



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0100051
 (43) 공개일자 2013년09월09일

- | | |
|---|--|
| (51) 국제특허분류(Int. Cl.)
C12Q 1/68 (2006.01) C12N 15/11 (2006.01)
GOIN 33/53 (2006.01)
(21) 출원번호 10-2012-7029015
(22) 출원일자(국제) 2011년04월05일
심사청구일자 없음
(85) 번역문제출일자 2012년11월05일
(86) 국제출원번호 PCT/US2011/031308
(87) 국제공개번호 WO 2011/127099
국제공개일자 2011년10월13일
(30) 우선권주장
13/080,616 2011년04월05일 미국(US)
61/321,124 2010년04월05일 미국(US) | (71) 출원인
프로그노시스 바이오사이언스, 인코포레이티드
미국 캘리포니아 라호이아 사우스 코스트 블러바드 505 (우: 92037)
(72) 발명자
체, 마크, 에스.
미국 캘리포니아 라호이아 사우스 코스트 블러바드 505 (우: 92037) 프로그노시스 바이오사이언스, 인코포레이티드 (내)
(74) 대리인
특허법인 남앤드남 |
|---|--|

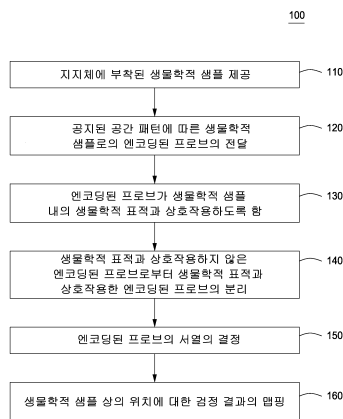
전체 청구항 수 : 총 57 항

(54) 발명의 명칭 공간적으로 엔코딩된 생물학적 검정

(57) 요약

본 발명은 공간적으로 엔코딩된 생물학적 검정에 사용하기 위한 검정 및 검정 시스템을 제공한다. 본 발명은 시약이 규정된 공간 패턴 내에 생물학적 샘플로 제공되는 높은 수준의 다중화가 가능한 검정; 공간 패턴에 따라 시약의 조절된 전달이 가능한 장비; 및 본질적으로 디지털인 판독값을 제공하는 디코딩 도식을 포함하는 검정 시스템을 제공한다.

대표도 - 도1



특허청구의 범위

청구항 1

샘플 내의 다수의 부위에서 다수의 생물학적 표적의 풍부함(abundance) 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴(spatial pattern)을 결정하는 검정 시스템으로서,

지지체에 부착된 샘플을 제공하는 단계;

공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 생물학적 표적에 대한 엔코딩된 프로브(encoded probe)를 전달하는 단계로서, 각각의 엔코딩된 프로브가 생물학적 표적과 상호작용할 수 있는 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그(tag)를 포함하는 단계;

엔코딩된 프로브가 생물학적 표적과 상호작용하도록 하는 단계;

생물학적 표적과 상호작용하지 않는 엔코딩된 프로브로부터 생물학적 표적과 상호작용하는 엔코딩된 프로브를 분리시키는 단계;

엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및

샘플 내의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행하는, 검정 시스템.

청구항 2

제 1항에 있어서, 생물학적 표적이 핵산이고, 엔코딩된 프로브가 올리고뉴클레오티드인 검정 시스템.

청구항 3

제 2항에 있어서, 다수의 핵산 표적 각각에 대한 2개의 엔코딩된 프로브가 존재하는 검정 시스템.

청구항 4

제 1항에 있어서, 다수의 생물학적 표적이 단백질이고, 엔코딩 프로브의 프로브 영역이 단백질이고, 코딩 태그가 올리고뉴클레오티드를 포함하는 검정 시스템.

청구항 5

제 4항에 있어서, 다수의 생물학적 표적이 효소를 포함하는 검정 시스템.

청구항 6

제 1항에 있어서, 엔코딩된 프로브의 프로브 영역이 항체를 포함하는 검정 시스템.

청구항 7

제 1항에 있어서, 엔코딩된 프로브의 프로브 영역이 앵타머(aptamer)인 검정 시스템.

청구항 8

제 1항에 있어서, 엔코딩된 프로브의 프로브 영역이 소분자인 검정 시스템.

청구항 9

제 1항에 있어서, 분리 단계와 결정 단계 사이에 증폭 단계를 추가로 포함하는 검정 시스템.

청구항 10

제 1항에 있어서, 결정 단계가 핵산 시퀀싱(sequencing)에 의해 수행되는 검정 시스템.

청구항 11

제 10항에 있어서, 시퀀싱이 고-처리량 디지털 핵산 시퀀싱인 검정 시스템.

청구항 12

제 1항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 20개 초과인 검정 시스템.

청구항 13

제 12항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 50개 초과인 검정 시스템.

청구항 14

제 13항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 75개 초과인 검정 시스템.

청구항 15

제 14항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 100개 초과인 검정 시스템.

청구항 16

제 15항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 1,000개 초과인 검정 시스템.

청구항 17

제 16항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 10,000개 초과인 검정 시스템.

청구항 18

제 17항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 100,000개 초과인 검정 시스템.

청구항 19

제 18항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 1,000,000개 초과인 검정 시스템.

청구항 20

제 11항에 있어서, 적어도 100,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 21

제 20항에 있어서, 적어도 500,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 22

제 21항에 있어서, 적어도 1,000,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 23

제 1항에 있어서, 공지된 공간 패턴이 샘플의 조직학적 특징에 의해 결정되는 검정 시스템.

청구항 24

제 1항에 있어서, 소프트웨어에 의해 프로그래밍된 하드웨어가 전달 단계, 분리 단계, 결정 단계 및 결부 단계 중 적어도 2개의 단계를 수행하는 검정 시스템.

청구항 25

제 1항에 있어서, 엔코딩된 프로브의 프로브 영역이 단백질이고, 분리 단계가 친화성 포획 작용제에 의해 포획되는 생물학적 표적과 상호작용하는 엔코딩된 프로브에 의해 달성되는 검정 시스템.

청구항 26

제 1항에 있어서, 엔코딩된 프로브의 프로브 영역이 핵산이고, 분리 단계가 샘플의 세척에 의해 달성되는 검정 시스템.

청구항 27

샘플 내의 다수의 부위에서 다수의 핵산 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템으로서,

지지체에 부착된 샘플을 제공하는 단계;

공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 핵산 표적에 대한 올리고뉴클레오티드 프로브를 전달하는 단계;

올리고뉴클레오티드 프로브가 핵산 표적과 하이브리드화되도록 하는 단계;

샘플로부터 하이브리드화되지 않은 엔코딩된 올리고뉴클레오티드 프로브를 세척하는 단계;

공지된 공간 패턴에 따라 샘플 내의 다수의 부위의 위치에 하나 이상의 엔코딩 작용제를 전달하는 단계로서, 각각의 부위에 전달되는 엔코딩 작용제의 조합이 상이한, 단계;

엔코딩된 프로브를 형성시키기 위해 엔코딩 작용제 및 올리고뉴클레오티드 프로브를 커플링시키는 단계;

고-처리량 시퀀싱을 이용하여 엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및

샘플 내의 다수의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행하는, 검정 시스템.

청구항 28

제 27항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 20개 초과인 검정 시스템.

청구항 29

제 28항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 50개 초과인 검정 시스템.

청구항 30

제 29항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 75개 초과인 검정 시스템.

청구항 31

제 30항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 10,000개 초과인 검정 시스템.

청구항 32

제 31항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 100,000개 초과인 검정 시스템.

청구항 33

제 32항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 1,000,000개 초과인

검정 시스템.

청구항 34

제 33항에 있어서, 적어도 100,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 35

제 34항에 있어서, 적어도 1,000,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 36

제 27항에 있어서, 2개의 올리고뉴클레오티드 프로브가 다수의 핵산 표적 각각에 대해 전달되는 검정 시스템.

청구항 37

제 27항에 있어서, 커플링 단계가 라이게이션(ligation)에 의해 수행되는 검정 시스템.

청구항 38

제 27항에 있어서, 커플링 단계가 신장(extension) 후의 라이게이션에 의해 수행되는 검정 시스템.

청구항 39

제 27항에 있어서, 커플링 단계와 결정 단계 사이에 증폭 단계를 추가로 포함하는 검정 시스템.

청구항 40

샘플 내의 다수의 부위에서 다수의 단백질 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템으로서,

지지체에 부착된 샘플을 제공하는 단계;

공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 단백질 표적에 대한 엔코딩된 프로브를 전달하는 단계로서, 각각의 엔코딩된 프로브가 단백질 표적과 상호작용할 수 있는 단백질 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그를 포함하고, 엔코딩 프로브의 코딩 태그 및 단백질 프로브 영역이 일부분인 단계;

엔코딩된 프로브가 단백질 표적과 상호작용하도록 하는 단계;

단백질 표적과 상호작용하지 않는 엔코딩된 프로브로부터 단백질 표적과 상호작용하는 엔코딩된 프로브를 분리시키는 단계;

엔코딩된 프로브의 서열의 전부 또는 일부를 고 처리량 시퀀싱에 의해 결정하는 단계; 및

샘플 내의 다수의 부위의 위치에 대해 다수의 단백질 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행하는, 검정 시스템.

청구항 41

제 40항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 20개 초과인 검정 시스템.

청구항 42

제 41항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 50개 초과인 검정 시스템.

청구항 43

제 42항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 75개 초과인 검정 시스템.

청구항 44

제 43항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 100개 초과인 검정 시스템.

청구항 45

제 44항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 500개 초과인 검정 시스템.

청구항 46

제 45항에 있어서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위가 1000개 초과인 검정 시스템.

청구항 47

제 46항에 있어서, 적어도 10,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 48

제 40항에 있어서, 적어도 100,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 49

제 48항에 있어서, 적어도 1,000,000개의 엔코딩 프로브의 서열이 동시에 결정되는 검정 시스템.

청구항 50

제 40항에 있어서, 커플링 단계와 결정 단계 사이에 증폭 단계를 추가로 포함하는 검정 시스템.

청구항 51

제 40항에 있어서, 단백질 표적이 효소이고, 엔코딩된 프로브의 프로브 영역이 기질, 둘 모두의 효소에 대한 추정상 기질인 검정 시스템.

청구항 52

제 40항에 있어서, 엔코딩된 프로브의 프로브 영역이 친화성 포획 작용제인 검정 시스템.

청구항 53

제 52항에 있어서, 엔코딩된 프로브의 프로브 영역이 항체인 검정 시스템.

청구항 54

제 52항에 있어서, 엔코딩된 프로브의 프로브 영역이 앵타머인 검정 시스템.

청구항 55

제 40항에 있어서, 단백질 표적과 상호작용하지 않는 엔코딩된 프로브로부터 단백질 표적과 상호작용하는 엔코딩된 프로브를 분리시키는 것이 단백질 표적과 상호작용하지 않는 엔코딩된 프로브로부터 단백질 표적과 상호작용하는 엔코딩된 프로브를 구별하는 친화성 포획 작용제에 의해 달성되는 검정 시스템.

청구항 56

제 36항에 있어서, 코딩 태그가 올리고뉴클레오티드인 검정 시스템.

청구항 57

샘플 내의 다수의 부위에서 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템으로서,

지지체에 부착된 샘플을 제공하는 단계;

공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 생물학적 표적에 대한 엔코딩된 프로브를 전달하는 단계로서, 각각의 엔코딩된 프로브가 생물학적 표적과 상호작용할 수 있는 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그를 포함하고, 각각의 엔코딩된 프로브가 생물학적 표적을 확인하는 단계;

엔코딩된 프로브가 생물학적 표적과 상호작용하도록 하는 단계;

엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및

샘플 내의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행하는, 검정 시스템.

명세서

기술분야

[0001] 관련 출원에 대한 전후-참조

[0002] 본 출원은 2011년 4월 5일에 출원된 미국 특허 출원 제13/080,616호, 및 2010년 4월 5일에 출원된 미국 가특허 출원 제61/321,124호를 우선권으로 주장하며, 상기 출원들은 참조로서 본원에 포함된다.

[0003] 발명의 분야

[0004] 본 발명은 생물학적 분자의 검정, 더욱 특히 고체 샘플 내의 많은 수의 생물학적 분자의 공간적 분포를 동시에 결정하기 위한 검정에 관한 것이다.

배경기술

[0005] 발명의 배경

[0006] 하기 논의에서, 배경 및 소개의 목적상 특정 문헌 및 방법이 기재될 것이다. 본원에 포함된 어떤 것도 종래 분야의 "시인(admission)"으로 간주되어선 안된다. 출원은 특히 적절한 경우 본원에 언급된 문헌 및 방법이 적용 가능한 법률상 규정하에서 종래 분야를 구성하지 않는 것을 입증하는 권리를 유보한다.

[0007] 포괄적인 유전자 발현 분석 및 단백질 분석은 생물학의 메커니즘을 이해하는 유용한 수단이었다. 이러한 수단의 이용은 발달 및 다양한 질병, 예를 들어, 암 및 자가면역 질병과 관련된 유전자 및 단백질의 확인을 가능케 하였다. 인 시츄 하이브리드화(in situ hybridization) 및 상이한 전사체의 다른 다중화된 검출과 같은 통상적인 방법은 유전자 발현의 공간적 패턴을 나타내었고, 발달 및 질병의 분자적 기초를 설명하는데 도움이 되었다. 샘플 당 많은 RNA 서열의 정량적 분석을 가능케 하는 다른 기술은 마이크로어레이(Shi, et al., Nature Biotechnology, 24(9): 1151-61 (2006); 및 Slonim and Yanai, Plos Computational Biology, 5(10):e1000543 (2009) 참조); 유전자 발현의 연쇄 분석(serial analysis of gene expression, SAGE)(Velculescu, et al., Science, 270(5235):484-87 (1995) 참조), qPCR의 고-처리량 수행(Spurgeon, et al., Plos ONE, 3(2):e1662 (2008) 참조) 및 인 시츄 PCR(Nuovo, Genome Res., 4: 151-67 (1995) 참조)을 포함한다. 상기 방법은 유용하지만, 이들은 샘플 내의 많은 공간적 위치에서 많은 유전자의 발현 또는 다수의 단백질의 존재 및/또는 활성의 동시 측정이 불가능하다. 레이저 포획 미세절개는 적은 수의 위치에서 많은 유전자의 분석을 가능케 하나, 이는 매우 고비용이고, 고된 작업이고, 평가가 잘 이루어지지 않는다. 2D 포맷의 특정 PCR 검정은 공간적 정보를 보존(Armani, et al., Lab on a Chip, 9(24): 3526-34 (2009) 참조)하지만, 이러한 방법은 낮은 공간적 해상도를 갖는데, 이는 이들이 조직 샘플에 대한 무작위 접근 및 높은 수준의 다중화를 또한 방지하는 웰로의 조직의 물리적 전이에 의지하기 때문이다.

[0008] 현재, 많은 수의 유전자, 단백질, 또는 다른 생물학적 활성 분자의 공간적 발현 패턴을 고해상도로 동시에 분석하는 실제 방법이 존재하지 않는다. 따라서, 조직 내의 생물학적 분자의 재현가능한 고-해상도 공간 맵이 필요하다. 본 발명은 상기 필요를 충족시킨다.

발명의 내용

[0009] 발명의 개요

[0010] 본 개요는 하기 상세한 설명에서 추가로 기재되는 간소화된 형태의 개념의 선택을 소개하기 위해 제공된다. 이

러한 개요는 청구된 주제의 핵심 또는 필수 특징을 확인하기 위한 것이 아니며, 청구된 주제의 범위를 제한하기 위해 이용되는 것이 아니다. 청구된 주제의 다른 특징, 세부사항, 유용성, 및 장점은 수반되는 도면에서 예시되고, 첨부된 청구항에서 정의된 양태를 포함하는 하기 기재된 상세한 설명으로부터 명백해질 것이다.

- [0011] 본 발명은 조직에서의 생물학적 활성의 고-해상도 공간 맵을 제공하는 검정 시스템을 포함한다. 상기 검정 시스템은 엔코딩된 프로브가 정의된 공간 패턴의 생물학적 샘플에 제공되는 높은 수준의 다중화가 가능한 검정; 공간 패턴에 따라 시약의 조절된 전달이 가능한 장비; 및 사실상 디지털인 판독값(readout)을 제공하는 디코딩 도식(decoding scheme)을 포함한다. 요컨대, 본 발명은 시퀀싱의 고도로-동시적인 데이터 분석과 함께 인 시추 하이브리드화의 분석을 제공하는, 많은 위치의 많은 생물학적 표적을 고찰하는 능력을 제공한다.
- [0012] 따라서, 일부 구체예에서, 본 발명은 샘플 내의 다수의 부위에서 다수의 생물학적 표적의 풍부함(abundance) 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템을 제공하며, 이러한 검정 시스템은 지지체에 부착된 샘플을 제공하는 단계; 공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 생물학적 표적에 대한 엔코딩된 프로브(encoded probe)를 전달하는 단계로서, 각각의 엔코딩된 프로브가 생물학적 표적과 상호작용할 수 있는 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그(tag)를 포함하는 단계; 엔코딩된 프로브가 생물학적 표적과 상호작용하도록 하는 단계; 생물학적 표적과 상호작용하지 않는 엔코딩된 프로브로부터 생물학적 표적과 상호작용하는 엔코딩된 프로브를 분리시키는 단계; 엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및 샘플 내의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행한다.
- [0013] 본 발명의 특정 양태에서, 생물학적 표적은 핵산을 포함하고, 엔코딩된 프로브는 올리고뉴클레오티드이고, 일부 양태에서, 다수의 핵산 표적 각각에 대해 2개의 엔코딩된 프로브가 존재한다. 일부 양태에서, 다수의 생물학적 표적은 단백질을 포함하고, 엔코딩된 프로브의 프로브 영역은 단백질이고, 코딩 태그는 올리고뉴클레오티드를 포함한다. 일부 양태에서, 다수의 생물학적 표적은 효소를 포함한다. 일부 양태에서, 엔코딩된 프로브의 프로브 영역은 항체, 앵타머(aptamer) 또는 소분자를 포함한다.
- [0014] 검정 시스템의 일부 양태는 분리 단계와 결정 단계 사이에 증폭 단계를 추가로 포함한다. 일부 양태에서, 결정 단계는 핵산 시퀀싱에 의해 수행되고, 바람직한 양태에서, 시퀀싱은 고-처리량 디지털 핵산 시퀀싱이다.
- [0015] 본 발명의 일부 양태에서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위는 20개를 초과하고, 일부 양태에서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위는 50개를 초과하고, 일부 양태에서, 다수의 생물학적 표적의 생성물이 검정되고, 샘플 내의 다수의 부위는 75개, 100개, 150개, 500개, 750개, 1,000개, 5,000개, 10,000개, 25,000개, 50,000개, 100,000개, 500,000개, 또는 1,000,000개 또는 이 이상을 초과한다. 다른 양태에서, 적어도 50,000개의 엔코딩 프로브의 서열이 동시에 결정되고, 다른 양태에서, 적어도 100,000개의 엔코딩 프로브의 서열이 동시에 결정되고, 일부 양태에서, 적어도 500,000개의 엔코딩 프로브의 서열이 동시에 결정되고, 일부 양태에서, 적어도 1,000,000개, 10,000,000개, 100,000,000개, 1,000,000,000개, 10,000,000,000개, 100,000,000,000개 또는 이 이상의 엔코딩 프로브의 서열이 동시에 결정된다.
- [0016] 일부 양태에서, 공지된 공간 패턴은 샘플의 조직학적 특징에 의해 결정된다. 또한, 일부 양태에서, 소프트웨어에 의해 프로그래밍된 하드웨어가 전달 단계, 분리 단계, 결정 단계 및 결부 단계 중 적어도 2개의 단계를 수행한다.
- [0017] 일부 양태에서, 엔코딩된 프로브의 프로브 영역은 단백질이고, 분리 단계는 친화성 포획 작용체에 의해 포획되는 생물학적 표적과 상호작용하는 엔코딩된 프로브에 의해 달성된다. 일부 양태에서, 엔코딩 프로브의 프로브 영역은 핵산이고, 분리 단계는 샘플의 세척에 의해 달성된다.
- [0018] 다른 구체예에서, 샘플 내의 다수의 위치에서 다수의 핵산 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템이 제공되며, 이러한 검정 시스템은 지지체에 부착된 샘플을 제공하는 단계; 공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 핵산 표적에 대한 올리고뉴클레오티드 프로브를 전달하는 단계; 올리고뉴클레오티드 프로브가 핵산 표적과 하이브리드화되도록 하는 단계; 샘플로부터 하이브리드화되지 않은 엔코딩된 올리고뉴클레오티드 프로브를 세척하는 단계; 공지된 공간 패턴에 따라 샘플 내의 다수의 부위의 위치에 하나 이상의 엔코딩 작용체를 전달하는 단계로서, 각각의 부위에 전달되는 엔코딩 작용체의 조합이 상이한, 단계; 엔코딩된 프로브를 형성시키기 위해 엔코딩 작용체 및 올리고뉴클레오티드 프로브를 커플링시키는 단계; 고-처리량 시퀀싱을 이용하여 엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및 샘플

내의 다수의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행한다.

[0019] 본 발명의 다른 구체예는 샘플 내의 다수의 부위에서 다수의 단백질 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템을 제공하며, 이러한 검정 시스템은 지지체에 부착된 샘플을 제공하는 단계; 공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 단백질 표적을 위한 엔코딩된 프로브를 전달하는 단계로서, 각각의 엔코딩된 프로브가 단백질 표적과 상호작용할 수 있는 단백질 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그를 포함하고, 코딩 태그의 엔코딩 프로브의 단백질 프로브 영역이 일부분인 단계; 엔코딩된 프로브가 단백질 표적과 상호작용하도록 하는 단계; 단백질 표적과 상호작용하지 않는 엔코딩된 프로브로부터 단백질 표적과 상호작용하는 엔코딩된 프로브를 분리시키는 단계; 엔코딩된 프로브의 서열의 전부 또는 일부를 고 처리량 시퀀싱에 의해 결정하는 단계; 및 샘플 내의 다수의 부위의 위치에 대해 다수의 단백질 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행한다.

[0020] 다른 구체예는 샘플 내의 다수의 부위에서 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두의 공간 패턴을 결정하는 검정 시스템을 제공하며, 이러한 검정 시스템은 지지체에 부착된 샘플을 제공하는 단계; 공지된 공간 패턴의 샘플 내의 다수의 부위에 다수의 생물학적 표적에 대한 엔코딩된 프로브를 전달하는 단계로서, 각각의 엔코딩된 프로브가 생물학적 표적과 상호작용할 수 있는 프로브 영역 및 엔코딩된 프로브가 전달되는 부위의 위치를 확인하는 코딩 태그를 포함하고, 각각의 엔코딩된 프로브가 생물학적 표적을 확인하는 단계; 엔코딩된 프로브가 생물학적 표적과 상호작용하도록 하는 단계; 엔코딩된 프로브의 서열의 전부 또는 일부를 결정하는 단계; 및 샘플 내의 부위의 위치에 대해 다수의 생물학적 표적의 풍부함 또는 활성 또는 풍부함과 활성 둘 모두를 결부시키는 단계를 수행한다.

[0021] 본 발명의 검정 시스템은 검출되는 분자 및 상기 검출 시스템에 필요한 시약을 기초로 하여 다양한 검출 메커니즘을 이용할 수 있다. 본 발명의 검정 시스템과 함께 사용될 수 있는 예시적 방법은 하기에 보다 상세히 기재된다.

[0022] **도면의 설명**

[0023] 도 1은 본 발명의 검정 시스템의 간소화된 개관을 제공한다.

[0024] 도 2는 핵산을 검출하기 위한 본 발명의 검정 시스템의 일 구체예의 간소화된 개관을 제공한다.

[0025] 도 3은 도 2에서 개관된 검정의 일 구체예의 대표적 도면이다.

[0026] 도 4는 본 발명의 검정 시스템의 조합 엔코딩 도식의 일 구체예에 대한 일반적 메커니즘을 예시한다.

[0027] 도 5는 도 4에서 도시된 조합 엔코딩 도식의 구체예의 간소화된 특정 예를 제공한다.

[0028] **정의**

[0029] 본원에서 사용되는 용어는 보통의 의미 및 당업자에 의해 이해되는 것과 같은 통상적인 의미를 갖는다. 하기 정의는 독자가 본 발명을 이해하는데 도움을 주며, 특별히 표시하지 않는 한 상기 용어의 의미를 변화시키거나 달리 제한하고자 하는 것은 아니다.

[0030] 본원에서 사용되는 용어 "항체"는 완전한 면역글로불린 또는 항체 또는 항원에 특이적으로 결합할 수 있는 면역글로불린 분자의 임의의 기능적 단편(항체 및 항원은 "결합 파트너"로 본원에서 정의된다)을 의미한다. 본원에서 사용되는 용어 "항체"는 완전한 항체 뿐만 아니라 관심 항원 또는 항원성 단편에 결합할 수 있는 임의의 항체 단편을 포함한다. 이러한 펩티드의 예는 완전한 항체 분자, 항체 단편, 예를 들어, Fab, F(ab')₂, CDRS, VL, VH, 및 항원에 특이적으로 결합할 수 있는 항체의 임의의 다른 부분을 포함한다. 본 발명의 검정을 위한 항체는 본 발명의 검정에서 검출(즉, 생물학적 표적)되거나 검출에 사용(즉, 프로브)되는 단백질에 대해 면역반응성이거나 면역특이적이고, 이에 따라 상기 단백질에 특이적 및 선택적으로 결합한다.

[0031] 본원에서 사용되는 용어 "결합 작용제"는 관심 생물학적 분자에 특이적으로 결합하는 임의의 작용제를 의미한다.

[0032] "상보적" 또는 "실질적으로 상보적"은 뉴클레오티드 또는 핵산 사이, 예를 들어, 이중-가닥 DNA 분자의 2개의 가닥 사이 또는 올리고뉴클레오티드 프라이머와 단일-가닥 핵산 상의 프라이머 결합 부위 사이의 하이브리드화 또는 염기 페어링(pairing) 또는 듀플렉스(duplex)의 형성을 의미한다. 상보적 뉴클레오티드는 일반적으로 A와 T(또는 A와 U), 또는 C와 G이다. 2개의 단일-가닥 RNA 또는 DNA 분자는 하나의 가닥의 뉴클레오티드가 최적으로

로 배열되고, 적절한 뉴클레오티드 삽입 또는 결실과 비교되고, 나머지 가닥의 적어도 약 80%, 보통 적어도 약 90% 내지 약 95%, 심지어 약 98% 내지 약 100%와 쌍을 이루는 경우에 실질적으로 상보적이라고 언급된다.

[0033] "하이브리드화"는 2개의 단일-가닥 폴리뉴클레오티드가 비공유적으로 결합하여 안정적인 이중-가닥 폴리뉴클레오티드를 형성하는 과정을 의미한다. 결과로서 발생된 (보통) 이중-가닥의 폴리뉴클레오티드는 "하이브리드" 또는 "듀플렉스"이다. "하이브리드화 조건"은 통상적으로 약 1M 미만, 종종 약 500 mM 미만의 염 농도를 포함할 것이나, 이는 약 200 mM 미만일 수 있다. "하이브리드화 완충액"은 완충된 염 용액, 예를 들어, 5% SSPE, 또는 당 분야에 공지된 상기 다른 완충액이다. 하이브리드화 온도는 5°C만큼 낮을 수 있으나, 통상적으로 22°C를 초과하고, 더욱 통상적으로 약 30°C를 초과하고, 통상적으로 37°C를 초과한다. 하이브리드화는 종종 엄격한 조건, 즉, 프라이머가 이의 표적 하위서열(subsequence)에 하이브리드화되나, 다른 비-상보적 서열에는 하이브리드화되지 않는 조건하에서 수행된다. 엄격한 조건은 서열-의존성이고, 상이한 환경에서 상이하다. 예를 들어, 보다 긴 단편은 특이적 하이브리드화를 위해 짧은 단편보다 높은 하이브리드화 온도를 필요로 할 수 있다. 상보적 가닥의 염기 조성 및 길이, 유기 용매의 존재, 및 염기 미스매치의 정도를 포함하는 다른 요인이 하이브리드화의 엄격성에 영향을 미칠 수 있음에 따라, 임의의 하나의 파라미터 단독의 확실한 측정보다 파라미터의 조합이 더욱 중요하다. 일반적으로, 엄격한 조건은 규정된 이온 강도 및 pH에서 특정 서열에 대한 Tm보다 약 5°C 낮도록 선택된다. 예시적인 엄격한 조건은 약 7.0 내지 약 8.3의 pH 및 적어도 25°C의 온도에서 적어도 0.01M 내지 1M 이하의 소듐 이온 농도(또는 다른 염)의 염 농도를 포함한다. 예를 들어, 5xSSPE(750 mM NaCl, 50 mM 소듐 포스페이트, 5 mM EDTA (pH 7.4)) 및 약 30°C의 온도의 조건이 대립유전자-특이적 하이브리드화에 적합하나, 적합한 온도는 하이브리드화되는 영역의 길이 및/또는 GC 함량에 좌우된다.

[0034] "라이게이션"은 주형-유도 반응에서 2개 이상의 핵산, 예를 들어, 올리고뉴클레오티드 및/또는 폴리뉴클레오티드의 말단 사이에서 공유 결합 또는 연결을 형성하는 것을 의미한다. 결합 또는 연결의 특성은 매우 다양할 수 있고, 라이게이션은 효소적 또는 화학적으로 수행될 수 있다. 본원에서 사용되는 바와 같이, 라이게이션은 보통 효소적으로 수행되어 하나의 올리고뉴클레오티드의 5' 탄소 말단 뉴클레오티드와 또 다른 뉴클레오티드의 3' 탄소 사이에 포스포디에스테르 연결을 형성한다.

[0035] "핵산", "올리고뉴클레오티드", "올리고" 또는 본원에서 사용되는 문법상 상당 어구는 일반적으로 공유적으로 함께 연결된 적어도 2개의 뉴클레오티드를 의미한다. 핵산은 일반적으로 포스포디에스테르 결합을 함유할 것이나, 일부 경우에, 포스포라미다이트, 포스포로디티오에이트, 또는 메틸포스포로아미다이트 연결과 같은 대안적 백본, 또는 펩티드 핵산 백본 및 연결을 갖는 핵산 유사체가 포함될 수 있다. 다른 유사체 핵산은 잠금 핵산(locked nucleic acid), 양성 백본, 비이온성 백본 및 비-리보오스 백본을 포함하는 바이시클릭 구조를 갖는 것을 포함한다. 분자의 안정성을 증가시키기 위해 리보오스-포스페이트 백본의 변형이 수행될 수 있고, 예를 들어, PNA:DNA 하이브리드는 일부 환경에서 보다 높은 안정성을 나타낼 수 있다.

[0036] "프라이머"는 핵산 합성의 개시점으로 작용하여 폴리뉴클레오티드 주형과 듀플렉스를 형성할 수 있고, 주형을 따라 이의 3' 말단으로부터 신장될 수 있어 신장된 듀플렉스가 형성되는 천연이거나 합성인 올리고뉴클레오티드를 의미한다. 신장 과정 동안 첨가된 뉴클레오티드의 서열은 주형 폴리뉴클레오티드의 서열에 의해 결정된다. 프라이머는 보통 DNA 중합효소에 의해 신장된다.

[0037] 용어 "SNP" 또는 "단일 뉴클레오티드 다형태(single nucleotide polymorphism)"는 개체 사이의 유전적 변화, 예를 들어, 가변적인 유기체의 DNA 내의 단일 질소 함유 염기 위치를 의미한다. SNP는 유전체 전체에 걸쳐 발견되고, 개체 사이의 유전적 변화 중 많은 변화는 SNP 유전자좌에서의 변화로 인한 것이며, 종종 이러한 유전적 변화는 개체 사이의 표현형 변화를 발생시킨다. 본 발명에서 이용하기 위한 SNP 및 이의 각각의 대립유전자는 임의의 수의 소스, 예를 들어, 공적인 데이터베이스(U.C. Santa Cruz Human Genome Browser Gateway (<http://genome.ucsc.edu/cgi-bin/hgGateway>) 또는 NCBI dbSNP 웹사이트 (<http://www.ncbi.nlm.nih.gov/SNP/>))로부터 유래될 수 있거나, 미국 특허 번호 6,969,589호; 및 "인간 유전체 다형태(Human Genomic Polymorphisms)"를 제목으로 하는 미국 공개공보 번호 2006/0188875호에 기재된 바와 같이 실험적으로 결정될 수 있다. SNP의 이용이 본원에 제시된 구체예 중 일부에 기재되어 있으나, 다른 이대립유전자(biallelic) 또는 다중-대립유전자(multi-allelic) 유전성 마커가 또한 이용될 수 있음이 이해될 것이다. 이대립유전자 유전성 마커는 2개의 다형태, 또는 대립유전자를 갖는 것이다. 상기 언급된 바와 같이, 소질(trait)과 관련된 이대립유전자 유전성 마커에 대해, 대조군에 비한 병증(case) 그룹의 유전성 조성물에 더욱 풍부한 대립유전자가 "관련 대립유전자"로 언급되며, 다른 대립유전자가 "관련되지 않은 대립유전자"로 언급될 수 있다. 따라서, 제공된 소질(예를 들어, 질병 또는 약물 반응)과 관련된 각각의 이대립유전자 다형태에 대해, 해당하는 관련 대립유전자가 존재한다. 본원에 제시된 방법과 함께 이용될 수 있는 다른 이대립유전자

다형태는 다중뉴클레오티드 변화, 삽입, 결실, 및 전위를 포함하나, 이에 제한되지는 않는다. 본원에서 DNA에 대한 언급은 유전체 DNA, 미토콘드리아 DNA, 에피솜 DNA, 및/또는 DNA의 유도체, 예를 들어, 앰플리콘, RNA 전사체, cDNA, DNA 유사체 등을 포함할 수 있음이 추가로 인지될 것이다. 관련 연구에서 스크리닝되는 다형태 유전자좌는 이배체 또는 일배체 상태로 존재할 수 있고, 이상적으로는, 유전체 전체에 걸친 부위로부터 유래될 것이다.

[0038] 결합 파트너(예를 들어, 단백질, 핵산, 항체 또는 다른 친화성 포획 작용제 등)를 언급하는 경우 본원에서 사용되는 용어 "선택적으로 결합하다", "선택적으로 결합하는" 등은 지정된 검정 조건하에서의 선택적 하이브리드화를 보장하기 위해 높은 친화성 및/또는 상보성을 갖는 2개 이상의 결합 파트너의 결합 반응을 의미한다. 통상적으로, 특이적 결합은 백그라운드 신호의 표준 편차의 적어도 3배일 것이다. 따라서, 지정된 조건하에서, 결합 파트너는 이의 특정 "표적" 분자에 결합하고, 샘플에 존재하는 다른 분자에는 유의한 양으로 결합하지 않는다.

[0039] "시퀀싱", "서열 결정" 등은 핵산의 뉴클레오티드 염기 서열과 관련된 정보의 결정을 의미한다. 이러한 정보는 핵산의 부분적 서열 정보 뿐만 아니라 완전한 서열 정보의 확인 또는 결정을 포함할 수 있다. 서열 정보는 다양한 정도의 통계적 확실성 또는 신뢰성으로 결정될 수 있다. 일 양태에서, 상기 용어는 핵산에서의 다수의 연속된 뉴클레오티드의 확인 및 순서의 결정을 포함한다. "고 처리량 디지털 시퀀싱" 또는 "차세대 시퀀싱"은 많은(통상적으로, 수천 내지 수십억) 핵산 서열을 본질적으로 동시 방식으로 결정하는 방법, 즉, DNA 주형이 시퀀싱을 위해 한번에 하나가 아니라 전체 과정으로 제조되고, 많은 서열이 바람직하게는 동시에 판독되거나 대안적으로 자체가 동시화될 수 있는 초-고 처리량의 연속 과정을 이용하여 판독되는 방법을 이용한 서열 결정을 의미한다. 이러한 방법은 파이로시퀀싱(pyrosequencing)(예를 들어, 454 Life Sciences, Inc., Branford, CT에 의해 시판됨); 라이게이션에 의한 시퀀싱(예를 들어, SOLiD™ technology, Life Technology, Inc., Carlsbad, CA로 시판됨); 변형된 뉴클레오티드를 이용한 합성에 의한 시퀀싱(예를 들어, TruSeq™ and HiSeq™ technology by Illumina, Inc., San Diego, CA, HeliScope™ by Helicos Biosciences Corporation, Cambridge, MA, 및 PacBio RS by Pacific Biosciences of California, Inc., Menlo Park, CA로 시판됨), 이온 검출 기술에 의한 시퀀싱(Ion Torrent, Inc., South San Francisco, CA); DNA 나노볼(nanoball)의 시퀀싱(Complete Genomics, Inc., Mountain View, CA); 나노포어(nanopore)-기반 시퀀싱 기술(예를 들어, Oxford Nanopore Technologies, LTD, Oxford, UK에 의해 개발됨), 및 유사한 고도로 동시화된 시퀀싱 방법을 포함하나, 이에 제한되지는 않는다.

[0040] 용어 " T_m "은 "융해 온도"와 관련하여 사용된다. 융해 온도는 이중-가닥 핵산 분자 집단이 반으로 해리되어 단일 가닥이 되는 온도이다. 핵산의 T_m 을 계산하기 위한 여러 식이 당 분야에 널리 공지되어 있다. 표준 참고문헌에 의해 표시되는 바와 같이, T_m 값의 간단한 측정은 핵산이 1M NaCl의 수용액에 존재하는 경우 식 $T_m=81.5+0.41(\% G+C)$ 에 의해 계산될 수 있다(예를 들어, Anderson and Young, Quantitative Filter Hybridization, in *Nucleic Acid Hybridization* (1985) 참조). 다른 참고문헌(예를 들어, Allawi and SantaLucia, Jr., *Biochemistry*, 36: 10581-94 (1997))은 T_m 의 계산을 위해 구조적 및 환경적 특징 뿐만 아니라 서열 특징을 고려하는 대안적 계산 방법을 포함한다.

[0041] **발명의 상세한 설명**

[0042] 본원에 기재된 기술의 실시는 달리 표시하지 않는 한 당업자의 기술 범위 내인 유기화학, 중합체 기술, 분자생물학(재조합 기술을 포함함), 세포생물학, 생화학, 및 시퀀싱 기술의 통상적인 기술 및 기재를 이용할 수 있다. 이러한 통상적인 기술은 중합체 어레이 합성, 폴리뉴클레오티드의 하이브리드화 및 라이게이션, 및 라벨을 이용한 하이브리드화의 검출을 포함한다. 적합한 기술의 특이적 예시는 본원의 예에 대한 참조가 될 수 있다. 그러나, 물론 다른 동등한 통상적인 절차가 또한 이용될 수 있다. 이러한 통상적인 기술 및 기재는 문헌[Green, et al., Eds., *Genome Analysis: A Laboratory Manual Series* (Vols. I-IV) (1999); Weiner, Gabriel, Stephens, Eds., *Genetic Variation: A Laboratory Manual* (2007); Dieffenbach, Dveksler, Eds., *PCR Primer: A Laboratory Manual* (2003); Bowtell and Sambrook, *DNA Microarrays: A Molecular Cloning Manual* (2003); Mount, *Bioinformatics: Sequence and Genome Analysis* (2004); Sambrook and Russell, *Condensed Protocols from Molecular Cloning: A Laboratory Manual* (2006); 및 Sambrook and Russell, *Molecular Cloning: A Laboratory Manual* (2002) (모두 Cold Spring Harbor Laboratory Press); Stryer, *Biochemistry* (4th Ed.) (1995) W.H. Freeman, New York N.Y.; Gait, *Oligonucleotide Synthesis: A Practical Approach*

(2002) IRL Press, London; Nelson and Cox, *Lehninger, Principles of Biochemistry* (2000) 3rd Ed., W. H. Freeman Pub., New York, N.Y.; 및 Berg, et al., *Biochemistry* (2002) 5th Ed., W.H. Freeman Pub., New York, N.Y.]과 같은 표준 실험실 메뉴얼에서 발견될 수 있고, 상기 모두는 모든 목적상 전체내용이 참조로서 본원에 포함된다.

[0043] 본원 및 첨부된 청구항에서 사용되는 바와 같은, 단수 형태는 문맥이 명백히 달리 기술하지 않는 한 복수의 지시대상을 포함함을 주의하라. 따라서, 예를 들어, "핵산"의 언급은 하나 이상의 핵산들을 의미하고, "검정"의 언급은 당업자에게 공지된 동등한 단계 및 방법의 언급을 포함하며, 다른 것도 마찬가지이다.

[0044] 달리 정의되지 않는 한, 본원에서 사용되는 모든 기술 및 과학 용어는 본 발명이 속하는 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 갖는다. 본원에 언급된 모든 간행물은 본원에 기재된 발명과 관련하여 사용될 수 있는 장치, 제형 및 방법을 기재하고 개시하기 위한 목적을 위해 참조로서 본원에 포함된다.

[0045] 일정 범위의 값이 제공되는 경우, 상기 범위의 상한 및 하한 및 언급된 범위 내의 임의의 다른 언급된 값 또는 사이에 존재하는 값 사이의 각각의 사이에 존재하는 값이 본 발명의 범위에 포함되는 것이 이해된다. 상기 보다 작은 범위의 상한 및 하한은 보다 작은 범위 내에 독립적으로 포함될 수 있고, 이는 또한 언급된 범위에서 임의의 특별히 배제된 한계에 종속되는 본 발명의 범위 내에 포함된다. 언급된 범위가 한계 중 하나 또는 둘 모두를 포함하는 경우, 둘 모두의 포함된 한계 중 어느 하나를 배제하는 범위가 또한 본 발명에 포함된다.

[0046] 하기 기재에서, 본 발명의 더욱 철저한 이해를 제공하기 위해 다수의 특정 세부사항이 기재된다. 그러나, 본 발명은 이러한 특정 세부사항 중 하나 이상 없이 실시될 수 있음이 당업자에게 명백할 것이다. 다른 예에서, 널리 공지된 특징 및 당업자에게 널리 공지된 절차는 본 발명을 불명료하게 하는 것을 피하기 위해 기재되지 않았다.

[0047] **전반적인 본 발명**

[0048] 본 발명의 검정 시스템은 1) 효과적인 공간 엔코딩 도식을 이용하여 높은 수준의 다중화가 가능한 검정; 2) 공간 패턴에 따라 시약을 전달할 수 있는 장비; 및 3) 사실상 디지털인 판독값(readout)에 의해 결정된 디코딩을 포함하는 공간적으로 엔코딩된 다중화 검정을 제공한다. 본 발명의 검정 시스템은 생물학적 표적 또는 생물학적 표적을 나타내는 생물학적 활성의 존재 또는 부재 및 상대량, 뿐만 아니라 생물학적 샘플, 예를 들어, 현미경 슬라이드 또는 배양 접시와 같은 지지체 상에 배치된 조직 섹션 또는 다른 생물학적 구조에서의 생물학적 표적 또는 활성의 위치를 검출한다.

[0049] 검정 시스템은 추가로 공간적으로 정의된 패턴으로 시약을 전달하는 능력을 갖는 장비를 제공한다. 소프트웨어, 시약 및 프로토콜과 함께 상기 장비는 본 발명의 고도로 혁신적인 검정 시스템의 주요 구성요소를 제공하여 다수의 생물학적 표적의 측정 또는 유전자 발현 및 펩티드 국소화를 포함하는 유의한 공간 환경에서의 활성의 측정을 가능케 한다. 이러한 검정 시스템에서 사용되는 엔코딩 도식은 다중화된 검정의 생성물이 생물학적 샘플로부터 분리되고 분석을 위해 풀링(pooling)된 후에 생물학적 샘플에서의 생물학적 표적 또는 활성 (또는 이의 결핍)의 위치를 결정하는 것을 가능케 한다. 엔코딩 도식의 디코딩은, 예를 들어, 저비용으로 수백만 내지 수조의 데이터포인트를 용이하게 제공하는 차세대 시퀀싱에 의해 수행될 수 있다. 생물학적 표적의 양 또는 활성과 같은 검정 결과는 이후에 생물학적 샘플 내의 특정 위치로 다시 맵핑될 수 있다. 검정 시스템은 신규한 분석 윈도우(window)가 생물학적 샘플 내에서의 세포 기능 및 조절의 복잡한 공간 패턴으로 통하게 한다.

[0050] 본 발명의 검정 시스템(100)의 간소화된 개관이 도 1에 제공된다. 단계(110)에서, 지지체에 부착된 생물학적 샘플이 제공된다. 생물학적 샘플은 관심 생물학적 표적을 함유한다. 생물학적 표적은 임의의 관심 분자, 예를 들어, 핵산(예를 들어, RNA 전사체, 유전체 DNA 서열, cDNA, 앰플리콘, 또는 다른 핵산 서열을 포함함) 및 단백질, 효소 등을 포함할 수 있다. 단계(120)에서, 엔코딩된 프로브는 공지된 공간 패턴에 따라 생물학적 샘플로 전달된다. 엔코딩된 프로브는 관심 생물학적 표적과 상호작용할 수 있는 프로브, 및 검정되는 생물학적 표적의 샘플 내의 위치를 확인하는 코딩 태그를 포함하며, 따라서 검정 결과를 샘플 내의 위치와 다시 연관시키는데 이용될 수 있다. 대부분의 구체예의 코딩 태그는 올리고뉴클레오티드이다. 그러나, 코딩 태그는 또한 대량의 태그, 형광 라벨, 또는 다른 모이어티(moiety)일 수 있다.

[0051] 일부 구체예에서, 프로브 및 엔코딩된 프로브의 코딩 태그 위치는 생물학적 샘플로 전달되기 전에 예비-커플링된다. 예를 들어, 엔코딩된 프로브가 올리고뉴클레오티드인 경우, 프로브 및 코딩 태그 서열 둘 모두는 단일한

올리고뉴클레오타이드로 합성될 수 있다. 대안적으로, 프로브 및 엔코딩 프로브의 코딩 태그 부분은 별개로 합성되거나 수득될 수 있고, 생물학적 샘플로의 전달 전에 조합될 수 있다(예를 들어, 2개의 별개의 올리고뉴클레오타이드가 합성되고, 예를 들어, 라이게이션에 의해 커플링될 수 있거나, 항체 및 올리고뉴클레오타이드는 개별적으로 제조될 수 있고, 생물학적 샘플로의 전달 전에 컨주게이션될 수 있다). 또한, 도 2-5에 기재된 바와 같이, 프로브 및 코딩 태그(엔코딩 올리고뉴클레오타이드 내)는 별개로 합성되고, 검정에서 상이한 단계에서 생물학적 샘플로 전달된다(예를 들어, 프로브가 먼저 및 코딩 태그가 이후, 또는 이와 반대).

[0052] 단계(130)에서, 엔코딩된 프로브는 생물학적 표적과 반응하거나 상호작용하게 되고, 즉, 예를 들어, 올리고뉴클레오타이드가 핵산 표적에 하이브리드화되고, 효소가 단백질 표적과의 반응을 촉매하게 하고, 항체가 에피토프에 결합하게 하는 조건 등이 제공된다. 생물학적 표적이 핵산인 경우, 엔코딩된 프로브는 통상적으로 올리고뉴클레오타이드이고, 표적 핵산에 하이브리드화된다. 생물학적 표적이 단백질인 경우, 엔코딩된 프로브는 통상적으로 표적 단백질과의 결합 또는 표적 단백질과의 반응에 의해 표적 단백질과 상호작용하는 앵타머, 소분자, 또는 올리고뉴클레오타이드-컨주게이션된 단백질이다(즉, 단백질 중 하나는 다른 단백질에 대한 기질이다). 엔코딩 올리고뉴클레오타이드는 적합한 기를 통한 컨주게이션, 화학적 또는 광-가교 등에 의해 프로브(단백질)에 커플링될 수 있다.

[0053] 엔코딩된 프로브가 생물학적 표적과 상호작용한 후, 생물학적 표적과 상호작용한 엔코딩된 프로브는 단계(140)에서 생물학적 표적과 상호작용하지 않은 엔코딩된 프로브로부터 분리되어야 한다. 생물학적 표적이 핵산이고, 엔코딩된 프로브가 올리고뉴클레오타이드인 경우, 분리는, 예를 들어, 샘플로부터 하이브리드화되지 않은 엔코딩된 프로브를 세척함으로써 달성될 수 있다. 유사하게, 앵타머, 소분자, 및 단백질 프로브를 이용한 검정을 포함하는 친화성 결합을 기초로 하는 다른 검정에 대해, 낮은 친화성 결합물을 제거하기 위해 세척 단계가 이용될 수 있다. 프로브가 표적과의 상호작용을 통해 변화되는 경우, 예를 들어, 프로테아제에 의한 분해 또는 키나제에 의한 인산화를 통하는 펩티드의 경우, 모든 엔코딩된 프로브, 즉, 생물학적 표적과 상호작용하고 변화된 엔코딩된 프로브 및 변화되지 않은 엔코딩된 프로브 둘 모두를 수거하는 것이 편리하다. 수집 또는 푸어링 후, 모이어터(예를 들어, 포스페이티기)의 첨가에 의해 변화된 프로브를 포획하는데 항체 또는 다른 친화성 포획 작용제가 사용될 수 있다. 프로브가 분해를 통해 변화된 경우, 변화된 프로브는, 예를 들어, 변화(예를 들어, 분해에 의한) 동안 변화된 프로브로부터 분리되는 태그를 통해 변화되지 않은 프로브를 포획하거나, 분해 부위에 신규한 태그를 첨가함으로써 분리될 수 있다.

[0054] 반응(변화)되거나 상호작용된 엔코딩된 프로브가 반응되지 않거나 상호작용하지 않은 엔코딩된 프로브로부터 분리된 후, 반응되고/되거나 상호작용을 한 엔코딩된 프로브의 서열은 바람직하게는 시퀀싱에 의해 결정된다. 엔코딩된 프로브의 서열은 다시 생물학적 샘플 내의 위치로의 검정 결과의 맵핑을 가능케 한다.

[0055] 도 2는 공간 정보의 엔코딩을 위한 조합 코딩 도식의 효과적인 이행을 구체화시키는 본 발명의 검정 시스템의 간소한 개관을 제공한다. 이러한 개관의 목적상, 프로브는 올리고뉴클레오타이드이나, 다른 곳에 설명된 바와 같이 다른 유형의 프로브가 또한 이용될 수 있다. 단계(210)에서, 지지체에 부착된 생물학적 샘플, 예를 들어, 조직 샘플 또는 다른 생물학적 구조가 제공된다. 단계(220)에서, 하나 이상의 올리고뉴클레오타이드 프로브가 생물학적 샘플로 전달되고, 이러한 올리고뉴클레오타이드 프로브는 생물학적 샘플 내의 생물학적 표적과 하이브리드화될 수 있다. 단계(230)에서, 올리고뉴클레오타이드 프로브는 핵산 표적과 상호작용(이에 하이브리드화)하는 것이 허용되며, 즉, 올리고뉴클레오타이드 프로브가 표적 핵산에 하이브리드화될 수 있는 적절한 조건이 제공된다.

[0056] 단계(240)에서, 표적 핵산과 하이브리드화되지 않은 올리고뉴클레오타이드 프로브가 분리되고, 이에 의해 표적 핵산에 하이브리드화된 올리고뉴클레오타이드 프로브로부터 분리된다. 이러한 구체예에서, 하이브리드화되지 않은 올리고뉴클레오타이드 프로브를 분리시키기 위해 분리는, 예를 들어, 샘플을 세척함으로써 달성될 수 있다. 다음으로, 단계(250)에서, 엔코딩 올리고뉴클레오타이드(엔코딩 작용제)가 선택된 공간 패턴에 따라 생물학적 샘플로 전달되고, 이러한 엔코딩 올리고뉴클레오타이드는 생물학적 샘플 내의 생물학적 표적의 위치를 엔코딩하는데 사용되는 코딩 태그를 포함한다. 도 1의 검정 시스템과 대조적으로, 여기서 프로브 및 엔코딩 작용제(엔코딩 올리고뉴클레오타이드)는 별개의 단계에서 전달되는 것을 주의하라. 단계(260)에서, 엔코딩 올리고뉴클레오타이드는 올리고뉴클레오타이드 프로브에 커플링되어 엔코딩된 프로브를 생성시킨다. 프로브가 올리고뉴클레오타이드인 경우, 엔코딩 올리고뉴클레오타이드는, 예를 들어, 라이게이션에 의해 올리고뉴클레오타이드 프로브에 커플링될 수 있다. 대안적으로, 엔코딩 올리고뉴클레오타이드 내의 정보는 프라이머로 작용함으로써 엔코딩 올리고뉴클레오타이드의 서열을 복제하고 통합시키는 프로브 올리고뉴클레오타이드를 신장시키는 DNA 증합효소를 이용함으로써 전달될 수 있다.

[0057] 단계(270)에서, 엔코딩된 프로브 내의 코딩 태그의 서열 뿐만 아니라 프로브 자체의 서열 또는 프로브 자체의 서열의 일부가 결정되고, 단계(280)에서, 표적 핵산이 생물학적 샘플로 다시 맵핑된다. 일부 구체예에서, 서열의 풍부함은 위치에서의 생물학적 표적의 상대량을 나타낸다. 이러한 구체예는 본 발명을 보다 잘 설명하기 위해 특정 순서의 개별적 단계를 나타내지만, 단계의 정확한 순서는 다양할 수 있다. 예를 들어, 단계(220 및 250)는 조합될 수 있어, 프로브 및 엔코딩 올리고뉴클레오티드의 혼합물이 선택된 공간 패턴에 따라 전달된다. 이후, 커플링 단계(260)가 조합 단계(220 및 250) 직후, 또는 조합 단계(220 및 250)와 동시에 수행될 수 있다. 이러한 경우, 단계(240)은 단계(260) 이후에 발생할 것이다. 따라서, 상기 일련의 단계의 2개의 주요 결과, 즉, 프로브 분자의 위치-특이적 엔코딩 및 상응하는 표적 분자와 상호작용하는 프로브 분자의 능력을 기초로 한 프로브 분자의 분리가 특정 단계의 이행에서 일부 유연성을 가지면서 달성될 수 있음이 인지될 수 있다. 유사하게, 코딩 도식의 설계에서 상당한 유연성이 존재한다. 하기에 기재되는 바와 같이, 본 발명의 검정은 조합 방법에 특히 적용가능하다.

[0058] 따라서, 본 발명은 시퀀싱의 고도로-동시적인 데이터 분석과 함께 인 시츄 하이브리드화의 분석을 제공하는, 많은 위치의 많은 생물학적 표적을 고찰하는 능력을 제공한다. 일부 구체예에서, 검정되는 다수의 생물학적 표적 및 생물학적 샘플 내의 다수의 부위의 합계는 20개를 초과하고, 다른 구체예에서, 검정되는 다수의 생물학적 표적 및 생물학적 샘플 내의 다수의 부위의 합계는 50개를 초과하고, 다른 구체예에서, 검정되는 다수의 생물학적 표적 및 생물학적 샘플 내의 다수의 부위의 합계는 100, 500, 1,000, 10,000, 25,000, 100,000, 500,000, 1,000,000개를 초과한다. 본 발명의 공간 엔코딩 치수로 인해, 훨씬 큰 수가 고려될 수 있음이 인지될 것이다. 예를 들어, 위치 x 10,000 위치 당 10,000개의 표적을 검정하는 것은 10^8 개의 상이한 검정을 발생시킬 것이며, 특히 단일 세포의 해상도와 거의 유사한 해상도를 갖는 공간 위치가 이용되는 경우에 상기 값보다 훨씬 큰 수가 용이하게 고려될 수 있다. 추가로, 고-처리량 디지털 시퀀싱이 사용되는 구체예에서, 적어도 1,000개의 엔코딩 프로브의 서열이 통상적으로 동시에 결정된다. 더욱 통상적으로, 디지털 판독값을 이용하여, 각각의 검정에 대한 다수의 서열 판독(프로브 및 공간 위치 코드에 의해 규정됨)을 수득하는 것이 바람직하다. 실험의 설계 및 검정의 필요조건에 따라 검정 당 적어도 3개의 카피, 더욱 통상적으로 검정 당 적어도 10개 또는 적어도 30개의 카피의 평균을 수득하는 것이 바람직하다. 적합한 동적 범위를 갖는 정량적 판독값을 위해, 검정 당 적어도 1,000개의 판독을 수득하는 것이 바람직할 수 있다. 따라서, 1,000,000개의 검정이 수행되는 경우, 서열 판독의 수는 10억 또는 그 이상일 수 있다. 고-처리량 디지털 시퀀싱 및 중복(redundancy)에 대한 허용과 함께, 적어도 10,000개의 엔코딩 프로브의 서열이 동시에 결정되거나, 적어도 100,000, 500,000, 1,000,000, 10,000,000, 100,000,000, 1,000,000,000개 또는 이 이상의 엔코딩 프로브의 서열이 동시에 결정된다.

[0059] **검정**

[0060] 본 발명의 검정 시스템의 검정 부분은 다음과 같은 일반적 단계를 포함한다: 프로브 및 엔코딩 작용제를 전달하는 단계로서, 이러한 엔코딩 작용제(일부 구체예에서, 프로브에 예비-커플링됨)가 공지된 공간 패턴에 따라 샘플로 전달되는 단계, 프로브가 샘플 내의 생물학적 표적과 상호작용하거나 반응하도록 하는 단계, 및 프로브 및 엔코딩 작용제가 예비-커플링되지 않은 경우, 엔코딩 작용제를 프로브에 커플링시키는 단계.

[0061] 본 발명의 샘플은 지지체에 부착되거나, 본질적으로 2-차원 방식으로 제공될 수 있는 사실상 임의의 생물학적 샘플을 포함하며, 여기서 검정된 생물학적 표적 또는 활성을 다시 생물학적 샘플 내의 위치에 고정시키는 능력이 중요하다. 예시적 생물학적 샘플은 조직 섹션(예를 들어, 전체 동물 섹셔닝 및 조직 생검을 포함함), 슬라이드 또는 배양 접시 상의 세포 집단 등을 포함한다. 본 발명의 검정 시스템은 이들이 신선한 샘플, 예를 들어, 일차 조직 섹션, 및 동결된 샘플 및 파라포름알린-고정, 파라핀-포매(paraformalin-fixed, paraffin-embedded, FFPE) 샘플을 포함하나 이에 제한되지는 않는 보존된 샘플을 포함하는 다수의 생물학적 샘플 유형과 상용된다는 점에서 특히 유리하다. 본 발명의 검정 시스템의 중요한 양태는 생물학적 샘플이 별개의 독립적으로 측정가능한 영역을 갖는 기관 표면 상에 고정된다는 점이다.

[0062] 검출되는 생물학적 표적은 단백질, 핵산, 지질, 탄수화물, 이온, 또는 이들 중 임의의 것을 함유하는 다성분 복합체를 포함하나 이에 제한되지는 않는 임의의 생물학적 분자일 수 있다. 세포하 표적의 예는 소기관, 예를 들어, 미토콘드리아, 골지체, 소포체, 엽록체, 세포내이입 소포, 세포외배출 소포, 액포, 리소좀 등을 포함한다.

[0063] 일부 특정 구체예에서, 핵산을 분석하기 위해 유전형 분석, DNA 카피수 또는 RNA 전사체의 정량, 샘플 내의 특정 전사체의 국소화 등에 의한 검정 시스템이 이용된다. 도 3은, 예를 들어, 본 발명의 검정 시스템과 함께 사용될 수 있는 단일 뉴클레오티드 다형체(SNP)를 검출하기 위한 예시적 검정에 대한 전반적 도식을 예시한다. 도 3에서, 2개의 올리고뉴클레오티드 프로브가 제공된다. 각각의 올리고뉴클레오티드 프로브는 (305) 및 (30

7)에서 관찰되는 표적-특이적 영역(분석되는 SNP의 어느 한 면에 위치됨), 및 (301) 및 (303)에서 관찰되는 라이게이션 영역을 포함한다. 올리고뉴클레오티드 프로브는 생물학적 샘플 내의 표적 핵산(제시되지 않음)에 하이브리드화된다. 단계(302)에서, 올리고뉴클레오티드 프로브 중 하나가 신장되어 SNP 서열을 통합시키고, 다른 프로브에 라이게이션되어 표적 핵산 영역(309) 및 라이게이션 영역(301 및 303)을 포함하는 신장된 프로브를 형성시킨다.

[0064] 둘 모두 코딩 태그(315 및 317에서 관찰됨), 라이게이션 영역(311 및 313에서 관찰됨), 및 프라이머 영역(319 및 321에서 관찰됨)을 포함하는 2개의 엔코딩 작용제가 조합되고, 단계(304)에서 신장된 프로브로 라이게이션되어 엔코딩된 표적-특이적 올리고뉴클레오티드를 형성시킨다. 다시, 도 1과 대조적으로, 프로브 및 엔코딩 작용제는 별개의 단계에서 전달된다. 이렇게 하는 것은 하기에 기재되는 조합 구체예의 이용을 가능케 한다. 바람직한 구체예에서, 엔코딩 올리고뉴클레오티드의 쌍 내의 엔코딩 올리고뉴클레오티드는 표적 서열의 한면 또는 다른면(즉, 표적 서열의 5' 또는 3')으로 특이적으로 라이게이션된다. 또한, 통상적으로, 엔코딩 올리고뉴클레오티드 및 프로브의 라이게이션 및 프라이머 영역은 보편적이며, 즉, 프로브 및 엔코딩 올리고뉴클레오티드를 작제하는데 사용되는 라이게이션 및 프라이머 영역의 세트는 일정하고, 프로브 및 엔코딩 올리고뉴클레오티드의 코딩 태그의 표적 특이적 영역만이 상이하다. 그러나, 다시 대안적 구체예에서, 라이게이션 및 프라이머 영역은 보편적이지 않고, 프로브와 엔코딩 작용제 사이에서 상이하다.

[0065] 라이게이션 후, 엔코딩된 프로브가 용리되고, 풀링되고, 임의로, 시퀀싱 어댑터가 PCR을 통해 엔코딩된 프로브에 첨가된다. 대안적 구체예에서, 시퀀싱 프라이머가 엔코딩 올리고뉴클레오티드에 라이게이션될 수 있거나, 시퀀싱 프라이머 서열이 엔코딩 올리고뉴클레오티드의 일부로서 포함될 수 있다. 도 3에서 관찰되는 바와 같이, 각각의 시퀀싱 어댑터는 엔코딩된 프로브 상의 프라이머 영역(319 및 321)과 상용되는 프라이머 영역(319 또는 321)을 포함한다. 제 1 어댑터(327), 제 1 프라이머 영역(319), 제 1 코딩 태그(315), 라이게이션 영역(311 및 301), 표적 영역(309), 라이게이션 영역(313 및 303), 제 2 코딩 태그(317), 제 2 프라이머 영역(325) 및 제 2 어댑터(329)를 포함하는 최종 작제물이 이제 디지털 고-처리량 시퀀싱 과정으로 투입 준비된다.

[0066] 신장 및 라이게이션 반응의 조합이 도 3에 예시되어 있으나, 라이게이션만을 포함하는, 엔코딩 올리고뉴클레오티드를 표적-특이적 올리고뉴클레오티드(예를 들어, 표적 핵산 서열의 연속 부분에 하이브리드화되는 올리고뉴클레오티드에 대해)에 커플링시키기 위해 다양한 반응이 이용될 수 있음이 인지되어야 한다. 대안적으로, GOLDENGATE® 검정에서와 같이 추가적인 올리고뉴클레오티드를 이용하는 검정(예를 들어, Fan, et al., Cold Spring Symp. Quant. Biol., 68:69-78 (2003); (Illumina, Inc., San Diego, CA))이 이용될 수 있다.

[0067] 다른 구체예에서, 본 발명의 검정 시스템은 또한 생물학적 샘플 내의 펩티드 또는 단백질, 항체의 존재, 효소 및 다른 단백질 활성, 번역후 변형, 펩티드의 활성 및 비활성 형태, 뿐만 아니라 펩티드 이소형(isoform)을 분석하는데 사용될 수 있다. 따라서, 프로브는 효소의 활성 영역, 번역글로불린의 결합 도메인, 단백질의 규정된 도메인, 전체 단백질, 합성 펩티드, 도입된 돌연변이를 갖는 펩티드, 앵타머 등을 포함할 수 있다.

[0068] 특정 양태에서, 프로브는 키나제, 포스파타제, 효소원, 프로테아제, 또는 이의 단편과 같은 효소 또는 효소전구체에 대한 기질이다. 특정 양태에서, 프로브는 키나제 또는 포스파타제와 같은 하나 이상의 신호전달 경로와 관련된 단백질을 검출하는데 사용되는 인산화 기질이다. 본 발명의 또 다른 특정 양태에서, 프로브는 개별적 프로테아제 또는 프로테아제의 부류만 결합하는 특이적 프로테아제 기질이다. 다른 양태에서, 프로브는 효소의 다양한 가공된 형태, 이소형 및/또는 도메인이다. 단백질-기반 프로브는 통상적으로 올리고뉴클레오티드 엔코딩 작용제에 컨주게이션되거나 달리 연결된다. 이러한 경우의 올리고뉴클레오티드 엔코딩 작용제는 또한 단백질 프로브의 확인을 가능케 하는 뉴클레오티드 서열 성분을 포함할 것이다.

[0069] 특정 양태에서, 본 발명은 샘플 내 및/또는 샘플 사이의 상이한 위치 사이의 생물학적 표적의 양 및/또는 활성에서의 차이를 평가하기 위한 검정을 제공한다. 상기 방법은 생물학적 샘플로부터 다수의 엔코딩된 결과를 결정하고, 생물학적 샘플 내의 각각의 위치의 생물학적 표적의 양에서의 차이를 평가하는 것을 포함한다.

[0070] **조합 구체예**

[0071] 엔코딩의 효율을 최대화시키기 위해, 엔코딩 올리고뉴클레오티드 내의 코딩 태그의 쌍을 이용하는 조합 방법이 이용될 수 있다. 표적-특이적 정보 및 코딩 태그를 분리시킴으로써, 필요한 올리고뉴클레오티드의 수는 급격히 감소되고, 부수적으로 비용이 감소된다.

[0072] 도 4는 본 발명의 검정 시스템의 조합 엔코딩 도식의 한 구체예에 대한 일반적 메커니즘을 예시하며, 여기서 대 표적 조직 섹션 내의 핵산(416에 제시됨)이 검정된다. 도 4의 A는 관심 표적 핵산(402)에 특이적으로 결합된 2

개의 표적-특이적/엔코딩 올리고뉴클레오티드 작제물(420 및 422)(예를 들어, 도 3의 단계(302)와 단계(304) 사이에서 형성됨)을 도시한다. 제 1의 엔코딩된 프로브(420)는, 예를 들어, 검정 생성물의 증폭을 위한 유니버설(universal) 프라이밍 부위와 결합된 코딩 태그(408) 또는 시퀀싱 기술(404)을 이용한 코딩 인식자의 확인을 가능케 하는 어댑터를 포함한다. 제 2의 엔코딩된 프로브(422)는, 예를 들어, 검정 생성물의 증폭을 위한 유니버설 프라이밍 부위와 결합된 코딩 태그(406) 또는 시퀀싱 기술(410)을 이용한 코딩 인식자의 확인을 가능케 하는 어댑터를 포함한다.

[0073] 도 4의 B는 a1으로부터 a10(엔코딩된 프로브(420) 상의 코딩 태그(406)) 및 b1으로부터 b10(코딩 태그(408) 엔코딩된 프로브(422))의 20개의 상이한 코딩 태그에 대해 사용될 수 있는 공간 패턴을 도시한다. 예를 들어, 코딩 태그 a1은 10개의 별개의 영역 또는 스폿에 생물학적 샘플 상에 배치된다((412)에서 스폿의 제 1 수평선으로 제시됨). 코딩 태그 a2는 (412)의 제 2 수평선 상의 10개의 스폿에 생물학적 샘플 상에 배치된다. 코딩 태그 a3은 (412)의 제 3 수평선 상의 10개의 스폿에 생물학적 샘플 상에 배치되고, 나머지도 마찬가지이다. "a" 태그는 수평 열에 배치되는 반면, "b" 태그는 (414)에 도시된 바와 같은 10개의 수직 열에 배치된다. 예를 들어, 코딩 태그 b1은 (414)의 제 1 수직 열의 10개의 별개의 스폿에 생물학적 샘플 상에 배치되고, 코딩 태그 b2는 (414)의 제 2 수직 열의 10개의 별개의 스폿에 생물학적 샘플 상에 배치되며, 나머지도 마찬가지이다. 상기 형태를 이용하는 것은 20개의 코딩 태그가 생물학적 샘플 상의 100개의 상이한 위치를 독특하게 규정하도록 한다.

[0074] 도 4의 C는 코딩 태그 격자(418)과 일치하는 대표적 조직 섹션(416)을 도시한다. 화살표는 "a" 코딩 태그 및 "b" 코딩 태그가 조직 섹션(416)과 일치하는 격자(418) 상에 배치되는 방법을 표시한다. 시퀀싱 후, 예를 들어, 코딩 태그 a1 및 b4가 표적 핵산 서열과 결합되는 경우, 표적 핵산 서열(즉, 생물학적 표적)은 위치 a1, b4의 조직 섹션에 존재하였다.

[0075] 도 5는 본 발명의 검정 시스템의 엔코딩 도식의 간소화된 특정 예를 제공한다. 도 5는 a1, a2, a3, a4 및 b1, b2, b3 및 b4를 포함하는 엔코딩 올리고뉴클레오티드(510)를 도시한다. 표적 특이적 올리고뉴클레오티드(TSO)(프로브) 1 및 2가 (520)에 제시된다. 배치 또는 분배 도식이 (530)에 제시된다. 도 4에서 예시된 격자와 같이, 엔코딩 올리고뉴클레오티드 a1 내지 a4가 패턴 내(여기서, 수직 패턴 내)의 스폿에 배치되고, 엔코딩 올리고뉴클레오티드 b1 및 b4가 패턴 내(여기서, 수평 패턴)의 스폿에 배치된다. 스폿을 갖는 정사각형으로 제시된 격자는 실제로 도 4에 제시된 조직 섹션(416)과 같은 생물학적 샘플(제시되지 않음) 상의 배치 패턴이다.

[0076] 표적 특이적 올리고뉴클레오티드가 생물학적 샘플에 전달되고, 여기서 표적 핵산이 존재하는 경우 표적 특이적 올리고뉴클레오티드는 생물학적 샘플 내의 표적 핵산에 하이브리드화된다. 하이브리드화되지 않은 표적 특이적 올리고뉴클레오티드는 이후에, 예를 들어, 세척에 의해 제거된다. 엔코딩 올리고뉴클레오티드는 이후 (530)에 제시된 공간 패턴에 따라 생물학적 샘플로 전달된다. 엔코딩 올리고뉴클레오티드는 생물학적 샘플 내의 표적 핵산에 하이브리드화되는 임의의 표적 특이적 올리고뉴클레오티드에 라이게이션(또는, 예를 들어, 신장 및 라이게이션)되고, 라이게이션된 작제물은 이후 생물학적 샘플로부터 용리되고, 푸올링되고, 시퀀싱 어댑터 서열이 엔코딩 올리고뉴클레오티드 내에 미리 포함되지 않은 경우 시퀀싱 어댑터가, 예를 들어, PCR 또는 라이게이션을 통해 첨가된다. 라이게이션된 작제물은, 예를 들어, 고 처리량 또는 "차세대" 시퀀싱에 의해 시퀀싱된다.

[0077] 생성된 서열의 푸올(pool)이 (540)에 제시된다. a4b1, a4b2, a1b3, a2b3, a3b3, a4b3 및 a4b4(위치는 수평선과 함께 제시됨)에서 표적 특이적 올리고뉴클레오티드 1에 대해서만 서열 판독값이 획득되었다. a1b1(위치는 수직선과 함께 제시됨)에서 표적 특이적 올리고뉴클레오티드 2에 대해서만 서열 판독값이 획득되었다. 위치 a2b1, a3b1, a1b2, a2b2, 및 a3b2(위치는 크로스-해칭(cross-hatching)과 함께 제시됨)에서 표적 특이적 올리고뉴클레오티드 1 및 2 둘 모두에 대해 서열 판독값이 획득되었다. a1b4, a2b4 또는 a3b4(위치는 음영이 없이 제시됨)에서 표적 특이적 올리고뉴클레오티드에 대해 서열 판독값이 획득되지 않았다. 따라서, 검정이 발생된 생물학적 샘플에서, 제 1 표적 핵산이 생물학적 샘플의 좌측면의 많은 부분 및 하부에서 검출되었고, 제 2 표적 핵산이 생물학적 샘플의 좌측 상부에서만 검출되었고, 생물학적 샘플의 우측 상부에서는 표적 핵산이 검출되지 않았다. 2개의 표적 핵산의 차별적 발현은 이제 생물학적 샘플 및 생물학적 샘플 내의 상기 위치의 생물학적 구조 또는 세포 유형으로 다시 맵핑될 수 있다.

[0078] 위치 정보에 더하여, 엔코딩된 태그의 상대적 풍부함과 관련된 정보가 획득될 수 있다. 예를 들어, 데이터 세트에서 발생하는 a4T1b1 서열이 a4T1b2 서열에 비해 10배나 많이 존재하는 것으로 발견되는 경우, 이는 표적 핵산 서열 1이 a4T1b2 위치보다 a4T1b1 위치에서 10배 더 풍부한 것을 나타낼 것이다.

[0079] 도 3에 도시된 바와 같은 뉴클레오티드 분석의 경우, 표적 특이적 올리고뉴클레오티드로의 직접적인 코딩 태그

의 라이게이션에 의해, n개의 표적에 대해 단지 2n개의 표적 특이적 올리고뉴클레오티드가 필요하다. 예를 들어, 도 2에 기재된 조합 방법을 이용하여, 10,000개의 공간 위치에서 100개의 상이한 표적을 검정하는 것은 2 x 100개의 표적 특이적 올리고뉴클레오티드 및 2 x 100개의 엔코딩 올리고뉴클레오티드를 필요로 할 것이다. 검정 올리고뉴클레오티드의 전체 수는 유니버설 프라이머를 계수하지 않고 단지 400개(200개의 표적 특이적 및 200개의 엔코딩)일 것이다. 대조적으로, 코딩 올리고뉴클레오티드가 표적 특이적 올리고뉴클레오티드로부터 분리되지 않는 경우, 유니버설 프라이머 서열을 계수하지 않고 (n x X 위치 코드) + (n x Y 위치 코드)가 필요하다거나, 상기 예에서 20,000개의 올리고뉴클레오티드가 필요할 것이다. 더욱이, 비록 도 2-5에 도시된 구체예는 2개의 엔코딩 작용제(코딩 태그)를 이용한 조합 도식을 도시하지만, 3개, 4개 또는 이 이상의 엔코딩 작용제 및 코딩 태그가 이용될 수 있고, 다양한 수단 및 다양한 단계의 조합에 의해 프로브에 부착되거나 서로 부착될 수 있다.

[0080] 본 발명의 검정 시스템의 공간 엔코딩 양태로 인해, 적당한 수의 검정에서도 많은 양의 정보가 생성될 수 있다. 예를 들어, 샘플 내의 5개 이상의 위치에서 검정된 5개 이상의 생물학적 표적은 25개 이상의 조합을 발생시킨다. 관독값으로서 디지털 시퀀싱을 이용하여, 조합 당 서열 관독의 최적 수는 필요한 민감성 및 동적 범위에 좌우되며, 이는 조정될 수 있다. 예를 들어, 각각의 조합에 대해 평균 100개의 관독이 샘플링되는 경우, 25개 조합에 대한 전체는 25,000개의 관독이다. 1,000개의 표적이 1,000의 평균 샘플링 깊이로 1,000개의 위치에서 검정되는 경우, 10⁹개의 관독이 필요하다. 이러한 수는 비록 크지만 적당한 기간 동안 관독 당 매우 낮은 비용으로 수십억 또는 수조의 데이터셋을 발생시킬 수 있는 본질적으로 동시의 디지털 시퀀싱 방법의 능력 범위 내이다. 따라서, 질의되는 위치의 수 또는 검정되는 생물학적 표적, 또는 질의되는 위치의 수 및 검정되는 생물학적 표적 둘 모두를 다양화시키고, 디지털 시퀀싱을 이용함으로써, 많은 양의 정보가 수득될 수 있다. 특정 양태에서, 다수의 위치가 2개 이상의 생물학적 분자에 대해 질의된다.

[0081] **시약 전달 시스템**

[0082] 본 발명의 시약 전달 시스템은 생물학적 샘플의 별개의 부분으로의 시약의 전달을 가능케 하여, 엔코딩 도식의 공간 패턴의 온전성을 유지시키는 장비를 포함한다. 본 발명의 검정 시스템의 시약 전달 시스템은 선택적 이미징(imaging) 수단, 시약 전달 하드웨어 및 조절 소프트웨어를 포함한다. 시약 전달은 다수의 상이한 방식으로 달성될 수 있다. 시약 전달은 한번에 많은 상이한 생물학적 샘플로 이루어질 수 있음을 주의해야 한다. 단일 조직 섹션이 본원에서 예시되었으나, 다수의 생물학적 샘플이 부착될 수 있고, 동시에 분석될 수 있다. 예를 들어, 조직 샘플의 연속 섹션이 동시에 분석될 수 있고, 데이터가 조합되어 3D 맵이 만들어질 수 있다.

[0083] 본 발명의 검정 시스템의 필수적인 것은 생물학적 샘플로의 시약의 공간 패턴화를 가능케 하는 장비이다. 생물학적 분자(예를 들어, 올리고뉴클레오티드 또는 항체) 및 화학 시약(예를 들어, 소분자 또는 dNTP) 둘 모두를 포플레이팅시키고 전달하는 기술은 당 분야에 공지되어 있고, 상기 장비 시스템의 사용은 당업자에게 공지되어 있으며, 본 발명의 검정 시스템에 용이하게 적용가능하다. 적합한 시약 전달 시스템의 한 예는 높은 정밀도 및 재현성으로 나노리터 규모의 생물학적 분자를 함유하는 비말을 전달하는데 사용될 수 있는 Labcyte™ Echo 어쿠스틱 액체 핸들러(acoustic liquid handler)이다. 당업자는 시약이 전달되어야 하는 위치를 특정하는 소프트웨어를 이용하여 상기 시약 전달 장치를 전체 시스템에 통합시킬 수 있다.

[0084] 생물학적 샘플로의 작용제 및/또는 코딩 인식자의 배치에 사용될 수 있는 다른 수단은 잉크 젯 스폿팅(ink jet spotting); 핀, 펜 또는 모세관에 의한 기계적 스폿팅; 미세 접촉 프린팅(micro contact printing); 광화학 또는 광식각(lithographic) 방법 등을 포함하나, 이에 제한되지는 않는다. 여러 적용을 위해, 생물학적 샘플의 특정 영역을 상이한 시약 분배 및/또는 생물학적 표적 결정을 위한 하나 이상의 검정 영역으로 분할하거나 격리시키는 것이 바람직할 수 있다. 검정 영역은 장벽 또는 채널을 이용하여 물리적으로 분리될 수 있다.

[0085] 한 예시적 양태에서, 시약 전달 시스템은 유동-기반 시스템일 수 있다. 본 발명에서 시약 전달을 위한 유동-기반 시스템은 하나 이상의 펌프, 밸브, 유체 저장소, 채널, 및/또는 시약 저장 셀과 같은 장비를 포함할 수 있다. 시약 전달 시스템은 유체를 이동시켜 생물학적 샘플의 별개의 섹션과 접촉하도록 배열된다. 시약의 이동은, 예를 들어, 유체 시약의 다운스트림에 배치된 펌프에 의해 유도될 수 있다. 펌프는 각각의 유체 시약을 반응 구획으로 유도(및 이를 통과)시킬 수 있다. 대안적으로, 시약은 중력에 의해 유체를 통해 유도될 수 있다. 미국 공개공보 번호 20070166725호 및 20050239192호에는 비교적 간단한 하드웨어를 이용하여 매우 복잡한 분석 조작을 달성하기 위해 가스, 액체 및 고체의 정밀한 조작을 가능케 하는, 본 발명의 검정 시스템과 함께 사용될 수 있는 특정한 범용 유체공학 수단이 개시되어 있다.

[0086] 더욱 특정한 예에서, 하나 이상의 유동-셀이 상기로부터의 기질-부착된 생물학적 샘플에 부착될 수 있다. 유동

-셀은 이에 연결된 입구 및 출구 튜브를 포함할 수 있고, 임의로 생물학적 샘플을 가로질러 유동-셀로 시약을 전달하기 위해 외부 펌프가 사용된다. 유동 셀은 생물학적 샘플의 특정 부분에만 시약을 전달하여 생물학적 샘플의 임의의 특정 섹션으로 전달되는 시약의 양 및 유형을 제한하도록 배열된다.

[0087] 또 다른 양태에서, 미세유체 시스템이 생물학적 샘플이 배치되는 기관에 통합되거나, 기관의 상부에 외부적으로 부착될 수 있다. 유체를 유지하고 이동시키기 위한 미세유체 통로가 기관에 인접한 유체공학 층에 의해 평면 기관 및/또는 기관 위에 형성될 수 있다. 유체 시약은 시약 저장소 사이에 배치된 밸브의 선택적 개방 및 폐쇄에 따라 선택되고 전달될 수 있다.

[0088] 펌프는 일반적으로 유체 및/또는 유체에 배치된 시약을 이동시키기 위한 임의의 메커니즘을 포함한다. 일부 예에서, 펌프는 작은 부피를 갖는 통로(즉, 미세유체 구조)를 통해 유체 및/또는 시약을 이동시키도록 배열될 수 있다. 펌프는 다른 수단 중에서 다른 수단 중에서 유체 및/또는 유체를 갖는 구조에 대한 양성 또는 음성 압력 적용, 전기적으로 전기장(들)의 적절한 적용, 또는 압력 적용 및 전기장(들)의 적용 둘 모두에 의해 기계적으로 작동할 수 있다. 예시적인 기계 펌프는 주사기 펌프, 연동 펌프, 회전 펌프, 가압 가스, 파이페터(pipettors) 등을 포함할 수 있다. 기계적 펌프는 마이크로기계화되거나 구조될 수 있다. 예시적 전기 펌프는 전극을 포함할 수 있고, 이는 전기영동, 전기삼투압(electroosmosis), 전기모세관(electrocapillarity), 유전영동(dielectrophoresis)(이의 진행과 형태를 포함함) 등에 의해 작동될 수 있다.

[0089] 밸브는 일반적으로 채널을 통한 유체의 통과를 조절하기 위한 임의의 메커니즘을 포함한다. 밸브는, 예를 들어, 채널에 부분적 또는 완전히 밀접하도록 선택적으로 변형될 수 있는 변형가능한 부재(member), 채널을 부분적 또는 완전히 차단하도록 채널로 선택적으로 연장될 수 있는 이동가능한 돌출부(projection), 전기모세관 구조 등을 포함할 수 있다.

[0090] 개방 개스킷이 생물학적 샘플의 상부에 부착될 수 있고, 샘플 및 시약이 개스킷으로 주입될 수 있다. 적합한 개스킷 물질은 네오프렌, 니트릴, 및 실리콘 고무를 포함하나, 이에 제한되지는 않는다. 대안적으로, 방수 반응 챔버는 기관 상의 생물학적 샘플과, 비제한적인 예로, 블랙-애노다이징 알루미늄(black-anodized aluminum), 열가소성 물질(예를 들어, 폴리스티렌, 포틸카르보네이트 등), 유리 등과 같은 화학적으로 비활성인 내수 물질(water resistant material) 사이에 삽입된 개스킷에 의해 형성될 수 있다.

[0091] 한 임의의 구체예에서, 검정 시스템은 관심 생물학적 샘플의 특징 및 구성을 결정하기 위한 이미징 수단을 포함한다. 수득된 이미지는, 예를 들어, 시약의 배치 패턴을 설계하는데 사용될 수 있다. 이미징 수단은 선택적인데, 이는 개인이, 예를 들어, 현미경을 이용하여 생물학적 샘플을 대신 관찰할 수 있고, 생물학적 샘플의 구성을 분석할 수 있고, 전달 검정 시약에 대한 공간 패턴을 특징할 수 있기 때문이다. 포함되는 경우, 전달 시스템은 참조로서 본원에 포함되는 미국 공개공보 번호 20090197326호에 기재된 것과 같은 이미지, 예를 들어, CCD 또는 IGFET-기반(예를 들어, CMOS-기반) 이미지 및 시약 전달을 위한 초음파 스프레이어를 포함하는 마이크로회로 배열을 포함할 수 있다. 또한, 도 4 및 5는 x,y 격자 형태를 이용하여 예시하지만, 예를 들어, 조직 샘플의 토폴로지(topology)를 따르고, 조직 내의 세포, 세포 층 및/또는 세포 유형의 특정 군을 표적으로 하는 것 등과 같은 다른 형태가 이용될 수 있다.

[0092] 또 다른 대안에서, 시약 전달 시스템은 차폐 및 스프레이와 같은 반도체 기술을 이용하여 생물학적 샘플 표면 상의 특정 패턴으로의 시약의 전달을 조절한다. 생물학적 샘플의 특정 영역은 노출로부터 특정 영역을 보호하는 마스크의 사용을 통해 시약에 대한 노출로부터 보호될 수 있다. 시약은 스프레이 또는 유체 유동과 같은 통상적인 기술을 이용하여 생물학적 샘플로 도입될 수 있다. 차폐된 전달의 이용은 기관 표면 상에 패턴화된 전달 도식을 발생시킨다.

[0093] 본 발명의 바람직한 양태에서, 시약 전달 장치는 잉크젯 프린팅 기술을 기반으로 한다. 매우 다양한 잉크-제팅(ink-jetting) 메커니즘(예를 들어, 열, 압전)이 존재하고, 수성 및 유기 잉크 포물레이션과의 상용성이 밝혀졌다. 동시에 다수의 시약을 전달하기 위해 독립적으로 작동되는 노즐의 세트가 이용될 수 있고, 매우 높은 분리가 달성된다.

[0094] 관심 특정 부위를 표적화하기 위해, 시약 전달 방법 및 관련 인코딩 도식을 돕기 위해 검정되는 생물학적 샘플의 정보 제공 이미지가 이용될 수 있다. 생물학적 샘플의 샘플 영역은 검정 시스템의 다른 특징과 통합된 이미지 처리(예를 들어, 면역조직화학 또는 다른 염색 화학에 의해 구별된 세포 유형의 이미지)를 이용하여 확인될 수 있다. 일부 양태에서, 이미지 정보를 시약 전달 패턴으로 자동적으로 번역하기 위해 소프트웨어가 사용된다. 따라서, 시약 전달을 위한 생물학적 샘플을 매우 정밀하게 기록하고 정렬시키는 메커니즘이 본 발명

의 검정 시스템의 중요한 구성요소이다. 슬라이드 및/또는 다른 매우 정확한 물리적 위치결정 시스템 상에서의 기점(fiducial) 마커의 사용과 같은 메커니즘이 상기 목적을 위해 적합화될 수 있다.

[0095] 본 발명은 바람직하게는 검정 시스템에 맞춤형된 완전한 한벌의 소프트웨어를 포함한다. 임의로, 수행되는 특정 검정을 위한 인코딩 뉴클레오티드(및 핵산이 검정되는 구체예에서는, 표적 특이적 올리고뉴클레오티드)를 설계하는데 올리고뉴클레오티드 설계 소프트웨어가 사용되며, 이는 시스템의 일부로서 통합될 수 있다. 또한, 임의로, 시약 전달 및 데이터 분석(즉, 서열 분석)을 위한 알고리즘 및 소프트웨어가 검정 결과를 결정하기 위해 통합될 수 있다. 생성되는 데이터셋의 유형이 규모의 결과로서 부피가 클 수 있으므로 통합된 데이터 분석이 특히 유용하다. 패턴-분석 소프트웨어 및 시각화 수단을 포함하는 검정 시스템에 의해 발생된 공간 관련 데이터의 분석을 위해 특별히 설계된 알고리즘 및 소프트웨어 도구는 검정 시스템에 의해 발생된 데이터의 값을 향상시킨다.

[0096] 특정 양태에서, 검정 시스템은 시약의 품질 관리, 예를 들어, 올리고뉴클레오티드 풀의 온전성 및 서열 정확도를 이루고 수행하기 위한 과정을 포함한다. 특히, 시약은 휘발성, 중요 온도에서의 안정성, 및 시약 전달 장비와의 상용성을 위한 화학적 상용성과 같은 요인에 따라 제형화되며, 이는 검정 시스템 내에 통합된 장비에 의해 분석될 수 있다.

[0097] 시퀀싱

[0098] 본 발명의 검정 시스템의 인코딩된 프로브에서 코딩 태그 및 프로브 서열을 확인하기 위해 다수의 방법이 이용될 수 있다. 코딩 태그는 질량분석법(예를 들어, MALDI-ToF, LC-MS/MS), 핵 자기 공명 이미징, 또는, 바람직하게는 핵산 시퀀싱과 같은 기술을 이용하여 검출될 수 있다. 본 발명의 코딩 태그를 디코딩하기 위한 기술의 예는, 예를 들어, 참조로서 본원에 포함되는 미국 공개공보 번호 20080220434호에서 발견될 수 있다. 예를 들어, 코딩 태그는 올리고뉴클레오티드 질량 태그(OMT 또는 massTag)일 수 있다. 이러한 태그는, 예를 들어, 전체내용이 참조로서 포함되는 미국 공개공보 번호 20090305237호에 기재되어 있다. 또 다른 대안에서, 인코딩된 프로브는 증폭될 수 있고, 마이크로어레이에 하이브리드화될 수 있다. 이는 각각의 증폭이 코드 특이적 프라이머를 이용하여 달성되는 특정 공간 코드 또는 코드의 서브셋에 대해 특이적인, 별개의 증폭 반응이 수행되는 것을 필요로 한다. 각각의 증폭은 또한 상이한 분리가능한 라벨(예를 들어, 플루오로포어(fluorophor))을 포함할 것이다. 하이브리드화 후, 샘플 내의 상이한 공간 위치에 대한 특정 표적 맵핑의 상대량은 분리가능한 라벨의 상대적 풍부함에 의해 결정될 수 있다.

[0099] 한 특히 바람직한 양태에서, 검정 시스템에 따른 결과로서 발생된 코딩 태그는 고-처리량의 차세대 시퀀싱에 대한 기질이며, 코딩 태그의 서열을 확인하기 위해, 예를 들어, SOLiD™ 기술(Life Technologies, Inc.) 또는 Genome Analyzer (Illumina, Inc.)와 같은 고도로 동시적인 차세대 시퀀싱 방법이 사용된다. 이러한 차세대 시퀀싱 방법은, 예를 들어, 원 패스(one pass) 시퀀싱 방법 또는 페어드-엔드(paired-end) 시퀀싱을 이용하여 수행될 수 있다. 차세대 시퀀싱 방법은, 예를 들어, 드마낙(Drmanac)의 미국 특허 번호 6,864,052호; 6,309,824호; 및 6,401,267호; 및 드마낙 외의 미국 특허 공개공보 2005/0191656호에 개시된 것과 같은 하이브리드화-기반 방법; 합성에 의한 시퀀싱(sequencing-by-synthesis) 방법, 예를 들어, 미국 특허 번호 6,210,891호; 6,828,100호; 6,969,488호; 6,897,023호; 6,833,246호; 6,911,345호; 6,787,308호; 7,297,518호; 7,462,449호 및 7,501,245호; 미국 공개공보 번호 20110059436호; 20040106110호; 20030064398호; 및 20030022207호; 문헌[Ronaghi, et al, Science, 281:363-365 (1998)]; 및 문헌[Li, et al, Proc. Natl. Acad. Sci., 100: 414-419 (2003)]; 라이게이션-기반 방법, 예를 들어, 미국 특허 번호 5,912,148호 및 6,130,073호; 및 미국 특허 출원 번호 20100105052호, 20070207482호 및 20090018024호; 나노포어 시퀀싱, 예를 들어, 미국 특허 출원 번호 20070036511호; 20080032301호; 20080128627호; 20090082212호; 및 문헌[Soni and Meller, Clin Chem 53: 1996-2001 (2007)], 뿐만 아니라 다른 방법, 예를 들어, 미국 특허 출원 번호 20110033854호; 20090264299호; 20090155781호; 및 20090005252호; 또한, 문헌[McKernan, et al., Genome Res., 19: 1527-41 (2009)] 및 [Bentley, et al., Nature 456:53-59 (2008)]에 개시된 것을 포함하나, 이에 제한되지는 않으며, 상기 참고문헌 모두는 모든 목적상 이의 전체내용이 본원에 포함된다.

[0100] 검정 시스템의 적용

[0101] 본 발명의 개시의 정독 후, 생물학적 샘플 내의 생물학적 표적의 양 및 공간적 위치를 동시에 측정할 수 있는 고 처리량의 다중화 검정 시스템으로부터 이득을 얻을 수 있는 생물학적 연구, 진단, 및 약물 개발의 다수의 중요한 영역이 존재하는 것이 당업자에게 명백할 것이다. 예를 들어, 상이한 RNA 전사체의 상대적 풍부함을 측정하는 능력과 조직 내에서 가능한 작거나 심지어 개별적 세포보다 작을 수 있는 많은 위치에 걸쳐 풍부함의 공간

적 패턴의 이미지를 구성하는 능력을 조합시키는 것은 많은 상이한 영역의 기초 연구를 가능케 한다. 하기는 예시적 용도이며, 범위를 이로 제한하는 것을 의미하지는 않는다.

[0102] 일 예에서, CT 스캐닝에서의 이미지 재구성과 유사한 방식으로 일련의 조직 섹션을 분석함으로써 유전자 발현의 3-차원 패턴이 결정된다. 이러한 방법은 질병 병리, 예를 들어, 암성 조직 및/또는 손상, 염증 또는 감염 후 조직에서의 유전자 발현의 변화를 측정하기 위해 이용될 수 있다. 본 발명의 검정 시스템을 이용하여, 복잡한 조직에서의 유전자 발현 및 단백질 국소화에 대한 보다 상세한 정보가 수득되고, 이는 정상 및 질병을 갖는 상태 둘 모두에서의 기능 및 조절에 대한 새로운 통찰을 발생시키고, 시험될 수 있는 새로운 가설을 제공한다. 예를 들어, 본 발명의 검정 시스템은 많은 개별적 연구 및 조직 수준으로 통합되는 ENCODE(Birney, et al., Nature, 447:799-816 (2007)) 및 modENCODE와 같은 보다 큰 프로그램으로부터 통찰 중 일부를 얻는 것을 가능케 할 수 있다. 검정 시스템은 또한 시스템 생물학 분야에서 유전자 발현의 상호작용 네트워크를 모델링하기 위한 계산적 노력을 돕는다.

[0103] 검정 시스템은 또한 암에서의 체세포 돌연변이 또는 감염 유기체에 대한 반응에서의 변동성과 같은 체세포 변화의 분석을 위한 신규한 방법을 제공한다. 예를 들어, 종양은 통상적으로 고도로 이종성이며, 이는 암세포 뿐만 아니라 비정상적인 국소적 환경에서의 유전적으로 정상인 세포를 함유한다. 암세포는 돌연변이 및 선택을 겪으며, 이러한 방법에서, 국소적 클론이 발달하는 것은 이상한 것이 아니다. 종양의 상황에서 비교적 드문 체세포 돌연변이를 확인하는 것은 클론 변이체의 선택에서 중요 돌연변이의 역할의 연구를 가능케 할 수 있다. 암 및 유전적으로 정상인 세포 둘 모두에서 혈관형성, 염증, 또는 다른 암-관련 과정과 관련된 전사 패턴은 암 생물학을 통찰하기 위해 분석될 수 있고, 이는 암의 치료를 위한 새로운 치료제의 개발을 돕는다. 또 다른 예에서, 개체는 감염 유기체에 대해 다양한 감수성을 가지며, 본 발명의 검정 시스템은 미생물과 조직 또는 조직 내의 다양한 세포 유형 사이의 상호작용을 연구하기 위해 이용될 수 있다.

[0104] 중요하게는, 공간적으로 관련된 정보를 제공하는 것에 더하여, 본 발명은 임의의 제공된 반응에서 단지 적은 위치가 검정되어 신호 대 노이즈가 급격하게 증가될 수 있으므로 희귀한 돌연변이를 검출하는 민감성을 크기 증가시킨다. 혼합된 샘플에서의 희귀한 돌연변이에 대한 통상적 검정에서, 샘플은 대량으로 처리되고, 즉, 핵산이 많은 세포로부터 단일한 풀로 추출된다. 따라서, 돌연변이가 10,000개 중 하나의 세포에 존재하는 경우, 이는 약 10,000개의 세포로부터의 정상 DNA의 백그라운드에 대해 검출되어야 한다. 대조적으로, 본 발명의 검정 시스템을 이용하여 많은 세포가 분석될 수 있으나, 개별적 세포 또는 세포의 작은 그룹은 공간 코딩 시스템에 의해 확인될 것이다. 따라서, 본 발명의 검정 시스템에서, 수십배까지 크기 증가된 민감성에 의해 백그라운드 감소된다. 더욱이, 돌연변이 세포의 공간적 구성이 관찰될 수 있고, 이는 암의 조직 섹션에서 중요한 돌연변이를 검출하는데 특히 중요할 수 있다. 이미, 분자 조직학적 분석이 암 생물학에 대한 통찰을 제공하며 있으며, 이는 진단학에 사용하기 위한 잠재성을 가질 수 있다. 본 발명의 기술은 상기 방법의 힘을 크게 증가시킬 가망이 있다.

발명을 실시하기 위한 구체적인 내용

[0105] 실시예

[0106] 하기 실시예는 본 발명을 제조하고 사용하는 방법의 완전한 개시 및 기재를 당업자에게 제공하기 위해 기재하며, 이는 본 발명의 발명자가 그의 발명으로 간주하는 것의 범위를 제한하는 것이 아니거나, 하기 기재된 실험이 전부이거나 수행되는 유일한 실험인 것을 나타내거나 의미하는 것이 아니다. 광범위하게 기재된 바와 같은 본 발명의 사상 또는 범위를 벗어남이 없이 특정 구체예에 제시된 바와 같은 본 발명에 대해 다수의 변화 및/또는 변형이 이루어질 수 있음이 당업자에 의해 인식될 것이다. 따라서, 본 발명의 구체예는 모든 점에서 예시적인 것이며, 제한적인 것이 아닌 것으로 간주되어야 한다.

[0107] 사용되는 수(예를 들어, 양, 온도 등)와 관련하여 정확성을 보장하고자 노력하였으나, 일부 실험 오차 및 편차가 고려되어야 한다. 달리 표시하지 않는 한, 부(parts)는 중량부이고, 분자량은 중량 평균 분자량이고, 온도는 섭씨 온도이고, 압력은 대기압 또는 대기압 근처이다.

[0108] 실시예 1: 엔코딩 도식 개념의 최초 증명

[0109] 개념의 최초 증명으로서, 작업 단일-플렉스(single-plex) 검정을 입증하기 위해 마이크로어레이를 이용하는 모델 시스템을 개발하였다. 기본 설계는 검정의 개념을 확인하였고, 보다 복잡한 생물학적 샘플의 분석과 관련된 문제를 다루기 전에 작업 검정을 확립하였다. 상기 개념의 증명을 위한 관독값으로서 통상적인 시퀀싱을 이용

하였다.

[0110] 조직 섹션을 위한 대용물로서 마이크로어레이를 이용하였다. 마이크로어레이의 표적 서열은 완전히 상술되어서, 표적의 조성은 공지되어 있고, 체계적으로 변화될 수 있다. 합성 올리고뉴클레오티드 주형을 5' 아미노 변형을 통해 유리 슬라이드에 부착시켰다. 각각의 슬라이드는 단일한 올리고뉴클레오티드 주형 서열을 가지며, 수행되는 검정은 라이게이션, 또는 특정 다형태를 결정하는데 유용할 수 있음에 따라 신장 후 라이게이션을 이용할 수 있다.

[0111] 검정의 인 시츄(in situ) 부분이 완료된 후, 반응 생성물을 용리시키고, qPCR에 의해 분석하여 생성물의 존재 또는 부재를 결정하고 수율을 평가하였으며, 통상적인 시퀀싱에 의해 검정 생성물의 구조를 결정하였다. 시험되는 단일-플렉스 검정은 적절한 양성 및 음성 대조군, 및 단일 염기 변화를 구별하는 능력을 확인하기 위한 단일 뉴클레오티드 변형(single nucleotide variant, SNV)을 포함한다.

[0112] **실시예 2: 확장성(Scalability)**

[0113] 고 처리량 연구에서 사용하기 위한 검정의 확장성을 확립하기 위해 검정 시스템의 복잡도를 증가시켰다. 공간적 엔코딩 및 검정 시스템 둘 모두의 확장성을 마이크로어레이 모델 시스템을 이용하여 24-플렉스 x 24-부위 검정을 수행함으로써 입증하였다.

[0114] 각각의 검정 위치에서의 생물학적 표적, 여기서는 DNA 표적 서열의 양은 마이크로어레이 기관 상에서 체계적으로 다양하였다. 예를 들어, 50 마이크로미터의 스폿 크기(중심-중심(center to center))를 갖는 마이크로어레이에서, 1 mm² 영역은 약 400개의 스폿을 함유하였다. 각각의 부위 주위의 영역은 임의로 표적 서열의 개별적 분리성을 가능케 하는 상이 스폿이 결여된 영역에 의해 점유된다. 대안적으로, 스폿은 군집되어 있을 수 있고, 2개 이상의 바로 인접한 스폿은 표적 서열이 결여된 영역에 의해 둘러싸여 있거나 이에 인접하여 있다.

[0115] 공간 엔코딩이 정확한 것을 입증하기 위해, 부위는 검정 판독값이 각각의 부위의 예상 조성과 일치하는 것을 나타내기 위해 상이한 표적 조성을 포함한다. 24개의 표적 서열을 이용하여, 각각의 부위가 바이너리 코드(binary code)(0 = 부재, 1 = 존재)를 만들기 위해 12개의 표적이 존재하는 세트 및 12개의 표적이 부재하는 세트의 상이한 세트를 갖는 간단한 디지털 패턴을 만들었다. 이후, 검출 영역이 공간 디코딩 후에 예상 신호와 일치하는 것을 나타내기 위해 검정 판독값을 결정하였다. 이러한 특정 예에서, 코드 공간은 약간의 오차가 있더라도 혼잡되어 있는 상이한 코드를 발생시키지 않도록 충분히 컸다(2²⁴). 더욱이, 이러한 설계는 오차의 확인을 가능케 하며, 공간 엔코딩의 정확성 평가 뿐만 아니라 표적 서열의 존재 또는 부재를 알려주는 정확성의 평가를 가능케 한다.

[0116] 정량적 차이를 검출하는 능력을 24-부위 검정에서 각각의 부위에서 수행되는 24 검정 각각에 대한 용량-반응 곡선을 발생시킴으로써 평가하였다. 이는 검출 한도, 동적 범위, 및 범위 전체에 걸쳐 제공된 배수-변화를 검출하는 능력의 평가를 가능케 한다.

[0117] 일 양태에서, 각각의 표적에 대한 특징의 수를 변화시킴으로써 상이한 비에서의 개별적 표적을 나타내기 위해 라틴 방진(latin square) 설계를 이용하였다. 즉, 부위에서 다수의 스폿을 이용하여, 24 표적 서열 각각에 할당된 스폿의 수가 변화될 수 있고, 24 부위 각각이 상이한 조성을 가질 수 있다. 1 x 3 인치의 마이크로어레이가 다수의 복제를 허용하기에 충분히 크다. 이러한 24 서열의 보다 큰 세트는 디컨볼루션(deconvolution)을 필요로 할 것이며, 이는 차세대 시퀀싱 기술(예를 들어, SOLiD™ 기술(Life Technologies, Inc., Carlsbad, CA) 또는 Genome Analyzer(Illumina, Inc., San Diego, CA)과 같은 고 처리량 기술을 이용하여 달성된다. 24-플렉스 검정의 이용은 공간 엔코딩 및 디코딩의 정확성, 및 검정 시스템의 정량적 반응 둘 모두를 나타내었다.

[0118] **실시예 3: 보존된 샘플에 대한 검정의 적합화.**

[0119] 단일-카피 참조 신호를 확립하기 위한 방식을 제공함에 따라 RNA를 검정하기 위한 개념의 증명으로서 유전체 DNA를 검정하였다. FFPE 샘플에 대한 작업 검정을 개발한 후, 이를 RNA 검정으로 적합화시켰다. 이를 위해, 높은 다중화와의 상용성을 보장하기 위해 검정 올리고뉴클레오티드 농도를 검정하였다. 10 마이크로미터의 세포 직경, 및 개별적 세포에 대한 10 마이크로미터 직경의 시약 비말 전달을 가정하여, 비말의 부피는 약 500 μl이며, 이는 1 μM의 농도에서 약 3 x 10¹¹개의 분자를 함유할 수 있다. 10,000개의 세포에서 1,000개의 표적 서열을 검정하기 위해, 비말 내에 약 2,000개의 표적 올리고뉴클레오티드가 필요할 것이다. 따라서, 각각의 비말은 세포에서 수천개의 표적 서열에 대해 광범위한 과량인 각 검정 올리고의 약 1억 6천만개의 카피를 함유할 수

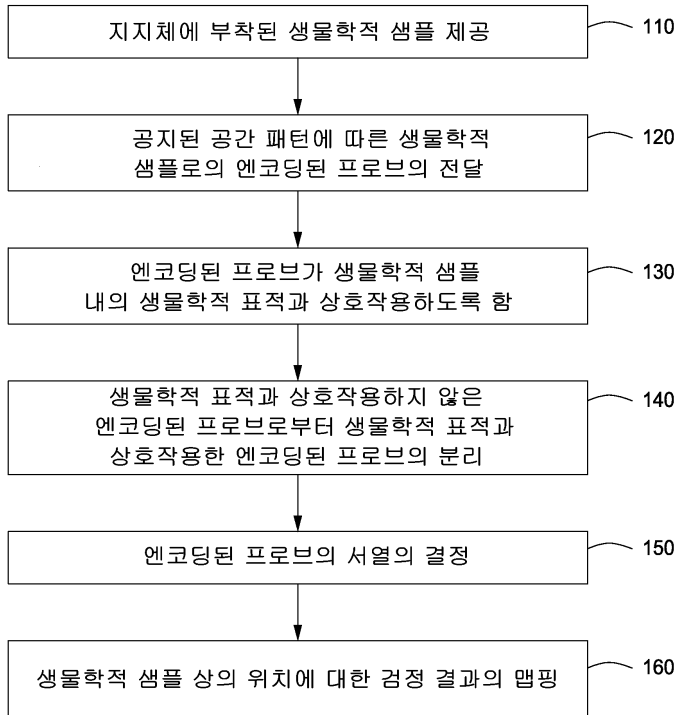
있다.

- [0120] 매우 적거나 손상된 샘플로부터 발생된 생성물 분자의 적은 절대 수의 취급은 낮은 회수 효율의 문제와 조우할 가능성이 높고, 즉, 용리가 효과적이고, 표면에 대한 분자의 흡착으로부터 발생하는 손실이 방지된다. 후자의 문제를 처리하기 위한 방법은 글리코젠 또는 담체 핵산과 같은 담체 물질을 포함시키는 것이다.
- [0121] **실시예 4: 생물학적 샘플에 대한 검정의 적합화.**
- [0122] 인공 시스템을 생성시키기 위해 대조군 RNA 주형을 고체 지지체에 고정시켰다. DNA/RNA 하이브리드에서 닉(nick)을 복구시킬 수 있는 T4 DNA 리가아제를 이용하여 검정을 수행하였다. 검정을 일치된 슬라이드, 또는 동일 슬라이드의 상이한 섹션에서 수행하였고, 여기서 한 경우에 gDNA가 검정되고, 다른 경우에 RNA가 검정된다. gDNA를 검정하는 경우, 슬라이드는 RNase로 전처리될 수 있고, RNA를 검정하는 경우, 슬라이드는 DNase로 전처리될 수 있다. gDNA 또는 RNA를 추출하고, qPCR 또는 RT-qPCR에 의해 각각 상대량을 정량함으로써 검정 결과를 확인하였다.
- [0123] 조직 섹션 RNA 검정을 가능한 한 정보를 제공하도록 만들기 위해, 검정 설계에서 다양한 풍부함에 걸친 표적 전사체에 대한 특정 조직 내의 발현 수준의 미리 존재하는 정보를 이용하였다. 높은 풍부함의 전사체 뿐만 아니라 일부 중간 및 낮은 풍부함의 전사체 둘 모두를 표적으로 하여 검정의 정량적 성능 특성의 최초 평가를 가능케 하였다.
- [0124] 상기 기재는 단지 본 발명의 원리를 예시한다. 당업자는 비록 본원에 명백히 기재되거나 제시되지 않았지만 본 발명의 원리를 구체화하는 다양한 배열을 고안할 수 있을 것이며, 이는 본 발명의 사상 및 범위에 포함되는 것이 인지될 것이다. 또한, 본원에 언급된 모든 예 및 조건적 언어는 주로 독자가 본 발명의 원리 및 본 발명자에 의해 제공된 개념을 더 추가로 당 분야까지 이해하는 것을 돕기 위한 것으로, 상기 특별히 언급된 예 및 조건으로 제한하는 것이 아닌 것으로 간주되어야 한다. 더욱이, 본 발명의 원리, 양태, 및 구체에 뿐만 아니라 이의 특정 예를 언급하는 본원의 모든 언급은 이의 구조적 및 기능적 동등부 둘 모두를 포함한다. 추가로, 상기 동등부는 현재 공지된 동등부 및 미래에 개발될 동등부, 즉 구조와 관계 없이 동일 기능을 수행하는 개발된 임의의 구성요소 둘 모두를 포함한다. 따라서, 본 발명의 범위는 본원에 제시되고 기재된 예시적 구체예로 제한되지 않는다. 오히려, 본 발명의 범위 및 사상은 첨부된 청구항에 의해 구체화된다. 하기 청구항에서, 용어 "수단"이 사용되지 않는 한, 청구항에서 언급된 특징 또는 구성요소 중 어느 것도 35 U.S.C. § 112, ¶6에 따라 수단-기능(means-plus-function) 제한으로 간주되어선 안된다.

도면

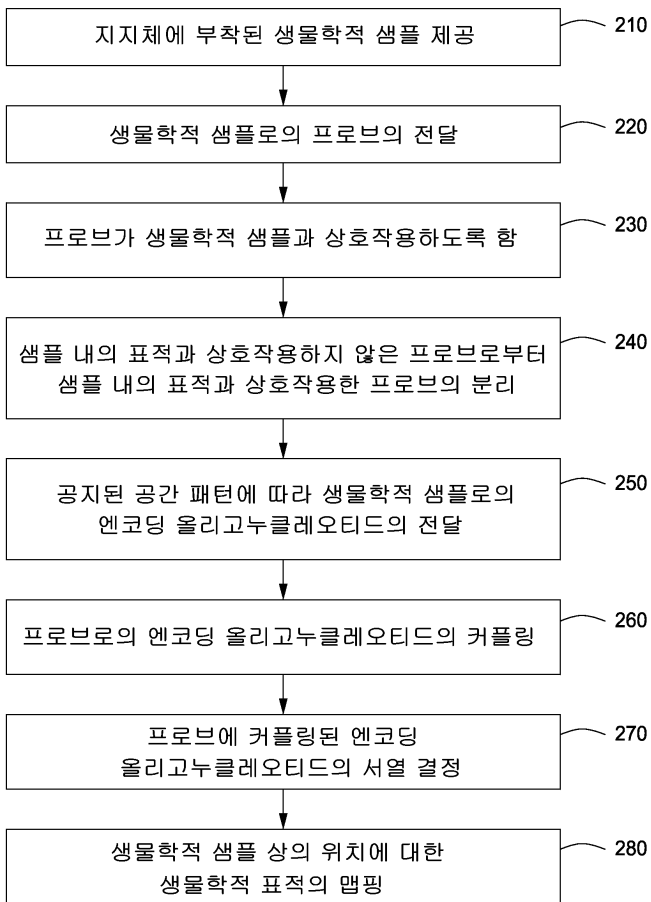
도면1

100

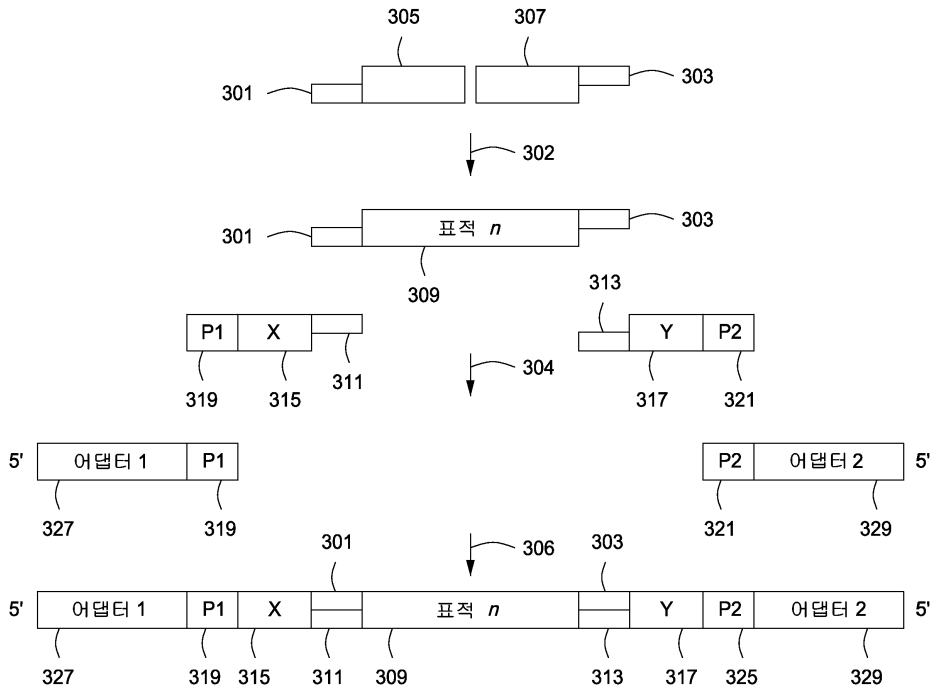


도면2

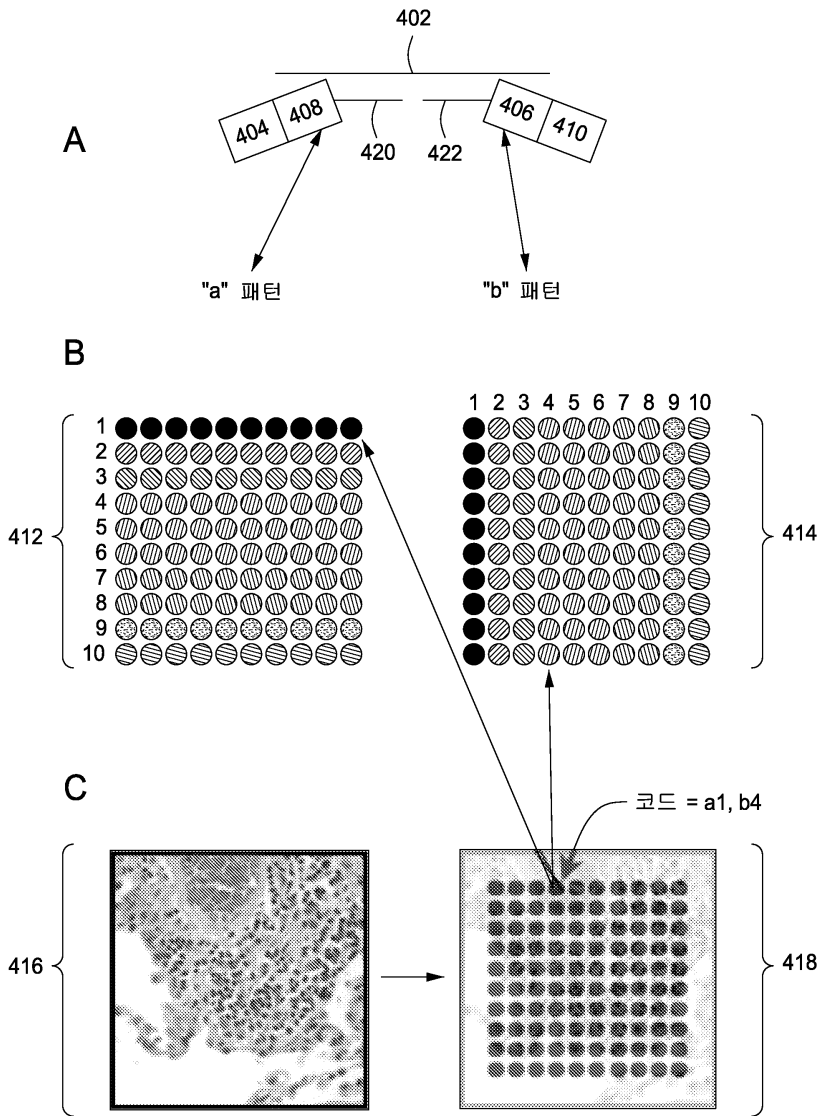
200



도면3



도면4



도면5

