



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년12월01일
(11) 등록번호 10-2333255
(24) 등록일자 2021년11월26일

- (51) 국제특허분류(Int. Cl.)
B01J 19/00 (2018.01) *B01L 3/00* (2006.01)
CA0B 20/04 (2006.01) *CA0B 50/16* (2006.01)
GO1N 33/53 (2006.01) *GO1N 33/58* (2006.01)
GO1N 33/68 (2006.01)
- (52) CPC특허분류
B01J 19/0046 (2013.01)
B01L 3/5085 (2013.01)
- (21) 출원번호 10-2020-7031596(분할)
- (22) 출원일자(국제) 2016년05월09일
 심사청구일자 2020년11월02일
- (85) 번역문제출일자 2020년11월02일
- (65) 공개번호 10-2020-0129164
- (43) 공개일자 2020년11월17일
- (62) 원출원 특허 10-2017-7035427
 원출원일자(국제) 2016년05월09일
 심사청구일자 2017년12월07일
- (86) 국제출원번호 PCT/US2016/031524
- (87) 국제공개번호 WO 2016/183029
 국제공개일자 2016년11월17일
- (30) 우선권주장
 62/159,710 2015년05월11일 미국(US)
- (56) 선행기술조사문헌
 WO2013096777 A2*
 *는 심사관에 의하여 인용된 문헌

- (73) 특허권자
 일루미나, 인코포레이티드
 미국 캘리포니아 92122 샌디에고 일루미나 웨이 5200
- (72) 발명자
 히 몰리
 미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200
 프레바이트 마이클
 미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200
 (뒷면에 계속)
- (74) 대리인
 특허법인아주김장리

전체 청구항 수 : 총 25 항

심사관 : 신창훈

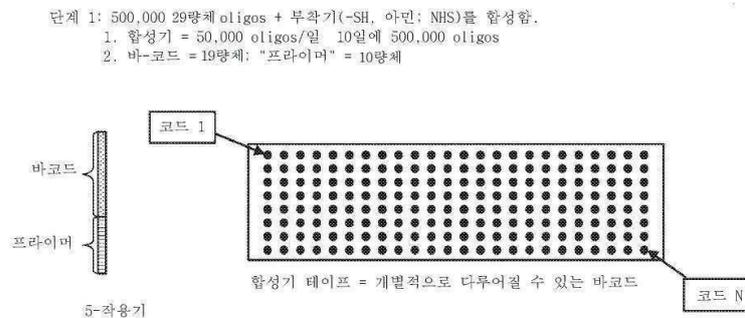
(54) 발명의 명칭 치료제의 발견 및 분석을 위한 플랫폼

(57) 요약

(a) 핵산 태그에 부착된 후보 약제의 라이브러리를 제공하는 단계; (b) 라이브러리를 고체 지지체와 접촉시켜 고체 지지체에 후보 약제를 부착시키고, 이에 의해, 후보 약제의 어레이가 형성되는 단계; (c) 어레이를 스크리닝제와 접촉시키는 단계로서, 여기서, 어레이에서의 하나 이상의 후보 약제가 스크리닝제와 반응하는 단계; (d) 어

(뒷면에 계속)

대표도 - 도1a



레이에서의 적어도 하나의 후보 약제가 스크리닝제와 반응하는지를 결정하기 위해 어레이를 검출하는 단계; (e) 핵산 태그를 시퀀싱하여 어레이에서의 후보 약제에 부착된 태그 서열을 결정하는 단계; 및 (f) 적어도 하나의 후보 약제에 부착된 태그 서열을 기초로 하여 스크리닝제와 반응하는 어레이에서의 적어도 하나의 후보 약제를 식별하는 단계를 포함하는 후보 약제를 특성 규명하는, 방법.

(52) CPC특허분류

C40B 20/04 (2013.01)

C40B 50/16 (2013.01)

G01N 33/53 (2018.05)

G01N 33/58 (2020.05)

G01N 33/68 (2013.01)

B01J 2219/005 (2013.01)

B01J 2219/00626 (2013.01)

B01J 2219/00743 (2013.01)

(72) 발명자

고린스키 미샤

미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200

켈링거 매튜 윌리엄

미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200

파이사조비치 세르지오

미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200

부텔 조나단 마크

미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200

명세서

청구범위

청구항 1

어레이로서,

(a) mRNA 분자의 라이브러리로서, 상기 라이브러리에서의 개별 mRNA 분자가 표적 서열 및 태그 서열을 포함하는, 상기 라이브러리, 및

(b) 상기 태그 서열의 보체를 갖는 핵산을 포함하는 고체 지지체를 포함하되,

상기 핵산이 상기 고체 지지체 상에서의 개별 피처에 부착되며,

상기 개별 mRNA 분자의 상기 태그 서열이 상기 고체 지지체 상에서의 상기 개별 피처에서 개개 상보적인 태그 서열에 하이브리드화되며,

상기 mRNA 분자의 번역에 의해 유도된 단백질이 개개 mRNA 분자에 부착되는, 어레이.

청구항 2

제1항에 있어서, 상기 단백질이 리보솜을 통해 상기 개개 mRNA 분자에 부착된 어레이.

청구항 3

제2항에 있어서, 상기 단백질이 상기 리보솜에 공유적으로 부착되어 있는, 어레이.

청구항 4

제1항에 있어서, 상기 단백질이 발광 표지되는, 어레이.

청구항 5

제4항에 있어서, 발광 표지된 스크리닝제가 단백질의 결합 부위에 특이적으로 결합되며, 이에 의해 상기 단백질이 발광 표지된 어레이.

청구항 6

제1항에 있어서, 상기 고체 지지체가 적어도 1×10^6 개의 상기 피처를 포함하는, 어레이.

청구항 7

제1항에 있어서, 상기 고체 지지체 상에서의 상기 피처의 평균 피치가 10 마이크로론 미만인, 어레이.

청구항 8

제1항에 있어서, 상기 피처가 100 제곱 마이크로론 미만의 평균 면적을 포함하는, 어레이.

청구항 9

제1항에 있어서, 상기 mRNA 분자의 라이브러리가 한 유전자의 복수의 변형체를 포함하는, 어레이.

청구항 10

제1항에 있어서, 상기 단백질이 항체, 효소, 수용체, 키나아제, 포스파타제, 폴리머라제, 프로테아제, 에스테라제, 히스톤 개질 효소 및 핵 호르몬 수용체로 이루어진 군으로부터 선택되는, 어레이.

청구항 11

제1항에 있어서, 상기 고체 지지체가 흐름셀 내에 위치되어 있는, 어레이.

청구항 12

단백질의 어레이를 생산하는 방법으로서,

- (a) mRNA 분자의 라이브러리를 제공하는 단계로서, 상기 라이브러리에서의 개별 mRNA 분자는 표적 서열 및 태그 서열을 포함하는, 상기 라이브러리를 제공하는 단계;
- (b) 상기 라이브러리로부터 제1 서브-라이브러리를 유도하는 단계로서, 상기 제1 서브-라이브러리는 상기 태그 서열 또는 이의 보체를 갖는 핵산을 포함하고, 상기 핵산은 고체 지지체 상의 개별 피처에 부착되는, 상기 제1 서브-라이브러리를 유도하는 단계;
- (c) 상기 라이브러리로부터 제2 서브-라이브러리를 유도하는 단계로서, 상기 제2 서브-라이브러리는 상기 표적 서열 및 태그 서열 또는 이의 보체를 갖는 핵산을 포함하는, 상기 제2 서브-라이브러리를 유도하는 단계;
- (d) 상기 제2 서브-라이브러리를 상기 제1 서브-라이브러리와 접촉시켜, 상기 태그 서열 및 이의 상기 보체의 하이브리드화를 통해 상기 고체 지지체에 상기 제2 서브-라이브러리의 핵산을 부착시키는 단계; 및
- (e) 상기 고체 지지체 상에서 상기 표적 서열을 번역하여 상기 개별 피처에 부착된 단백질의 어레이를 생산하는 단계를 포함하는, 방법.

청구항 13

제12항에 있어서, 상기 고체 지지체 상에서 태그 서열 또는 이의 보체를 시퀀싱하여, 상기 고체 지지체 상의 개별 피처에서 상기 태그 서열 또는 이의 보체의 위치를 결정하는 단계를 더 포함하는, 방법.

청구항 14

제12항 또는 제13항에 있어서, 상기 제1 서브-라이브러리의 핵산이 상기 표적 서열을 더 포함하는, 방법.

청구항 15

제14항에 있어서, 상기 고체 지지체 상에서 상기 표적 서열을 시퀀싱하는 단계를 더 포함하는, 방법.

청구항 16

제12항에 있어서, 상기 제1 서브-라이브러리가 상기 라이브러리의 mRNA 분자를 상기 고체 지지체와 접촉시켜 상기 고체 지지체에 상기 mRNA 분자를 부착시키는 것을 포함하는 방법에 의해 유도되는, 방법.

청구항 17

제16항에 있어서, 상기 mRNA 분자가 상기 태그 서열의 상기 보체를 생산하기 위해 상기 고체 지지체 상에서 증폭되는, 방법.

청구항 18

제12항에 있어서, 상기 제1 서브-라이브러리가, 상기 태그 서열을 포함하는 상기 개별 mRNA 분자 또는 상기 개별 mRNA 분자의 일부를 역전사하는 것을 포함하는 방법에 의해 유도되는, 방법.

청구항 19

제18항에 있어서, 상기 역전사된 mRNA 분자, 또는 이의 부분이 상기 태그 서열의 보체를 생산하기 위해 상기 고체 지지체 상에서 증폭되는, 방법.

청구항 20

제12항에 있어서, 상기 제1 서브-라이브러리가 상기 태그 서열의 보체를 갖는 핵산을 포함하며, 상기 제2 서브-라이브러리가 상기 표적 서열 및 상기 태그 서열을 갖는 RNA분자를 포함하며, 단계 (d)가 상기 제2 서브-라이브러리를 상기 제1 서브-라이브러리와 접촉시켜 상기 태그 서열 및 이의 보체의 하이브리드화를 통해 상기 고체 지지체에 상기 제2 서브-라이브러리의 mRNA 분자를 부착시키는 단계를 포함하는, 방법.

청구항 21

제12항에 있어서, 상기 제1 서브-라이브러리가 상기 태그 서열을 갖는 핵산을 포함하며,
 상기 제2 서브-라이브러리가 상기 표적 서열 및 상기 태그 서열의 보체를 갖는 cDNA 분자를 포함하며,
 단계 (d)가 상기 제2 서브-라이브러리를 상기 제1 서브-라이브러리와 접촉시켜, 상기 태그 서열 및 이의 보체의 하이브리드화를 통해 상기 고체 지지체에 상기 제2 서브-라이브러리의 cDNA 분자를 부착시키는 단계를 포함하며,
 단계 (e)가 상기 cDNA 분자를 역전사하여 상기 고체 지지체 상에 mRNA 분자를 생산하고, 상기 고체 지지체 상에서 상기 mRNA 분자를 번역하여 상기 개별 피처에 부착된 단백질 어레이를 생산하는 단계를 포함하는, 방법.

청구항 22

제12항에 있어서, 상기 표적 서열이 상기 고체 지지체 상에서 리보솜과 함께 번역되며, 리보솜이 상기 개별 피처에 부착된 단백질의 어레이를 생산하기 위해 푸로마이신으로 처리되는, 방법.

청구항 23

단백질을 스크리닝하는 방법으로서,
 (i) 제12항에 따른 단백질의 어레이를 생산하는 단계;
 (ii) 상기 단백질의 어레이를 스크리닝제와 접촉시키는 단계로서, 상기 어레이에서의 하나 이상의 단백질은 상기 스크리닝제와 반응하는, 상기 접촉시키는 단계; 및
 (iii) 상기 스크리닝제와 접촉시키는 동안 또는 후에 상기 단백질의 어레이를 검출하여, 상기 어레이에서의 적어도 하나의 단백질이 상기 스크리닝제와 반응하는지를 결정하는 단계를 포함하는, 방법.

청구항 24

제23항에 있어서, 상기 고체 지지체 상에 상기 태그 서열 또는 이의 보체를 시퀀싱하여, 상기 고체 지지체 상의 상기 개별 피처에서 상기 태그 서열 또는 이의 보체의 위치를 결정하는 단계를 더 포함하는, 방법.

청구항 25

제24항에 있어서, 상기 적어도 하나의 단백질에 부착된 상기 태그 서열을 기초로 하여 스크리닝제와 반응하는 상기 어레이에서의 상기 적어도 하나의 단백질을 식별하는 단계를 더 포함하는, 방법.

발명의 설명

기술 분야

[0001] 본 발명은 일반적으로, 의학, 농업, 및 산업적 용도를 갖는 약제(agent)(예를 들어, 소분자, 단백질 또는 세포)의 발견, 및 보다 상세하게는, 이러한 용도를 위한 후보 약제를 스크리닝하기 위한 플랫폼(platform)에 관한 것이다.

배경 기술

[0002] 현재, 고처리량 스크리닝 약물 발견은 특정 표적에 대한 효능 및 효력뿐만 아니라 분자 표적의 패널에 대한 바람직한 독성 프로파일을 나타내는 "리드(lead)" 또는 "히트(hit)"에 대한 소분자의 큰 라이브러리(library)를 스크리닝하기 위한 여러 단계 및 플랫폼을 사용한다.

[0003] 본 방법에서 제1 단계에서, 소분자 라이브러리는 사전-결정된 표적에 대해 고처리량 방식으로 스크리닝된다. 이러한 라이브러리는 통상적으로 수 십만개의 화합물로 구성된다. 현재의 고처리량 방법은 자동화될 수 있다. 로봇학 및 자동화의 진보의 결과로서, 고가의 기기는 낮은 마이크로리터 범위의 부피의 입력 시약을 사용하여 100,000개의 분자/일을 스크리닝할 수 있다. 검정은 동종(homogenous)이거나 이종(heterogeneous)일 수 있으며, 전자는 비교적 간단하고 저렴하며, 후자는 더욱 복잡하고 시간 소비적이고 고가이지만, 더욱 민감하다.

- [0004] 초기 소분자 스크리닝 검정이 "히트" 또는 "리드"를 식별하는데 효과적이지만, 이러한 검정은 통상적으로 단지 제1 단계이다. 이러한 리드 또는 히트는 통상적으로, 불리한 임상 효과를 가질 수 있는 후보물질을 제거하기 위해 잘 규정된 표적에 대한 프로파일링 검정을 이용하여 유전독성, 약리학적 독성, 세포 독성, 및 유기체 세포독성에 대해 추가로 스크리닝된다. 이러한 검정 각각은 독립적으로, 히트 또는 리드의 유용성에 대한 컨센서스 (consensus)를 발생시키기 위해 권장되고 데이터 마이닝되는 독립적인 워크플로우(workflow)를 이용하는 다양한 플랫폼 또는 키트화된 검정(kitted assay)을 이용하여 수행된다.
- [0005] 단백질 진화 방법은 다른 타입의 스크린 t를 구성한다. 이러한 방법은 광범위한 순열을 다루었다. 이러한 방법의 처리량, 속도 및 낮은 비용이 빠른 단백질 진화에 대한 선도적인 선택으로 만드는 경향이 있지만, 이러한 방법은 종종 복잡하고 어려움을 겪는다. 예를 들어, 에멀전 기반 스크린은 액적의 출현/혼합(멀티플렉스 스크리닝의 경우)의 문제, 에멀전 파괴의 문제(고려되는 성분을 회수), 다양한 액적 크기로 인한 농도 편차로부터 발생하는 문제, 및 액적의 교차 오염을 일으킨다.
- [0006] 세포-기반 치료법은 소분자-기반 또는 단백질-기반 치료법에 비해 잠재적으로 유리한 치료 능력을 지닌다. 중요한 장점 중에는, 세포가 외부 신호를 감지하고, 신체 내의 특정 부위로 이동하고, 여러 자극을 통합하고, 복잡한 거동(예를 들어, 특정 이펙터 분자의 방출)으로 반응한다는 것이다. 세포-기반 치료법의 가능성의 완전한 달성은 치료 세포를 정확하게 조작하는 것으로부터 이익을 얻을 것이며, 이에 따라, 이의 "거동"은 공간 및 시간 둘 모두에 있어서 제어될 수 있다.
- [0007] 세포 공학에서 현 워크플로우는 통상적으로 하기 단계를 포함한다: (i) 자극에 대한 감지, 통합 및 반응을 담당하는 세포내 신호전달 회로의 설계; (ii) 세포에서 그러한 기능을 매개하기 위해 담당하는 유전자를 도입하기 위한 복잡한 유전 공학; 및 (iii) 라이브러리에서의 모든 세포 중에서, 요망되는 기능을 더 잘 수행하는 유전자 보체를 지니는 것을 식별하기 위한 스크리닝 방법.요망되는 세포 거동의 복잡성으로 인하여, 현 스크리닝 방법은 이상적이지 않다. 예를 들어, 형광 활성화된 세포 분류(FACS)-기반 방법은 높은 처리량을 가지지만, 단지 세포 거동의 스냅샷(snapshot)을 조사한다. 다른 한편으로, 형광 현미경은 세포 거동의 동역학을 상세히 추적할 수 있지만, 처리량은 낮다.
- [0008] 이에 따라, 유익한 성질을 위한 소분자, 단백질, 세포 및 다른 체제를 스크리닝하기 위한 플랫폼 및 방법에 대한 필요성이 존재한다. 본 발명은 이러한 필요성을 다루고, 또한 다른 장점을 제공한다.

발명의 내용

- [0009] 본 발명은 후보 약제를 특성 규명하는 방법을 제공한다. 본 방법은 (a) 후보 약제의 라이브러리를 제공하는 단계로서, 각 후보 약제는 태그 서열을 갖는 핵산 태그에 부착되는 단계; (b) 후보 약제의 라이브러리를 고체 지지체와 접촉시켜 고체 지지체에 후보 약제를 부착시키고, 이에 의해 후보 약제의 어레이는 각각이 라이브러리로부터 개별 후보 약제에 부착하는 고체 지지체 상의 개별 피쳐(feature)를 포함하는 후보 약제의 어레이가 형성되는 단계; (c) 후보 약제의 어레이를 스크리닝제와 접촉시키는 단계로서, 어레이에서의 하나 이상의 후보 약제가 스크리닝제와 반응하는 단계; (d) 어레이를 스크리닝제와 접촉하는 동안 또는 후에 어레이를 검출하여, 어레이에서의 적어도 하나의 후보 약제가 스크리닝제와 반응하는지를 결정하는 단계; (e) 어레이 상에서 핵산 태그를 시퀀싱하여 후보 약제 각각에 부착되는 태그 서열을 결정하는 단계; 및 (f) 적어도 하나의 후보 약제에 부착되는 태그 서열을 기반으로 하여 스크리닝제와 반응하는 어레이에서의 적어도 하나의 후보 약제를 식별하는 단계를 포함할 수 있다.
- [0010] 본 발명은 또한, 단백질의 어레이를 생산하는 방법을 제공한다. 본 방법은 (a) 고체 지지체에 부착된 cDNA 분자의 라이브러리를 제공하는 단계; (b) 고체 지지체 상에서 cDNA 분자를 증폭시켜 클러스터를 형성시키는 단계로서, 각 클러스터는 라이브러리로부터의 특정 cDNA 분자의 다수의 복사체를 포함하는 단계; (c) 클러스터에서 다수의 복사체를 전사하여 클러스터 각각에 부착된 다수의 mRNA 분자를 생산하는 단계; 및 (d) 클러스터에서 mRNA 분자를 번역하여 클러스터 각각에 부착된 다수의 단백질을 생산하는 단계를 포함할 수 있다.
- [0011] 본 발명은 또한, (a) mRNA 분자의 라이브러리를 제공하는 단계로서, 라이브러리에서 개별 mRNA 분자는 표적 서열 및 태그 서열을 포함하는 단계, (b) 라이브러리로부터 제1 서브-라이브러리를 유도하는 단계로서, 제1 서브-라이브러리는 태그 서열 또는 이의 보체를 갖는 핵산을 포함하며, 핵산은 고체 지지체 상에서 개별 피쳐에 부착되는 단계, (c) 라이브러리로부터 제2 서브-라이브러리를 유도하는 단계로서, 제2 서브-라이브러리는 표적 서열 및 태그 서열 또는 이의 보체를 갖는 핵산을 포함하는 단계, (d) 제2 서브-라이브러리를 제1 서브-라이브러리와 접촉시켜, 제2 서브-라이브러리의 핵산을 태그 서열 및 이의 보체의 하이브리드화를 통해 고체 지지체에 부착시

키는 단계, 및 (e) 고체 지지체 상에서 표적 서열을 번역하여 개별 피처에 부착된 단백질의 어레이를 생산하는 단계를 포함하는, 단백질의 어레이를 생산하는 방법을 제공한다.

[0012] 또한, 본 발명에 의해 세포를 스크리닝하는 방법이 제공된다. 본 방법은 (a) 복수의 상이한 세포를 제공하는 단계로서, 상이한 세포 각각은 태그 서열을 갖는 핵산 태그를 포함하는 단계; (b) 상이한 세포의 혼합물을 고체 지지체와 접촉시켜 고체 지지체 상에 부착된 다수의 세포를 형성시키는 단계; (c) 적어도 하나의 광학적 특징에 대하여, 고체 지지체 상의 세포의 어레이를 스크리닝하는 단계로서, 스크리닝 반응은 고체 지지체에 부착된 개별 세포를 검출하는 것을 포함하는 단계; (d) 고체 지지체에 부착된 핵산 태그의 태그 서열을 시퀀싱하는 단계; 및 (e) 어레이에서 적어도 하나의 세포를 광학적 특징 및 후보 세포의 태그 서열을 기초로 하여 후보 세포로서 식별하는 단계를 포함할 수 있다.

[0013] 본 발명은 (a) 고체 지지체; (b) 고체 지지체에 부착된 상이한 cDNA 분자의 라이브러리로서, 각 상이한 cDNA 분자는 고체 지지체 상의 개별 피처에 부착되며 각 피처는 특정 cDNA 분자의 다수의 복사체를 포함하는, 라이브러리; (c) cDNA 분자에 부착된 mRNA 분자로서, cDNA 분자 각각은 개개의 부착된 mRNA 분자에 대해 상보적인, mRNA 분자; 및 (d) mRNA 분자에 부착된 단백질 분자로서, 단백질 분자 각각은 개개의 부착된 mRNA 분자에 의해 엔코딩되는, 단백질을 포함하는 어레이를 제공한다.

[0014] 본 발명은 또한, (a) mRNA 분자의 라이브러리로서, 라이브러리에서 개별 mRNA 분자는 표적 서열 및 태그 서열을 포함하는 라이브러리, (b) 태그 서열의 보체를 갖는 핵산을 포함하는 고체 지지체를 포함하는 어레이를 제공하며, 여기서, 핵산은 고체 지지체 상의 개별 피처에 부착되며, 개별 mRNA 분자의 태그 서열은 고체 지지체 상의 개별 피처에서 개개 상보적 태그 서열에 하이브리드화되며, mRNA 분자의 번역에 의해 유도된 단백질은 개개 mRNA 분자에 부착된다.

도면의 간단한 설명

[0015] 도 1a 내지 도 1f는 후보 약제의 고처리량 스크리닝을 위한 공정에서의 단계를 도시한 것이다.
 도 2는 후보 단백질 약제의 고처리량 합성 및 스크리닝을 위한 공정을 도시한 것이다.
 도 3a 및 도 3b는 후보 단백질 약제의 고처리량 태그화, 합성 및 스크리닝을 위한 공정을 도시한 것이다.
 도 4는 흐름셀에 부착된 세포에 스크리닝제의 노출, 및 스크리닝제에 대한 다양한 세포 반응으로부터 예상되는 시간에 대한 형광의 플롯을 도시한 것이다.
 도 5는 개별 마이크로웰에 분류되고 이후에, 마이크로웰-특이적 태그로 태그화된 세포를 도시한 것이다.
 도 6은 코딩된 비드를 사용하여 고체 지지체 상의 세포 위치의 디코딩(decoding)을 도시한 것이다.
 도 7은 핵산 태그를 사용하여 고체 지지체 상의 세포 위치의 디코딩을 도시한 것이다.
 도 8은 세포에서 고체 지지체 상의 부위로 핵산 태그의 이동 및 고체 지지체 상의 핵산 태그의 시퀀싱을 도시한 것이다.
 도 9는 세포막에서 지방산에 한 쌍의 핵산의 부착에 의해 태그화된 세포를 도시한 것이다.
 도 10은 흐름셀 표면 상의 핵산 태그화된 세포의 포획을 도시한 것이다.
 도 11은 핵산 태그의 제한 엔도뉴클레아제 분해를 통한 흐름셀 표면으로부터의 세포의 탈착을 도시한 것이다.
 도 12는 세포가 수득되는 용기에서 디코딩된 태그의 위치를 기초로 하여 흐름셀 표면 상의 세포의 식별을 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명은 치료 용도를 위한 후보물질과 같은 후보 약제의 고처리량 스크리닝을 위한 장치 및 방법을 제공한다. 본 명세서에 기술된 특정 구현에는 핵산 시퀀싱 기술 및 장치를 유리하게 사용한다. 본 명세서에 기술된 바와 같은 시퀀싱 기술 및 장치를 사용하는 장점은, 각 후보 약제와 관련된 핵산이 구별될 수 있고 특정 자극에 대한 각 후보 약제의 반응이 고체 지지체 상에서 개별적으로 검출될 수 있도록 다수의 상이한 후보 약제를 공간적으로 배열하는 능력이다.

[0017] 특정 구현예에서, 복수의 상이한 후보 약제는 라이브러리에 제공되며, 여기서, 각 후보 약제는 독특한 핵산 태

그와 함께 부착된다(또는 달리 결합된다). 후보 약제는 핵산, 단백질, 세포 및 소분자를 포함하지만, 이들로 제한되지 않는 임의의 다양한 항목일 수 있다. 이러한 태그화된 후보 약제는, 개별 일원이 다수의 공간적으로 분리된 후보 약제를 형성하도록, 고체 지지체에 부착될 수 있다. 이후에, 후보 약제의 어레이는 스크리닝제(또는 다른 자극)에 대한 노출에 의해 스크리닝될 수 있으며, 후보 약제의 반응은 고체 지지체 상의 공간적으로 분해된 위치에서 검출될 수 있다. 하나 이상의 "히트"의 위치는 어레이 상에 예상되거나, 요망되거나, 독특한 신호의 공간적으로 분해된 검출을 기초로 하여 식별될 수 있다. 시퀀싱 반응은 또한, 어레이에서 이의 위치에 대해 각 태그를 식별하기 위해 (스크린 단계 전 또는 후에) 어레이 상에서 수행될 수 있다. 후보 약제는 "히트"의 위치를 그러한 위치에서의 태그의 동일성과 연관시킴으로써 식별될 수 있다.

[0018] 본 명세서에 기술된 방법 또는 장치에서 사용되는 일부 후보 약제는 예를 들어, 단백질을 포함하는, 핵산-기반이다. 단백질을 사용하는 여러 구현예에서, 단백질을 엔코딩하는 DNA 또는 RNA 분자의 서열은 하나의 단백질을 다른 단백질과 구별하기 위해 결정될 수 있다. 또한, 개별 단백질의 아미노산 서열은 공지된 유전자 코드를 기초로 하여 RNA 서열로부터 추론될 수 있다. 그러나, 일부 구현예에서, 단백질, 또는 단백질을 엔코딩하는 핵산은 시퀀싱될 필요가 없다. 오히려, 각 단백질은 단백질의 서열과 선형적으로 연관된 태그와 부착(또는 달리 결합)될 수 있다. 이에 따라, 태그의 시퀀싱은 하나의 단백질을 다른 단백질과 구별하기에 충분할 수 있다. 유사하게, 세포-기반 후보 약제는 집단에서 개별 세포를 식별하기 위해 시퀀싱될 수 있는 핵산을 함유한다. 또한, 핵산 서열은 개별 세포의 유용한 특징을 결정하기 위해 평가될 수 있다. 또한, 사용 또는 선형적으로 지정된 태그는 세포의 다른 함유물을 시퀀싱할 필요 없이 세포 특징을 구별할 수 있다.

[0019] 본 발명의 방법 및 장치가 본 명세서에서 치료 기능에 대한 후보 약제를 스크리닝 하는 것과 관련하여 예시되었지만, 다른 기능적 또는 구조적 특징이 스크리닝될 수 있는 것으로 이해될 것이다. 예를 들어, 본 방법 및 장치는 독성, 농업 용도(예를 들어, 살충제, 성장 인자, 호르몬, 등), 산업적 용도(예를 들어, 촉매, 염료, 플라스틱, 등), 영양학(예를 들어, 풍미제, 보존제, 등), 환경 정화, 등에 대해 스크리닝하기 위해 사용될 수 있다. 일반적으로, 본 방법 및 장치는 생물학적 기능 또는 비-생물학적 기능에 대해 스크리닝하기 위해 사용될 수 있다.

[0020] 본 명세서에 기술된 방법 및 장치는 고처리량 스크리닝 동안 정량적 측정을 수행하는 능력에 있어서 장점을 제공한다. 예를 들어, 상업적으로 입수 가능한 시퀀싱 플랫폼, 예를 들어, Illumina, Inc.(캘리포니아주 샌디에이고 소재)에 의해 상업화된 시퀀싱 플랫폼은 형광 신호를 정량화하기 위한 비교적 넓은 동적 범위를 갖는 정밀한 옵틱(precision optics)을 포함한다. 다른 장점은 후보 약제의 어레이에서 다수의 위치에서 스크리닝 반응의 일시적 동역학(temporal dynamics)을 따르는 능력이다. 상반되게, 여러 전통적인 스크린에서 사용되는 유체 분류 방법은 단지 통과하는 후보 약제의 스냅샷을 제공하고, 이에 의해 측정을 단일 시점으로 제한한다. 본 명세서에 기술된 본 방법 및 장치는 유체 분류 기술과 동등하거나 일부 경우에, 이보다 더 양호한 고처리량을 제공하지만, 스크리닝 결과의 시간 기반 측정의 추가적인 이점을 제공한다. 본 명세서에 기술된 장치 및 방법의 추가 장점은 스크린 결과를 검출하고, 스크린으로부터 "히트"를 식별하기 위해 스크린 결과와 태그의 공간적 상관성을 허용하는 방식으로 태그를 검출하는 능력이다.

[0021] 본 명세서에서 사용되는 용어는 달리 특정하지 않는 한, 관련 분야에서의 이의 일반적인 의미를 갖는 것으로 이해될 것이다. 본 명세서에서 사용되는 여러 용어 및 이들의 의미는 하기에 기술된다.

[0022] 본 명세서에서 이용되는 용어 "앰플리콘"은 핵산에 대해 이용되는 경우, 핵산 복제 산물을 의미하며, 여기서 산물은 핵산의 적어도 일부 뉴클레오타이드 서열과 동일하거나 상보적인 뉴클레오타이드 서열을 갖는다. 앰플리콘은, 예를 들어 폴리머라제 확장, 폴리머라제 연쇄 반응(PCR), 롤링 서클 증폭(RCA), 다중 변위 증폭(MDA), 결찰 확장, 또는 결찰 연쇄 반응을 포함하는, 주형으로서 핵산 또는 이들의 앰플리콘을 이용하는 다양한 임의의 증폭 방법에 의해 생성될 수 있다. 앰플리콘은 특정 뉴클레오타이드 서열의 단일 복사체(예를 들어, PCR 산물) 또는 뉴클레오타이드 서열의 다중 사본(예로 RCA의 콘카타머 산물)을 갖는 핵산 분자일 수 있다. 표적 핵산의 첫 번째 앰플리콘은 전형적으로 상보적 복사체이다. 후속 앰플리콘은 첫 번째 앰플리콘의 생성 후 표적 핵산으로부터 또는 앰플리콘으로부터 생성되는 복사체이다. 후속 앰플리콘은 표적 핵산에 실질적으로 상보적이거나 표적 핵산과 실질적으로 동일한 서열을 가질 수 있다.

[0023] 본 명세서에서 이용되는 용어 "어레이"는 상대적인 위치에 따라 서로 구분될 수 있는 피쳐 또는 부위의 집단을 지칭한다. 어레이의 상이한 부위에 있는 상이한 분자는 어레이에서 부위의 위치에 따라 서로 구분될 수 있다. 어레이의 개별 부위는 특정 타입의 하나 이상의 분자(또는 다른 항목)를 포함할 수 있다. 예를 들어, 부위에는 특정 서열을 갖는 단일 표적 핵산 분자가 포함될 수 있거나 부위에는 동일한 서열(및/또는 이들의 상보적 서

열)을 갖는 몇몇 핵산 분자가 포함될 수 있다. 어레이의 부위는 동일한 기재 상에 배치된 상이한 특징부일 수 있다. 예시적 피처는 비제한적으로 기재 내의 웰, 기내 내 또는 상의 비드(또는 다른 입자), 기재로부터의 돌출부, 기재 상의 용기 또는 기재 내의 채널을 포함한다. 어레이의 부위는 각각 상이한 분자(또는 다른 항목)를 보유하는 별도의 기재일 수 있다. 별도의 기재에 부착된 상이한 분자는 기재가 연합되는 표면 상 기재의 위치에 따라 또는 액체 또는 겔 중 기재의 위치에 따라 확인될 수 있다. 별도의 기재가 표면 상에 배치된 예시적 어레이에는 비제한적으로 웰 내 비드를 갖는 것들이 포함된다.

[0024] 본 명세서에서 사용되는 용어 "부착된(attached)"은 두 개의 사물이 서로 연결되거나, 고정되거나, 접촉되거나, 연결되거나, 결합되어 있는 상태를 지칭한다. 예를 들어, 피분석물, 예를 들어, 핵산은 공유 또는 비-공유 결합에 의해 겔 또는 고체 지지체와 같은 물질에 부착될 수 있다. 공유 결합은 원자들 간에 전자 쌍을 공유함으로써 특성화된다. 비-공유 결합은 전자 쌍의 공유를 수반하지 않고 예를 들어, 수소 결합, 이온 결합, 반 데르 발스력, 친수성 상호작용 및 소수성 상호작용을 포함할 수 있는 화학적 결합이다. 일부 구현예에서, 부착은 포획제를 통해 일어날 수 있다. 포획제는 예를 들어, 항체, 수용체, 핵산, 리간드, 렉틴, 탄수화물, 아비딘, 바이오틴, 또는 이의 유사체를 포함할 수 있다.

[0025] 본 명세서에서 사용되는 용어 "후보 약제"는 특정 구조 또는 기능을 갖는 것으로 의심되는 항목을 의미하도록 의도된다. 예시적인 항목은 분자, 세포 및 세포내 성분(subcellular component)을 포함하지만, 이들로 제한되지 않는다. 분자는 선택적으로, 생물학적 활성 분자, 예를 들어, 단백질, 아미노산, 핵산(예를 들어, DNA 또는 RNA), 뉴클레오타이드, 다당류, 사카라이드, 대사산물, 비타민, 효소 보조인자, 등일 수 있다. 다른 후보 약제는 마크로사이클(macrocyclic), 사이클릭 펩티드, 융합 분자(예를 들어, 핵산-단백질 방식) 또는 디스플레이된 작제물(예를 들어, 표지 상의 펩티드)을 포함한다. 후보 약제가 가질 것으로 의심되는 예시적인 기능은 다른 제제의 활성화, 다른 제제의 억제, 다른 제제의 화학적 개질, 다른 제제의 분해, 다른 제제의 합성을 포함하지만, 이들로 제한되지 않으며, 여기서, 다른 제제는 선택적으로, 후보 약제인 것으로서 상기에 예시된 항목들 중 임의의 하나 이상일 수 있다. 후보 약제의 구조는 상기 항목 또는 당해 분야에 공지된 다른 항목에 대해 임의의 공지되거나 의심된 구조일 수 있다.

[0026] 핵산을 참조로 하여 사용될 때, 본 명세서에서 사용되는 용어 "상이한"은, 핵산이 서로 동일하지 않은 뉴클레오타이드 서열을 가짐을 의미한다. 둘 이상의 핵산은 이들의 전체 길이에 걸쳐 상이한 뉴클레오타이드 서열을 가질 수 있다. 대안적으로 둘 이상의 핵산은 이들 길이의 상당 부분에 걸쳐 상이한 뉴클레오타이드 서열을 가질 수 있다. 예를 들어, 둘 이상의 핵산은 서로 상이한 표적 뉴클레오타이드 서열 부분을 갖는 반면, 또한 서로 동일한 공통 서열 영역을 가질 수 있다. 이러한 용어는 아미노산 서열 차이를 기초로 하여 서로 상이한 것으로서 구별될 수 있는 단백질에 유사하게 적용될 수 있다.

[0027] 본 명세서에서 이용되는 용어 "각각"은 항목의 집합에 대해 이용되는 경우, 집합에서의 개별 항목을 나타내기 위한 것이지만 문맥 상 명백히 달리 나타나지 않는 한, 반드시 집합에서의 모든 항목을 나타내는 것은 아니다. 명시적 기술 또는 문맥이 달리 명확하게 기술하지 않는 경우에, 예외가 발생할 수 있다.

[0028] 핵산과 관련하여 사용될 때 본 명세서에서 사용되는 용어 "확장하다"는 핵산에 적어도 하나의 뉴클레오타이드 또는 올리고뉴클레오타이드의 첨가를 의미하는 것으로 의도된다. 특정 구현예에서, 하나 이상의 뉴클레오타이드는 예를 들어, 폴리머라제 촉매작용(예를 들어, DNA 폴리머라제, RNA 폴리머라제 또는 역전사 효소)을 통해 핵산의 3' 단부에 첨가될 수 있다. 핵산의 3' 단부 또는 5' 단부에 하나 이상의 뉴클레오타이드를 첨가하기 위해 화학적 방법 또는 효소 방법이 사용될 수 있다. 하나 이상의 올리고뉴클레오타이드는 예를 들어, 화학적 방법 또는 효소(예를 들어, 리가제 촉매작용) 방법을 통해, 핵산의 3' 단부 또는 5' 단부에 첨가될 수 있다. 핵산은 주형 유도 방식으로 확장될 수 있으며, 이에 의해, 확장 산물은 확장된 핵산에 하이브리드화되는 주형 핵산에 대해 상보적이다.

[0029] 본 명세서에서 사용되는 용어 "피처(feature)"는 특정 종의 분자 또는 세포를 위한 어레이에서의 위치를 의미한다. 피처는 단지 단일 분자(또는 세포)를 함유할 수 있거나, 이는 동일한 종의 여러 분자(또는 세포)의 집단을 함유할 수 있다. 일부 구현예에서, 피처는 분자 또는 세포를 부착하기 전에 고체 지지체 상에 존재한다. 다른 구현예에서, 피처는 고체 지지체에 분자 또는 세포의 부착에 의해 생성된다. 어레이의 피처는 통상적으로 별개이다. 별개의 피처는 인접할 수 있거나, 이러한 것은 서로 이격되어 있다. 피처의 크기 및/또는 피처 간의 간격은, 어레이가 고밀도, 중간 밀도, 또는 저밀도일 수 있게 다양할 수 있다. 고밀도 어레이는 약 15 μm 미만으로 분리된 부위를 갖는 것으로서 특성화된다. 중간 밀도의 어레이는 약 15 내지 30 μm 로 분리된 부위를 가지며, 저밀도 어레이는 30 μm 초과로 분리된 부위를 갖는다. 본 명세서에서 유용한 어레이는 예를 들어, 100 μm , 50 μm , 10

μm , $5\mu\text{m}$, $1\mu\text{m}$, 또는 $0.5\mu\text{m}$ 미만으로 분리된 부위를 가질 수 있다. 본 발명의 장치 또는 방법은 상기 밀도 또는 밀도 범위에서 부위를 구별하기에 충분한 해상도로 어레이를 검출하기 위해 사용될 수 있다.

[0030] 본 명세서에서 사용되는 용어 "흐름셀(flow cell)"은 반응이 수행될 수 있는 챔버, 챔버에 시약을 전달하기 위한 유입구, 및 챔버로부터 시약을 제거하기 위한 유출구를 갖는 용기를 의미하도록 의도된다. 일부 구현예에서, 챔버는 챔버에서 일어나는 반응의 검출을 위해 구성된다. 예를 들어, 챔버는 챔버에서 어레이, 광학적으로 표지된 분자, 등의 광학적 검출을 가능하게 하는 하나 이상의 투명한 표면을 포함할 수 있다. 예시적인 흐름셀은 핵산 시퀀싱 장치에서 사용되는 것, 예를 들어, Illumina, Inc.(캘리포니아주 샌디에이고 소재)에 의해 상업화된 Genome Analyzer[®], MiSeq[®], NextSeq[®] 또는 HiSeq[®] 플랫폼을 위한 흐름셀; 또는 Life Technologies(Carlsbad, CA)에 의해 상업화된 SOLiD[™] 또는 Ion Torrent[™] 시퀀싱 플랫폼을 위한 흐름셀을 포함하지만, 이들로 제한되지 않는다. 예시적인 흐름셀 및 이를 제작하고 사용하기 위한 방법은 또한, 예를 들어, WO 2014/142841 A1호; 미국 특허 출원 공개 제2010/0111768 A1호 및 미국 특허 제8,951,781호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0031] 본 명세서에서 사용되는 용어 "라이브러리(library)"는 여러 상이한 항목을 포함하는 집단을 의미하는 것으로 의도된다. 집단에서 이러한 항목은 구조 및/또는 기능에 있어서 상이할 수 있다. 예를 들어, 집단은 상이한 뉴클레오타이드 서열을 갖는 핵산을 포함할 수 있거나, 집단은 상이한 1차 구조(즉, 아미노산 서열), 2차 구조, 3차 구조 또는 4차 구조를 갖는 핵산을 포함할 수 있다. 그러나, 라이브러리에서 항목의 일부 중복이 존재할 수 있는 것으로 이해될 것이다. 예를 들어, 특정 핵산 또는 단백질의 다수의 복사체는 그룹에도 불구하고, 매우 다양한 상이한 핵산 또는 단백질을 포함하는 라이브러리에 존재할 수 있다. 라이브러리에 존재할 수 있는 예시적인 타입의 항목은 후보 약제 또는 스크리닝제에 대하여 본 명세서에 기술된 것을 포함한다.

[0032] 본 명세서에서 사용되는 용어 "발광(luminescent)"은 저온 바디 방사선(cold body radiation)을 방출하는 것을 의미한다. 이러한 용어는 열의 결과로서 물질로부터 방출된 방사선인 고온 발광(incandescence)과는 구별되는 것으로 의도된다. 일반적으로, 발광은 에너지원이 이의 가장 낮은 에너지 기저 상태에서 보다 높은 에너지 여기된 상태로 원자의 전자를 전위시킬 때 나타나며, 이후에, 전자는 에너지를 방사선 형태로 되돌려 놓고, 이에 따라, 이의 기저 상태로 되돌아 갈 수 있다. 특히 유용한 타입의 발광 항목은 에너지가 여기 방사선에 의해 제공될 때 저온 바디 방사선을 방출시키는 것이다. 이러한 항목은 "형광" 또는 "광발광"으로서 지칭된다. 형광 또는 광발광은 다른 과정에서 항목을 조사하는 결과인 하나의 과정에서 항목에 의한 방사선의 방출로서 인지될 수 있다.

[0033] 본 명세서에서 사용되는 용어 "핵산" 및 "뉴클레오타이드"는 당해 분야에서 이의 사용과 일치하는 것이고 천연 발생 종 또는 이의 기능적 유사체를 포함하도록 의도된다. 핵산의 특히 유용한 기능성 유사체는 서열 특이적 방식으로 핵산에 하이브리드화 가능하거나, 특정 뉴클레오타이드 서열의 복제를 위한 주형으로서 사용될 수 있다. 천연 발생 핵산은 일반적으로, 포스포디에스터 결합을 함유한 골격을 갖는다. 유사체 구조는 당해 분야에 공지된 임의의 다양한 것을 포함하는 교대 골격 연결을 가질 수 있다. 천연 발생 핵산은 일반적으로, 데옥시리보스 당(예를 들어, 데옥시리보핵산(DNA)에서 발견됨) 또는 리보스 당(예를 들어, 리보핵산(RNA)에서 발견됨)을 갖는다. 핵산은 당해 분야에 공지된 이러한 당 모이어티의 임의의 다양한 유사체를 갖는 뉴클레오타이드를 함유할 수 있다. 핵산에는 천연 또는 비천연 염기가 포함될 수 있다. 이와 관련하여, 천연 데옥시리보핵산은 아데닌, 티민, 시토신 또는 구아닌으로 이루어진 군에서 선택되는 하나 이상의 염기를 가질 수 있고 리보핵산은 우라실, 아데닌, 시토신 또는 구아닌으로 이루어진 군에서 선택되는 하나 이상의 염기를 가질 수 있다. 핵산에 포함될 수 있는 유용한 비천연 염기는 당해 분야에 공지되어 있다.

[0034] 어레이의 피쳐와 관련하여 사용될 때 본 명세서에서 사용되는 용어 "피치(pitch)"는 인접한 피쳐에 대한 중심-대-중심 간격을 지칭하기 위해 의도된다. 피쳐의 패턴은 평균 피치에 관하여 특성화될 수 있다. 패턴은, 평균 피치 주위의 변동 계수가 작도록 정렬될 수 있거나 패턴은 랜덤일 수 있으며, 그러한 경우에, 변동 계수는 비교적 클 수 있다. 어느 한 경우에, 평균 피치는 예를 들어, 적어도 약 10nm, $0.1\mu\text{m}$, $0.5\mu\text{m}$, $1\mu\text{m}$, $5\mu\text{m}$, $10\mu\text{m}$, $100\mu\text{m}$ 이상일 수 있다. 대안적으로 또는 추가적으로, 평균 피치는 예를 들어, 최대 약 $100\mu\text{m}$, $10\mu\text{m}$, $5\mu\text{m}$, $1\mu\text{m}$, $0.5\mu\text{m}$, $0.1\mu\text{m}$ 이하일 수 있다. 물론, 피쳐의 특정 패턴에 대한 평균 피치는 상기 범위로부터 선택된 하한치들 중 하나와 상한치들 중 하나 사이일 수 있다.

[0035] 본 명세서에서 사용되는 용어 "단백질" 또는 "아미노산"은 당해 분야에서 이의 사용과 일치하고 천연 발생 종 또는 이의 기능성 유사체를 포함하도록 의도된다. 천연 발생 단백질은 일반적으로, 펩티드 결합을 함유한 골격을 갖는다. 유사체 구조는 당해 분야에 공지된 임의의 다양한 것을 포함하는 교대 골격 연결을 가질 수 있다.

천연 발생 단백질은 일반적으로 아르기닌, 히스티딘, 라이신, 아스파르트산, 글루탐산, 세린, 트레오닌, 아스파라긴, 글루타민, 시스테인, 글리신, 프롤린, 알라닌, 발린, 아이소류신, 류신, 메티오닌, 페닐알라닌, 티로신, 트립토판 및 이의 천연 발생 변형체로부터 선택된 천연 아미노산을 갖는다. 일부 천연 발생 변형체는 포스포릴화(예를 들어, 세린, 트레오닌, 히스티딘, 아스파르트산 및 글루탐산), 프레닐화, 이소프레닐화, 아실화, 알킬화, 글리코실화, 바이오틸화, 유비퀴틴화, 등을 포함한다. 단백질은 비-천연 모이어티를 갖는 천연 아미노산을 포함할 수 있다. 단백질은 천연 아미노산 또는 비-천연 아미노산을 포함할 수 있다.

[0036] 제1 제제 및 제2 제제와 관련하여 사용될 때, 본 명세서에서 사용되는 용어 "반응하다"는 제제들 중 하나 또는 둘 모두의 화학적 구조를 개질시키거나, 두 제제들 간에 하나 이상의 공유 결합을 생성시키거나, 시약들 중 하나가 다른 제제의 화학적 구조의 개질을 촉매화시킬 수 있게 하거나, 두 제제를(예를 들어, 비-공유 상호작용을 통해) 특이적으로 결합시키는 행동을 지칭하도록 의도된다. 예시적인 반응은 환원, 산화, 첨가, 제거, 재배열, 에스터화, 아미드화, 에터화, 환형화, 또는 치환과 같은 화학적 반응; 제1 제제가 특이적 친화력으로 제2 제제에 결합하는 결합 상호작용; 둘 이상의 제제가 서로 탈착하는 해리 반응; 형광; 발광; 화학발광; 및 생물학적 반응, 예를 들어, 핵산 복제, 핵산 증폭, 핵산 하이브리드화, 핵산 결합, 포스포릴화, 효소 촉매작용, 수용체 결합, 또는 리간드 결합을 포함하지만, 이들로 제한되지 않는다.

[0037] 본 명세서에서 사용되는 용어 "제조합"은 비-천연 발생 유전적 작제물을 지칭하도록 의도된다. 일 예는 하나 초과 기원으로부터의 유전 물질을 결합하는 생성물이다. 예시적인 제조합 분자는 DNA, RNA 및 단백질을 포함하지만, 이들로 제한되지 않는다. 제조합 분자가 유도될 수 있는 기원은, 예를 들어, 상이한 유기체로부터의 유사한 유전적 구성요소, 동일한 유기체로부터의 상이한 유전적 구성요소, 상이한 유기체로부터의 상이한 유전적 구성요소, 합성 유전적 구성요소, 또는 합성 및 천연 유전적 구성요소의 조합을 포함한다.

[0038] 본 명세서에서 사용되는 용어 "스크리닝제"는 제2 후보 약제와 비교하여 제1 후보 약제에 대해 선택적인 구조 또는 기능을 갖는 항목을 의미하는 것으로 의도된다. 여러 구현예에서, 스크리닝제의 구조 또는 기능은 본 발명의 방법 또는 조성물에서의 이의 사용 이전에 공지되어 있다. 예시적인 항목은 분자, 세포, 및 세포내 성분을 포함하지만, 이들로 제한되지 않는다. 분자는 선택적으로, 생물학적 활성 분자, 예를 들어, 단백질, 아미노산, 핵산(예를 들어, DNA 또는 RNA), 뉴클레오타이드, 당류, 사카라이드, 대사산물, 비타민, 효소 보조인자, 동일 수 있다. 예시적인 선택적 기능은 다른 제제의 활성화, 다른 제제의 억제, 다른 제제의 화학적 개질, 다른 제제의 분해, 다른 제제의 합성을 포함하지만, 이들로 제한되지 않으며, 여기서, 다른 제제는 후보 약제로서 상기에서 예시된 항목들 중 임의의 하나 이상일 수 있다. 스크리닝제의 구조는 상기 항목 또는 당해 분야에 공지된 다른 항목에 대해 임의의 공지된 또는 의심나는 구조일 수 있다.

[0039] 본 명세서에서 사용되는, 제2의 것과 비교하여 제1의 것을 "선택적으로" 조작하는 것(또는 ~의 "선택적" 조작)에 대한 언급은 조작이 제2의 것에 대한 영향과 비교하여 제1의 것에 대해 보다 큰 효과를 갖는 것을 의미하도록 의도된다. 조작은 제2 것에 대해 임의의 영향을 가질 필요는 없다. 조작은 제2의 것에 대한 효과에 비해 적어도 1%, 10%, 50%, 90%, 또는 99% 보다 큰, 제1의 것에 대한 영향을 가질 수 있다. 조작은 제2의 것에 대한 효과에 비해 적어도 2배, 5배, 10배, 100배, 1×10^3 배, 1×10^4 배 또는 1×10^6 배 더 높은 제1의 것에 대한 영향을 가질 수 있다. 조작은 예를 들어, 개질, 접촉, 처리, 변경, 분열(예를 들어, 화학 결합의), 광-화학적 분열(예를 들어, 화학 결합의), 형성(예를 들어, 화학 결합의), 광-화학적 형성(예를 들어, 화학 결합의), 공유 개질, 비-공유 개질, 파괴, 광-절제(photo-ablating), 제거, 합성, 중합, 광-중합, 증폭(예를 들어, 핵산의), 복사(예를 들어, 핵산의), 확장(예를 들어, 핵산의), 결합(예를 들어, 핵산의), 또는 본 명세서에 기술되거나 당해 분야에서 달리 공지된 다른 조작을 포함할 수 있다.

[0040] 본 명세서에서 사용되는 용어 "소분자"는 대략 1000 달톤 미만인 분자량을 갖는 화합물을 의미하도록 의도된다. 특정 구현예에서, 소분자는 비-폴리머이다. 그러나, 소분자는 다른 구현예에서, 다이머 또는 트라이머일 수 있다. 또한, 소분자가 폴리머에 도입될 수 있는 모노머일 수 있는 것으로 이해될 것이다. 특히 유용한 소분자는 유기 화합물이다. 유용한 소분자는 900, 800, 600, 400, 200 또는 100 달톤 미만인 분자량을 가질 수 있다.

[0041] 본 명세서에서 사용되는 용어 "고체 지지체"는 수성 액체에서 불용성인 단단한 기체를 지칭한다. 이러한 기체는 비다공성 또는 다공성일 수 있다. 기체는 선택적으로, (예를 들어, 다공성으로 인하여) 액체를 흡수할 수 있지만, 통상적으로, 기관이 액체를 흡수하였을 때 실질적으로 팽창하지 않고 액체가 건조에 의해 제거될 때 실질적으로 수축하지 않게 충분히 강성일 것이다. 비다공성 고체 지지체는 일반적으로 액체 또는 가스에 대해 투과 가능하지 않다. 예시적인 고체 지지체는 유리 및 개질되거나 작용화된 유리, 플라스틱(예를 들어, 아크릴, 폴리스타이렌 및 스타이렌과 다른 물질의 코폴리머, 폴리프로필렌, 폴리에틸렌, 폴리부틸렌, 폴리우레탄, TeflonTM, 사

이클릭 올레핀, 폴리이미드, 등), 나일론, 세라믹, 수지, Zeonor, 실리콘 및 개질된 실리콘을 포함하는 실리카 또는 실리카-기반 물질, 탄소, 금속, 무기 유리, 광섬유 번들, 및 폴리머를 포함하지만, 이들로 제한되지 않는다. 일부 구현예에 대해 특히 유용한 고체 지지체는 흐름셀의 성분이거나 흐름셀 장치 내에 위치된다. 예시적인 흐름셀은 본 명세서에서 더욱 상세히 기술된다.

[0042] 본 명세서에서 사용되는 "서브-라이브러리(sub-library)"는 라이브러리로부터의 항목 또는 라이브러리로부터의 항목의 복사체를 포함하는 표현(representation)을 갖는 집합을 의미하도록 의도된다. 서브-라이브러리에서의 표현은 라이브러리에서의 표현과 비교하여 완전하거나 부분적일 수 있다. 서브-라이브러리는 라이브러리로부터 항목들 중 적어도 일부를 분리시킴으로써 또는 라이브러리에서 항목들 중 적어도 일부의 복사체를 제조함으로써 유도될 수 있다.

[0043] 본 명세서에서 사용되는 용어 "태그 서열"은 핵산에 부착된(또는 핵산과 결합된) 체제를 식별하거나 특성화하기 위해 사용될 수 있는 핵산 중 일련의 뉴클레오타이드를 의미하는 것으로 의도된다. 태그 서열은 자연 발생 서열, 또는 핵산이 수득되는 유기체에서 자연적으로 발생하지 않는 서열일 수 있다. 특정 구현예에서, 생물학적 샘플과 함께 사용되는 하나 이상의 태그 서열은 생물학적 샘플의 게놈, 전사체, 또는 다른 핵산에서 자연적으로 존재하지 않는다. 예를 들어, 태그 서열은 특정 생물학적 샘플에서 핵산 서열에 대해 80%, 70%, 60%, 50% 또는 40% 미만의 서열 동일성을 가질 수 있다.

[0044] 본 명세서에서 사용되는 용어 "유니버설 서열"은 분자가 또한 서로 상이한 서열의 영역을 가짐에도 불구하고, 둘 이상의 핵산 서열에 대해 공통인 일련의 뉴클레오타이드를 지칭하는 것이다. 분자 집합의 상이한 일원에 존재하는 유니버설 서열은 유니버설 서열에 대해 상보적인 유니버설 포획 핵산의 집단을 사용하여 다수의 상이한 핵산의 포획을 가능하게 할 수 있다. 유사하게, 분자의 집합 중 상이한 일원에 존재하는 유니버설 서열은 유니버설 서열에 대해 상보적인 유니버설 프라이머의 집단을 사용하여 다수의 상이한 핵산의 복제 또는 증폭을 가능하게 할 수 있다. 이에 따라, 유니버설 포획 핵산 또는 유니버설 프라이머는 유니버설 서열에 특이적으로 하이브리드화할 수 있는 서열을 포함한다.

[0045] 하기에 기술되고 청구범위에서 인용되는 구현예는 상기 정의를 고려하여 이해될 수 있다.

[0046] 본 발명은 후보 약제를 특성 규명하는 방법을 제공한다. 본 방법은 (a) 후보 약제의 라이브러리를 제공하는 단계로서, 각 후보 약제는 태그 서열을 갖는 핵산 태그에 부착되는 단계; (b) 후보 약제의 라이브러리를 고체 지지체와 접촉시켜 고체 지지체에 후보 약제를 부착시키는 단계로서, 이에 의해 라이브러리로부터 개별 후보 약제에 각각 부착하는 고체 지지체 상의 개별 피처를 포함하는 후보 약제의 어레이가 형성되는 단계; (c) 후보 약제의 어레이를 스크리닝제와 접촉시키는 단계로서, 여기서, 어레이에서 하나 이상의 후보 약제가 스크리닝제와 반응하는 단계; (d) 어레이를 스크리닝제와 접촉시키는 동안 또는 후에 어레이를 검출하는 단계로서, 이에 의해, 어레이에서 적어도 하나의 후보 약제가 스크리닝제와 반응하는지를 결정하는 단계; (e) 어레이 상에서 핵산 태그를 시퀀싱하여 후보 약제 각각에 부착된 태그 서열을 결정하는 단계; 및 (f) 적어도 하나의 후보 약제에 부착된 태그 서열을 기초로 하여 스크리닝제와 반응하는 어레이에서의 적어도 하나의 후보 약제를 식별하는 단계들을 포함할 수 있다.

[0047] 임의의 다양한 후보 약제는 본 명세서에 기술된 방법에서 사용될 수 있다. 일부 구현예에서, 후보 약제는 단백질, 핵산, 세포 또는 소분자로 이루어진 군으로부터 선택된다. 일부 구현예에서, 사용되는 후보 약제가 단백질, 핵산, 세포 또는 소분자와 같은 하나 이상의 타입의 항목을 배제하는 것으로 이해될 것이다. 방법 또는 조성물로부터 배제되거나 포함될 수 있는 이러한 타입의 항목의 예는 본 발명 전반에 걸쳐 기술된다.

[0048] 유용한 후보 약제는 생물학적 활성을 갖거나, 이러한 활성을 갖는 것으로 의심되는 분자일 수 있다. 예시적인 생물학적 활성은 치료 활성, 독성, 호르몬 활성, 생물학적 분자 또는 세포의 활성화, 생물학적 분자 또는 세포의 억제, 항생제 활성, 항바이러스제 활성, 살충제 활성, 유기체 또는 이의 기관에 대한 효과, 예를 들어, 생리 약리학적 효과 또는 면역학적 효과, 등을 포함하지만, 이들로 제한되지 않는다. 특히 유용한 타입의 후보 약제에는 예를 들어, 본 명세서에 기술된 효소에 대해 표적화된 것을 포함하는, 효소 억제제 또는 효소 활성제가 있다. 또한, 세포 신호전달의 후보 활성제, 또는 세포 신호전달의 후보 억제제가 유용하다. 그러나, 일부 구현예에서, 후보 약제는 생물학적 활성을 가지거나 가지는 것으로 의심될 필요는 없다. 일부 경우에, 후보 약제는 비-생물학적 활성을 가지거나 갖는 것으로 의심될 것이다. 비-생물학적일 수 있는 활성의 예는 비제한적으로 산업적 촉매작용, 식품 보존, 원유 처리, 폴리머 합성, 등을 포함한다.

[0049] 후보 약제는 폴리머 또는 비-폴리머일 수 있다. 특히 유용한 폴리머는 단백질, 핵산, 다당류, 단백질 핵산(PNA)

및 플라스틱을 포함하지만, 이들로 제한되지 않는다. 유용한 비-폴리머 분자는 예를 들어, 지질, 아미노산, 뉴클레오타이드, 효소 보조인자, 대사물질, 단당류 및 다른 소분자를 포함한다.

[0050] 후보 약제로서 유용한 단백질의 예는 항체; 효소, 예를 들어, 옥시도환원효소, 트랜스퍼라제, 하이드롤라제, 리아제, 이소머라제, 리가제, 키나아제, 포스포타제, 폴리머라제, 프로테아제, 뉴클레아제, 셀룰라제, 리그니나아제, 아밀라제, 리파아제, 만나아제, 아밀라제, 글루카나제, 파파인, 레닌, 히스톤 개질 효소 또는 에스테라제; 또는 수용체, 예를 들어, G-결합된 수용체, 세포 표면 수용체, 면역수용체, 감각 수용체 및 핵 호르몬 수용체를 포함하지만, 이들로 제한되지 않는다.

[0051] 다른 특히 유용한 타입의 후보 약제는 예를 들어 포유류, 예를 들어, 설치류, 마우스, 래트, 토끼, 기니아피, 유제류, 말, 양, 돼지, 염소, 소, 고양이, 개, 영장류(즉, 인간 또는 비인간 영장류); 식물, 예를 들어, 아라비답시스 탈리아나(*Arabidopsis thaliana*), 옥수수, 수수, 귀리, 밀, 쌀, 카놀라 또는 대두; 해조류, 예를 들어, 클라미도모나스 레인하르티(*Chlamydomonas reinhardtii*), 선충, 예를 들어, 캐노르합디티스 엘레간스(*Caenorhabditis elegans*); 곤충, 예를 들어, 드로소필라 멜라노가스터(*Drosophila melanogaster*), 모기, 초파리, 꿀벌 또는 거미; 어류, 예를 들어, 제브라피시(zebrafish); 파충류; 양서류, 예를 들어, 개구리 또는 제노푸스 래비스(*Xenopus laevis*); 디티오스텔리움 디스코이드움(*dictyostelium discoideum*); 진균, 예를 들어, 뉴모시스티스 카리니(*pneumocystis carinii*), 타키푸구 루브리페스(*Takifugu rubripes*), 효모, 사카로마이세스 세레비시애(*Saccharomyces cerevisiae*) 또는 스킨조사카로마이세스 폼베(*Schizosaccharomyces pombe*); 또는 플라스모디움 팔시파룸(plasmodium falciparum) 유래의 것들이 포함된다. 다른 유기체는 또한 원핵생물, 예를 들어, 박테리아, 에스체리키아 콜리(*Escherichia coli*), 스태필로코키(*staphylococci*) 또는 미코플라스마 뉴모니아(*mycoplasma pneumoniae*); 원시세균; 바이러스, 예를 들어, C형 간염 바이러스 또는 인간 면역결핍 바이러스; 또는 바이로이드에서 유래될 수 있다. 세포는 상기 유기체의 균질한 배양물 또는 집단으로부터 또는 대안적으로 여러 상이한 유기체의 집합으로부터, 예를 들어, 혼합된 배양물, 커뮤니티(community) 또는 생태계에서 유래될 수 있다.

[0052] 본 명세서에서 사용되는 후보 약제는 천연 발생될 수 있고, 예를 들어, 원집단(native population)으로부터 수확될 수 있다. 예를 들어, 후보 약제는 다세포 유기체 또는 군집으로부터 단리된 유전적 자연 세포일 수 있다. 유사하게, 하나 이상의 유전적 자연 세포 또는 유기체로부터 얻어진 단백질, 핵산 또는 다른 생물학적 분자가 사용될 수 있다. 대안적으로, 후보 약제는 합성될 수 있거나, 이러한 것은 천연 발생 약제의 조작된 변형체일 수 있다. 예를 들어, 유전적으로 조작된 세포, 단백질 또는 핵산은 본 명세서에 기술된 방법에서 사용될 수 있다. 후보 약제로서 사용되는 세포 또는 세포 성분은 단세포 유기체 또는 다세포 유기체로부터 유도될 수 있다. 특정 구현예에서, 이러한 세포 타입 중 하나 이상으로부터 유도된 줄기 세포, 면역 세포 또는 생물학적 성분은 본 명세서에서의 후보 약제로서 사용될 수 있다.

[0053] 일부 구현예에서, 핵산 태그는 후보 약제의 합성과 병행하여 합성될 수 있다. 예를 들어, 후보 약제의 라이브러리는 조합 화학 방법을 이용하여 합성될 수 있으며, 특정 뉴클레오타이드(또는 뉴클레오타이드의 서열)는 다수의 상이한 화학적 모이어티가 라이브러리의 각 일원에 첨가되는지를 지시하기 위해 첨가될 수 있다. 이에 따라, 특정 후보 약제에 부착된 태그에서 뉴클레오타이드의 서열은 특정 후보 약제에 대한 합성 히스토리를 제공하는데, 이러한 정보는 그러한 특정 후보 약제의 화학적 구조를 결정하기 위해 선택적으로 사용될 수 있다. 부착 및 조합 합성 방법을 위해 사용되는 화학은 예를, 미국 특허 제5,565,324호; 제5,573,905호 또는 제6,060,596호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0054] 본 명세서에 기술된 방법의 특정 구현예는 후보 약제의 라이브러리를 조합적으로 합성하는 단계를 포함할 수 있으며, 여기서, 각 후보 약제 상에서 수행되는 조합 합성의 개별 반응은 후보 약제 각각에 부착되는 핵산 태그에 하나 이상의 뉴클레오타이드의 독특한 특징의 첨가에 의해 추적되며, 이에 의해 후보 약제의 라이브러리를 제공하며, 여기서, 각 후보 약제는 독특한 핵산 태그에 부착된다.

[0055] 핵산 태그는 핵산 및 사용 중의 특정 타입의 후보 약제와 함께 사용하기 위해 적절한 당해 분야에 공지된 임의의 다양한 화학을 이용하여 후보 약제에 부착될 수 있다. 특정 구현예에서, 핵산 태그는 후보 약제에 공유적으로 부착될 것이다. 후보 약제가 핵산이거나 핵산에 의해 엔코딩되는 경우에, 연속 핵산은 태그 서열 및 후보 약제 서열을 포함할 수 있다. 또한, 후보 약제로서 제공되거나 이를 엔코딩하는 다른 핵산에 핵산 태그를 공유적으로 부착시키는 화학적 방법을 사용하는 것이 가능하다. 후보 약제(후보 약제가 핵산 또는 다른 중이든지)에 핵산 태그를 부착시키기 위해 적합한 화학은 예를 들어, N-히드록시숙신이미드 에스터(NHS 에스터), 이미도에스터, 히드라진, 카보디이미드, 말레이미드, 할로아세틸, 피리디닐 디설파이드, 디아지린, 클릭 화학(click

chemistry)(예를 들어, 미국 특허 제6,737,236호; 제7,427,678호; 제7,375,234호; 제7,763,736호; 또는 제 8,129,542호, 이러한 문헌 각각은 본 명세서에 참고로 포함됨) 또는 실프히드릴을 포함한다. 다른 유용한 화학은 미국 특허 제7,259,258호 또는 제7,504,499호에 기술된 바와 같은 비드 또는 다른 고체 지지체에 핵산을 부착시키기 위해 사용된 것을 포함하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0056] 태그 서열은 임의의 다양한 길이를 가질 수 있다. 보다 긴 서열은 일반적으로, 집단에 대해 보다 큰 수 및 다양한 태그를 수용할 수 있다. 일반적으로, 많은 수의 모든 프로브는 동일한 길이의 태그(다른 서열을 가짐)를 가질 것이지만, 또한, 상이한 프로브에 대하여 상이한 길이의 태그를 사용하는 것이 가능하다. 태그 서열은 길이에 있어서, 적어도 2, 4, 6, 8, 10, 12, 15, 20개 또는 그 이상의 뉴클레오타이드일 수 있다. 대안적으로, 또는 추가적으로, 태그 서열의 길이는 최대 20, 15, 12, 10, 8, 6, 4개 또는 보다 적은 뉴클레오타이드일 수 있다. 사용될 수 있는 태그 서열의 예는 예를 들어, 미국 특허 출원 공개 제2014/0342921 A1호 및 미국 특허 제 8,460,865호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0057] 본 발명의 방법은 후보 약제의 라이브러리를 고체 지지체와 접촉시켜 고체 지지체에 후보 약제를 부착하는 단계를 포함할 수 있다. 결과적으로, 후보 약제의 어레이는 지지체 상에 형성될 수 있으며, 이러한 어레이는 각각이 라이브러리로부터 개별 후보 약제에 부착하는 개별 피처를 포함한다.

[0058] 임의의 다양한 고체 지지체가 사용될 수 있다. 특히 유용한 고체 지지체는 핵산 어레이를 위해 사용되는 것이다. 예는 유리, 개질된 유리, 작용화된 유리, 무기 유리, 미세구체(예를 들어, 불활성 및/또는 자성 입자), 플라스틱, 다당류, 나일론, 니트로셀룰로오스, 세라믹, 수지, 실리카, 실리카-기반 물질, 탄소, 금속, 광학 섬유 또는 광학 섬유 번들, 폴리머 및 다중웰(예를 들어, 마이크로적정) 플레이트를 포함한다. 예시적인 플라스틱은 아크릴, 폴리스타이렌, 스타이렌과 다른 물질의 코폴리머, 폴리프로필렌, 폴리에틸렌, 폴리부틸렌, 폴리우레탄, 및 TeflonTM을 포함한다. 예시적인 실리카-기반 물질은 실리콘 및 다양한 형태의 개질된 실리콘을 포함한다.

[0059] 특정 구현예에서, 고체 지지체는 용기 내에 존재할 수 있거나, 용기의 일부, 예를 들어, 웰, 튜브, 채널, 큐벳, 페트리 플레이트, 병, 등일 수 있다. 특히 유용한 용기는 예를 들어, WO 2014/142841 A1호; 미국 특허 출원 공개 제2010/0111768 A1호 및 미국 특허 제8,951,781호 또는 문헌[Bentley et al., *Nature* 456:53-59 (2008)]에 기술된 바와 같은, 흐름-셀(flow-cell)이며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 예시적인 흐름-셀은 Genome Analyzer[®], MiSeq[®], NextSeq[®] 또는 HiSeq[®] 플랫폼과 같은, 시퀀싱 플랫폼과 함께 사용하기 위한, Illumina, Inc.(캘리포니아주 샌디에이고 소재)로부터 상업적으로 입수 가능한 것이다. 다른 특히 유용한 용기는 다중웰 플레이트 또는 마이크로적정 플레이트에서의 웰이다.

[0060] 선택적으로, 고체 지지체는 겔 코팅을 포함할 수 있다. 겔을 통한 고체 지지체에 대한 핵산의 부착은 Illumina Inc.(캘리포니아주 샌디에이고 소재)로부터 상업적으로 입수 가능하거나 미국 특허 출원 공개 제2011/0059865 A1호, 제2014/0079923 A1호, 또는 제2015/0005447 A1호; 또는 PCT 공개번호 WO 2008/093098호에 기술된 흐름셀에 의해 예시되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 본 명세서에 기술된 방법 및 장치에서 사용될 수 있는 예시적인 겔은 콜로이드성 구조, 예를 들어, 아가로오스; 폴리머 메시 구조, 예를 들어 젤라틴; 또는 가교된 폴리머 구조, 예를 들어 폴리아크릴아미드, SFA(예를 들어, 미국 특허 출원 공개 제2011/0059865 A1호 참조, 이는 본 명세서에 참고로 포함됨) 또는 PAZAM(예를 들어, 미국 특허 출원 공개 제2014/0079923 A1호, 또는 제2015/0005447 A1호 참조, 이러한 문헌 각각은 본 명세서에 참고로 포함됨)을 갖는 것을 포함하지만, 이들로 제한되지 않는다. 겔은 직접적으로(예를 들어, 겔과 후보 약제 간의 공유 결합을 통해) 또는 후보 약제에 이미 부착된 상보적 핵산에 대한 겔 부착된 핵산의 하이브리드화를 통해 후보 약제를 부착하기 위해 사용될 수 있다.

[0061] 일부 구현예에서, 고체 지지체는 핵산 및/또는 후보 약제가 부착될 수 있는 피처의 어레이로서 구성될 수 있다. 특정 구현예에서, 각 피처는 하나 이하의 후보 약제를 수용할 것이거나, 그렇지 않으면, 후보 약제의 특정 혼합물의 단일 종을 함유하도록 구성될 것이다. 피처는 임의의 다양한 요망되는 포맷으로 존재할 수 있다. 예를 들어, 피처는 웰(well), 피트(pit), 채널(channel), 리지(ridge), 상승된 영역(raised region), 펙(peg), 포스트(post), 등일 수 있다. 일부 구현예에서, 피처는 비드를 함유할 수 있다. 그러나, 특정 구현예에서, 피처는 비드 또는 입자를 함유할 필요가 없다. 예시적인 피처는 454 LifeSciences(Roche의 자회사, 스위스 바젤 소재) 또는 Ion Torrent(Life Technologies의 자회사, 캘리포니아주 칼스배드 소재)에 의해 판매되는 상업적 시퀀싱 플랫폼을 위해 사용되는 기재(substrate)에 존재하는 웰을 포함한다. 웰을 갖는 다른 기재는 예를 들어, 미국 특허 제6,266,459호; 제6,355,431호; 제6,770,441호; 제6,859,570호; 제6,210,891호; 제6,258,568호; 제 6,274,320호; 미국 특허 출원 공개 제2009/0026082 A1호; 제2009/0127589 A1호; 제2010/0137143 A1호; 제

2010/0282617 A1호 또는 PCT 공개번호 WO 00/63437호에 기술된 에칭된 섬유 옵틱 및 다른 기체를 포함하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 일부 구현예에서, 기체의 웰은 미국 특허 출원 공개 제 2014/0243224 A1호에 기술된 바와 같은 겔 물질(비드를 갖거나 이를 가지지 않음)을 포함할 수 있으며, 이러한 문헌은 본 명세서에 참고로 포함된다.

[0062] 고체 지지체 상의 피처는 비-금속성 표면, 예를 들어, 유리, 플라스틱 또는 상기에 예시된 다른 물질 상의 금속 피처일 수 있다. 금속 피처를 갖는 예시적인 고체 지지체, 및 이의 제조 방법은 미국 특허 제8,895,249호 또는 미국출원공개번호 제2014/0243224 A1호에 제공되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0063] 피처는 스폿(spot) 또는 패치(patch)의 그리드로서 고체 지지체 상에 나타날 수 있다. 피처는 반복 패턴으로 또는 불규칙, 비-반복 패턴으로 위치될 수 있다. 특히 유용한 반복 패턴에는 육각형 패턴, 직선 패턴, 그리드 패턴, 반사 대칭을 갖는 패턴, 회전 대칭을 갖는 패턴, 등이 있다. 비대칭 패턴이 또한 유용할 수 있다. 피처는 상이한 쌍의 가장 가까운 이웃 피처 사이에 동일할 수 있거나, 피처는 상이한 쌍의 가장 가까운 이웃 피처 사이에서 다를 수 있다.

[0064] 고밀도 어레이는 약 15 μm 미만의 평균 피치(이웃 피처에 대해)를 갖는 것으로서 특성화된다. 중간 밀도 어레이는 약 15 내지 30 μm 의 평균 피치를 가지며, 저밀도 어레이는 30 μm 보다 큰 평균 피치를 갖는다. 본 발명에서 유용한 어레이는 100 μm , 50 μm , 10 μm , 5 μm , 1 μm 또는 0.5 μm 미만인 평균 피치를 가질 수 있다. 상기 기술되거나 본 명세서에서의 다른 곳에 기술된 평균 피치 값 및 범위는 정렬된 어레이 또는 랜덤 어레이에 적용되도록 의도된다.

[0065] 특정 구현예에서, 고체 지지체 상의 피처는 각각, 약 100nm², 250nm², 500nm², 1 μm ², 2.5 μm ², 5 μm ², 10 μm ², 100 μm ², 또는 500 μm ² 또는 그 이상 보다 더욱 큰 면적을 가질 수 있다. 대안적으로 또는 추가적으로, 피처 각각은 약 1mm², 500 μm ², 100 μm ², 25 μm ², 10 μm ², 5 μm ², 1 μm ², 500nm², 또는 100nm² 또는 그 미만 보다 작은 면적을 가질 수 있다. 상기 범위는 위에서 볼 때 또는 이미지화될 때, 고체 지지체 상의 비드 또는 다른 입자의 겹보기 면적을 기술할 수 있다.

[0066] 특정 구현예에서, 고체 지지체는 비드 또는 다른 입자의 집합을 포함할 수 있다. 표면 상에 위치된 비드를 갖는 어레이의 예는 비드가 웰에 위치되어 있는 것, 예를 들어 BeadChipTM 어레이(Illumina Inc., 캘리포니아주 샌디에이고 소재), 454 LifeSciences(Roche의 자회사, 스위스 바젤 소재)로부터의 시퀀싱 플랫폼에서 사용되는 기재, 또는 Ion Torrent(Life Technologies의 자회사, 캘리포니아주 칼스배드 소재)로부터의 시퀀싱 플랫폼에서 사용되는 기체를 포함한다. 표면 상에 위치된 비드를 갖는 다른 고체 지지체는 미국 특허 제6,266,459호; 제6,355,431호; 제6,770,441호; 제6,859,570호; 제6,210,891호; 제6,258,568호; 또는 제6,274,320호; 미국 특허 출원 공개 제2009/0026082 A1호; 제2009/0127589 A1호; 제2010/0137143 A1호; 또는 제2010/0282617 A1호 또는 PCT 공개번호 WO 00/63437호에 기술되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 여러 상기 참고문헌에는 고체 지지체에 또는 고체 지지체 상에 비드를 로딩하기 전에 비드에 핵산을 부착시키기 위한 방법이 기술되어 있다. 이와 같이, 비드의 집합은 각각 독특한 핵산이 부착된 상이한 비드를 포함할 수 있다. 그러나, 비드가 유니버설 프라이머를 포함하기 위해 제조될 수 있으며, 비드가 이후에 어레이 상에 로딩될 수 있으며, 이에 의해 본 명세서에 기술된 방법에서 사용하기 위한 유니버설 어레이를 형성시키는 것으로 이해될 것이다. 후보 약제는 고체 지지체 상에 비드가 로딩되기 전 또는 후에 비드에 부착될 수 있다. 본 명세서에서 이미 기술된 바와 같이, 비드 어레이를 위해 통상적으로 사용되는 고체 지지체는 비드 없이 사용될 수 있다. 예를 들어, 핵산(예를 들어, 프로브 또는 프라이머) 또는 후보 약제는 웰에 또는 웰에서의 겔 물질에 직접적으로 부착될 수 있다. 이에 따라, 상기 참조문헌에는 본 명세서에 기술된 방법 및 조성물에서 사용하기 위해 개질될 수 있는, 물질, 조성물, 또는 장치가 예시되어 있다.

[0067] 이에 따라, 본 명세서에 기술된 방법에서 사용되는 고체 지지체는 비드의 어레이를 포함할 수 있으며, 여기서, 상이한 후보 약제 또는 상이한 핵산은 어레이에서 상이한 비드에 부착된다. 이러한 구현예에서, 각 비드는 상이한 후보 약제 또는 핵산에 부착될 수 있으며, 비드는 고체 지지체에 상이한 핵산을 효과적으로 부착시키기 위해 고체 지지체 상에 랜덤하게 분포될 수 있다. 선택적으로, 고체 지지체는 하나 이하의 단일 비드 또는 단일 후보 약제를 수용하는 치수를 갖는 웰을 포함할 수 있다. 이러한 구성에서, 비드는 웰에 비드의 고정(fit)으로부터 형성된 힘으로 인해 웰에 부착될 수 있다. 또한, 웰에 비드를 보유하기 위해 부착 화학물질 또는 접착제를 사용하는 것이 가능하다.

- [0068] 고체 지지체는 복수의 핵산 또는 후보 약제를 포함할 수 있거나, 복수의 핵산 또는 후보 약제를 부착하기 위해 본 명세서에 기술된 방법에 의해 제조될 수 있다. 예를 들어, 고체 지지체는 적어도 10, 100, 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 개 또는 그 이상의 상이한 핵산 또는 후보 약제를 포함할 수 있다. 대안적으로 또는 추가적으로, 고체 지지체는 최대 1×10^9 , 1×10^8 , 1×10^7 , 1×10^6 , 1×10^5 , 1×10^4 , 1×10^3 , 100, 10 개 또는 보다 작은 상이한 핵산 또는 후보 약제를 포함할 수 있다. 상이한 핵산 또는 후보 약제 각각이, 예를 들어, 후보 약제의 핵산 성분이 클러스터를 형성하기 위해 증폭되었을 때, 수 개의 복사체로 존재할 수 있는 것으로 이해될 것이다. 이에 따라, 상기 범위는 고체 지지체 상에서 상이한 후보 약제 또는 핵산 클러스터의 수를 기술할 수 있다. 또한, 상기 범위가 상이한 태그, 또는 특정 핵산 또는 후보 약제에 대해 독특한 것으로서 본 명세서에 기술된 다른 서열 구성요소의 수를 기술할 수 있는 것으로 이해될 것이다. 대안적으로 또는 추가적으로, 이러한 범위는 본 명세서에 기술된 방법을 이용하여 고체 지지체 상에 생성된 확장된 핵산 또는 개질된 후보 약제의 수를 기술할 수 있다.
- [0069] 피처는 고체 지지체를 핵산 또는 후보 약제와 접촉하기 전에 고체 지지체 상에 존재할 수 있다. 예를 들어, 핵산 또는 후보 약제가 프라이머에 대한 하이브리드화를 통해 지지체에 부착되는 구현예에서, 프라이머는 피처에서 부착되며, 피처 외측의 간질 영역(interstitial area)에는 임의의 프라이머가 실질적으로 결여되어 있다. 미국 특허 제8,895,249호, 미국 특허 제8,778,849호 또는 미국 특허 출원 공개 제2014/0243224 A1호에 기술된 방법을 이용하여, 핵산 또는 후보 약제는 고체 지지체 상에 사전형성된 피처에서 포획될 수 있고, 선택적으로, 고체 지지체 상에서 증폭될 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.
- [0070] 일부 구현예에서, 피처는 고체 지지체에 핵산 태그 및/또는 후보 약제의 부착 동안 또는 후에 형성된다. 예를 들어, 고체 지지체는 프라이머의 론(lawn)을 가질 수 있거나, 그렇지 않으면 피처가 결여될 수 있다. 이러한 경우에, 피처는 고체 지지체 상에 핵산 또는 후보 약제의 부착에 의해서 형성될 수 있다. 선택적으로, 포획된 핵산은, 얻어진 클러스터가 피처가 되도록, 고체 지지체 상에서 증폭될 수 있다. 부착이 상기에서 프라이머와 다른 핵산의 상보적 부분 사이에서의 포획으로서 예시되지만, 프라이머 이외의 포획 모이어티가 사전형성된 피처에서 또는 론으로서 존재할 수 있는 것으로 이해될 것이다. 다른 예시적인 포획 모이어티는 공유 결합을 생성시키기 위해 핵산 또는 후보 약제와 반응할 수 있는 화학적 모이어티, 또는 핵산 또는 후보 약제 상의 리간드에 비-공유적으로 결합할 수 있는 수용체를 포함하지만, 이들로 제한되지 않는다.
- [0071] 특정 구현예에서, 핵산 프라이머는 고체 지지체에 부착될 것이다. 부착된 핵산 태그를 갖는, 후보 약제의 라이브리리는 부착된 핵산 프라이머에 대한 핵산 태그의 하이브리드화를 통해 고체 지지체에 부착될 수 있다. 예를 들어, 핵산 태그는 유니버설 프라이머 결합 서열을 포함할 수 있으며, 핵산 프라이머는 유니버설 프라이머 서열을 포함할 수 있으며, 후보 약제는 유니버설 프라이머 서열에 대한 유니버설 프라이머 결합 서열의 하이브리드화를 통해 고체 지지체에 부착할 수 있다. 유니버설 프라이머에 대한 대안으로서, 고체 지지체는 특이적 태그 서열에 하이브리드화하는 표적 특이적 프라이머를 포함할 수 있다.
- [0072] 피처 당 단일 후보 약제의 어레이는 고체 지지체에 대한 후보 약제의 부착에 의해 형성될 수 있다. 이에 따라, 고체 지지체 상의 하나 이상의 피처는 각각 단일 후보 약제(예를 들어, 단일 분자, 단세포 또는 다른 단일 항목)를 포함할 수 있다. 피처는 일부 구현예에서, 하나 이하의 특정 타입의 단일 후보 약제를 수용하도록 구성될 수 있다. 그러나, 피처가 하나 초과인 후보 약제를 수용할 수 있는지의 여부와는 관계없이, 피처는 하나 이하의 후보 약제, 하나 이하의 단일 핵산 태그, 또는 하나 이하의 단일 후보 약제 및 단일 핵산 태그 둘 모두를 포함할 수 있다. 대안적으로, 개별 피처는 복수의 후보 약제 및/또는 핵산 태그를 포함할 수 있다. 예를 들어, 개별 피처는 핵산 분자의 앙상블(ensemble) 및/또는 서로 동일한 서열을 갖는 단백질의 앙상블을 포함할 수 있다. 특정 구현예에서, 앙상블은 예를 들어, 각 피처에 부착된 클러스터로서, 앰플리콘을 생산하기 위해 단일 핵산 주형으로부터 증폭에 의해 생산될 수 있다.
- [0073] 본 명세서에 기술된 방법은 임의의 다양한 증폭 기술을 이용할 수 있다. 사용될 수 있는 예시적인 기술은 폴리머라제 연쇄 반응(PCR), 롤링 서클 증폭(rolling circle amplification; RCA), 다중 변위 증폭(multiple displacement amplification; MDA), 또는 랜덤 프라임 증폭(random prime amplification; RPA)을 포함하지만, 이들로 제한되지 않는다. 일부 구현예에서, 증폭은 용액 중에서 수행될 수 있다. 바람직하게, 본 발명의 방법에서 사용되는 증폭 기술은 고체상으로 수행될 것이다. 특히, 고체 지지체에 부착되고 핵산 태그에 하이브리드화하는 하나 이상의 프라이머 중(예를 들어, 핵산 태그에 존재하는 하나 이상의 유니버설 프라이머 결합 부위에 대한 유니버설 프라이머)은 증폭 기술에서 확장될 수 있다. 일 예로서 고체상 PCR 구현예를 고려하여, 증폭을

위해 사용되는 프라이머 중 하나 또는 둘 모두는 (예를 들어, 겔을 통해) 고체 지지체에 부착될 수 있다. 고체 지지체에 부착되는 두 개의 프라이머 중을 사용하는 포맷은, 이중 가닥 앰플리콘이 복사된 주형 서열 측면에 있는 두 개의 고체 지지체-부착된 프라이머 사이에 브릿지-유사 구조를 형성하기 때문에, 종종 브릿지 증폭로서 지칭된다. 브릿지 증폭을 위해 사용될 수 있는 예시적인 시약 및 조건은 예를 들어, 미국 특허 제5,641,658호, 제7,115,400호, 또는 제8,895,249호; 또는 미국 특허 공개 제2002/0055100 A1호, 제2004/0096853 A1호, 제2004/0002090 A1호, 제2007/0128624 A1호 또는 제2008/0009420 A1호에 기술되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 고체상 PCR 증폭은 또한, 고체 지지체에 부착된 증폭 프라이머 및 용액 중 제2 프라이머 중 하나로 수행될 수 있다. 고체 지지체 부착된 프라이머 및 가용성 프라이머의 조합을 사용하는 예시적인 포맷은 예를 들어, 문헌[Dressman et al., *Proc. Natl. Acad. Sci. USA* 100:8817-8822 (2003)], WO 05/010145호, 또는 미국 특허 출원 공개 제2005/0130173 A1호 또는 제2005/0064460 A1호에 기술된 바와 같이, 에멀전 PCT에서 사용되는 포맷이며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 에멀전 PCR은 이러한 포맷의 예시이며, 본 명세서에 기술된 방법의 목적을 위하여, 에멀전의 사용이 선택적인 것이며 실제로 여러 구현예를 위하여 에멀전이 사용되지 않는 것으로 이해될 것이다.

[0074] RCA 기술은 본 발명의 방법에서 사용하기 위해 변경될 수 있다. PCR 반응에서 사용될 수 있는 예시적인 성분, 및 RCA가 앰플리콘을 생산하는 원리는 예를 들어, 문헌[Lizardi et al., *Nat. Genet.* 19:225-232 (1998)] 및 미국 특허 출원 공개 제2007/0099208 A1호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. RCA를 위해 사용되는 프라이머는 용액 중에 존재할 수 있거나 고체 지지체에 부착될 수 있다. 프라이머는 본 명세서에 기술된 유니버설 프라이머 중 하나 이상일 수 있다.

[0075] MDA 기술은 본 발명의 방법에서 사용하기 위해 변경될 수 있다. MDA를 위한 일부 기본 원리 및 유용한 조건은 예를 들어, 문헌[Dean et al., *Proc Natl. Acad. Sci. USA* 99:5261-66 (2002); Lage et al., *Genome Research* 13:294-307 (2003); Walker et al., *Molecular Methods for Virus Detection, Academic Press, Inc.*, 1995; Walker et al., *Nucl. Acids Res.* 20:1691-96 (1992)]; US 5,455,166호; US 5,130,238호; 및 US 6,214,587호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. MDA를 위해 사용되는 프라이머는 용액 중에 존재할 수 있거나, 증폭 부위에서 고체 지지체에 부착될 수 있다. 또한, 프라이머는 본 명세서에 기술된 유니버설 프라이머 중 하나 이상일 수 있다.

[0076] 특정 구현예에서, 상기 예시된 증폭 기술의 조합이 사용될 수 있다. 예를 들어, RCA 및 MDA는 RCA가 (예를 들어, 용액상 프라이머를 사용하여) 용액 중 콘카타머 앰플리콘을 발생시키기 위해 사용되는 조합으로 사용될 수 있다. 이후에, 앰플리콘은 고체 지지체에 부착된 프라이머(예를 들어, 유니버설 프라이머)를 사용하여 MDA를 위한 주형으로서 사용될 수 있다. 이러한 예에서, 결합된 RCA 및 MDA 단계 후에 생산된 앰플리콘은 고체 지지체에 부착될 것이다.

[0077] 본 발명의 방법은 후보 약제의 어레이를 스크리닝제 또는 자극제와 접촉하는 단계를 포함할 수 있다. 어레이에서 하나 이상의 후보 약제는 선택적으로, 스크리닝제와 반응하거나 자극제에 반응할 수 있다. 결과적으로, 하나 이상의 후보 약제는 히트로서 분류될 수 있다.

[0078] 일부 구현예에서, 스크리닝제는 하나 이상의 후보 약제에 결합시킴으로써 또는 후보 약제에 대한 친화력을 갖는 피분석물과 후보 약제 간에 결합을 차단함으로써 어레이에서 하나 이상의 후보 약제와 반응할 것이다. 어레이는 스크리닝제가 결합되는 피처를 식별하기 위해 검출될 수 있다. 예를 들어, 스크리닝제는 예를 들어, 스크리닝제 상에 존재하는 라벨로 인하여, 검출될 수 있는 신호를 생성시킬 수 있다.

[0079] 일부 구현예에서, 스크리닝제는 후보 약제가 있는 어레이 상에 하나 이상의 피처를 개질시킬 것이다. 예를 들어, 스크리닝제는 후보 약제, 스크리닝제, 또는 둘 모두를 화학적으로 개질시킴으로써 후보 약제와 반응할 수 있다. 어레이는 개질이 이루어진 피처를 식별하기 위해 검출될 수 있다. 특히, 개질은 예를 들어, 후보 약제가 있는 피처에 부착된 라벨로 인해 검출 가능한 신호를 생성시킬 수 있다. 예시적인 개질은 후보 약제가 있는 피처에 라벨 모이어티의 첨가; 후보 약제가 있는 피처에 친화력 모이어티의 첨가(여기서, 친화력 모이어티가 라벨에 결합함); 후보 약제가 있는 피처에서 형광 켜치의 제거(여기서, 발광성 신호는 켜치의 제거로 인해 발생함); 후보 약제가 있는 피처에 포스터 공명 에너지 전달(Forster resonance energy transfer; FRET) 공여체 또는 수용체의 첨가(여기서, 첨가는 피처로부터 방출된 겔보기 발광 파장을 변경시킴); 후보 약제가 있는 피처로부터 라벨 모이어티의 제거; 후보 약제가 있는 피처로부터 친화력 모이어티의 제거(여기서, 친화력 모이어티는 라벨에 결합함); 후보 약제가 있는 피처로부터 포스터 공명 에너지 전달(FRET) 공여체 또는 수용체 모이어티의 제거(여기서, 제거는 피처로부터 방출된 겔보기 발광 파장을 변경시킴), 등을 포함한다.

- [0080] 일부 구현예에서, 스크리닝제는 피분석물을 생산하기 위해 어레이에서 하나 이상의 후보 약제와 반응할 것이다. 어레이는 피분석물이 생산된 피처를 식별하기 위해 검출될 수 있다. 예를 들어, 피분석물은 검출 가능한 신호를 형성시킬 수 있다. 생성될 수 있는 예시적인 피분석물은 발광성 라벨, 검출 가능한 수용체를 위한 리간드, 검출 가능한 생성물을 생성시키기 위해 효소에 의해 사용되는 기질, 등을 포함한다.
- [0081] 일부 구현예에서, 스크리닝제는 후보 약제의 라이브러리와 직접 접촉될 필요는 없다. 오히려, 포획제는 포획 생성물을 수득하기 위해 스크리닝제와 접촉될 수 있으며, 이후에, 포획 생성물은 후보 약제의 집합과 접촉될 수 있다. 다른 구현예에서, 포획제는 포획 생성물을 수득하기 위해 후보 약제의 집합과 접촉될 수 있으며, 이후에, 포획 생성물은 스크리닝제와 접촉될 수 있다. 포획 생성물을 사용하는 구현예에 의해 예시되는 바와 같이, 스크리닝제는 후보 약제에 대한 요망되거나 의심되는 구조 또는 기능을 지시하는데 스크리닝제가 유용하게 하기 위해 후보 약제와 직접 접촉시킬 필요는 없다.
- [0082] 본 명세서에 기술된 방법은 어레이를 스크리닝제와 접촉하는 동안 또는 후에 후보 약제의 어레이를 검출하고, 이에 의해 어레이에서 적어도 하나의 후보 약제가 스크리닝제와 반응하는 것을 결정하는 단계를 추가로 포함할 수 있다. 어레이가 부착되는 고체 지지체는 본 명세서에 기술된 임의의 다양한 상태일 수 있다. 예를 들어, 고체 지지체는 여기에 부착된 핵산 태그와 함께 후보 약제를 포함할 수 있다. 대안적으로, 고체 지지체는 핵산 태그를 포함하지 않을 수 있으며, 대신에, 후보 약제로부터 핵산 태그의 제거를 수행하는 상태이다. 핵산(예를 들어, 태그 및/또는 핵산-기반 후보 약제)은 검출 단계 동안 단일 분자 형태 또는 앙상블 형태일 수 있다. 또 다른 구현예에서, 고체 지지체는 후보 약제를 포함하지 않을 수 있으며, 대신에, 후보 약제의 제거를 따르는 상태이다. 이에 따라, 검출은 본 명세서에 기술된 방법에서 임의의 다양한 포인트에서 일어날 수 있다.
- [0083] 예를 들어, 광학적 신호, 예를 들어, 방사전의 흡광도, 발광 방출, 발광 수명, 발광 편광, 등; 레일리 및/또는 미에 산란(Rayleigh and/or Mie scattering); 자성 성질; 전기적 성질; 전하; 질량; 방사능, 등을 포함하는 임의의 다양한 신호는 본 명세서에 기술된 스크리닝 단계에서 검출될 수 있다. 본 명세서에 기술된 방법에서 검출될 수 있는 예시적인 라벨은 비제한적으로, 형광 물질, 발광체, 발색단, 나노입자(예를 들어, 금, 은, 카본 나노튜브), 중원자, 방사성 동위원소, 질량 라벨, 전하 라벨, 스핀 라벨, 수용체, 리간드, 등을 포함한다.
- [0084] 특정 구현예는 이미징 기술을 이용한다. 이미지는 당해 분야에 공지된 검출 장치를 이용하여 얻어질 수 있다. 예는 광, 명시야, 암시야, 위상차, 형광, 반사, 간섭, 또는 공초점 이미징을 위해 구성된 현미경을 포함한다. 특정 구현예에서, 형광 현미경(예를 들어, 공초점 형광 현미경)은 예를 들어, 형광 라벨에 의해 형광성을 나타내는 스크리닝제(또는 다른 피분석물)를 검출하기 위해 사용될 수 있다. 형광 시편은 또한, Illumina, Inc.(캘리포니아주 샌디에이고 소재)에 의해 상업화된 Genome Analyzer[®], MiSeq[®], NextSeq[®] 또는 HiSeq[®] 플랫폼 장치; 또는 Life Technologies(Carlsbad, CA)에 의해 상업화된 SOLiD[™] 시퀀싱 플랫폼과 같은 형광 검출을 위한 옵션을 갖는 핵산 시퀀싱 장치를 이용하여 이미징될 수 있다. 사용될 수 있는 다른 이미징 옵션은 문헌[Bentley et al., *Nature* 456:53-59 (2008)], PCT 공개번호 WO 91/06678호, WO 04/018497호 또는 WO 07/123744호; 미국 특허 제7,057,026호, 제7,329,492호, 제7,211,414호, 제7,315,019호 또는 제7,405,281호, 및 미국 특허 출원 공개 제2008/0108082호에 기술된 검출 장치에서 확인된 것을 포함하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.
- [0085] 고체 지지체의 이미지는 예를 들어, 고체 지지체 상에서 어레이에서의 후보 약제를 구별하기 위해 요망되는 해상도로 얻어질 수 있다. 이에 따라, 해상도는 적어도 0.5 μ m, 1 μ m, 5 μ m, 10 μ m, 50 μ m, 100 μ m, 500 μ m, 1 mm 또는 그 이상으로 분리되는 어레이의 피처를 구별하기에 충분할 수 있다. 대안적으로 또는 추가적으로, 해상도는 최대 1 mm, 500 μ m, 100 μ m, 50 μ m, 10 μ m, 5 μ m, 1 μ m, 0.5 μ m 또는 그 미만으로 분리되는 어레이의 피처를 구별하도록 셋팅될 수 있다.
- [0086] 본 발명의 방법은 고체 지지체(예를 들어, 어레이) 상에서 후보 약제에 부착되는 태그 서열을 결정하기 위해 핵산 태그를 시퀀싱하는 단계를 포함할 수 있다. 다수의 구현예에서, 후보 약제는 어레이에서 랜덤하게 위치되며, 시퀀싱 반응은 상이한 후보 약제 각각을 위치시키기 위한 정보를 제공한다. 시퀀싱은 후보 약제 및 이의 핵산 태그가 부착된 고체 지지체 상에서 수행될 수 있다. 이에 따라, 핵산 태그는 이의 서열을 결정하기 위해 고체 지지체로부터 제거될 필요는 없다. 일부 경우에, 후보 약제는 선택적으로, 핵산 태그 상에서 시퀀싱 반응을 수행하기 전에 고체 지지체로부터 제거될 수 있다.
- [0087] 일부 구현예에서, 후보 약제는 핵산이거나, 후보 약제를 엔코딩하는 핵산에 부착된다. 이러한 구현예에서, 시퀀싱 기술은 후보 약제의 서열을 결정하기 위해 사용될 수 있다. 시퀀싱은 후보 약제가 부착된 고체 지지체 상에

서 수행될 수 있다. 후보 약제 및 이의 핵산 태그는 함께(예를 들어, 연속 시퀀싱 관독에서) 또는 별도로(예를 들어, 비-서열화 단계는 태그의 시퀀싱과 후보 약제의 시퀀싱 사이에 수행될 수 있다) 시퀀싱될 수 있다.

[0088] 본 발명의 방법은 후보 약제 어레이의 시간 기반 검출 또는 동력학 측정을 이용할 수 있다. 이에 따라, 어레이의 검출은 어레이 상에서 개별 피쳐 중 하나 이상에 대한 여러 시점에서 신호를 획득하는 것을 포함할 수 있다. 임의의 다양한 본 명세서에 기술된 신호 및 라벨은 시간 기반 또는 동력학 분석에서 검출될 수 있다.

[0089] 본 발명의 방법은 어레이에서 후보 약제에 부착되는 태그 서열을 결정하기 위해 핵산 태그를 시퀀싱하는 단계를 포함할 수 있다. 일부 구현예에서, 후보 약제는 핵산이며, 후보 약제는 부착되는 핵산(예를 들어, mRNA 또는 cDNA에 부착된 단백질 후보 약제) 또는 핵산을 함유한 후보 약제(예를 들어, 핵산을 함유한 세포)에 의해 엔코딩된다. 시퀀싱 기술, 예를 들어, 합성에 의한 시퀀싱(SBS; sequencing-by-synthesis) 기술이 특히 유용한 방법이다.

[0090] SBS는 하기와 같이 수행될 수 있다. 제1 SBS 사이클을 개시하기 위하여, 하나 이상의 표지된 뉴클레오타이드, DNA 폴리머라제, SBS 프라이머 등은 고체 지지체 상에서의 하나 이상의 피쳐(예를 들어, 핵산이 고체 지지체에 부착된 피쳐(들))와 접촉될 수 있다. SBS 프라이머 확장이 표지된 뉴클레오타이드를 도입할 수 있게 하는 그러한 피쳐는 검출될 수 있다. 선택적으로, 뉴클레오타이드는, 뉴클레오타이드가 SBS 프라이머에 첨가된 직후에, 추가의 프라이머 확장을 종결시키는 가역적 종결 모이어티를 포함할 수 있다. 예를 들어, 가역적 종결자 모이어티를 갖는 뉴클레오타이드 유사체는, 차단방지체가 모이어티를 제거하기 위해 전달될 때까지 후속 확장이 일어나지 않을 수 있도록, 프라이머에 첨가될 수 있다. 이에 따라, 가역적 종결을 이용하는 구현예에 대하여, 차단방지 시약은(검출이 일어나기 전 또는 후) 고체 지지체에 전달될 수 있다. 세척은 다양한 전달 단계 사이에서 수행될 수 있다. 사이클은 이후에, n 뉴클레오타이드에 의해 프라이머를 확장하기 위해 n회 반복될 수 있고, 이에 의해, 길이 n의 서열을 검출할 수 있다. 본 발명의 조성물, 장치, 또는 방법과 함께 사용하기 위해 용이하게 구성될 수 있는 예시적인 SBS 절차, 유체 시스템 및 검출 플랫폼은 예를 들어, 문헌[Bentley et al., *Nature* 456:53-59 (2008)], PCT 공개번호 WO 91/06678호, WO 04/018497호 또는 WO 07/123744호; 미국 특허 제 7,057,026호, 제7,329,492호, 제7,211,414호, 제7,315,019호 또는 제7,405,281호, 및 미국 특허 출원 공개 제 2008/0108082호에 기술되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0091] 사이클릭 반응을 이용하는 다른 시퀀싱 절차, 예를 들어, 파이로시퀀싱(pyrosequencing)이 사용될 수 있다. 파이로시퀀싱은 특정 뉴클레오타이드가 초기 핵산 가닥에 도입될 때 무기 피로포스페이트(PPi)의 방출을 검출한다 [Ronaghi, et al., *Analytical Biochemistry* 242(1), 84-9 (1996); Ronaghi, *Genome Res.* 11(1), 3-11 (2001); Ronaghi et al. *Science* 281(5375), 363 (1998); 또는 미국 특허 제6,210,891호, 제6,258,568호 또는 제6,274,320호, 이러한 문헌 각각은 본 명세서에 참고로 포함됨]. 파이로시퀀싱에서, 방출된 PPi는 ATP 설퍼릴라제에 의해 아데노신 트리포스페이트(ATP)로 즉시 전환됨으로써 검출될 수 있으며, 방출된 ATP의 수준은 루시페라제-생성 광자를 통해 검출될 수 있다. 이에 따라, 시퀀싱 반응은 발광 검출 시스템을 통해 모니터링될 수 있다. 형광 기반 검출 시스템에 대해 사용되는 여기 방사선원은 파이로시퀀싱 절차를 위한 것일 필요는 없다. 본 발명의 장치, 조성물 또는 방법에 대한 파이로시퀀싱의 적용을 위해 사용될 수 있는 유용한 유체 시스템, 검출기 및 절차는 예를 들어, PCT 특허출원공개번호 WO2012/058096호, 미국 특허 출원 공개 제2005/0191698 A1호, 또는 미국 특허 제7,595,883호 또는 제7,244,559호에 기술되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0092] 예를 들어, 문헌[Shendure et al. *Science* 309:1728-1732 (2005)], 또는 미국 특허 제5,599,675호 또는 제 5,750,341호에 기술된 것을 포함하는 결찰 반응에 의한 시퀀싱(sequencing-by-ligation reaction)이 또한 유용하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 일부 구현예는 예를 들어, 문헌[Bains et al., *Journal of Theoretical Biology* 135(3), 303-7 (1988); Drmanac et al., *Nature Biotechnology* 16, 54-58 (1998); Fodor et al., *Science* 251(4995), 767-773 (1995)]; 또는 PCT 특허출원공개번호 WO 1989/10977호에 기술된 바와 같은 하이브리드화 절차에 의한 시퀀싱을 포함할 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 결찰에 의한 시퀀싱 및 하이브리드화에 의한 시퀀싱 절차 둘 모두에서, 어레이의 부위에 존재하는 핵산은 올리고뉴클레오타이드 전달 및 검출의 반복된 사이클로 수행된다. 본 명세서에 기술되거나 본 명세서에 기술된 문헌에 기술된 조성물, 장치 또는 방법은 결찰에 의한 시퀀싱 또는 하이브리드화에 의한 시퀀싱 절차를 위해 용이하게 구성될 수 있다. 통상적으로, 올리고뉴클레오타이드는 형광 표지되고, 본원 또는 본 명세서에 인용된 참고문헌에서 SBS 절차와 관련하여 기술된 것과 유사한 형광 검출기를 이용하여 검출될 수 있다.

[0093] 일부 시퀀싱 구현예는 DNA 폴리머라제 활성의 실시간 모니터링을 수반하는 방법을 사용할 수 있다. 예를 들어,

뉴클레오타이드 도입은 형광물질-함유 폴리머라제와 γ -포스페이트-표지된 뉴클레오타이드 간의 형광 공명 에너지 전달(FRET) 상호작용을 통해 그리고 제로모드 도파관(ZMW; zeromode waveguide)과 함께, 검출될 수 있다. FRET-기반 시퀀싱을 위한 기술 및 시약은 예를 들어, 문헌[Levene et al. *Science* 299, 682-686 (2003); Lundquist et al. *Opt. Lett.* 33, 1026-1028 (2008); Korlach et al. *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008)]에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0094] 일부 시퀀싱 구현에는 확장 생성물에 뉴클레오타이드의 도입 시에 방출되는 광자의 검출을 포함한다. 예를 들어, 방출된 광자의 검출을 기초로 한 시퀀싱은 전기 검출기 및 Ion Torrent(Guilford, CT, Thermo Fisher 회사)로부터 상업적으로 입수 가능한 관련된 기술, 또는 미국 특허 출원 공개 제2009/0026082 A1호; 제 2009/0127589 A1호; 제2010/0137143 A1호; 또는 US 2010/0282617 A1호에 기술된 시퀀싱 방법 및 시스템을 사용할 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0095] 본 발명의 방법은 고체 지지체로부터 하나 이상의 후보 약제를 제거하는 단계를 추가로 포함할 수 있다. 특정 구현예에서, 후보 약제는 피처에서 핵산 태그를 시퀀싱하기 전에 어레이의 피처로부터 제거될 수 있다. 이는 후보 약제가 시퀀싱 기술을 간섭할 때 또는 시퀀싱 기술에서 사용되는 시약으로부터 후보 약제를 보호하는 것이 요망될 때 유리할 수 있다. 일부 구현예에서, 스크린에서 히트로서 식별된 하나 이상의 후보 약제는 어레이로부터 선택적으로 제거될 수 있다.

[0096] 일부 구현예에서, 후보 약제는 고체 지지체에 이를 부착시키는 테더(tether)를 분열시킴으로써 제거될 수 있다. 예를 들어, 후보 약제는 이러한 것이 부착되는 핵산 태그를 통해 고체 지지체에 테더링될 수 있다. 태그 핵산이 상보적 핵산에 대한 하이브리드화에 의해 지지체에 부착되는 경우에, 제거는 하이브리드 착물을 변성시킴으로써 달성될 수 있다. 대안적으로, 핵산 태그는 후보 약제 및 태그 서열에 대한 부착 포인트 사이에 위치한 분열 부위를 포함할 수 있다. 이와 같이, 후보 약제는 분열 부위에 표적화된 분열 반응에 의해 태그 서열로부터 분리될 수 있다. 일부 방법에서, 분열 부위는 증폭 또는 개질 단계 동안 핵산에 도입될 수 있다. 예를 들어, 분열 부위는 확장 단계 동안 확장된 프라이머에 도입될 수 있다.

[0097] 예시적인 분열 부위는 결합 파괴를 야기시키는 화학적, 효소적, 또는 물리적 공정에 민감한 모이어티를 포함하지만, 이들로 제한되지 않는다. 예를 들어, 이러한 위치는 엔도뉴클레아제에 의해 인식되는 뉴클레오타이드 서열일 수 있다. 적합한 엔도뉴클레아제 및 이의 인식 서열은 당해 분야에 널리 공지되어 있고, 여러 경우에, 심지어, 상업적으로 입수 가능하다[예를 들어, New England Biolabs, Beverley MA; ThermoFisher, Waltham, MA 또는 Sigma Aldrich, St. Louis MO로부터]. 특히 유용한 엔도뉴클레아제는 핵산에서 이의 결합 부위에 대해 3'-원격인 부위에 핵산 가닥에서 결합을 파괴할 것이며, 이의 예는 타입 II 또는 타입 II 제한 엔도뉴클레아제를 포함한다.

[0098] 광-불안정한 모이어티는 유용한 분열 부위를 제공한다. 예시적인 광-불안정한 모이어티는 (1-(4,5-디메톡시-2-니트로페닐)에틸)에터(즉, DMNPE) 및 (1-(2-니트로페닐)에틸)에스터(즉, NPE)를 포함하지만, 이들로 제한되지 않는다[문헌[Meth. Enzymol. 291:307-347 (1998)] 참조, 이러한 문헌은 본 명세서에 참고로 포함됨]. 다른 광-불안정한 모이어티, 및 이의 합성 및 사용하기 위한 방법은 WO 91/06678호 또는 미국 특허 출원 공개 제 20100092957 A1호에 기술되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 광-불안정한 모이어티는 유리하게, 개별 피처로부터 후보 약제의 표적화된 분열을 위해 사용될 수 있다. 공간 분해된 방사선(예를 들어, 레이저, 포커싱된 광원 및/또는 공간 필터/마스크로부터)은 특정 후보 약제가 광-불안정한 모이어티를 통해 테더링되는 개별 피처를 다루기 위해 사용될 수 있다. 결과적으로, 그러한 특정 후보 약제는 (동일한 어레이 상에서 다른 피처에서의 후보 약제와 비교하여) 선택적으로 제거될 수 있다.

[0099] 일부 구현예에서, 분열 부위는 무염기 부위, 또는 무염기 부위를 생성시키기 위해 제거되기에 용이한 염기를 갖는 뉴클레오타이드이다. 무염기 부위로부터 제거되기에 용이한 뉴클레오타이드의 예는 우라실 및 8-옥소-구아닌을 포함한다. 특정 구현예에서, New England Biolabs로부터 상업적으로 입수 가능한 USERTM 시약이 사용된다. 핵산을 분열시키기 위해 사용될 수 있는 분열 부위 및 방법의 다른 예는 예를 들어, 미국 특허 제7,960,120호에 기술되어 있으며, 이러한 문헌은 본 명세서에 참고로 포함된다.

[0100] 본 발명은 또한, 단백질의 어레이를 생산하는 방법을 제공한다. 본 방법은 (a) 고체 지지체에 부착된 cDNA 분자의 라이브러리를 제공하는 단계; (b) 고체 지지체 상에 cDNA 분자를 증폭시켜, 클러스터를 형성시키는 단계로서, 각 클러스터는 라이브러리로부터 특정 cDNA 분자의 다수의 복사체를 포함하는 단계; (c) 클러스터에서 다수의 복사체를 전사하여 클러스터 각각에 부착된 다수의 mRNA 분자를 생산하는 단계; 및 (d) 클러스터에서

mRNA 분자를 번역하여 클러스터 각각에 부착된 다수의 단백질을 생성시키는 단계를 포함할 수 있다.

- [0101] 본 발명은 또한, (a) mRNA 분자의 라이브러리를 제공하는 단계로서, 라이브러리에서 개별 mRNA 분자는 표적 서열 및 태그 서열을 포함하는 단계, (b) 라이브러리로부터 제1 서브-라이브러리를 유도하는 단계로서, 제1 서브-라이브러리는 태그 서열 또는 이의 보체를 갖는 핵산을 포함하고, 핵산은 고체 지지체 상에서 개별 피처에 부착되는 단계; (c) 라이브러리로부터 제2 서브-라이브러리를 유도하는 단계로서, 제2 서브-라이브러리는 표적 서열 및 태그 서열 또는 이의 보체를 갖는 핵산을 포함하는 단계; (d) 제2 서브-라이브러리를 제1 서브-라이브러리와 접촉시켜, 태그 서열 및 이의 보체의 하이브리드화를 통해 고체 지지체에 제2 서브-라이브러리의 핵산을 부착하는 단계; 및 (e) 고체 지지체 상에서 표적 서열을 번역하여 개별 피처에 부착된 단백질의 어레이를 생성시키는 단계를 포함하는 단백질의 어레이를 생산하는 단계를 포함한다.
- [0102] 게놈 DNA(gDNA), 메신저 RNA(mRNA) 또는 복사 DNA(cDNA)의 라이브러리는 고려되는 하나 이상의 단백질을 발현시키는 하나 이상의 유기체로부터 유도될 수 있다. 예를 들어, 라이브러리는 요망되는 단백질 또는 요망되는 활성을 갖는 단백질을 발현시키기 용이한 유기체로부터 취득될 수 있다. 본 명세서에 기술된 유기체와 같은 당해 분야에 공지된 임의의 유기체가 사용될 수 있다. 유기체는 산업적, 치료적, 진단적 또는 예측적 관심의 대상일 수 있다. 예를 들어, 라이브러리는 인간 유기체로부터 얻어지고, 예후 또는 진단 목적을 위해 스크리닝될 수 있다. 대안적으로, 유기체는 치료 효과를 형성시키거나 산업적 적용을 갖는, 유전자를 발현시키기에 용이할 수 있다. 일부 구현예에서, 라이브러리는 비-인간 유기체로부터 얻어지거나, 달리 비-인간 기원이다.
- [0103] 일부 적용에서, gDNA, mRNA 또는 cDNA 분자의 라이브러리는 제조할 유기체의 라이브러리에서 발현되는 단일 유전자의 변형체를 포함한다. 이러한 변형체는 문헌[Sambrook et al., *Molecular Cloning: A Laboratory Manual, Third Ed.*, Cold Spring Harbor Laboratory, New York (2001) 및 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999)]에 기술된 것을 포함하지만, 이들로 제한되지 않는 공지된 단백질 조작 기술을 이용하여 생성될 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 예를 들어, 변형체는 모든 유전자 또는 유전자의 일부의 랜덤 돌연변이 유발에 의해 생성될 수 있다. 반-랜덤 돌연변이 유발 기술이 또한 사용될 수 있다.
- [0104] 본 명세서에 기술된 방법에서 발현되는 유전자(예를 들어, mRNA, cDNA 또는 단백질로서)는 항체, 효소, 수용체, 키나아제, 포스포타제, 폴리머라제, 프로테아제, 에스테라제, 핵 호르몬 수용체, 히스톤 개질 효소 및 본 명세서에 예시된 다른 것들로 이루어진 군으로부터 선택될 수 있다. 유전자는 제조할 수 있고, 예를 들어, 유전자의 천연 공급원과 상이하거나 유전자에 자연적이지 않은 개질을 갖는 유기체에서 발현된다. 특정 구현예에서, gDNA, mRNA, cDNA 또는 단백질 라이브러리는 클로닝 또는 발현을 위해 일반적으로 사용되는 박테리아, 효모, 곤충, 포유동물 또는 다른 세포계로부터 단리될 수 있다.
- [0105] 상이한 핵산(예를 들어, gDNA, mRNA 또는 cDNA)의 라이브러리는, 라이브러리의 각 개별 일원이 단백질 코딩 서열 및 태그 서열을 포함하도록 구성될 수 있다. 태그 서열은 랜덤하게 또는 체계적으로(즉, 공지된 태그가 공지된 변형체와 선형적으로 연관되는 경우) 각 단백질 코딩 서열로 지정될 수 있다. 일 예로서, 태그는 또한 단백질을 엔코딩하는 핵산 작제물을 합성할 때 랜덤 뉴클레오타이드를 태그 영역에서 하나 이상의 위치에 도입함으로써 랜덤하게 지정될 수 있다. 다른 예에서, 상이한 핵산 태그의 집단은 평균, 각 단백질 엔코딩 핵산이 독특한 태그 서열에 결합하도록 상이한 단백질 엔코딩 핵산의 집단에 랜덤하게 결합될 수 있다. 체계적 태그화의 예로서, 특이적 태그는 작제물에서 단백질 엔코딩 서열에서 돌연변이 또는 변형체와 선형적으로 연관되는 방식으로 핵산 작제물에 합성될 수 있다.
- [0106] 본 명세서에 기술된 방법은 라이브러리로부터 제1 및 제2 서브-라이브러리를 유도하는 단계를 포함할 수 있다. 일부 구현예에서, 서브-라이브러리는 라이브러리로부터 분획을 제거함으로써 또는 라이브러리를 물리적으로 분할시킴으로써 유도된다. 일반적으로, 라이브러리는 각 후보 약제의 다수의 복사체를 포함하는 유체일 것이다. 임의의 다양한 후보 약제, 예를 들어, gDNA, mRNA, cDNA, 마크로사이클, 사이클릭 펩티드, 융합 분자(예를 들어, 핵산-단백질 융합), 디스플레이된 작제물(예를 들어, 표지 상의 펩티드), 또는 단백질, 등이 사용될 수 있다. 이와 같이, 유체를 분획화하거나 분할함으로써 유도되는 서브-라이브러리는 대략 동일한 함유물을 갖는 것으로 예상될 것이다. 일부 경우에서, 증폭 단계는 다수의 복사체를 생성시키기 위해 핵산 라이브러리(예를 들어, DNA 또는 RNA 라이브러리) 상에서 수행될 수 있다.
- [0107] gDNA, mRNA, cDNA 또는 단백질 라이브러리는 본 명세서에서의 다른 곳에 기술된 방법을 이용하여 고체 지지체에 부착될 수 있다. 특정 구현예에서, mRNA 분자의 라이브러리가 사용되며, 여기서, mRNA 분자 각각은 태그 서열과 함께 후보 약제를 엔코딩하는 서열을 포함한다. 제1 서브-라이브러리는 고체 지지체에 mRNA 분자를 부착하기 위

해 라이브러리의 mRNA 분자를 고체 지지체와 접촉시키는 것을 포함하는 방법에 의해 유도된다. 예를 들어, 고체 지지체는 고체 지지체-부착된 핵산 프라이머를 가질 수 있으며, mRNA 분자는 핵산 프라이머에 대한 하이브리드화를 통해 고체 지지체에 부착할 수 있다. 더욱 특정 예로서, mRNA 분자는 유니버설 프라이머 결합 서열을 포함할 수 있으며, 핵산 프라이머는 유니버설 프라이머 서열을 포함하며, mRNA 분자는 유니버설 프라이머 서열에 대한 유니버설 프라이머 결합 서열의 하이브리드화를 통해 고체 지지체에 부착할 수 있다. 부착된 mRNA 분자는 이후에, 태그 서열의 보체를 생성시키기 위해 고체 지지체 상에 복사되거나 증폭될 수 있다. 이후에, 제1 서브-라이브러리와 동일한 라이브러리로부터 유도된 제2 서브-라이브러리는 제1 서브-라이브러리의 앰플리콘(또는 복사체)와 접촉될 수 있고, 이에 의해, 태그 서열 및 이의 보체의 하이브리드화를 통해 고체 지지체에 제2 서브-라이브러리의 핵산을 부착시킬 수 있다. 이러한 방식으로, 고체 지지체는 제2 서브-라이브러리의 mRNA 분자에 의해 발현되는 단백질의 라이브러리를 생성시키기 위한 스크리닝 반응을 위해 또는 번역을 위해 이용 가능한 mRNA의 분자를 포함하도록 개질된다.

[0108] 일부 구현예에서, 제1 핵산 서브-라이브러리 또는 제2 핵산 서브-라이브러리 중 하나는 복사될 수 있거나 용액 중에서 증폭될 수 있다. 이와 같이, 증폭되거나 복사된 서브-라이브러리는 다른 서브-라이브러리에서 태그 서열에 하이브리드화될 수 있는 상보적 태그 서열을 포함할 것이다. 서브-라이브러리 중 어느 하나는 고체 지지체에 부착될 수 있으며, 다른 서브-라이브러리는 스크리닝 반응, mRNA를 생성시키기 위한 전사 반응 및/또는 단백질이 라이브러리를 생성시키기 위한 번역 반응을 위해 이용 가능한 핵산 분자의 라이브러리를 갖는 고체 지지체를 생성시키기 위해 하이브리드화될 수 있다.

[0109] mRNA 라이브러리를 사용하는 구현예에 대하여, 제1 서브-라이브러리는 cDNA를 생성시키기 위해 라이브러리에서 개별 mRNA 분자의 일부 또는 모두를 역전사하는 것을 포함하는 방법에 의해 유도될 수 있다. 역전사는 용액 중에서 또는 고체 지지체 상에서 일어날 수 있다. 이후에, 얻어진 cDNA 분자는 용액 중에서 또는 고체 지지체 상에서 증폭될 수 있다. 이에 따라, 라이브러리에 있는 태그 서열의 보체를 갖는 cDNA 분자가 생성될 수 있다. 일부 경우에, cDNA 분자는 mRNA 분자의 일부 이하, 예를 들어, 다중도메인 단백질로부터 특정 도메인 또는 도메인의 서브셋을 발현시키는 부분을 포함할 수 있다.

[0110] 본 발명의 방법은 cDNA를 mRNA로 전사하는 단계를 포함할 수 있다. 전사는 문헌[Sambrook et al., *Molecular Cloning: A Laboratory Manual, Third Ed.*, Cold Spring Harbor Laboratory, New York (2001) 및 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999)]에 기술된 것 또는 상업적으로 입수 가능한 것과 같은 공지된 각태일을 이용하여 수행될 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 특정 구현예에서, 전사는 지지체-부착된 cDNA 분자 상에서 수행된다. 전사 반응이 mRNA 생성물은 이러한 것이 생성되는 피처에 부착될 수 있다. 예를 들어, 부착은 본 명세서에서의 실시예 II에 기술된 기술을 이용하여 달성될 수 있다. 또한, (예를 들어, 직접 부착 또는 에멀전 액적에서의 공동-국소화에 의해) 이러한 것을 엔코딩하는 DNA 분자에 단백질을 부착하는 것이 유용할 수 있다. 단백질은 또한, 이러한 것을 엔코딩하는 DNA 분자에 가깝게 부착될 수 있다(예를 들어, 핵산 프로그램 가능한 단백질 어레이(NAPPA)).

[0111] 본 발명의 방법은 mRNA를 단백질로 번역하는 단계를 포함할 수 있다. 번역은 문헌[Sambrook et al., *Molecular Cloning: A Laboratory Manual, Third Ed.*, Cold Spring Harbor Laboratory, New York (2001) 및 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999)]에 기술된 것 또는 상업적으로 입수 가능한 것과 같은 공지된 각태일을 사용하여 사용될 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 특정 구현예에서, 번역은 지지체-부착된 mRNA 분자 상에서 수행된다. 번역 반응의 단백질 생성물은 이러한 것이 생성되는 피처에 부착될 수 있다. 예를 들어, 부착은 본 명세서에서의 실시예 II에 기술된 기술을 이용하여 달성될 수 있다.

[0112] 단백질 어레이를 생산하기 위한 본 명세서에 기술된 방법은 단백질을 스크리닝하는 방법에서 사용하기 위해 포함될 수 있다. 본 방법은 (i) 단백질의 어레이를 생산하는 단계; (ii) 단백질의 어레이를 스크리닝제와 접촉시키는 단계로서, 어레이에서 하나 이상의 단백질은 스크리닝제(또는 자극제)와 반응하는 단계; 및 (iii) 스크리닝제(또는 자극제)와 접촉하는 동안 또는 후에 단백질의 어레이를 검출하는 단계로서, 이에 의해 어레이에서 적어도 하나의 단백질이 스크리닝제(또는 자극제)와 반응하는지를 결정하는 단계를 포함할 수 있다.

[0113] 특정 구현예에서, 스크리닝제는 하나 이상의 단백질에 결합시킴으로써 또는 단백질에 대한 친화력을 갖는 피분 석물과 단백질 간의 결합을 차단함으로써 하나 이상의 단백질과 반응한다. 어레이의 검출은 하나 이상의 단백질에 결합되는 스크리닝제를 검출하는 것을 포함할 수 있다. 예를 들어, 스크리닝제는 발광성일 수 있으며, 검출은 어레이 상에서 발광을 검출함으로써 수행될 수 있다.

- [0114] 일부 구현예에서, 스크리닝제는 하나 이상의 단백질을 화학적으로 개질시킴으로써 하나 이상의 단백질과 반응한다. 이러한 경우에, 어레이의 검출은 하나 이상의 개질된 단백질을 검출함으로써 달성될 수 있다. 예를 들어, 하나 이상의 개질된 단백질은 발광성일 수 있으며, 검출은 어레이 상에서 발광을 검출함으로써 다량될 수 있다.
- [0115] 일부 구현예에서, 스크리닝제는 피분석물 산물을 생성시킴으로써 하나 이상의 단백질과 반응한다. 이러한 경우에, 어레이의 검출은 피분석물 산물을 검출하는 것을 포함한다. 예를 들어, 피분석물 산물은 발광성일 수 있으며, 검출은 어레이 상에서 발광을 검출함으로써 수행될 수 있다.
- [0116] 본 명세서에 기술된 바와 같이, 어레이의 검출은 어레이 상에서 개별 피처 중 하나 이상에 대한 여러 시점에서 신호를 획득하는 것을 포함할 수 있다.
- [0117] gDNA, cDNA 또는 mRNA 분자의 일부인 태그 서열은 본 명세서에 기술되거나 당해 분야에 공지된 방법을 이용하여 시퀀싱될 수 있다. 예를 들어, 시퀀싱 반응은 스크린이 수행되거나, 수행되었거나, 수행될 고체 지지체 상에서 수행될 수 있다. 이에 따라, 본 명세서에 기술된 방법은 고체 지지체 상에서 태그 서열 또는 이의 보체를 시퀀싱하여, 고체 지지체 상에서의 개별 피처에서, 태그 서열 또는 이의 보체의 위치를 결정하는 단계를 포함할 수 있다. 이러한 방법은 적어도 하나의 단백질에 부착된 태그 서열을 기초로 하여 스크리닝제와 반응하는 어레이에서의 적어도 하나의 단백질을 식별하는 단계를 추가로 포함할 수 있다.
- [0118] 본 명세서에 기술된 방법은 고체 지지체 상에서 후보 gDNA, cDNA 또는 mRNA를 시퀀싱하는 단계를 추가로 포함할 수 있다. 이에 따라, 하나 이상의 시퀀싱 반응은 태그 서열의 식별 및/또는 후보 서열의 식별을 결정하기 위해 고체 지지체 상에서 수행될 수 있으며, 이에 의해 고체 지지체 상의 개별 피처에서 요망되는 후보 약제의 위치를 결정할 수 있다.
- [0119] 일부 경우에서, gDNA, cDNA 또는 mRNA는 시퀀싱 방법의 평균 판독 길이 보다 더욱 길 수 있다. 이러한 경우에, 후보 핵산은 각각 상이한 프라이머를 사용하여 여러 시퀀싱 실행으로 처리될 수 있다. 예를 들어, 제1 시퀀싱 실행은 본 명세서에 그리고 본 명세서에 도입된 참고문헌에 기술된 시퀀싱 방법 및 플랫폼에서 일반적으로 사용되는 유니버설 프라이밍 서열에 하이브리드화하는 프라이머를 사용할 수 있다. 후속 시퀀싱 실행은 제1 시퀀싱 프라이머가 하이브리드화되는 위치의 다운스트림에서 하이브리드화하는 제2 시퀀싱 프라이머를 사용하여 수행될 수 있다. 프라이머는, 주형 핵산의 중첩 또는 인접한 판독이 얻어질 수 있도록 시퀀싱 실행의 예상되는 판독 길이에 따라 이격될 수 있다. 여러 시퀀싱 실행은 후보 핵산의 요망되는 영역(들)을 포함하기 위해 충분한 수의 시퀀싱 프라이머를 위해 반복될 수 있다.
- [0120] 후보 핵산 시퀀싱 방법의 평균 판독 길이 보다 더욱 길 때의 다른 옵션은 카세트 방법을 이용하는 것이다. 카세트 방법에서, 핵산의 작은 영역 또는 도메인은 후보 핵산에서 돌연변이되거나 개질된다. 카세트는 단일 시퀀싱 실행에서 판독하기에 충분히 짧도록 선택될 수 있다. 카세트의 업스트림에 있는 후보 핵산의 영역이 라이브러리의 모든 일원에서 균일하기만 하면, 이는 유니버설 프라이밍 부위로서 효과적으로 역할을 할 수 있으며, 시퀀싱 프라이머는 카세트에서 국소화된 시퀀싱을 위해 이러한 영역에 하이브리드화할 수 있다.
- [0121] 일부 구현예에서, 후보 gDNA, cDNA 또는 mRNA는 시퀀싱될 필요가 없다. 오히려, 태그 서열은, 태그의 시퀀싱이 후보 핵산을 식별하기에 충분하도록, 공지된 후보 핵산 서열과 선택적으로 연관될 수 있다. 특정 피처에 존재하는 태그의 선택적 지식은 후보 핵산, 이를 엔코딩하는 단백질, 또는 단백질의 성질을 식별하기 위해 스크리닝 단계에서 피처로부터 관찰된 신호와 연관될 수 있다.
- [0122] 일부 구현예에서, 후보 핵산(예를 들어, gDNA, cDNA 또는 mRNA) 또는 단백질은 고체 지지체에 이를 부착시키는 테더를 분열시킴으로써 제거될 수 있다. 예를 들어, 이러한 후보 약제는 태그 서열을 통해 고체 지지체에 테더링될 수 있다. 분열 부위는 후보 약제 및 태그 서열에 대한 부착 포인트 사이에 위치될 수 있다. 이와 같이, 후보 약제는 분열 부위에 표적화된 분열 반응에 의해 태그 서열로부터 분리될 수 있다. 예시적인 분열 부위는 본 명세서에서의 다른 곳에 기술되거나 그렇지 않으면 당해 분야에 공지된 것을 포함하지만, 이들로 제한되지 않는다.
- [0123] 이에 따라, 본 발명의 방법은 어레이의 하나 이상의 피처에 부착된 gDNA, cDNA, mRNA 또는 단백질 분자를 선택적으로 제거하는 단계를 포함할 수 있다. 선택적 제거는 예를 들어, 본 명세서에서의 다른 곳에 기술된 시약 및 방법을 사용하여, 피처에 분자를 부착시키는 광-불안정한 결합의 광 매개 분열을 포함할 수 있다.
- [0124] 본 발명은 (a) 고체 지지체; (b) 고체 지지체에 부착된 상이한 cDNA 분자의 라이브러리로서, 각 상이한 cDNA 분자가 고체 지지체 상의 개별 피처에 부착되며 각 피처가 특정 cDNA 분자의 다수의 복사체를 포함하는 라이브러

리; (c) cDNA 분자에 부착된 mRNA 분자로서, cDNA 각각이 개개 부착된 mRNA 분자에 상보적인 mRNA 분자; 및 (d) mRNA 분자에 부착된 단백질 분자로서, 단백질 분자 각각이 개개 부착된 mRNA 분자에 의해 엔코딩된 단백질 분자를 포함하는 어레이를 제공한다.

[0125] 본 발명은 (a) mRNA 분자의 라이브러리로서, 라이브러리에서 개별 mRNA 분자가 표적 서열 및 태그 서열을 포함하는 라이브러리, (b) 태그 서열의 보체를 갖는 핵산을 포함하는 고체 지지체로서, 핵산이 고체 지지체 상의 개별 피처에 부착되는 고체 지지체를 포함하는 어레이로서, 개별 mRNA 분자의 태그 서열이 고체 지지체 상의 개별 피처에서 개개 상보적인 태그 서열에 하이브리드화되며, mRNA 분자의 번역에 의해 유도된 단백질이 개개 mRNA 분자에 부착되는, 어레이를 추가로 제공한다.

[0126] 본 발명의 어레이는 본 명세서에 기술된 방법의 하나 이상의 단계로부터 형성된 성분을 포함할 수 있다. 예를 들어, 핵산, 예를 들어, gDNA, cDNA 또는 mRNA 종은 본 명세서에 기술된 방법에 따라 고체 지지체에 부착될 수 있다. 유사하게, 단백질은 본 명세서에 기술된 방법에 따라 고체 지지체에 부착될 수 있다. 특정 구현예에서, 실시예 II 및 실시예 III에 기술된 바와 같이, cDNA는 고체 지지체에 부착될 수 있으며, 엔코딩된 mRNA는 cDNA에 부착될 수 있으며, 엔코딩된 단백질은 mRNA에 부착될 수 있다. 그러한 실시예에 기술된 바와 같이, cDNA를 통한 부착은 선택적이며, 대신에, mRNA는 (예를 들어, 상보적 태그에 대한 하이브리드화를 통해 또는 다른 부착 방법을 통해) 고체 지지체에 부착될 수 있으며, 엔코딩된 단백질은 mRNA에 부착될 수 있다.

[0127] 선택적으로, mRNA 분자는 RNA 폴리머라제와 함께 형성된 착물을 통해 이의 엔코딩 cDNA 분자에 부착될 수 있다. 부착은 착화된 cDNA, mRNA 및 RNA 폴리머라제 간의 공유 가교에 의해 매개될 수 있다. 유사하게, 단백질 분자는 리보솜과 함께 형성된 착물을 통해 이의 엔코딩 mRNA에 부착될 수 있다. 부착은 착화된 mRNA, 단백질 및 리보솜 간의 공유 가교에 의해 매개될 수 있다.

[0128] 본 명세서에 기술된 스크리닝 단계의 임의의 다양한 스크리닝제, 라벨 또는 생성물은 본 발명의 어레이에 존재할 수 있다. 예를 들어, 단백질이 위치되는 피처는 선택적으로, 단백질에 대한 발광성 분자(예를 들어, 스크리닝제)의 특이적, 비-공유 결합을 통해 또는 단백질에 공유적으로 부착된 발광성 모이어티를 통해, 발광적으로 표지될 수 있다. 모든 피처가 반드시 표지될 필요는 없다. 오히려, 요망되거나 선택된 활성을 갖는 단백질을 함유한 피처의 서브세트는 선택적으로 표지될 수 있으며, 다른 피처는 표지되지 않는다. 예를 들어, 피처의 50%, 25%, 10%, 5% 또는 1% 이하는 라이브러리의 함량 및 스크린의 특성에 따라 표지될 수 있다.

[0129] 임의의 다양한 본 명세서에 기술된 후보 약제 및/또는 핵산 태그는 본 발명의 어레이에 존재할 수 있다. 어레이 상의 상이한 종의 수, 어레이 상의 종을 함유한 피처의 밀도, 또는 특정 피처에 부착된 각 종의 수는 어레이를 제조하고 사용하는 방법과 관련하여 기술된 범위일 수 있다.

[0130] 또한, 본 발명에 의해 세포를 스크리닝하는 방법이 제공된다. 본 방법은 (a) 복수의 상이한 세포를 제공하는 단계로서, 상이한 세포 각각은 태그 서열을 갖는 핵산 태그를 포함하는 단계; (b) 상이한 세포의 혼합물을 고체 지지체와 접촉시켜 고체 지지체 상에 부착된 다수의 세포를 형성시키는 단계; (c) 적어도 하나의 광학적 특징에 대해 고체 지지체 상의 세포의 어레이를 스크리닝하는 단계로서, 스크리닝 반응은 고체 지지체에 부착된 개별 세포를 검출하는 것을 포함하는 단계; (d) 고체 지지체에 부착된 핵산 태그의 태그 서열을 시퀀싱하는 단계; 및 (e) 후보 세포의 광학적 특징 및 태그 서열을 기초로 하여 후보 세포로서 어레이에서의 적어도 하나의 세포를 식별하는 단계를 포함할 수 있다.

[0131] 특정 구현예에서, 세포는 본 명세서에 기술된 방법에서 후보 약제로서 사용된다. 세포는 다세포 유기체로부터 단리된 천연 세포, 단세포 유기체를 포함하는 천연 세포, 다세포 유기체 또는 유전적을 조작된 단세포화된 유기체로부터 유전적으로 조작된 세포일 수 있다. 본 명세서에 기술된 방법 또는 조성물에서 사용되는 세포는 천연 공급원으로부터 수득될 수 있거나, 이러한 것은 세포외 배양물로부터 수득될 수 있다.

[0132] 세포외에서 복사되거나 배양되거나 확장될 수 있는 세포가 특히 유용하다. 예를 들어, 스크리닝 단계 전 또는 후에, 후보 세포의 하나 이상의 복사체를 제조하는 것이 유용하다. 이에 따라, 스크린(또는 세포의 클론)에서 히트로서 식별된 세포는 사전형성된 모액으로부터 또는 스크린에서 사용되는 고체 지지체로부터 단리될 수 있다. 단리된 세포는 예를 들어, 세포를 보다 충분히 특성화하거나, 치료 절차에서 세포를 사용하기 위해 추가로 사용되거나 조작될 수 있다.

[0133] 유전적으로 개질된 세포를 사용하는 구현예에 대하여, 유전적 개질은 예를 들어, 비-천연 재조합 단백질을 발현 시키거나, 돌연변이 재조합 단백질을 발현시키거나, 천연 발생 단백질의 코딩 서열 모두 또는 이의 일부를 삭제하거나, 천연 발생 단백질의 발현을 억제하거나, 천연 발생 단백질의 발현을 향상시키거나, 비-천연 피분석물을

생성시키거나, 천연 피분석물의 생산을 억제하는 것을 포함하는, 당해 분야에 공지된 임의의 다양한 것일 수 있다. 예를 들어, 후보 세포의 라이브러리에서 하나 이상의 유전자에 대한 코딩 서열은 포인트 돌연변이, 결실(예를 들어, 전체 단백질 코딩 서열, 단백질의 도메인 또는 다른 부분의 제거), 또는 삽입(예를 들어, 키메라)을 함유할 수 있다.

- [0134] 라이브러리에 존재하는 후보 세포는 핵산 태그를 포함할 수 있다. 세포를 태그화하기 위한 여러 예시적인 방법은 실시예 IV에 기술되고/거나 하기에 기술된다.
- [0135] 일부 구현예에서, 핵산 태그는 세포의 표면에 공유적으로 부착될 수 있다. 일반적으로, 단지 단일 핵산 태그 서열과 함께 라이브러리에서의 각 세포를 태그화하기 위한 방법이 사용된다(각각이 동일한 태그 서열이 복사체를 갖는 다수의 핵산 분자는 각 세포에 또는 각 세포 상에 존재할 수 있다). 예를 들어, 세포는, 각 세포가 핵산 태그 분자와 개별적으로 반응될 수 있도록, 물리적으로 단리될 수 있다. 세포를 물리적으로 분리하기 위한 방법은 예를 들어, 개별 용기, 마이크로플레이트 상의 웰, 어레이 상의 피쳐, 비드, 액적 액츄에이터 장치에서의 유체 액적, 에멀전에서의 유체 액적, 또는 소낭으로의 분리를 포함한다.
- [0136] 핵산 태그가 전달될 수 있는 액적을 생성시키기 위한 특히 유용한 방법은 예를 들어, RainDance Technologies (매세추세즈주 빌레리카 소재)에 의해 상업화되거나 미국 특허 제9,017,623호 또는 제8,857,462호에 기술된 것을 포함하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 태그를 생성시키고 액적에 태그를 첨가하기 위한 다른 방법은 10X Genomics에 의해 상업화되거나 미국 특허 출원 공개 제2014/0155295 A1호; 제2014/0206554 A1호; 제2014/0227684 A1호 또는 제2014/0378322 A1호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 이러한 방법은, 세포(또는 본 명세서에 기술된 다른 후보 약제)가 개별 액적에 로딩되도록 개질될 수 있으며, 로딩된 액적은, 개별 액적이 결국 단일 태그 핵산 중 및 단일 후보 약제 종이 되도록 핵산 태그를 함유하는 유체와 상호작용된다. 물론, 핵산 중의 다수의 복사체 또는 후보 약제 중의 다수의 복사체는 개별 액적에 존재할 수 있다. 일부 구현예에서, 핵산 태그는 개별 액적에 전달되는 비드에 부착된다. 핵산 태그는 예를 들어, 실시예 IV에 기술된 부착 화학물질을 사용하여 세포의 표면에 부착될 수 있다.
- [0137] 세포(또는 다른 후보 약제)를 분리하기 위해, 예를 들어, 세포를 태그화하기 위해 사용될 수 있는 특히 유용한 액적 조작 장치는 예를 들어, 미국 특허 제8,637,242호, 미국 특허 제6,911,132호(발명의 명칭: "Apparatus for Manipulating Droplets by Electrowetting-Based Techniques," 2005년 6월 28일자로 발행); Pamula et al., 미국 특허 공개 제20060194331호(발명의 명칭: "Apparatuses and Methods for Manipulating Droplets on a Printed Circuit Board," 2006년 8월 31일자로 공개됨); Pollack et al., 국제특허공개번호 WO/2007/120241호(발명의 명칭: "Droplet-Based Biochemistry," 2007년 10월 25일자로 공개됨); Shenderov, 미국 특허 제6,773,566호(발명의 명칭: "Electrostatic Actuators for Microfluidics and Methods for Using Same," 2004년 8월 10일자로 발행됨); Shenderov, 미국 특허 제6,565,727호(발명의 명칭: "Actuators for Microfluidics Without Moving Parts," 2003년 5월 20일자로 발행됨); Kim et al., 미국 특허 공개 제20030205632호(발명의 명칭: "Electrowetting-driven Micropumping," 2003년 11월 6일자로 공개됨); Kim et al., 미국 특허 공개 제20060164490호(발명의 명칭: "Method and Apparatus for Promoting the Complete Transfer of Liquid Drops from a Nozzle," 2006년 7월 27일자로 공개됨); Kim et al., 미국 특허 공개 제20070023292호(발명의 명칭: "Small Object Moving on Printed Circuit Board," 2007년 2월 1일자로 공개됨); Shah et al., 미국 특허 공개 제20090283407호(발명의 명칭: "Method for Using Magnetic Particles in Droplet Microfluidics," 2009년 11월 19일자로 공개됨); Kim et al., 미국 특허 공개 제20100096266호(발명의 명칭: "Method and Apparatus for Real-time Feedback Control of Electrical Manipulation of Droplets on Chip," 2010년 4월 22일자로 공개됨); Velev, 미국 특허 제7,547,380호(발명의 명칭: "Droplet Transportation Devices and Methods Having a Fluid Surface," 2009년 6월 16일자로 발행됨); Sterling et al., 미국 특허 제7,163,612호(발명의 명칭: "Method, Apparatus and Article for Microfluidic Control via Electrowetting, for Chemical, Biochemical and Biological Assays and the Like," 2007년 1월 16일자로 발행됨); Becker et al., 미국 특허 제7,641,779호(발명의 명칭: "Method and Apparatus for Programmable Fluidic Processing," 2010년 1월 5일자로 발행됨); Becker et al., 미국 특허 제6,977,033호(발명의 명칭: "Method and Apparatus for Programmable Fluidic Processing," 2005년 12월 20일자로 발행됨); Decre et al., 미국 특허 제7,328,979호(발명의 명칭: "System for Manipulation of a Body of Fluid," 2008년 2월 12일자로 발행됨); Yamakawa et al., 미국 특허 공개 제20060039823호(발명의 명칭: "Chemical Analysis Apparatus," 2006년 2월 23일자로 공개됨); Wu, 미국 특허 공개 제20110048951호(발명의 명칭: "Digital Microfluidics Based Apparatus for Heat-exchanging Chemical Processes," 2011년 3월 3일자로 공개됨); Fouillet et al., 미국 특허 공개 제20090192044호(발명의 명칭:

"Electrode Addressing Method," 2009년 7월 30일자로 공개됨); Fouillet et al., 미국 특허 제7,052,244호(발명의 명칭: "Device for Displacement of Small Liquid Volumes Along a Micro-catenary Line by Electrostatic Forces," 2006년 5월 30일자로 발행됨); Marchand et al., 미국 특허 공개 제20080124252호(발명의 명칭: "Droplet Microreactor," 2008년 5월 29일자로 공개됨); Adachi et al., 미국 특허 공개 제20090321262호(발명의 명칭: "Liquid Transfer Device," 2009년 12월 31일자로 공개됨); Roux et al., 미국 특허 공개 제20050179746호(발명의 명칭: "Device for Controlling the Displacement of a Drop Between Two or Several Solid Substrates," 2005년 8월 18일자로 공개됨); 및 문헌[Dhindsa et al., "Virtual Electrowetting Channels: Electronic Liquid Transport with Continuous Channel Functionality," Lab Chip, 10:832-836 (2010)]에 기술된 바와 같은 액적 액츄에이터(droplet actuator)이며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

- [0138] 특정 구현예에서, 라이브러리에서의 세포는 핵산 태그를 포함하기 위해 유전적으로 개질될 수 있다. 예를 들어, 집단에서의 개별 세포는 태그 서열을 엔코딩하는 플라스미드를 수반하거나, 개별 세포의 게놈은 태그 서열을 포함하도록 개질될 수 있다. 일부 경우에, 태그 서열은 본 명세서에 기술된 방법에서 스크리닝되는 유전적 변형체를 또한 포함하는 핵산 작제물에 엔코딩된다.
- [0139] 상이한 세포의 라이브러리는, 각 개별 세포가 랜덤하게 지정된 태그 서열을 획득하도록 작제화될 수 있거나, 각 세포는 공지된 태그 서열을 포함하도록 개질될 수 있다. 예를 들어, 랜덤 태그화는, 스크리닝될 유전적 변형체를 또한 엔코딩하는 핵산 작제물을 합성할 때, 태그 영역에서 랜덤 뉴클레오타이드를 하나 이상의 부분에 도입함으로써 수행될 수 있다. 다른 예에서, 상이한 핵산 태그의 집단은 평균적으로 각 변형체 핵산이 독특한 태그 서열에 결합하도록, 상이한 변형체 엔코딩 핵산의 집단에 랜덤하게 결합될 수 있다. 이후에, 작제물은 라이브러리에서의 세포에 첨가될 수 있다. 유사하게, 상이한 핵산 태그의 집단은, 평균 각 세포가 독특한 태그 서열에 부착하도록 상이한 세포의 집단에 공유적으로 부착될 수 있다.
- [0140] 다른 구현예에서, 상이한 세포의 라이브러리는, 각 개별 세포가 공지된 세포와 선형적으로 관련이 있는 공지된 태그를 획득하도록 작제화될 수 있다. 선형적 태그화의 일 예로서, 특이적 태그는 그러한 작제물에서 돌연변이 또는 변형체와 연관된 방식으로 핵산 작제물로 합성될 수 있다. 일부 구현예에서, 세포는 서로 물리적으로 분리되며, 공지된 태그 서열을 갖는 핵산은 단세포와 단일 타입의 태그의 공유 부착을 형성하기 위해 세포와 접촉된다.
- [0141] 특정 구현예에서, 핵산 태그는 비드에 부착될 수 있으며, 비드는 세포에 결합할 수 있다. 선택적으로, 비드는 특이적 결합 친화력을 갖는 항체 또는 세포에 부착될 수 있다. 그러나, 다른 부착 양식은, 예를 들어, 후보 약제 또는 고체 지지체에 핵산을 부착시키는 상황에서 본 명세서에 예시된 것을 포함하는, 세포에 비드를 부착하기 위해 사용될 수 있다.
- [0142] 일부 구현예에서, 핵산 태그는 세포의 원형질 막 지질 또는 지방산에 핵산 태그의 공유 부착에 의해 세포에 부착될 수 있다. 대안적으로, 핵산 태그는 세포의 원형질 막 지질에서 단백질에 대한 공유 부착에 의해 세포에 부착될 수 있다. 공유 부착에 대한 대안예로서, 핵산 태그는 세포 표면 상의 리간드에 결합하는 수용체를 포함할 수 있거나, 핵산 태그는 세포 표면 상에 수용체에 결합하는 리간드를 포함할 수 있다. 예시적인 부착 방법은 하기 실시예 IV에 기술되어 있다.
- [0143] 세포의 라이브러리는 고체 지지체에 다른 타입의 후보 약제를 부착시키기 위한 본 명세서에 기술된 방법 중 하나 이상을 이용하여 어레이에 부착될 수 있다. 예를 들어, 고체 지지체는 핵산 프라이머를 포함할 수 있으며, 세포는 핵산 프라이머에 대한 핵산 태그의 하이브리드화를 통해 고체 지지체에 부착할 수 있다. 일부 경우에, 핵산 태그는 유니버설 프라이머 결합 서열을 포함할 수 있으며, 핵산 프라이머는 유니버설 프라이머 서열을 포함할 수 있으며, 후보 약제는 유니버설 프라이머 서열에 대한 유니버설 프라이머 결합 서열의 하이브리드화를 통해 고체 지지체에 부착할 수 있다.
- [0144] 본 발명의 방법은 핵산, 예를 들어, 핵산 태그가 하이브리드화되는 고체 지지체-부착된 프라이머를 확장시키는 단계를 포함할 수 있다. 얻어진 확장된 프라이머는 핵산(그림에도 불구하고, 상보적 형태)으로부터의 태그 서열 및 다른 서열을 포함할 것이다. 이에 따라, 확장된 프라이머는 후보 약제로부터 핵산의 공간적으로 태그화된 버전이다. 핵산에 존재하는, 태그 서열 이외의 서열 엘리먼트가 또한 확장된 프라이머에 포함되는 것으로 이해될 것이다. 이러한 엘리먼트는 예를 들어, 프라이머 결합 부위, 분열 부위, 다른 태그 서열(예를 들어, 샘플 식별 태그), 포획 서열, 핵산 결합 단백질 또는 핵산 효소에 대한 인식 부위, 등을 포함한다.

- [0145] 프라이머의 확장은 핵산의 증폭 또는 핵산의 시퀀싱을 위해 본 명세서에 예시되거나 달리 당해 분야에 공지된 방법을 이용하여 수행될 수 있다. 특정 구현예에서, 하나 이상의 뉴클레오타이드는 폴리머라제 촉매작용(예를 들어, DNA 폴리머라제, RNA 폴리머라제 또는 역전사 효소)을 통해, 프라이머의 3' 단부에 첨가될 수 있다. 화학적 또는 효소적 방법은 프라이머의 3' 또는 5' 단부에 하나 이상의 뉴클레오타이드를 첨가하기 위해 사용될 수 있다. 하나 이상의 올리고뉴클레오타이드는 예를 들어, 화학적 또는 효소적(예를 들어, 리가제 촉매작용) 방법을 통해, 프라이머의 3' 또는 5' 단부에 첨가될 수 있다. 프라이머는 주형 유도 방식으로 확장될 수 있으며, 이에 의해 확장의 생성물은 프라이머에 하이브리드화되는 주형 핵산에 대해 상보적이다. 일부 구현예에서, DNA 프라이머는 RNA 주형을 이용하여 역전사 효소에 의해 확장되며, 이에 의해 cDNA를 생성시킨다. 이에 따라, 본 명세서에 기술된 방법에서 제조된 확장된 프로브는 역전사된 DNA 분자일 수 있다. 핵산을 확장시키기 위한 예시적인 방법은 미국 특허 출원 공개 US 2005/0037393 A1호 또는 미국 특허 제8,288,103호 또는 제8,486,625호에 기술되어 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.
- [0146] 프라이머에 하이브리드화되는 모든 핵산 또는 핵산의 일부는 확장에 의해 복사될 수 있다. 예를 들어, 확장된 프로브는 핵산으로부터 복사된 적어도 1, 2, 5, 10, 25, 50, 100, 200, 500, 1000개 또는 그 이상의 뉴클레오타이드를 포함할 수 있다. 확장 생성물의 길이는 예를 들어, 확장 반응에서 가역적으로 중결된 뉴클레오타이드를 사용하고 제한된 수의 확장 사이클을 진행시켜 조절될 수 있다. 사이클은 SBS 기술에 대해 예시된 바와 같이 실행될 수 있으며, 표지된 뉴클레오타이드의 사용은 필수적인 것은 아니다. 이에 따라, 본 명세서에 기술된 방법에서 생성된 확장된 프라이머는 핵산으로부터 복사된 1000, 500, 200, 100, 50, 25, 10, 5, 2 또는 1개 이하의 뉴클레오타이드를 포함할 수 있다. 물론, 확장된 프로브는 상기 기술된 범위 내에 또는 이를 벗어난 임의의 길이일 수 있다.
- [0147] 어레이에 부착된 세포의 라이브러리는 다른 후보 약제를 스크리닝하는 것과 관련하여 본 명세서에 기술된 방법을 이용하여 스크리닝될 수 있다. 특정 구현예에서, 세포의 어레이의 스크리닝은 세포를 스크리닝제로 처리하는 단계를 포함할 수 있다. 결과적으로, 스크리닝제는 어레이 상에서 적어도 하나의 후보 세포에 결합할 수 있다. 선택적으로, 스크리닝제는 발광성이며, 스크리닝 반응은 적어도 하나의 후보 세포의 발광을 검출함으로써 수행될 것이다.
- [0148] 일부 구현예에서, 어레이로 전달되는 스크리닝제는 어레이에서 적어도 하나의 후보 세포를 개질시킬 것이다. 예를 들어, 스크리닝제는 어레이 상에서 적어도 하나의 후보 세포를 자극시킬 수 있다. 대안적으로, 스크리닝제는 어레이 상에서 적어도 하나의 후보 세포를 억제하거나 심지어 사멸시킬 수 있다.
- [0149] 다수의 세포와 접촉된 스크리닝제는 적어도 하나의 후보 세포의 발광을 증가시키거나 감소시킬 수 있으며, 스크리닝 반응은 적어도 하나의 후보 세포의 발광을 검출하는 단계를 포함할 수 있다.
- [0150] 본 명세서에서의 다른 곳에 기술된 바와 같이, 검출 단계는 동력학 또는 시간 기반 측정을 포함할 수 있다. 예를 들어, 다수의 세포의 검출은 어레이 상의 개별 피쳐 중 하나 이상에 대해 여러 시점에서 신호를 획득하는 단계를 포함할 수 있다.
- [0151] 세포에 부착되거나 세포에 존재하는 핵산 태그는 본 명세서에 기술된 방법에서 시퀀싱될 수 있다. 세포에 부착된 핵산은 세포 상에서 시퀀싱될 수 있다. 대안적으로, 핵산은 실시예 IV에 예시된 것과 같은 프라이머 확장 방법을 통해 고체 지지체로 옮겨질 수 있으며, 태그 서열(또는 이의 보체)을 함유한 확장된 프라이머의 영역은 고체 지지체 상에서 시퀀싱될 수 있다. 본 명세서에 다른 곳에 기술된 고체 지지체-부착된 핵산을 시퀀싱하는 방법은 핵산 태그가 유도되는 세포에 가까운 피쳐 또는 세포에 부착된 핵산 태그를 위해 사용될 수 있다.
- [0152] 일부 경우에, 고체 지지체에 부착된 핵산 태그(또는 태그의 복사체)를 남기면서, 고체 지지체로부터 세포를 제거하는 것이 요망될 수 있다. 이후에, 태그 서열은 세포의 부재 하에 시퀀싱될 수 있다.
- [0153] 핵산 태그가 세포에 존재하는 구현예에 대하여, 세포는 어레이의 표면 상에서 용리되어, 세포 함유물의 배출 및 핵산 태그의 국소화된 포획을 야기시킬 수 있다. 핵산 태그는 고체 지지체 상의 프라이머에 대해 상보적인 유니버설 서열 영역을 포함할 수 있다. 이에 따라, 핵산 태그는 증폭 및/또는 시퀀싱하기 용이한 포맷에서 고체 지지체 상에 포획될 수 있다. 고체 지지체 상에 포획된 핵산은 선택적으로 증폭될 수 있다. 고체 지지체 상에 포획된 핵산 또는 이로부터 생성된 앰플리콘은 본 명세서에 기술된 방법을 이용하여 고체 지지체 상에서 시퀀싱될 수 있다.
- [0154] 본 발명의 방법은 어레이로부터 하나 이상의 세포를 제거하는 단계를 포함할 수 있다. 예를 들어, 스크리닝 단계에서 히트로서 식별된 하나 이상의 세포는 선택적으로 제거될 수 있다. 선택적으로, 세포의 생존능력

(viability)를 보유하는 조건이 사용될 수 있다. 이와 같이, 제거된 세포(들)는 배양되거나, 복사되거나, 확장될 수 있다. 세포(들)는 다른 후보 약제에 대해 본 명세서에서의 다른 곳에 기술된 기술 및 시약을 사용하여 제거될 수 있다. 예를 들어, 광-불안정한 링커 및 공간적으로 필터링된 광 빔의 사용은 어레이에 다른 세포로부터 특정 세포를 단리시키기 위해 특히 유용할 수 있다.

[0155] 일부 구현예에서, 생존 세포는 스크리닝 단계 이후에 어레이로부터 제거될 필요는 없다. 오히려, 태그 서열은, 태그의 시퀀싱이 추가로 세포를 단리하거나 특성화할 필요 없이, 세포를 식별하기에 충분하도록, 선형적으로 공지된 세포와 결합될 수 있다. 특정 피처에 존재하는 태그의 선형적 지식은 세포를 식별하기 위해 스크리닝 단계에서 피처로부터 관찰될 신호와 연관될 수 있다.

[0156] 하기 실시예는 본 발명을 예시하기 위한 것으로서 제한하고자 의도된 것은 아니다.

[0157] **실시예 I**

[0158] **소분자 스크리닝**

[0159] 차세대 시퀀싱 플랫폼, 예를 들어, Illumina(캘리포니아주 샌디에이고 소재)에 의해 상업화된 것은 약물 발견을 위한 통합형 고처리량 스크리닝 플랫폼을 형성시키기 위한 기반을 제공한다. 핵산을 표적 분자에 부착을 가능하게 하는 작용기로 개질시킴으로써, 고처리량 화합물 스크리닝을 위한 어레이를 형성시키기 위해 기질로서 시퀀싱 흐름셀을 사용하는 것이 가능하다. 다양한 스크리닝 검정은 부착된 표적에 대한 변경과 함께 동일한 플랫폼을 사용하여 실행될 수 있다. 시퀀싱 플랫폼 상에서 소분자를 스크리닝하기 위한 예시적인 방법은 도 1a 내지 도 1f에 도식화되어 있고, 하기에 기술되어 있다.

[0160] 도 1a는 각각이 29개 뉴클레오타이드 길이인, 50,000개의 상이한 핵산 분자를 테이프-기반 합성 기기 상에서 합성하는 제1 단계를 도시한 것이다. 테이프는 개별 핵산 종이 각각 합성되는 개별적으로 다루어질 수 있는 부위를 포함한다. 핵산 분자는 또한, 5' 단부에 설프히드릴, 아민 또는 N-히드록시숙신이미드 기와 같은 작용기(FG)를 포함한다. 핵산의 5' 영역은 유니버설 프라이머(50,000개의 핵산 모두에 대해 동일함)를 엔코딩하는 10개의 뉴클레오타이드를 포함하며, 3' 영역은 50,000개의 상이한 태그 서열(또한, "코드" 서열로 지칭됨) 중 하나를 엔코딩한다. 예시적인 테이프-기반 DNA 합성 기기는 미국 특허 출원 공개 제2011/0178285 A1호에 기술되어 있으며, 이러한 문헌은 본 명세서에 참고로 포함된다.

[0161] 제2 단계는 도 1b에 도시되어 있는데, 여기서, 모노주형 하이브리드 핵산(monotemplate hybrid nucleic acid)을 합성 테이프의 각 웰에서 핵산에 결합시켜 유니버설 오버행 프라이머를 생성시킨다. 다이어그램에 도시되어 있는 바와 같이, 모노주형 하이브리드 상에서 오버행과 상이한 핵산의 유니버설 프라이머 서열의 하이브리드화 후에 결합이 일어난다. 선택적으로, 상이한 핵산은 결합 사건 이전에 (예를 들어, 핵산을 정제하거나 개질시키기 위해) 테이프로부터 제거될 수 있다. 다른 선택적 구성은 상이한 핵산 대신에 모노주형 하이브리드에 부착된 작용기를 갖는 것을 포함한다. 대안적으로 또는 추가적으로, 모노주형 하이브리드는 바이오틴, 아자이드 또는 알킨기와 같은 표면 부착 모이어티를 가질 수 있다.

[0162] 도 1c에 도시된 바와 같이, 제3 단계를 합성기 테이프 상에서 개별 웰 각각에 특정 후보 약제(예를 들어, "화합물")를 로보트로 첨가하도록 수행한다. 후보 약제 각각은 상이한 핵산 상의 작용기와 반응하는 반응기를 포함한다. 결과적으로, 후보 약제 각각은 합성기 테이프 상에서 특정 핵산 태그에 부착되며, 이에 의해, 엔코딩된 후보 약제를 생산한다.

[0163] 도 1d에 도시된 바와 같은, 단계 4는 흐름셀 표면 상에서 구별 가능한 부위에서 엔코딩된 후보 약제를 고정하도록 수행된다. 고정화는 흐름셀 표면 상에 있는 상보적 유니버설 프라이머에 대한 핵산의 모노주형 부분의 하이브리드화를 통해 일어날 수 있다. 프라이머는 예를 들어, 미국 특허 출원 공개 제2011/0059865 A1호, 또는 미국 특허 제9,012,022호에 기술된 바와 같이, PAZAM 겔 또는 다른 하이드로겔을 통해 흐름셀 표면에 부착될 수 있으며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0164] 도 1e에 도시된 바와 같이, 단계 5에서, 스크리닝제(또한 "표적 분자"로 지칭됨)는 흐름셀의 표면 상에 있는 고정화된 엔코딩된 후보 약제와 접촉된다. 스크리닝제는 Illumina 시퀀싱 기기와 같은 광학 장치를 이용하여 검출될 수 있는 형광 라벨을 포함한다. 이미지는 예를 들어, 구별 가능한 부위 각각에서 스크리닝제에 대한 결합 동력학을 결정하기 위해 실시간으로 얻어질 수 있다. 도 1e의 예에서, 코드 N에 부착된 후보 약제는 요망되는 결합 동력학 프로파일에 의해 결정되는 바와 같이 "히트"이다.

[0165] 도 1f에 도시된 바와 같이, 코드 N에 대한 서열은 흐름셀 상에서 수행되는 시퀀싱 프로토콜을 기초로 하여 결정

될 수 있다. 코드 N에 대한 서열은 이러한 서열을 합성하는 합성기 테이프 상의 위치와 연관될 수 있으며, 또한, "히트"의 동일성은 후보 약제가 코드 N이 합성되는 합성기 테이프 상의 위치로 전달되는 지식을 기초로 하여 결정될 수 있다.

[0166]

실시에 II

[0167]

단백질 스크리닝

[0168]

고체상 증폭 방법은 핵산의 고도로 다중화된 프리젠테이션을 가능하게 한다. 특히 유용한 고체상 증폭 방법은 예를 들어, 미국 특허 제5,641,658호; 미국 특허 출원 공개 제2002/0055100호; 미국 특허 제7,115,400호; 미국 특허 출원 공개 제2004/0096853호; 미국 특허 출원 공개 제2004/0002090호; 미국 특허 출원 공개 제2007/0128624호; 및 미국 특허 출원 공개 제2008/0009420호에 기술된 바와 같이 수행될 수 있는 브릿지 증폭(또한 클러스터 형성으로 지칭됨)이며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0169]

이러한 실시에는 흐름셀 상에 단백질 어레이를 형성시키기 위해 고체상 증폭으로부터 형성되는 핵산 클러스터를 사용하는 것을 기술하는 것이다. 이러한 기술에 대한 일부 중요한 장점은 (a) 이는 매우 다양한 상이한 단백질이 활성화에 대해 스크리닝될 수 있도록 표면 상에서 비교적 작은 크기를 갖지만 광학적으로 구별 가능한 다수의 단백질-함유 피처를 제공하고; (b) 피처가 단백질을 엔코딩하는 핵산을 함유하며, 핵산이 시퀀싱되기 용이하여, 흐름셀의 각 피처에서 단백질이 식별될 수 있다는 것이다.

[0170]

도 2에 도식화되어 있는 바와 같이, cDNA 라이브러리는 전사 및 번역 부위와 함께 P5 및 P7 프라이머 결합 부위를 갖는 어댑터를 갖도록 구성된다. 라이브러리는 흐름셀에 부착되며, 라이브러리 일원은 클러스터를 형성하기 위해 브릿지 증폭을 이용하여 흐름셀의 표면 상에서 증폭된다. P5 및 P7 프라이머 결합 서열, 어댑터를 갖는 핵산 라이브러리를 제조하는 방법, 및 클러스터를 제조하는 방법을 설명하기 위해, 예를 들어, 브릿지 증폭 기술과 관련하여 상기에서 인용된 문헌, 문헌[Bentley et al., *Nature* 456:53-59 (2008)], 미국 특허 출원 공개 제2011/0059865 A1호 및 미국 특허 제7,741,463호가 참조되며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다.

[0171]

연어진 cDNA 클러스터는 RNA 폴리머라제를 사용하여 mRNA로 전사된다. mRNA는 cDNA 클러스터 가까이 유지되며, 이로부터, 이는 (a) RNA 폴리머라제 stall 및 mRNA 전사체가 cDNA에 결합된 채로 유지되도록 cDNA 작제물의 단부에서 전사 종결 부위를 가지지 않음으로써, (b) RNA 폴리머라제를 유사하게 정지시키는 공지된 전사 휴지 부위(예를 들어, *E. coli*의 Trp 오페론으로부터)의 사용에 의해, (c) 전사를 휴지하기 위한 회학적 방법(예를 들어, 비시클로마이신)을 사용함으로써, 또는 (d) 전사된 RNA가 하이브리드화하는 cDNA 클러스터에 상보적 서열을 가짐으로써 전사된다.

[0172]

전사되는 cDNA 클러스터에 국소화된 mRNA는 이후에 번역되어 단백질을 제조한다. 이는 mRNA에서 출발 코돈에 의해 촉진될 것이다(이는 cDNA 라이브러리 작제물의 5' 단부에서 어댑터에 전사 개시 서열과 함께 도입될 수 있다). 또한, mRNA는 3' 단부에서 정지 코돈을 가지지 않을 것이다. 도 2에 도식화된 바와 같이, 정지 코돈의 결여는 mRNA의 단부에서 정지하는 리보솜을 야기시키며, 이에 따라, 단백질은 리보솜을 통해 mRNA에 부착된 채로 유지된다(mRNA는 또한, 정지된 RNA 폴리머라제에 의해 cDNA 클러스터에 부착된다). 대안적으로, 리보솜은 화학적 수단(예를 들어, 클로람페니콜/푸로마이신)에 의해 stall 될 수 있다. 커플링된 전사 및 번역은 용이하게 이용 가능한 토끼 망상적혈구 용해질, 박테리아 S30 또는 맥아 추출물 시스템을 사용하여 수행될 수 있다.

[0173]

대안적인 방법에서 단백질은 용액 중에서 발생되고, 이후에, 흐름셀 또는 다른 표면에 부착될 수 있다. 이러한 방법이 표면 상에서 증폭을 포함하지 않지만, 용액 기반 기술이 더욱 용이하게 입수 가능하거나 평가될 단백질에 적용 가능할 때 유익할 수 있다. 단백질의 연어진 어레이는 히트를 식별하기 위해 스크리닝제 또는 다른 자극제로 스크리닝된다. 클러스터에서의 cDNA는 단백질 히트를 식별하기 위해 서열화된다.

[0174]

실시에 III

[0175]

태그를 사용한 단백질 스크리닝

[0176]

차세대 시퀀싱 플랫폼, 예를 들어, Illumina(캘리포니아주 샌디에이고 소재)에 의해 상업화된 것은 단백질 진화를 위한 통합된 고처리량 스크리닝 플랫폼을 형성시키기 위한 기반을 제공한다. 예시적인 방법은 도 3a 및 도 3b에 도식화된다.

[0177]

제1 단계로서, 벡터의 라이브러리는 고려되는 표적 단백질 및 랜덤 태그(또는 "코드"로 지칭됨)에 대한 랜덤 또는 반-랜덤 돌연변이를 갖도록 생성된다. 벡터는 단백질 및 태그의 코딩 서열을 포함하는 mRNA를 발현시키도록 구성된다. 상이한 랜덤 태그의 수는, 각 변형체 mRNA가 독특한 태그를 포함하는 높은 확률을 갖도록 라이브러리

에서의 단백질 변형체의 수를 초과한다. 예를 들어, 10개의 뉴클레오타이드의 랜덤화된 서열은 10^6 개의 상이한 태그를 제공할 것이며, 20개의 뉴클레오타이드의 랜덤화된 서열은 10^{12} 개의 상이한 태그를 제공할 것이다.

- [0178] 돌연변이의 랜덤 도입 및 랜덤 태그의 사용에 대해 대안적으로, 제1 단계는 사전규정된 돌연변이가 이루어지고 공지된 태그와 결합되는 방식으로 수행될 수 있다.
- [0179] 제2 단계에서, 벡터와 함께 변형된 박테리아의 라이브러리는 배양되어 개개 벡터로부터 mRNA를 생성시킨다. 라이브러리는 라이브러리(즉, mRNA 라이브러리)에서 다양한 벡터로부터 mRNA 전사체의 혼합물을 방출하도록 용리된다.
- [0180] 제3 단계에서, mRNA 라이브러리는 분할되어 두 개의 서브-라이브러리를 생성시킨다. 라이브러리는 라이브러리의 분할이 두 개의 서브-라이브러리에 존재하는 동일한 일원을 야기시키도록 각 일원의 다수의 복사체를 포함한다. 또한, 독특한 태그 서열은 각 단백질 변형체 서열에 부착된다.
- [0181] 제4 단계에서, 도 3b에 도시된 바와 같이, 제1 서브-라이브러리는 흐름셀에 부착되고, 흐름셀 상에 증폭되며, 코딩 서열 및 얻어진 앰플리콘의 개개 태그는 서열화된다. 이러한 방식으로, 각 일원이 부착되는 부위가 위치된다. 이후에, 제2 서브-라이브러리의 일원은 흐름셀과 접촉되어 제2 서브-라이브러리로부터의 일원을 태그 서열의 상보성을 통해 제1 서브-라이브러리로부터의 동일한 일원으로부터 발생된 앰플리콘에 하이브리드화시킬 수 있다. 이러한 방식으로, 제2 서브-라이브러리의 각 일원이 위치한 부위는 시퀀싱 결과 및 태그와 이의 보체 간의 예상된 하이브리드화를 기초로 하여 추론될 수 있다. 흐름셀 상에 포획된, 제2 서브-라이브러리의 일원은 이후에, 흐름셀 상에서 단백질 표적을 발현시키도록 번역된다. 단백질 각각은 번역된 RNA에 부착된 채로 유지되는 방식으로 발현된다. 예를 들어, 부착은 실시예 II에 기술된 기술을 이용하여 달성될 수 있다. 단백질은 이후에, 스크리닝제 또는 다른 자극제로 스크리닝되어, 고처리량 방식으로 요망하는 기능을 식별할 수 있게 한다. 결과는 단백질을 엔코딩하는 RNA 서열과 연관된 흐름셀 상의 부위에 국소화된 단백질 활성 신호이다.
- [0182] 이러한 방법은 흐름셀 레인 당 3천만개 이상의 단백질의 고처리량을 가능하게 할 수 있다. 이러한 스크리닝 과정은 에멀전 기반 스크리닝과 관련된 문제를 피한다. 예를 들어, 정량적 및 동력학-기반 데이터는 전통적인 에멀전 기반 기술을 이용하여 이용 가능한 것 보다 더욱 정제되고 선택적인 스크리닝 기준을 가능하게 하기 위해 Illumina 시퀀싱 기기를 이용하여 흐름셀 상의 각 피처에서 수득될 수 있다.

실시예 IV

세포 스크리닝

- [0185] 본 실시예는 집단에서 개별 세포를 식별하기 위한 차세대 시퀀싱의 장점을 부가하면서, 통상적으로 형광 현미경법과 관련된 장점인, 실시간으로 동적 공정을 따르는 능력과 함께 형광 활성화된 세포 분류(FACS)의 고처리량 장점을 제공하는 세포를 스크리닝하기 위한 방법을 기술한다. 결과적으로, 본 명세서에 기술된 방법은 요망되는 거동을 디스플레이하는 세포의 회수 및 특성화를 가능하게 한다. 상세하게, 본 실시예는 초-고처리량에서 특정 자극제에 반응하는 형광, 표현형 거동(예를 들어, 형광 리포터의 발현에 의해 측정함)을 모니터링하고, 이후에, 후속 식별 및 회수를 위한 각 개별 세포의 표면 상에 디스플레이된 태그의 시퀀싱에 의해 모니터링하기 위해 Illumina Inc.(캘리포니아주 샌디에이고 소재)로부터의 차세대 시퀀싱 플랫폼을 구성한다. 이러한 플랫폼은 세포 조작을 위한 고처리량 스크리닝을 위한 표준 방법을 제공할 수 있다.
- [0186] 세포는 소분자 및 단백질의 것과는 상이하고, 다수의 경우에, 이러한 것에 비해 개선된 잠재적인 치료 능력을 갖는다. 세포는 신체에서 특정 위치로 능동적으로 이동할 수 있으며, 이러한 것은 다수의 외부 자극제를 감지할 수 있고, 이러한 것은 정확한 출력으로 반응하는 다수의 정보 소스를 통합할 수 있다. 또한, 세포-기반 반응은 사전-프로그래밍된 기능에 의해 또는 외부 인자의 부가에 의해 조절될 수 있는 복잡한 동적 패턴을 갖는다. 현재, 세포-기반 치료법은 크론병을 치료하기 위한 조작된 미생물의 사용에서 암을 치료하기 위한 환자-유래 조작된 인간 면역 세포의 사용에 이르기까지의 광범위한 질환의 치료를 위해 연구되고 있다. 중요하게, 미생물 또는 인간 세포 중 어느 하나는 치료학적으로 유용하도록 조작될 수 있다.

[0187] **세포 조작을 위한 워크플로우**

- [0188] 세포 조작(미생물 또는 인간 세포 중 어느 하나에서)은 고려될 여러 변수를 갖는 복잡한 공정이다. 이에 따라, 단일 디자인을 시도하기 보다는 오히려, 이후에 선택되거나 스크리닝될 수 있는 후보 세포의 라이브러리를 디자인하는 것이 유익하다. 상세하게, 세포 조작은 하기 단계를 포함할 수 있다: (1) 외부 신호에 대한 하나 이상의

센서, 정보 통합 네트워크, 및 신호-의존 반응(반응은 종종 유전자 발현을 수반함)을 포함할 수 있는 유전 회로의 설계; (2) (예를 들어, CRISPR-Cas9를 사용하여) 후보 세포의 라이브러리를 가로질러 상기 언급된 구성요소에 대해 하나 이상의 유전 표적을 개질화시킴; 및 (3) 요망되는 세포 거동을 부여하는, 여러 상이한 후보 세포들 중에서, 히트를 식별하기 위한 고처리량 스크리닝.

[0189] 강력하고 고처리량의 스크리닝 기술은 시험될 필요가 있는 세포의 수가 많고(종종, 수천개의 상이한 디자인이 시도됨) 세포 거동의 크고 다양한 집합이 조작 공정에서 도입된 개질로 인하여 나타나기 때문에, 세포 조작에 대한 상당한 장점을 제공할 수 있다. 요망되는 스크린은 또한, 매우 고처리량을 달성하면서 복잡한 시간적 동역학을 따를 수 있을 것이다. 본 명세서에 기술된 스크린은 이러한 장점을 제공한다.

[0190] 세포 스크리닝 방법은 하기 2 스테이지를 갖는다:

[0191] **스테이지 1: 표현형 모니터링.** 적합한 형광 리포터를 갖는 살아있는 조작된 세포의 라이브러리는 도 4에 도시된 바와 같이, 시퀀싱 플랫폼(예를 들어, Illumina MiSeq®, NextSeq®, HiSeq® 또는 Genome Analyzer® 플랫폼)을 위해 흐름셀에 로딩된다. 개별 세포는 랜덤하게 분포되고(분석될 특정 세포에 따라 특정 부착 화학물질을 사용하여) 흐름셀의 표면에 부착될 것이다. 각 개별 세포의 형광은 시퀀싱 동안 Illumina 흐름셀에 대해 현재 수행되는 바와 같이 흐름셀 표면을 스캐닝하여, 규정된 시간 간격으로 기록된다. 개별 세포가 현 시퀀싱 클러스터와 유사한(또는 이보다 더 큰) 평균 크기를 가질 것이며, 이에 의해 시퀀싱 플랫폼의 해상도를 스크린에 대해 매우 적합하게 만든다. 또한, 세포는 수 천개의 형광 리포터 분자를 운반할 수 있으며, 이에 따라, 신호 세기는 플랫폼을 시퀀싱함으로써 현재에 검출된 것과 유사(또는 보다 큰)할 것이다. 특정 시간에, 세포는 스크리닝제(또는 다른 자극제)에 노출되며, 세포의 반응(예를 들어, 리포터 유전자의 발현의 변화, 또는 임의의 다른 타입의 세포 리포터, 예를 들어, 세포질에서의 칼슘 농도, 하위세포 국소화에서의 변화에 의한 형광체 켜짐, 등의 형광 세기의 변화에서 반영되는 바와 같이)은 검출된다. 시간에 따른 형광 리포터의 변화를 모니터링함으로써, 다양한 자극(또는 자극의 조합)에 대해 반응에서 표현형 변화의 시간적 동역학을 결정하는 것이 가능할 것이다.

[0192] **스테이지 2: 개별 태그의 디코딩에 의한 단세포 식별.** 집단에서 각 세포는 태그를 지정할 것이다. 세포를 태그화하기 위한 세 가지 상이한 기술은 하기에 기술된다. 제1 단계에서, 태그는 개별 세포에 부착된다. 태그 타입(예를 들어, 비드, 막-부착 핵산, 등)에 독립적으로, 세포는 먼저, 하나의 세포가 각 용기에 위치되도록, 개별 용기(예를 들어, 마이크로웰)에 분류된다. 이후에, 독특한 태그는 도 5에 도시된 바와 같이 각 용기에서 세포에 제공된다.

[0193] **2.1. 비드로 단세포 바코딩.** 집단에서의 각 세포(예를 들어, T 세포의 경우 ~10 μ m 직경)는 1 내지 10개의 비드로 태그화된다. 비드는 약 1 내지 2 μ m의 직경을 가질 수 있다. 유용한 비드는 예를 들어, Illumina's BeadArray Technology에서 사용되거나 미국 특허 제7,622,294호 또는 제8,741,630호에 기술된 것을 포함하며, 이러한 문헌 각각은 본 명세서에 참고로 포함된다. 각 세포는 이전에 개별 세포를 개별 용기에 분리시키고 각 용기에 동일한 독특한 태그를 분할하는 비드를 첨가함으로써, 단일 태그를 엔코딩하는 비드로 태그화될 것이다(즉, 각 용기와 각 태그 사이에 공지된 관련성(correspondence)이 존재할 것이다). 세포 태그화는 다수의 방식으로, 예를 들어, 표적 세포의 표면 상에서 자연적으로 존재하는 특이적 에피토프를 인식하는 항체와 비드를 공유적으로 연결시킴으로써 달성될 수 있다. 이러한 방식으로, 각 비드-태그화된 세포가 흐름셀의 표면 상에 랜덤하게 분포되지만, 각 세포가 자체 비드(들)로 공간적으로 공동-국소화될 것이라는 것이 주지된다(도 6). 이에 따라, 스테이지 1 이후에, 각 세포의 동일성은 흐름셀 상에서 비드(들)를 디코딩하고 상응하는 비드(들)에 가장 가까운 세포에 식별된 태그를 지정함으로써 나타낼 것이다. 비드 상의 태그는 합성 기술[예를 들어, 문헌[Bentley et al., *Nature* 456:53-59 (2008)], PCT 공개번호 WO 91/06678호, WO 04/018497호 또는 WO 07/123744호; 미국 특허 제 7,057,026호, 제7,329,492호, 제7,211,414호, 제7,315,019호 또는 제7,405,281호, 및 미국 특허 출원 공개 2008/0108082 A1호 참조, 문헌들 각각은 본 명세서에 참고로 포함됨] 또는 미국 특허 제8,460,865호 또는 문헌[Gunderson et al., *Genome Research* 14:870-877 (2004)](이러한 문헌 각각은 본 명세서에 참고로 포함됨)에 기술된 디코딩 기술에 의한 시퀀싱을 이용하여 디코딩될 수 있다.

[0194] **2.2. 서열 태그로 단세포 식별.** 비드 태그화에 대한 대안 또는 추가는 각 세포의 표면 상에 직접적으로 나타낸 독특한 서열 식별자(sequence identifier)를 사용하는 것이다. 그러한 것을 위하여, 집단에서 각 세포는 흐름셀에 세포 로딩 전에(예를 들어, 스테이지 1 이전에) 표면-디스플레이된 개별 단일-가닥 DNA 태그로 개질될 것이다. 핵산 태그 부착 화학은 특정 세포 타입에 따를 것이다. 예를 들어, 인간 세포는 문헌[Sleden et al., *J Am Chem Soc* 134: 765-8 (2012)]에 의해 보고된 바와 같이, 원형질 막 지질에 공유적으로 연결된 단일 가닥 DNA 분자로 태그화될 수 있으며, 이러한 문헌은 본 명세서에 참고로 포함된다. 다른 예로서, 효모 세포는 Aga1-Aga2

세포벽 착물 상에 디스플레이된 HaloTag 단백질에 공유적으로 연결된 단일 가닥 DNA 분자로 태그화될 수 있다.

- [0195] 각 개별 태그는 적어도 세 개의 영역으로 이루어진다. 영역 1은 부착 모이어티(예를 들어, 세포 타입에 따라, 지질 또는 단백질)에 연결된다. 영역 2는 요망되는 플렉시티(plexity)에 따르는 규정된 수의 염기로 이루어진다 (모두 4개의 뉴클레오타이드를 사용하여, 길이 "n"의 서열에 대한 4^n 조합을 발생시키는 것이 가능하며, 예를 들어, 4개 염기-길이 서열은 256개의 상이한 개별 세포의 식별을 위한 태그를 포함하며, 5는 1024개의 상이한 태그를 야기시키는 등등을 한다). 이에 따라, 9-염기 태그 서열만큼 적은 경우에, 250,000개 초과 상이한 세포를 개별적으로 식별하는 것이 가능할 것이다. 영역 3은 시퀀싱 프라이머에 대해 상보적인 DNA 서열로 이루어진다. 각 세포가, 미리 용기에 개별 세포를 분리시키고 각 용기에 동일한 독특한 태그 서열을 분할시키는 핵산을 첨가함으로써, 단일 태그를 인코딩하는 핵산으로 태그화될 것이라는 것이 주지된다(이전과 같이, 각 용기와 각 태그 간의 공지된 관련성이 존재할 것이다).
- [0196] 별도의 단계에서(표현형 측정이 수행되기 전 또는 후에 수행됨), 각 개별 세포의 동일성은 예를 들어, 하기에 기술된, 두 가지 방식 중 하나에서 결정될 것이다.
- [0197] **2.2.1. 각 세포에 부착된 특정 태그의 인시튜 시퀀싱.** 각 세포는 특정 태그의 다수의 복사체를 나타낼 것이다. 이에 따라, 신호는 시퀀싱 기기 상에서 용이하게 검출 가능하다(도 7 참조). 또한, 각 태그를 식별하기 위한 판독 길이는 상업적으로 입수 가능한 시퀀싱 기기로부터 통상적인 수백 뉴클레오타이드의 판독 길이 내에서 비교적 짧다(예를 들어, 250,000개 초과 태그를 식별하기 위해 9개 염기가 충분함).
- [0198] 인시튜 태그 시퀀싱은 다수의 방식으로 달성될 수 있다. 하나의 가능성은 살아있는 세포의 표면 상에 직접적으로 시퀀싱하기 위한 것이다. 다른 가능성은 (살아있는 세포로부터 누출된 뉴클레오타이드가 시퀀싱 반응을 방해하는 경우) 시퀀싱 이전에 세포를 고정시키는 것이다. 제3 가능성은 각 세포로부터 핵산 태그를 탈착시키고 포획하는 것이다. 이를 위하여, 흐름셀 표면은 모든 세포-부착된 핵산에 공통인 영역에 대해 상보적인 영역을 함유하는 포획 프로브(예를 들어, 시퀀싱 프라이머)로 개질되어, 표면에 세포의 초기 부착을 가능하게 하고, 시퀀싱 이전에 탈착된 핵산의 포획을 촉진시킨다. 선택적으로, 세포는 핵산 태그의 표면 포획 후에 제거될 수 있다.
- [0199] **2.2.2. 흐름셀에 부착된 태그의 복사체의 인시튜 시퀀싱.** 도 8에 도시된 바와 같이, 세포-표면 디스플레이된 핵산 태그는 흐름셀-표면 부착된 프라이머(부분 상보적인 서열을 가짐)에 하이브리드화된다. 세포-표면 핵산 태그는 특정 제한 효소로 소화에 의해 세포로부터 탈착되며, 세포는 이후에, 흐름셀로부터 세척된다. 흐름셀-표면 부착된 핵산 태그는 이후에, 확장되어, 본래 세포 태그의 복사체를 효과적으로 생성시키며, 이는 이후에 시퀀싱될 것이다. 다수의 경우에, 세포가 수천개의 핵산으로 이루어진 약 $1\mu\text{m}$ 직경의 "스팟"을 떠날 가능성이 있기 때문에, 복사체를 브릿지-증폭할 필요는 없을 것이다. 그러나, 브릿지 증폭은 요망되는 경우에 수행될 수 있다(그러한 경우에, 적절한 프라이머 결합 부위는 핵산 태그에 첨가될 수 있으며, 제2 프라이머 타입은 흐름셀 표면에 부착될 수 있다).
- [0200] 일부 경우에, 태그는 도 9에 도시된 바와 같이, 막-고정 핵산을 사용하여 개별 세포에 부착될 수 있다. 막 고정 핵산은 링커 영역, 제한 부위(RE), 시퀀싱 프라이머 결합 부위(SBS3), 태그 서열(바코드) 및 포획 서열을 포함할 것이다. 세포는 이후에, 문헌[Sleden et al., *J Am Chem Soc* 134: 765-8 (2012)]에 기술된 바와 같이, 또는 도 10에 도시된 바와 같이, 포획 서열에 대해 상보적인 서열을 갖는 표면-부착된 핵산에 의해 표면-포획될 수 있으며, 이러한 문헌은 본 명세서에 참고로 포함된다.
- [0201] 스크리닝 단계(예를 들어, 소정 시간에 걸쳐 및 특정 배지 조건 하에서 리포터로부터의 형광을 측정함)에서 세포 표현형 거동을 지시하는 형광 신호를 기록한 후에, 핵산 태그는 흐름셀 표면 상에 복사될 수 있으며, 이후에, 세포는 세척될 수 있다. 표면 상의 복사체는 도 11에 도시된 바와 같이, 흐름셀 표면 상에 세포가 위치된 것과 관련한 공간 정보를 효율적으로 보존한다.
- [0202] 표면 상에 각 세포의 공간 위치를 기록함으로써, 기록된 형광 시계열 각각에 태그 동일성을 지정하여, 이에 따라 세포 동일성과 표현형 반응을 효과적으로 연결하는 것이 가능할 것이다. 이러한 동일성은 이후에, 단일 용기(여기에서, 도 12에 도시된 바와 같이, 세포는 초기에 태그화됨)에서 개별 클론을 식별하거나, 선택된 세포의 표면 상에 존재하는 태그에 대해 상보적인 핵산과 함께 풀 다운에 의해 세포 혼합물로부터 선택된 세포를 회복하기 위해 사용될 수 있다.
- [0203] 본 출원에 걸쳐 다양한 공보, 특허 및 특허 출원이 참조되었다. 이들 공보의 개시는 본 발명이 속하는 분야의 상태를 보다 상세히 설명하기 위해 그 전문이 본 출원에서 참조로 도입된다.

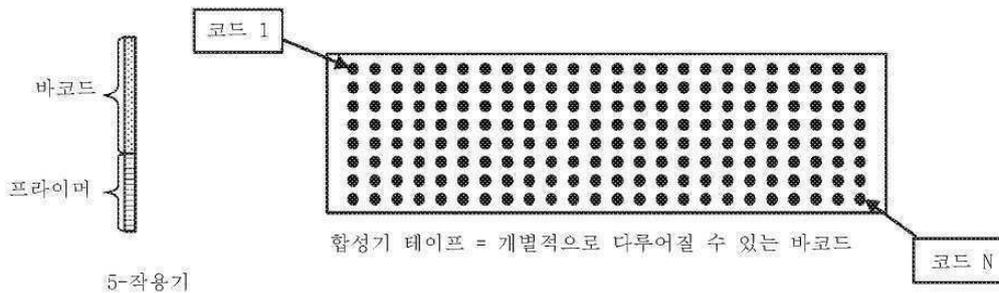
[0204] 용어 "포함하는"은 본 명세서에서 언급된 성분뿐만 아니라 임의의 추가적 성분을 추가로 포괄하여 포함하는 개방형 종결부를 나타낸다.

[0205] 본 발명이 상기 제공된 실시예를 참조하여 기재되었으나, 본 발명에서 벗어나지 않고 다양한 개질이 수행될 수 있음이 이해되어야 한다. 따라서 본 발명은 특허청구범위에 의해서만 제한된다.

도면

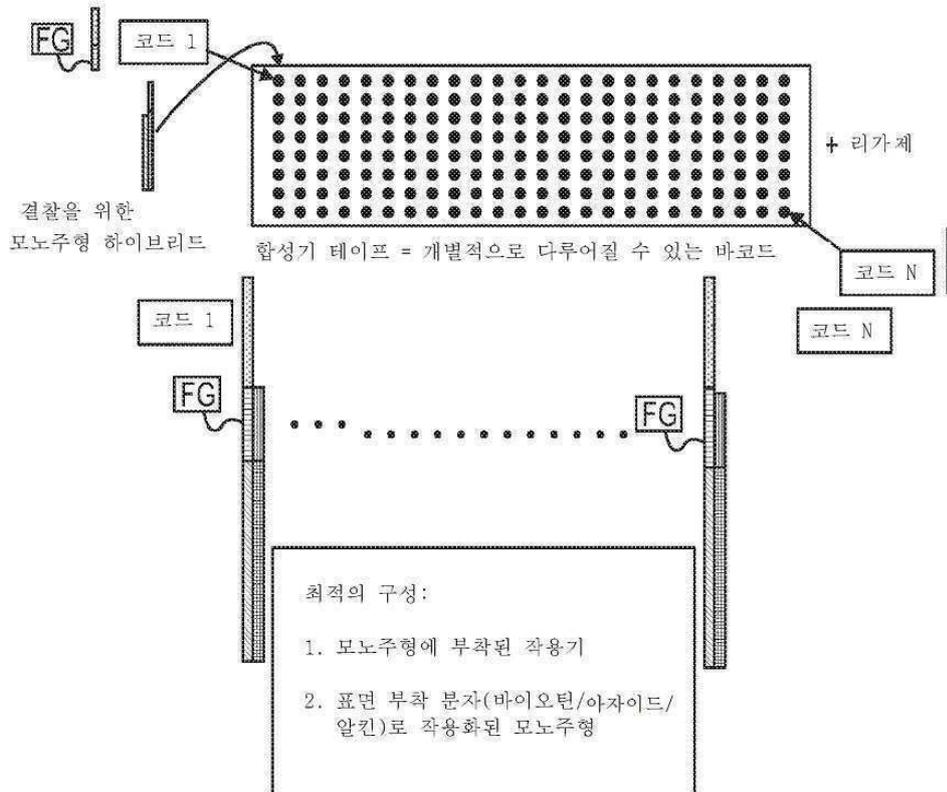
도면1a

단계 1: 500,000 29량체 oligos + 부착기(-SH, 아민; NHS)를 합성함.
 1. 합성기 = 50,000 oligos/일 10일에 500,000 oligos
 2. 바-코드 = 19량체; "프라이머" = 10량체



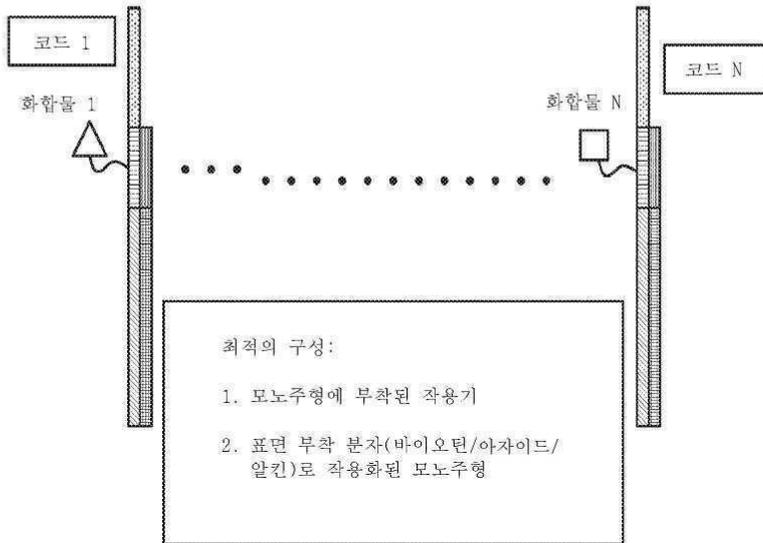
도면1b

단계 2: 오버행 프라이머를 개별 웰에서 모노주형 서열과 결합시킴.



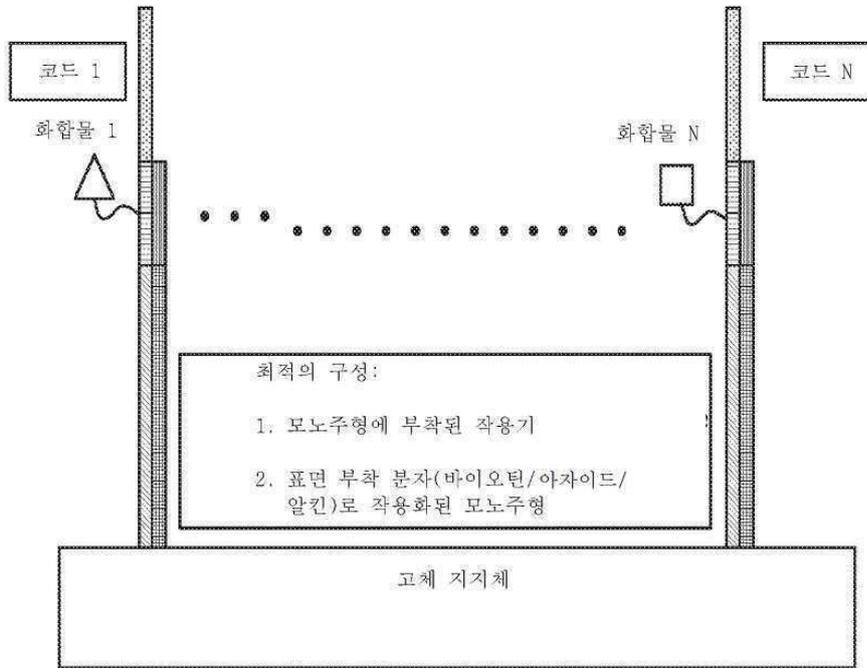
도면1c

단계 3: 반응기(RG)를 갖는 화합물을 작용화된 코드를 갖는 개별 웰에 로봇으로 첨가함.



도면1d

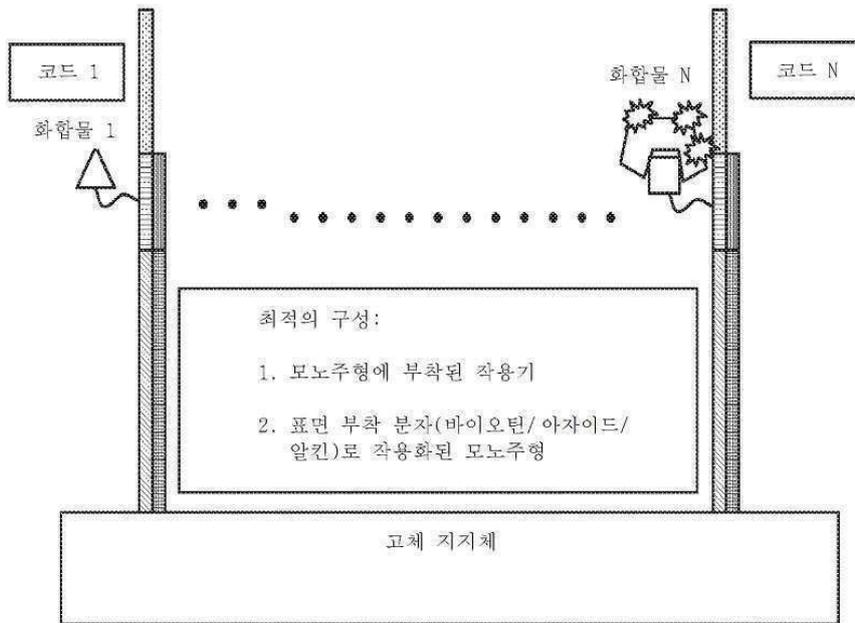
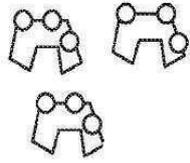
단계 4: 화합물을 하이브리드화 또는 표면 작용기를 통해 단일 분자 표면(또는 Pizam 흐름셀 표면)에 고정시킴



도면1e

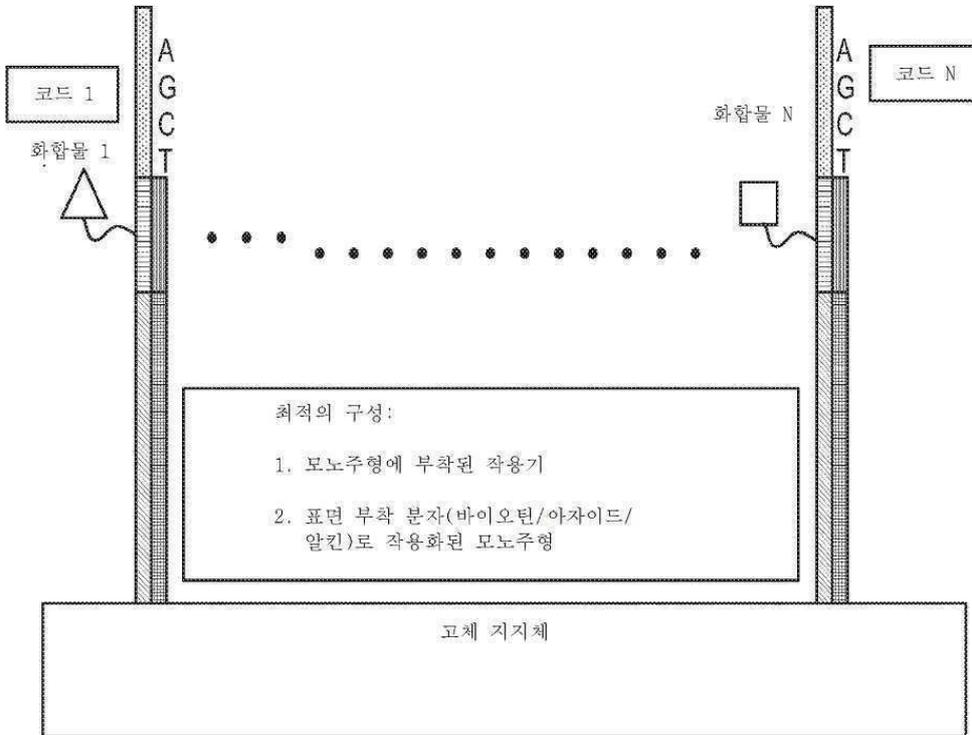
단계 5: 표적 분자를 유입시키고 화합물에 결합 동력학의 실시간 이미지를 기록함.

다중-표지된 표적

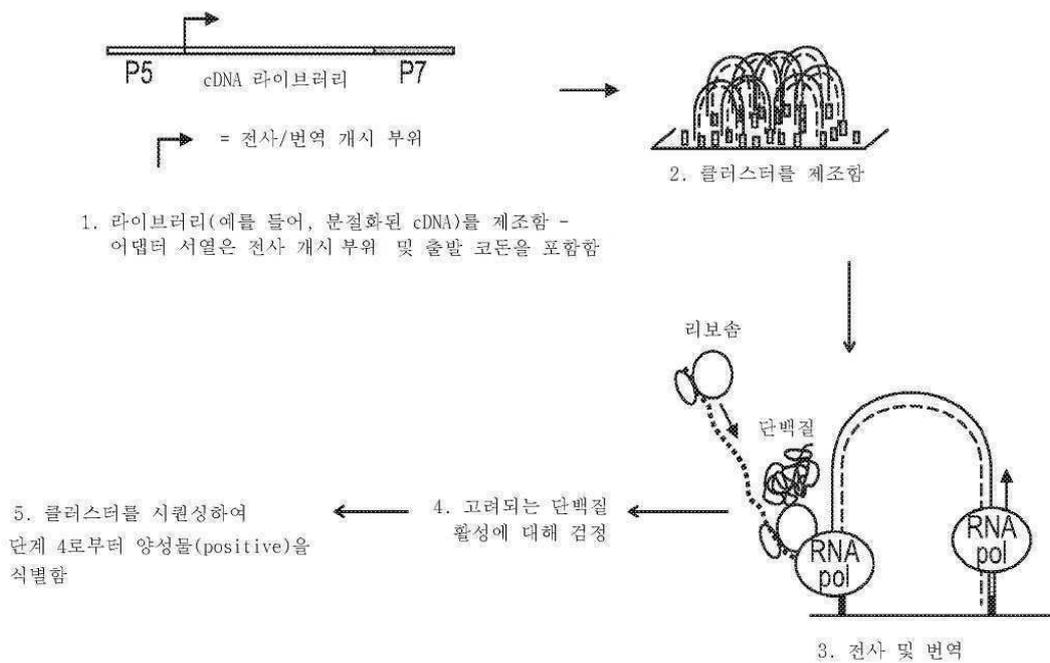


도면1f

단계 6: 엔코딩된 분자를 시퀀싱하여 "리드" 또는 "히트"를 식별함.

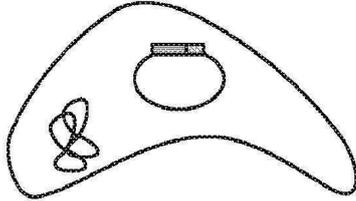


도면2

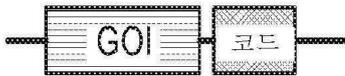


도면3a

1. mRNA 전사체 + 지수를 갖는 박테리아 라이브러리를 생성시킴
(지수 : 10bp = 10^6 코드 ; 20 bp = 10^{12} 코드)



2. 박테리아를 용해함 - mRNA 전사체를 수집함

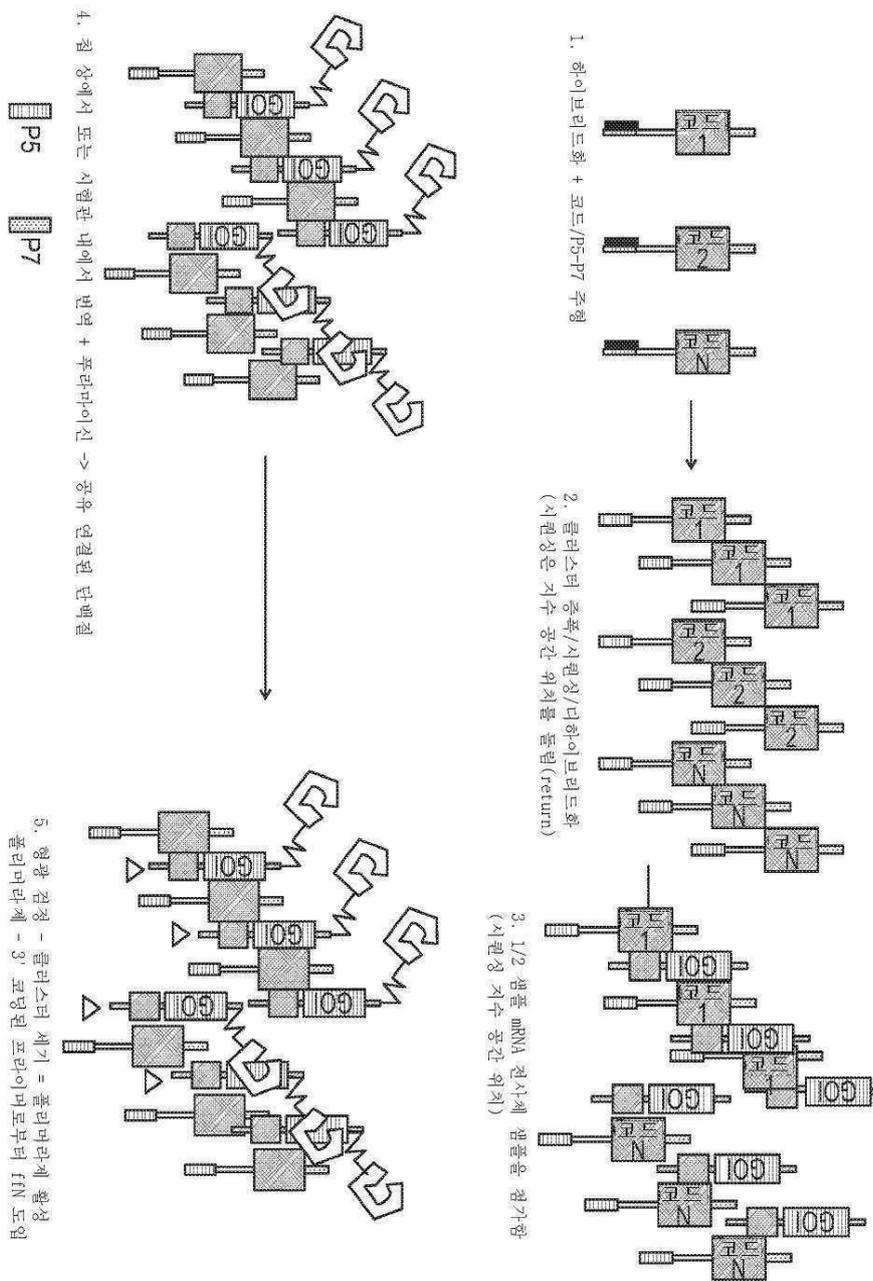


3. 1/2 샘플 -> 시퀀싱은 돌연변이에 대한 지수를 결정함

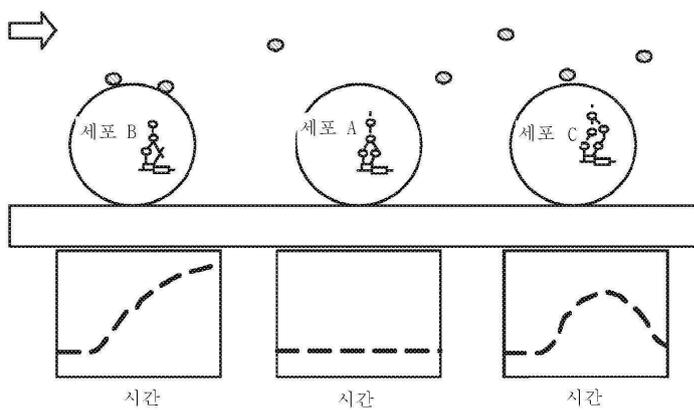
$$\begin{aligned} \text{돌연변이 } Y_1 &= \text{코드 } X_1 \\ \text{돌연변이 } Y_2 &= \text{코드 } X_2 \\ \text{돌연변이 } Y_N &= \text{코드 } X_N \end{aligned}$$

4. 1/2 샘플 -> 흐름셀 상에서 번역/검정

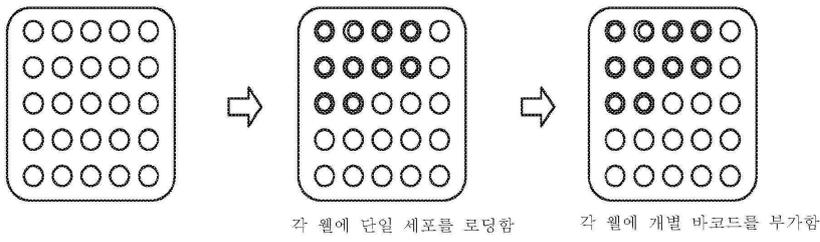
도면3b



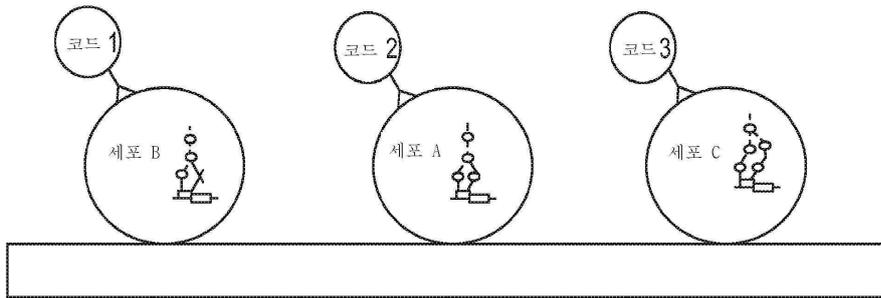
도면4



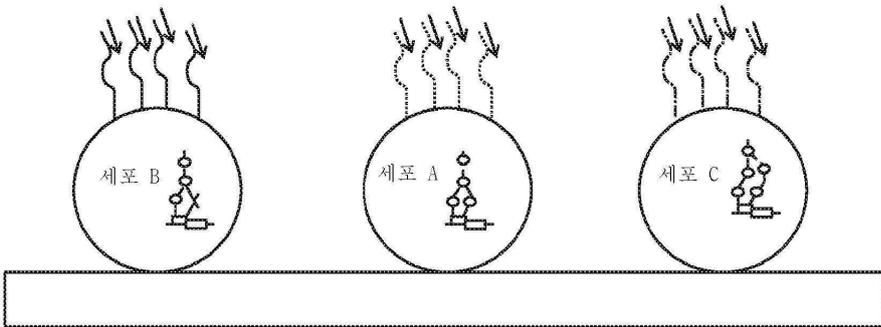
도면5



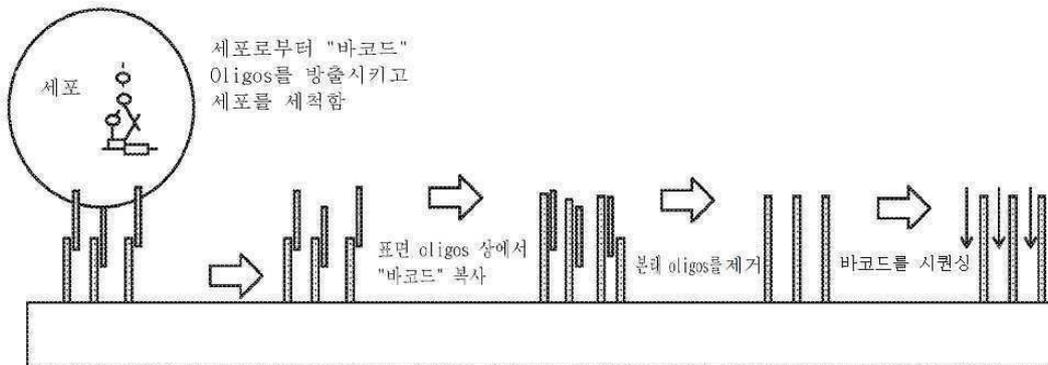
도면6



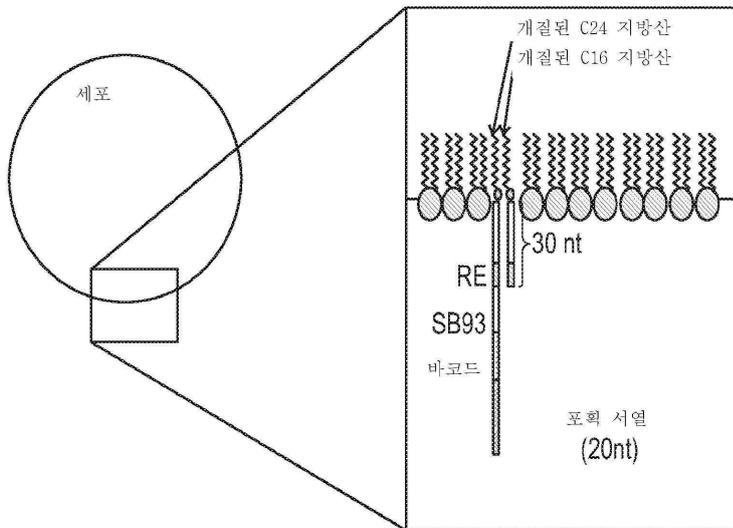
도면7



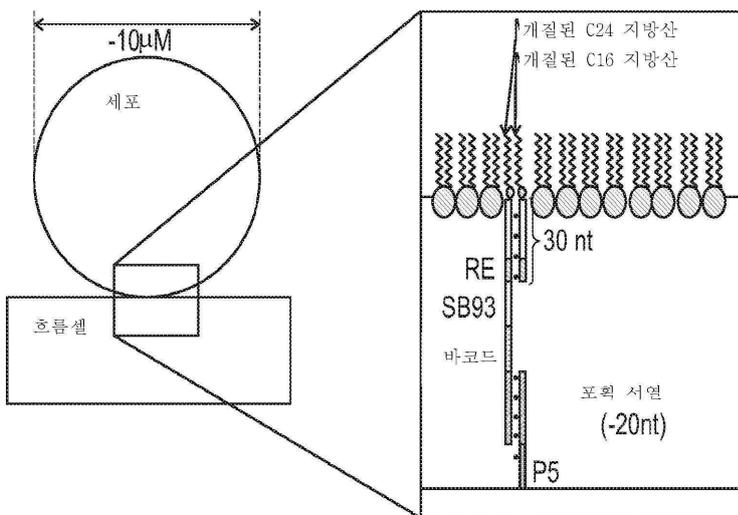
도면8



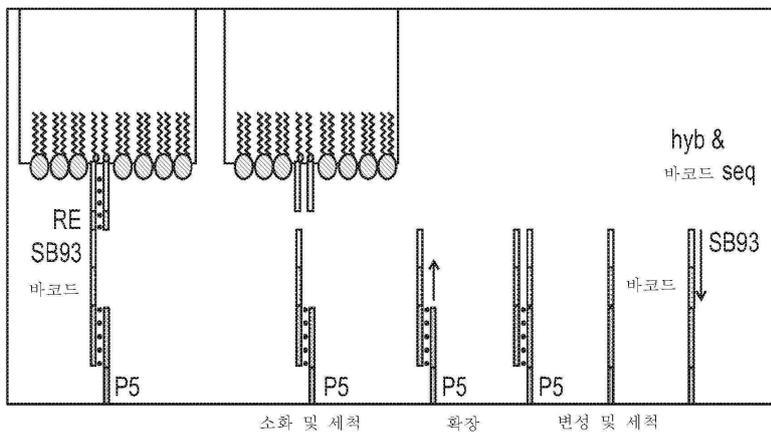
도면9



도면10



도면11



도면12

