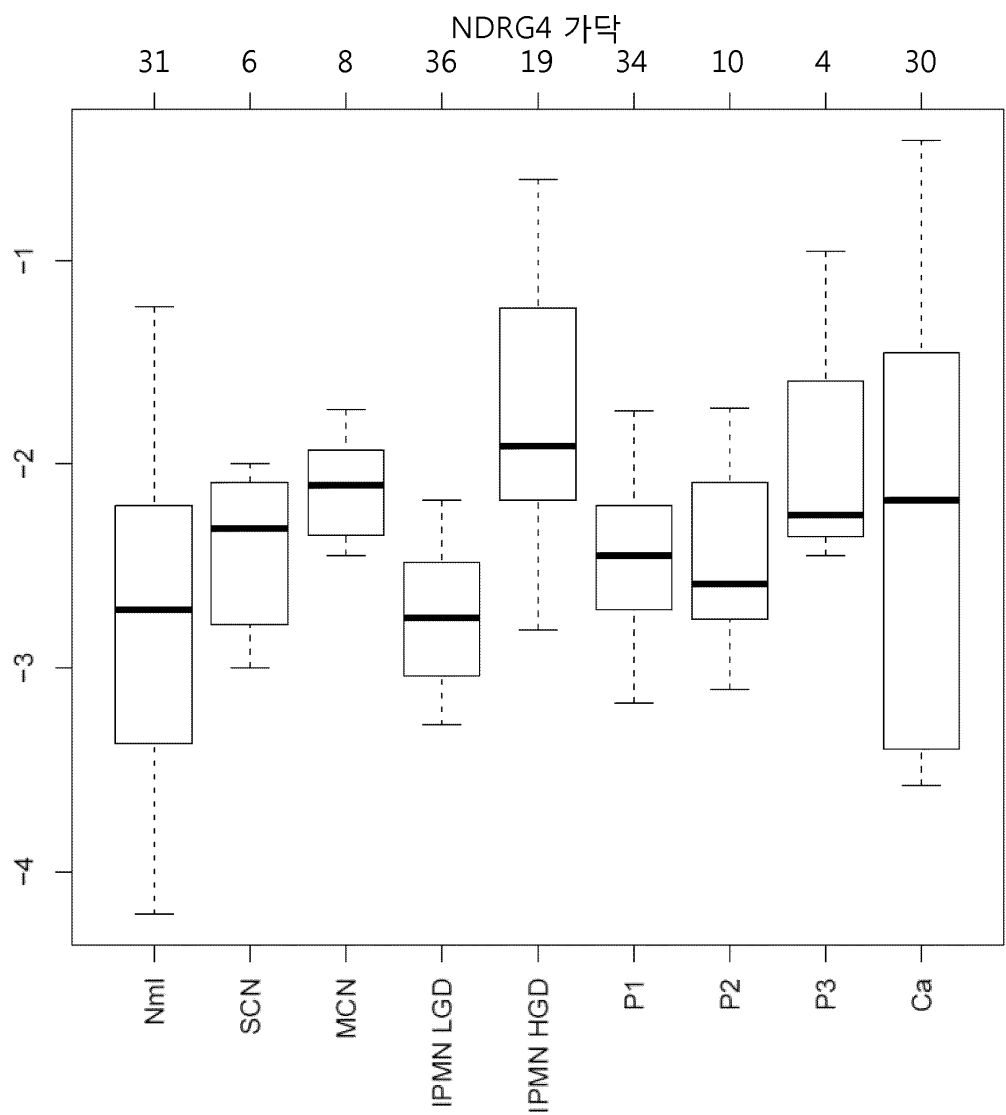
[0274] 상위 5개의 MDM (BMP3, EMX1, CLEC11A, ST8IA1, VWC2)은 0.90 이상의 ROC 곡선 하의 영역을 개별적으로 달성하여 환자군을 대조군으로부터 구별하였다. 상위 2개의 MDM (BMP3, EMX1)의 낭액 수준 구별은 높은 구별을 도시한다 (도 3을 참고). 90%의 특이도에서 (95% CI 80-96%), 이 2개의 MDM 패널은 환자군의 93% (66-100 %)를 검출하였다. 중간 등급 이형성증이 있는 낭포에서 MDM 수준은 일반적으로 HGD와 LGD의 사이에 떨어지고(도 3을 참고); 2개의 MDM 패널은 95% 특이도에서 양성인 이러한 병변의 38%를 호출했을 것이다.

[0275] 따라서, 이러한 실험은 췌장 낭액으로부터 신규한 MDM의 검정이 LGD가 있거나 이형성증이 없는 대조군으로부터 암 또는 HGD가 있는 환자군을 정확히 구분하는 것을 보여준다.

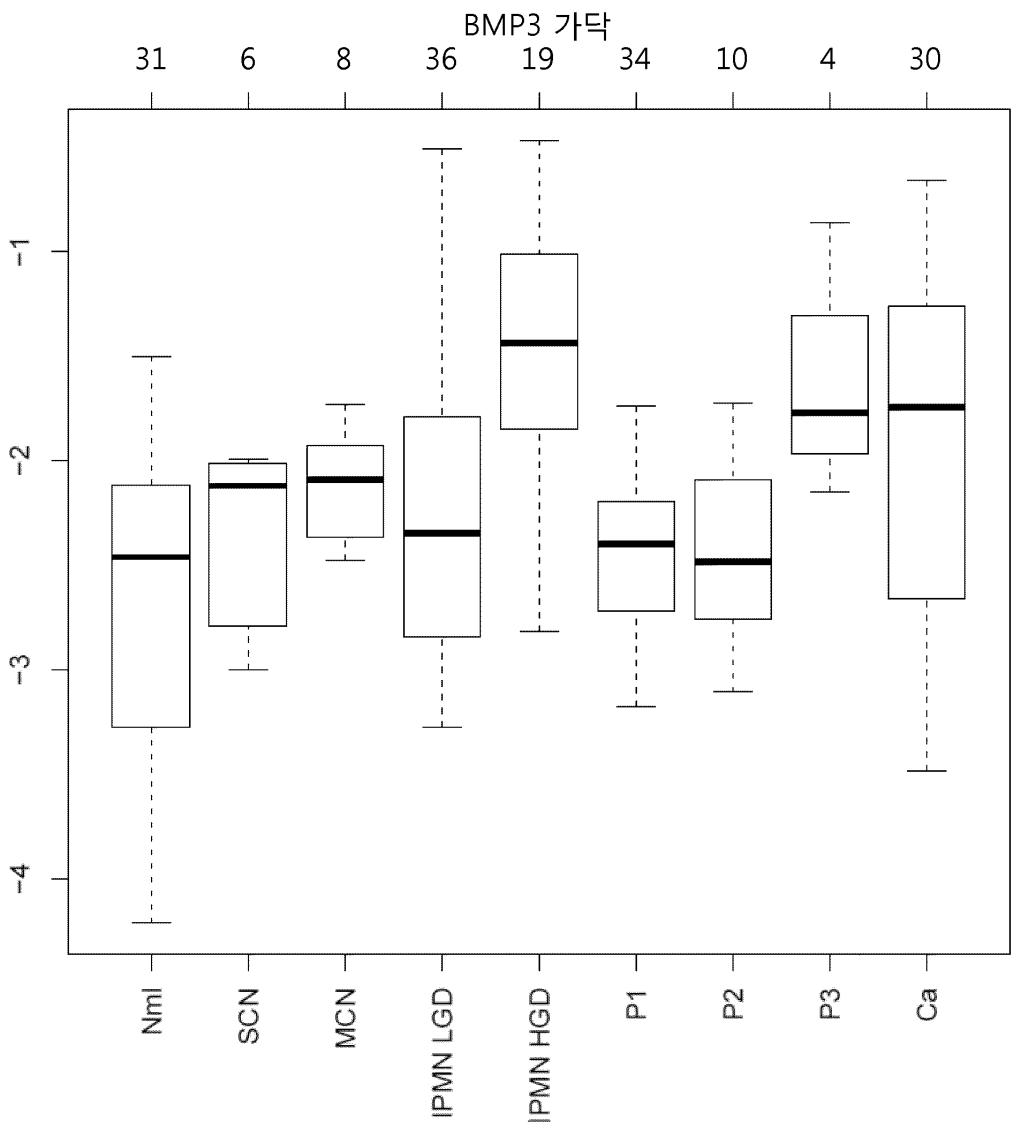
[0276] 상기 명세서에 언급된 모든 간행물 및 특허는 모든 목적을 위해 이들 전체가 참고문헌에 의해 본 명세서에 포함된다. 본 발명의 기술의 기재된 조성물, 방법 및 용도의 다양한 변경 및 변형이 기재된 바와 같은 기술의 범위 및 사상을 벗어나지 않고 통상의 기술자에게 명백할 것이다. 본 발명의 기술이 특정한 예시적 양태와 관련하여 기재되었음에도, 청구된 본 발명이 이러한 특정 양태에 과도하게 제한되어서는 안 된다는 것을 이해해야 한다. 실제로, 약리학, 생화학, 의학 또는 관련 분야의 통상적 기술자에게 자명한 본 발명을 수행하기 위한 기재된 방식의 다양한 변경은 하기 청구 범위의 범주 내에 있을 것으로 의도된다.

도면

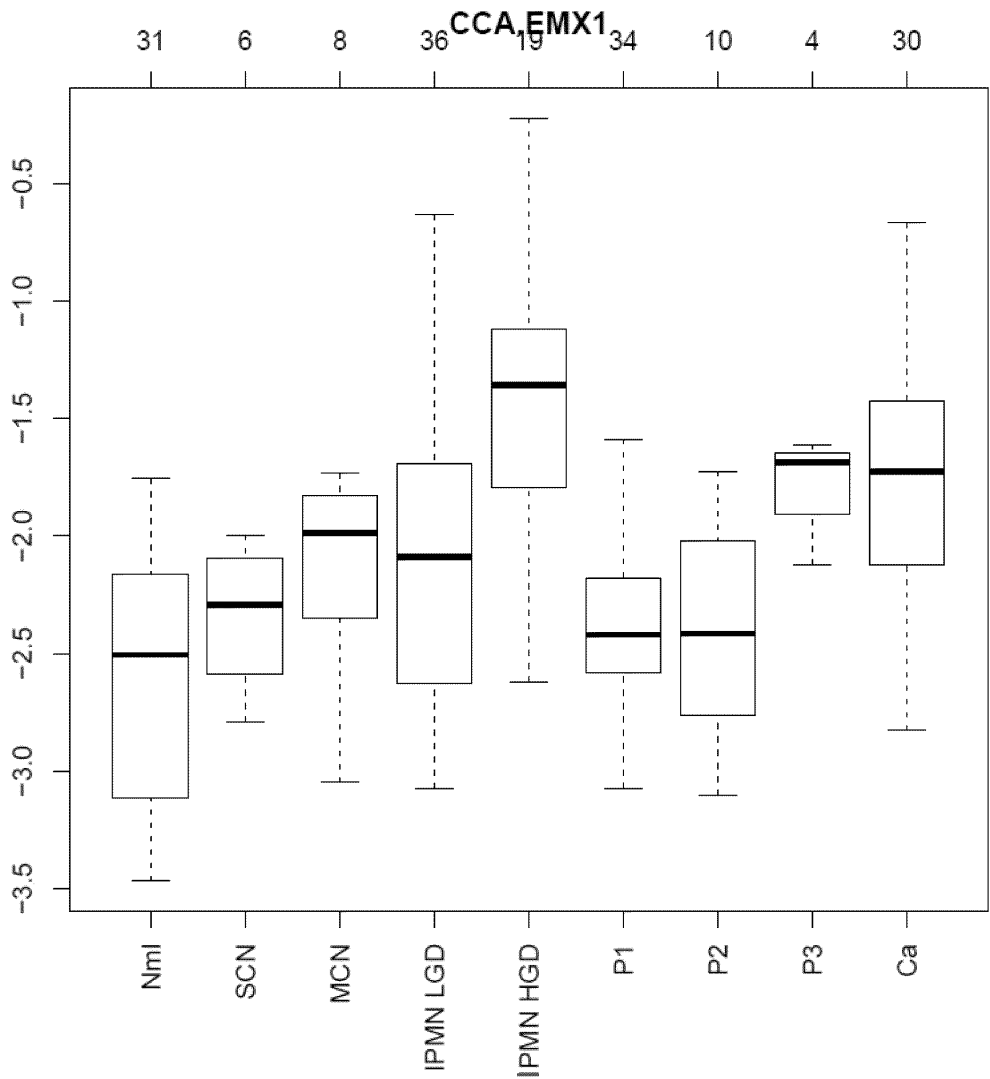
도면1a



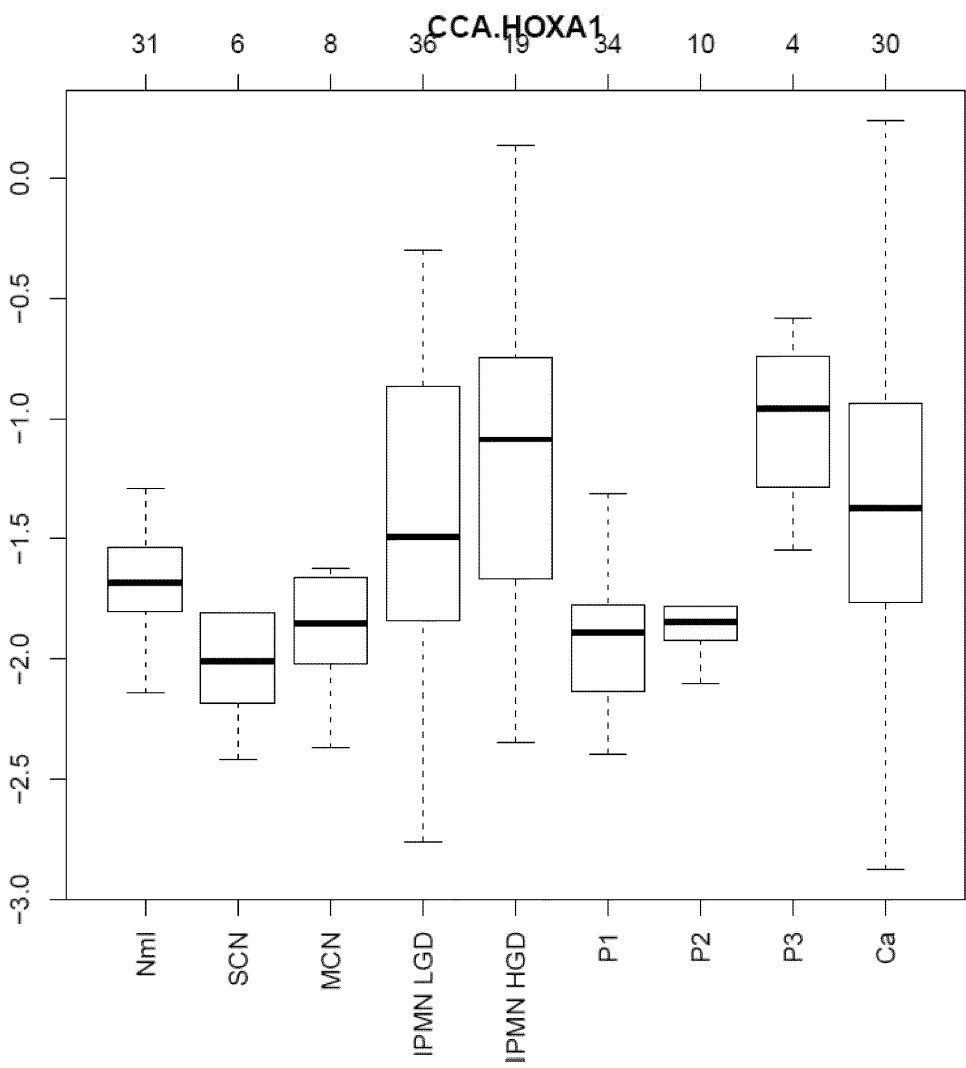
도면1b



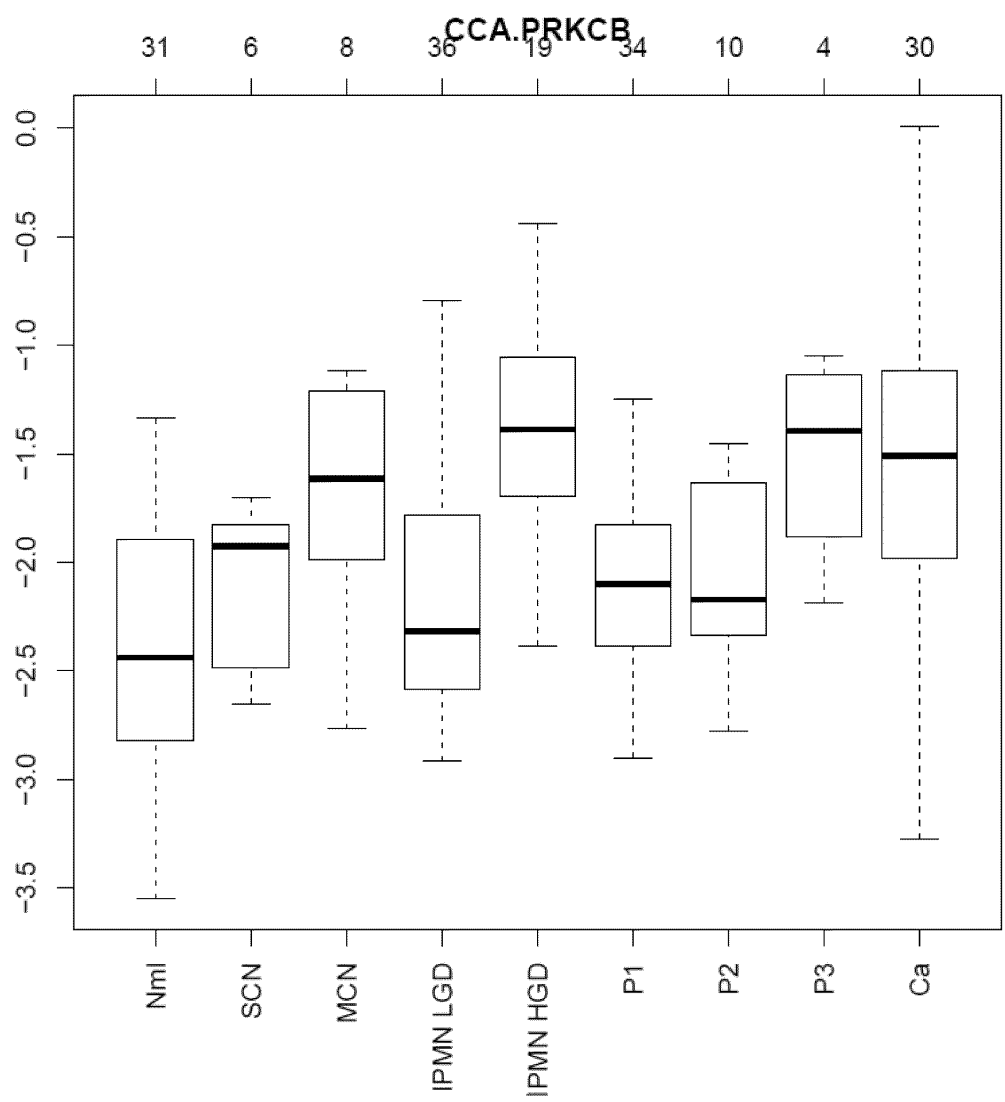
도면1c



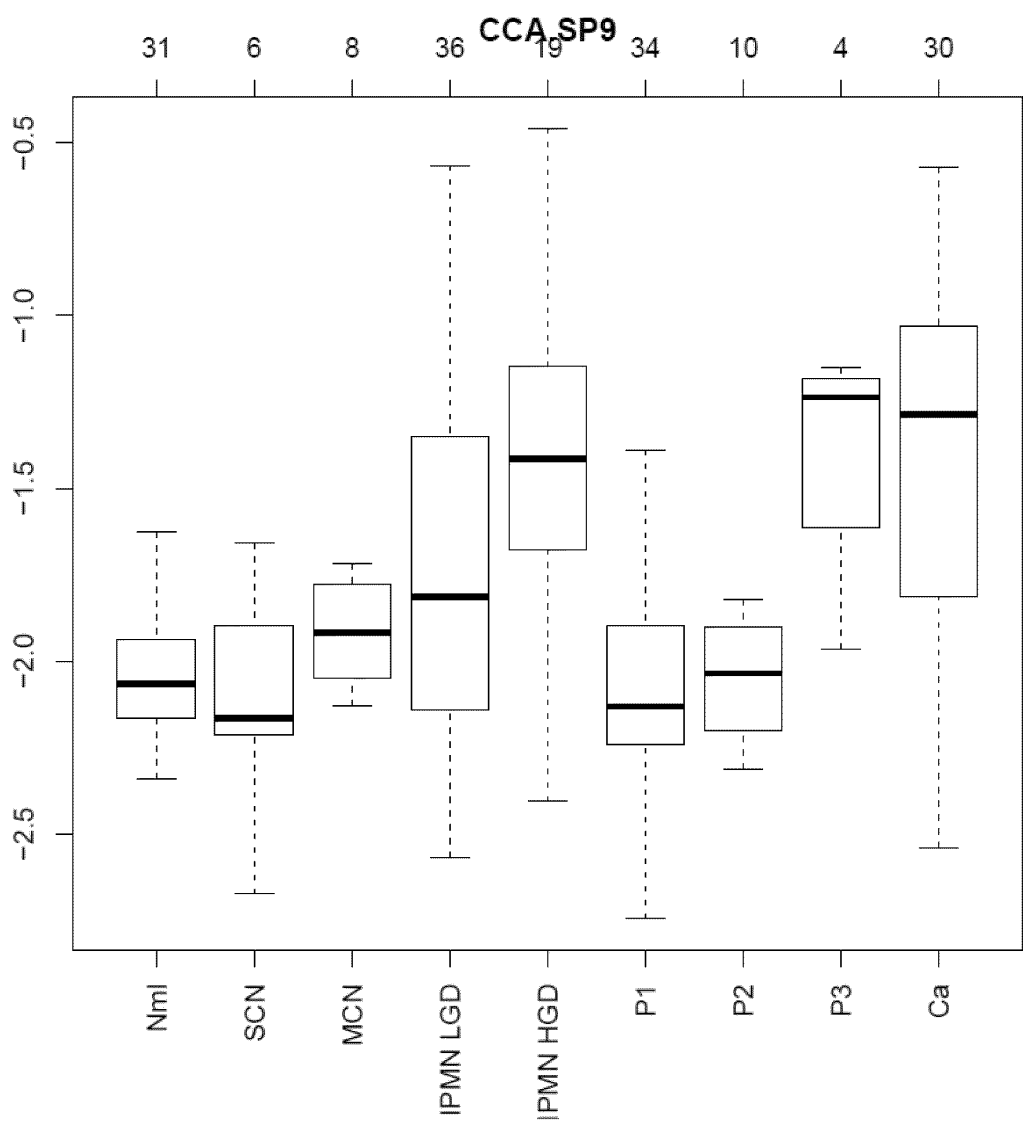
도면1d



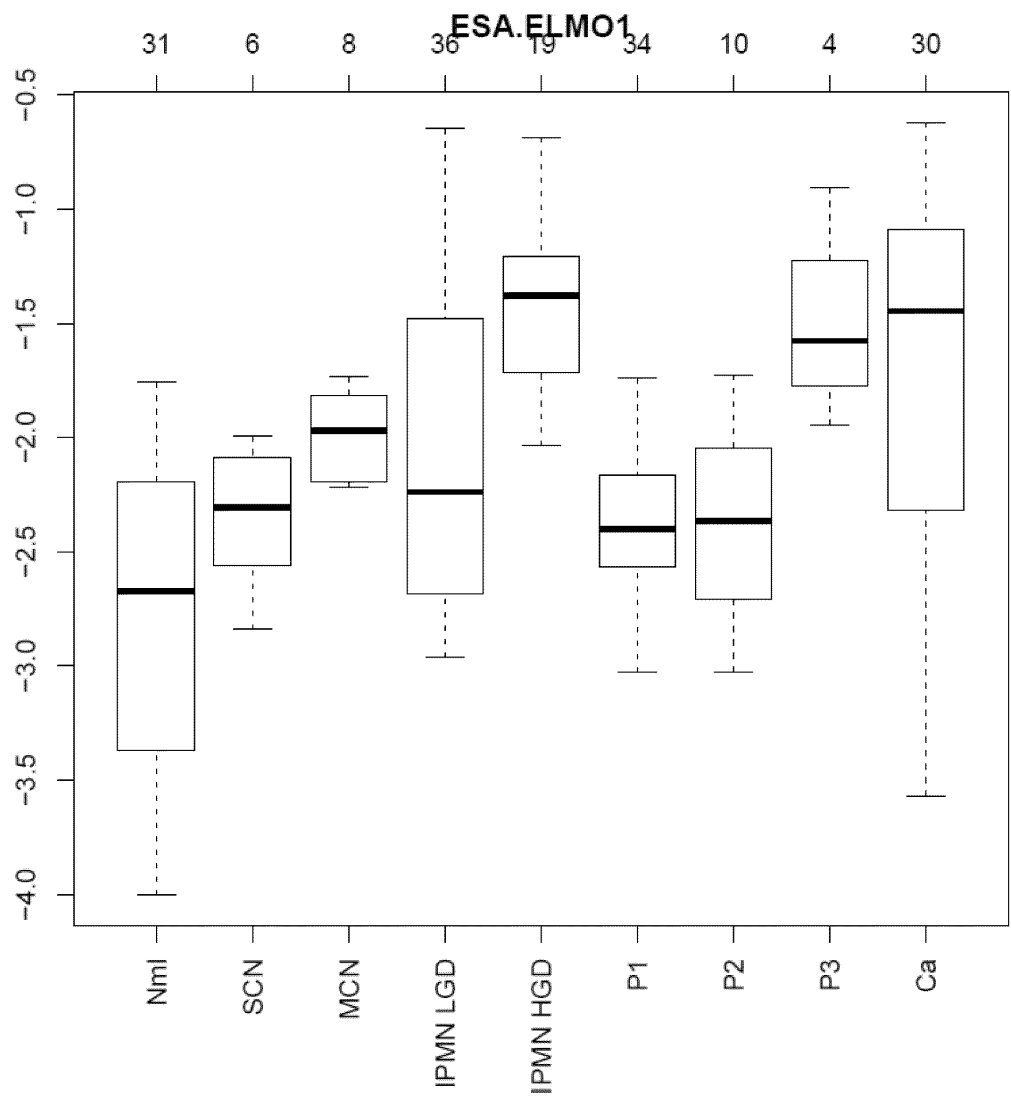
도면1e



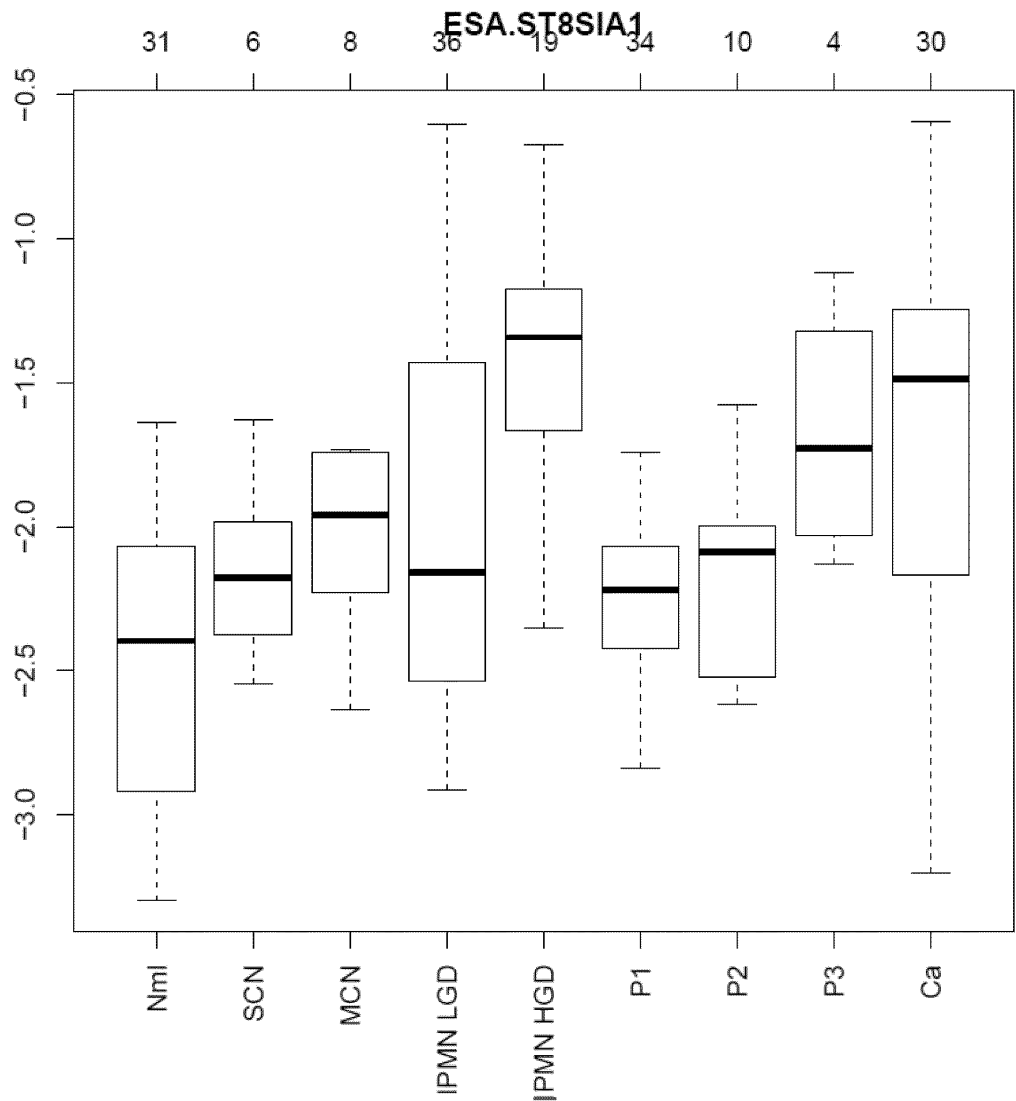
도면1f



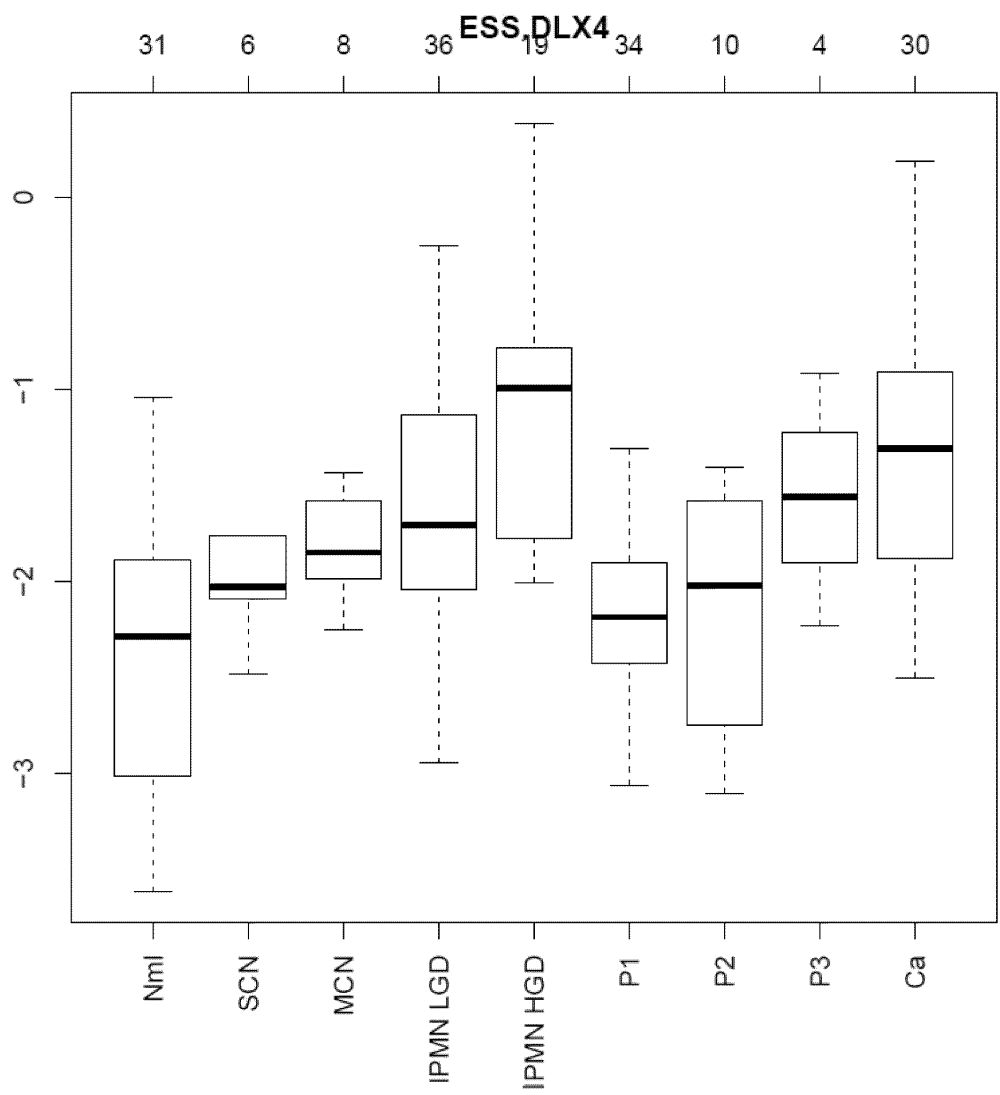
도면1g



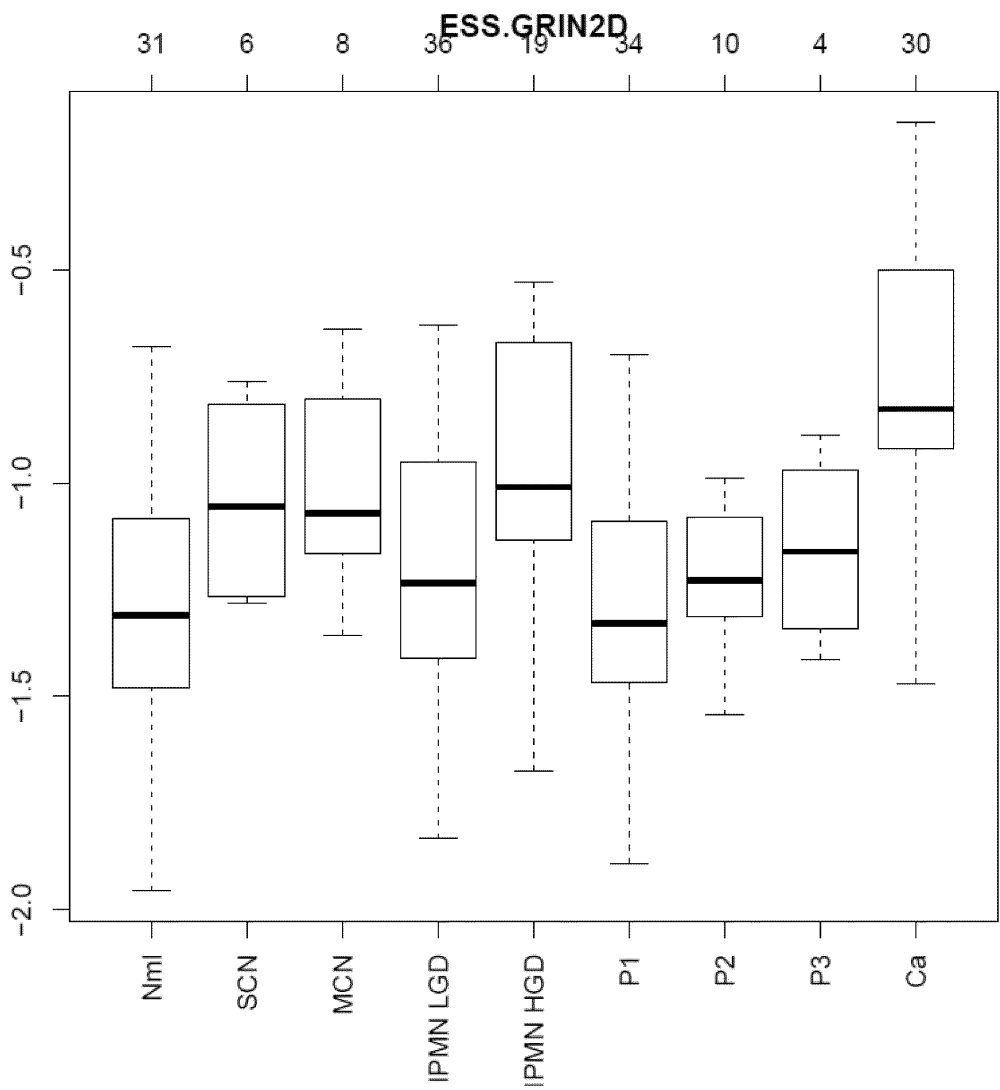
도면 1h



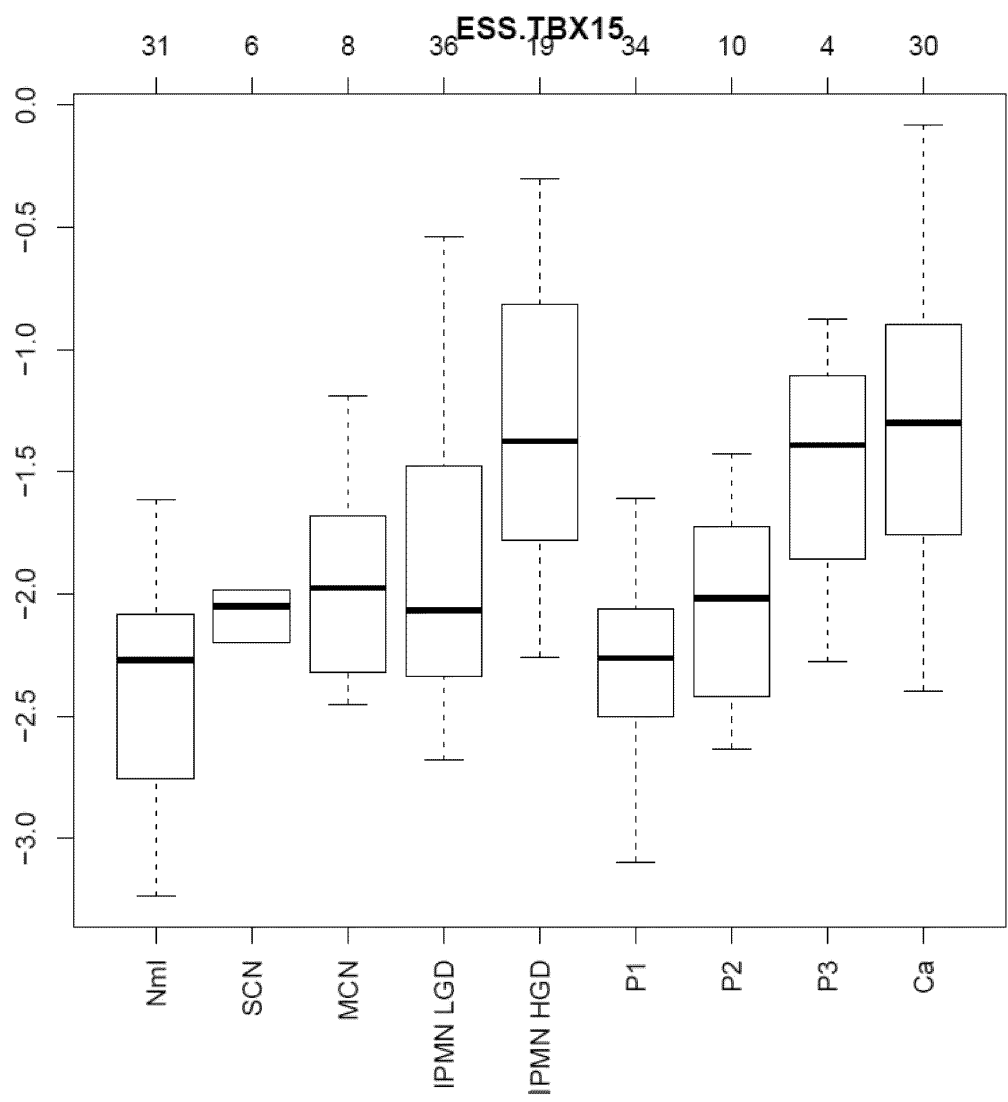
도면1i



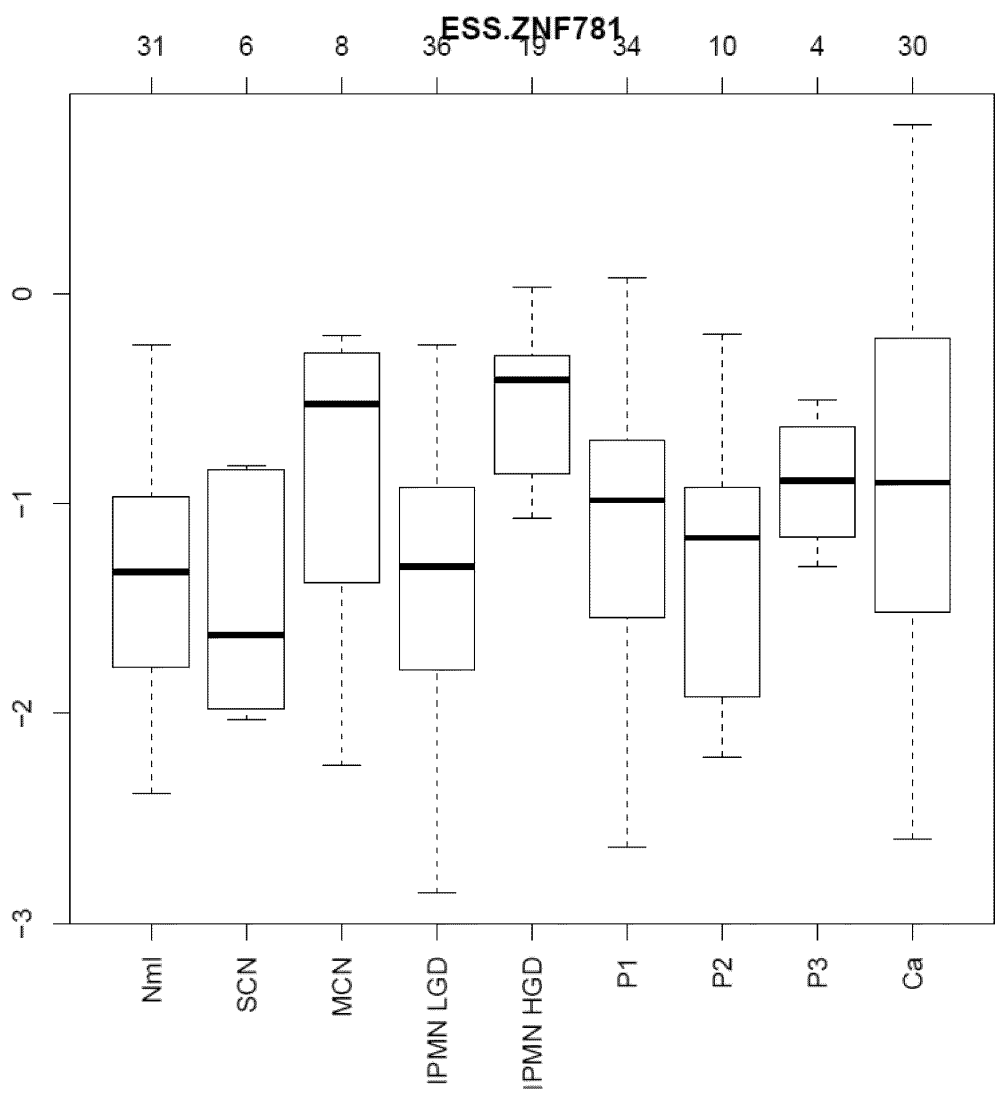
도면1j



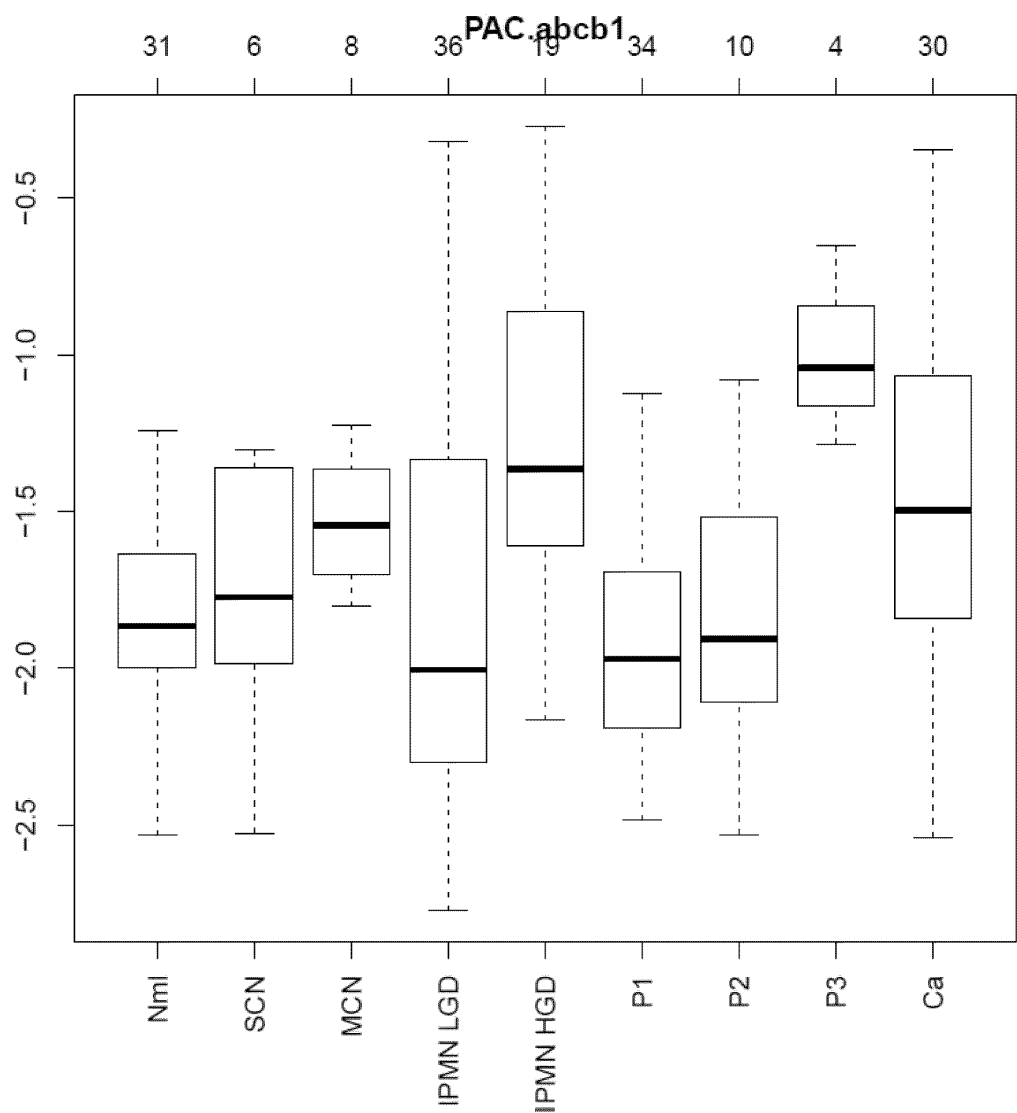
도면1k



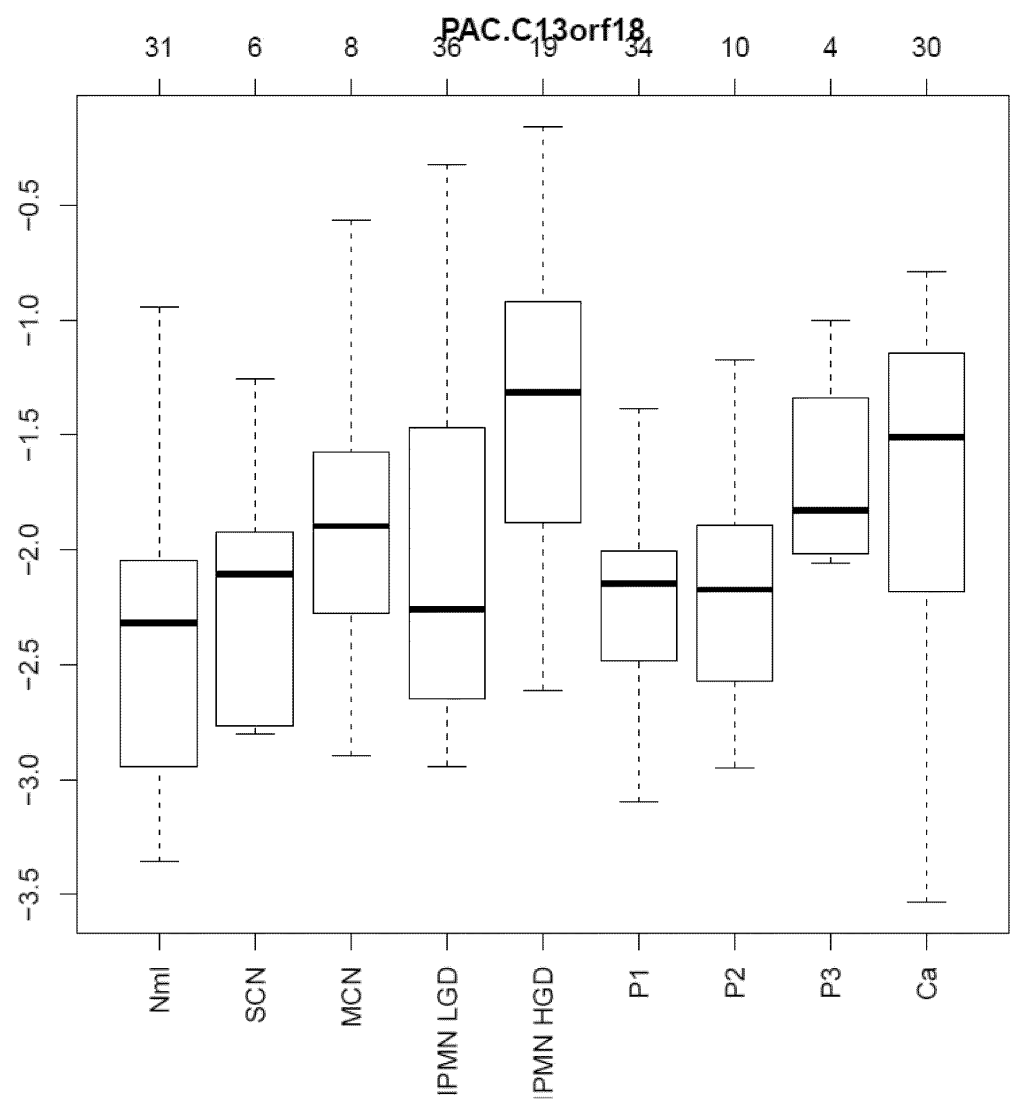
도면11



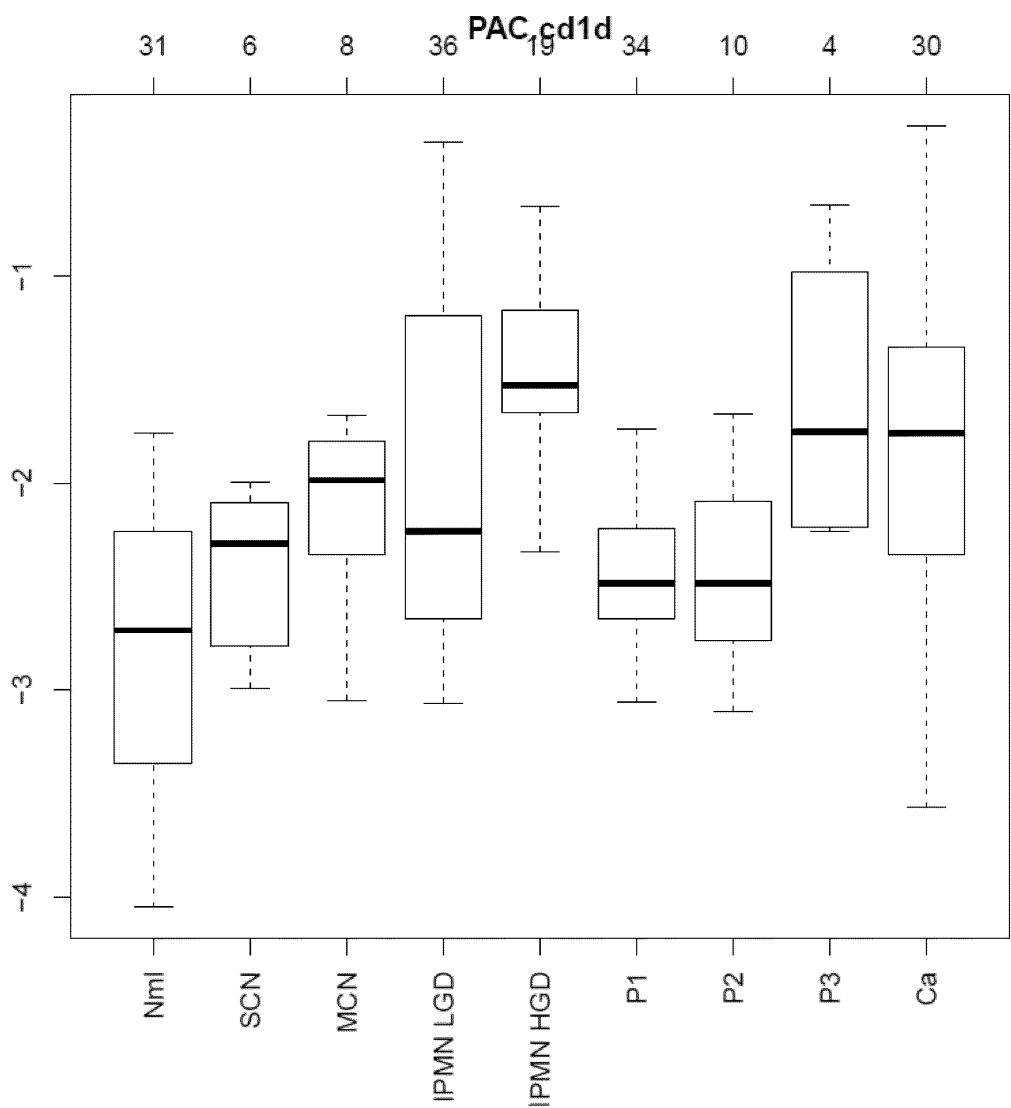
도면 1m



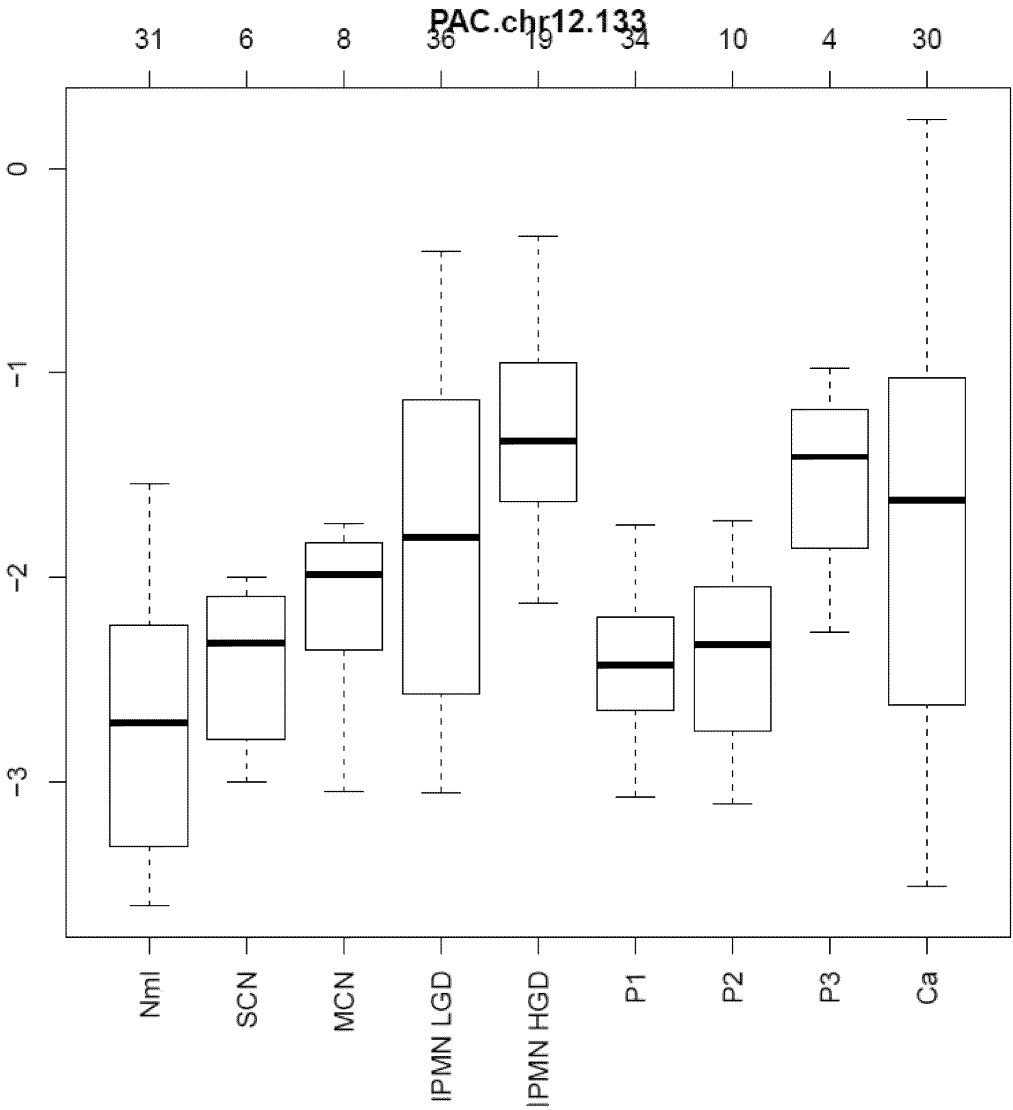
도면1n



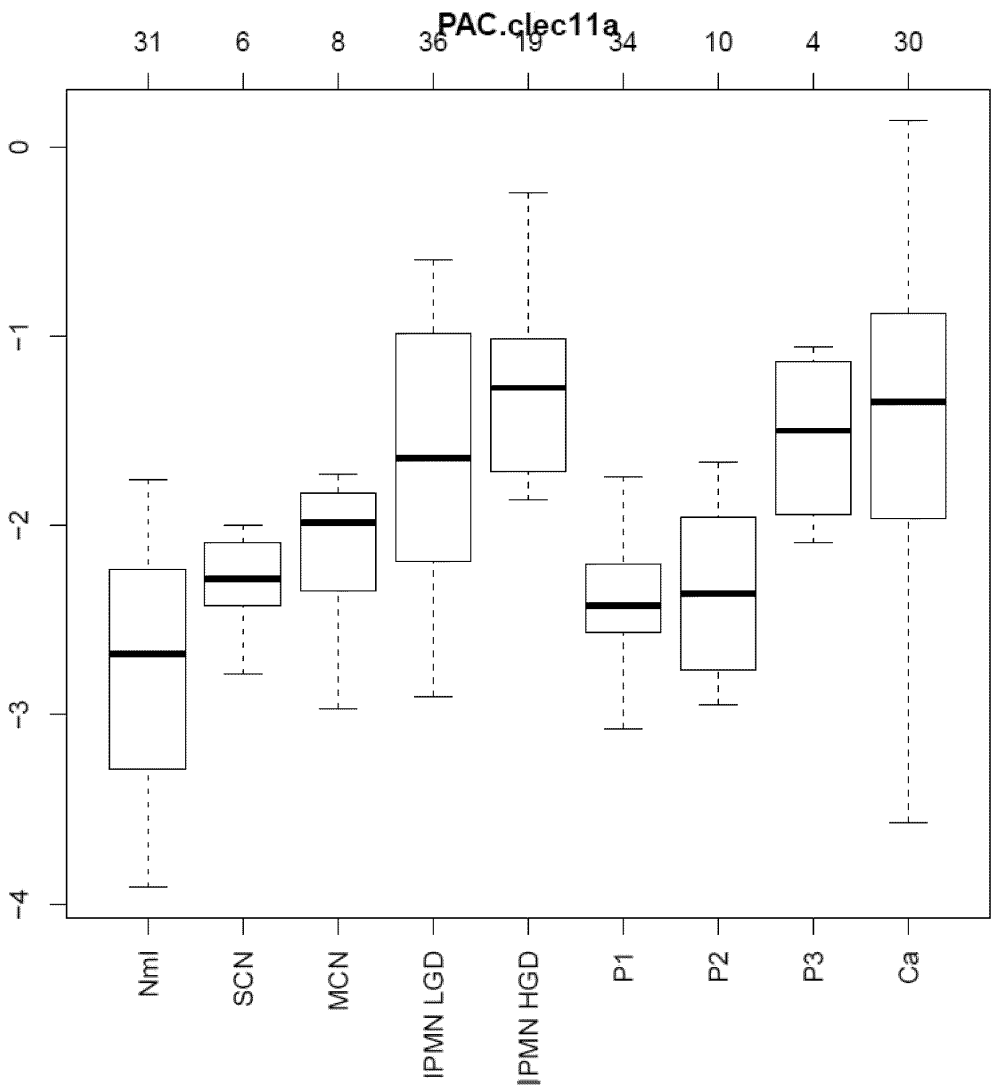
도면10



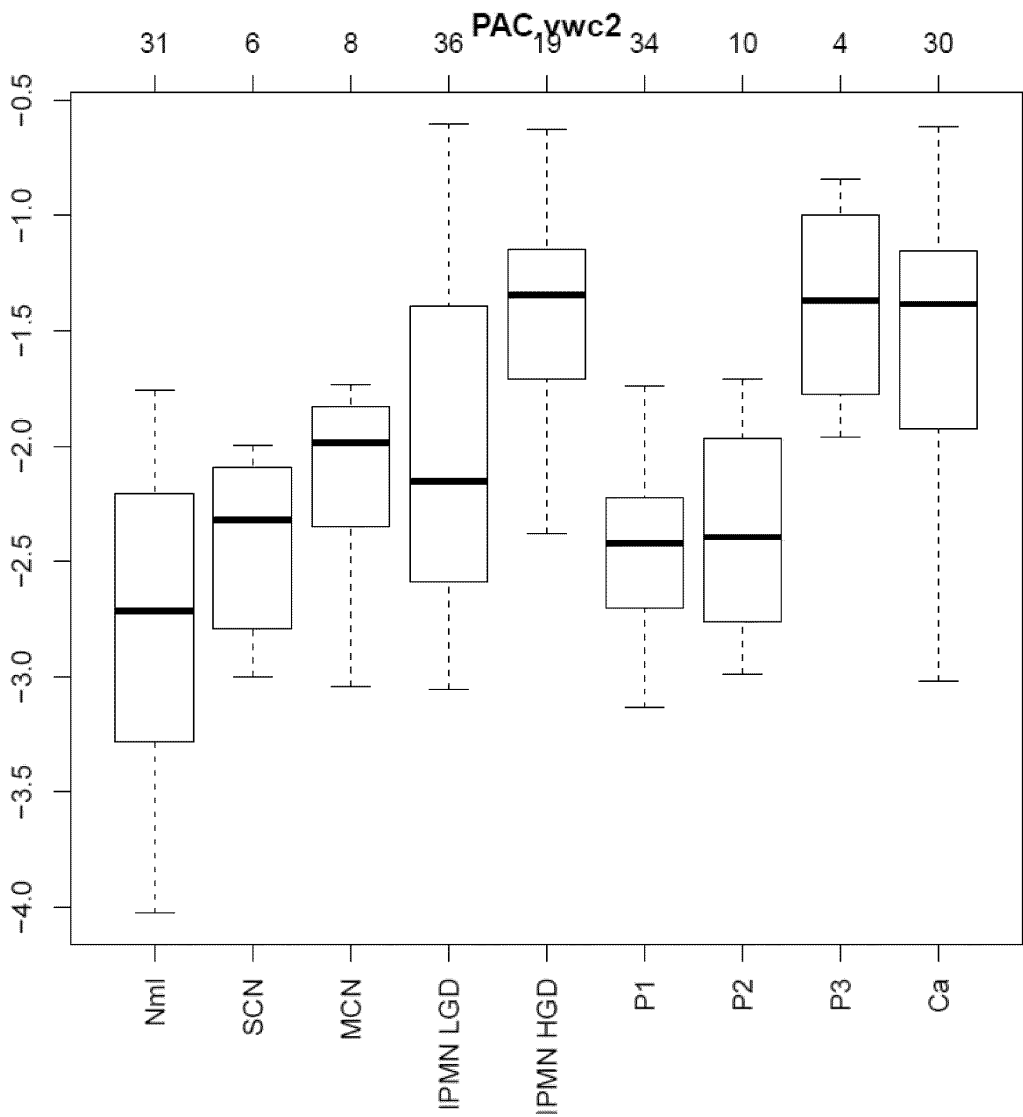
도면1p



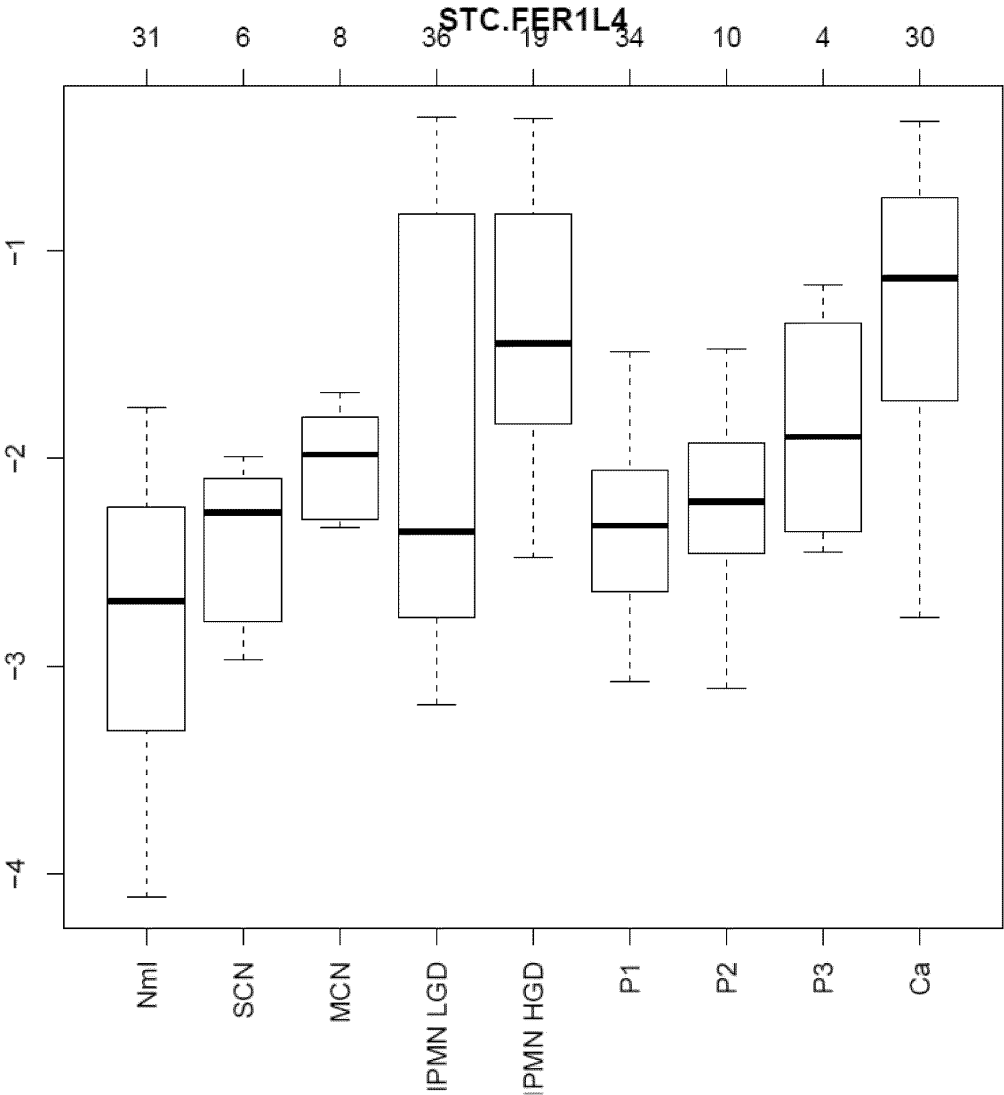
도면1q



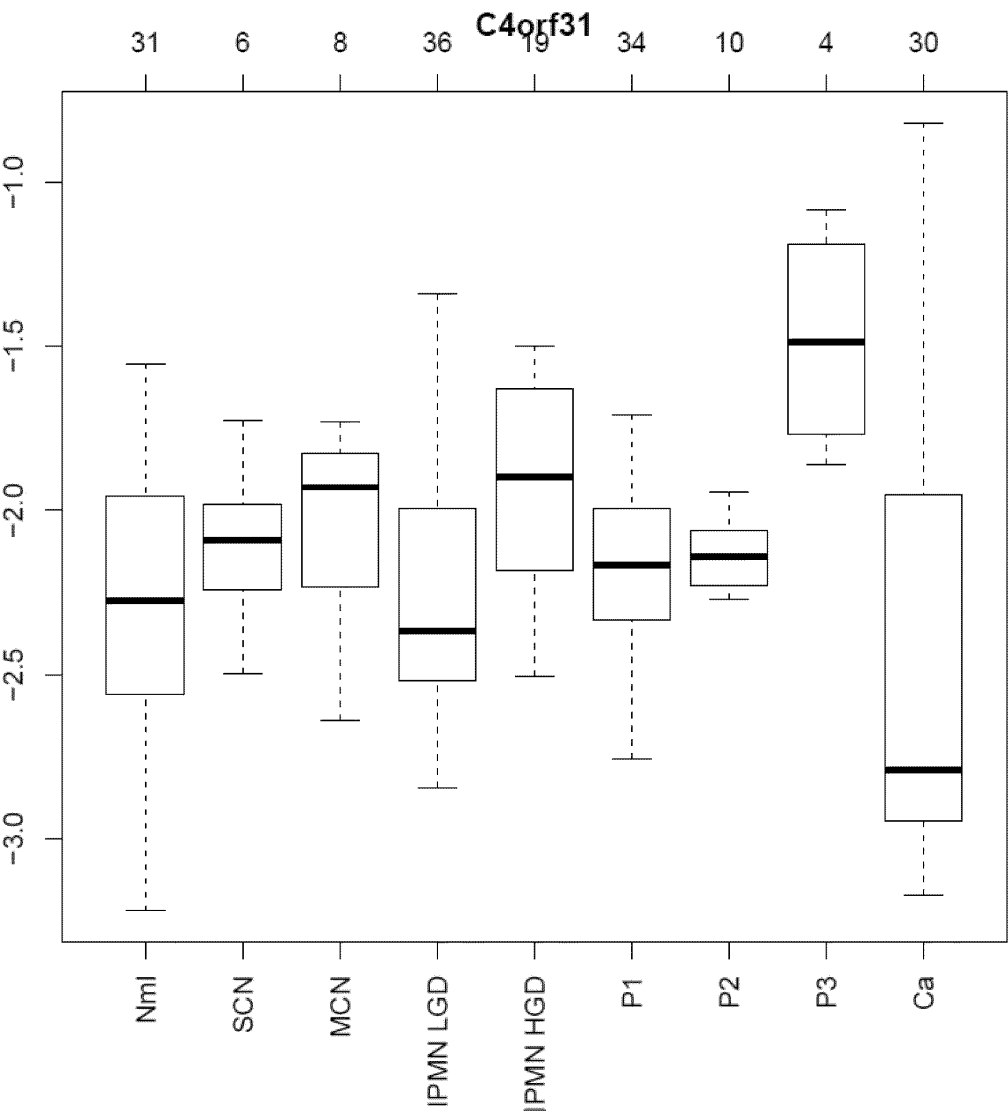
도면1r



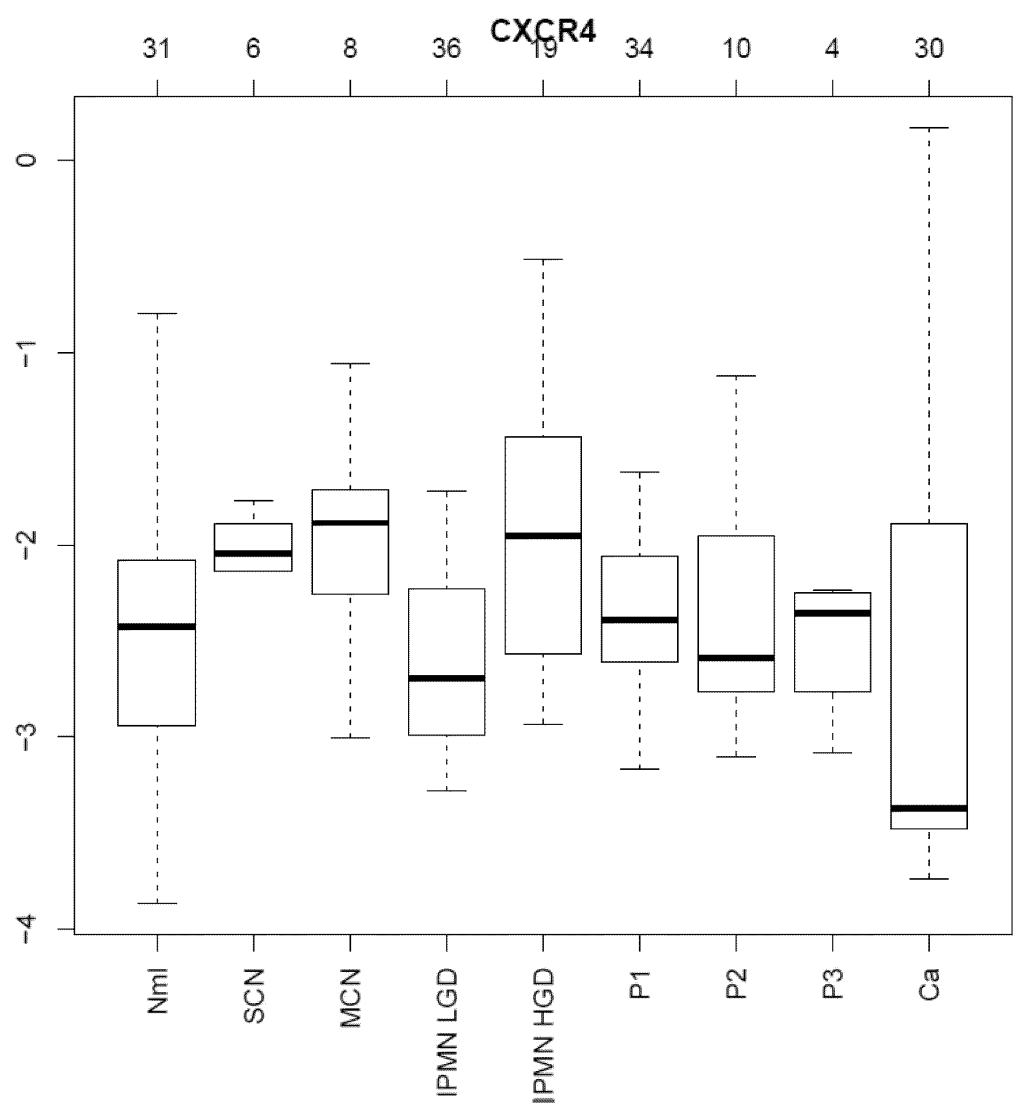
도면1s



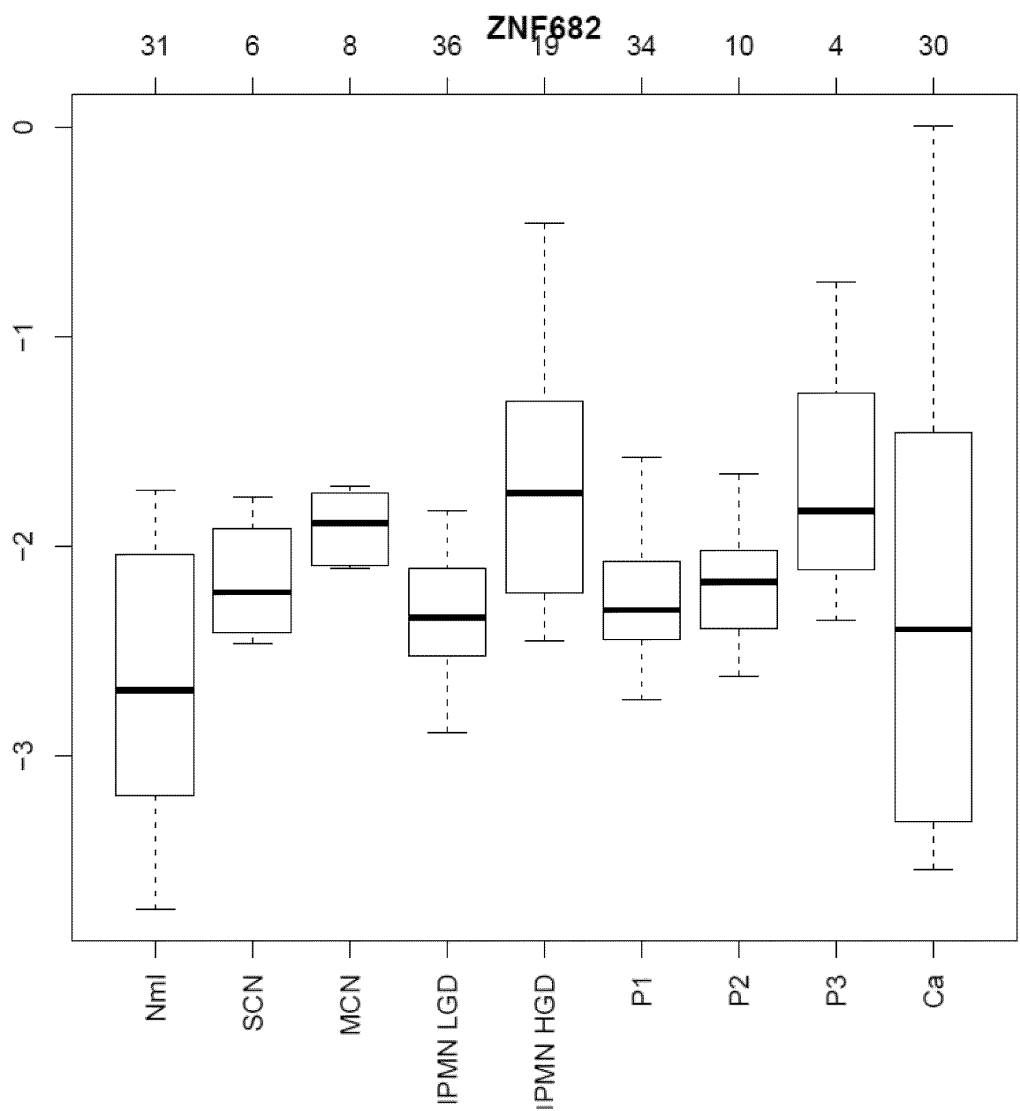
도면1t



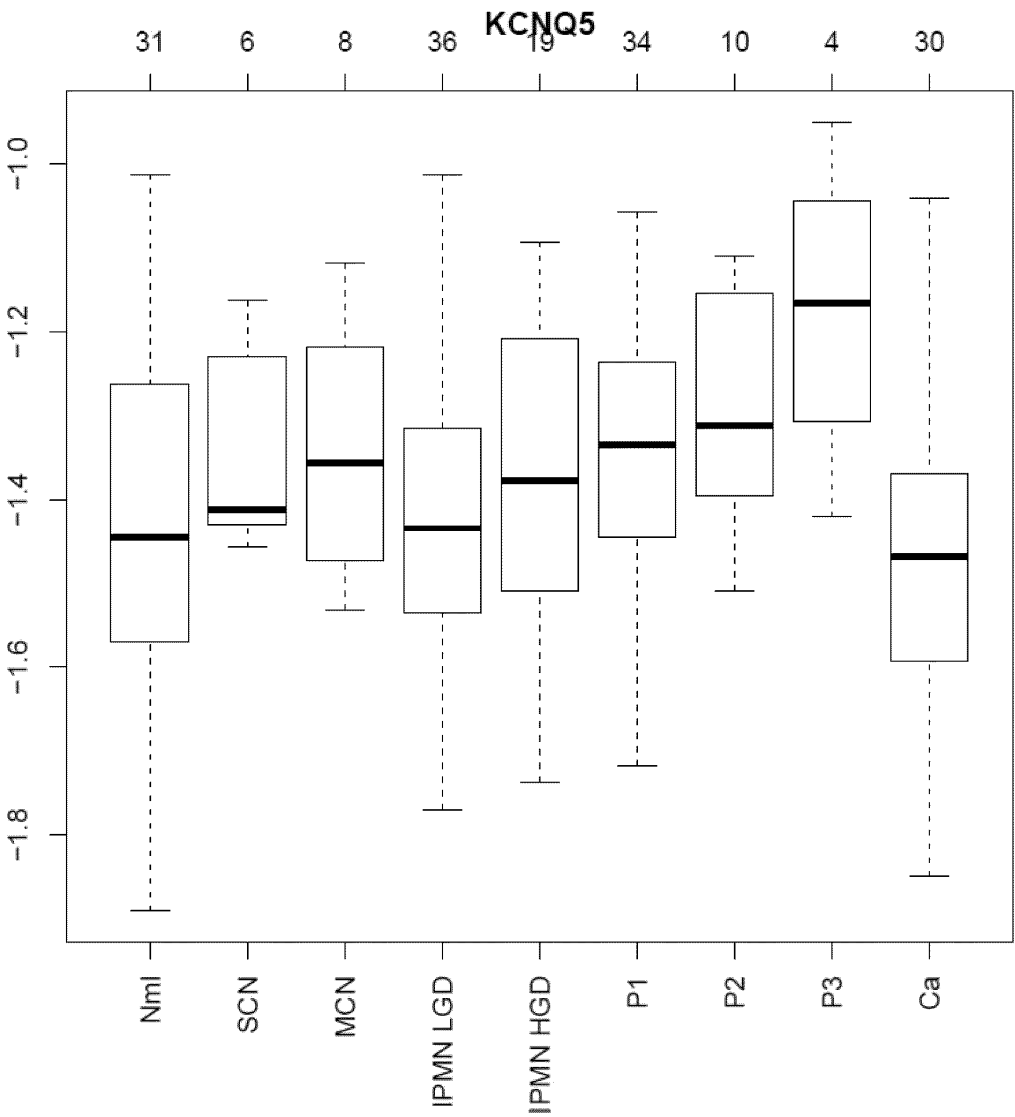
도면1u



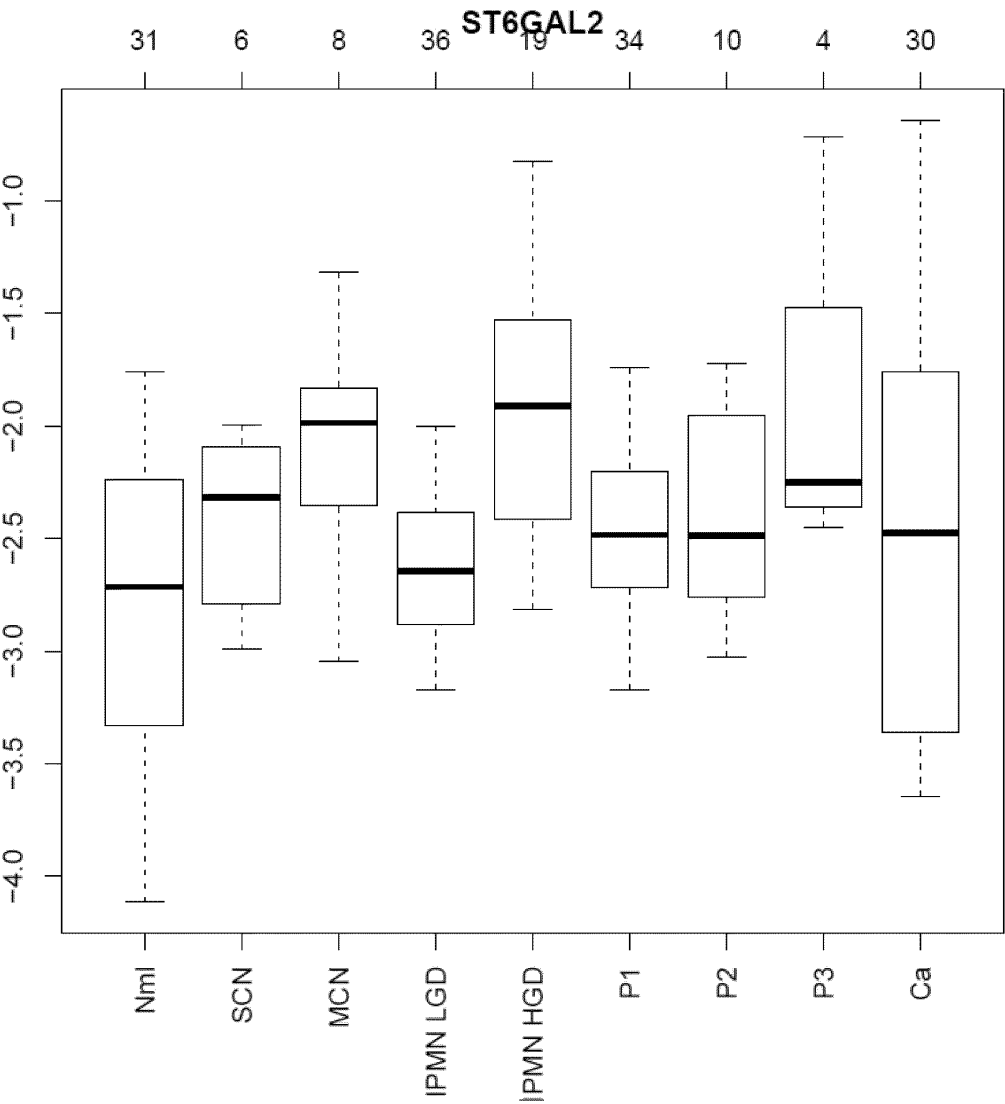
도면1v



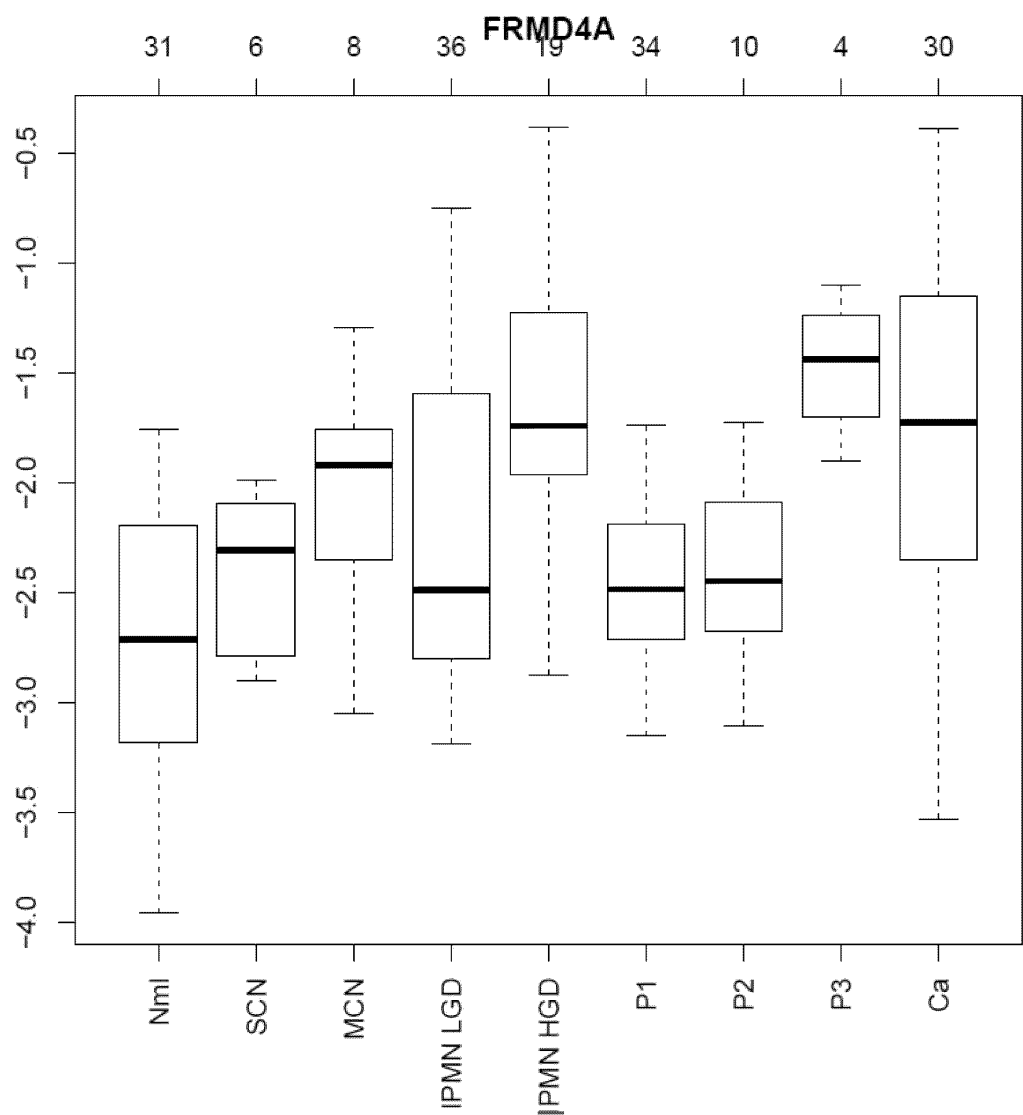
도면1w



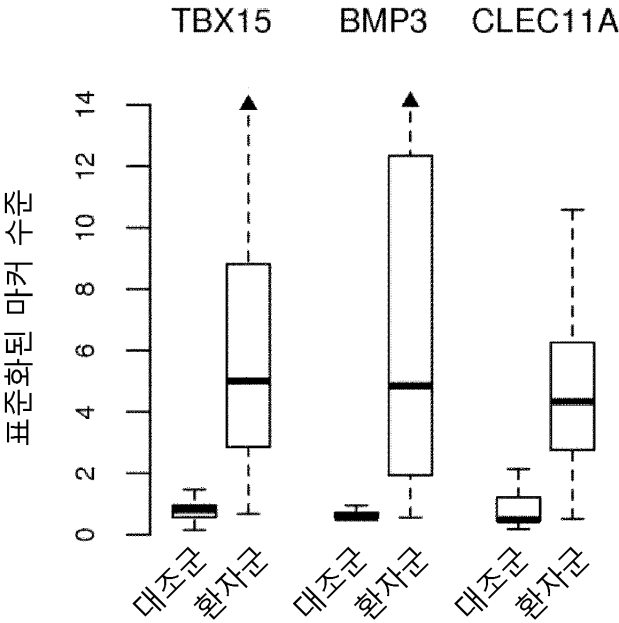
도면1x



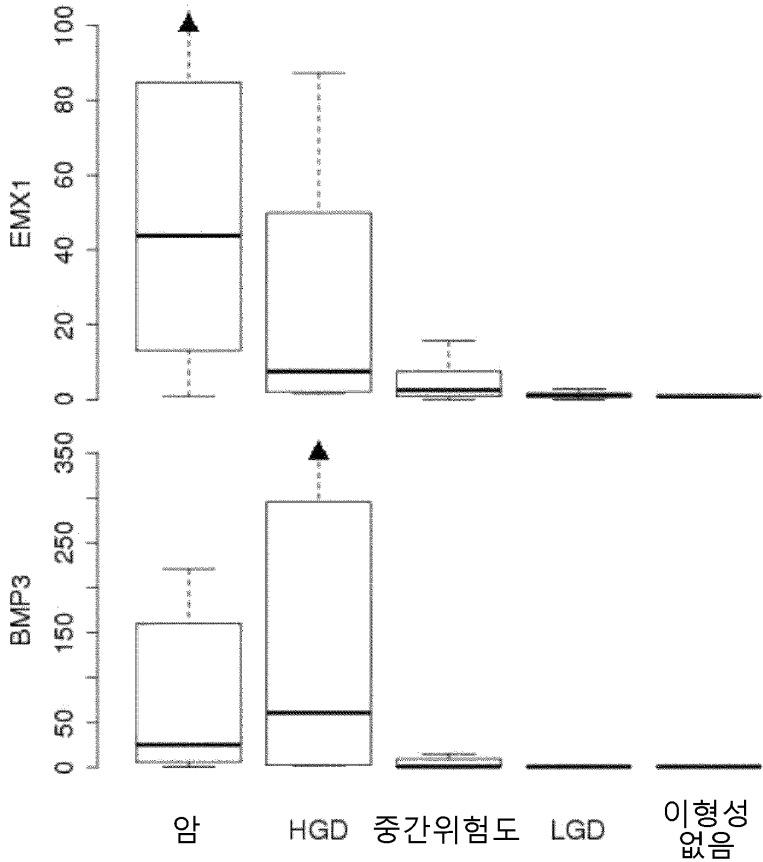
도면1y



도면2



도면3



서열 목록

SEQUENCE LISTING

<110> MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH

<120> DETECTING PANCREATIC HIGH-GRADE DYSPLASIA
 <130> EXCTM34830
 <150> US 62/322,610
 <151> 2016-04-14
 <160> 50
 <170> PatentIn version 3.5
 <210> 1
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 1
 gattttgttc gtcgtagtg c 21
 <210> 2
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400>
 > 2
 tctctaaacc cgcaacga 19
 <210> 3
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 3
 tcggcgtatt tttcgtagac gc 22
 <210> 4
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 4
 cgcaatctta aacgtacgct tcga 24

<210> 5
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223>
 synthetic
 <400> 5
 tgtttttcgcg gtcgttatat attacgtcgt 30
 <210> 6
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 6
 gaactatccc cgaactccga ctcca 25
 <210> 7
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 7
 tttagggaag taaagcgtcg ttttc 25
 <210> 8
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 8
 aacgacgtct cgatacctac ga 22
 <210> 9
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic

<400> 9	
agtttggcgt agtcggtaga tc	22
<210> 10	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 10	
gcgcgcaaat accgaataaa cg	22
<210> 11	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 11	
gtatttttaa tatttggtga gtgcg	25
<210> 12	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 12	
tactaaaacg tacgataaac atcgt	25
<210> 13	
<211> 34	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 13	
ttatatTTTT cgttttagt aatttcgct tagc	34
<210> 14	
<211> 23	
<212> DNA	
<213>	

Artificial Sequence

<220><223> synthetic

<400> 14

gaaaaccgcg cgaacattt cga 23

<210> 15

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 15

cgggttttag cgatgtttat ttagtttcg t 31

<210> 16

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 16

ccttttcgtt cgtataaaat ttcgtt 26

<210> 17

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 17

tggagttttt ttcgtcgttg acgc 24

<210> 18

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 18

aataccgcgt cgacaaaaa cg 22

<210> 19

<211> 25

<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 19	
tgtaagcggg tggtatttcg gaacg	25
<210> 20	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 20	
caaaaacctt cgcccgaacg aa	22
<210> 21	
<211> 34	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 21	
ttatatTTTT cgtttttagt aatttcgcgt tagc	34
<210> 22	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 22	
gaaaaccgc cgaacattt cga	23
<210> 23	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 23	
tgggttatcg gttttttaag ttcgg	25

<210> 24	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 24	
gaattcctcc caaccaaccc tctacg	26
<210> 25	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 25	
cgtggtgtgc gtttgaata tcgc	24
<210> 26	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 26	
ataatccccg ctaactaacg acg	23
<210> 27	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 27	
tcgctaggtg atagggaggg gttac	25
<210> 28	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 28	

acaaaacaca tcctattaac gcgaa	25
<210> 29	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 29	
gtcgtttttg ggcgttttag aggc	24
<210> 30	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 30	
taatcccaaa cgaaccgccg	20
<210> 31	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 31	
cgggggtaaa aagggtagcg atagtatac	30
<210> 32	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 32	
cgaaaaattc gaaacgaaac gtc	23
<210> 33	
<211> 24	
<212> DNA	
<213> Artificial Sequence	

<220><223> synthetic

<400> 33

gttagttttc ggtcgggatc gagc 24

<210> 34

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 34

cgttcgtaac gtcgacgatt accgta 26

<210> 35

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 35

gacgtttgtc gtcgggttcg ttc 23

<210> 36

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 36

aaaaaccctc cgctaccact tcgc 24

<210> 37

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> synthetic

<400> 37

cgcgtaggtc ggattgattt ttcgt 25

<210> 38

<211> 23

<212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 38
 ttaattaacg cccgaacacc cgc 23
 <210> 39
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 39
 ttggagagtt tttcgaattt tttcgc 26
 <210> 40
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 40
 gaaaaccacc ctaacgccg 19

 <210> 41
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 41
 tcgttaggaa ttagtggttg gatttatcgc 30
 <210> 42
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 42
 cgttctctta tttccgaat acgcgaa 27
 <210> 43

<211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 43
 gcgcgtagcg gcgtttc 17

<210> 44
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 44
 cccatatcgc ccgacgtaa 19

<210> 45
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 45
 gtttaatttt cggtttcgtc gtc 23

<210> 46
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 46
 cgctacgaaa cactccga 18

<210> 47
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic
 <400> 47

cggttttcgt tcgttttttc g	21
<210> 48	
<211> 18	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 48	
ccgccttcta cgcgacta	18
<210> 49	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 49	
cgccgaggcg gttttttgcg c	21
<210> 50	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic	
<400> 50	
ccacggacgg ttcgtttatc gc	22