



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년09월01일

(11) 등록번호 10-2438495

(24) 등록일자 2022년08월26일

- (51) 국제특허분류(Int. Cl.)
C12Q 1/6804 (2018.01) **C12Q 1/6869** (2018.01)
G01N 33/53 (2006.01)
- (52) CPC특허분류
C12Q 1/6804 (2018.05)
C12Q 1/6869 (2018.05)
- (21) 출원번호 10-2019-7038794
- (22) 출원일자(국제) 2018년03월27일
 심사청구일자 2021년03월26일
- (85) 번역문제출일자 2019년12월30일
- (65) 공개번호 10-2020-0014370
- (43) 공개일자 2020년02월10일
- (86) 국제출원번호 PCT/US2018/024602
- (87) 국제공개번호 WO 2018/226293
 국제공개일자 2018년12월13일
- (30) 우선권주장
 62/515,285 2017년06월05일 미국(US)
 (뒷면에 계속)
- (56) 선행기술조사문헌
 W02012048341 A1*
 W02017053905 A1*
 *는 심사관에 의하여 인용된 문헌

- (73) 특허권자
백톤 디킨슨 앤드 컴퍼니
 미합중국, 뉴저지주, 프랭클린 레이크스, 1백톤
 드라이브 (우:07417-1880)
- (72) 발명자
장 크리스티나
 미국 94025 캘리포니아주 멘로 파크 캠벨 에비뉴
 4040 스위트 110
판 크리스티나
 미국 94025 캘리포니아주 멘로 파크 캠벨 에비뉴
 4040 스위트 110
 (뒷면에 계속)
- (74) 대리인
양영준, 김윤기

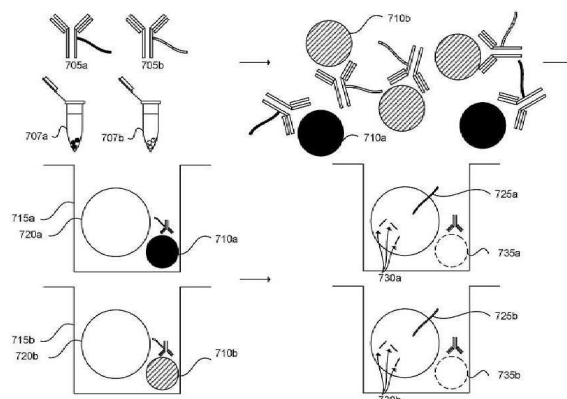
전체 청구항 수 : 총 48 항

심사관 : 김현태

(54) 발명의 명칭 단일 세포를 위한 샘플 인덱싱

(57) 요약

본 명세서의 개시 내용은 샘플 식별을 위한 시스템, 방법, 조성물, 및 키트를 포함한다. 샘플 인덱싱 조성물은, 예를 들어 샘플 인덱싱 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함할 수 있다. 상이한 샘플 인덱싱 조성물은 상이한 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드를 포함할 수 있다. 세포의 샘플 기원은 샘플 인덱싱 올리고뉴클레오타이드의 서열에 기초하여 식별될 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 바코드를 사용하여 바코딩되고 테이지-체이닝 프라이머(daisy-chaining primer)를 사용하여 신장될 수 있다.

대표도

(52) CPC특허분류

G01N 33/5308 (2013.01)
C12Q 2525/179 (2013.01)
C12Q 2563/131 (2013.01)
C12Q 2563/149 (2013.01)
C12Q 2563/179 (2013.01)
C12Q 2563/185 (2013.01)
G01N 2458/10 (2013.01)

(72) 발명자

슘 엘린

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

마틴 조디

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

반살 니드한잘리

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

가디알리 제임스

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

라자룩 캐썬

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

람 그렛첸 인본

미국 94025 캘리포니아주 덴로 파크 캠펠 예비뉴
 4040 스위트 110

(30) 우선권주장

62/532,905	2017년07월14일	미국(US)
62/532,949	2017년07월14일	미국(US)
62/532,971	2017년07월14일	미국(US)
62/554,425	2017년09월05일	미국(US)
62/578,957	2017년10월30일	미국(US)
62/645,703	2018년03월20일	미국(US)

명세서

청구범위

청구항 1

(a) 제1 복수의 세포를 포함하는 제1 샘플 및 제2 복수의 세포를 포함하는 제2 샘플을 제1 올리고뉴클레오타이드-접합된 항체 및 제2 올리고뉴클레오타이드-접합된 항체와 각각 접촉시켜 제1 바코딩된 샘플 및 제2 바코딩된 샘플을 각각 형성하는 단계로서, 각각의 샘플은 제1 올리고뉴클레오타이드-접합된 항체 및 제2 올리고뉴클레오타이드-접합된 항체와 회합된 세포를 각각 포함하고,

제1 항체는 제1 복수의 세포 각각의 단백질 표적과 특이적으로 결합할 수 있고, 제1 샘플 인덱싱(indexing) 서열 및 폴리(A) 테일(tail)을 포함하는 제1 샘플 인덱싱 올리고뉴클레오타이드와 접합될 수 있고,

제2 항체는 제2 복수의 세포 각각의 단백질 표적과 특이적으로 결합할 수 있고, 제2 샘플 인덱싱 서열 및 폴리(A) 테일을 포함하는 제2 샘플 인덱싱 올리고뉴클레오타이드와 접합될 수 있고,

제1 샘플 인덱싱 서열 및 제2 샘플 인덱싱 서열은 상이한 서열을 포함하고,

제1 복수의 세포 각각의 단백질 표적 및 제2 복수의 세포 각각의 단백질 표적은 동일한 단계;

(b) 제1 바코딩된 샘플 및 제2 바코딩된 샘플을 풀링(pooling)하여 제1 올리고뉴클레오타이드-접합된 항체와 회합된 제1 복수의 세포 및 제2 올리고뉴클레오타이드-접합된 항체와 회합된 제2 복수의 세포를 포함하는 조합되고 바코딩된 샘플을 형성하는 단계;

(c) 조합되고 바코딩된 샘플의 제1 올리고뉴클레오타이드-접합된 항체와 회합된 제1 복수의 세포 및 제2 올리고뉴클레오타이드-접합된 항체와 회합된 제2 복수의 세포를 복수의 파티션(partition)으로 분배하는 단계로서, 복수의 파티션 중 파티션 각각은 조합되고 바코딩된 샘플의 단일 세포를 포함하는 단계;

(d) 복수의 파티션 중 파티션 각각에서, 바코딩 입자를 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드와 접촉시키는 단계로서, 바코딩 입자는 각각이 폴리(T) 영역을 포함하는 복수의 올리고뉴클레오타이드 프로브 및 바코딩 입자의 복수의 올리고뉴클레오타이드 프로브에 공통적인 바코드 서열을 포함하는 단계;

(e) 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드의 폴리(A) 테일과 올리고뉴클레오타이드 프로브의 폴리(T) 영역 사이의 혼성화를 통해 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 인덱싱 올리고뉴클레오타이드와 혼성화된 각각의 올리고뉴클레오타이드 프로브를 연장하여 제1 표지된 핵산 또는 제2 표지된 핵산을 생성하는 단계로서, 제1 표지된 핵산 및 제2 표지된 핵산은 제1 샘플 인덱싱 서열 및 제2 샘플 인덱싱 서열을 각각, 또는 이들의 상보적인 서열, 및 바코드 서열을 포함하는 단계; 및

(f) 제1 샘플 인덱싱 서열을 포함하는 제1 표지된 핵산 또는 제2 샘플 인덱싱 서열을 포함하는 제2 표지된 핵산, 또는 이들의 일부의 서열 정보를 획득하는 단계; 및

(g) 서열 정보 내의 제1 샘플 인덱싱 서열 또는 제2 샘플 인덱싱 서열 각각, 이들의 상보적 서열, 또는 이들의 일부에 기초하여 조합되고 바코딩된 샘플의 복수의 파티션 중 파티션 내의 각각의 단일 세포가 제1 샘플 또는 제2 샘플로부터 유래되었음을 결정하는 단계

를 포함하는, 샘플 식별을 위한 방법.

청구항 2

제1항에 있어서, 단계 (a) 후에, 각각 제1 복수의 세포 및 제2 복수의 세포와 회합되지 않은 제1 올리고뉴클레오타이드-접합된 항체 및 제2 올리고뉴클레오타이드-접합된 항체를 제거하는 단계를 포함하는, 방법.

청구항 3

제1항에 있어서, 제1 복수의 세포 각각의 단백질 표적과 특이적으로 결합할 수 있는 제1 항체와 제2 복수의 세포 각각의 단백질 표적과 특이적으로 결합할 수 있는 제2 항체는 동일한, 방법.

청구항 4

제1항에 있어서, 단계 (c) 내지 (e) 중 어느 하나에서, 제1 올리고뉴클레오타이드-접합된 항체 및 제2 올리고뉴클레오타이드-접합된 항체로부터 제1 샘플 인덱싱 올리고뉴클레오타이드 및 제2 샘플 인덱싱 올리고뉴클레오타이드를 각각 분리하는 단계를 포함하는, 방법.

청구항 5

제1항에 있어서, 제1 샘플 인덱싱 서열과 제2 샘플 인덱싱 서열 중 적어도 하나는 6 내지 60개의 뉴클레오타이드 길이인, 방법.

청구항 6

제1항에 있어서, 제1 샘플 인덱싱 올리고뉴클레오타이드와 제2 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나는 50 내지 500개의 뉴클레오타이드 길이인, 방법.

청구항 7

제1항에 있어서, 제1 복수의 세포 각각의 단백질 표적과 제2 복수의 세포 각각의 단백질 표적 중 적어도 하나는 세포 표면 단백질, 세포내 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 조합인, 방법.

청구항 8

제1항에 있어서, 단계 (c)에서 조합되고 바코딩된 샘플의 제1 복수의 세포 및 제2 복수의 세포를 분배하는 단계는 조합된 샘플의 제1 올리고뉴클레오타이드-접합된 항체와 회합된 제1 복수의 세포 및 제2 올리고뉴클레오타이드-접합된 항체와 회합된 제2 복수의 세포 및 바코딩 입자를 포함하는 복수의 바코딩 입자를 복수의 파티션으로 분배하는 단계를 포함하는, 방법.

청구항 9

제1항에 있어서, 제1 표지된 핵산 또는 제2 표지된 핵산, 또는 이들의 일부의 서열 정보를 획득하는 단계 (f)는 제1 표지된 핵산 또는 제2 표지된 핵산이 하나 이상의 반응을 거치도록 하여 핵산 서열에 대한 제1 세트의 핵산 또는 제2 세트의 핵산을 생성하는 단계를 포함하는, 방법.

청구항 10

제1항에 있어서, 각각의 올리고뉴클레오타이드 프로브는 세포 표지, 범용 프라이머를 위한 결합 부위, 증폭 어댑터, 서열 어댑터, 또는 이들의 조합을 포함하는, 방법.

청구항 11

제1항에 있어서, 복수의 올리고뉴클레오타이드 프로브 중 적어도 하나의 올리고뉴클레오타이드 프로브는 바코딩 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 조합인, 방법.

청구항 12

제1항에 있어서, 단계 (d)는 단일 세포를 포함하는 파티션의 바코딩 입자로부터 복수의 올리고뉴클레오타이드 프로브를 방출하는 단계를 포함하는, 방법.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 바코딩 입자는 세파로스(Sepharose) 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 하이드로겔 비드, 또는 붕괴성(disruptable) 하이드로겔 비드인, 방법.

청구항 14

제1항 내지 제12항 중 어느 한 항에 있어서, 복수의 파티션은 웰(well) 또는 액적(droplet)을 포함하는, 방법.

청구항 15

제1항 내지 제12항 중 어느 한 항에 있어서, 제1 복수의 세포 및 제2 복수의 세포는 T 세포, B 세포, 종양 세포, 골수 세포, 혈액 세포, 정상 세포, 태아 세포, 모체 세포, 또는 이들의 조합을 포함하는, 방법.

청구항 16

제1항 내지 제12항 중 어느 한 항에 있어서, 복수의 파티션의 파티션 중 하나의 단일 세포는 관심 대상인 복수의 단백질을 포함하고,

단계 (a)는 제3 복수의 올리고뉴클레오타이드-접합된 항체를 단일 세포와 접촉시키는 단계를 추가로 포함하고, 제3 복수의 올리고뉴클레오타이드-접합된 항체 각각은 폴리(A) 테일 및 접합된 항체에 대한 고유 식별자를 포함하는 항체 특이적 올리고뉴클레오타이드와 접합된 항체를 포함하고, 항체는 관심 대상인 복수의 단백질 중 적어도 하나와 특이적으로 결합할 수 있고,

단계 (e)는 항체 특이적 올리고뉴클레오타이드의 폴리(A)테일과 올리고뉴클레오타이드 프로브의 폴리(T) 영역 사이의 혼성화를 통해 항체 특이적 올리고뉴클레오타이드와 혼성화된 올리고뉴클레오타이드 프로브를 연장하여 제1 복수의 표지된 핵산을 생성하는 단계를 추가로 포함하고, 제1 복수의 표지된 핵산 각각은 고유 식별자, 또는 이들의 상보적인 서열, 및 바코드 서열을 포함하고,

단계 (f)는 제1 복수의 표지된 핵산 또는 이들의 일부의 서열 정보를 획득하여 단일 세포의 관심 대상인 복수의 단백질 중 하나 이상의 양을 결정하는 단계를 추가로 포함하는, 방법.

청구항 17

제16항에 있어서, 항체 특이적 올리고뉴클레오타이드는 분자 표지, 세포 표지, 범용 프라이머를 위한 결합 부위, 증폭 어댑터, 서열 어댑터 또는 이들의 조합을 포함하는, 방법.

청구항 18

제16항에 있어서, 관심 대상인 복수의 단백질 중 관심 대상인 단백질은 세포 표면 단백질, 세포내 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 조합인, 방법.

청구항 19

제16항에 있어서, 단일 세포는 복수의 mRNA 표적 분자를 포함하고,

단계 (d)는 mRNA 표적 분자의 폴리(A) 테일과 올리고뉴클레오타이드 프로브의 폴리(T) 영역 사이의 복수의 올리고뉴클레오타이드 프로브 중 올리고뉴클레오타이드 프로브와 단일 세포로부터의 mRNA 표적 분자를 혼성화하는 단계를 추가로 포함하고,

단계 (e)는 mRNA 표적 분자의 폴리(A) 테일과 올리고뉴클레오타이드 프로브의 폴리(T) 영역 사이의 혼성화를 통해 mRNA 표적 분자와 혼성화된 올리고뉴클레오타이드 프로브를 연장하여 제2 복수의 표지된 핵산을 생성하는 단계를 추가로 포함하고, 제2 복수의 표지된 핵산 각각은 바코드 서열과 mRNA 표적 분자 중 하나의 상보적 서열을 포함하고,

단계 (f)는 제2 복수의 표지된 핵산 또는 이들의 일부의 서열 정보를 획득하여 단일 세포의 복수의 mRNA 표적 분자 중 하나 이상의 양을 결정하는 단계를 추가로 포함하는, 방법.

청구항 20

(a) 제1 복수의 세포를 포함하는 제1 샘플 및 제2 복수의 세포를 포함하는 제2 샘플을 제1 샘플 인덱싱 조성물 및 제2 샘플 인덱싱 조성물과 각각 접촉시켜 제1 표지된 샘플 및 제2 표지된 샘플을 각각 형성하는 단계로서, 각각의 샘플은 제1 샘플 인덱싱 조성물 및 제2 샘플 인덱싱 조성물과 회합된 세포를 각각 포함하고,

제1 샘플 인덱싱 조성물은 제1 샘플 인덱싱 올리고뉴클레오타이드와 회합된 제1 세포 성분 결합 시약을 포함하고, 제1 세포 성분 결합 시약은 제1 복수의 세포 각각의 세포 성분 표적과 특이적으로 결합할 수 있고,

제2 샘플 인덱싱 조성물은 제2 샘플 인덱싱 올리고뉴클레오타이드와 회합된 제2 세포 성분 결합 시약을 포함하고, 제2 세포 성분 결합 시약은 제2 복수의 세포 각각의 세포 성분 표적과 특이적으로 결합할 수 있고,

제1 및 제2 샘플 인덱싱 올리고뉴클레오타이드는 제1 샘플 인덱싱 서열 및 제2 샘플 인덱싱 서열을 각각 포함하고, 제1 샘플 인덱싱 서열 및 제2 샘플 인덱싱 서열은 상이한 서열을 포함하고,

제1 복수의 세포 각각의 세포 성분 표적 및 제2 복수의 세포 각각의 세포 성분 표적은 동일한 단계;

(b) 제1 표지된 샘플 및 제2 표지된 샘플을 풀링하여 제1 샘플 인덱싱 조성물과 회합된 세포 및 제2 샘플 인덱싱 조성물과 회합된 세포를 각각 포함하는 조합되고 표지된 샘플을 형성하는 단계;

(c) 조합되고 표지된 샘플의 제1 샘플 인덱싱 조성물과 회합된 세포 및 제2 샘플 인덱싱 조성물과 회합된 세포를 복수의 파티션으로 분배하는 단계로서, 복수의 파티션 중 파티션 각각은 조합되고 표지된 샘플의 단일 세포를 포함하는 단계;

(d) 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드가 하나 이상의 반응을 거치도록 하여 제1 샘플 인덱싱 서열 또는 이들의 상보체를 포함하는 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드, 또는 제2 샘플 인덱싱 서열 또는 이들의 상보체를 포함하는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계;

(e) 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드, 또는 이들의 일부의 서열 정보를 획득하는 단계; 및

(f) 단계 (e)에서 획득한 서열 정보 내의 제1 샘플 인덱싱 서열 또는 제2 샘플 인덱싱 서열 각각, 이들의 상보적 서열, 또는 이들의 일부에 기초하여, 조합되고 표지된 샘플 내의 세포가 제1 샘플 또는 제2 샘플로부터 유래되었음을 결정하는 단계

를 포함하는, 샘플 식별을 위한 방법.

청구항 21

제20항에 있어서, 세포 성분 결합 시약은 단백질 결합 시약을 포함하고, 단백질 결합 시약은 상응하는 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드와 회합되고, 세포 성분 표적은 단백질 표적을 포함하는, 방법.

청구항 22

제20항에 있어서, 단계 (d)의, 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드가 하나 이상의 반응을 거치도록 하여 제1 샘플 인덱싱 서열 또는 이들의 상보체를 포함하는 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드, 또는 제2 샘플 인덱싱 서열 또는 이들의 상보적 서열을 포함하는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는,

복수의 바코드를 사용하여 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 올리고뉴클레오타이드를 바코딩하여 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함하는, 방법.

청구항 23

제20항에 있어서, 조합되고 표지된 샘플 내의 세포가 제1 샘플 또는 제2 샘플로부터 유래되었음을 결정하는 단계는 바코딩된 샘플 인덱싱 조성물 내의 제1 샘플 인덱싱 서열, 제2 샘플 인덱싱 서열, 이들의 상보적 서열, 또는 이들의 일부의 존재 또는 부재를 식별하는 단계를 포함하는, 방법.

청구항 24

제20항에 있어서, 제1 샘플 인덱싱 서열과 제2 샘플 인덱싱 서열 중 적어도 하나는 6 내지 60개의 뉴클레오타이드 길이인, 방법.

청구항 25

제20항에 있어서, 제1 복수의 세포 각각의 세포 성분 표적과 제2 복수의 세포 각각의 세포 성분 표적 중 적어도

하나는 세포 표면 상에 존재하는, 방법.

청구항 26

제20항에 있어서, 단계 (a) 및 (b) 중 어느 하나에서, 제1 복수의 세포 및 제2 복수의 세포와 회합되지 않은 제1 및 제2 샘플 인텍싱 조성물을 제거하는 단계를 포함하는, 방법.

청구항 27

제20항에 있어서, 제1 샘플 인텍싱 올리고뉴클레오타이드가 제1 세포 성분 결합 시약으로부터 탈착 불가능하도록 구성되거나, 또는 제2 샘플 인텍싱 올리고뉴클레오타이드가 제2 성분 결합 시약으로부터 탈착 불가능하도록 구성되거나, 또는 둘 모두인, 방법.

청구항 28

제20항에 있어서, 단계 (c) 내지 (e) 중 어느 하나에서, 제1 세포 성분 결합 시약으로부터 제1 샘플 인텍싱 올리고뉴클레오타이드를 탈착하는 단계, 제2 세포 성분 결합 시약으로부터 제2 샘플 인텍싱 올리고뉴클레오타이드를 탈착하는 단계, 또는 둘 모두를 포함하는, 방법.

청구항 29

제22항에 있어서, 제1 및/또는 제2 샘플 인텍싱 올리고뉴클레오타이드는 제1 및/또는 제2 샘플 인텍싱 올리고뉴클레오타이드의 서열을 포획하도록 구성된 포획 서열에 상보적인 서열을 포함하고, 각각의 바코드는 포획 서열을 포함하는 표적-결합 영역을 포함하는, 방법.

청구항 30

제29항에 있어서, 표적-결합 영역은 폴리(dT) 영역을 포함하고, 포획 서열에 상보적인 샘플 인텍싱 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함하는, 방법.

청구항 31

제20항에 있어서, 제1 샘플 인텍싱 올리고뉴클레오타이드와 제2 샘플 인텍싱 올리고뉴클레오타이드 중 적어도 하나는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 둘 모두를 포함하는, 방법.

청구항 32

제20항에 있어서, 제1 복수의 세포 각각의 세포 성분 표적, 제2 복수의 세포 각각의 세포 성분 표적, 또는 둘 모두는 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 세포내 단백질 또는 이들의 임의의 조합을 포함하는, 방법.

청구항 33

제22항에 있어서, 복수의 바코드는 입자와 회합되는, 방법.

청구항 34

제33항에 있어서, 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 조합인, 방법.

청구항 35

제33항에 있어서, 입자는 붕괴성(disruptable) 비드 또는 붕괴성 하이드로겔 비드인, 방법.

청구항 36

제33항에 있어서, 복수의 바코드를 사용하여 제1 샘플 인텍싱 올리고뉴클레오타이드 또는 제2 샘플 인텍싱 올리고뉴클레오타이드를 바코딩하는 단계는

복수의 바코드를 제1 샘플 인텍싱 올리고뉴클레오타이드 또는 제2 샘플 인텍싱 올리고뉴클레오타이드와 접촉시켜 바코드가 제1 샘플 인텍싱 올리고뉴클레오타이드 또는 제2 샘플 인텍싱 올리고뉴클레오타이드와 혼성화되도록 하는 단

계로서, 각각의 바코드는 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드, 또는 둘 모두와 특이적으로 결합할 수 있는 표적 결합 영역을 포함하는 단계; 및

제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드와 혼성화된 바코드를 연장하여 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계

를 포함하는, 방법.

청구항 37

제33항에 있어서, 복수의 바코드 각각은 동일한 바코드 서열을 포함하는, 방법.

청구항 38

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 샘플 인덱싱 올리고뉴클레오타이드와 제2 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나는 50 내지 500개의 뉴클레오타이드 길이인, 방법.

청구항 39

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 세포 성분 결합 시약과 제2 세포 성분 결합 시약 중 적어도 하나는 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함하는, 방법.

청구항 40

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 샘플 인덱싱 올리고뉴클레오타이드가 링커(linker)를 통해 제1 세포 성분 결합 시약에 접합되거나, 또는 제2 샘플 인덱싱 올리고뉴클레오타이드가 링커를 통해 제2 세포 성분 결합 시약에 접합되거나, 또는 둘 모두인, 방법.

청구항 41

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일한, 방법.

청구항 42

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 복수의 세포 각각의 세포 성분 표적과 제2 복수의 세포 각각의 세포 성분 표적 중 적어도 하나는 세포 표면 단백질, 세포내 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 조합을 포함하는, 방법.

청구항 43

제20항 내지 제37항 중 어느 한 항에 있어서, 복수의 파티션은 웰 또는 액적을 포함하는, 방법.

청구항 44

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 복수의 세포, 제2 복수의 세포, 또는 둘 모두는 T 세포, B 세포, 종양 세포, 골수 세포, 혈액 세포, 정상 세포, 태아 세포, 모체 세포, 또는 이들의 조합을 포함하는, 방법.

청구항 45

제22항에 있어서, 복수의 바코드 각각은 세포 표지, 범용 프라이머를 위한 결합 부위, 증폭 어댑터, 서열 어댑터, 또는 이들의 조합을 포함하는, 방법.

청구항 46

제33항에 있어서, 단계 (c)의, 조합되고 바코딩된 샘플의 제1 샘플 인덱싱 조성물과 회합된 제1 복수의 세포 및 제2 샘플 인덱싱 조성물과 회합된 제2 복수의 세포를 복수의 파티션으로 분배하는 단계는,

조합되고 바코딩된 샘플의 제1 샘플 인덱싱 조성물과 회합된 제1 복수의 세포, 제2 샘플 인덱싱 조성물과 회합된 제2 복수의 세포, 및 복수의 바코드를 포함하는 입자를 포함하는 복수의 입자를 분배하는 단계를 포함하는,

방법.

청구항 47

제33항에 있어서, 단계 (d) 전에 단일 세포를 포함하는 파티션의 입자로부터 복수의 바코드 중 하나 이상을 방출하는 단계를 포함하는, 방법.

청구항 48

제20항 내지 제37항 중 어느 한 항에 있어서, 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드가 하나 이상의 반응을 거치도록 하는 단계는, 복수의 파티션 중 파티션에서, 제1 샘플 인덱싱 올리고뉴클레오타이드 또는 제2 샘플 인덱싱 올리고뉴클레오타이드가 하나 이상의 반응을 거치도록 하여 제1 샘플 인덱싱 서열 또는 이들의 상보체를 포함하는 제1 바코딩된 샘플 인덱싱 올리고뉴클레오타이드, 또는 제2 샘플 인덱싱 서열 또는 이들의 상보체를 포함하는 제2 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함하는, 방법.

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

삭제

청구항 120

삭제

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

삭제

청구항 128

삭제

청구항 129

삭제

청구항 130

삭제

청구항 131

삭제

청구항 132

삭제

청구항 133

삭제

청구항 134

삭제

청구항 135

삭제

청구항 136

삭제

청구항 137

삭제

청구항 138

삭제

청구항 139

삭제

청구항 140

삭제

청구항 141

삭제

청구항 142

삭제

청구항 143

삭제

청구항 144

삭제

청구항 145

삭제

청구항 146

삭제

청구항 147

삭제

청구항 148

삭제

청구항 149

삭제

청구항 150

삭제

청구항 151

삭제

청구항 152

삭제

청구항 153

삭제

청구항 154

삭제

청구항 155

삭제

청구항 156

삭제

청구항 157

삭제

청구항 158

삭제

청구항 159

삭제

청구항 160

삭제

청구항 161

삭제

청구항 162

삭제

청구항 163

삭제

청구항 164

삭제

청구항 165

삭제

청구항 166

삭제

청구항 167

삭제

청구항 168

삭제

청구항 169

삭제

청구항 170

삭제

청구항 171

삭제

청구항 172

삭제

청구항 173

삭제

청구항 174

삭제

청구항 175

삭제

청구항 176

삭제

청구항 177

삭제

청구항 178

삭제

청구항 179

삭제

청구항 180

삭제

청구항 181

삭제

청구항 182

삭제

청구항 183

삭제

청구항 184

삭제

청구항 185

삭제

청구항 186

삭제

청구항 187

삭제

청구항 188

삭제

청구항 189

삭제

청구항 190

삭제

청구항 191

삭제

청구항 192

삭제

청구항 193

삭제

청구항 194

삭제

청구항 195

삭제

청구항 196

삭제

청구항 197

삭제

청구항 198

삭제

청구항 199

삭제

청구항 200

삭제

청구항 201

삭제

청구항 202

삭제

청구항 203

삭제

청구항 204

삭제

청구항 205

삭제

청구항 206

삭제

청구항 207

삭제

청구항 208

삭제

청구항 209

삭제

청구항 210

삭제

청구항 211

삭제

청구항 212

삭제

청구항 213

삭제

청구항 214

삭제

청구항 215

삭제

청구항 216

삭제

청구항 217

삭제

청구항 218

삭제

청구항 219

삭제

청구항 220

삭제

청구항 221

삭제

청구항 222

삭제

청구항 223

삭제

청구항 224

삭제

청구항 225

삭제

청구항 226

삭제

청구항 227

삭제

청구항 228

삭제

청구항 229

삭제

청구항 230

삭제

청구항 231

삭제

청구항 232

삭제

청구항 233

삭제

청구항 234

삭제

청구항 235

삭제

청구항 236

삭제

청구항 237

삭제

청구항 238

삭제

청구항 239

삭제

청구항 240

삭제

청구항 241

삭제

청구항 242

삭제

청구항 243

삭제

청구항 244

삭제

청구항 245

삭제

청구항 246

삭제

청구항 247

삭제

청구항 248

삭제

청구항 249

삭제

청구항 250

삭제

청구항 251

삭제

청구항 252

삭제

청구항 253

삭제

청구항 254

삭제

청구항 255

삭제

청구항 256

삭제

청구항 257

삭제

청구항 258

삭제

청구항 259

삭제

청구항 260

삭제

청구항 261

삭제

청구항 262

삭제

청구항 263

삭제

청구항 264

삭제

청구항 265

삭제

청구항 266

삭제

청구항 267

삭제

청구항 268

삭제

청구항 269

삭제

청구항 270

삭제

청구항 271

삭제

청구항 272

삭제

청구항 273

삭제

청구항 274

삭제

청구항 275

삭제

청구항 276

삭제

청구항 277

삭제

청구항 278

삭제

청구항 279

삭제

청구항 280

삭제

청구항 281

삭제

청구항 282

삭제

청구항 283

삭제

청구항 284

삭제

청구항 285

삭제

청구항 286

삭제

청구항 287

삭제

청구항 288

삭제

청구항 289

삭제

청구항 290

삭제

청구항 291

삭제

청구항 292

삭제

청구항 293

삭제

청구항 294

삭제

청구항 295

삭제

청구항 296

삭제

청구항 297

삭제

청구항 298

삭제

청구항 299

삭제

청구항 300

삭제

청구항 301

삭제

청구항 302

삭제

청구항 303

삭제

청구항 304

삭제

청구항 305

삭제

청구항 306

삭제

청구항 307

삭제

청구항 308

삭제

청구항 309

삭제

청구항 310

삭제

청구항 311

삭제

청구항 312

삭제

청구항 313

삭제

청구항 314

삭제

청구항 315

삭제

청구항 316

삭제

청구항 317

삭제

청구항 318

삭제

청구항 319

삭제

청구항 320

삭제

청구항 321

삭제

청구항 322

삭제

청구항 323

삭제

청구항 324

삭제

청구항 325

삭제

청구항 326

삭제

청구항 327

삭제

청구항 328

삭제

청구항 329

삭제

청구항 330

삭제

청구항 331

삭제

청구항 332

삭제

청구항 333

삭제

청구항 334

삭제

청구항 335

삭제

청구항 336

삭제

청구항 337

삭제

청구항 338

삭제

청구항 339

삭제

청구항 340

삭제

청구항 341

삭제

청구항 342

삭제

청구항 343

삭제

청구항 344

삭제

청구항 345

삭제

청구항 346

삭제

청구항 347

삭제

청구항 348

삭제

청구항 349

삭제

청구항 350

삭제

청구항 351

삭제

청구항 352

삭제

청구항 353

삭제

청구항 354

삭제

청구항 355

삭제

청구항 356

삭제

청구항 357

삭제

청구항 358

삭제

청구항 359

삭제

청구항 360

삭제

청구항 361

삭제

청구항 362

삭제

청구항 363

삭제

청구항 364

삭제

청구항 365

삭제

청구항 366

삭제

청구항 367

삭제

청구항 368

삭제

청구항 369

삭제

청구항 370

삭제

청구항 371

삭제

청구항 372

삭제

청구항 373

삭제

청구항 374

삭제

청구항 375

삭제

청구항 376

삭제

청구항 377

삭제

청구항 378

삭제

청구항 379

삭제

청구항 380

삭제

청구항 381

삭제

청구항 382

삭제

청구항 383

삭제

청구항 384

삭제

청구항 385

삭제

청구항 386

삭제

청구항 387

삭제

청구항 388

삭제

청구항 389

삭제

청구항 390

삭제

청구항 391

삭제

청구항 392

삭제

청구항 393

삭제

청구항 394

삭제

청구항 395

삭제

청구항 396

삭제

청구항 397

삭제

청구항 398

삭제

청구항 399

삭제

청구항 400

삭제

청구항 401

삭제

청구항 402

삭제

청구항 403

삭제

청구항 404

삭제

청구항 405

삭제

청구항 406

삭제

청구항 407

삭제

청구항 408

삭제

청구항 409

삭제

청구항 410

삭제

청구항 411

삭제

청구항 412

삭제

청구항 413

삭제

청구항 414

삭제

청구항 415

삭제

청구항 416

삭제

청구항 417

삭제

청구항 418

삭제

청구항 419

삭제

청구항 420

삭제

청구항 421

삭제

청구항 422

삭제

청구항 423

삭제

청구항 424

삭제

청구항 425

삭제

청구항 426

삭제

청구항 427

삭제

청구항 428

삭제

청구항 429

삭제

청구항 430

삭제

청구항 431

삭제

청구항 432

삭제

청구항 433

삭제

청구항 434

삭제

청구항 435

삭제

청구항 436

삭제

청구항 437

삭제

청구항 438

삭제

청구항 439

삭제

청구항 440

삭제

청구항 441

삭제

청구항 442

삭제

청구항 443

삭제

청구항 444

삭제

청구항 445

삭제

청구항 446

삭제

청구항 447

삭제

청구항 448

삭제

청구항 449

삭제

청구항 450

삭제

청구항 451

삭제

청구항 452

삭제

청구항 453

삭제

청구항 454

삭제

청구항 455

삭제

청구항 456

삭제

청구항 457

삭제

청구항 458

삭제

청구항 459

삭제

청구항 460

삭제

청구항 461

삭제

청구항 462

삭제

청구항 463

삭제

청구항 464

삭제

청구항 465

삭제

청구항 466

삭제

청구항 467

삭제

청구항 468

삭제

청구항 469

삭제

청구항 470

삭제

청구항 471

삭제

청구항 472

삭제

청구항 473

삭제

청구항 474

삭제

청구항 475

삭제

청구항 476

삭제

청구항 477

삭제

청구항 478

삭제

청구항 479

삭제

청구항 480

삭제

청구항 481

삭제

청구항 482

삭제

청구항 483

삭제

청구항 484

삭제

청구항 485

삭제

청구항 486

삭제

청구항 487

삭제

청구항 488

삭제

청구항 489

삭제

청구항 490

삭제

청구항 491

삭제

청구항 492

삭제

청구항 493

삭제

청구항 494

삭제

청구항 495

삭제

청구항 496

삭제

청구항 497

삭제

청구항 498

삭제

청구항 499

삭제

청구항 500

삭제

청구항 501

삭제

청구항 502

삭제

청구항 503

삭제

청구항 504

삭제

청구항 505

삭제

청구항 506

삭제

청구항 507

삭제

청구항 508

삭제

청구항 509

삭제

청구항 510

삭제

청구항 511

삭제

청구항 512

삭제

청구항 513

삭제

청구항 514

삭제

청구항 515

삭제

청구항 516

삭제

청구항 517

삭제

청구항 518

삭제

청구항 519

삭제

청구항 520

삭제

청구항 521

삭제

청구항 522

삭제

청구항 523

삭제

청구항 524

삭제

청구항 525

삭제

청구항 526

삭제

청구항 527

삭제

청구항 528

삭제

청구항 529

삭제

청구항 530

삭제

청구항 531

삭제

청구항 532

삭제

청구항 533

삭제

청구항 534

삭제

청구항 535

삭제

청구항 536

삭제

청구항 537

삭제

청구항 538

삭제

청구항 539

삭제

청구항 540

삭제

청구항 541

삭제

청구항 542

삭제

청구항 543

삭제

청구항 544

삭제

청구항 545

삭제

청구항 546

삭제

청구항 547

삭제

청구항 548

삭제

청구항 549

삭제

청구항 550

삭제

청구항 551

삭제

청구항 552

삭제

청구항 553

삭제

청구항 554

삭제

청구항 555

삭제

청구항 556

삭제

청구항 557

삭제

청구항 558

삭제

청구항 559

삭제

청구항 560

삭제

청구항 561

삭제

청구항 562

삭제

청구항 563

삭제

청구항 564

삭제

청구항 565

삭제

청구항 566

삭제

청구항 567

삭제

청구항 568

삭제

청구항 569

삭제

청구항 570

삭제

청구항 571

삭제

청구항 572

삭제

청구항 573

삭제

청구항 574

삭제

청구항 575

삭제

청구항 576

삭제

청구항 577

삭제

청구항 578

삭제

청구항 579

삭제

청구항 580

삭제

청구항 581

삭제

청구항 582

삭제

청구항 583

삭제

청구항 584

삭제

청구항 585

삭제

청구항 586

삭제

청구항 587

삭제

청구항 588

삭제

청구항 589

삭제

청구항 590

삭제

청구항 591

삭제

청구항 592

삭제

청구항 593

삭제

청구항 594

삭제

청구항 595

삭제

발명의 설명

기술 분야

[0001] 관련 출원

[0002] 본 출원은 2017년 6월 5일에 출원된 미국 가출원 62/515,285; 2017년 7월 14일에 출원된 미국 가출원 62/532,905; 2017년 7월 14일에 출원된 미국 가출원 62/532,949; 2017년 7월 14일에 출원된 미국 가출원 62/532,971; 2017년 9월 5일에 출원된 미국 가출원 62/554,425; 2017년 10월 30일에 출원된 미국 가출원 62/578,957; 및 2018년 3월 20일에 출원된 미국 가출원 62/645,703에 대한 우선권을 주장한다. 이들 관련 출원의 각각의 내용은 전체가 본 명세서에 참고로 포함된다.

[0003] 서열 목록의 참조

[0004] 본 출원은 전자 형식의 서열 목록과 함께 제출된다. 서열 목록은 2018년 3월 20일 생성된 3,181 바이트 크기의, Sequence_Listing_BDCRI_033A.txt라는 제목의 파일로서 제공된다. 서열 목록의 전자 형식에서의 정보는 전체가 본 명세서에 참고로 포함된다.

[0005] 기술분야

[0006] 본 개시 내용은 일반적으로 분자 생물학 분야에 관한 것으로, 예를 들어 상이한 샘플의 세포들을 식별하고, 분자 바코딩을 사용하여 세포 성분 사이의 상호작용을 검출하는 것에 관한 것이다.

배경 기술

[0007] 현재 기술은, 각각의 세포가 구획 내에서 바코딩된 시약 비드(bead)와 공동-국재화됨(co-localized)에 따라, 세포 특이적 올리고뉴클레오타이드 바코드를 개별 세포로부터의 폴리(A) mRNA 분자에 부착시킴으로써 대량 병렬(massively parallel) 방식(예를 들어, 10000개 초과 세포)으로 단일 세포의 유전자 발현의 측정을 가능하게 한다. 유전자 발현은 단백질 발현에 영향을 줄 수 있다. 단백질-단백질 상호작용은 유전자 발현 및 단백질 발현에 영향을 줄 수 있다. 단백질 발현을 정량적으로 분석하고, 세포 내에서의 단백질 발현과 유전자 발현을 동시에 측정하고, 세포 내에서의 단백질-단백질 상호작용을 결정할 수 있는 시스템 및 방법에 대한 필요성이 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0008] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱(indexing) 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포는 하나 이상의 항원 표적을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 하나의 샘플 인덱싱 조성물은 2개 이상의 항원 결합 시약(예를 들어, 단백질 결합 시약 및 항체)을 포함하고, 2개 이상의 항원 결합 시약 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합되고, 2개 이상의 항원 결합 시약 중 적어도 하나는 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는(예를 들어, 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는) 단계를 포함한다. 방법은, 예를 들어 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.
- [0009] 일부 구현예에서, 샘플 인덱싱 서열은 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이이다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 100개 또는 1000개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.
- [0010] 일부 구현예에서, 항원 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 항원 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 항원 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.
- [0011] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 항원 표적 중 적어도 하나는 세포 표면 상에 또는 세포 내부에 존재할 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.
- [0012] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 항원 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.
- [0013] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일(tail)을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.
- [0014] 일부 구현예에서, 항원 표적은 세포외 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양

항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다. 항원 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 항원 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 항원 결합 시약을 포함할 수 있다. 항원 결합 시약과 제2 항원 결합 시약은 동일할 수 있다.

[0015] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0016] 일부 구현예에서, 복수의 바코드는 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 입자는 비드일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스(Sepharose), 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.

[0017] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0018] 일부 구현예에서, 방법은 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0019] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적(stochastic) 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0020] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물(copy)을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 복수의 바코딩된 표적의 서열분석 데이터를 획득하기

전에, 방법은 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 주계적 바코드를 사용하여 세포의 복수의 표적을 주계적으로 바코딩하여 복수의 주계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0021] 일부 구현예에서, 복수의 샘플 인덱싱 조성물 각각은 항원 결합 시약을 포함한다. 2개 이상의 항원 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 동일할 수 있다. 2개 이상의 항원 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 항원 결합 시약을 포함할 수 있다.

[0022] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 하나의 샘플 인덱싱 조성물은 2개 이상의 세포 성분 결합 시약(예를 들어, 항원 결합 시약 또는 항체)을 포함하고, 2개 이상의 세포 성분 결합 시약 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합되고, 2개 이상의 세포 성분 결합 시약 중 적어도 하나는 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 방법은 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0023] 일부 구현예에서, 샘플 인덱싱 서열은 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이이다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0024] 일부 구현예에서, 세포 성분 결합 시약은 세포 표면 결합 시약, 항체, 테트라머, 압타머, 단백질 스캐폴드, 인테그린, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0025] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 세포 성분 표적 중 적어도 하나는 세포 표면 상에 발현될 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0026] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0027] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일을 포함할 수 있다.

다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.

[0028] 일부 구현예에서, 세포 성분 표적은 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다. 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 세포 성분 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0029] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함할 수 있다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0030] 일부 구현예에서, 복수의 바코드는 입자 내에 봉입된다. 입자는 비드일 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.

[0031] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0032] 일부 구현예에서, 방법은 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0033] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0034] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물을

바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0035] 일부 구현예에서, 복수의 샘플 인덱싱 조성물 각각은 세포 성분 결합 시약을 포함한다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 동일할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다.

[0036] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 하나의 샘플 인덱싱 조성물은 2개 이상의 세포 성분 결합 시약을 포함하고, 2개 이상의 세포 성분 결합 시약 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합되고, 2개 이상의 세포 성분 결합 시약 중 적어도 하나는 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 및 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 방법은, 예를 들어 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0037] 일부 구현예에서, 샘플 인덱싱 서열은 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이이다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 100개 또는 1000개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0038] 일부 구현예에서, 항원 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 항원 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 항원 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0039] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 항원 표적 중 적어도 하나는 세포 표면 상에 발현될 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0040] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 항원 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0041] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일을 포함할 수 있다.

다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.

- [0042] 일부 구현예에서, 항원 표적은 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다. 항원 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 항원 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 항원 결합 시약을 포함할 수 있다. 항원 결합 시약과 제2 항원 결합 시약은 동일할 수 있다.
- [0043] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 복수의 바코드를 사용하여 복수의 샘플 인덱싱 조성물의 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함한다.
- [0044] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함할 수 있다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.
- [0045] 일부 구현예에서, 복수의 바코드는 입자 상에 고정화된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 입자는 비드일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.
- [0046] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0047] 일부 구현예에서, 방법은 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.
- [0048] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 체계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열의 존재 또는 부재를 식별하는 단계를 포함할 수 있다. 샘플 인덱싱 서열의 존재 또는 부재를 식별하는 단계는 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하

여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 서열분석 데이터에서 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 상응하는 복수의 샘플 인덱싱 올리고뉴클레오타이드 중 복제된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함할 수 있다.

[0049] 일부 구현예에서, 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하기 전에, 복제 어댑터(replicating adaptor)를 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 결합하는 (ligating) 단계를 포함한다. 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 결합된 복제 어댑터를 사용하여 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0050] 일부 구현예에서, 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하기 전에, 포획 프로브를 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 연장하여 포획 프로브와 회합된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하는 단계는 포획 프로브와 회합된 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0051] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 체계적 바코드를 사용하여 세포의 복수의 표적을 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0052] 일부 구현예에서, 복수의 샘플 인덱싱 조성물 각각은 항원 결합 시약을 포함한다. 2개 이상의 항원 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 동일할 수 있다. 2개 이상의 항원 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 항원 결합 시약을 포함할 수 있다.

[0053] 본 명세서의 개시 내용은 복수의 샘플 인덱싱 조성물을 포함한다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 항원 결합 시약을 포함할 수 있다. 2개 이상의 항원 결합 시약 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 2개 이상의 항원 결합 시약 중 적어도 하나는 적어도 하나의 항원 표적에 특이적으로 결합할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 샘플의 하나 이상의 세포의 샘플 기원을 식별하기 위한 샘플 인덱싱 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0054] 일부 구현예에서, 샘플 인덱싱 서열은 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이의 뉴클레오타이드 서열을 포함한다. 복수의 샘플 인덱싱 조성물 중 적어도 10, 100, 또는 1000개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0055] 일부 구현예에서, 항원 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 항원 결합 시약에 접합될 수 있다. 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 항원 결합 시약의

분자에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

- [0056] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 종의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포, 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.
- [0057] 일부 구현예에서, 항원 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다. 항원 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 항원 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 항원 결합 시약을 포함할 수 있다. 항원 결합 시약과 제2 항원 결합 시약은 동일할 수 있다.
- [0058] 본 명세서의 개시 내용은 제어 입자 조성물을 포함한다. 일부 구현예에서, 제어 입자 조성물은 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함하고, 복수의 제어 입자 올리고뉴클레오타이드 각각은 제어 바코드 서열 및 폴리(dA) 영역을 포함한다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2개는 상이한 제어 바코드 서열을 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 분자 표지 서열을 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 범용 프라이머를 위한 결합 부위를 포함할 수 있다.
- [0059] 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 5, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 약 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 3, 5, 10, 100개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.
- [0060] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제1 제어 바코드 서열을 각각 포함하는 복수의 제1 제어 입자 올리고뉴클레오타이드, 및 제2 제어 바코드 서열을 각각 포함하는 복수의 제2 제어 입자 올리고뉴클레오타이드를 포함하고, 제1 제어 바코드 서열과 제2 제어 바코드 서열은 상이한 서열을 갖는다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 대략 동일할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 상이할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수는 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 적어도 2배, 10배, 100배, 또는 그 이상일 수 있다.
- [0061] 일부 구현예에서, 제어 바코드 서열은 종의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.
- [0062] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 통해 제어 입자와 회합된다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결(linkage), 또는 이들의 임의의 조합을 포함할 수 있다.
- [0063] 일부 구현예에서, 제어 입자의 직경은 약 1 내지 1000 마이크로미터, 약 10 내지 100 마이크로미터, 7.5 마이크로미터, 또는 이들의 조합이다.
- [0064] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 고정화된다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 부분적으로 고정화될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 봉입될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 부분적으로 봉입될 수

있다. 제어 입자는 붕괴성일 수 있다. 제어 입자는 비드일 수 있다. 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 제어 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 제어 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0065] 일부 구현예에서, 제어 입자는 검출가능한 모이어티(moiety)와 회합된다. 제어 입자 올리고뉴클레오타이드는 검출가능한 모이어티와 회합될 수 있다.

[0066] 일부 구현예에서, 제어 입자는 복수의 제1 단백질 결합 시약과 회합되고, 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 하나와 회합된다. 제1 단백질 결합 시약은 제1 항체를 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 링커를 통해 제1 단백질 결합 시약에 접합될 수 있다. 제1 제어 입자 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 단백질 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0067] 일부 구현예에서, 제1 단백질 결합 시약은 동일한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드 중 2개 이상과 회합된다. 제1 단백질 결합 시약은 상이한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드 중 2개 이상과 회합될 수 있다. 일부 구현예에서, 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 어느 것과도 회합되지 않는다. 제어 입자 올리고뉴클레오타이드와 회합된 제1 단백질 결합 시약 및 임의의 제어 입자 올리고뉴클레오타이드와 회합되지 않은 제1 단백질 결합 시약은 동일한 단백질 결합 시약일 수 있다.

[0068] 일부 구현예에서, 제어 입자는 복수의 제2 단백질 결합 시약과 회합된다. 복수의 제2 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 하나와 회합될 수 있다. 제1 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오타이드와 제2 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오타이드는 상이한 제어 바코드 서열을 포함할 수 있다. 제1 단백질 결합 시약과 제2 단백질 결합 시약은 동일한 단백질 결합 시약일 수 있다.

[0069] 일부 구현예에서, 제1 단백질 결합 시약은 파트너 결합 시약과 회합될 수 있고, 제1 단백질 결합 시약은 파트너 결합 시약을 사용하여 제어 입자와 회합된다. 파트너 결합 시약은 파트너 항체를 포함할 수 있다. 파트너 항체는 항-고양이 항체, 항-닭 항체, 항-소 항체, 항-개 항체, 항-당나귀 항체, 항-염소 항체, 항-기니피그 항체, 항-햄스터 항체, 항-말 항체, 항-인간 항체, 항-라마 항체, 항-원숭이 항체, 항-마우스 항체, 항-돼지 항체, 항-토끼 항체, 항-랫 항체, 항-양 항체, 또는 이들의 조합을 포함할 수 있다. 파트너 항체는 면역글로불린 G(IgG), F(ab') 단편, F(ab')₂ 단편, 이들의 조합, 또는 이들의 단편을 포함할 수 있다.

[0070] 일부 구현예에서, 제1 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다. 제2 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다.

[0071] 본 명세서의 개시 내용은 표적의 수를 결정하기 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 추계적 바코드를 사용하여, 복수의 세포 중의 세포의 복수의 표적 및 복수의 제어 입자 올리고뉴클레오타이드를 추계적으로 바코딩하여, 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 생성하는 단계를 포함하며, 복수의 추계적 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 추계적 바코드 중 적어도 2개의 추계적 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 추계적 바코드 중 적어도 2개의 추계적 바코드는 동일한 세포 표지 서열을 포함하고, 제어 입자 조성물은 복수의 제어 입자 올리고뉴클레오타이드와 회합된 제어 입자를 포함하고, 복수의 제어 입자 올리고뉴클레오타이드 각각은 제어 바코드 서열, 및 복수의 추계적 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적(pseudo-target) 영역을 포함한다. 방법은, 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 서열분석 데이터에서 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 계수하는 단계를 포함할 수 있다. 방법은, 복수의 표적 중 적어도 하나의 표적에 대하여, 서열분석 데이터에서 표적과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 계수하는 단계; 및 표적의 수를 예측하는 단계로서, 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 분자 표지 서열의 수 및

제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수와 상관되는 단계를 포함한다.

- [0072] 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다. 의사-표적 영역은 표적의 하위서열(subsequence)을 포함할 수 있다. 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 임의의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 5, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 3, 5, 10, 100개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.
- [0073] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제1 제어 바코드 서열을 각각 포함하는 복수의 제1 제어 입자 올리고뉴클레오타이드, 및 제2 제어 바코드 서열을 각각 포함하는 복수의 제2 제어 입자 올리고뉴클레오타이드를 포함한다. 제1 제어 바코드 서열과 제2 제어 바코드 서열은 상이한 서열을 가질 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 대략 동일할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 상이할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수는 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 적어도 2배, 10배, 100배, 또는 그 이상일 수 있다.
- [0074] 일부 구현예에서, 서열분석 데이터에서 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 계수하는 단계는, 서열분석 데이터에서 제1 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 계수하는 단계; 및 서열분석 데이터에서 제2 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 계수하는 단계를 포함한다. 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 분자 표지 서열의 수, 제1 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수, 및 제2 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수와 상관될 수 있다. 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 분자 표지 서열의 수, 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수, 및 제어 바코드 서열을 포함하는 복수의 제어 입자 올리고뉴클레오타이드의 수와 상관될 수 있다. 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 분자 표지 서열의 수, 및 제어 바코드 서열을 포함하는 복수의 제어 입자 올리고뉴클레오타이드의 수와 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수의 비와 상관될 수 있다.
- [0075] 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 세포의 유전체 서열과 상동성이 아니다. 제어 입자 올리고뉴클레오타이드는 종의 유전체 서열과 상동성이 아닐 수 있다. 제어 입자 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.
- [0076] 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 링커를 통해 제어 입자에 접합될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 통해 제어 입자와 회합될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 포함할 수 있다. 화학 기는 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.
- [0077] 일부 구현예에서, 제어 입자의 직경은 약 1 내지 1000 마이크로미터, 약 10 내지 100 마이크로미터, 약 7.5 마이크로미터, 또는 이들의 조합이다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 고정화된다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 부분적으로 고정화될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 제어 입자 내에 봉입될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 부분적으로 봉입될 수 있다.
- [0078] 일부 구현예에서, 방법은 복수의 표적과 제어 입자 및 복수의 제어 입자 올리고뉴클레오타이드를 추계적으로 바코딩하기 전에, 제어 입자로부터 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나를 방출하는 단계를 포함한다.
- [0079] 일부 구현예에서, 제어 입자는 붕괴성이다. 제어 입자는 제어 입자 비드일 수 있다. 제어 입자 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드,

드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 제어 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 제어 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0080] 일부 구현예에서, 제어 입자는 검출가능한 모이어티와 회합된다. 제어 입자 올리고뉴클레오티드는 검출가능한 모이어티와 회합될 수 있다.

[0081] 일부 구현예에서, 제어 입자는 복수의 제1 단백질 결합 시약과 회합될 수 있고, 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오티드 중 하나와 회합될 수 있다. 제1 단백질 결합 시약은 제1 항체를 포함할 수 있다. 제어 입자 올리고뉴클레오티드는 링커를 통해 제1 단백질 결합 시약에 접합될 수 있다. 제1 제어 입자 올리고뉴클레오티드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 단백질 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0082] 일부 구현예에서, 제1 단백질 결합 시약은 동일한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오티드 중 2개 이상과 회합될 수 있다. 제1 단백질 결합 시약은 상이한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오티드 중 2개 이상과 회합될 수 있다. 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오티드 중 어느 것과도 회합될 수 없다. 제어 입자 올리고뉴클레오티드와 회합된 제1 단백질 결합 시약과 임의의 제어 입자 올리고뉴클레오티드와 회합되지 않은 제1 단백질 결합 시약은 동일한 단백질 결합 시약일 수 있다. 제어 입자는 복수의 제2 단백질 결합 시약과 회합될 수 있다. 복수의 제2 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오티드 중 하나와 회합될 수 있다. 제1 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오티드와 제2 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오티드는 상이한 제어 바코드 서열을 포함할 수 있다. 제1 단백질 결합 시약과 제2 단백질 결합 시약은 동일한 단백질 결합 시약일 수 있다.

[0083] 일부 구현예에서, 제1 단백질 결합 시약은 파트너 결합 시약과 회합되고, 제1 단백질 결합 시약은 파트너 결합 시약을 사용하여 제어 입자와 회합된다. 파트너 결합 시약은 파트너 항체를 포함할 수 있다. 파트너 항체는 항-고양이 항체, 항-닭 항체, 항-소 항체, 항-개 항체, 항-당나귀 항체, 항-염소 항체, 항-기니피그 항체, 항-햄스터 항체, 항-말 항체, 항-인간 항체, 항-라마 항체, 항-원숭이 항체, 항-마우스 항체, 항-돼지 항체, 항-토끼 항체, 항-랫 항체, 항-양 항체, 또는 이들의 조합을 포함할 수 있다. 파트너 항체는 면역글로불린 G(IgG), F(ab') 단편, F(ab')₂ 단편, 이들의 조합, 또는 이들의 단편을 포함할 수 있다.

[0084] 일부 구현예에서, 제1 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다. 제2 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다.

[0085] 일부 구현예에서, 추계적 바코드는 범용 프라이머를 위한 결합 부위를 포함한다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0086] 일부 구현예에서, 복수의 추계적 바코드는 바코딩 입자와 회합된다. 복수의 추계적 바코드 중 적어도 하나의 추계적 바코드는 바코딩 입자 상에 고정화될 수 있다. 복수의 추계적 바코드 중 적어도 하나의 추계적 바코드는 바코딩 입자 상에 부분적으로 고정화될 수 있다. 복수의 추계적 바코드 중 적어도 하나의 추계적 바코드는 바코딩 입자 내에 봉입될 수 있다. 복수의 추계적 바코드 중 적어도 하나의 추계적 바코드는 바코딩 입자 내에 부분적으로 봉입될 수 있다.

[0087] 일부 구현예에서, 바코딩 입자는 붕괴될 수 있다. 바코딩 입자는 바코딩 비드일 수 있다. 바코딩 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 바코딩 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 바코딩 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0088] 일부 구현예에서, 바코딩 입자의 추계적 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함한다. 일부 구현예에서, 추계적 바코드의 분자 표지 서열은 랜덤 서열

을 포함한다. 일부 구현예에서, 바코딩 입자는 적어도 10000개의 추계적 바코드를 포함한다.

- [0089] 일부 구현예에서, 복수의 추계적 바코드를 사하여 복수의 표적 및 복수의 제어 입자 올리고뉴클레오타이드를 추계적으로 바코딩하는 단계는, 복수의 추계적 바코드를 복수의 표적 중 표적 및 복수의 제어 입자 올리고뉴클레오타이드 중 제어 입자 올리고뉴클레오타이드와 접촉시켜 표적 및 제어 입자 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 생성하는 단계; 및 표적 및 제어 입자 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 연장하여 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 추계적 바코드를 연장하는 단계는 DNA 중합효소, 역전사효소, 또는 이들의 조합을 사용하여 추계적 바코드를 연장하는 단계를 포함할 수 있다.
- [0090] 일부 구현예에서, 방법은 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분 또는 분자 표지 서열의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분, 또는 분자 표지 서열의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.
- [0091] 본 명세서의 개시 내용은 또한 서열분석 제어용 키트를 포함한다. 일부 구현예에서, 키트는, 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함하는 제어 입자 조성물을 포함하며, 복수의 제어 입자 올리고뉴클레오타이드의 각각은 제어 바코드 서열 및 폴리(dA) 영역을 포함한다.
- [0092] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2개는 상이한 제어 바코드 서열을 포함한다. 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 임의의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 5, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 3, 5, 10, 100개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.
- [0093] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제1 제어 바코드 서열을 각각 포함하는 복수의 제1 제어 입자 올리고뉴클레오타이드, 및 제2 제어 바코드 서열을 각각 포함하는 복수의 제2 제어 입자 올리고뉴클레오타이드를 포함한다. 제1 제어 바코드 서열과 제2 제어 바코드 서열은 상이한 서열을 가질 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 대략 동일할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 상이할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수는 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 적어도 2배, 10배, 100배, 또는 그 이상일 수 있다.
- [0094] 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 세포의 유전체 서열과 상동성이 아니다. 제어 입자 올리고뉴클레오타이드는 종의 유전체 서열과 상동성이 아닐 수 있다. 제어 입자 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.
- [0095] 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 링커를 통해 제어 입자에 접합될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 통해 제어 입자에 회합될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 포함할 수 있다. 화학 기는 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트랩타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.
- [0096] 일부 구현예에서, 제어 입자의 직경은 약 1 내지 1000 마이크로미터, 약 10 내지 100 마이크로미터, 약 7.5 마이크로미터, 또는 이들의 조합이다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 고정화된다. 복수

의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 부분적으로 고정화될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 봉입될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 부분적으로 봉입될 수 있다.

[0097] 일부 구현예에서, 키트는 복수의 바코드를 포함한다. 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함할 수 있고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함할 수 있다. 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함한다.

[0098] 일부 구현예에서, 복수의 바코드는 바코딩 입자와 회합될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 부분적으로 고정화된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 부분적으로 봉입될 수 있다. 바코딩 입자는 붕괴성일 수 있다. 바코딩 입자는 제2 비드일 수 있다. 비드는 세파로스 비드, 스트랩타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 바코딩 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 바코딩 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0099] 일부 구현예에서, 바코딩 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함한다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다. 바코딩 입자는 적어도 10000개의 바코드를 포함할 수 있다. 키트는 DNA 중합효소를 포함할 수 있다. 키트는 중합효소 연쇄 반응(PCR)을 위한 시약을 포함할 수 있다.

[0100] 본 명세서에 개시된 세포 식별을 위한 방법은 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 단백질 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 하나 이상의 단백질 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 회합된 세포 표지 서열을 식별하는 단계; 및 획득된 서열분석 데이터로부터 세포 표지 서열과 연관된 서열분석 데이터를 제거하는 단계를 포함할 수 있다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다. 본 명세서에 기재된 바와 같이, 제1 복수의 세포는 제2 복수의 세포와 상이한 조직 또는 기관으로부터 획득되거나 유래될 수 있고, 제1 복수의 세포와 제2 복수의 세포는 동일하거나 상이한 대상체(예를 들어, 포유동물)로부터 유래될 수 있다. 예를 들어, 제1 복수의 세포는 제2 복수의 세포와 상이한 인간 대상체로부터 획득되거나 유래될 수 있다. 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포는 동일한 인간 대상체의 상이한 조직으로부터 획득되거나 유래된다.

[0101] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계는 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제1 샘플 인덱싱 조성물과 접촉시키는 단계; 및 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제2 샘플 인덱싱 조성물과 접촉시키는 단계를 포함한다.

[0102] 본 명세서에 기재된 바와 같이, 샘플 인덱싱 서열은, 예를 들어 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 일

부 구현예에서, 복수의 샘플 인덱싱 조성물 중 적어도 10, 100, 1000개, 또는 그 이상의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

- [0103] 일부 구현예에서, 단백질 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 단백질 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 단백질 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결 또는 이들의 임의의 조합을 포함할 수 있다.
- [0104] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 단백질 표적 중 적어도 하나는 세포 표면 상에 존재할 수 있다.
- [0105] 일부 구현예에서, 방법은 2개의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함한다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 제1 복수의 세포 및 제2 복수의 세포 중의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 샘플 인덱싱 조성물의 적어도 하나의 단백질 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함한다.
- [0106] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 단백질 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 단백질 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리, 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 단백질 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.
- [0107] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 하나 이상의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성이 아닐 수 있다. 종은 미-포유동물 종일 수 있다. 미-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합이다.
- [0108] 일부 구현예에서, 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함한다. 복수의 샘플은 포유동물 세포, 세균 세포, 바이러스 세포, 효모 세포, 진균 세포, 또는 이들의 임의의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드는 포획 서열을 포함하는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다.
- [0109] 일부 구현예에서, 단백질 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 단백질 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 단백질 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 단백질 표적은 10 내지 100개의 상이한 단백질 표적을 포함하는 군으로부터 선택될 수 있다.
- [0110] 일부 구현예에서, 단백질 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합된다. 단백질 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 단백질 결합 시약을 포함할 수 있다. 단백질 결합 시약과 제2 단백질 결합 시약은 동일할 수 있다.
- [0111] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.
- [0112] 일부 구현예에서, 복수의 바코드는 입자와 회합될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다. 입자는 비드일 수 있다. 비드는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단

백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0113] 일부 구현예에서, 단백질 결합 시약은 검출가능한 모이어티와 회합된다. 일부 구현예에서, 입자는 검출가능한 모이어티와 회합된다. 샘플 인덱싱 올리고뉴클레오타이드는 광학 모이어티와 회합된다.

[0114] 일부 구현예에서, 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.

[0115] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0116] 일부 구현예에서, 방법은 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열 분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표지의 샘플 기원을 식별하는 단계를 포함한다.

[0117] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0118] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다.

[0119] 일부 구현예에서, 방법은 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함한다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0120] 본 명세서의 개시 내용은 서열분석 제어에 사용될 수 있는 방법 및 조성물을 포함한다. 일부 구현예에서, 서열분석 제어를 위한 방법은, 복수의 세포 중 하나 이상의 세포를 복수의 제어 조성물 중 하나의 제어 조성물과 접촉시키는 단계로서, 복수의 세포 중 세포는 복수의 표적 및 복수의 단백질 표적을 포함하고, 복수의 제어 조성물 각각은 제어 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 복수의 단백질

표적 중 적어도 하나에 특이적으로 결합할 수 있고, 제어 올리고뉴클레오타이드는 제어 바코드 서열, 및 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적 영역을 포함하는, 단계; 복수의 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는 단계; 복수의 바코딩된 제어 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 적어도 하나의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계를 포함한다. 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다.

[0121] 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.

[0122] 일부 구현예에서, 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 결정하는 단계; 및 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 방법은, 결정된 하나 이상의 세포의 수에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함한다. 방법은, 결정된 하나 이상의 세포의 수와 복수의 세포의 수의 비에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함한다.

[0123] 일부 구현예에서, 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서의 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계; 및 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계는, 서열분석 데이터에서의 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 최대수의 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계를 포함할 수 있다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계는 최대수의 별개의 서열을 갖는 분자 표지 서열의 수와, 최대수의 별개의 서열을 갖는 분자 표지 서열의 수와 연관된 서열분석 데이터에서의 세포 표지의 수의 플롯을 생성하는 단계; 및 하나 이상의 세포의 수로서 플롯에서의 컷오프를 결정하는 단계를 포함할 수 있다.

[0124] 일부 구현예에서, 제어 올리고뉴클레오타이드는 복수의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.

[0125] 일부 구현예에서, 방법은 제어 올리고뉴클레오타이드를 바코딩하기 전에 단백질 결합 시약으로부터 제어 올리고뉴클레오타이드를 방출하는 단계를 포함한다. 일부 구현예에서, 방법은 복수의 제어 조성물 중 결합되지 않은 제어 조성물을 제거하는 단계를 포함한다. 결합되지 않은 제어 조성물을 제거하는 단계는 복수의 세포 중 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 제어 조성물의 적어도 하나의 단백질 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다.

[0126] 일부 구현예에서, 복수의 단백질 표적 중 적어도 하나는 세포 표면 상에 존재한다. 복수의 단백질 표적 중 적어도 하나는 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합을 포함할 수 있다. 단백질 결합 시약은 항체를 포함할 수 있다. 제어 올리고뉴클레오타이드는 링커를 통해 단백질 결합 시약에 접합될 수 있다. 제어 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 단백질 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함

할 수 있다.

- [0127] 일부 구현예에서, 단백질 결합 시약은 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합된다. 단백질 결합 시약은 상이한 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합될 수 있다. 일부 구현예에서, 복수의 제어 조성물의 제2 단백질 결합 시약은 제어 올리고뉴클레오타이드와 회합되지 않는다. 단백질 결합 시약과 제2 단백질 결합 시약은 동일할 수 있다.
- [0128] 일부 구현예에서, 바코드는 범용 프라이머를 위한 결합 부위를 포함한다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 일부 구현예에서, 복수의 바코드는 바코딩 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 봉입된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 부분적으로 봉입된다. 바코딩 입자는 봉괴성일 수 있다. 바코딩 입자는 바코딩 비드일 수 있다. 바코딩 비드는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 바코딩 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세포로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 바코딩 입자는 봉괴성 하이드로겔 입자를 포함할 수 있다.
- [0129] 일부 구현예에서, 바코딩 입자는 광학 모이어티와 회합된다. 제어 올리고뉴클레오타이드는 광학 모이어티와 회합될 수 있다.
- [0130] 일부 구현예에서, 바코딩 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함한다. 일부 구현예에서, 바코드의 분자 표지 서열은 랜덤 서열을 포함한다. 일부 구현예에서, 바코딩 입자는 적어도 10000개의 바코드를 포함한다.
- [0131] 일부 구현예에서, 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 추계적 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 추계적으로 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 일부 구현예에서, 복수의 바코드를 사용하여 복수의 제어 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 복수의 제어 조성물의 제어 올리고뉴클레오타이드와 접촉시켜 제어 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 제어 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 연장하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소, 역전사효소, 또는 이들의 조합을 사용하여 바코드를 연장하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함한다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.
- [0132] 본 명세서의 개시 내용은 서열분석 제어를 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 세포 중 하나 이상의 세포를 복수의 제어 조성물 중의 제어 조성물과 접촉시키는 단계로서, 복수의 세포 중의 세포는 복수의 표적 및 복수의 결합 표적을 포함하고, 복수의 제어 조성물 각각은 제어 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 복수의 결합 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 제어 올리고뉴클레오타이드는 제어 바코드 서열, 및 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적 영역을 포함하는, 단계; 복수의 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 제어 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 적어도 하나의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계를 포함한다. 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다.
- [0133] 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약

128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.

[0134] 일부 구현예에서, 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는: 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 결정하는 단계; 및 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 일부 구현예에서, 방법은, 결정된 하나 이상의 세포의 수에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함한다. 일부 구현예에서, 방법은, 결정된 하나 이상의 세포의 수와 복수의 세포의 수의 비에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함한다.

[0135] 일부 구현예에서, 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계; 및 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함할 수 있다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계는, 서열분석 데이터에서 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 최대 수의 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계를 포함할 수 있다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계는, 최대 수의 별개의 서열을 갖는 분자 표지 서열의 수와, 최대 수의 별개의 서열을 갖는 분자 표지 서열의 수와 연관된 서열분석 데이터에서 세포 표지의 수의 플롯을 생성하는 단계; 및 하나 이상의 세포의 수로서 플롯에서의 컷오프를 결정하는 단계를 포함할 수 있다.

[0136] 일부 구현예에서, 제어 올리고뉴클레오타이드는 복수의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.

[0137] 일부 구현예에서, 방법은, 제어 올리고뉴클레오타이드를 바코딩하기 전에, 세포 성분 결합 시약으로부터 제어 올리고뉴클레오타이드를 방출하는 단계를 포함한다. 복수의 결합 표적 중 적어도 하나는 세포 표면 상에 발현될 수 있다. 복수의 결합 표적 중 적어도 하나는 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합을 포함할 수 있다. 세포 성분 결합 시약은 세포 표면 결합 시약, 항체, 테트라머, 압타머, 단백질 스캐폴드, 침입체(invasion), 또는 이들의 조합을 포함할 수 있다.

[0138] 일부 구현예에서, 세포 성분 결합 시약의 결합 표적은 10 내지 100개의 상이한 결합 표적을 포함하는 군으로부터 선택된다. 세포 성분 결합 시약의 결합 표적은 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 세포내 단백질, 또는 이들의 임의의 조합을 포함할 수 있다. 제어 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 제어 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 세포 성분 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트랩타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0139] 일부 구현예에서, 세포 성분 결합 시약은 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합된다. 세포 성분 결합 시약은 상이한 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합될 수 있다. 일부 구현예에서, 복수의 제어 조성물 중 제2 세포 성분 결합 시약은 제어 올리고뉴클레오타이드와 회합되지 않는다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0140] 일부 구현예에서, 바코드는 범용 프라이머를 위한 결합 부위를 포함한다. 일부 구현예에서, 표적-결합 영역은 폴리(dT) 영역을 포함한다.

[0141] 일부 구현예에서, 복수의 바코드는 바코딩 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어

도 하나의 바코드는 바코딩 입자 내에 부분적으로 봉입될 수 있다. 바코딩 입자는 붕괴성일 수 있다. 바코딩 입자는 바코딩 비드일 수 있다. 일부 구현예에서, 바코딩 비드는 세파로스 비드, 스트랩타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함한다. 바코딩 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 바코딩 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다. 바코딩 입자는 광학 모이어티와 회합될 수 있다.

[0142] 일부 구현예에서, 제어 올리고뉴클레오타이드는 광학 모이어티와 회합될 수 있다. 일부 구현예에서, 바코딩 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함한다. 일부 구현예에서, 바코드의 분자 표지 서열은 랜덤 서열을 포함한다. 바코딩 입자는 적어도 10000개의 바코드를 포함할 수 있다.

[0143] 일부 구현예에서, 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 추계적 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 추계적으로 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 복수의 제어 조성물의 제어 올리고뉴클레오타이드와 접촉시켜 제어 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 제어 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 연장하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 DNA 중합효소, 역전사효소, 또는 이들의 조합을 사용하여 바코드를 연장하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0144] 일부 구현예에서, 서열분석 제어를 위한 방법은 복수의 세포 중 하나 이상의 세포를 복수의 제어 조성물 중의 제어 조성물과 접촉시키는 단계로서, 복수의 세포 중의 세포는 복수의 표적 및 복수의 단백질 표적을 포함하고, 복수의 제어 조성물 각각은 제어 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 복수의 단백질 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 제어 올리고뉴클레오타이드는 제어 바코드 서열, 및 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적 영역을 포함하는, 단계; 및 복수의 제어 올리고뉴클레오타이드의 적어도 하나의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계를 포함할 수 있다. 의사-표적 영역은 폴리(dA) 영역을 포함할 수 있다.

[0145] 일부 구현예에서, 제어 바코드 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 제어 입자 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2, 10, 100, 1000개, 또는 그 이상은 상이한 제어 바코드 서열을 포함할 수 있다.

[0146] 일부 구현예에서, 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 결정하는 단계; 및 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 방법은, 결정된 하나 이상의 세포의 수에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함할 수 있다. 방법은, 결정된 하나 이상의 세포의 수와 복수의 세포의 수의 비에 기초하여 단일 세포 포획 효율을 결정하는 단계를 포함할 수 있다.

[0147] 일부 구현예에서, 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 특징을 사용하여 하나 이상

의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서의 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계; 및 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계는, 서열분석 데이터에서의 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 최대수의 별개의 서열을 갖는 분자 표지 서열의 수를 결정하는 단계를 포함할 수 있다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 분자 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계는, 최대수의 별개의 서열을 갖는 분자 표지 서열의 수와, 최대수의 별개의 서열을 갖는 분자 표지 서열의 수와 연관된 서열분석 데이터에서의 세포 표지의 수의 플롯을 생성하는 단계; 및 하나 이상의 세포의 수로서 플롯에서의 컷오프를 결정하는 단계를 포함할 수 있다.

[0148] 일부 구현예에서, 제어 올리고뉴클레오타이드는 복수의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.

[0149] 일부 구현예에서, 방법은 제어 올리고뉴클레오타이드를 바코딩하기 전에 단백질 결합 시약으로부터 제어 올리고뉴클레오타이드를 방출하는 단계를 포함한다. 일부 구현예에서, 방법은 복수의 제어 조성물 중 결합되지 않은 제어 조성물을 제거하는 단계를 포함한다. 결합되지 않은 제어 조성물을 제거하는 단계는 복수의 세포 중 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 제어 조성물의 적어도 하나의 단백질 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다.

[0150] 일부 구현예에서, 복수의 단백질 표적 중 적어도 하나는 세포 표면 상에 존재한다. 복수의 단백질 표적 중 적어도 하나는 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합을 포함할 수 있다. 단백질 결합 시약은 항체를 포함할 수 있다. 제어 올리고뉴클레오타이드는 링커를 통해 단백질 결합 시약에 접합될 수 있다. 제어 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 단백질 결합 시약에 가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0151] 일부 구현예에서, 단백질 결합 시약은 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합된다. 단백질 결합 시약은 상이한 동일한 제어 바코드 서열을 갖는 2개 이상의 제어 올리고뉴클레오타이드와 회합될 수 있다. 일부 구현예에서, 복수의 제어 조성물의 제2 단백질 결합 시약은 제어 올리고뉴클레오타이드와 회합되지 않는다. 단백질 결합 시약과 제2 단백질 결합 시약은 동일할 수 있다.

[0152] 일부 구현예에서, 바코드는 범용 프라이머를 위한 결합 부위를 포함한다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 일부 구현예에서, 복수의 바코드는 바코딩 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 봉입된다. 복수의 바코드 중 적어도 하나의 바코드는 바코딩 입자 내에 부분적으로 봉입된다. 바코딩 입자는 봉입성일 수 있다. 바코딩 입자는 바코딩 비드일 수 있다. 바코딩 비드는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 바코딩 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세포로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 바코딩 입자는 봉입성 하이드로겔 입자를 포함할 수 있다.

[0153] 일부 구현예에서, 바코딩 입자는 광학 모이어티와 회합된다. 제어 올리고뉴클레오타이드는 광학 모이어티와 회합될 수 있다.

[0154] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 복수의 바코딩된 제어

올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계를 포함한다.

- [0155] 일부 구현예에서, 바코딩 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함한다. 일부 구현예에서, 바코드의 분자 표지 서열은 랜덤 서열을 포함한다. 일부 구현예에서, 바코딩 입자는 적어도 10000개의 바코드를 포함한다.
- [0156] 일부 구현예에서, 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 추계적 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 추계적으로 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 일부 구현예에서, 복수의 바코드를 사용하여 복수의 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 복수의 제어 조성물의 제어 올리고뉴클레오타이드와 접촉시켜 제어 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 제어 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 연장하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소, 역전사효소, 또는 이들의 조합을 사용하여 바코드를 연장하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함한다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.
- [0157] 일부 구현예에서, 세포 식별하기 위한 방법은, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 항원 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 항원 결합 시약을 포함하고, 항원 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드의 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 회합된 세포 표지 서열을 식별하는 단계; 및 획득된 서열분석 데이터로부터 세포 표지 서열과 연관된 서열분석 데이터를 제거하는 단계 및/또는 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석에서 제외하는 단계를 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함한다.
- [0158] 본 명세서의 개시 내용은 또한 멀티플렛 식별을 위한 방법을 포함한다. 멀티플렛 발현 프로파일, 또는 멀티플렛은 다중 세포(multiplet cell)들의 발현 프로파일들을 포함하는 발현 프로파일일 수 있다. 단일 세포의 발현 프로파일을 결정할 때, n 개의 세포가 하나의 세포로서 식별될 수 있고, n 개의 세포의 발현 프로파일은 하나의 세포에 대한 발현 프로파일(멀티플렛 또는 n -플렛 발현 프로파일로 지칭됨)로서 식별될 수 있다. 예를 들어, 바코딩(예를 들어, 추계적 바코딩)을 사용하여 2개의 세포의 발현 프로파일들을 결정할 때, 2개의 세포의 mRNA 분자는 동일한 세포 표지를 갖는 바코드와 회합될 수 있다. 또 다른 예로서, 2개의 세포는 하나의 입자(예를 들어, 비드)와 회합될 수 있다. 입자는 동일한 세포 표지를 갖는 바코드를 포함할 수 있다. 세포를 용해시킨 후에, 2개의 세포 내의 mRNA 분자는 입자의 바코드와 회합될 수 있으며, 이에 따라 동일한 세포 표지와 회합될 수 있다. 더블렛 발현 프로파일은 발현 프로파일의 해석을 왜곡할 수 있다. 멀티플렛은 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 복수의 멀티플렛은 더블렛, 트리플렛, 쿼텟, 퀸텟, 섀스텟, 셉텟, 옥텟, 노넷, 또는 이들의 임의의 조합을 포함할 수 있다.
- [0159] 멀티플렛은 임의의 n -플렛일 수 있다. 일부 구현예에서, n 은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 이들 중 임의의 두 값 사이의 범위이거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 이들 중 임의의 두 값 사이의 범위이다. 일부 구현예에서, n 은 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20이다. 싱글렛은 멀티플렛 발현 프로파일일 아닌 발현 프로파일일 수 있다.
- [0160] 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 샘플을 인덱싱하고 멀티플렛을 식별하는 수행은 합성 멀티플렛 발

현 프로파일을 사용하여 멀티플렛을 식별하는 수행에 비견될 수 있다(2018년 3월 20일에 출원되고 발명의 명칭이 "SYNTHETIC MULTIPLETS FOR MULTIPLETS DETERMINATION"인 미국 출원 15/926977에 기재되어 있으며, 이의 내용은 전체가 본 명세서에 참고로 포함됨). 일부 구현예에서, 멀티플렛은 샘플 인텍싱 올리고뉴클레오타이드 및 합성 멀티플렛 발현 프로파일 둘 모두를 사용하여 식별될 수 있다.

[0161] 일부 구현예에서, 본 명세서에 개시된 멀티플렛 식별 방법은 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인텍싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 항원 표적을 포함하고, 2개의 샘플 인텍싱 조성물 각각은 샘플 인텍싱 올리고뉴클레오타이드와 회합된 항원 결합 시약을 포함하고, 항원 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인텍싱 올리고뉴클레오타이드는 샘플 인텍싱 서열을 포함하고, 2개의 샘플 인텍싱 조성물의 샘플 인텍싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인텍싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인텍싱 서열과 각각 회합된 하나 이상의 다중 세포 표지 서열을 식별하는 단계를 포함한다. 일부 구현예에서, 방법은, 하나 이상의 다중 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 하나 이상의 다중 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석에서 제외하는 단계를 포함한다. 일부 구현예에서, 샘플 인텍싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함한다.

[0162] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인텍싱 조성물과 각각 접촉시키는 단계는, 제1 복수의 세포를 2개의 샘플 인텍싱 조성물 중 제1 샘플 인텍싱 조성물과 접촉시키는 단계; 및 제1 복수의 세포를 2개의 샘플 인텍싱 조성물 중 제2 샘플 인텍싱 조성물과 접촉시키는 단계를 포함한다.

[0163] 일부 구현예에서, 샘플 인텍싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인텍싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 일부 구현예에서, 복수의 샘플 인텍싱 조성물 중 적어도 10, 100, 1000개, 또는 그 이상의 샘플 인텍싱 조성물의 샘플 인텍싱 서열은 상이한 서열을 포함한다.

[0164] 일부 구현예에서, 항원 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인텍싱 올리고뉴클레오타이드는 링커를 통해 항원 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 항원 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트랩타비딘, 비오틴, 아민, 이황화 연결 또는 이들의 임의의 조합을 포함할 수 있다.

[0165] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포 중 적어도 하나는 단일 세포를 포함한다. 하나 이상의 항원 표적 중 적어도 하나는 세포 표면 상에 존재할 수 있다.

[0166] 일부 구현예에서, 방법은, 2개의 샘플 인텍싱 조성물 중 결합되지 않은 샘플 인텍싱 조성물을 제거하는 단계를 포함한다. 결합되지 않은 샘플 인텍싱 조성물을 제거하는 단계는 제1 복수의 세포 및 제2 복수의 세포 중의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인텍싱 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 샘플 인텍싱 조성물의 적어도 하나의 항원 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 제1 복수의 세포 및 제2 복수의 세포 중 하나 이상의 세포를 용해시키는 단계를 포함한다.

[0167] 일부 구현예에서, 샘플 인텍싱 올리고뉴클레오타이드는 항원 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 샘플 인텍싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인텍싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이틀을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인텍싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

- [0168] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 하나 이상의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성이 아닐 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합이다.
- [0169] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포는 종양 세포, 포유동물 세포, 세균 세포, 바이러스 세포, 효모 세포, 진균 세포, 또는 이들의 임의의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드는 포획 서열을 포함하는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다.
- [0170] 일부 구현예에서, 항원 표적은 세포외 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다.
- [0171] 일부 구현예에서, 항원 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합된다. 항원 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 항원 결합 시약을 포함할 수 있다. 항원 결합 시약과 제2 항원 결합 시약은 동일할 수 있다.
- [0172] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.
- [0173] 일부 구현예에서, 복수의 바코드는 입자와 회합될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다. 입자는 비드일 수 있다. 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.
- [0174] 일부 구현예에서, 항원 결합 시약은 검출가능한 모이어티와 회합된다. 일부 구현예에서, 입자는 검출가능한 모이어티와 회합된다. 샘플 인덱싱 올리고뉴클레오타이드는 광학 모이어티와 회합된다. 일부 구현예에서, 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.
- [0175] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0176] 일부 구현예에서, 방법은, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을

증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열 분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함한다.

[0177] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0178] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다.

[0179] 일부 구현예에서, 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함한다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0180] 일부 구현예에서, 세포 식별을 위한 방법은, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열을 식별하는 단계; 및 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석에서 제외하는 단계를 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다.

[0181] 본 명세서의 개시 내용은 멀티플렉트 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱

싱 서열과 각각 회합된 하나 이상의 다중 세포 표지 서열을 식별하는 단계를 포함한다. 일부 구현예에서, 방법은, 하나 이상의 다중 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 하나 이상의 다중 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석에서 제외하는 단계를 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다.

[0182] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계는, 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제1 샘플 인덱싱 조성물과 접촉시키는 단계; 및 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제2 샘플 인덱싱 조성물과 접촉시키는 단계를 포함한다.

[0183] 일부 구현예에서, 샘플 인덱싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인덱싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 일부 구현예에서, 복수의 샘플 인덱싱 조성물 중 적어도 10, 100, 1000개, 또는 그 이상의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0184] 일부 구현예에서, 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약에 가역적 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 이황화 연결 또는 이들의 임의의 조합을 포함할 수 있다.

[0185] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포 중 적어도 하나는 단일 세포를 포함한다. 하나 이상의 세포 성분 표적 중 적어도 하나는 세포 표면 상에 존재할 수 있다.

[0186] 일부 구현예에서, 방법은, 2개의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함한다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 제1 복수의 세포 및 제2 복수의 세포 중의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 샘플 인덱싱 조성물의 적어도 하나의 세포 성분 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포 중 하나 이상의 세포를 용해시키는 단계를 포함한다.

[0187] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 세포 성분 결합 시약으로부터 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프라이트를 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 세포 성분 결합 시약으로부터 샘플 인덱싱 올리고뉴클레오타이드를 분리하는 단계를 포함할 수 있다.

[0188] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 임의의 하나 이상의 세포의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성이 아닐 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합이다.

[0189] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포는 종양 세포, 포유동물 세포, 박테리아 세포, 세균 세포, 효모 세포, 균류 세포, 또는 이들의 임의의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드는 포획 서열을 포함하는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다.

[0190] 일부 구현예에서, 항원 표적은 세포외 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다.

[0191] 일부 구현예에서, 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회

합된다. 세포 성분 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0192] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0193] 일부 구현예에서, 복수의 바코드는 입자와 회합될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다. 입자는 비드일 수 있다. 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0194] 일부 구현예에서, 세포 성분 결합 시약은 검출가능한 모이어티와 회합된다. 일부 구현예에서, 입자는 검출가능한 모이어티와 회합된다. 샘플 인덱싱 올리고뉴클레오타이드는 광학 모이어티와 회합된다.

[0195] 일부 구현예에서, 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.

[0196] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0197] 일부 구현예에서, 방법은, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열 분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열 분석 데이터를 획득하는 단계를 포함할 수 있다. 서열 분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함한다.

[0198] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0199] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를

포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다.

[0200] 일부 구현예에서, 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함한다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0201] 본 명세서의 개시 내용은 세포 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포 각각으로부터의 하나 이상의 세포를 복수의 2개의 샘플 인덱싱 조성물들 중의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 항원 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 항원 결합 시약을 포함하고, 항원 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 및 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포를 식별하는 단계를 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다.

[0202] 본 명세서의 개시 내용은 멀티플렛 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포 각각으로부터의 하나 이상의 세포를 복수의 2개의 샘플 인덱싱 조성물들 중의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 항원 표적을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 항원 결합 시약을 포함하고, 항원 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 및 다중 세포로서 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포를 식별하는 단계를 포함한다.

[0203] 일부 구현예에서, 2개 이상의 샘플 인덱싱 서열과 각각 회합된 세포를 식별하는 단계는, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 분자 표지 서열, 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열을 식별하는 단계를 포함한다. 방법은 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석에서 제외하는 단계를 포함할 수 있다.

[0204] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계는, 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제1 샘플 인덱싱 조성물과 접촉시키는 단계; 및 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제2 샘플 인덱싱 조성물과 접촉시키는 단계를 포함한다.

[0205] 일부 구현예에서, 샘플 인덱싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 60개의 뉴클레오타이드 길이(예를 들어, 45개의 뉴클레오타이드 길이), 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인덱싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 일부 구현예에서, 복수의 샘플 인덱싱 조성물 중 적어도 10, 100, 1000개, 또는 그 이상의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0206] 일부 구현예에서, 항원 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 항원 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커

를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 항원 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 이황화 연결 또는 이들의 임의의 조합을 포함할 수 있다.

[0207] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포 중 적어도 하나는 단일 세포를 포함한다. 하나 이상의 항원 표적 중 적어도 하나는 세포 표면 상에 존재할 수 있다. 일부 구현예에서, 방법은, 2개의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함한다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 제1 복수의 세포 및 제2 복수의 세포 중의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 샘플 인덱싱 조성물의 적어도 하나의 항원 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포 중 하나 이상의 세포를 용해시키는 단계를 포함한다.

[0208] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 항원 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오트레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 항원 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0209] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 하나 이상의 세포 중 임의의 것의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성이 아닐 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합이다.

[0210] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포는 종양 세포, 포유동물 세포, 세균 세포, 바이러스 세포, 효모 세포, 진균 세포, 또는 이들의 임의의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드는 포획 서열을 포함하는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다.

[0211] 일부 구현예에서, 항원 표적은 세포외 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 항원 표적은 10 내지 100개의 상이한 항원 표적을 포함하는 군으로부터 선택될 수 있다.

[0212] 일부 구현예에서, 항원 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합된다. 항원 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 항원 결합 시약을 포함할 수 있다. 항원 결합 시약과 제2 항원 결합 시약은 동일할 수 있다.

[0213] 일부 구현예에서, 복수의 바코드 중 바코드는 표적-결합 영역 및 분자 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0214] 일부 구현예에서, 복수의 바코드는 입자와 회합될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다.. 입자는 비드일 수 있다. 비드는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세포로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

- [0215] 일부 구현예에서, 항원 결합 시약은 검출가능한 모이어티와 회합된다. 일부 구현예에서, 입자는 검출가능한 모이어티와 회합된다. 샘플 인덱싱 올리고뉴클레오타이드는 광학 모이어티와 회합된다.
- [0216] 일부 구현예에서, 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.
- [0217] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0218] 일부 구현예에서, 방법은, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 방법을 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열 분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함한다.
- [0219] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 체계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.
- [0220] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다.
- [0221] 일부 구현예에서, 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함한다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 복수의 체계적 바코드를 사용하여 세포의 복수의 표적을 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.
- [0222] 본 명세서의 개시 내용은 세포 성분 표적 사이의 상호작용, 예를 들어 단백질들 사이의 상호작용을 결정하기 위한 시스템, 방법, 및 키트를 포함한다. 일부 구현예에서, 방법은, 세포를 제1 쌍의 상호작용 결정 조성물과 접촉시키는 단계로서, 세포는 제1 단백질 표적 및 제2 단백질 표적을 포함하고, 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 제1 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약은 제1 단백질 표적에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약은 제2 단백질 표적에 특이적으로 결합할 수 있고, 상호작용 결정 올리고뉴클레오타이드는 상호작용 결정 서열 및 가교 올리고뉴클레오타이드 혼성 영역을 포함하고, 제1 쌍의 상호작용 결정 조성물들 중 상호작용 결정 서열은 상이한 서열을 포함하는, 단계; 가교 올리고뉴클레오타이드를 사용하여 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 올리고뉴클레오타이드를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계로서, 가교 올리고뉴클레오타이드는 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 포함하는, 단계; 복수의 바코드를 사

용하여 결합된 상호작용 결정 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 바코드 서열 및 포획 서열을 포함하는, 단계; 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열의 회합에 기초하여 제1 단백질 표적과 제2 단백질 표적 사이의 상호작용을 결정하는 단계를 포함한다.

[0223] 본 명세서의 개시 내용은 세포 성분 표적 사이의 상호작용을 결정하기 위한 시스템, 방법, 및 키트를 포함한다. 일부 구현예에서, 방법은, 세포를 제1 쌍의 상호작용 결정 조성물과 접촉시키는 단계로서, 세포는 제1 세포 성분 표적 및 제2 세포 성분 표적을 포함하고, 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 제1 쌍의 상호작용 결정 조성물 중 하나의 세포 성분 결합 시약은 제1 세포 성분 표적에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 세포 성분 결합 시약은 제2 세포 성분 표적에 특이적으로 결합할 수 있고, 상호작용 결정 올리고뉴클레오타이드는 상호작용 결정 서열 및 가교 올리고뉴클레오타이드 혼성화 영역을 포함하고, 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함하는, 단계; 가교 올리고뉴클레오타이드를 사용하여 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 올리고뉴클레오타이드를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계로서, 가교 올리고뉴클레오타이드는 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 포함하는, 단계; 복수의 바코드를 사용하여 결합된 상호작용 결정 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 바코드 서열 및 포획 서열을 포함하는, 단계; 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열의 회합에 기초하여 제1 세포 성분 표적 및 제2 세포 성분 표적 사이의 상호작용을 결정하는 단계를 포함한다. 2개의 세포 성분 결합 시약 중 적어도 하나는 단백질 결합 시약을 포함할 수 있다. 단백질 결합 시약은 2개의 상호작용 결정 올리고뉴클레오타이드 중 하나와 회합될 수 있다. 하나 이상의 세포 성분 표적은 적어도 하나의 단백질 표적을 포함할 수 있다.

[0224] 일부 구현예에서, 세포를 제1 쌍의 상호작용 결정 조성물과 접촉시키는 단계는, 세포를 제1 쌍의 상호작용 결정 조성물 각각과 차례로 또는 동시에 접촉시키는 단계를 포함한다. 제1 단백질 표적은 제2 단백질 표적과 동일할 수 있거나, 제1 단백질 표적은 제2 단백질 표적과 상이할 수 있다.

[0225] 상호작용 결정 서열의 길이는 변동될 수 있다. 예를 들어, 상호작용 결정 서열은 2개의 뉴클레오타이드 내지 약 1000개의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 상호작용 결정 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 60개의 뉴클레오타이드 길이, 약 45개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 임의의 조합일 수 있다.

[0226] 일부 구현예에서, 방법은, 세포를 제2 쌍의 상호작용 결정 조성물과 접촉시키는 단계를 포함하며, 세포는 제3 단백질 표적 및 제4 단백질 표적을 포함하고, 제2 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 제2 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약은 제3 단백질 표적에 특이적으로 결합할 수 있고, 제2 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약은 제4 단백질 표적에 특이적으로 결합할 수 있다. 제3 단백질 표적 및 제4 단백질 표적 중 적어도 하나는 제1 단백질 표적 및 제2 단백질 표적 중 하나와 상이할 수 있다. 일부 구현예에서, 제3 단백질 표적 및 제4 단백질 표적 중 적어도 하나와 제1 단백질 표적 및 제2 단백질 표적 중 적어도 하나는 동일할 수 있다.

[0227] 일부 구현예에서, 방법은, 세포를 3개 이상의 쌍의 상호작용 결정 조성물과 접촉시키는 단계를 포함한다. 복수의 쌍의 상호작용 결정 조성물 중 적어도 10개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다. 복수의 쌍의 상호작용 결정 조성물 중 적어도 100개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다. 복수의 쌍의 상호작용 결정 조성물 중 적어도 1000개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다.

[0228] 일부 구현예에서, 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역은 상이한 서열을 포함한다. 가교 올리고뉴클레오타이드 혼성화 영역 중 적어도 하나는 가교 올리고뉴클레오타이드의 2개의 혼성화 영역 중 적어도 하나에 상보적일 수 있다.

[0229] 일부 구현예에서, 가교 올리고뉴클레오타이드를 사용하여 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 올리고

뉴클레오타이드를 결합하는 단계는: 가교 올리고뉴클레오타이드의 제1 혼성화 영역을 상호작용 결정 올리고뉴클레오타이드의 가교 올리고뉴클레오타이드 혼성화 영역 중 제1 가교 올리고뉴클레오타이드 혼성화 영역과 혼성화하는 단계; 가교 올리고뉴클레오타이드의 제2 혼성화 영역을 상호작용 결정 올리고뉴클레오타이드의 가교 올리고뉴클레오타이드 혼성화 영역 중 제2 가교 올리고뉴클레오타이드 혼성화 영역과 혼성화하는 단계; 및 가교 올리고뉴클레오타이드에 혼성화된 상호작용 결정 올리고뉴클레오타이드를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0230] 일부 구현예에서, 단백질 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 인테그린, 또는 이들의 조합을 포함한다. 상호작용 결정 올리고뉴클레오타이드는 링커를 통해 단백질 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 단백질 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 이황화 연결 또는 이들의 임의의 조합을 포함할 수 있다.

[0231] 세포 내의 단백질 표적의 위치는 변동될 수 있으며, 예를 들어 세포 표면 상에 또는 세포 내부에 존재할 수 있다. 일부 구현예에서, 하나 이상의 단백질 표적 중 적어도 하나는 세포 표면 상에 존재한다. 일부 구현예에서, 하나 이상의 단백질 표적 중 적어도 하나는 세포내 단백질이다. 일부 구현예에서, 하나 이상의 단백질 표적 중 적어도 하나는 막관통 단백질이다. 일부 구현예에서, 하나 이상의 단백질 표적 중 적어도 하나는 세포외 단백질이다. 일부 구현예에서, 방법은, 세포를 제1 쌍의 상호작용 결정 조성물들과 접촉시키기 전에, 세포를 고정시키는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 제1 쌍의 상호작용 결정 조성물 중 결합되지 않은 상호작용 결정 조성물을 제거하는 단계를 포함할 수 있다. 결합되지 않은 상호작용 결정 조성물을 제거하는 단계는 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 상호작용 결정 조성물을 제거하는 단계는 유세포분석을 사용하여 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 세포를 용해시키는 단계를 포함할 수 있다.

[0232] 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드는 단백질 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은, 단백질 결합 시약으로부터 상호작용 결정 올리고뉴클레오타이드를 탈착시키는 단계를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리, 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 단백질 결합 시약으로부터 상호작용 결정 올리고뉴클레오타이드를 탈착시키는 단계를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드는 세포의 유전체 서열과 상동성이 아닐 수 있다. 상호작용 결정 올리고뉴클레오타이드는 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다. 일부 구현예에서, 제1 쌍의 상호작용 결정 조성물 중 하나의 상호작용 결정 올리고뉴클레오타이드는 포획 서열에 상보적인 서열을 포함한다. 포획 서열은 폴리(dT) 영역을 포함할 수 있다. 포획 서열에 상보적인 상호작용 결정 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다. 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드는 제2 바코드 서열을 포함한다. 제1 쌍의 상호작용 결정 조성물들 중 다른 하나의 상호작용 결정 올리고뉴클레오타이드는 범용 프라이머를 위한 결합 부위를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드는 검출가능한 모이어티와 회합될 수 있다.

[0233] 일부 구현예에서, 단백질 결합 시약은 상이한 상호작용 결정 서열을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드와 회합될 수 있다. 일부 구현예에서, 복수의 상호작용 결정 조성물 중 하나는 상호작용 결정 올리고뉴클레오타이드와 회합되지 않은 제2 단백질 결합 시약을 포함한다. 단백질 결합 시약과 제2 단백질 결합 시약은 동일할 수 있다. 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다.

[0234] 일부 구현예에서, 세포는 종양 세포 또는 비-종양 세포이다. 예를 들어, 세포는 포유동물 세포, 세균 세포, 바이러스 세포, 효모 세포, 진균 세포, 또는 이들의 임의의 조합일 수 있다. 일부 구현예에서, 방법은, 2개 이상의 세포를 제1 쌍의 상호작용 결정 조성물과 접촉시키는 단계를 포함하며, 2개 이상의 세포 각각은 제1 및 제2 단백질 표적을 포함한다. 2개 이상의 세포 중 적어도 하나는 단일 세포를 포함한다.

[0235] 일부 구현예에서, 단백질 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 단백질 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 단백질 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 단백질 표적은 10 내지 100개의 상이한 단백질 표적을 포함하는 군으로부터 선택될 수 있다. 단백질 결합 시약은 동일한 서열을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드와 회합될 수 있다.

- [0236] 일부 구현예에서, 바코드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함한다. 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함할 수 있다. 일부 구현예에서, 복수의 바코드는 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다.. 입자는 비드를 포함할 수 있다. 입자는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로 비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세포로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다. 입자는 검출가능한 모이어티와 회합될 수 있다.
- [0237] 일부 구현예에서, 입자의 바코드는 적어도 1000개의 상이한 바코드 서열로부터 선택되는 바코드 서열을 포함한다. 입자의 바코드는 적어도 10000개의 상이한 바코드 서열로부터 선택되는 바코드 서열을 포함할 수 있다. 바코드의 바코드 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.
- [0238] 일부 구현예에서, 복수의 바코드를 사용하여 상호작용 결정 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 상호작용 결정 올리고뉴클레오타이드와 접촉시켜 상호작용 결정 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 상호작용 결정 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 몰로니 뮤린 백혈병 바이러스(Moloney Murine Leukemia Virus(M-MLV)) 역전사효소 또는 Taq DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 결합된 상호작용 결정 올리고뉴클레오타이드로부터 가교 올리고뉴클레오타이드를 변위시키는 것(displacing)을 포함할 수 있다.
- [0239] 일부 구현예에서, 방법은, 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 바코드 서열의 적어도 일부분 및 상호작용 결정 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 바코드 서열의 적어도 일부분 및 상호작용 결정 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.
- [0240] 일부 구현예에서, 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드의 부분적인 그리고/또는 완전한 서열을 획득하는 단계를 포함한다.
- [0241] 일부 구현예에서, 복수의 바코드는 복수의 추계적 바코드를 포함한다. 복수의 추계적 바코드의 각각의 바코드 서열은 분자 표지 서열을 포함할 수 있다. 복수의 추계적 바코드 중 적어도 2개의 추계적 바코드의 분자 표지 서열은 상이한 서열을 포함할 수 있다. 복수의 바코드를 사용하여 상호작용 결정 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 상호작용 결정 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0242] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을

증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 바코딩된 표적을 중합효소 연쇄 반응(PCR)에 의해 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 주계적 바코드를 사용하여 세포의 복수의 표적을 주계적으로 바코딩하여 복수의 주계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0243] 본 명세서의 개시 내용은 세포 성분 사이의 상호작용, 예를 들어 단백질-단백질 상호작용을 식별하기 위한 키트를 포함한다. 일부 구현예에서, 키트는, 제1 쌍의 상호작용 결정 조성물로서, 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 제1 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약은 제1 단백질 표적에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약은 제2 단백질 표적에 특이적으로 결합할 수 있고, 상호작용 결정 올리고뉴클레오타이드는 상호작용 결정 서열 및 가교 올리고뉴클레오타이드 혼성화 영역을 포함하고, 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함하는, 제1 쌍의 상호작용 결정 조성물; 및 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 각각 포함하는 복수의 가교 올리고뉴클레오타이드를 포함한다.

[0244] 상호작용 결정 서열의 길이는 변동될 수 있다. 일부 구현예에서, 상호작용 결정 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 60개의 뉴클레오타이드 길이, 약 45개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 적어도 128개의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 임의의 조합일 수 있다.

[0245] 일부 구현예에서, 키트는, 제2 쌍의 상호작용 결정 조성물로서, 제2 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 제2 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약은 제3 단백질 표적에 특이적으로 결합할 수 있고, 제2 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약은 제4 단백질 표적에 특이적으로 결합할 수 있는, 제2 쌍의 상호작용 결정 조성물을 추가로 포함한다. 제3 단백질 표적 및 제4 단백질 표적 중 적어도 하나는 제1 단백질 표적 및 제2 단백질 표적 중 하나와 상이할 수 있다. 제3 단백질 표적 및 제4 단백질 표적 중 적어도 하나와 제1 단백질 표적 및 제2 단백질 표적 중 적어도 하나는 동일할 수 있다.

[0246] 일부 구현예에서, 키트는, 3개 이상의 쌍의 상호작용 결정 조성물을 포함할 수 있다. 3개 이상의 쌍의 상호작용 결정 조성물 중 적어도 10개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다. 3개 이상의 쌍의 상호작용 결정 조성물의 적어도 100개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다. 3개 이상의 쌍의 상호작용 결정 조성물의 적어도 1000개의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함할 수 있다.

[0247] 일부 구현예에서, 복수의 상호작용 결정 조성물 중 2개의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역은 상이한 서열을 포함한다. 가교 올리고뉴클레오타이드 혼성화 영역 중 적어도 하나는 가교 올리고뉴클레오타이드의 2개의 혼성화 영역 중 적어도 하나에 상보적일 수 있다.

[0248] 일부 구현예에서, 단백질 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 상호작용 결정 올리고뉴클레오타이드는 링커를 통해 단백질 결합 시약에 접합될 수 있다. 적어도 하나의 상호작용 결정 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 단백질 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트랩타민, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드는 관심 대상의 어떠한 세포의 유전체 서열과도 상동성이 아닐 수 있다. 관심 대상의 세포는 종양 세포 또는 비-종양 세포일 수 있다. 관심 대상의 세포는 단일 세포, 포유동물 세포, 세균 세포, 바이러스 세포, 효모 세포, 진균 세포, 또는 이들의 임의의 조합일 수 있다.

[0249] 일부 구현예에서, 키트는, 복수의 바코드의 각각이 바코드 서열 및 포획 서열을 포함하는 복수의 바코드를 추가로 포함한다. 제1 쌍의 상호작용 결정 조성물 중 하나의 상호작용 결정 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 포획 서열은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 상호작용 결정 올리고뉴클레오타이드의 서열은 폴리(dA) 영역을 포함할 수 있다. 제1 쌍의 상호작용 식별 조성물 중 다른 하나의 상호작용 결정 올리고뉴클레오타이드는 세포 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 바코드는 복수의 주계적 바

코드를 포함할 수 있으며, 복수의 추계적 바코드의 각각의 바코드 서열은 분자 표지 서열을 포함하고, 복수의 추계적 바코드 중 적어도 2개의 추계적 바코드의 분자 표지 서열은 상이한 서열을 포함한다.

[0250] 일부 구현예에서, 단백질 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 단백질 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 중앙 항원, 수용체, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 단백질 표적은 10 내지 100 개의 상이한 단백질 표적을 포함하는 군으로부터 선택될 수 있다. 단백질 결합 시약은 동일한 서열을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드와 회합될 수 있다. 단백질 결합 시약은 상이한 상호작용 결정 서열을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드와 회합될 수 있다. 일부 구현예에서, 단백질 결합 시약은 검출가능한 모이어티와 회합될 수 있다.

[0251] 일부 구현예에서, 제1 쌍의 상호작용 결정 조성물 중 하나는 상호작용 결정 올리고뉴클레오타이드와 회합되지 않은 제2 단백질 결합 시약을 포함한다. 제1 단백질 결합 시약과 제2 단백질 결합 시약은 동일할 수 있다. 상호작용 결정 올리고뉴클레오타이드는 검출가능한 모이어티와 회합될 수 있다.

[0252] 일부 구현예에서, 복수의 바코드는 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다. 입자는 비드를 포함할 수 있다. 입자는 세포로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합을 포함할 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세포로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다. 입자는 검출가능한 모이어티와 회합될 수 있다.

[0253] 일부 구현예에서, 입자의 바코드는 적어도 1000개의 상이한 바코드 서열로부터 선택되는 바코드 서열을 포함한다. 입자의 바코드는 적어도 10000개의 상이한 바코드 서열로부터 선택되는 바코드 서열을 포함할 수 있다. 바코드의 바코드 서열은 랜덤 서열을 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다.

[0254] 일부 구현예에서, 키트는 DNA 중합효소, 역전사효소, 몰로니 뮤린 백혈병 바이러스(M-MLV) 역전사효소, Taq DNA 중합효소, 또는 이들의 임의의 조합을 포함한다. 키트는 고정제(fixation agent)를 포함할 수 있다.

[0255] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포는 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약(예를 들어, 항체)을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적(예를 들어, 단백질) 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 데이터-체이닝 증폭 프라이머(daisy-chaining amplification primer)를 사용하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 데이터-체이닝 신장된 앰플리콘을 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 포함하는 복수의 데이터-체이닝 신장된 앰플리콘의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 방법은 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0256] 일부 구현예에서, 샘플 인덱싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인덱싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 샘플 인덱싱 조성물

중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 100개 또는 1000개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0257] 일부 구현예에서, 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0258] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 세포 성분 표적 중 적어도 하나는 세포 표면 상에 존재할 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0259] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이트를 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0260] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.

[0261] 일부 구현예에서, 세포 성분 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 세포 성분 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다. 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 세포 성분 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0262] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함한다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0263] 일부 구현예에서, 복수의 바코드는 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 입자는 비드일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는

분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.

[0264] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0265] 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 데이지-체이닝 신장된 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0266] 일부 구현예에서, 복수의 데이지-체이닝 증폭 프라이머의 데이지-체이닝 증폭 프라이머 각각은 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역 및 오버행(overhang) 영역을 포함하며, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 샘플 인덱싱 올리고뉴클레오타이드의 데이지-체이닝 샘플 인덱싱 영역에 결합할 수 있다. 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 적어도 20개의 뉴클레오타이드 길이, 적어도 30개의 뉴클레오타이드 길이, 약 40개의 뉴클레오타이드 길이, 적어도 40개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 오버행 영역은 적어도 50개의 뉴클레오타이드 길이, 적어도 100개의 뉴클레오타이드 길이, 적어도 150개의 뉴클레오타이드 길이, 약 150개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 오버행 영역은 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 2개의 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머의 오버행 영역은 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 신장된 앰플리콘 중의 데이지-체이닝 신장된 앰플리콘은 적어도 250개의 뉴클레오타이드 길이, 적어도 300개의 뉴클레오타이드 길이, 적어도 350개의 뉴클레오타이드 길이, 적어도 400개의 뉴클레오타이드 길이, 약 400개의 뉴클레오타이드 길이, 적어도 450개의 뉴클레오타이드 길이, 적어도 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다.

[0267] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0268] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 방법은 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의

바코딩된 표적을 생성하는 단계는 복수의 주계적 바코드를 사용하여 세포의 복수의 표적을 주계적으로 바코딩하여 복수의 주계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0269] 일부 구현예에서, 복수의 샘플 인덱싱 조성물 각각은 세포 성분 결합 시약을 포함한다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 동일할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다. 일부 구현예에서, 세포 성분은 항원이다. 세포 성분 결합 시약은 항체일 수 있다.

[0270] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 결합 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 결합 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 데이지-체이닝 증폭 프라이머를 사용하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 방법은 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0271] 일부 구현예에서, 샘플 인덱싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인덱싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0272] 일부 구현예에서, 세포 성분 결합 시약은 세포 표면 결합 시약, 항체, 테트라머, 압타머, 단백질 스캐폴드, 인테그린, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0273] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 세포 성분 표적 중 적어도 하나는 세포 표면 상에 발현될 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0274] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0275] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일을 포함할 수 있다.

다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.

[0276] 일부 구현예에서, 세포 성분 표적은 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다. 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 세포 성분 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0277] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함할 수 있다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0278] 일부 구현예에서, 복수의 바코드는 입자 내에 봉입된다. 입자는 비드일 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.

[0279] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0280] 일부 구현예에서, 방법은 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 복수의 데이지-체이닝 증폭 프라이머를 사용하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 데이지-체이닝 신장된 앰플리콘을 생성하는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 데이지-체이닝 신장된 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0281] 일부 구현예에서, 복수의 데이지-체이닝 증폭 프라이머 중의 데이지-체이닝 증폭 프라이머는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역 및 오버행 영역을 포함하며, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 샘플 인덱싱 올리고뉴클레오타이드의 데이지-체이닝 샘플 인덱싱 영역에 결합할 수 있다. 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 적어도 20개의 뉴클레오타이드 길이, 적어도 30개의 뉴클레오타이드 길이, 약 40개의 뉴클레오타이드 길이, 적어도 40개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 오버행

영역은 적어도 50개의 뉴클레오타이드 길이, 적어도 100개의 뉴클레오타이드 길이, 적어도 150개의 뉴클레오타이드 길이, 약 150개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 오버행 영역은 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 2개의 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머의 오버행 영역은 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 신장된 앰플리콘 중의 데이지-체이닝 신장된 앰플리콘은 적어도 250개의 뉴클레오타이드 길이, 적어도 300개의 뉴클레오타이드 길이, 적어도 350개의 뉴클레오타이드 길이, 적어도 400개의 뉴클레오타이드 길이, 약 400개의 뉴클레오타이드 길이, 적어도 450개의 뉴클레오타이드 길이, 적어도 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다.

[0282] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 체계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0283] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 체계적 바코드를 사용하여 세포의 복수의 표적을 체계적으로 바코딩하여 복수의 체계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0284] 일부 구현예에서, 복수의 샘플 인덱싱 조성물 각각은 세포 성분 결합 시약을 포함한다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 동일할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다.

[0285] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 및 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열 및 복수의 데이지-체이닝 증폭 프라이머에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 적어도 하나의 세포의 샘플 기원을 식별하는 단계는, 복수의 바코드를 사용하여 복수의 샘플 인덱싱 조성물의 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 서열분석 데이터에서 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함할 수 있다. 방법은, 예를 들어 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0286] 일부 구현예에서, 샘플 인덱싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인덱싱 올리고뉴클레오타이드는 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이, 또는 이들의 조합이다.

고뉴클레오티드는 약 50개의 뉴클레오티드 길이, 약 100개의 뉴클레오티드 길이, 약 200개의 뉴클레오티드 길이, 적어도 200개의 뉴클레오티드 길이, 약 200 내지 300개 미만의 뉴클레오티드 길이, 약 200 내지 500개의 뉴클레오티드 길이, 약 500개의 뉴클레오티드 길이, 또는 이들의 조합일 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 10개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 100개 또는 1000개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0287] 일부 구현예에서, 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인덱싱 올리고뉴클레오티드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 올리고뉴클레오티드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0288] 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포를 포함한다. 하나 이상의 세포 성분 표적 중 적어도 하나는 세포 표면 상에 발현될 수 있다. 복수의 샘플 중의 샘플은 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 복수의 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0289] 일부 구현예에서, 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 복수의 샘플 각각으로부터의 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함한다. 방법은 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오티드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다. 방법은 샘플 인덱싱 올리고뉴클레오티드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오티드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오트레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오티드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.

[0290] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오티드는 복수의 샘플의 세포의 유전체 서열과 상동성이 아니다. 샘플 인덱싱 올리고뉴클레오티드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오티드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함할 수 있다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다. 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오티드의 서열은 폴리(A) 테일을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오티드는 분자 표지를 포함할 수 있다.

[0291] 일부 구현예에서, 세포 성분 표적은 탄수화물, 지질, 단백질, 세포의 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다. 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오티드와 회합될 수 있다. 세포 성분 결합 시약은 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오티드와 회합될 수 있다. 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물은 샘플 인덱싱 올리고뉴클레오티드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

[0292] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 복수의 바코드를 사용하여 복수의 샘플 인덱싱 조성물의 샘플 인덱싱 올리고뉴클레오티드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오티드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오티드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오티드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오티드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함한다.

[0293] 일부 구현예에서, 복수의 바코드 중의 바코드는 표적-결합 영역 및 분자 표지 서열을 포함한다. 복수의 바코드 중 적어도 2개의 바코드의 분자 표지 서열은 상이한 분자 표지 서열을 포함할 수 있다. 바코드는 세포 표지, 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함할 수 있다. 표적-결합 영역은 폴리(dT) 영역을 포함할 수 있다.

[0294] 일부 구현예에서, 복수의 바코드는 입자 상에 고정화된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나,

이들의 임의의 조합일 수 있다. 입자는 분해성일 수 있다. 입자는 비드일 수 있다. 비드는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다. 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되는 재료를 포함할 수 있다. 입자는 적어도 10000개의 바코드를 포함할 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000개 또는 10000개의 상이한 분자 표지 서열로부터 선택되는 분자 표지 서열을 포함할 수 있다. 바코드의 분자 표지 서열은 랜덤 서열을 포함할 수 있다.

[0295] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0296] 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 복수의 데이지-체이닝 증폭 프라이머를 사용하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 데이지-체이닝 신장된 앰플리콘을 생성하는 단계를 포함할 수 있다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 데이지-체이닝 신장된 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0297] 일부 구현예에서, 복수의 데이지-체이닝 증폭 프라이머 중의 데이지-체이닝 증폭 프라이머는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역 및 오버행 영역을 포함하며, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 샘플 인덱싱 올리고뉴클레오타이드의 데이지-체이닝 샘플 인덱싱 영역에 결합할 수 있다. 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역은 적어도 20개의 뉴클레오타이드 길이, 적어도 30개의 뉴클레오타이드 길이, 약 40개의 뉴클레오타이드 길이, 적어도 40개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 오버행 영역은 적어도 50개의 뉴클레오타이드 길이, 적어도 100개의 뉴클레오타이드 길이, 적어도 150개의 뉴클레오타이드 길이, 약 150개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 오버행 영역은 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 2개의 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머의 오버행 영역은 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 신장된 앰플리콘 중의 데이지-체이닝 신장된 앰플리콘은 적어도 250개의 뉴클레오타이드 길이, 적어도 300개의 뉴클레오타이드 길이, 적어도 350개의 뉴클레오타이드 길이, 적어도 400개의 뉴클레오타이드 길이, 적어도 450개의 뉴클레오타이드 길이, 적어도 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 데이지-체이닝 신장된 앰플리콘 중의 데이지-체이닝 신장된 앰플리콘은 약 200개의 뉴클레오타이드 길이, 약 250개의 뉴클레오타이드 길이, 약 300개의 뉴클레오타이드 길이, 약 350개의 뉴클레오타이드 길이, 약 400개의 뉴클레오타이드 길이, 약 450개의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 약 600개의 뉴클레오타이드 길이, 또는 이들 중 임의의 두 값 사이의 범위일 수 있다.

[0298] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플

플 인텍싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인텍싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 복수의 샘플 인텍싱 조성물의 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드의 샘플 인텍싱 서열의 존재 또는 부재를 식별하는 단계를 포함할 수 있다. 샘플 인텍싱 서열의 존재 또는 부재를 식별하는 단계는, 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 서열분석 데이터에서 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드에 상응하는 복수의 샘플 인텍싱 올리고뉴클레오타이드 중 복제된 샘플 인텍싱 올리고뉴클레오타이드의 샘플 인텍싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함할 수 있다.

[0299] 일부 구현예에서, 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계는, 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 복제하기 전에, 복제 어댑터를 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드에 연결하는 단계를 포함한다. 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 복제하는 단계는 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드에 연결된 복제 어댑터를 사용하여 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0300] 일부 구현예에서, 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계는, 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드를 복제하기 전에, 포획 프로브를 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드와 접촉시켜 샘플 인텍싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 생성하는 단계; 및 샘플 인텍싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 연장하여 포획 프로브와 회합된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드를 복제하는 단계는 포획 프로브와 회합된 샘플 인텍싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인텍싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0301] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지를 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인텍싱 올리고뉴클레오타이드의 샘플 인텍싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.

[0302] 일부 구현예에서, 복수의 샘플 인텍싱 조성물 각각은 세포 성분 결합 시약을 포함한다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인텍싱 올리고뉴클레오타이드의 샘플 인텍싱 서열은 동일할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 샘플 인텍싱 올리고뉴클레오타이드의 샘플 인텍싱 서열은 상이한 서열을 포함할 수 있다. 복수의 샘플 인텍싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다.

[0303] 본 명세서의 개시 내용은 키트를 포함하며, 상기 키트는 복수의 샘플 인텍싱 조성물; 및 복수의 데이지-체이닝 증폭 프라이머를 포함한다. 복수의 샘플 인텍싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다. 2개 이상의 세포 성분 결합 시약 각각은 샘플 인텍싱 올리고뉴클레오타이드와 회합될 수 있다. 2개 이상의 세포 성분 결합 시약 중 적어도 하나는 적어도 하나의 세포 성분 표적에 특이적으로 결합할 수 있다. 샘플 인텍싱 올리고뉴클레오타이드는 샘플의 하나 이상의 세포의 샘플 기원을 식별하기 위한 샘플 인텍싱 서열을 포함할 수 있다. 샘플 인텍싱 올리고뉴클레오타이드는 데이지-체이닝 증폭 프라이머 결합 서열을 포함할 수 있다. 복수의 샘플 인텍싱 조성물 중 적어도 2개의 샘플 인텍싱 조성물의 샘플 인텍싱 서열은 상이한 서열을 포함할 수 있다.

[0304] 복수의 데이지-체이닝 증폭 프라이머 중의 데이지-체이닝 증폭 프라이머는 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역 및 오버행 영역을 포함할 수 있으며, 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역은

샘플 인텍싱 올리고뉴클레오타이드의 데이지-체이닝 샘플 인텍싱 영역에 결합할 수 있다. 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역은 적어도 20개의 뉴클레오타이드 길이, 적어도 30개의 뉴클레오타이드 길이, 약 40개의 뉴클레오타이드 길이, 적어도 40개의 뉴클레오타이드 길이, 약 50개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 서열을 갖는 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 상이한 서열을 갖는 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역을 포함할 수 있다. 바코딩된 샘플 인텍싱 올리고뉴클레오타이드-결합 영역은 데이지-체이닝 증폭 프라이머 결합 서열, 이의 상보체, 이의 역상보체, 또는 이들의 조합을 포함할 수 있다. 복수의 샘플 인텍싱 조성물 중 적어도 2개의 샘플 인텍싱 조성물의 데이지-체이닝 증폭 프라이머 결합 서열은 동일한 서열을 포함할 수 있다. 복수의 샘플 인텍싱 조성물 중 적어도 2개의 샘플 인텍싱 조성물의 데이지-체이닝 증폭 프라이머 결합 서열은 상이한 서열을 포함한다. 오버행 영역은 적어도 50개의 뉴클레오타이드 길이, 적어도 100개의 뉴클레오타이드 길이, 적어도 150개의 뉴클레오타이드 길이, 약 150개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 오버행 영역은 적어도 200개의 뉴클레오타이드 길이일 수 있다. 오버행 영역은 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 동일한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머 중 2개의 데이지-체이닝 증폭 프라이머는 2개의 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역을 포함할 수 있다. 복수의 데이지-체이닝 증폭 프라이머의 오버행 영역은 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다.

[0305] 일부 구현예에서, 샘플 인텍싱 서열은 적어도 6개의 뉴클레오타이드 길이, 25 내지 45개의 뉴클레오타이드 길이, 약 128개의 뉴클레오타이드 길이, 또는 적어도 128개의 뉴클레오타이드 길이, 또는 이들의 조합이다. 샘플 인텍싱 올리고뉴클레오타이드는 약 50개의 뉴클레오타이드 길이, 약 100개의 뉴클레오타이드 길이, 약 200개의 뉴클레오타이드 길이, 적어도 200개의 뉴클레오타이드 길이, 약 200 내지 300개 미만의 뉴클레오타이드 길이, 약 200 내지 500개의 뉴클레오타이드 길이, 약 500개의 뉴클레오타이드 길이, 또는 이들의 조합일 수 있다. 복수의 샘플 인텍싱 조성물 중 적어도 10, 100, 또는 1000개의 샘플 인텍싱 조성물의 샘플 인텍싱 서열은 상이한 서열을 포함한다.

[0306] 일부 구현예에서, 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 샘플 인텍싱 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 적어도 하나의 샘플 인텍싱 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 결합 시약의 분자에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0307] 일부 구현예에서, 샘플 인텍싱 올리고뉴클레오타이드는 종의 유전체 서열과 상동성이 아니다. 샘플 인텍싱 올리고뉴클레오타이드는 분자 표지 서열, 폴리(A) 영역, 또는 이들의 조합을 포함할 수 있다. 일부 구현예에서, 복수의 샘플 중 적어도 하나의 샘플은 단일 세포, 복수의 세포, 조직, 종양 샘플, 또는 이들의 임의의 조합을 포함할 수 있다. 샘플은 포유동물 샘플, 세균 샘플, 바이러스 샘플, 효모 샘플, 진균 샘플, 또는 이들의 임의의 조합을 포함할 수 있다.

[0308] 일부 구현예에서, 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다. 세포 성분 결합 시약은 동일한 서열을 갖는 2개 이상의 샘플 인텍싱 올리고뉴클레오타이드와 회합될 수 있다. 세포 성분 결합 시약은 상이한 샘플 인텍싱 서열을 갖는 2개 이상의 샘플 인텍싱 올리고뉴클레오타이드와 회합될 수 있다. 복수의 샘플 인텍싱 조성물 중의 샘플 인텍싱 조성물은 샘플 인텍싱 올리고뉴클레오타이드와 접합되지 않은 제2 세포 성분 결합 시약을 포함할 수 있다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다.

도면의 간단한 설명

[0309] 도 1은 비제한적이고 예시적인 추계적 바코드를 예시한다.

도 2는 추계적 바코딩 및 디지털 계수의 비제한적이고 예시적인 작업 흐름을 나타낸다.

도 3은 복수의 표적으로부터 추계적으로 바코딩된 표적의 인텍싱된 라이브러리를 생성하기 위한 비제한적이고 예시적인 과정을 나타내는 개략도이다.

도 4는 단백질 결합 시약에 대한 고유 식별자를 포함하는 올리고뉴클레오타이드와 접합된 예시적인 단백질 결합

시약(여기서는 항체가 예시됨)의 개략도를 나타낸다.

도 5는 동일한 또는 상이한 샘플로부터의 세포를 결정하기 위한 샘플 인덱싱에 대한 고유 식별자를 포함하는 올리고뉴클레오타이드와 접합된 예시적인 결합 시약(여기서는 항체가 예시됨)의 개략도를 나타낸다.

도 6은 올리고뉴클레오타이드-접합된 항체를 사용하여 고처리량(high throughput) 방식으로 단백질 발현과 유전자 발현을 동시에 결정하는 예시적인 작업 흐름의 개략도를 나타낸다.

도 7은 샘플 인덱싱을 위한 올리고뉴클레오타이드-접합된 항체를 사용하는 예시적인 작업 흐름의 개략도를 나타낸다.

도 8aa 및 도 8ab는 바코딩된 올리고뉴클레오타이드를 데이지-체이닝하는 예시적인 작업 흐름의 개략도를 나타낸다. 도 8ba 및 도 8bb는 바코딩된 올리고뉴클레오타이드를 데이지-체이닝하는 또 다른 예시적인 작업 흐름의 개략도를 나타낸다.

도 9a 내지 도 9e는 올리고뉴클레오타이드로 작용화된 입자들의 비제한적이고 예시적인 개략도를 나타낸다.

도 10a는 단일 세포 서열분석 제어를 위한 올리고뉴클레오타이드로 작용화된 입자 사용의 예시적인 작업 흐름의 개략도이다. 도 10b는 단일 세포 서열분석 제어를 위한 올리고뉴클레오타이드로 작용화된 입자 사용의 또 다른 예시적인 또 다른 작업 흐름의 개략도이다.

도 11은 단일 세포 서열분석 효율을 결정하기 위한 제어 올리고뉴클레오타이드-접합된 항체 사용의 예시적인 작업 흐름의 개략도를 나타낸다.

도 12는 단일 세포 서열분석 효율을 결정하기 위한 제어 올리고뉴클레오타이드-접합된 항체 사용의 예시적인 작업 흐름의 또 다른 개략도를 나타낸다.

도 13a 내지 도 13c는 제어 올리고뉴클레오타이드가 세포 계수에 사용될 수 있음을 나타내는 플롯이다.

도 14a 내지 도 14f는 한 쌍의 상호작용 결정 조성물을 사용하여 세포 성분(예를 들어, 단백질)들 사이의 상호작용을 결정하는 예시적인 작업 흐름의 개략도를 나타낸다.

도 15a 내지 도 15d는 단백질 발현과 유전자 발현을 동시에 결정하기 위한 그리고 샘플 인덱싱을 위한 올리고뉴클레오타이드의 비제한적이고 예시적인 설계를 나타낸다.

도 16은 단백질 발현 및 유전자 발현을 동시에 결정하기 위한 비제한적이고 예시적인 올리고뉴클레오타이드 서열의 개략도를 나타낸다.

도 17a 내지 도 17f는 CD4 단백질 발현 및 유전자 발현을 동시에 고처리량 방식으로 측정하기 위한 올리고뉴클레오타이드-접합된 항체 사용의 결과를 나타내는 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 18a 내지 도 18f는 CD4 T 세포, CD8 T 세포, 및 골수 세포 내 CD4 mRNA 및 단백질의 발현을 나타내는 비제한적이고 예시적인 막대 차트이다.

도 19는, 유사한 서열분석 깊이로, CD4 단백질 수준에 대한 검출 감도가 항체:올리고뉴클레오타이드의 더 높은 비에 따라 증가하며, 1:3의 비가 1:1 및 1:2의 비보다 더 양호하게 수행됨을 나타내는 비제한적이고 예시적인 막대 차트이다.

도 20a 내지 도 20d는 유세포분석을 사용하여 분류된 세포의 세포 표면 상에서의 CD4 단백질 발현을 나타내는 플롯이다.

도 21a 내지 도 21f는 2개의 샘플의 CD4 T 세포, CD8 T 세포, 및 골수 세포에서 CD4 mRNA 및 단백질의 발현을 나타내는 비제한적이고 예시적인 막대 차트이다.

도 22는 1:3의 항체:올리고뉴클레오타이드 비를 갖는 상이한 샘플 제조 프로토콜을 사용하여 결정된 CD4 단백질 수준에 대한 검출 감도를 나타내는 비제한적이고 예시적인 막대 차트이다.

도 23은 샘플 인덱싱을 수행하고, 샘플 인덱싱에 대한 샘플 인덱싱 올리고뉴클레오타이드의 길이 및 샘플 인덱싱 올리고뉴클레오타이드의 절단성의 효과를 결정하기 위한 비제한적이고 예시적인 실험 설계를 나타낸다.

도 24a 내지 도 24c는 상이한 샘플 인덱싱 올리고뉴클레오타이드(절단 가능한 95량체, 절단 불가능한 95량체, 및 절단 가능한 200량체)와 접합된 3가지 유형의 항-CD147 항체가 CD147의 단백질 발현 수준을 결정하는 데 사용될

수 있음을 나타내는 비제한적이고 예시적인 tSNE 플롯이다.

도 25는 세포당 GAPDH 발현의 오버레이를 갖는 비제한적이고 예시적인 tSNE 플롯이다.

도 26a 내지 도 26c는 3가지 유형의 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 검출된 샘플 인덱싱 올리고뉴클레오타이드의 분자수를 나타내는 비제한적이고 예시적인 히스토그램이다.

도 27a 내지 도 27c는 CD147 발현이 분열성 세포(dividing cell)에서 더 높았음을 나타내는 비제한적이고 예시적인 플롯 및 막대 차트이다.

도 28a 내지 도 28c는 샘플 인덱싱이 상이한 샘플의 세포들을 식별하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 29a 내지 도 29c는 결정된 샘플 인덱싱 올리고뉴클레오타이드의 분자수에 기초한 세포당 샘플 인덱싱 서열의 비제한적이고 예시적인 히스토그램이다.

도 30a 내지 도 30d는 CD3D(주르카트(Jurkat) 세포의 경우) 및 JCHAIN(라모스(Ramos) 세포의 경우)의 mRNA 발현 및 주르카트 및 라모스 세포의 샘플 인덱싱에 기초하여 결정된 세포 유형의 주석(annotation)을 비교하는 비제한적이고 예시적인 플롯이다.

도 31a 내지 도 31c는 CD3D 및 JCHAIN(도 31a), JCHAIN(도 31b), 및 CD3D(도 31c)의 mRNA 발현의 오버레이를 갖는 주르카트 및 라모스 세포의 mRNA 발현 프로파일의 비제한적인 tSNE 투영 플롯이다.

도 32는 250의 DBEC 컷오프로 샘플 인덱싱을 사용하여 결정된 세포 유형의 오버레이를 갖는 주르카트 및 라모스 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 33a 내지 도 33c는 "단길이(Short) 3" 샘플 인덱싱 올리고뉴클레오타이드, "단길이 2" 및 "단길이 3" 샘플 인덱싱 올리고뉴클레오타이드, 및 "단길이 2" 샘플 인덱싱 올리고뉴클레오타이드로 표지되지 않거나 표지된 라모스 및 주르카트 세포(도 33), 라모스 세포(도 33b), 및 주르카트 세포(도 33c)에 대한 세포당 샘플 인덱싱 올리고뉴클레오타이드의 분자수의 비제한적이고 예시적인 막대 차트이다.

도 34a 내지 도 34c는 단일 세포의 1% 미만인 "단길이 2" 및 "단길이 3" 샘플 인덱싱 올리고뉴클레오타이드 둘 모두로 표지된 것을 나타내는 비제한적이고 예시적인 플롯이다.

도 35a 내지 도 35c는 도 23에 개략적으로 설명된 바와 같이 상이한 플로우셀(flowcell)을 사용하여 제조된 샘플 간의 주르카트 및 라모스 세포의 발현 프로파일에 대한 배치(batch) 효과를 나타내는 비제한적이고 예시적인 tSNE 플롯이다.

도 36aa 내지 도 36ai, 도 36b 및 도 36c는 회석 적정을 사용하여 항체 스톡의 최적 회석의 결정을 나타내는 비제한적이고 예시적인 플롯이다.

도 37은 항체 올리고뉴클레오타이드가 서열분석 데이터에서 총 판독물의 원하는 백분율을 차지하도록 하는 올리고뉴클레오타이드-접합된 항체의 염색 농도를 결정하기 위한 비제한적이고 예시적인 실험 설계를 나타낸다.

도 38a 내지 도 38d는 항체 올리고뉴클레오타이드의 예측된 크기와 일치하는 피크(화살표로 표시됨)를 나타내는 비제한적이고 예시적인 생체분석기 트레이스이다.

도 39aa 내지 도 39ac 및 도 39ba 내지 도 39bc는 상이한 항체 회석 및 항체 올리고뉴클레오타이드와 접합된 항체 분자("고온 항체")의 상이한 백분율로 염색된 샘플에 대해 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 비제한적이고 예시적인 히스토그램이다.

도 40a 내지 도 40c는 올리고뉴클레오타이드-접합된 항-CD147 항체 분자가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형은 혈액 패널 내 488개 유전자의 발현 프로파일을 사용하여 결정되었다(도 40a). 세포는 1:100 희석된 스톡을 사용하여 제조된 10% 고온 항체:90% 저온 항체의 혼합물로 염색되었으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 명확한 신호를 생성하였다(도 40b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 40c에 나타내어져 있다.

도 41a 내지 도 41c는 올리고뉴클레오타이드-접합된 항-CD147 항체가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형은 혈액 패널 내 488개 유전자의 발현 프로파일을 사용하여 결정되었다(도 41a). 세포는 1:100으로 희석된 스톡을 사용하여 제조된 1% 고온 항체:99% 저온 항체의

혼합물로 염색되었으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 어떠한 명확한 신호도 생성하지 않았다(도 41b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 41c에 나타내어져 있다.

도 42a 내지 도 42c는 올리고뉴클레오타이드-접합된 항-CD147 항체 분자가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형은 혈액 패널 내 488개 유전자의 발현 프로파일을 사용하여 결정되었다(도 42a). 세포는 1:800으로 희석된 스톡으로 염색되었으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 명확한 신호를 생성하였다(도 42b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 42c에 나타내어져 있다.

도 43a 및 도 43b는 염색 완충액 내 제어 입자 올리고뉴클레오타이드 및 도 10에 예시된 작업 흐름을 사용하여 검출된 제어 입자와 회합된 제어 입자 올리고뉴클레오타이드의 조성을 나타내는 플롯이다.

도 44a 및 도 44b는 혈구계(hemocytometer)에서 세포(도 44a, 백색 원) 및 제어 입자(도 44b, 흑색 원)의 명시야상(brightfield image)이다.

도 45a 및 도 45b는 형광단에 회합된 올리고뉴클레오타이드-접합된 항체에 결합된 제어 입자의 위상차(도 45a, 10 X) 및 형광(도 45b, 10 X) 이미지이다.

도 46은 카트리지의 마이크로웰 내로 로딩되는 세포 및 제어 입자를 나타내는 이미지이다.

도 47a 내지 도 47c는 샘플 인덱싱을 위한 항체 각테일을 사용함으로써 표지화 감도를 증가시킬 수 있음을 나타내는 플롯이다.

도 48은 멀티플렉스 샘플 인덱싱을 사용하여 식별되고 서열분석 데이터로부터 제거될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다.

도 49a는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 싱글렛 또는 멀티플렉스로서 식별된 2마리의 마우스로부터의 6개의 조직의 12개의 샘플로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 49b는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별되고 도 49a에 나타나 있는 멀티플렉스 제거된 2마리의 마우스로부터의 6개의 조직의 12개의 샘플로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 50a는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별된 멀티플렉스 제거된 2마리의 마우스로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다.

도 50b 및 도 50c는, 멀티플렉스 발현 프로파일을 제거한 후에, 생물학적 복제물인 2마리의 마우스가 유사한 발현 프로파일을 나타내었음을 나타내는 비제한적이고 예시적인 파이 차트이다.

도 51a 내지 도 51f는 서열분석 데이터에서의 멀티플렉스 발현 프로파일이 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별되고 제거된 6개의 상이한 조직의 면역 세포 프로파일을 나타내는 비제한적이고 예시적인 파이 차트이다.

도 52a 내지 도 52c는 서열분석 데이터에서의 멀티플렉스 발현 프로파일이 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별되고 제거된 후에 6개의 상이한 조직으로부터의 대식세포, T 세포, 및 B 세포의 발현 프로파일을 나타내는 비제한적이고 예시적인 그래프이다.

도 53a 내지 도 53d는 결장과 비장에서의 대식세포를 비교하는 비제한적이고 예시적인 플롯이다.

도 54는 샘플 인덱싱 조성물에 의한 세포의 태깅이 PBMC에서의 mRNA 발현 프로파일을 변경시키지 않았음을 나타내는 비제한적이고 예시적인 플롯이다.

발명을 실시하기 위한 구체적인 내용

하기 상세한 설명에서, 본 명세서의 일부를 이루는 첨부 도면을 참조한다. 도면에서, 문맥에서 달리 지시되지 않는 한, 유사한 부호는 통상적으로 유사한 구성요소를 나타낸다. 발명의 상세한 설명, 도면 및 청구범위에 기재된 예시적인 구현에는 제한을 의미하지 않는다. 본 명세서에 제시된 주제의 사상 또는 범주를 벗어나지 않으면서 다른 구현예가 사용될 수 있고, 다른 변화들이 이루어질 수 있다. 본 개시 내용의 양태는, 본 명세서에 일

[0310]

반적으로 기재되고, 도면에 예시된 바와 같이, 광범위한 다양한 상이한 구성으로 배열, 치환, 조합, 구분 및 설계될 수 있으며, 이들 모두는 본 명세서에서 명시적으로 고려되고, 본 명세서의 개시 내용의 일부를 형성함이 쉽게 이해될 것이다.

[0311] 모든 특허, 공개된 특허 출원, 기타 다른 간행물, 및 GenBank로부터의 서열, 및 본 명세서에서 언급된 기타 다른 데이터베이스는 관련 기술에 대하여 전체가 참고로 포함된다.

[0312] 적은 수의 핵산, 예를 들어 메신저 리보핵산(mRNA) 분자를 정량하는 것은, 예를 들어 개발의 상이한 단계에서 또는 상이한 환경 조건 하에서 세포 내에서 발현되는 유전자를 결정하는 데 임상적으로 중요하다. 그러나, 특히 분자 수가 매우 적을 경우, 핵산 분자(예를 들어, mRNA 분자)의 절대 수를 결정하는 것 또한 매우 어려운 도전이 될 수 있다. 샘플 내 분자의 절대 수를 결정하는 하나의 방법은 디지털 중합효소 연쇄 반응(PCR)이다. 이상적으로, PCR은 각각의 사이클에서 분자의 동일한 복제물을 생성한다. 그러나, PCR은 각각의 분자가 추계적 확률로 복제되고, 이러한 확률은 PCR 사이클 및 유전자 서열에 의해 변화되어, 증폭 편향 및 부정확한 유전자 발현 측정을 초래하게 하는 단점을 가질 수 있다. 고유 분자 표지를 갖는 추계적 바코드(분자 인덱스(MI)로도 지칭됨)는 분자 수의 계수와 증폭 편향에 대한 보정에 사용될 수 있다. 추계적 바코딩, 예컨대 Precise™ 분석(Cellular Research, Inc.(미국 캘리포니아주 팰로앨토 소재))은 역전사(RT) 동안 mRNA를 표지하기 위해 분자 표지(ML)를 사용함으로써, PCR 및 라이브러리 제조 단계에 의해 유도된 편향을 보정할 수 있다.

[0313] Precise™ 분석은 큰 수, 예를 들어, 6561 내지 65536를 갖는 추계적 바코드의 비-고갈 풀(non-depleting pool)인, RT 단계 동안 샘플에서 모든 폴리(A)-mRNA에 혼성화하는, 폴리(T) 올리고뉴클레오타이드 상의 고유 분자 표지를 이용할 수 있다. 추계적 바코드는 범용 PCR 프라이밍 부위를 포함할 수 있다. RT 동안, 표적 유전자 분자는 추계적 바코드와 랜덤으로 반응한다. 각각의 표적 분자는 추계적 바코드에 혼성화되어 추계적으로 바코딩된 상보적 리보핵산(cDNA) 분자를 생성할 수 있다. 표지 후, 마이크로웰 플레이트의 마이크로웰들로부터의 추계적으로 바코딩된 cDNA 분자는 PCR 증폭 및 서열분석을 위해 단일 튜브 내로 풀링될(pooled) 수 있다. 미가공 서열 분석 데이터를 분석하여 판독물 수, 고유 분자 표지를 갖는 추계적 바코드의 수, 및 mRNA 분자 수를 산출할 수 있다.

[0314] 단일 세포의 mRNA 발현 프로파일 결정 방법은 대량 병렬 방식으로 수행될 수 있다. 예를 들어, Precise™ 분석을 사용하여 동시에 10000개 초과 세포의 mRNA 발현 프로파일을 결정할 수 있다. 샘플 당 분석을 위한 단일 세포의 수(예를 들어, 100개 또는 1000개의 단일 세포들)는 현재의 단일 세포 기술의 용량보다 더 낮을 수 있다. 상이한 샘플로부터 세포를 풀링(pooling)하는 것은 현재의 단일 기술의 용량의 활용 개선을 가능하게 하여, 낭비되는 시약과 단일 세포 분석 비용을 절감할 수 있다. 본 개시 내용은 세포 분석, 예컨대 단일 세포 분석을 위한 cDNA 라이브러리 제조를 위하여 상이한 샘플의 세포들을 구별하기 위한 샘플 인덱싱 방법을 제공한다. 상이한 샘플로부터 세포를 풀링(pooling)하는 것은 현재의 단일 기술의 용량의 활용을 개선하여, 낭비되는 시약과 단일 세포 분석 비용을 절감할 수 있다. 상이한 샘플들로부터 세포를 풀링하는 것은 상이한 샘플의 세포의 cDNA 라이브러리 제조시 변형을 최소화할 수 있어, 상이한 샘플을 더 정확하게 비교할 수 있게 한다.

[0315] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 항원 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적 바코딩)하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(예를 들어, 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드)를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다.

[0316] 일부 구현예에서, 본 명세서에 개시된 샘플 식별을 위한 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합

된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계; 복수의 바코드(예를 들어, 추계적 바코드)를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적 바코딩)하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(예를 들어, 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드)를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다.

[0317] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 항원 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 하나 이상의 항원 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계; 및 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다.

[0318] 본 명세서의 개시 내용은 복수의 샘플 인덱싱 조성물이다. 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함할 수 있다. 단백질 결합 시약은 적어도 하나의 항원 표적에 특이적으로 결합할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 샘플의 하나 이상의 세포의 샘플 기원을 식별하기 위한 샘플 인덱싱 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0319] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포는 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약(예를 들어, 항체)을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적(예를 들어, 단백질) 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 데이터-체이닝 증폭 프라임을 사용하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 데이터-체이닝 신장된 앰플리콘을 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 포함하는 복수의 데이터-체이닝 신장된 앰플리콘의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다. 방법은 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함할 수 있다.

[0320] 일부 구현예에서, 본 명세서에 개시된 샘플 인덱싱 방법, 키트, 및 조성물은 (예를 들어, 낮은 세포 수의 샘플, 단리하기 어려운 세포, 및 불균질한 세포에 대해) 샘플 처리량을 증가시킬 수 있고/있거나, 시약 비용을 낮출 수 있고/있거나, 단일 튜브 반응으로 라이브러리 제조를 수행함으로써 기술적 오류 및 배치 효과를 감소시킬 수 있고/있거나, 데이터 분석 동안 샘플간 다중 세포를 식별할 수 있다. 일부 구현예에서, 상이한 림프계 기관 및 비-림프계 기관으로부터의 조직의 세포는, 예를 들어 샘플 처리량을 증가시키고 배치 효과를 감소시키거나 최소화하기 위해, 상이한 샘플 인덱싱 조성물을 사용하여 태깅될 수 있다. 일부 구현예에서, 면역 방어 및 조직 항상성 및 기능이 본 개시 내용의 방법, 키트, 및 조성물을 사용하여 연구될 수 있다.

[0321] 정의

[0322] 달리 정의되지 않으면, 본 명세서에서 사용되는 기술 용어 및 과학 용어는 본 개시 내용이 속하는 기술 분야의 당업자가 일반적으로 이해하는 것과 동일한 의미를 갖는다. 예를 들어, 문헌[Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, NY 1994)]; 문헌[Sambrook et

al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY 1989)]을 참조한다. 본 개시 내용의 목적상, 하기 용어가 아래에 정의되어 있다.

[0323] 본 명세서에 사용되는 바와 같이, 용어 "어댑터"는 회합된 핵산의 증폭 또는 서열분석을 용이하게 하는 서열을 의미할 수 있다. 회합된 핵산은 표적 핵산을 포함할 수 있다. 회합된 핵산은 공간 표지, 표적 표지, 샘플 표지, 인덱싱 표지, 또는 바코드 서열(예를 들어, 분자 표지) 중 하나 이상을 포함할 수 있다. 어댑터는 선형일 수 있다. 어댑터는 사전-아데닐화된 어댑터일 수 있다. 어댑터는 이중 가닥 또는 단일 가닥일 수 있다. 하나 이상의 어댑터는 핵산의 5' 또는 3' 말단 상에 위치할 수 있다. 어댑터가 5' 및 3' 말단 상에 공지된 서열을 포함하는 경우, 공지된 서열은 동일하거나 상이한 서열일 수 있다. 폴리뉴클레오티드의 5' 및/또는 3' 말단 상에 위치한 어댑터는 표면 상에 고정화된 하나 이상의 올리고뉴클레오티드에 혼성화될 수 있다. 어댑터는 일부 구현예에서, 범용 서열을 포함한다. 범용 서열은 2개 이상의 핵산 분자에 공통적인 뉴클레오티드 서열의 영역일 수 있다. 2개 이상의 핵산 분자는 또한 상이한 서열의 영역을 가질 수 있다. 따라서, 예를 들어, 5' 어댑터는 동일한 핵산 서열 및/또는 범용 핵산 서열을 포함할 수 있고, 3' 어댑터는 동일한 서열 및/또는 범용 서열을 포함할 수 있다. 복수의 핵산 분자의 상이한 구성원에 존재할 수 있는 범용 서열은, 범용 서열에 상보적인 단일 범용 프라이머를 사용하여 다수의 상이한 서열을 복제하거나 증폭시킬 수 있게 할 수 있다. 유사하게, 핵산 분자의 집합의 상이한 구성원에 존재할 수 있는 적어도 하나, 둘(예를 들어, 한 쌍), 또는 그 이상의 범용 서열은, 범용 서열에 상보적인 적어도 하나, 둘(예를 들어, 한 쌍), 또는 그 이상의 단일 범용 프라이머를 사용하여, 다수의 상이한 서열을 복제하거나 증폭시킬 수 있게 할 수 있다. 따라서, 범용 프라이머는 그러한 범용 서열에 혼성화될 수 있는 서열을 포함한다. 표적 핵산 서열-보유 분자는 범용 어댑터(예를 들어, 비-표적 핵산 서열)를 상이한 표적 핵산 서열의 한쪽 또는 양쪽 말단에 부착하도록 변형될 수 있다. 표적 핵산에 부착된 하나 이상의 범용 프라이머는 범용 프라이머의 혼성화를 위한 부위를 제공할 수 있다. 표적 핵산에 부착된 하나 이상의 범용 프라이머는 서로 동일하거나 상이할 수 있다.

[0324] 본 명세서에 사용되는 바와 같이, 항체는 전장(full-length)(예를 들어, 자연적으로 발생하거나 정상의 면역글로불린 유전자 단편 재조합 과정에 의해 형성됨) 면역글로불린 분자(예를 들어, IgG 항체) 또는 면역글로불린 분자의 면역학적으로 활성인(즉, 특이적으로 결합하는) 일부분, 예컨대 항체 단편일 수 있다.

[0325] 일부 구현예에서, 항체는 기능적 항체 단편이다. 예를 들어, 항체 단편은 항체의 일부분, 예컨대 F(ab')₂, Fab', Fab, Fv, sFv 등일 수 있다. 항체 단편은 전장 항체에 의해 인식되는 동일한 항원과 결합할 수 있다. 항체 단편은 항체의 가변 영역으로 이루어진 단리된 단편, 예컨대 중쇄 및 경쇄의 가변 영역으로 이루어진 "Fv" 단편, 및 경쇄 및 중쇄 가변 영역이 펩티드 링커에 의해 연결된 재조합 단일 사슬 폴리펩티드 분자("scFv 단백질")를 포함할 수 있다. 예시적인 항체는 암세포에 대한 항체, 바이러스에 대한 항체, 세포 표면 수용체(예를 들어, CD8, CD34, 및 CD45)에 결합하는 항체, 및 치료제 항체를 포함할 수 있으며, 이에 제한되지는 않는다.

[0326] 본 명세서에 사용되는 바와 같이 용어 "회합된" 또는 "와 회합된"은 2개 이상의 종이 일정 시점에서 공동-배치된 것으로 식별됨을 의미할 수 있다. 회합은 2개 이상의 종이 유사한 용기 내에 존재하거나 존재하였음을 의미할 수 있다. 회합은 정보적 회합일 수 있다. 예를 들어, 2개 이상의 종과 관련되는 디지털 정보가 저장될 수 있고, 하나 이상의 종이 일정 시점에 공동 배치되었음을 결정하는 데 사용될 수 있다. 회합은 또한 물리적 회합일 수 있다. 일부 구현예에서, 2개 이상의 회합된 종은 서로 또는 공통 고체 또는 반고체 표면에 "묶이거나(tethered)", "부착되거나", "고정된다". 회합은 표지를 고체 또는 반-고체 지지체, 예컨대 비드에 부착시키기 위한 공유결합 또는 비-공유결합 수단을 지칭할 수 있다. 회합은 표적과 표지 사이의 공유 결합일 수 있다. 회합은 2개의 분자(예컨대, 표적 분자와 표지) 사이의 혼성화를 포함할 수 있다.

[0327] 본 명세서에 사용되는 바와 같이, 용어 "상보적인"은 2개의 뉴클레오티드 사이의 정확한 페어링(pairing)을 위한 용량을 지칭할 수 있다. 예를 들어, 핵산의 소정의 위치에서의 뉴클레오티드가 또 다른 핵산의 뉴클레오티드와 수소 결합할 수 있는 경우, 2개의 핵산은 그 위치에서 서로 상보적인 것으로 여겨진다. 2개의 단일 가닥 핵산 분자 사이의 상보성은 단지 일부의 뉴클레오티드가 결합하는 "부분적"일 수 있거나, 단일 가닥 분자 사이에 전체 상보성이 존재할 때 완전할 수 있다. 제1 뉴클레오티드 서열은, 제1 뉴클레오티드 서열이 제2 뉴클레오티드 서열에 상보적인 경우 제2 서열의 "보체"라고 말할 수 있다. 제1 뉴클레오티드 서열은, 제1 뉴클레오티드 서열이 제2 서열의 역(즉, 뉴클레오티드의 순서가 역전됨)인 서열에 상보적인 경우, 제2 서열의 "역보체"라고 말할 수 있다. 본 명세서에 사용되는 바와 같이, 용어 "보체", "상보성", 및 "역보체"는 상호변경가능하게 사용될 수 있다. 분자가 또 다른 분자에 혼성화될 수 있는 경우, 분자는 혼성화하는 분자의 보체일 수 있음이 본 개시 내용으로부터 이해된다.

- [0328] 본 명세서에 사용되는 바와 같이, 용어 "디지털 계수"는 샘플 내 다수의 표적 분자를 추산하기 위한 방법을 지칭할 수 있다. 디지털 계수는 샘플 내 표적과 회합된 다수의 고유 표지를 결정하는 단계를 포함할 수 있다. 성질이 추계적일 수 있는 이러한 방법론은, 미리 정의된 표지의 세트의 검출과 관련하여 분자를 계수하는 문제를 동일한 분자의 배치 및 식별 중 하나로부터 일련의 예/아니오의 디지털 질문으로 변환시킨다.
- [0329] 본 명세서에 사용되는 바와 같이, 용어 "표지" 또는 "표지들"은 샘플 내에서 표적과 회합된 핵산 코드를 지칭할 수 있다. 표지는 예를 들어, 핵산 표지일 수 있다. 표지는 전체적으로 또는 부분적으로 증폭 가능한 표지일 수 있다. 표지는 전체적으로 또는 부분적으로 서열분석 가능한 표지일 수 있다. 표지는 별개의 것으로 식별될 수 있는 천연 핵산의 일부분일 수 있다. 표지는 공지된 서열일 수 있다. 표지는 핵산 서열의 접합점(junction), 예를 들어, 천연 및 비-천연 서열의 접합점을 포함할 수 있다. 본 명세서에 사용되는 바와 같이, 용어 "표지(label)"는 용어 "인덱스", "태그" 또는 "표지-태그"와 상호교환가능하게 사용될 수 있다. 표지는 정보를 전달할 수 있다. 예를 들어, 다양한 구현예에서, 표지는 샘플의 정체(identity), 샘플의 공급원, 세포의 정체, 및/또는 표적을 결정하는 데 사용될 수 있다.
- [0330] 본 명세서에 사용되는 바와 같이, 용어 "비-고갈 저장소"는 많은 상이한 표지로 이루어진 바코드(예를 들어, 추계적 바코드)의 풀을 지칭할 수 있다. 비-고갈 저장소는 많은 수의 상이한 바코드를 포함할 수 있으므로, 비-고갈 저장소가 표적의 풀과 회합되는 경우 각각의 표적은 고유 바코드와 회합될 가능성이 있다. 각각의 표지된 표적 분자의 고유성은 랜덤 선택의 통계에 의해 결정될 수 있으며, 표지의 다양성에 비해 집합 내 동일한 표적 분자의 복제물의 수에 따라 달라진다. 표지된 표적 분자의 생성된 세트의 크기는 바코딩 과정의 추계적 성질에 의해 결정될 수 있고, 이어서 검출된 바코드의 수를 분석함으로써 원래의 집합 또는 샘플 내에 존재하는 표적 분자 수를 계산할 수 있다. 고유 바코드의 수에 대해 존재하는 표적 분자의 복제물의 수의 비가 낮으면, 표지된 표적 분자는 매우 고유하다(즉, 하나 초과 표적 분자가 소정의 표지로 표지될 가능성이 매우 낮음).
- [0331] 본 명세서에 사용되는 바와 같이, 용어 "핵산"은 폴리뉴클레오타이드 서열, 또는 이의 단편을 지칭한다. 핵산은 뉴클레오타이드를 포함할 수 있다. 핵산은 세포에 대해 외인성이거나 내인성일 수 있다. 핵산은 무세포(cell-free) 환경에서 존재할 수 있다. 핵산은 유전자 또는 그의 단편일 수 있다. 핵산은 DNA일 수 있다. 핵산은 RNA일 수 있다. 핵산은 하나 이상의 유사체(예를 들어, 변경된 주쇄, 당, 또는 핵염기)를 포함할 수 있다. 유사체의 일부 비제한적 예는, 5-브로모우라실, 펩티드 핵산, 이종(xeno) 핵산, 모르폴리노, 고정(locked) 핵산, 글리콜 핵산, 트레오스 핵산, 디데옥시뉴클레오타이드, 코르디세핀(cordycepin), 7-데아자-GTP, 형광단(예를 들어, 당에 회합된 로다민 또는 플루오레세인), 티올-함유 뉴클레오타이드, 비오틴-연결 뉴클레오타이드, 형광 염기 유사체, CpG 섬(island), 메틸-7-구아노신, 메틸화 뉴클레오타이드, 이노신, 티오우리딘, 의사우리딘, 디하이드로우리딘, 큐에오신(queuosine), 및 와이오신(wyosine)을 포함한다. "핵산", "폴리뉴클레오타이드", "표적 폴리뉴클레오타이드", 및 "표적 핵산"은 상호교환가능하게 사용될 수 있다.
- [0332] 핵산은 신규하거나 향상된 특성(예를 들어, 개선된 안정성)을 핵산에 제공하기 위해 하나 이상의 변형(예를 들어, 염기 변형, 주쇄 변형)을 포함할 수 있다. 핵산은 핵산 친화성 태그를 포함할 수 있다. 뉴클레오타이드는 염기-당 조합일 수 있다. 뉴클레오타이드의 염기 부분은 헤테로시클릭 염기일 수 있다. 그러한 헤테로시클릭 염기의 가장 일반적인 2가지 부류는 퓨린 및 피리미딘이다. 뉴클레오타이드는 뉴클레오타이드의 당 부분에 공유결합된 포스페이트 기를 추가로 포함하는 뉴클레오타이드일 수 있다. 펜토푸라노실 당을 포함하는 뉴클레오타이드의 경우, 포스페이트 기는 당의 2', 3', 또는 5' 히드록실 모이어티에 연결될 수 있다. 핵산 형성에 있어서, 포스페이트 기는 인접 뉴클레오타이드들을 서로 공유결합시켜 선형 중합체 화합물을 형성할 수 있다. 결국, 이러한 선형 중합체 화합물의 개별적인 말단은 추가로 결합되어 원형 화합물을 형성할 수 있지만; 선형 화합물이 일반적으로 적합하다. 추가적으로, 선형 화합물은 내부 뉴클레오타이드 염기 상보성을 가질 수 있고, 따라서 완전히 또는 부분적으로 이중-가닥 화합물을 생성하는 방식으로 접힐 수 있다. 핵산 내에서, 포스페이트 기는 일반적으로 핵산의 뉴클레오타이드간 주쇄를 형성하는 것으로 언급될 수 있다. 결합 또는 주쇄는 3'-5' 포스포디에스테르 결합일 수 있다.
- [0333] 핵산은 변형된 주쇄 및/또는 변형된 뉴클레오타이드간 결합을 포함할 수 있다. 변형된 주쇄는 주쇄에 인 원자를 보유하는 주쇄와 주쇄에 인 원자를 갖지 않는 주쇄를 포함할 수 있다. 그 안에 인 원자를 함유하는 적합한 변형된 핵산 주쇄는, 예를 들어 포스포로티오에이트, 키랄 포스포로티오에이트, 포스포로디티오에이트, 포스포트리에스테르, 아미노알킬 포스포트리에스테르, 메틸 및 기타 다른 알킬 포스포네이트, 예컨대 3'-알킬렌 포스포네이트, 5'-알킬렌 포스포네이트, 키랄 포스포네이트, 포스포네이트, 3'-아미노포스포아미데이트 및 아미노알킬 포스포아미데이트를 포함하는 포스포아미데이트, 포스포로디아미데이트, 티오노포스포아미데이트, 티오노알킬포스포네이트, 티오노알킬포스포트리에스테르, 셀레노포스포에이트, 및 정상 3'-5' 연결을 갖는 보라노포스포에이트,

2'-5' 연결된 유사체, 및 하나 이상의 뉴클레오타이드간 연결이 3' 에서 3', 5' 에서 5' 또는 2' 에서 2' 결합인 역전된 극성을 갖는 것들을 포함할 수 있다.

[0334] 핵산은 단쇄(short chain) 알킬 또는 시클로알킬 뉴클레오시드간 연결, 혼합된 헤테로원자와 알킬 또는 시클로알킬 뉴클레오시드간 연결, 또는 하나 이상의 단쇄 헤테로원자 또는 헤테로시클릭 뉴클레오시드간 연결에 의해 형성된 폴리뉴클레오타이드 주쇄를 포함할 수 있다. 이들은 모르폴리노 연결(뉴클레오시드의 당 부분으로부터 부분적으로 형성됨); 실록산 주쇄; 설퍼드, 설퍼시드 및 설퍼 주쇄; 포름아세틸 및 티오포름아세틸 주쇄; 메틸렌 포름아세틸 및 티오포름아세틸 주쇄; 리보아세틸 주쇄; 알켄-함유 주쇄; 설파메이트 주쇄; 메틸렌이미노 및 메틸렌히드라지노 주쇄; 설퍼네이트 및 설퍼아미드 주쇄; 아미드 주쇄를 갖는 것들; 및 혼합된 N, O, S 및 CH₂ 성분 부분을 갖는 기타 다른 것들을 포함할 수 있다.

[0335] 핵산은 핵산 모방체를 포함할 수 있다. 용어 "모방체"는 단지 푸라노스 고리 또는 푸라노스 고리와 뉴클레오타이드간 연결 모두가 비-푸라노스 기로 치환된 폴리뉴클레오타이드를 포함하고자 할 수 있고, 푸라노스 고리만의 대체는 또한 당 대체물인 것으로 지칭될 수 있다. 헤테로시클릭 염기 모이어티 또는 변경된 헤테로시클릭 염기 모이어티는 적절한 표적 핵산과의 혼성화를 위해 유지될 수 있다. 그러한 하나의 핵산은 펩티드 핵산(PNA)일 수 있다. PNA에서, 폴리뉴클레오타이드의 당-주쇄는 아미드-함유 주쇄, 구체적으로 아미노에틸글리신 주쇄로 대체될 수 있다. 뉴클레오타이드는 보유될 수 있고, 주쇄의 아미드 부분의 아자 질소 원자에 직접적으로 또는 간접적으로 결합된다. PNA 화합물의 주쇄는 PNA에 아미드-함유 주쇄를 부여하는 2개 이상의 연결된 아미노에틸글리신 단위를 포함할 수 있다. 헤테로시클릭 염기 모이어티는 주쇄의 아미드 부분의 아자 질소 원자에 직접적으로 또는 간접적으로 결합될 수 있다.

[0336] 핵산은 모르폴리노 주쇄 구조를 포함할 수 있다. 예를 들어, 핵산은 리보스 고리 대신 6-원 모르폴리노 고리를 포함할 수 있다. 이들 일부 구현예에서, 포스포로디아미테이트 또는 기타 다른 비-포스포디에스테르 뉴클레오시드간 연결은 포스포디에스테르 결합을 대체할 수 있다.

[0337] 핵산은, 모르폴리노 고리에 부착된 헤테로시클릭 염기를 갖는 회합된 모르폴리노 단위(즉, 모르폴리노 핵산)를 포함할 수 있다. 연결 기는 모르폴리노 핵산에서 모르폴리노 단량체 단위를 연결할 수 있다. 비이온성 모르폴리노계 올리고머성 화합물은 세포 단백질과 원치않는 상호작용이 더 적을 수 있다. 모르폴리노계 폴리뉴클레오타이드는 핵산의 비이온성 모방체일 수 있다. 모르폴리노 부류에 속하는 다양한 화합물은 상이한 회합 기를 사용하여 연결될 수 있다. 폴리뉴클레오타이드 모방체의 추가의 부류는 시클로헥세닐 핵산(CeNA)으로 지칭될 수 있다. 핵산 분자 중에 정상적으로 존재하는 푸라노스 고리는 시클로헥세닐 고리로 대체될 수 있다. CeNA DMT 보호된 포스포아미다이트 단량체는 제조되어, 포스포아미다이트 화학을 사용하여 올리고머성 화합물 합성에 사용될 수 있다. CeNA 단량체의 핵산 사슬 내로의 혼입은 DNA/RNA 혼성체의 안정성을 증가시킬 수 있다. CeNA 올리고아데닐레이트는 천연 복합체와 유사한 안정성을 갖는 핵산 보체와 복합체를 형성할 수 있다. 추가의 변형은 2'-히드록실 기가 당 고리의 4' 탄소 원자에 연결되어, 2'-C, 4'-C-옥시메틸렌 연결을 형성하여, 바이시클릭 당 모이어티를 형성하는 고정 핵산(LNA)을 포함할 수 있다. 연결은 *n*이 1 또는 2이고, 2' 산소 원자와 4' 탄소 원자를 가 교하는 기인, 메틸렌(-CH₂)일 수 있다. LNA 및 LNA 유사체는 상보적 핵산과 함께 매우 높은 이중(duplex) 열 안정성(T_m+3 내지 +10°C), 3'-엑소뉴클레오타이드(exonucleolytic) 분해에 대한 안정성 및 양호한 용해도 특성을 나타낼 수 있다.

[0338] 핵산은 또한 핵염기(종종 간단히 "염기"로도 지칭됨) 변형 또는 치환을 포함할 수 있다. 본 명세서에 사용되는 바와 같이, "변형된" 또는 "천연" 핵염기는 퓨린 염기(예를 들어, 아데닌(A) 및 구아닌(G)), 및 피리미딘 염기(예를 들어, 티민(T), 시토신(C) 및 우라실(U))을 포함할 수 있다. 변형된 핵염기는 기타 다른 합성 및 천연 핵염기, 예컨대 5-메틸시토신(5-me-C), 5-히드록시메틸 시토신, 잔틴, 하이포잔틴, 2-아미노아데닌, 아데닌 및 구아닌의 6-메틸 및 기타 다른 알킬 유도체, 아데닌 및 구아닌의 2-프로필 및 기타 다른 알킬 유도체, 2-티오우라실, 2-티오티민 및 2-티오시토신, 5-할로우라실 및 시토신, 5-프로피닐(-C≡CH₃) 우라실 및 시토신과 피리미딘 염기의 기타 다른 알킬 유도체, 6-아조 우라실, 시토신 및 티민, 5-우라실(의사우라실), 4-티오우라실, 8-할로, 8-아미노, 8-티올, 8-티오알킬, 8-히드록실 및 기타 다른 8-치환된 아데닌 및 구아닌, 5-할로, 특히 5-브로모, 5-트리플루오로메틸 및 기타 다른 5-치환된 우라실 및 시토신, 7-메틸구아닌 및 7-메틸아데닌, 2-F-아데닌, 2-아미노아데닌, 8-아자구아닌 및 8-아자아데닌, 7-데아자구아닌과 7-데아자아데닌 및 3-데아자구아닌과 3-데아자아데닌을 포함할 수 있다. 변형된 핵염기는 트리시클릭 피리미딘, 예컨대 페녹사진 시티딘(1H-피리미도(5,4-b)(1,4)벤조티아진-2(3H)-온), 페노티아진 시티딘(1H-피리미도(5,4-b)(1,4)벤조티아진-2(3H)-온), G-클램프(clamp), 예컨대 치환된 페녹사진 시티딘(예를 들어, 9-(2-아미노에톡시)-H-피리미도(5,4-b)(1,4)벤조옥사진

-2(3H)-온), 페노티아진 시티딘(1H-피리미도(5,4-b)(1,4)벤조티아진-2(3H)-온), G-클램프, 예컨대 치환된 페녹사진 시티딘(예를 들어, 9-(2-아미노에톡시)-H-피리미도(5,4-b)(1,4)벤즈옥사진-2(3H)-온), 카르바졸 시티딘(2H-피리미도(4,5-b)인돌-2-온), 피리도인돌 시티딘(H-피리도(3',2':4,5)피롤로[2,3-d]피리미딘-2-온)을 포함할 수 있다.

[0339] 본 명세서에 사용되는 바와 같이, 용어 "샘플"은 표적을 포함하는 조성물을 지칭할 수 있다. 개시된 방법, 디바이스 및 시스템에 의한 분석에 적합한 샘플은 세포, 조직, 기관 또는 유기체를 포함한다.

[0340] 본 명세서에 사용되는 바와 같이, 용어 "샘플링 디바이스" 또는 "디바이스"는 샘플의 절단부를 취하고/취하거나 기재 상에서 그 절단부를 위치시킬 수 있는 디바이스를 지칭할 수 있다. 샘플 디바이스는, 예를 들어 형광 활성화 세포 분류(FACS)기, 세포 분류기, 생검 바늘, 생검 디바이스, 조직 절단 디바이스, 마이크로유체 디바이스, 블레이드 그리드(blade grid), 및/또는 마이크로톰(microtome)을 지칭할 수 있다.

[0341] 본 명세서에 사용되는 바와 같이, 용어 "고체 지지체"는 복수의 바코드(예를 들어, 추계적 바코드)가 부착될 수 있는 별개의 고체 또는 반고체 표면을 지칭할 수 있다. 고체 지지체는 핵산이 그 위에 (예를 들어, 공유결합으로 또는 비-공유결합으로) 고정화될 수 있는 임의의 유형의 고체, 다공성, 또는 중공 구, 볼, 베어링, 실린더, 또는 플라스틱, 세라믹, 금속 또는 중합체성 재료(예를 들어, 하이드로겔)로 구성된 기타 다른 유사한 구조물을 포함할 수 있다. 고체 지지체는 구형(예를 들어, 미소구체)일 수 있거나 비-구형 또는 불규칙한 형태, 예컨대 입방체, 직육면체, 피라미드형, 원통형, 원뿔형, 장타원형(oblong), 또는 원반형태 등을 가질 수 있는 별개의 입자를 포함할 수 있다. 비드는 비-구형 형태일 수 있다. 어레이 내에 이격된 복수의 고체 지지체는 기재를 포함하지 않을 수 있다. 고체 지지체는 용어 "비드"와 상호교환가능하게 사용될 수 있다.

[0342] 본 명세서에 사용되는 바와 같이, 용어 "추계적 바코드"는 본 개시 내용의 표지를 포함하는 폴리뉴클레오티드 서열을 지칭할 수 있다. 추계적 바코드는 추계적 바코딩에 사용될 수 있는 폴리뉴클레오티드 서열일 수 있다. 추계적 바코드는 샘플 내에서 표적을 정량화하는 데 사용될 수 있다. 추계적 바코드는 표지가 표적과 회합된 후 발생할 수 있는 오류를 제어하는 데 사용될 수 있다. 예를 들어, 추계적 바코드는 증폭 또는 서열분석 오류를 평가하는 데 사용될 수 있다. 표적과 회합된 추계적 바코드는 추계적 바코드-표적 또는 추계적 바코드-태그-표적으로 칭해질 수 있다.

[0343] 본 명세서에 사용되는 바와 같이, 용어 "유전자-특이적 추계적 바코드"는 유전자-특이적인 표적-결합 영역 및 표지를 포함하는 폴리뉴클레오티드 서열을 지칭할 수 있다. 추계적 바코드는 추계적 바코딩에 사용될 수 있는 폴리뉴클레오티드 서열일 수 있다. 추계적 바코드는 샘플 내에서 표적을 정량화하는 데 사용될 수 있다. 추계적 바코드는 표지가 표적과 회합된 후에 발생할 수 있는 오류를 제어하는 데 사용될 수 있다. 예를 들어, 추계적 바코드는 증폭 또는 서열분석 오류를 평가하는 데 사용될 수 있다. 표적과 회합된 추계적 바코드는 추계적 바코드-표적 또는 추계적 바코드-태그-표적으로 칭해질 수 있다.

[0344] 본 명세서에 사용되는 바와 같이, 용어 "추계적 바코딩"은 핵산의 랜덤 표지화(예를 들어, 바코딩)를 지칭할 수 있다. 추계적 바코딩은 순환 푸와송(recursive Poisson) 전략을 이용하여 표적과 회합된 표지를 회합하고 정량화할 수 있다. 본 명세서에 사용되는 바와 같이, 용어 "추계적 바코딩"은 "추계적 표지화"와 상호교환가능하게 사용될 수 있다.

[0345] 본 명세서에 사용되는 바와 같이, 용어 "표적"은 바코드(예를 들어, 추계적 바코드)와 회합될 수 있는 조성물을 지칭할 수 있다. 개시된 방법, 디바이스, 및 시스템에 의한 분석에 적합한 예시적인 표적은 올리고뉴클레오티드, DNA, RNA, mRNA, 마이크로RNA, tRNA 등을 포함한다. 표적은 단일 가닥 또는 이중 가닥일 수 있다. 일부 구현예에서, 표적은 단백질, 펩티드, 또는 폴리펩티드일 수 있다. 일부 구현예에서, 표적은 지질이다. 본 명세서에 사용되는 바와 같이, "표적"은 "종"과 상호교환가능하게 사용될 수 있다.

[0346] 본 명세서에 사용되는 바와 같이, 용어 "역전사효소"는 역전사효소 활성을 갖는 효소의 군(즉, RNA 주형으로부터 DNA의 합성을 촉진함)을 지칭할 수 있다. 대체로, 그러한 효소는 레트로바이러스 역전사효소, 레트로트랜스포존(retrotransposon) 역전사효소, 레트로플라스미드 역전사효소, 레트로 역전사효소, 세균 역전사효소, II군 인트론-유도된 역전사효소, 및 이들의 돌연변이, 변이체 또는 유도체를 포함하지만, 이에 제한되지는 않는다. 비-레트로바이러스 역전사효소는 비-LTR 레트로트랜스포존 역전사효소, 레트로플라스미드 역전사효소, 레트로 역전사효소, 및 II군 인트론 역전사효소를 포함한다. II군 인트론 역전사효소의 예는 락토코커스 락티스 (*Lactococcus lactis*) LI.LtrB 인트론 역전사효소, 써모시네코코커스 일롱가투스(*Thermosynechococcus elongatus*) TeI4c 인트론 역전사효소, 또는 게오바실러스 스테아로써모필러스(*Geobacillus*

stearothermophilus) GsI-IIC 인트론 역전사효소를 포함한다. 기타 다른 부류의 역전사효소는 많은 부류의 비-레트로바이러스 역전사효소(즉, 특히 레트론, II군 인트론, 및 다양성-생성 레트로요소)를 포함할 수 있다.

[0347] 용어 "범용 어댑터 프라이머", "범용 프라이머 어댑터" 또는 "범용 어댑터 서열"은 바코드(예를 들어, 추계적 바코드)에 혼성화하여 유전자-특이적 바코드를 생성하는 데 사용될 수 있는 뉴클레오타이드 서열을 지칭하는 데 상호교환가능하게 사용된다. 범용 어댑터 서열은, 예를 들어 본 개시 내용의 방법에서 사용된 모든 바코드에 걸쳐 보편적인 공지된 서열일 수 있다. 예를 들어, 다수의 표적이 본 명세서에 개시된 방법을 사용하여 표지되는 경우, 표적-특이적 서열 각각은 동일한 범용 어댑터 서열에 연결될 수 있다. 일부 구현예에서, 하나 초과와 범용 어댑터 서열이 본 명세서에 개시된 방법에 사용될 수 있다. 예를 들어, 다수의 표적이 본 명세서에 개시된 방법을 사용하여 표지되는 경우, 적어도 2개의 표적-특이적 서열이 상이한 범용 어댑터 서열에 연결된다. 범용 어댑터 프라이머 및 그의 보체는 2개의 올리고뉴클레오타이드 내에 포함될 수 있으며, 그 중 하나는 표적-특이적 서열을 포함하고, 나머지 하나는 바코드를 포함한다. 예를 들어, 범용 어댑터 서열은, 표적 핵산에 상보적인 뉴클레오타이드 서열을 생성하기 위한 표적-특이적 서열을 포함하는 올리고뉴클레오타이드의 일부일 수 있다. 범용 어댑터 서열의 바코드 및 상보성 서열을 포함하는 제2 올리고뉴클레오타이드는 뉴클레오타이드 서열과 혼성화되고 표적-특이적 바코드(예를 들어, 표적-특이적 추계적 바코드)를 생성할 수 있다. 일부 구현예에서, 범용 어댑터 프라이머는 본 개시 내용의 방법에서 사용되는 범용 PCR 프라이머와 상이한 서열을 갖는다.

[0348] 바코드

[0349] 바코딩, 예컨대 추계적 바코딩은, 예를 들어, US20150299784, W02015031691, 및 문헌[Fu et al, Proc Natl Acad Sci U.S.A. 2011 May 31;108(22):9026-31]에 기재되어 있으며, 이들 간행물의 내용은 전체가 본 명세서에 포함된다. 일부 구현예에서, 본 명세서에 개시된 바코드는 표적을 추계적으로 표지(예를 들어, 바코드, 태그)하는 데 사용될 수 있는 폴리뉴클레오타이드 서열일 수 있는 추계적 바코드일 수 있다. 추계적 바코드의 상이한 바코드 서열의 수와, 표지되는 표적의 임의의 것의 출현 수의 비가 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있는 경우, 추계적 바코드로 지칭될 수 있다. 표적은 동일하거나 거의 동일한 서열을 갖는 mRNA 분자를 포함하는 mRNA 중일 수 있다. 바코드는 추계적 바코드의 상이한 바코드 서열의 수와 표지되는 표적의 임의의 것의 출현 수의 비가 적어도, 또는 최대 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 또는 100:1인 경우 추계적 바코드로 지칭될 수 있다. 추계적 바코드의 바코드 서열은 분자 표지로 지칭될 수 있다.

[0350] 바코드, 예를 들어, 추계적 바코드는 하나 이상의 표지를 포함할 수 있다. 예시적인 표지는 범용 표지, 세포 표지, 바코드 서열(예를 들어, 분자 표지), 샘플 표지, 플레이트 표지, 공간 표지, 및/또는 전-공간(pre-spatial) 표지를 포함할 수 있다. 도 1은 공간 표지를 갖는 예시적인 바코드(104)를 예시한다. 바코드(104)는 고체 지지체(105)에 바코드를 연결할 수 있는 5' 아민을 포함할 수 있다. 바코드는 범용 표지, 차원(dimension) 표지, 공간 표지, 세포 표지 및/또는 분자 표지를 포함할 수 있다. 바코드에서 상이한 표지들(범용 표지, 치수 표지, 공간 표지, 세포 표지, 및 분자 표지를 포함하지만, 이에 제한되지 않음)의 순서는 변화될 수 있다. 예를 들어, 도 1에 나타난 바와 같이, 범용 표지는 가장 5' 쪽(5'-most)의 표지일 수 있고, 분자 표지는 가장 3' 쪽(3'-most)의 표지일 수 있다. 공간 표지, 치수 표지, 및 세포 표지는 임의의 순서일 수 있다. 일부 구현예에서, 범용 표지, 공간 표지, 치수 표지, 세포 표지, 및 분자 표지는 임의의 순서이다. 바코드는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 샘플 내에서 표적(예를 들어, 표적 핵산, RNA, mRNA, DNA)과 상호작용할 수 있다. 예를 들어, 표적-결합 영역은 mRNA의 폴리(A) 테일과 상호작용할 수 있는 올리고(dT) 서열을 포함할 수 있다. 일부 경우에서, 바코드의 표지(예를 들어, 범용 표지, 치수 표지, 공간 표지, 세포 표지, 및 바코드 서열)는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20개, 또는 그 이상의 뉴클레오타이드에 의해 분리될 수 있다.

[0351] 표지, 예를 들어 세포 표지는 정의된 길이, 예를 들어 각각 7개의 뉴클레오타이드(일부 해밍(Hamming) 오류 정정 코드에 사용된 비트의 수에 균등함)의 핵산 하위서열의 고유 세트를 포함할 수 있고, 이는 오류 정정능을 제공하도록 설계될 수 있다. 7개의 뉴클레오타이드 서열을 포함하는 오류 정정 하위서열의 세트는, 세트 내 서열의 임의의 쌍의(pairwise) 조합이 정의된 "유전적 거리"(또는 미스매치된 염기의 수)를 나타내도록 설계될 수 있으며, 예를 들어 오류 정정 하위서열의 세트가 3개의 뉴클레오타이드의 유전적 거리를 나타내도록 설계될 수 있

다. 이러한 경우, 표지된 표적 핵산 분자에 대한 서열분석 데이터의 세트에서 오류 정정 서열을 검토하면(아래에 더욱 충분히 설명됨), 증폭 또는 서열분석 오류를 검출하거나 정정할 수 있게 된다. 일부 구현예에서, 오류 정정 코드를 생성하는 데 사용되는 핵산 하위서열의 길이는 변화될 수 있고, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 31, 40, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 31, 40, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 다른 길이의 핵산 하위서열은 오류 정정 코드를 생성하는 데 사용될 수 있다.

[0352] 바코드는 표적-결합 영역을 포함할 수 있다. 표적-결합 영역은 샘플 내에서 표적과 상호작용할 수 있다. 표적은 리보핵산(RNA), 메신저 RNA(mRNA), 마이크로RNA, 소간섭 RNAs(siRNA), RNA 분해 산물, 각각 폴리(A) 테일을 포함하는 RNA, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 일부 구현예에서, 복수의 표적은 데옥시리보핵산(DNA)을 포함할 수 있다.

[0353] 일부 구현예에서, 표적-결합 영역은 mRNA의 폴리(A) 테일과 상호작용할 수 있는 올리고(dT) 서열을 포함할 수 있다. 바코드의 표지(예를 들어, 범용 표지, 치수 표지, 공간 표지, 세포 표지, 및 바코드 서열(예를 들어, 분자 표지)) 중 하나 이상은 스페이서에 의해 바코드의 남아있는 표지 중 또 다른 하나 또는 둘로부터 구분될 수 있다. 스페이서는, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20 개, 또는 그 이상의 뉴클레오타이드일 수 있다. 일부 구현예에서, 바코드의 표지 중 어느 것도 스페이서에 의해 구분되지 않는다.

[0354] 범용 표지

[0355] 바코드는 하나 이상의 범용 표지를 포함할 수 있다. 일부 구현예에서, 하나 이상의 범용 표지는 소정의 고체 지지체에 부착된 바코드의 세트 내 모든 바코드에 대해 동일할 수 있다. 일부 구현예에서, 하나 이상의 범용 표지는 복수의 비드에 부착된 모든 바코드에 대해 동일할 수 있다. 일부 구현예에서, 범용 표지는 서열분석 프라이머에 혼성화할 수 있는 핵산 서열을 포함할 수 있다. 서열분석 프라이머는 범용 표지를 포함하는 서열분석 바코드를 위해 사용될 수 있다. 서열분석 프라이머(예를 들어, 범용 서열분석 프라이머)는 고처리량 서열분석 플랫폼과 회합된 서열분석 프라이머를 포함할 수 있다. 일부 구현예에서, 범용 표지는 PCR 프라이머에 혼성화할 수 있는 핵산 서열을 포함할 수 있다. 일부 구현예에서, 범용 표지는 서열분석 프라이머 및 PCR 프라이머에 혼성화할 수 있는 핵산 서열을 포함할 수 있다. 서열분석 또는 PCR 프라이머에 혼성화할 수 있는 범용 표지의 핵산 서열은 프라이머 결합 부위로 지칭될 수 있다. 범용 표지는 바코드의 전사를 개시하는 데 사용될 수 있는 서열을 포함할 수 있다. 범용 표지는 바코드의 연장에 사용될 수 있는 서열 또는 바코드 내 영역을 포함할 수 있다. 범용 표지는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 예를 들어, 범용 표지는 적어도 약 10개의 뉴클레오타이드를 포함할 수 있다. 범용 표지는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 절단 가능한 링커 또는 변형된 뉴클레오타이드는 바코드가 지지체로부터 절단되어 떨어질 수 있도록 하는 범용 표지 서열의 일부일 수 있다.

[0356] 치수 표지

[0357] 바코드는 하나 이상의 치수 표지를 포함할 수 있다. 일부 구현예에서, 치수 표지는 표지(예를 들어, 추계적 표지)가 발생한 치수에 대한 정보를 제공하는 핵산 서열을 포함할 수 있다. 예를 들어, 치수 표지는 표적이 바코딩된 시간에 대한 정보를 제공할 수 있다. 치수 표지는 샘플에서 바코딩(예를 들어, 추계적 바코딩)의 시간과 연결될 수 있다. 치수 표지는 표지 시에 활성화될 수 있다. 상이한 치수 표지는 상이한 시기에 활성화될 수 있다. 치수 표지는 표적, 표적의 군, 및/또는 샘플이 바코딩된 순서에 대한 정보를 제공한다. 예를 들어, 세포의 개체군은 세포 사이클의 G0 상에서 바코딩될 수 있다. 세포는 세포 사이클의 G1 상에서 바코드(예를 들어, 추계적 바코드)로 다시 펄싱(pulsing)될 수 있다. 세포는 세포 사이클의 S 상 등에서 바코드로 다시 펄싱될 수 있다. 각 펄스(예를 들어, 세포 사이클의 각각의 상)에서 바코드는 상이한 치수 표지를 포함할 수 있다. 이러한 방식으로, 치수 표지는, 세포 사이클의 어떤 상에서 표적이 표지되었는지에 대한 정보를 제공한다. 치수 표지는 많은 다양한 생물학적 시간에서 정보를 얻을 수 있다. 예시적인 생물학적 시간은, 세포 사이클, 전사(예를 들어, 전사 개시), 및 전사 분해를 포함할 수 있지만, 이에 제한되지 않는다. 또 다른 예에서, 샘플(예를 들어, 세포, 세포의 개체군)은 약물 및/또는 요법으로 처리 전 및/또는 처리 후에 표지될 수 있다. 별개의 표적의 복제물 수의 변화는 샘플의 약물 및/또는 요법에 대한 반응을 나타낼 수 있다.

- [0358] 치수 표지는 활성화될 수 있다. 활성화 가능한 치수 표지는 특정 시점에 활성화될 수 있다. 활성화 가능한 표지는, 예를 들어 구성적으로 활성화(예를 들어, 꺼지지 않음)될 수 있다. 활성화 가능한 치수 표지는, 예를 들어 가역적으로 활성화(예를 들어, 활성화 가능한 치수 표지는 켜지고 꺼질 수 있음)될 수 있다. 치수 표지는, 예를 들어 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10배, 또는 그 이상으로 가역적으로 활성화될 수 있다. 치수 표지는, 예를 들어 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10배, 또는 그 이상으로 가역적으로 활성화될 수 있다. 일부 구현예에서, 치수 표지는 형광, 빛, 화학적 이벤트(예를 들어, 절단, 또 다른 분자의 결합, 변형(예를 들어, 폐기화, 수모일화(sumoylated), 아세틸화, 메틸화, 디아세틸화, 디메틸화)의 추가), 광화학적 이벤트(예를 들어, 광 케이징(photocaging)), 및 비-천연 뉴클레오타이드의 도입으로 활성화될 수 있다.
- [0359] 치수 표지는, 일부 구현예에서, 소정의 고체 지지체(예를 들어, 비드)에 부착된 모든 바코드(예를 들어, 추계적 바코드)에 대해 동일할 수 있지만, 상이한 고체 지지체(예를 들어, 비드)에 대해 상이할 수 있다. 일부 구현예에서, 동일한 고체 지지체 상에서 적어도 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99% 또는 100%의 바코드는 동일한 치수 표지를 포함할 수 있다. 일부 구현예에서, 동일한 고체 지지체 상에서 적어도 60%의 바코드는 동일한 치수 표지를 포함할 수 있다. 일부 구현예에서, 동일한 고체 지지체 상에서 적어도 95%의 바코드는 동일한 치수 표지를 포함할 수 있다.
- [0360] 복수의 고체 지지체(예를 들어, 비드)에 나타난 고유 치수 표지 서열이 10^6 이상만큼 많이 존재할 수 있다. 치수 표지는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 치수 표지는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오타이드 길이일 수 있다. 치수 표지는 약 5 내지 약 200개의 뉴클레오타이드를 포함할 수 있다. 치수 표지는 약 10 내지 약 150개의 뉴클레오타이드를 포함할 수 있다. 치수 표지는 약 20 내지 약 125개의 뉴클레오타이드를 포함할 수 있다.
- [0361] 공간 표지
- [0362] 바코드는 하나 이상의 공간 표지를 포함할 수 있다. 일부 구현예에서, 공간 표지는 바코드와 회합된 표적 분자의 공간 배향에 대한 정보를 제공하는 핵산 서열을 포함할 수 있다. 공간 표지는 샘플의 좌표와 연결될 수 있다. 좌표는 고정된 좌표일 수 있다. 예를 들어, 좌표는 기재에 관련되어 고정될 수 있다. 공간 표지는 2차원 또는 3차원 그리드에 관련될 수 있다. 좌표는 랜드마크(landmark)에 관련되어 고정될 수 있다. 랜드마크는 공간에서 식별가능할 수 있다. 랜드마크는 이미지화될 수 있는 구조물일 수 있다. 랜드마크는 생물학적 구조물, 예를 들어 해부학적 랜드마크일 수 있다. 랜드마크는 세포 랜드마크, 예를 들어 세포 소기관일 수 있다. 랜드마크는 비-천연 랜드마크, 예컨대 색상 코드, 바 코드, 자성 특성, 형광, 방사능 또는 고유 크기 또는 고유 형태와 같은 식별가능한 식별자를 갖는 구조물일 수 있다. 공간 표지는 물리적 칸막이(예를 들어, 웰, 용기 또는 액적(droplet))와 회합될 수 있다. 일부 구현예에서, 다중 공간 표지가 공간 내 하나 이상의 위치를 인코딩하는 데 함께 사용된다.
- [0363] 공간 표지는 소정의 고체 지지체(예를 들어, 비드)에 부착된 모든 바코드에 대해 동일할 수 있지만, 상이한 고체 지지체(예를 들어, 비드)에 대해 상이할 수 있다. 일부 구현예에서, 동일한 공간 표지를 포함하는 동일한 고체 지지체 상에서 바코드의 백분율은 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 공간 표지를 포함하는 동일한 고체 지지체 상에서 바코드의 백분율은 적어도, 또는 최대 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 또는 100%일 수 있다. 일부 구현예에서, 동일한 고체 지지체 상에서 적어도 60%의 바코드는 동일한 공간 표지를 포함할 수 있다. 일부 구현예에서, 동일한 고체 지지체 상에서 적어도 95%의 바코드는 동일한 공간 표지를 포함할 수 있다.
- [0364] 복수의 고체 지지체(예를 들어, 비드)에 나타난 고유한 공간 표지 서열이 10^6 이상만큼 많이 존재할 수 있다. 공간 표지는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 공간 표지는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오타이드 길이일 수 있다. 공간 표지는 약 5 내지 약 200개의 뉴클레오타이드를 포함할 수 있다. 공간 표지는 약 10 내지 약 150개

의 뉴클레오티드를 포함할 수 있다. 공간 표지는 약 20 내지 약 125개의 뉴클레오티드 길이를 포함할 수 있다.

[0365] 세포 표지

[0366] 바코드(예를 들어, 추계적 바코드)는 하나 이상의 세포 표지를 포함할 수 있다. 일부 구현예에서, 세포 표지는 어떤 표적 핵산이 어떤 세포로부터 기원되었는지를 결정하기 위한 정보를 제공하는 핵산 서열을 포함할 수 있다. 일부 구현예에서, 세포 표지는 소정의 고체 지지체(예를 들어, 비드)에 부착된 모든 바코드에 대해 동일할 수 있지만, 상이한 고체 지지체(예를 들어, 비드)에 대해 상이할 수 있다. 일부 구현예에서, 동일한 세포 표지를 포함하는 동일한 고체 지지체 상에서 바코드의 백분율은 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 세포 표지를 포함하는 동일한 고체 지지체 상에서 바코드의 백분율은 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 또는 100%, 또는 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 또는 100%일 수 있다. 예를 들어, 동일한 고체 지지체 상에서 적어도 60%의 바코드는 동일한 세포 표지를 포함할 수 있다. 또 다른 예로서, 동일한 고체 지지체 상에서 적어도 95%의 바코드는 동일한 세포 표지를 포함할 수 있다.

[0367] 복수의 고체 지지체(예를 들어, 비드)에 나타낸 고유 세포 표지 서열이 10^6 이상만큼 많이 존재할 수 있다. 세포 표지는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오티드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오티드 길이일 수 있다. 세포 표지는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오티드 길이일 수 있다. 예를 들어, 세포 표지는 약 5 내지 약 200개의 뉴클레오티드를 포함할 수 있다. 세포 표지는 약 10 내지 약 150개의 뉴클레오티드를 포함할 수 있다. 또 다른 예로서, 세포 표지는 약 20 내지 약 125개의 뉴클레오티드 길이를 포함할 수 있다.

[0368] 바코드 서열

[0369] 바코드는 하나 이상의 바코드 서열을 포함할 수 있다. 일부 구현예에서, 바코드 서열은 바코드에 혼성화된 특정 유형의 표적 핵산 중에 대한 식별 정보를 제공하는 핵산 서열을 포함할 수 있다. 바코드 서열은 바코드(예를 들어, 표적-결합 영역)에 혼성화된 표적 핵산 중의 특정 발생에 대한 계수기를 제공하는(예를 들어, 대략적 근사치를 제공하는) 핵산 서열을 포함할 수 있다.

[0370] 일부 구현예에서, 바코드 서열의 다양한 세트는 소정의 고체 지지체(예를 들어, 비드)에 부착된다. 일부 구현예에서, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 고유 분자 표지 서열이 존재할 수 있거나, 약 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 고유 분자 표지 서열이 존재할 수 있다. 예를 들어, 복수의 바코드는 별개의 서열을 갖는 약 6561개의 바코드 서열을 포함할 수 있다. 또 다른 예로서, 복수의 바코드는 별개의 서열을 갖는 약 65536개의 바코드 서열을 포함할 수 있다. 일부 구현예에서, 적어도, 또는 최대 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 또는 10^9 개의 고유 바코드 서열이 존재할 수 있다. 고유 분자 표지 서열은 소정의 고체 지지체(예를 들어, 비드)에 부착될 수 있다.

[0371] 바코드의 길이는 상이한 실시에서 상이할 수 있다. 예를 들어, 바코드는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오티드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오티드 길이일 수 있다. 또 다른 예로서, 바코드는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오티드 길이일 수 있다.

[0372] 분자 표지

[0373] 바코드(예를 들어, 추계적 바코드)는 하나 이상의 분자 표지를 포함할 수 있다. 분자 표지는 바코드 서열을 포함할 수 있다. 일부 구현예에서, 분자 표지는 바코드에 혼성화된 특정 유형의 표적 핵산 중에 대한 식별 정보를 제공하는 핵산 서열을 포함할 수 있다. 분자 표지는 바코드(예를 들어, 표적-결합 영역)에 혼성화된 표적 핵산 중의 특정 발생에 대한 계수기를 제공하는 핵산 서열을 포함할 수 있다.

- [0374] 일부 구현예에서, 분자 표지의 다양한 세트는 소정의 고체 지지체(예를 들어, 비드)에 부착된다. 일부 구현예에서, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 값 중 임의의 두 값 사이의 어떤 수 또는 범위인 고유 분자 표지 서열이 존재할 수 있거나, 약 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 값 중 임의의 두 값 사이의 어떤 수 또는 범위인 고유 분자 표지 서열이 존재할 수 있다. 예를 들어, 복수의 바코드는 별개의 서열을 갖는 약 6561개의 분자 표지를 포함할 수 있다. 또 다른 예로서, 복수의 바코드는 별개의 서열을 갖는 약 65536개의 분자 표지를 포함할 수 있다. 일부 구현예에서, 적어도, 또는 최대 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 또는 10^9 개의 고유 분자 표지 서열이 존재할 수 있다. 고유 분자 표지 서열을 갖는 바코드는 소정의 고체 지지체(예를 들어, 비드)에 부착될 수 있다.
- [0375] 복수의 추계적 바코드를 사용하여 추계적으로 바코딩하기 위해, 상이한 분자 표지 서열 수와 표적 중 임의의 것의 출현 수의 비는 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 표적은 동일하거나 거의 동일한 서열을 갖는 mRNA 분자를 포함하는 mRNA 종일 수 있다. 일부 구현예에서, 상이한 분자 표지 서열의 수와 복수의 표적 중 임의의 것의 출현 수의 비는 적어도, 또는 최대 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 또는 100:1이다.
- [0376] 분자 표지는 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 분자 표지는 적어도, 또는 최대 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, 또는 300개의 뉴클레오타이드 길이일 수 있다.
- [0377] 표적-결합 영역
- [0378] 바코드는 하나 이상의 표적-결합 영역, 예컨대 포획 프로브를 포함할 수 있다. 일부 구현예에서, 표적-결합 영역은 관심 대상의 표적과 혼성화할 수 있다. 일부 구현예에서, 표적-결합 영역은 표적(예를 들어, 표적 핵산, 표적 분자, 예를 들어 분석될 세포 핵산)에 특이적으로 혼성화하는 핵산 서열, 예를 들어 특정 유전자 서열을 포함할 수 있다. 일부 구현예에서, 표적-결합 영역은 특정 표적 핵산의 특정 위치에 부착(예를 들어, 혼성화)할 수 있는 핵산 서열을 포함할 수 있다. 일부 구현예에서, 표적-결합 영역은 제한 효소 부위 오버행(overhang)(예를 들어, EcoRI 점착성-말단(sticky-end) 오버행)에 특이적 혼성화가 가능한 핵산 서열을 포함할 수 있다. 이어서, 바코드는 제한 부위 오버행에 상보적인 서열을 포함하는 임의의 핵산 분자에 결합될 수 있다.
- [0379] 일부 구현예에서, 표적-결합 영역은 비특이적 표적 핵산 서열을 포함할 수 있다. 비특이적 표적 핵산 서열은 표적 핵산의 특정 서열에 독립적인, 다중 표적 핵산에 결합할 수 있는 서열을 지칭할 수 있다. 예를 들어, 표적-결합 영역은 랜덤 다량체 서열, 또는 mRNA 분자 상에서 폴리(A) 테일에 혼성화하는 올리고(dT) 서열을 포함할 수 있다. 랜덤 다량체 서열은, 예를 들어 랜덤 2량체, 3량체, 4량체, 5량체, 6량체, 7량체, 8량체, 9량체, 10량체 또는 임의의 길이의 상급 다량체 서열일 수 있다. 일부 구현예에서, 표적-결합 영역은 소정의 비드에 부착된 모든 바코드에 대해 동일하다. 일부 구현예에서, 소정의 비드에 부착된 복수의 바코드에 대한 표적-결합 영역은 2개 이상의 상이한 표적-결합 서열을 포함할 수 있다. 표적-결합 영역은 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 표적-결합 영역은 최대 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 그 이상의 뉴클레오타이드 길이일 수 있다.
- [0380] 일부 구현예에서, 표적-결합 영역은 폴리아데닐화 말단을 포함하는 mRNA와 혼성화할 수 있는 올리고(dT)를 포함할 수 있다. 표적-결합 영역은 유전자-특이적일 수 있다. 예를 들어, 표적-결합 영역은 표적의 특정 영역에 혼성화하도록 구성될 수 있다. 표적-결합 영역은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위

의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 표적-결합 영역은 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 또는 30개의 뉴클레오타이드 길이일 수 있다. 표적-결합 영역은 약 5 내지 30개의 뉴클레오타이드 길이일 수 있다. 바코드가 유전자 특이적 표적-결합 영역을 포함하는 경우, 바코드는 본 명세서에서 유전자-특이적 바코드로 지칭될 수 있다.

[0381] 배향 특성

[0382] 추계적 바코드(예를 들어, 추계적 바코드)는, 바코드를 배향(예를 들어, 정렬(align))하는 데 사용될 수 있는 하나 이상의 배향 특성을 포함할 수 있다. 바코드는 등전점 전기영동(isoelectric focusing)을 위한 모이어티를 포함할 수 있다. 상이한 바코드는 상이한 등전점 전기영동 지점을 포함할 수 있다. 이들 바코드가 샘플로 도입될 때, 샘플은 바코드를 공지된 방식으로 배향시키기 위하여 등전점 전기영동시킬 수 있다. 이러한 방식으로, 배향 특성은 샘플 내 바코드의 공지된 지도를 개발하는 데 사용될 수 있다. 예시적인 배향 특성은 전기영동 이동도(예를 들어, 바코드의 크기에 기초하여), 등전점, 스핀(spin), 전도성 및/또는 자기조립을 포함할 수 있다. 예를 들어, 자기조립의 배향 특성을 갖는 바코드는 활성화시 특정 배향(예를 들어, 핵산 나노구조)으로 자기조립할 수 있다.

[0383] 친화도 특성

[0384] 바코드(예를 들어, 추계적 바코드)는 하나 이상의 친화도 특성을 포함할 수 있다. 예를 들어, 공간 표지는 친화도 특성을 포함할 수 있다. 친화도 특성은 또 다른 실체(entity)(예를 들어, 세포 수용체)로의 바코드의 결합을 용이하게 할 수 있는 화학 모이어티 및/또는 생물학적 모이어티를 포함할 수 있다. 예를 들어, 친화도 특성은 항체, 예를 들어, 샘플 상에서 특정 모이어티(예를 들어, 수용체)에 특이적인 항체를 포함할 수 있다. 일부 구현예에서, 항체는 바코드를 특정 세포 유형 또는 분자로 안내(guide)할 수 있다. 특정 세포 유형 또는 분자에서 및/또는 그 가까이에서 표적은 표지화(예를 들어, 추계적으로 표지화)될 수 있다. 친화도 특성은, 일부 구현예에서, 공간 표지의 뉴클레오타이드 서열에 추가하여 공간 정보를 제공하는데, 이는 항체가 바코드를 특정 위치로 안내할 수 있기 때문이다. 항체는 치료용 항체, 예를 들어, 단클론성 항체 또는 다클론성 항체일 수 있다. 항체는 인간화되거나 키메라일 수 있다. 항체는 네이키드(naked) 항체 또는 융합 항체일 수 있다.

[0385] 항체는 전장(즉, 자연적으로 발생하거나 정상의 면역글로불린 유전자 단편 재조합 과정에 의해 형성됨) 면역글로불린 분자(예를 들어, IgG 항체) 또는 면역글로불린 분자의 면역학적으로 활성인(즉, 특이적으로 결합하는) 부분, 예컨대 항체 단편일 수 있다.

[0386] 항체 단편은, 예를 들어 항체의 일부분, 예컨대 F(ab')₂, Fab', Fab, Fv, sFv 등일 수 있다. 일부 구현예에서, 항체 단편은 전장 항체에 의해 인식되는 동일한 항원과 결합할 수 있다. 항체 단편은 항체의 가변 영역으로 이루어진 단리된 단편, 예컨대 중쇄 및 경쇄의 가변 영역으로 이루어진 "Fv" 단편, 및 경쇄 및 중쇄 가변 영역이 펩티드 링커에 의해 연결된 재조합 단일 사슬 폴리펩티드 분자("scFv 단백질")를 포함할 수 있다. 예시적인 항체는 암세포에 대한 항체, 바이러스에 대한 항체, 세포 표면 수용체(CD8, CD34, CD45)에 결합하는 항체, 및 치료제 항체를 포함할 수 있으며, 이에 제한되지는 않는다.

[0387] 범용 어댑터 프라이머

[0388] 바코드는 하나 이상의 범용 어댑터 프라이머를 포함할 수 있다. 예를 들어, 유전자-특이적 바코드, 예컨대 유전자-특이적 추계적 바코드는 범용 어댑터 프라이머를 포함할 수 있다. 범용 어댑터 프라이머는 모든 바코드에 걸쳐 범용인 뉴클레오타이드 서열을 지칭할 수 있다. 범용 어댑터 프라이머는 유전자-특이적 바코드를 구축하는 데 사용될 수 있다. 범용 어댑터 프라이머는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 범용 어댑터 프라이머는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 또는 30개의 뉴클레오타이드 길이일 수 있다. 범용 어댑터 프라이머는 5 내지 30개의 뉴클레오타이드 길이일 수 있다.

[0389] 링커

[0390] 바코드가 하나 초과와 유형의 표지(예를 들어, 하나 초과와 세포 표지 또는 하나 초과와 바코드 서열, 예컨대

하나의 분자 표지)를 포함하는 경우, 표지는 링커 표지 서열과 함께 산재될 수 있다. 링커 표지 서열은 적어도 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개 또는 그 이상의 뉴클레오타이드 길이일 수 있다. 링커 표지 서열은 최대 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50개, 또는 그 이상의 뉴클레오타이드 길이일 수 있다. 일부 경우에서, 링커 표지 서열은 12개의 뉴클레오타이드 길이이다. 링커 표지 서열은 바코드의 합성을 촉진하는 데 사용될 수 있다. 링커 표지는 오류-정정(예를 들어, 해밍) 코드를 포함할 수 있다.

[0391] 고체 지지체

[0392] 본 명세서에 개시된 바코드, 예컨대 추계적 바코드는, 일부 구현예에서 고체 지지체와 회합될 수 있다. 고체 지지체는, 예를 들어 합성 입자일 수 있다. 일부 구현예에서, 바코드 서열은, 예컨대 고체 지지체 상에서 복수의 바코드(예를 들어, 제1 복수의 바코드)의 추계적 바코드(예를 들어, 제1 바코드 서열)에 대한 분자 표지의 일부 또는 전부는 적어도 하나의 뉴클레오타이드만큼 상이하다. 동일한 고체 지지체 상에서 바코드의 세포 표지는 동일할 수 있다. 상이한 고체 지지체 상에서 바코드의 세포 표지는 적어도 하나의 뉴클레오타이드만큼 상이할 수 있다. 예를 들어, 제1 고체 지지체 상에서 제1 복수의 바코드의 제1 세포 표지는 동일한 서열을 가질 수 있고, 제2 고체 지지체 상에서 제2 복수의 바코드의 제2 세포 표지는 동일한 서열을 가질 수 있다. 제1 고체 지지체 상에서 제1 복수의 바코드의 제1 세포 표지 및 제2 고체 지지체 상에서 제2 복수의 바코드의 제2 세포 표지는 적어도 하나의 뉴클레오타이드만큼 상이할 수 있다. 세포 표지는, 예를 들어 약 5 내지 20개의 뉴클레오타이드 길이일 수 있다. 바코드 서열은, 예를 들어 약 5 내지 20개의 뉴클레오타이드 길이일 수 있다. 합성 입자는, 예를 들어 비드일 수 있다.

[0393] 비드는, 예를 들어 실리카 겔 비드, 제어된 기공 유리 비드, 자성 비드, 다이나비드(Dynabead), 세파텍스/세파로스 비드, 셀룰로스 비드, 폴리스티렌 비드, 또는 이들의 임의의 조합일 수 있다. 비드는 재료, 예컨대 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합을 포함할 수 있다.

[0394] 일부 구현예에서, 비드는, 바코드 또는 추계적 바코드로 작용화된 중합체 비드, 예를 들어 변형가능한 비드 또는 겔 비드(예컨대, 10X Genomics(미국 캘리포니아주 샌프란시스코 소재)의 겔 비드)일 수 있다. 일부 실시예에서, 겔 비드는 중합체 기반 겔을 포함할 수 있다. 겔 비드는, 예를 들어 하나 이상의 중합체 전구체를 액적 내로 캡슐화함으로써 생성될 수 있다. 중합체 전구체를 가속화제(예를 들어, 테트라메틸에틸렌디아민(TEMED))에 노출 시, 겔 비드가 생성될 수 있다.

[0395] 일부 구현예에서, 입자는 분해성일 수 있다. 예를 들어, 중합체 비드는 원하는 조건 하에서 용해, 용융 또는 분해될 수 있다. 원하는 조건은 환경 조건을 포함할 수 있다. 원하는 조건은 제어된 방식으로 중합체 비드를 용해, 용융 또는 분해시킬 수 있다. 겔 비드는 화학적 자극, 물리적 자극, 생물학적 자극, 열적 자극, 자기 자극, 전기 자극, 광 자극, 또는 이들의 임의의 조합으로 인해 용해, 용융, 또는 분해될 수 있다.

[0396] 피분석물 및/또는 시약, 예컨대 올리고뉴클레오타이드 바코드는, 겔 비드의 내부 표면(예를 들어, 올리고뉴클레오타이드 바코드의 확산을 통해 접근가능한 내부 및/또는 올리고뉴클레오타이드 바코드의 생성에 사용된 재료) 및/또는 본 명세서에 기재된 겔 비드 또는 임의의 기타 다른 마이크로캡슐의 외측 표면에, 예를 들어 커플링/고정화될 수 있다. 커플링/고정화는 임의의 형태의 화학적 결합(예를 들어, 공유결합, 이온결합) 또는 물리적 현상(예를 들어, 반데르 발스 힘, 쌍극자-쌍극자 상호작용 등)을 통해 이루어질 수 있다. 일부 구현예에서, 본 명세서에 기재된 겔 비드 또는 임의의 기타 다른 마이크로캡슐로의 시약의 커플링/고정화는, 예를 들어 불안정한 모이어티를 통해(예를 들어, 본 명세서에 기재된 화학 가교결합체를 포함한, 화학 가교결합체를 통해) 가역적일 수 있다. 자극의 적용 시, 불안정한 모이어티는 절단되고 고정화된 시약은 자유롭게 될 수 있다. 일부 구현예에서, 불안정한 모이어티는 이황화 결합이다. 예를 들어, 올리고뉴클레오타이드 바코드가 이황화 결합을 통해 겔 비드에 고정화된 경우, 이황화 결합이 환원제에 노출되면 이황화 결합은 절단되고 올리고뉴클레오타이드 바코드는 비드로부터 유리될 수 있다. 불안정한 모이어티는 겔 비드 또는 마이크로캡슐의 부분으로서, 시약 또는 피분석물을 겔 비드 또는 마이크로캡슐에 연결하는 화학 링커의 부분으로서, 그리고/또는 시약 또는 피분석물의 부분으로서 포함될 수 있다. 일부 구현예에서, 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다.

[0397] 일부 구현예에서, 겔 비드는 중합체, 열 민감성 중합체, 광 민감성 중합체, 자성 중합체, pH 민감성 중합체, 염-민감성 중합체, 화학 민감성 중합체, 고분자 전해질, 폴리사카라이드, 펩티드, 단백질 및/또는 플라스틱을 포

함할 수 있지만, 이에 제한되지 않는 넓은 범위의 상이한 중합체를 포함할 수 있다. 중합체는 재료, 예컨대 폴리(N-이소프로필아크릴아미드)(PNIPAAm), 폴리(스티렌 설포네이트)(PSS), 폴리(알릴 아민)(PAAm), 폴리(아크릴산)(PAA), 폴리(에틸렌 이민)(PEI), 폴리(디알릴디메틸-암모늄 클로라이드)(PDADMAC), 폴리(피롤)(PPy), 폴리(비닐피롤리돈)(PVPON), 폴리(비닐 피리딘)(PVP), 폴리(메트아크릴산)(PMAA), 폴리(메틸 메타크릴레이트)(PMMA), 폴리스티렌(PS), 폴리(테트라하이드로푸란)(PTHF), 폴리(프탈알데히드)(PTHF), 폴리(헥실 비올로겐)(PHV), 폴리(L-리신)(PLL), 폴리(L-아르기닌)(PARG), 폴리(락틱-코-글리콜산)(PLGA)를 포함할 수 있지만 이에 제한되지는 않는다.

[0398] 비드의 붕괴, 용해 또는 분해를 촉발하기 위해 수많은 화학적 자극이 사용될 수 있다. 이들 화학적 변화의 예는, 비드 벽에 대해 pH-매개된 변화, 가교결합의 화학적 절단을 통한 비드 벽의 붕괴, 비드 벽의 촉발된 해중합(depolymerization), 및 비드 벽 스위칭 반응을 포함할 수 있지만, 이에 제한되지는 않는다. 벌크 변화 또한 비드의 붕괴를 촉발하는 데 사용될 수 있다.

[0399] 다양한 자극을 통한 마이크로캡슐에 대한 벌크 또는 물리적 변화는 또한 시약을 방출하기 위한 캡슐의 설계에 많은 장점을 제공한다. 벌크 또는 물리적 변화는 거시적 규모로 일어나며, 여기에서 비드 파열은 자극에 의해 유도된 기계-물리적 힘의 결과이다. 이들 과정은 압력 유도된 파열, 비드 벽 용융, 또는 비드 벽의 다공성에서의 변화를 포함할 수 있지만, 이에 제한되지는 않는다.

[0400] 생물학적 자극은 또한 비드의 붕괴, 용해 또는 분해의 촉발에 사용될 수 있다. 일반적으로, 생물학적 촉발제는 화학적 촉발제와 비슷하지만, 많은 예들은 생명 시스템에서 흔히 발견되는 생체분자 또는 분자, 예컨대 효소, 펩티드, 당류, 지방산, 핵산 등을 사용한다. 예를 들어, 비드는 특정 프로테아제에 의한 절단에 민감한 펩티드 가교결합을 갖는 중합체를 포함할 수 있다. 더 구체적으로, 일 예는 GFLGK 펩티드 가교결합을 포함하는 마이크로캡슐을 포함할 수 있다. 생물학적 촉발제, 예컨대 프로테아제 카텝신 B의 첨가 시, 껍질(shell)의 펩티드 가교결합은 절단되고, 비드의 내용물이 방출된다. 다른 경우, 프로테아제는 열-활성화될 수 있다. 또 다른 예에서, 비드는 셀룰로스를 포함하는 껍질 벽을 포함한다. 가수분해성 효소 키토산의 첨가는 셀룰로스 결합의 절단, 껍질 벽의 해중합, 및 그 내부의 내용물의 방출에 대한 생물학적 촉발제로서의 역할을 한다.

[0401] 비드는 또한 열 자극의 적용 시 그 내용물을 방출하도록 유도될 수 있다. 온도의 변화는 비드에 다양한 변화를 유발할 수 있다. 열의 변화는 비드를 용융시켜 비드 벽이 붕괴하도록 할 수 있다. 다른 경우, 열은 비드의 내부 성분의 내부 압력을 증가시켜서 비드가 파열되거나 폭발하도록 할 수 있다. 또 다른 경우에, 열은 비드를 줄어든 탈수 상태로 변형시킬 수 있다. 열은 또한 비드의 벽 내부의 열-민감성 중합체 상에 작용하여 비드의 붕괴를 유발할 수 있다.

[0402] 자성 나노입자를 마이크로캡슐의 비드 벽에 포함하는 것은 비드의 촉발된 파열을 가능하게 할 뿐만 아니라, 비드를 어레이로 안내할 수 있다. 본 개시 내용의 디바이스는 어느 목적을 위해 자성 비드를 포함할 수 있다. 일 예에서, Fe_3O_4 나노입자를 다가전해질-함유 비드 내로 혼입하는 것은 진동하는 자기장 자극의 존재 하에서 파열을 촉발시킨다.

[0403] 비드는 또한, 전기 자극의 결과로서 붕괴, 용해 또는 분해될 수 있다. 이전의 섹션에 기재된 자성 입자와 유사하게, 전기적으로 민감성인 비드는 비드의 촉발된 파열 및 기타 다른 기능, 예컨대 전기장, 전기 전도성 또는 산화환원 반응에서의 정렬 모두를 허용할 수 있다. 일 예에서, 전기적으로 민감성인 재료를 함유하는 비드는 전기장에서, 내부 시약의 방출이 제어될 수 있도록 정렬된다. 다른 예에서, 전기장은 다공성을 증가시킬 수 있는 비드 벽 그 자체 내에서 산화환원 반응을 유도할 수 있다.

[0404] 광 자극은 또한 비드를 붕괴하는 데 사용될 수 있다. 수많은 광 촉발제가 가능하며, 다양한 분자, 예컨대 특정 파장 범위의 광자를 흡수할 수 있는 나노 입자 및 발색단을 사용하는 시스템을 포함할 수 있다. 예를 들어, 금속 산화물 코팅이 캡슐 촉발제로 사용될 수 있다. SiO_2 로 코팅된 다가전해질 캡슐의 UV 조사는 비드 벽의 붕괴를 초래할 수 있다. 또 다른 예에서, 광전환성 재료, 예컨대 아조벤젠 기가 비드 벽 내에 혼입될 수 있다. UV 또는 가시 광선의 적용 시, 이들과 같은 화합물들은 광자 흡수 시 가역적인 시스에서 트랜스형으로의 이성질화를 겪는다. 이러한 양태에서, 광자 스위치의 혼입은, 광 촉발제의 적용 시 분해되거나 더 다공성이 될 수 있는 비드 벽을 생성한다.

[0405] 예를 들어, 도 2에 예시된 바코딩(예를 들어, 추계적 바코딩)의 비제한적인 예에서, 단일 세포와 같은 세포를 블록(208)에서 마이크로웰 어레이의 복수의 마이크로웰 상으로 도입시킨 후, 비드는 블록(212)에서 마이크로웰 어레이의 복수의 마이크로웰 상으로 도입될 수 있다. 각각의 마이크로웰은 하나의 비드를 포함할 수 있다. 비드

는 복수의 바코드를 포함할 수 있다. 바코드는 비드에 부착된 5' 아민 영역을 포함할 수 있다. 바코드는 범용 표지, 바코드 서열(예를 들어, 분자 표지), 표적-결합 영역, 또는 이들의 임의의 조합을 포함할 수 있다.

[0406] 본 명세서에 개시된 바코드는 고체 지지체(예를 들어, 비드)와 회합될(예를 들어, 부착될) 수 있다. 고체 지지체와 회합된 바코드는 고유 서열을 갖는 적어도 100개 또는 1000개의 바코드 서열을 포함하는 군으로부터 선택되는 바코드 서열을 각각 포함할 수 있다. 일부 구현예에서, 고체 지지체와 회합된 상이한 바코드는 상이한 서열을 갖는 바코드를 포함할 수 있다. 일부 구현예에서, 고체 지지체와 회합된 바코드의 백분율은 동일한 세포 표지를 포함한다. 예를 들어, 백분율은 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 다른 예로서, 백분율은 적어도, 또는 최대 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 또는 100%일 수 있다. 일부 구현예에서, 고체 지지체와 회합된 바코드는 동일한 세포 표지를 가질 수 있다. 상이한 고체 지지체와 회합된 바코드는 고유 서열을 갖는 적어도 100개 또는 1000개의 세포 표지를 포함하는 군으로부터 선택되는 상이한 세포 표지를 가질 수 있다.

[0407] 본 명세서에 개시된 바코드는 고체 지지체(예를 들어, 비드)에 회합될(예를 들어, 부착될) 수 있다. 일부 구현예에서, 샘플에서 복수의 표적의 바코딩은 복수의 바코드과 회합된 복수의 합성 입자를 포함하는 고체 지지체를 이용하여 수행될 수 있다. 일부 구현예에서, 고체 지지체는 복수의 바코드와 회합된 복수의 합성 입자를 포함할 수 있다. 상이한 고체 지지체 상에서 복수의 바코드의 공간 표지는 적어도 하나의 뉴클레오티드만큼 상이할 수 있다. 고체 지지체는, 예를 들어 2차원 또는 3차원의 복수의 바코드를 포함할 수 있다. 합성 입자는 비드일 수 있다. 비드는 실리카 겔 비드, 제어된 기공 유리 비드, 자성 비드, 다이내비드, 세파텍스/세파로스 비드, 셀룰로스 비드, 폴리스티렌 비드, 또는 이들의 임의의 조합일 수 있다. 고체 지지체는 중합체, 매트릭스, 하이드로겔, 니들 어레이 디바이스, 항체, 또는 이들의 임의의 조합을 포함할 수 있다. 일부 구현예에서, 고체 지지체는 자유 유동할 수 있다. 일부 구현예에서, 고체 지지체는 반고체 또는 고체 어레이 내에 매립될(embedded) 수 있다. 바코드는 고체 지지체와 회합되지 않을 수 있다. 바코드는 개별적인 뉴클레오티드일 수 있다. 바코드는 기재와 회합될 수 있다.

[0408] 본 명세서에 사용되는 바와 같이, 용어 "묶인(tethered)", "부착된", 또는 "고정화된"은 상호교환가능하게 사용되며, 바코드를 고체 지지체에 부착하기 위한 공유결합 또는 비-공유결합 수단을 지칭할 수 있다. 다양한 상이한 고체 지지체 중 임의의 것이 미리 합성된 바코드를 부착하기 위한 또는 바코드의 동일계내(in situ) 고체-상 합성을 위해 사용될 수 있다.

[0409] 일부 구현예에서, 고체 지지체는 비드이다. 비드는 하나 이상의 유형의 고체, 다공성, 또는 중공 스피어, 볼, 베어링, 실린더, 또는 핵산이 (예를 들어, 공유결합적으로 또는 비-공유결합적으로) 고정화될 수 있는 기타 다른 유사한 구성을 포함할 수 있다. 비드는, 예를 들어 플라스틱, 세라믹, 금속, 중합체 재료, 또는 이들의 임의의 조합으로 구성될 수 있다. 비드는 구형(예를 들어, 미소구체)이거나 비-구형 또는 불규칙한 형태, 예컨대 입방체, 직육면체, 피라미드형, 원통형, 원뿔형, 장타원형, 또는 원반형태 등을 가지는 별개의 입자일 수 있거나, 이를 포함할 수 있다. 일부 구현예에서, 비드는 비-구형 형태일 수 있다.

[0410] 비드는 상자성체 재료(예를 들어, 마그네슘, 몰리브덴, 리튬 및 탄탈럼), 초상자성체 재료(예를 들어, 페라이트(Fe_3O_4 ; 자철석) 나노입자), 강자성 재료(예를 들어, 철, 니켈, 코발트, 이들의 일부 합금, 및 일부 희토류 금속 화합물), 세라믹, 플라스틱, 유리, 폴리스티렌, 실리카, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 아가로스, 하이드로겔, 중합체, 셀룰로스, 나일론, 또는 이들의 임의의 조합을 포함하지만, 이에 제한되지는 않는 다양한 재료를 포함할 수 있다.

[0411] 일부 구현예에서, 비드(예를 들어, 표지가 부착된 비드)는 하이드로겔 비드이다. 일부 구현예에서, 비드는 하이드로겔을 포함한다.

[0412] 본 명세서에 개시된 일부 구현예는 하나 이상의 입자(예를 들어, 비드)를 포함한다. 각각의 입자는 복수의 올리고뉴클레오티드(예를 들어, 바코드)를 포함할 수 있다. 복수의 올리고뉴클레오티드 각각은 바코드 서열(예를 들어, 분자 표지 서열), 세포 표지, 및 표적-결합 영역(예를 들어, 올리고(dT) 서열, 유전자-특이적 서열, 랜덤 다량체, 또는 이들의 조합)을 포함할 수 있다. 복수의 올리고뉴클레오티드 각각의 세포 표지 서열은 동일할 수 있다. 상이한 입자 상의 올리고뉴클레오티드의 세포 표지 서열은, 상이한 입자 상의 올리고뉴클레오티드가 식별될 수 있도록 상이할 수 있다. 상이한 세포 표지 서열의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 세포 표지 서열의 수는 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000,

10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 그 이상일 수 있거나, 약 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 그 이상일 수 있다. 일부 구현예에서, 세포 표지 서열의 수는 적어도, 또는 최대 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 또는 10^9 개일 수 있다. 일부 구현예에서, 복수의 입자 중 단지 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 그 이상은 동일한 세포 서열을 갖는 올리고뉴클레오타이드를 포함한다. 일부 구현예에서, 동일한 세포 서열을 갖는 올리고뉴클레오타이드를 포함하는 복수의 입자는 최대 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 또는 그 이상일 수 있다. 일부 구현예에서, 복수의 입자 중 어느 것도 동일한 세포 표지 서열을 갖지 않는다.

[0413] 각 입자 상의 복수의 올리고뉴클레오타이드는 상이한 바코드 서열(예를 들어, 분자 표지)을 포함할 수 있다. 일부 구현예에서, 바코드 서열의 수는 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 바코드 서열의 수는 적어도, 또는 최대 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 또는 10^9 개일 수 있다. 예를 들어, 복수의 올리고뉴클레오타이드 중 적어도 100개는 상이한 바코드 서열을 포함한다. 또 다른 예로서, 단일 입자 내에, 복수의 올리고뉴클레오타이드 중 적어도 100, 500, 1000, 5000, 10000, 15000, 20000, 50000개, 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 그 이상은 상이한 바코드 서열을 포함한다. 일부 구현예에는 바코드를 포함하는 복수의 입자를 제공한다. 일부 구현예에서, 표지되는 표적의 출현(또는 복제물 또는 수)와 상이한 바코드 서열의 비는 적어도 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:30, 1:40, 1:50, 1:60, 1:70, 1:80, 1:90, 또는 그 이상일 수 있다. 일부 구현예에서, 복수의 올리고뉴클레오타이드의 각각은 샘플 표지, 범용 표지, 또는 둘 모두를 추가로 포함한다. 입자는, 예를 들어 나노입자 또는 마이크로입자일 수 있다.

[0414] 비드의 크기는 다양할 수 있다. 예를 들어, 비드의 직경은 0.1 마이크로미터 내지 50 마이크로미터의 범위일 수 있다. 일부 구현예에서, 비드의 직경은 0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 마이크로미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 마이크로미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다.

[0415] 비드의 직경은 기재의 웰의 직경에 관련될 수 있다. 일부 구현예에서, 비드의 직경은 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 웰의 직경보다 더 길거나 더 짧을 수 있거나, 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 웰의 직경보다 더 길거나 더 짧을 수 있다. 비드의 직경은 세포(예를 들어, 기재의 웰에 의해 포획된 단일 세포)의 직경에 관련될 수 있다. 일부 구현예에서, 비드의 직경은 적어도, 또는 최대 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 100%만큼 웰의 직경보다 더 길거나 짧을 수 있다. 비드의 직경은 세포(예를 들어, 기재의 웰에 의해 포획된 단일 세포)의 직경에 관련될 수 있다. 일부 구현예에서, 비드의 직경은 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 300%만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 세포의 직경보다 더 길거나 더 짧을 수 있거나, 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 300%만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 세포의 직경보다 더 길거나 더 짧을 수 있다. 일부 구현예에서, 비드의 직경은 적어도, 또는 최대 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 또는 300%만큼 세포의 직경보다 더 길거나 짧을 수 있다.

[0416] 비드는 기재에 부착되고/부착되거나 기재 내에 매립될 수 있다. 비드는 겔, 하이드로겔, 중합체 및/또는 매트릭

스에 부착되고/부착되거나 그 안에 매립될 수 있다. 기재(예를 들어, 겔, 매트릭스, 스캐폴드, 또는 중합체) 내 비드의 공간 위치는 위치 주소로서의 역할을 할 수 있는 비드 상의 바코드 상에 존재하는 공간 표지를 사용하여 식별될 수 있다.

[0417] 비드의 예는 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 다이나비드[®], MACS[®] 마이크로비드, 항체 접합된 비드(예를 들어, 항-면역글로불린 마이크로비드), 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 및 BcMag[™] 카르복실레이트-말단화 자성 비드를 포함할 수 있지만 이에 제한되지 않는다.

[0418] 비드는 양자점(quantum dot)또는 형광 염료와 회합되어(예를 들어, 이들로 함침되어) 하나의 형광 광학 채널 또는 다수의 광학 채널에서 비드를 형광으로 만들 수 있다. 비드는 산화 철 또는 산화 크롬과 회합되어 비드를 상자성 또는 강자성으로 만들 수 있다. 비드는 식별가능할 수 있다. 예를 들어, 비드는 카메라를 사용하여 이미지화될 수 있다. 비드는 비드와 회합된 검출가능한 코드를 가질 수 있다. 예를 들어, 비드는 바코드를 포함할 수 있다. 예를 들어, 유기 또는 무기 용액 내에서의 팽창으로 인해, 비드는 크기가 변할 수 있다. 비드는 소수성일 수 있다. 비드는 친수성일 수 있다. 비드는 생체적합성일 수 있다.

[0419] 고체 지지체(예를 들어, 비드)는 시각화될 수 있다. 고체 지지체는 시각화 태그(예를 들어, 형광 염료)를 포함할 수 있다. 고체 지지체(예를 들어, 비드)는 식별자(예를 들어, 숫자)로 애칭될 수 있다. 식별자는 비드를 이미지화함으로써 시각화될 수 있다.

[0420] 고체 지지체는 불용성, 반-수용성, 또는 불용성 재료를 포함할 수 있다. 고체 지지체가 링커, 스캐폴드, 빌딩 블록, 또는 그에 부착된 기타 다른 반응성 모이어티를 포함하는 경우, "작용화된" 것으로 지칭될 수 있는 반면, 고체 지지체가 그에 부착된 그러한 반응성 모이어티가 없는 경우 고체 지지체는 "비작용화될" 수 있다. 고체 지지체는, 예컨대 마이크로적정 웰 포맷으로; 플로-스루(flow-through) 포맷으로, 예컨대 컬럼으로; 또는 딥스틱(dipstick)으로, 용액 중에서 자유롭게 이용될 수 있다.

[0421] 고체 지지체는 막, 종이, 플라스틱, 코팅된 표면, 평면, 유리, 슬라이드, 칩(chip) 또는 이들의 임의의 조합을 포함할 수 있다. 고체 지지체는 수지, 겔, 미소구체, 또는 기타 다른 기하학적 형태의 형태를 취할 수 있다. 고체 지지체는 실리카 칩, 마이크로입자, 나노입자, 플레이트, 어레이, 모세관, 평면 지지체, 예컨대 유리 섬유 필터, 유리 표면, 금속 표면(강철, 금은, 알루미늄, 규소 및 구리), 유리 지지체, 플라스틱 지지체, 규소 지지체, 칩, 필터, 막, 마이크로웰 플레이트, 슬라이드, 다중웰 플레이트 또는 막(예를 들어, 폴리에틸렌, 폴리프로필렌, 폴리아미드, 폴리비닐리덴디플루오라이드로 형성됨), 및/또는 웨이퍼, 빗, 핀 또는 바늘(예를 들어, 조합 합성 또는 분석에 적합한 핀들의 어레이), 또는 웨이퍼(예를 들어, 규소 웨이퍼), 필터 바닥이 있거나 없는 피트(pit)를 가진 웨이퍼와 같은 평면의 피트 또는 나노리터 웰의 어레이 내 비드를 포함하는 플라스틱 재료를 포함할 수 있다.

[0422] 고체 지지체는 중합체 매트릭스(예를 들어, 겔, 하이드로겔)를 포함할 수 있다. 중합체 매트릭스는 세포내 공간(예를 들어, 기관 주위)을 침투할 수 있다. 중합체 매트릭스는 순환계를 통해 펌프될 수 있다.

[0423] 기재 및 마이크로웰 어레이

[0424] 본 명세서에 사용되는 바와 같이, 기재는 고체 지지체의 유형을 지칭할 수 있다. 기재는 본 개시 내용의 바코드 또는 추계적 바코드를 포함할 수 있는 고체 지지체를 지칭할 수 있다. 기재는, 예를 들어 복수의 마이크로웰을 포함한다. 예를 들어, 기재는 2개 이상의 마이크로웰을 포함하는 웰 어레이일 수 있다. 일부 구현예에서, 마이크로웰은 정의된 부피의 작은 반응 챔버를 포함할 수 있다. 일부 구현예에서, 마이크로웰은 하나 이상의 세포를 포획할 수 있다. 일부 구현예에서, 마이크로웰은 하나의 세포만을 포획할 수 있다. 일부 구현예에서, 마이크로웰은 하나 이상의 고체 지지체를 포획할 수 있다. 일부 구현예에서, 마이크로웰은 하나의 고체 지지체만을 포획할 수 있다. 일부 구현예에서, 마이크로웰은 단일 세포 및 단일 고체 지지체(예를 들어, 비드)를 포획한다. 마이크로웰은 본 개시 내용의 바코드 시약을 포함할 수 있다.

[0425] 바코딩 방법

[0426] 본 개시 내용은 물리적 샘플(예를 들어, 조직, 기관, 종양, 세포) 내 별개의 위치에 별개의 표적의 수를 추산하기 위한 방법을 제공한다. 방법은 바코드(예를 들어, 추계적 바코드)를 샘플과 근접하여 위치시키는 단계, 샘플을 용해시키는 단계, 별개의 표적을 바코드와 회합하는 단계, 표적을 증폭하는 단계 및/또는 표적을 디지털 계

수하는 단계를 포함할 수 있다. 방법은 바코드 상에서 공간 표지로부터 취득된 정보를 분석하고/분석하거나 시각화하는 단계를 추가로 포함할 수 있다. 일부 구현예에서, 방법은 샘플 내 복수의 표적을 시각화하는 단계를 포함한다. 샘플의 맵 상에 복수의 표적을 매핑하는 단계는 샘플의 2차원 맵 또는 3차원 맵을 생성하는 단계를 포함할 수 있다. 2차원 맵 및 3차원 맵은 샘플 내 복수의 표적을 바코딩(예를 들어, 추계적 바코딩)하기 전 또는 후에 생성될 수 있다. 샘플 내 복수의 표적을 시각화하는 단계는 샘플의 맵 상에 복수의 표적을 매핑하는 단계를 포함할 수 있다. 샘플의 맵 상에 복수의 표적을 매핑하는 단계는 샘플의 2차원 맵 또는 3차원 맵을 생성하는 단계를 포함할 수 있다. 2차원 맵 및 3차원 맵은 샘플 내 복수의 표적을 바코딩하기 전 또는 후에 생성될 수 있다. 일부 구현예에서, 2차원 맵 및 3차원 맵은 샘플을 용해시키기 전 또는 후에 생성될 수 있다. 2차원 맵 또는 3차원 맵의 생성 전 또는 후에 샘플을 용해시키는 단계는 샘플을 가열하는 단계, 샘플을 세제와 접촉시키는 단계, 샘플의 pH를 변화시키는 단계, 또는 이들의 임의의 조합을 포함할 수 있다.

[0427] 일부 구현예에서, 복수의 표적을 바코딩하는 단계는 복수의 바코드를 복수의 표적과 혼성화하여 바코딩된 표적(예를 들어, 추계적으로 바코딩된 표적)을 생성하는 단계를 포함한다. 복수의 표적을 바코딩하는 단계는 바코딩된 표적의 인덱싱된 라이브러리를 생성하는 단계를 포함할 수 있다. 바코딩된 표적의 인덱싱된 라이브러리를 생성하는 단계는 복수의 바코드(예를 들어, 추계적 바코드)를 포함하는 고체 지지체를 이용하여 수행될 수 있다.

[0428] 샘플과 바코드의 접촉

[0429] 본 개시 내용은 샘플(예를 들어, 세포)을 본 개시 내용의 기재와 접촉시키는 방법을 제공한다. 예를 들어, 세포, 기관, 또는 조직 박편을 포함하는 샘플은 바코드(예를 들어, 추계적 바코드)에 접촉될 수 있다. 세포는, 예를 들어, 중력 흐름에 의해 접촉될 수 있으며, 세포가 침전되어 단일 층을 생성할 수 있다. 샘플은 조직 박편일 수 있다. 박편은 기재 상에 위치될 수 있다. 샘플은 1차원(예를 들어, 편평한 표면을 형성함)일 수 있다. 샘플(예를 들어, 세포)은, 예를 들어 기재 상에서 세포를 성장/배양시킴으로써 기재에 걸쳐 퍼질 수 있다.

[0430] 바코드가 표적과 근접하여 있는 경우, 표적은 바코드에 혼성화될 수 있다. 바코드는 각각의 별개의 표적이 본 개시 내용의 별개의 바코드와 회합될 수 있도록 비-고갈성 비로 접촉될 수 있다. 표적과 바코드 사이의 효율적인 회합을 보장하기 위하여, 표적은 바코드에 가교결합될 수 있다.

[0431] 세포 용해

[0432] 세포 및 바코드의 분포 후, 세포는 용해되어 표적 분자를 자유롭게 할 수 있다. 세포 용해는 다양한 수단 중 임의의 것에 의해, 예를 들어, 화학적 또는 생화학적 수단에 의해, 삼투압 충격에 의해, 또는 열 용해, 기계적 용해 또는 광학적 용해를 통해 달성될 수 있다. 세포는 세제(예를 들어, SDS, Li 도데실 설페이트, 트리톤(Triton) X-100, 트윈(Tween)-20, 또는 NP-40), 유기 용매(예를 들어, 메탄올 또는 아세톤), 또는 소화 효소(예를 들어, 프로테이나제 K, 펩신, 또는 트립신), 또는 이들의 임의의 조합을 포함하는 세포 용해 완충액을 첨가함으로써 용해될 수 있다. 표적 및 바코드의 회합을 증가시키기 위하여, 표적 분자의 확산 속도는, 예를 들어, 온도를 감소시키고/감소시키거나 용해물의 점도를 증가시킴으로써 변경될 수 있다.

[0433] 일부 구현예에서, 샘플은 여과지를 사용하여 용해될 수 있다. 여과지는 여과지의 맨 위에서 용해 완충액로 적셔질 수 있다. 여과지는 샘플의 용해 및 샘플의 표적을 기재에 혼성화하는 것을 용이하게 할 수 있는 압력으로 샘플에 적용될 수 있다.

[0434] 일부 구현예에서, 용해는 기계적 용해, 열 용해, 광학 용해, 및/또는 화학적 용해에 의해 수행될 수 있다. 화학적 용해는 소화 효소, 예컨대 프로테이나제 K, 펩신, 및 트립신의 사용을 포함할 수 있다. 용해는 용해 완충액을 기재에 첨가함으로써 수행될 수 있다. 용해 완충액은 Tris HCl을 포함할 수 있다. 용해 완충액은 적어도 약 0.01, 0.05, 0.1, 0.5, 또는 1 M, 또는 그 이상의 Tris HCl을 포함할 수 있다. 용해 완충액은 최대 약 0.01, 0.05, 0.1, 0.5, 또는 1 M, 또는 그 이상의 Tris HCl을 포함할 수 있다. 용해 완충액은 약 0.1 M Tris HCl을 포함할 수 있다. 용해 완충액의 pH는 적어도 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 또는 그 이상일 수 있다. 용해 완충액의 pH는 최대 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 또는 그 이상일 수 있다. 일부 구현예에서, 용해 완충액의 pH는 약 7.5이다. 용해 완충액은 염(예를 들어, LiCl)을 포함할 수 있다. 용해 완충액 중 염의 농도는 적어도 약 0.1, 0.5, 또는 1 M, 또는 그 이상일 수 있다. 용해 완충액 중 염의 농도는 최대 약 0.1, 0.5, 또는 1 M, 또는 그 이상일 수 있다. 일부 구현예에서, 용해 완충액 중 염의 농도는 약 0.5M이다. 용해 완충액은 세제(예를 들어, SDS, Li 도데실 설페이트, 트리톤 X, 트윈, NP-40)를 포함할 수 있다. 용해 완충액 중 세제의 농도는 적어도 약 0.0001%, 0.0005%, 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 또는 7%, 또는 그 이상일 수 있다. 용해 완충액 중 세제의 농도는 최대 약 0.0001%, 0.0005%, 0.001%, 0.005%,

0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 또는 7%, 또는 그 이상일 수 있다. 일부 구현예에서, 용해 완충액 중 세제의 농도는 약 1% Li 도데실 설페이트이다. 용해 방법에 사용되는 시간은 사용되는 세제의 양에 따라 달라질 수 있다. 일부 구현예에서, 세제가 더 많이 사용될수록, 용해에 필요한 시간은 더 줄어든다. 용해 완충액은 킬레이트화제(예를 들어, EDTA, EGTA)를 포함할 수 있다. 용해 완충액 중 킬레이트화제의 농도는 적어도 약 1, 5, 10, 15, 20, 25, 또는 30 mM, 또는 그 이상일 수 있다. 용해 완충액 중 킬레이트화제의 농도는 최대 약 1, 5, 10, 15, 20, 25, 또는 30 mM, 또는 그 이상일 수 있다. 일부 구현예에서, 용해 완충액 중 킬레이트화제의 농도는 약 10 mM이다. 용해 완충액은 환원 시약(예를 들어, 베타-메르캅토에탄올, DTT)을 포함할 수 있다. 용해 완충액 중 환원 시약의 농도는 적어도 약 1, 5, 10, 15, 또는 20 mM, 또는 그 이상일 수 있다. 용해 완충액 중 환원 시약의 농도는 최대 약 1, 5, 10, 15, 또는 20 mM, 또는 그 이상일 수 있다. 일부 구현예에서, 용해 완충액 중 환원 시약의 농도는 약 5 mM이다. 일부 구현예에서, 용해 완충액은 약 0.1 M TrisHCl, 약 pH 7.5, 약 0.5 M LiCl, 약 1% 리튬 도데실 설페이트, 약 10 mM EDTA, 및 약 5 mM DTT를 포함할 수 있다.

[0435] 용해는 약 4, 10, 15, 20, 25, 또는 30℃의 온도에서 수행될 수 있다. 용해는 약 1, 5, 10, 15, 또는 20분, 또는 그 이상 동안 수행될 수 있다. 용해된 세포는 적어도 약 100000, 200000, 300000, 400000, 500000, 600000, 또는 700000개, 또는 그 이상의 표적 핵산 분자를 포함할 수 있다. 용해된 세포는 최대 약 100000, 200000, 300000, 400000, 500000, 600000, 또는 700000개, 또는 그 이상의 표적 핵산 분자를 포함할 수 있다.

[0436] 표적 핵산 분자에 대한 바코드의 부착

[0437] 세포의 용해 및 그로부터의 핵산 분자의 방출 후, 핵산 분자는 공동-국제화된 고체 지지체의 바코드와 랜덤으로 회합될 수 있다. 회합은, 바코드의 표적 인식 영역을 표적 핵산 분자의 상보적인 부분에 혼성화하는 것(예를 들어, 바코드의 올리고(dT)는 표적의 폴리(A) 테일과 상호작용할 수 있음)을 포함할 수 있다. 혼성화에 사용되는 분석 조건(예를 들어, 완충액 pH, 이온 강도, 온도 등)은 특정한 안정적인 혼성화물의 형성을 촉진하도록 선택될 수 있다. 일부 구현예에서, 용해된 세포로부터 방출된 핵산 분자는 기재 상에서 복수의 프로브와 회합될 수 있다(예를 들어, 기재 상에서 프로브와 혼성화됨). 프로브가 올리고(dT)를 포함하는 경우, mRNA 분자는 프로브에 혼성화되고, 역전사될 수 있다. 올리고뉴클레오타이드의 올리고(dT) 부분은 cDNA 분자의 제1 가닥 합성을 위한 프라이머로서 작용할 수 있다. 예를 들어, 도 2에 예시된 바코딩의 비제한적인 예에 있어서, 블록(216)에서, mRNA 분자는 비드 상에서 바코드에 혼성화될 수 있다. 예를 들어, 단일 가닥 뉴클레오타이드 단편은 바코드의 표적-결합 영역에 혼성화될 수 있다.

[0438] 부착은 바코드의 표적 인식 영역 및 표적 핵산 분자 부분의 결합을 추가로 포함할 수 있다. 예를 들어, 표적-결합 영역은 제한 부위 오버행(예를 들어, EcoRI 점착성-말단 오버행)에 특이적 혼성화가 가능할 수 있는 핵산 서열을 포함할 수 있다. 분석 절차는 표적 핵산을 제한 효소(예를 들어, EcoRI)로 처리하여 제한 부위 오버행을 생성하는 단계를 추가로 포함할 수 있다. 이어서 바코드는 제한 부위 오버행에 상보적인 서열을 포함하는 임의의 핵산 분자에 결합될 수 있다. 2개의 단편을 접합하기 위해 리가제(예를 들어, T4 DNA 리가제)가 사용될 수 있다.

[0439] 예를 들어, 도 2에 예시된 바코딩의 비제한적인 예에서, 블록(220)에서, 복수의 세포(또는 복수의 샘플)로부터 표지된 표적(예를 들어, 표적-바코드 분자)은, 예를 들어, 튜브 내로 연속적으로 풀링될 수 있다. 표지된 표적은, 예를 들어 표적-바코드 분자가 부착된 바코드 및/또는 비드를 회수함으로써 풀링될 수 있다.

[0440] 부착된 표적-바코드 분자의 고체 지지체-기반 집합의 회수는 자성 비드 및 외부-인가 자기장을 사용하여 실시될 수 있다. 표적-바코드 분자가 일단 풀링되면, 모든 추가 가공은 단일 반응 용기에서 진행될 수 있다. 추가 가공은, 예를 들어 역전사 반응, 증폭 반응, 절단 반응, 분해 반응, 및/또는 핵산 연장 반응을 포함할 수 있다. 추가 가공 반응은 마이크로 웰 내에서, 즉 표지된 표적 핵산 분자를 복수의 세포로부터 먼저 풀링하지 않고, 수행될 수 있다.

[0441] 역전사

[0442] 본 개시 내용은 역전사를 사용하여 표적-바코드 접합체를 생성하는 방법을 제공한다(예를 들어, 도 2의 블록(224)). 표적-바코드 접합체는 표적 핵산(즉, 바코딩된 cDNA 분자, 예컨대 추계적으로 바코딩된 cDNA 분자)의 전부 또는 일부분의 바코드 및 상보적인 서열을 포함할 수 있다. 회합된 RNA 분자의 역전사는 역전사효소와 함께 역전사 프라이머의 첨가에 의해 일어날 수 있다. 역전사 프라이머는 올리고(dT) 프라이머, 랜덤 핵산뉴클레오타이드 프라이머, 또는 표적-특이적 올리고뉴클레오타이드 프라이머일 수 있다. 올리고(dT) 프라이머는 12 내지 18개의 뉴클레오타이드 길이일 수 있거나, 약 12 내지 18개의 뉴클레오타이드 길이일 수 있고, 포유동물 mRNA의 3'

말단에서 내인성 폴리(A) 테일에 결합한다. 랜덤 핵산뉴클레오타이드 프라이머는 다양한 상보적인 부위에서 mRNA에 결합할 수 있다. 표적-특이적 올리고뉴클레오타이드 프라이머는 통상적으로 관심 대상의 mRNA를 선택적으로 프라이밍한다.

[0443] 일부 구현예에서, 표지된-RNA 분자의 역전사는 역전사 프라이머의 추가에 의해 일어날 수 있다. 일부 구현예에서, 역전사 프라이머는 올리고(dT) 프라이머, 랜덤 핵산뉴클레오타이드 프라이머, 또는 표적-특이적 올리고뉴클레오타이드 프라이머이다. 일반적으로, 올리고(dT) 프라이머는 12 내지 18개의 뉴클레오타이드 길이이고, 포유동물 mRNA의 3' 말단에서 내인성 폴리(A) 테일에 결합한다. 랜덤 핵산뉴클레오타이드 프라이머는 다양한 상보적인 부위에서 mRNA에 결합할 수 있다. 표적-특이적 올리고뉴클레오타이드 프라이머는 통상적으로 관심 대상의 mRNA를 선택적으로 프라이밍한다.

[0444] 역전사는 다수의 표지된-cDNA 분자를 생성하기 위해 반복적으로 일어날 수 있다. 본 명세서에 개시된 방법은 적어도 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20회의 역전사 반응을 수행하는 단계를 포함할 수 있다. 방법은 적어도 약 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100회의 역전사 반응을 수행하는 단계를 포함할 수 있다.

[0445] 증폭

[0446] 표지된 표적 핵산 분자의 다수의 복제물을 생성하기 위해, 하나 이상의 핵산 증폭 반응(예를 들어, 도 2의 블록(228))이 수행될 수 있다. 증폭은 다중화된 방식으로 수행될 수 있으며, 다중 표적 핵산 서열은 동시에 증폭된다. 증폭 반응은 서열분석 어댑터를 핵산 분자에 첨가하는 데 사용될 수 있다. 증폭 반응은 존재하는 경우, 샘플 표지의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 증폭 반응은 세포 표지 및/또는 바코드 서열(예를 들어, 분자 표지)의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 증폭 반응은 샘플 태그, 세포 표지, 공간 표지, 바코드 서열(예를 들어, 분자 표지), 표적 핵산, 또는 이들의 조합의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 증폭 반응은 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의, 복수의 핵산을 증폭시키는 단계를 포함할 수 있다. 방법은 하나 이상의 cDNA 합성 반응을 수행하여, 샘플 표지, 세포 표지, 공간 표지, 및/또는 바코드 서열(예를 들어, 분자 표지)을 포함하는 표적-바코드 분자의 하나 이상의 cDNA 복제물을 생성하는 단계를 추가로 포함할 수 있다.

[0447] 일부 구현예에서, 증폭은 중합효소 연쇄 반응(PCR)을 사용하여 수행될 수 있다. 본 명세서에 사용되는 바와 같이, PCR은 DNA의 상보적 가닥의 동시 프라이머 연장에 의해 특정 DNA 서열을 시험관 내 증폭하기 위한 반응을 지칭할 수 있다. 본 명세서에 사용되는 바와 같이, PCR은, RT-PCR, 실시간 PCR, 네스티드(nested) PCR, 정량 PCR, 다중 PCR, 디지털 PCR, 및 어셈블리 PCR을 포함하지만, 이에 제한되지는 않는 반응의 유도 형태들을 포괄할 수 있다.

[0448] 표지된 핵산의 증폭은 비-PCR 기반 방법을 포함할 수 있다. 비-PCR 기반 방법의 예는 다중 변위 증폭(MDA), 전사-매개 증폭(TMA), 핵산 서열-기반 증폭(NASBA), 가닥 변위 증폭(SDA), 실시간 SDA, 회전 원형 증폭, 또는 써클-투-써클(circle-to-circle) 증폭을 포함하지만, 이에 제한되지는 않는다. 기타 다른 비-PCR-기반 증폭 방법은, DNA 또는 RNA 표적을 증폭시키기 위한 DNA-의존성 RNA 중합효소-구동된 RNA 전사 증폭 또는 RNA-유도(directed) DNA 합성 및 전사의 다중 사이클, 리가제 사슬 반응(LCR), 및 Q β 복제효소(Q β) 방법, 회문식 프로브의 사용, 가닥 변위 증폭, 제한 엔도뉴클레아제를 사용하는 올리고뉴클레오타이드-구동된 증폭, 프라이머가 핵산 서열에 혼성화되고 생성된 듀플렉스가 연장 반응 및 증폭 전에 절단되는 증폭 방법, 5' 엑소뉴클레아제 활성이 없는 핵산 중합효소를 사용하는 가닥 변위 증폭, 회전 원형 증폭, 및 분지 연장 증폭(RAM)을 포함한다. 일부 구현예에서, 증폭은 원형화된 전사물을 생성하지 않는다.

[0449] 일부 구현예에서, 본 명세서에 개시된 방법은 표지된 핵산(예를 들어, 표지된-RNA, 표지된-DNA, 표지된-cDNA) 상에서 중합효소 연쇄 반응을 수행하여 표지된 앰플리콘(예를 들어, 추계적으로 표지된-앰플리콘)을 생성하는 단계를 추가로 포함한다. 표지된 앰플리콘은 이중 가닥 분자일 수 있다. 이중 가닥 분자는 이중 가닥 RNA 분자, 이중 가닥 DNA 분자, 또는 DNA 분자에 혼성화된 RNA 분자를 포함할 수 있다. 이중 가닥 분자 중 하나 또는 모두는 샘플 표지, 공간 표지, 세포 표지, 및/또는 바코드 서열(예를 들어, 분자 표지)을 포함할 수 있다. 표지된 앰플리콘은 단일 가닥 분자일 수 있다. 단일 가닥 분자는 DNA, RNA, 또는 이들의 조합을 포함할 수 있다. 본 개시 내용의 핵산은 합성 또는 변경된 핵산을 포함할 수 있다.

[0450] 증폭은 하나 이상의 비-천연 뉴클레오타이드의 사용을 포함할 수 있다. 비-천연 뉴클레오타이드는 광불안정성 또는

촉발성 뉴클레오티드를 포함할 수 있다. 비-천연 뉴클레오티드의 예는, 펩티드 핵산(PNA), 모르폴리노 및 고정 핵산(LNA)뿐만 아니라, 글리콜 핵산(GNA) 및 트레오스 핵산(TNA)을 포함할 수 있지만 이에 제한되지 않는다. 비-천연 뉴클레오티드는 1회 이상의 사이클의 증폭 반응에 첨가될 수 있다. 비-천연 뉴클레오티드의 첨가는 증폭 반응에서 특정 사이클 또는 시점으로서 산물을 식별하는 데 사용될 수 있다.

[0451] 하나 이상의 증폭 반응의 수행은 하나 이상의 프라이머의 사용을 포함할 수 있다. 하나 이상의 프라이머는, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 또는 15개, 또는 그 이상의 뉴클레오티드를 포함할 수 있다. 하나 이상의 프라이머는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 또는 15개, 또는 그 이상의 뉴클레오티드를 포함할 수 있다. 하나 이상의 프라이머는 12 내지 15개 미만의 뉴클레오티드를 포함할 수 있다. 하나 이상의 프라이머는 복수의 표지된 표적(예를 들어, 추가적으로 표지된 표적)의 적어도 일부분에 어닐링될 수 있다. 하나 이상의 프라이머는 복수의 표지된 표적의 3' 말단 또는 5' 말단에 어닐링될 수 있다. 하나 이상의 프라이머는 복수의 표지된 표적의 내부 영역에 어닐링될 수 있다. 내부 영역은 복수의 표지된 표적의 3' 말단으로부터 적어도 약 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 또는 1000개의 뉴클레오티드일 수 있다. 하나 이상의 프라이머는 프라이머의 고정된 패널을 포함할 수 있다. 하나 이상의 프라이머는 적어도 하나 이상의 커스텀(custom) 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 적어도 하나 이상의 제어 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 적어도 하나 이상의 유전자-특이적 프라이머를 포함할 수 있다.

[0452] 하나 이상의 프라이머는 범용 프라이머를 포함할 수 있다. 범용 프라이머는 범용 프라이머 결합 부위에 어닐링될 수 있다. 하나 이상의 커스텀 프라이머는 제1 샘플 표지, 제2 샘플 표지, 공간 표지, 세포 표지, 바코드 서열(예를 들어, 분자 표지), 표적, 또는 이들의 임의의 조합에 어닐링될 수 있다. 하나 이상의 프라이머는 범용 프라이머와 커스텀 프라이머를 포함할 수 있다. 커스텀 프라이머는 하나 이상의 표적을 증폭시키도록 설계될 수 있다. 표적은 하나 이상의 샘플 중 총 핵산의 하위세트를 포함할 수 있다. 표적은 하나 이상의 샘플 중 총 표지된 표적의 하위세트를 포함할 수 있다. 하나 이상의 프라이머는 적어도 96개 이상의 커스텀 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 적어도 960개 이상의 커스텀 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 적어도 9600개 이상의 커스텀 프라이머를 포함할 수 있다. 하나 이상의 커스텀 프라이머는 2개 이상의 상이한 표지된 핵산에 어닐링될 수 있다. 2개 이상의 상이한 표지된 핵산은 하나 이상의 유전자에 상응할 수 있다.

[0453] 임의의 증폭 체계가 본 개시 내용의 방법에 사용될 수 있다. 예를 들어, 하나의 체계에서, PCR의 제1회차는 유전자 특이적 프라이머, 및 범용 일루미나(Illumina) 서열분석 프라이머 1 서열에 대한 프라이머를 사용하여 비드에 부착된 분자를 증폭시킬 수 있다. PCR의 제2회차는, 일루미나 서열분석 프라이머 2 서열에 의해 측점된(flanked) 네스티드 유전자 특이적 프라이머, 및 범용 일루미나 서열분석 프라이머 1 서열에 대한 프라이머를 사용하여 제1 PCR 산물을 증폭시킬 수 있다. PCR의 제3회차는 P5와 P7 및 샘플 인덱스를 추가하여 PCR 산물을 일루미나 서열분석 라이브러리로 변화시킨다. 150 bp X 2 서열분석을 사용한 서열분석은 판독물 1 상에서 세포 표지 및 바코드 서열(예를 들어, 분자 표지), 판독물 2 상에서 유전자, 및 인덱스 1 판독물 상에서 샘플 인덱스를 나타낼 수 있다.

[0454] 일부 구현예에서, 핵산은 화학적 절단을 사용하여 기체로부터 제거될 수 있다. 예를 들어, 핵산 내 존재하는 화학 기 또는 변형된 염기가 고체 지지체로부터 핵산의 제거를 용이하게 하기 위해 사용될 수 있다. 예를 들어, 효소는 기체로부터 핵산을 제거하는 데 사용될 수 있다. 예를 들어, 핵산은 제한 엔도뉴클레아제 분해를 통해 기체로부터 제거될 수 있다. 예를 들어, 우라실-d-글리코실라제(UDG)로 dUTP 또는 ddUTP를 함유하는 핵산을 처리하는 것이 기체로부터 핵산을 제거하는 데 사용될 수 있다. 예를 들어, 핵산은 뉴클레오티드 절제를 수행하는 효소, 예컨대 염기 절제 복구 효소, 예컨대 아푸린/아피리미딘(apurinic/aprimidinic)(AP) 엔도뉴클레아제를 사용하여 기체로부터 제거될 수 있다. 일부 구현예에서, 핵산은 광절단성 기 및 광을 사용하여 기체로부터 제거될 수 있다. 일부 구현예에서, 절단 가능한 링커가 기체로부터 핵산을 제거하는 데 사용될 수 있다. 예를 들어, 절단 가능한 링커는 적어도 하나의 비오틴/아비딘, 비오틴/스트렙타비딘, 비오틴/뉴트라비딘, Ig-단백질 A, 광 불안정성 링커, 산 또는 염기 불안정성 링커 기, 또는 애타머를 포함할 수 있다.

[0455] 프로브가 유전자-특이적인 경우, 분자는 프로브에 혼성화될 수 있고, 역전사되고/역전사되거나 증폭될 수 있다. 일부 구현예에서, 핵산은 합성된(예를 들어, 역전사된) 후, 증폭될 수 있다. 증폭은 다중 방식으로 수행될 수

있는 한편, 다중 표적 핵산 서열은 동시에 증폭된다. 증폭은 서열분석 어댑터를 핵산에 첨가할 수 있다.

- [0456] 일부 구현예에서, 증폭은 기재 상에서, 예를 들어 가고 증폭을 이용하여 수행될 수 있다. cDNA는 기재 상에서 올리고(dT) 프로브를 사용하여 가고 증폭에 상용가능한 말단을 생성하기 위한 테일이 있는 단일중합체일 수 있다. 가고 증폭에서, 주형 핵산의 3' 말단에 상보적인 프라이머는 고체 입자에 공유적으로 부착된 각 쌍의 제1 프라이머일 수 있다. 주형 핵산을 함유하는 샘플이 입자와 접촉되고 단일 열 사이클이 수행되는 경우, 주형 분자는 제1 프라이머에 어닐링될 수 있고, 제1 프라이머는 뉴클레오타이드의 첨가에 의해 순방향으로 신장되어 주형 분자, 및 주형에 상보적인 새로 형성된 DNA 가닥으로 이루어지는 듀플렉스 분자를 형성한다. 다음 사이클의 가열 단계에서, 듀플렉스 분자는 변성되어, 입자로부터 주형 분자를 방출시키고, 제1 프라이머를 통해 입자에 부착된 상보적 DNA 가닥을 남길 수 있다. 이어지는 어닐링 및 신장 단계의 어닐링 단계에서, 상보적 가닥은 제2 프라이머에 혼성화될 수 있으며, 여기서 제2 프라이머는 제1 프라이머로부터 제거된 위치에서 상보적 가닥의 절편에 상보적이다. 이러한 혼성화는, 공유결합에 의해 제1 프라이머에, 그리고 혼성화에 의해 제2 프라이머에 고정되는, 제1 프라이머 및 제2 프라이머 사이의 가교를 형성하는 상보적 가닥을 야기할 수 있다. 신장 단계에서, 제2 프라이머는 동일한 반응 혼합물 내 뉴클레오타이드의 첨가에 의해 역방향으로 신장됨으로써, 가교를 이중 가닥 가교로 전환시킬 수 있다. 이어서 다음 사이클이 시작되고, 이중 가닥 가교는 변성되어, 각각 제1 프라이머 및 제2 프라이머를 통해 입자 표면에 부착된 한쪽 말단과, 미부착된 각각의 다른 말단을 각각 갖는, 2개의 단일 가닥 핵산 분자를 산출할 수 있다. 이 제2 사이클의 어닐링 및 신장 단계에서, 각각의 가닥은 동일 입자 상에서 이전에 사용되지 않은 추가의 상보적 프라이머에 혼성화되어, 신규의 단일 가닥 가교를 형성할 수 있다. 이제 혼성화되는 이전에 사용되지 않은 2개의 프라이머는 신장되어, 2개의 새로운 가교를 이중 가닥 가교로 전환시킨다.
- [0457] 증폭 반응은 적어도 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 또는 100%의 복수의 핵산을 증폭시키는 단계를 포함할 수 있다.
- [0458] 표지된 핵산의 증폭은 PCR-기반 방법 또는 비-PCR 기반 방법을 포함할 수 있다. 표지된 핵산의 증폭은 표지된 핵산의 지수적 증폭을 포함할 수 있다. 표지된 핵산의 증폭은 표지된 핵산의 선형 증폭을 포함할 수 있다. 증폭은 중합효소 연쇄 반응(PCR)에 의해 수행될 수 있다. PCR은 DNA의 상보적 가닥의 동시 프라이머 연장에 의한 특정 DNA 서열의 시험관 내 증폭을 위한 반응을 지칭할 수 있다. PCR은 RT-PCR, 실시간 PCR, 네스티드 PCR, 정량 PCR, 다중 PCR, 디지털 PCR, 억제 PCR, 반-억제성 PCR 및 어셈블리 PCR을 포함하지만, 이에 제한되지는 않는 반응의 유도 형태들을 포괄할 수 있다.
- [0459] 일부 구현예에서, 표지된 핵산의 증폭은 비-PCR 기반 방법을 포함한다. 비-PCR 기반 방법의 예는 다중 변위 증폭(MDA), 전사-매개 증폭(TMA), 핵산 서열-기반 증폭(NASBA), 가닥 변위 증폭(SDA), 실시간 SDA, 회전 원형 증폭, 또는 씨클-투-씨클 증폭을 포함하지만, 이에 제한되지는 않는다. 기타 다른 비-PCR-기반 증폭 방법은, DNA 또는 RNA 표적을 증폭시키기 위한 DNA-의존성 RNA 중합효소-구동된 RNA 전사 증폭 또는 RNA-유도 DNA 합성 및 전사의 다중 사이클, 리가제 사슬 반응(LCR), 및 Q β 복제효소(Q β), 회문식 프로브의 사용, 가닥 변위 증폭, 제한 엔도뉴클레아제를 사용하는 올리고뉴클레오타이드-구동된 증폭, 프라이머가 핵산 서열에 혼성화되고 생성된 듀플렉스가 연장 반응 및 증폭 전에 절단되는 증폭 반응, 5' 엑소뉴클레아제 활성이 없는 핵산 중합효소를 사용하는 가닥 변위 증폭, 회전 원형 증폭, 및 분지 연장 증폭(RAM)을 포함한다.
- [0460] 일부 구현예에서, 본 명세서에 개시된 방법은 증폭된 앰플리콘(예를 들어, 표적) 상에서 네스티드 중합효소 연쇄 반응을 수행하는 단계를 추가로 포함한다. 앰플리콘은 이중 가닥 분자일 수 있다. 이중 가닥 분자는 이중 가닥 RNA 분자, 이중 가닥 DNA 분자, 또는 DNA 분자에 혼성화된 RNA 분자를 포함할 수 있다. 이중 가닥 분자의 가닥 하나 또는 둘 모두는 샘플 태그 또는 분자 식별자 표지를 포함할 수 있다. 대안적으로, 앰플리콘은 단일 가닥 분자일 수 있다. 단일 가닥 분자는 DNA, RNA, 또는 이들의 조합을 포함할 수 있다. 본 발명의 핵산은 합성 또는 변형된 핵산을 포함할 수 있다.
- [0461] 일부 구현예에서, 방법은 표지된 핵산을 반복적으로 증폭시켜 다수의 앰플리콘을 생성하는 단계를 포함한다. 본 명세서에 개시된 방법은 적어도 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20회의 증폭 반응을 수행하는 단계를 포함할 수 있다. 대안적으로는, 방법은 적어도 약 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100회의 증폭 반응을 수행하는 단계를 포함한다.
- [0462] 증폭은 하나 이상의 제어 핵산을, 복수의 핵산을 포함하는 하나 이상의 샘플에 첨가하는 단계를 추가로 포함할 수 있다. 증폭은 하나 이상의 제어 핵산을 복수의 핵산에 첨가하는 단계를 추가로 포함할 수 있다. 제어 핵산은

제어 표지를 포함할 수 있다.

[0463] 증폭은 하나 이상의 비-천연 뉴클레오타이드의 사용을 포함할 수 있다. 비-천연 뉴클레오타이드는 광불안정성 및/또는 촉발성 뉴클레오타이드를 포함할 수 있다. 비-천연 뉴클레오타이드의 예는, 펩티드 핵산(PNA), 모르폴리노 및 고정 핵산(LNA)뿐만 아니라, 글리콜 핵산(GNA) 및 트레오스 핵산(TNA)을 포함할 수 있지만, 이에 제한되지 않는다. 비-천연 뉴클레오타이드는 1회 이상 사이클의 증폭 반응에 첨가될 수 있다. 비-천연 뉴클레오타이드의 첨가는 증폭 반응에서 특정 사이클 또는 시점으로서 산물을 식별하는 데 사용될 수 있다.

[0464] 하나 이상의 증폭 반응을 수행하는 단계는 하나 이상의 프라이머의 사용을 포함할 수 있다. 하나 이상의 프라이머는 하나 이상의 올리고뉴클레오타이드를 포함할 수 있다. 하나 이상의 올리고뉴클레오타이드는 적어도 약 7 내지 9개의 뉴클레오타이드를 포함할 수 있다. 하나 이상의 올리고뉴클레오타이드는 12 내지 15개 미만의 뉴클레오타이드를 포함할 수 있다. 하나 이상의 프라이머는 복수의 표지된 핵산의 적어도 일부분에 어닐링될 수 있다. 하나 이상의 프라이머는 복수의 표지된 핵산의 3' 말단 및/또는 5' 말단에 어닐링될 수 있다. 하나 이상의 프라이머는 복수의 표지된 핵산의 내부 영역에 어닐링될 수 있다. 내부 영역은 복수의 표지된 핵산의 3' 말단으로부터 적어도 약 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 또는 1000개의 뉴클레오타이드일 수 있다. 하나 이상의 프라이머는 프라이머의 고정된 패들을 포함할 수 있다. 하나 이상의 프라이머는 적어도 하나 이상의 커스텀 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 적어도 하나 이상의 하우스키핑(housekeeping) 유전자 프라이머를 포함할 수 있다. 하나 이상의 프라이머는 범용 프라이머를 포함할 수 있다. 범용 프라이머는 범용 프라이머 결합 부위에 어닐링될 수 있다. 하나 이상의 커스텀 프라이머는 제1 샘플 태그, 제2 샘플 태그, 분자 식별자 표지, 핵산 또는 이의 산물에 어닐링될 수 있다. 하나 이상의 프라이머는 범용 프라이머 및 커스텀 프라이머를 포함할 수 있다. 커스텀 프라이머는 하나 이상의 표적 핵산을 증폭시키도록 설계될 수 있다. 표적 핵산은 하나 이상의 샘플 중 총 핵산의 하위세트를 포함할 수 있다. 일부 구현예에서, 프라이머는 본 개시 내용의 어레이에 부착된 프로브이다.

[0465] 일부 구현예에서, 샘플 내 복수의 표적을 바코딩(예를 들어, 추계적 바코딩)하는 단계는 바코딩된 표적(예를 들어, 추계적으로 바코딩된 표적) 또는 표적의 바코딩된 단편의 인덱싱된 라이브러리를 생성하는 단계를 추가로 포함한다. 상이한 바코드의 바코드 서열(예를 들어, 상이한 추계적 바코드의 분자 표지)은 서로 상이할 수 있다. 바코딩된 표적의 인덱싱된 라이브러리를 생성하는 단계는 샘플 내 복수의 표적로부터 복수의 인덱싱된 폴리뉴클레오타이드들을 생성하는 단계를 포함한다. 예를 들어, 제1 인덱싱된 표적 및 제2 인덱싱된 표적을 포함하는 바코딩된 표적의 인덱싱된 라이브러리의 경우, 제1 인덱싱된 폴리뉴클레오타이드의 표지 영역은 제2 인덱싱된 폴리뉴클레오타이드의 표지 영역과, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50개만큼, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50개만큼, 적어도 이들 값만큼, 또는 최대 이들 값 만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드만큼 상이할 수 있다. 일부 구현예에서, 바코딩된 표적의 인덱싱된 라이브러리를 생성하는 단계는, 복수의 표적, 예를 들어, mRNA 분자를, 폴리(T) 영역 및 표지 영역을 포함하는 복수의 올리고뉴클레오타이드와 접촉시키는 단계; 및 역전사효소를 사용하여 제1 가닥 합성을 수행하여, cDNA 영역 및 표지 영역을 각각 포함하는 단일 가닥 표지된 cDNA 분자를 생성하는 단계를 포함하며, 여기서 복수의 표적은 상이한 서열의 적어도 2개의 mRNA 분자를 포함하고, 복수의 올리고뉴클레오타이드는 상이한 서열의 적어도 2개의 올리고뉴클레오타이드를 포함한다. 바코딩된 표적의 인덱싱된 라이브러리를 생성하는 단계는, 단일 가닥 표지된 cDNA 분자를 증폭시켜 이중 가닥 표지된 cDNA 분자를 생성하는 단계; 및 이중 가닥 표지된 cDNA 분자 상에서 네스티드 PCR을 수행하여 표지된 앰플리콘을 생성하는 단계를 추가로 포함할 수 있다. 일부 구현예에서, 방법은 어댑터-표지된 앰플리콘을 생성하는 단계를 포함할 수 있다.

[0466] 바코딩(예를 들어, 추계적 바코딩)은 개별적인 핵산(예를 들어, DNA 또는 RNA) 분자를 표지하기 위해 핵산 바코드 또는 태그를 사용하는 것을 포함할 수 있다. 일부 구현예에서, DNA 바코드 또는 태그가 mRNA로부터 생성됨에 따라, 이들을 cDNA 분자에 첨가하는 단계를 포함한다. 네스티드 PCR이 PCR 증폭 편향을 최소화하기 위해 수행될 수 있다. 어댑터는, 예를 들어 차세대 서열분석(NGS)를 사용하여 서열분석하기 위해 첨가될 수 있다. 서열분석 결과는, 예를 들어 도 2의 블록(232)에서, 세포 표지, 분자 표지, 및 표적의 하나 이상의 복제물의 뉴클레오타이드 단편의 서열을 결정하는 데 사용될 수 있다.

[0467] 도 3은 바코딩된 표적(예를 들어, 추계적으로 바코딩된 표적), 예컨대 바코딩된 mRNA 또는 이의 단편의 인덱싱된 라이브러리를 생성하는 비제한적이고 예시적인 과정을 나타내는 개략도이다. 단계 1에 나타난 바와 같이, 역전사 과정은 고유 분자 표지, 세포 표지, 및 범용 PCR 부위를 갖는 mRNA 분자 각각을 인코딩할 수 있다. 구체적인

으로, RNA 분자(302)는 역전사되어 표지된 cDNA 분자(304)를 생성할 수 있으며, 여기서 표지된 cDNA 분자(304)는 바코드(예를 들어, 추계적 바코드)(310)의 세트를 RNA 분자(302)의 폴리(A) 테일 영역(308)에 혼성화(예를 들어, 추계적 혼성화)함으로써 cDNA 영역(306)을 포함한다. 바코드(310) 각각은 표적-결합 영역, 예를 들어 폴리(dT) 영역(312), 표지 영역(314)(예를 들어, 바코드 서열 또는 분자), 및 범용 PCR 영역(316)을 포함할 수 있다.

[0468] 일부 구현예에서, 세포 표지는 3 내지 20개의 뉴클레오타이드를 포함할 수 있다. 일부 구현예에서, 분자 표지는 3 내지 20개의 뉴클레오타이드를 포함할 수 있다. 일부 구현예에서, 복수의 추계적 바코드 각각은 범용 표지 및 세포 표지 중 하나 이상을 추가로 포함하며, 여기서 범용 표지는 고체 지지체 상에서 복수의 추계적 바코드에 대해 동일하고, 세포 표지는 고체 지지체 상에서 복수의 추계적 바코드에 대해 동일하다. 일부 구현예에서, 범용 표지는 3 내지 20개의 뉴클레오타이드를 포함할 수 있다. 일부 구현예에서, 세포 표지는 3 내지 20개의 뉴클레오타이드를 포함한다.

[0469] 일부 구현예에서, 표지 영역(314)은 바코드 서열 또는 분자 표지(318) 및 세포 표지(320)를 포함할 수 있다. 일부 구현예에서, 표지 영역(314)은 범용 표지, 치수 표지, 및 세포 표지 중 하나 이상을 포함할 수 있다. 바코드 서열 또는 분자 표지(318)는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 세포 표지(320)는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 범용 표지는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 범용 표지는 고체 지지체 상에서 복수의 추계적 바코드에 대해 동일하며, 세포 표지는 고체 지지체 상에서 복수의 추계적 바코드에 대해 동일하다. 치수 표지는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다.

[0470] 일부 구현예에서, 표지 영역(314)은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 표지, 예컨대 바코드 서열 또는 분자 표지(318) 및 세포 표지(320)를 포함할 수 있다. 각각의 표지는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 또는 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 바코드 또는 추계적 바코드(310)의 세트는 10, 20, 40, 50, 70, 80, 90, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{20} 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 바코드 또는 추계

적 바코드(310)를 함유할 수 있거나, 약 10, 20, 40, 50, 70, 80, 90, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{20} 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 바코드 또는 추계적 바코드(310)를 함유할 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 바코드 또는 추계적 바코드(310)를 함유할 수 있다. 바코드 또는 추계적 바코드(310)의 세트는, 예를 들어 고유 표지 영역(314)을 각각 함유할 수 있다. 표지된 cDNA 분자(304)는 과량의 바코드 또는 추계적 바코드(310)를 제거하기 위해 정제될 수 있다. 정제는 암퓨어(Ampure) 비드 정제를 포함할 수 있다.

[0471] 단계 2에 나타난 바와 같이, 단계 1에서 역전사 공정으로부터의 산물은 1 개의 튜브 내로 풀링되고, 제1 PCR 프라이머 풀 및 제1 범용 PCR 프라이머를 이용하여 PCR 증폭될 수 있다. 고유 표지 영역(314) 때문에 풀링이 가능하다. 구체적으로, 표지된 cDNA 분자(304)는 증폭되어 네스티드 PCR 표지된 앰플리콘(322)을 생성할 수 있다. 증폭은 다중 PCR 증폭을 포함할 수 있다. 증폭은 단일 반응 부피 내 96개의 다중 프라이머를 이용한 다중 PCR 증폭을 포함할 수 있다. 일부 구현예에서, 다중 PCR 증폭은 단일 반응 부피 내 10, 20, 40, 50, 70, 80, 90, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{20} 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 다중 프라이머를 사용할 수 있거나, 약 0, 20, 40, 50, 70, 80, 90, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{20} 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 다중 프라이머를 사용할 수 있거나, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 다중 프라이머를 사용할 수 있거나, 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 다중 프라이머를 사용할 수 있다. 증폭은 특정 유전자를 표적화하는 커스텀 프라이머(326A~C) 및 범용 프라이머(328)를 포함하는 제1 PCR 프라이머 풀(324)을 사용하는 것을 포함할 수 있다. 커스텀 프라이머(326)는 표지된 cDNA 분자(304)의 cDNA 부분(306') 내의 영역에 혼성화될 수 있다. 범용 프라이머(328)는 표지된 cDNA 분자(304)의 범용 PCR 영역(316)에 혼성화될 수 있다.

[0472] 도 3의 단계 3에 나타난 바와 같이, 단계 2에서 PCR 증폭으로부터의 산물은 네스티드 PCR 프라이머 풀 및 제2 범용 PCR 프라이머를 이용하여 증폭될 수 있다. 네스티드 PCR은 PCR 증폭 편향을 최소화할 수 있다. 구체적으로, 네스티드 PCR 표지된 앰플리콘(322)은 네스티드 PCR에 의해 추가로 증폭될 수 있다. 네스티드 PCR은 단일 반응 부피 내에서 네스티드 PCR 프라이머(332a~c)의 네스티드 PCR 프라이머 풀(330) 및 제2 범용 PCR 프라이머(328')를 이용하는 다중 PCR을 포함할 수 있다. 네스티드 PCR 프라이머 풀(328)은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 적어도 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 최대 이들 값, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 네스티드 PCR 프라이머(330)를 함유할 수 있다. 네스티드 PCR 프라이머(332)는 어댑터(334)를 함유할 수 있고, 표지된 앰플리콘(322)의 cDNA 부분(306") 내 영역에 혼성화될 수 있다. 범용 프라이머(328')는 어댑터(336)를 함유할 수 있고, 표지된 앰플리콘(322)의 범용 PCR 영역(316)에 혼성화될 수 있다. 따라서, 단계 3은 어댑터-표지된 앰플리콘(338)을 생성한다. 일부 구현예에서, 네스티드 PCR 프라이머(332) 및 제2 범용 PCR 프라이머(328')는 어댑터(334 및 336)를 함유하지 않을 수 있다. 대신에 어댑터(334 및 336)는 네스티드 PCR의 산물에 결합되어 어댑터-표지된 앰플리콘(338)을 생성할 수 있다.

[0473] 단계 4에 나타난 바와 같이, 단계 3으로부터의 PCR 산물은 라이브러리 증폭 프라이머를 사용하여 서열분석을 위해 PCR 증폭될 수 있다. 구체적으로, 어댑터(334 및 336)는 어댑터-표지된 앰플리콘(338) 상에서 하나 이상의 추가 분석을 수행하는 데 사용될 수 있다. 어댑터(334 및 336)는 프라이머(340 및 342)에 혼성화될 수 있다. 하나 이상의 프라이머(340 및 342)는 PCR 증폭 프라이머일 수 있다. 하나 이상의 프라이머(340 및 342)는 서열분석 프라이머일 수 있다. 하나 이상의 어댑터(334 및 336)는 어댑터-표지된 앰플리콘(338)의 추가의 증폭에 사용될 수 있다. 하나 이상의 어댑터(334 및 336)는 어댑터-표지된 앰플리콘(338)의 서열분석에 사용될 수 있다. 프라이머(342)는 플레이트 인덱스(344)를 함유하여, 바코드 또는 추계적 바코드(310)의 동일한 세트를 사용하여 생성된 앰플리콘이 차세대 서열분석(NGS)를 사용하는 하나의 서열분석 반응에서 서열분석될 수 있도록 한다.

[0474] 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하는 조성물

- [0475] 본 명세서에 개시된 일부 구현에는 올리고뉴클레오타이드와 접합된 세포 성분 결합 시약(예컨대, 단백질 결합 시약)을 각각 포함하는 복수의 조성물을 제공하며, 올리고뉴클레오타이드는 접합되는 세포 성분 결합 시약에 대한 고유 식별자를 포함한다. 일부 구현예에서, 세포 성분 결합 시약은 세포 성분 표적에 특이적으로 결합할 수 있다. 예를 들어, 세포 성분 결합 시약의 결합 표적은 탄수화물, 지질, 단백질, 세포외 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 세포내 단백질, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 일부 구현예에서, 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)은 항원 표적 또는 단백질 표적에 특이적으로 결합할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 바코드, 예컨대 추계적 바코드를 포함할 수 있다. 바코드는 바코드 서열(예를 들어, 분자 표지), 세포 표지, 샘플 표지, 또는 이들의 임의의 조합을 포함할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 링커를 포함할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 올리고뉴클레오타이드 프로브, 예컨대 폴리(A) 테일에 대한 결합 부위를 포함할 수 있다. 예를 들어, 폴리(A) 테일은, 예를 들어 고체 지지체에 고착되지 않거나 고체 지지체에 고착될 수 있다. 폴리(A) 테일은 약 10 내지 50개의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 폴리(A) 테일은 18개의 뉴클레오타이드 길이일 수 있다. 올리고뉴클레오타이드는 데옥시리보뉴클레오타이드, 리보뉴클레오타이드 또는 이들 모두를 포함할 수 있다.
- [0476] 고유 식별자는, 예를 들어 임의의 적합한 길이, 예를 들어 약 4개의 뉴클레오타이드 내지 약 200개의 뉴클레오타이드를 갖는 뉴클레오타이드 서열일 수 있다. 일부 구현예에서, 고유 식별자는 25개의 뉴클레오타이드 내지 약 45개의 뉴클레오타이드 길이의 뉴클레오타이드 서열이다. 일부 구현예에서, 고유 식별자는 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9 개의 뉴클레오타이드, 10개의 뉴클레오타이드, 15개의 뉴클레오타이드, 20개의 뉴클레오타이드, 25개의 뉴클레오타이드, 30개의 뉴클레오타이드, 35개의 뉴클레오타이드, 40개의 뉴클레오타이드, 45개의 뉴클레오타이드, 50개의 뉴클레오타이드, 55개의 뉴클레오타이드, 60개의 뉴클레오타이드, 70개의 뉴클레오타이드, 80개의 뉴클레오타이드, 90개의 뉴클레오타이드, 100개의 뉴클레오타이드, 200개의 뉴클레오타이드, 또는 상기 값 중 임의의 두 값 사이의 어떤 수 또는 범위, 약 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9 개의 뉴클레오타이드, 10개의 뉴클레오타이드, 15개의 뉴클레오타이드, 20개의 뉴클레오타이드, 25개의 뉴클레오타이드, 30개의 뉴클레오타이드, 35개의 뉴클레오타이드, 40개의 뉴클레오타이드, 45개의 뉴클레오타이드, 50개의 뉴클레오타이드, 55개의 뉴클레오타이드, 60개의 뉴클레오타이드, 70개의 뉴클레오타이드, 80개의 뉴클레오타이드, 90개의 뉴클레오타이드, 100개의 뉴클레오타이드, 200개의 뉴클레오타이드, 또는 상기 값 중 임의의 두 값 사이의 어떤 수 또는 범위, 이들 값, 또는 상기 값 중 임의의 두 값 사이의 어떤 수 또는 범위 미만, 이들 값, 또는 상기 값 중 임의의 두 값 사이의 어떤 수 또는 범위 초과를 가질 수 있다.
- [0477] 일부 구현예에서, 고유 식별자는 고유 식별자의 다양한 세트로부터 선택된다. 고유 식별자의 다양한 세트는 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 고유 식별자를 포함할 수 있거나, 약 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 고유 식별자를 포함할 수 있다. 고유 식별자의 다양한 세트는 적어도, 또는 최대 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 또는 5000개의 상이한 고유 식별자를 포함할 수 있다. 일부 구현예에서, 고유 식별자의 세트는 분석될 샘플의 DNA 또는 RNA 서열과의 최소 서열 상동성을 갖도록 설계된다. 일부 구현예에서, 고유 식별자의 세트의 서열은, 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 또는 약 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼 서로 상이하거나, 이들의 상보체이다. 일부 구현예에서, 고유 식별자의 세트의 서열은 적어도, 또는 최대 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드만큼 서로 상이하거나, 이들의 상보체이다. 일부 구현예에서, 고유 식별자의 세트의 서열은 적어도 3%, 적어도 5%, 적어도 8%, 적어도 10%, 적어도 15%, 적어도 20%, 또는 그 이상에 의해, 서로 상이하거나, 그의 보체이다.
- [0478] 일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머에 대한 결합 부위를 포함할 수 있다. 일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머에 대한 적어도 2개의 결합 부위를 포함할 수 있다.

일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머에 대한 적어도 3개의 결합 부위를 포함할 수 있다. 프라이머는, 예를 들어 PCR 증폭에 의해, 고유 식별자의 증폭에 사용될 수 있다. 일부 구현예에서, 프라이머는 네스티드 PCR 반응에 사용될 수 있다.

[0479] 임의의 적합한 세포 성분 결합 시약이 본 개시 내용에서 고려되는데, 이에는, 예를 들어 단백질 결합 시약, 항체 또는 이의 단편, 압타머, 소분자, 리간드, 펩티드, 올리고뉴클레오타이드 등, 또는 이들의 임의의 조합이 있다. 일부 구현예에서, 세포 성분 결합 시약은 다중클론 항체, 단일클론 항체, 재조합 항체, 단쇄 항체(sc-Ab), 또는 이들의 단편, 예컨대 Fab, Fv 등일 수 있다. 일부 구현예에서, 복수의 세포 성분 결합 시약은 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 시약을 포함할 수 있거나, 약 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 시약을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 결합 시약은 적어도, 또는 최대 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 또는 5000개의 상이한 세포 성분 시약을 포함할 수 있다.

[0480] 올리고뉴클레오타이드는 다양한 메커니즘을 통해 세포 성분 결합 시약과 접합될 수 있다. 일부 구현예에서, 올리고뉴클레오타이드는 세포 성분 결합 시약과 공유적으로 접합될 수 있다. 일부 구현예에서, 올리고뉴클레오타이드는 세포 성분 결합 시약과 비-공유적으로 접합될 수 있다. 일부 구현예에서, 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약과 접합된다. 링커는, 예를 들어 세포 성분 결합 시약 및/또는 올리고뉴클레오타이드로부터 절단 가능하거나 탈착 가능할 수 있다. 일부 구현예에서, 링커는 올리고뉴클레오타이드를 세포 성분 결합 시약에 가역적으로 부착시키는 화학 기를 포함할 수 있다. 화학 기는, 예를 들어 아민 기를 통해 링커에 접합될 수 있다. 일부 구현예에서, 링커는 세포 성분 결합 시약에 접합된 또 다른 화학 기와 안정적인 결합을 형성하는 화학 기를 포함할 수 있다. 예를 들어, 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민 등일 수 있다. 일부 구현예에서, 화학 기는 아미노산, 예컨대 리신 상의 1차 아민, 또는 N-말단을 통해 세포 성분 결합 시약에 접합될 수 있다. 구매 가능한 컨주게이션 키트, 예컨대 단백질-올리고 컨주게이션 키트(미국 캘리포니아주 샌디에고 소재의 Solulink, Inc.), Thunder-Link[®] 올리고 컨주게이션 시스템(영국 캠브리지 소재의 Innova Biosciences) 등이 올리고뉴클레오타이드를 세포 성분 결합 시약에 접합하는 데 사용될 수 있다.

[0481] 올리고뉴클레오타이드는, 그것이 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)과 그의 세포 성분 표적 사이의 특이적 결합을 방해하지 않는 한, 세포 성분 결합 시약의 임의의 적합한 부위에 접합될 수 있다. 일부 구현예에서, 세포 성분 결합 시약은 단백질, 예컨대 항체이다. 일부 구현예에서, 세포 성분 결합 시약은 항체가 아니다. 일부 구현예에서, 올리고뉴클레오타이드는 항원-결합 부위 이외의 어느 곳에서도, 예를 들어 Fc 영역, C_H1 도메인, C_H2 도메인, C_H3 도메인, C_L 도메인 등에서 항체에 접합될 수 있다. 올리고뉴클레오타이드를 세포 성분 결합 시약(예를 들어, 항체)에 접합하는 방법은, 예를 들어 미국 특허 6,531,283에 이전에 개시되어 있으며, 이의 내용은 전체가 본 명세서에 명시적으로 참고로 포함된다. 올리고뉴클레오타이드 대 세포 성분 결합 시약의 화학량은 변동될 수 있다. 서열분석에서 세포 성분 결합 시약 특이적 올리고뉴클레오타이드의 검출 감도를 증가시키기 위하여, 접합 동안 올리고뉴클레오타이드 대 세포 성분 결합 시약의 비를 증가시키는 것이 유리할 수 있다. 일부 구현예에서, 각각의 세포 성분 결합 시약은 단일 올리고뉴클레오타이드 분자와 접합될 수 있다. 일부 구현예에서, 각각의 세포 성분 결합 시약은 하나 초과 올리고뉴클레오타이드 분자, 예를 들어 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 올리고뉴클레오타이드 분자와 접합될 수 있으며, 여기서 올리고뉴클레오타이드 분자 각각은 동일하거나 상이한 고유 식별자를 포함한다. 일부 구현예에서, 각각의 세포 성분 결합 시약은 하나 초과 올리고뉴클레오타이드 분자, 예를 들어 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000개의 올리고뉴클레오타이드 분자와 접합될 수 있으며, 여기서 올리고뉴클레오타이드 분자 각각은 동일하거나 상이한 고유 식별자를 포함한다.

[0482] 일부 구현예에서, 복수의 세포 성분 결합 시약은 샘플, 예컨대 단일 세포, 복수의 세포, 조직 샘플, 종양 샘플, 혈액 샘플 등 내의 복수의 세포 성분 표적에 특이적으로 결합할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합을 포함한다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 세포 성분을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 세포 성분을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 세포 또는 유기체 내의 모든 세포 성분(예를 들어, 단백질)의 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두

값 사이의 어떤 수 또는 범위일 수 있거나, 약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 세포 성분은 세포 또는 유기체 내의 모든 세포 성분(예를 들어, 단백질)의 적어도, 또는 최대 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 또는 99%일 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있거나, 약 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개의 상이한 세포 성분 표적을 포함할 수 있다.

[0483] 도 4는 항체에 대한 고유 식별자 서열을 포함하는 올리고뉴클레오타이드와 접합되어 있는, 예시적인 세포 성분 결합 시약, 예를 들어 항체의 개략도를 나타낸다. 세포 성분 결합 시약과 접합된 올리고뉴클레오타이드, 세포 성분 결합 시약과의 접합을 위한 올리고뉴클레오타이드, 또는 세포 성분 결합 시약과 미리 접합된 올리고뉴클레오타이드는 본 명세서에서 항체 올리고뉴클레오타이드(결합 시약 올리고뉴클레오타이드로 약기됨)로 지칭될 수 있다. 항체와 접합된 올리고뉴클레오타이드, 항체와 접합하기 위한 올리고뉴클레오타이드, 또는 항체와 미리 접합된 올리고뉴클레오타이드는 본 명세서에서 항체 올리고뉴클레오타이드("Ab올리고" 또는 "AbO"로 약기됨)로 지칭될 수 있다. 또한, 올리고뉴클레오타이드는 추가의 성분을 포함할 수 있으며, 이에에는 하나 이상의 링커, 항체에 대한 하나 이상의 고유 식별자, 선택적으로 하나 이상의 바코드 서열(예를 들어, 분자 표지), 및 폴리(A) 테일이 포함되지만 이에 제한되지는 않는다. 일부 구현예에서, 올리고뉴클레오타이드는, 5'에서 3'으로, 링커, 고유 식별자, 바코드 서열(예를 들어, 분자 표지), 및 폴리(A) 테일을 포함할 수 있다. 항체 올리고뉴클레오타이드는 mRNA 모방체일 수 있다.

[0484] 도 5는 항체에 대한 고유 식별자 서열을 포함하는 올리고뉴클레오타이드와 접합되어 있는, 예시적인 세포 성분 결합 시약, 예를 들어 항체의 개략도를 나타낸다. 세포 성분 결합 시약은 적어도 하나의 세포 성분 표적, 예컨대 항원 표적 또는 단백질 표적에 특이적으로 결합할 수 있다. 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드, 또는 항체 올리고뉴클레오타이드)는 본 개시 내용의 방법을 수행하기 위한 서열(예를 들어, 샘플 인덱싱 서열)을 포함할 수 있다. 예를 들어, 샘플 인덱싱 올리고뉴클레오타이드는 샘플의 하나 이상의 세포의 샘플 기원을 식별하기 위한 샘플 인덱싱 서열을 포함할 수 있다. 세포 성분 결합 시약을 포함하는 복수의 조성물 중 2개의 세포 성분 결합 시약을 포함하는 적어도 2개의 조성물(예를 들어, 샘플 인덱싱 조성물)의 인덱싱 서열(예를 들어, 샘플 인덱싱 서열)은 상이한 서열을 포함할 수 있다. 일부 구현예에서, 결합 시약 올리고뉴클레오타이드는 종의 유전체 서열과 상동성이 아니다. 결합 시약 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성될 수 있다.

[0485] 세포 성분 결합 시약에 접합된 올리고뉴클레오타이드는, 예를 들어 바코드 서열(예를 들어, 분자 표지 서열), 폴리(A) 테일, 또는 이들의 조합을 포함할 수 있다. 세포 성분 결합 시약에 접합된 올리고뉴클레오타이드는 mRNA 모방체일 수 있다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드 중 적어도 하나의 바코드의 포획 서열에 상보적인 서열을 포함한다. 바코드의 표적-결합 영역은 포획 서열을 포함할 수 있다. 표적-결합 영역은, 예를 들어 폴리(dT) 영역을 포함할 수 있다. 일부 구현예에서, 바코드의 포획 서열에 상보적인 샘플 인덱싱 올리고뉴클레오타이드의 서열은 폴리(A) 테일을 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지를 포함할 수 있다.

[0486] 일부 구현예에서, 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 올리고뉴클레오타이드)는 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 서열, 또는 약 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870,

880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 결합 시약 올리고뉴클레오타이드는 적어도, 또는 최대 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990 또는 1000개의 뉴클레오타이드 길이의 뉴클레오타이드 서열을 포함한다.

[0487] 일부 구현예에서, 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함한다. 결합 시약 올리고뉴클레오타이드는, 예를 들어 링커를 통해 세포 성분 결합 시약에 접합될 수 있다. 결합 시약 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 세포 성분 단백질 결합 시약의 분자에 가역적으로, 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[0488] 일부 구현예에서, 세포 성분 결합 시약은 ADAM10, CD156c, ANO6, ATP1B2, ATP1B3, BSG, CD147, CD109, CD230, CD29, CD298, ATP1B3, CD44, CD45, CD47, CD51, CD59, CD63, CD97, CD98, SLC3A2, CLDN1, HLA-ABC, ICAM1, ITFG3, MPZL1, NA K ATP아제 알파1, ATP1A1, NPTN, PMCA ATP아제e, ATP2B1, SLC1A5, SLC29A1, SLC2A1, SLC44A2, 또는 이들의 임의의 조합에 결합할 수 있다.

[0489] 일부 구현예에서, 단백질 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 일부 구현예에서, 항원 또는 단백질 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 항원 또는 단백질 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함한다. 단백질 표적은 다수의 단백질 표적을 포함하는 군으로부터 선택될 수 있다. 항원 표적 또는 단백질 표적의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 단백질 표적의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 또는 10000개일 수 있다.

[0490] 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)은 동일한 서열을 갖는 2개 이상의 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 회합될 수 있다. 세포 성분 결합 시약은 상이한 서열을 갖는 2개 이상의 결합 시약 올리고뉴클레오타이드와 회합될 수 있다. 세포 성분 결합 시약과 회합된 결합 시약 올리고뉴클레오타이드의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 동일한 서열을 갖는 상이한 서열을 갖는 어느 것이든 간에 결합 시약 올리고뉴클레오타이드의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 결합 시약 올리고뉴클레오타이드의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다.

[0491] 세포 성분 결합 시약을 포함하는 복수의 조성물(예를 들어, 복수의 샘플 인덱싱 조성물)은 결합 시약 올리고뉴클레오타이드(예컨대, 샘플 인덱싱 올리고뉴클레오타이드)와 접합되지 않은 하나 이상의 추가의 세포 성분 결합 시약을 포함할 수 있으며, 이것은 본 명세서에서 결합 시약 올리고뉴클레오타이드-부재 세포 성분 결합 시약(예컨대, 샘플 인덱싱 올리고뉴클레오타이드-부재 세포 성분 결합 시약)으로도 지칭된다. 복수의 조성물 내의 추가의 세포 성분 결합 시약의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 추가의 세포 성분 결합 시약의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 추가의 세

포 성분 결합 시약의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100개일 수 있다. 일부 구현예에서, 세포 성분 결합 시약과 임의의 추가의 세포 성분 결합 시약은 동일할 수 있다.

[0492] 일부 구현예에서, 하나 이상의 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 접합된 세포 성분 결합 시약(들) 및 결합 시약 올리고뉴클레오타이드와 접합되지 않은 세포 성분 결합 시약(들)을 포함하는 혼합물이 제공된다. 본 명세서에 개시된 방법의 일부 구현예에서, 상기 혼합물은, 예를 들어 샘플(들) 및/또는 세포(들)와 접촉하는 데 사용될 수 있다. 혼합물 내의 (1) 결합 시약 올리고뉴클레오타이드와 접합된 세포 성분 결합 시약의 수와 (2) 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드) 또는 기타 다른 결합 시약 올리고뉴클레오타이드(들)와 접합되지 않은 또 다른 세포 성분 결합 시약(예를 들어, 동일한 세포 성분 결합 시약)의 수의 비는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 비는 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 비는 적어도, 또는 최대 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 또는 1:10000일 수 있다.

[0493] 일부 구현예에서, 비는 1:1, 1.1:1, 1.2:1, 1.3:1, 1.4:1, 1.5:1, 1.6:1, 1.7:1, 1.8:1, 1.9:1, 2:1, 2.5:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, 50:1, 51:1, 52:1, 53:1, 54:1, 55:1, 56:1, 57:1, 58:1, 59:1, 60:1, 61:1, 62:1, 63:1, 64:1, 65:1, 66:1, 67:1, 68:1, 69:1, 70:1, 71:1, 72:1, 73:1, 74:1, 75:1, 76:1, 77:1, 78:1, 79:1, 80:1, 81:1, 82:1, 83:1, 84:1, 85:1, 86:1, 87:1, 88:1, 89:1, 90:1, 91:1, 92:1, 93:1, 94:1, 95:1, 96:1, 97:1, 98:1, 99:1, 100:1, 200:1, 300:1, 400:1, 500:1, 600:1, 700:1, 800:1, 900:1, 1000:1, 2000:1, 3000:1, 4000:1, 5000:1, 6000:1, 7000:1, 8000:1, 9000:1, 10000:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 1.1:1, 1.2:1, 1.3:1, 1.4:1, 1.5:1, 1.6:1, 1.7:1, 1.8:1, 1.9:1, 2:1, 2.5:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, 50:1, 51:1, 52:1, 53:1, 54:1, 55:1, 56:1, 57:1, 58:1, 59:1, 60:1, 61:1, 62:1, 63:1, 64:1, 65:1, 66:1, 67:1, 68:1, 69:1, 70:1, 71:1, 72:1, 73:1, 74:1, 75:1, 76:1, 77:1, 78:1, 79:1,

80:1, 81:1, 82:1, 83:1, 84:1, 85:1, 86:1, 87:1, 88:1, 89:1, 90:1, 91:1, 92:1, 93:1, 94:1, 95:1, 96:1, 97:1, 98:1, 99:1, 100:1, 200:1, 300:1, 400:1, 500:1, 600:1, 700:1, 800:1, 900:1, 1000:1, 2000:1, 3000:1, 4000:1, 5000:1, 6000:1, 7000:1, 8000:1, 9000:1, 10000:1, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 비는 적어도, 또는 최대 1:1, 1.1:1, 1.2:1, 1.3:1, 1.4:1, 1.5:1, 1.6:1, 1.7:1, 1.8:1, 1.9:1, 2:1, 2.5:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, 50:1, 51:1, 52:1, 53:1, 54:1, 55:1, 56:1, 57:1, 58:1, 59:1, 60:1, 61:1, 62:1, 63:1, 64:1, 65:1, 66:1, 67:1, 68:1, 69:1, 70:1, 71:1, 72:1, 73:1, 74:1, 75:1, 76:1, 77:1, 78:1, 79:1, 80:1, 81:1, 82:1, 83:1, 84:1, 85:1, 86:1, 87:1, 88:1, 89:1, 90:1, 91:1, 92:1, 93:1, 94:1, 95:1, 96:1, 97:1, 98:1, 99:1, 100:1, 200:1, 300:1, 400:1, 500:1, 600:1, 700:1, 800:1, 900:1, 1000:1, 2000:1, 3000:1, 4000:1, 5000:1, 6000:1, 7000:1, 8000:1, 9000:1, 또는 10000:1일 수 있다.

[0494] 세포 성분 결합 시약은 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 접합될 수 있거나, 그렇지 않을 수 있다. 일부 구현예에서, 결합 시약 올리고뉴클레오타이드와 접합된 세포 성분 결합 시약 및 결합 시약 올리고뉴클레오타이드와 접합되지 않은 세포 성분 결합 시약(들)을 포함하는 혼합물 내의 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 접합된 세포 성분 결합 시약의 백분율은 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 혼합물 내의 샘플 인덱싱 올리고뉴클레오타이드와 접합된 세포 성분 결합 시약의 백분율은 적어도, 또는 최대 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%일 수 있다.

[0495] 일부 구현예에서, 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 접합된 세포 성분 결합 시약 및 샘플 인덱싱 올리고뉴클레오타이드와 접합되지 않은 세포 성분 결합 시약을 포함하는 혼합물 내의 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 접합되지 않은 세포 성분 결합 시약의 백분율은 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%일 수 있다.

70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 혼합물 내의 결합 시약 올리고뉴클레오타이드와 접합되지 않은 세포 성분 결합 시약의 백분율은 적어도, 또는 최대 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%일 수 있다.

[0496] 세포 성분 콕테일

[0497] 일부 구현예에서, 본 명세서에 개시된 방법에서 표지화 감도를 증가시키기 위하여, 세포 성분 결합 시약의 콕테일(예를 들어, 항체 콕테일)이 사용될 수 있다. 어떠한 특정 이론에 의해 구제되지 않지만, 이것은, 세포 성분 발현 또는 단백질 발현이 세포 유형간에 그리고 세포 상태간에 변동될 수 있어서, 시험되는 모든 세포 유형을 표지하는 범용 세포 성분 결합 시약 또는 항체를 찾을 수 있게 하기 때문인 것으로 여겨진다. 예를 들어, 세포 성분 결합 시약의 콕테일은 더 많은 샘플 유형의 더 감도 높고 효율적인 표지화를 가능하게 하는 데 사용될 수 있다. 세포 성분 결합 시약의 콕테일은 2개 이상의 상이한 유형의 세포 성분 결합 시약, 예를 들어 더 넓은 범위의 세포 성분 결합 시약 또는 항체를 포함할 수 있다. 상이한 세포 성분 표적을 표지하는 세포 성분 결합 시약은 함께 풀링되어 모든 세포 유형 또는 관심 대상의 하나 이상의 세포 유형을 충분히 표지하는 콕테일을 생성할 수 있다.

[0498] 일부 구현예에서, 복수의 조성물(예를 들어, 샘플 인덱싱 조성물) 각각은 세포 성분 결합 시약을 포함한다. 일부 구현예에서, 복수의 조성물 중의 조성물은 2개 이상의 세포 성분 결합 시약을 포함하고, 2개 이상의 세포 성분 결합 시약 각각은 결합 시약 올리고뉴클레오타이드(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드)와 회합되고, 2개 이상의 세포 성분 결합 시약 중 적어도 하나는 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 결합 시약 올리고뉴클레오타이드의 서열은 동일할 수 있다. 2개 이상의 세포 성분 결합 시약과 회합된 결합 시약 올리고뉴클레오타이드의 서열은 상이한 서열을 포함할 수 있다. 복수의 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다.

[0499] 조성물 내의 상이한 유형의 세포 성분 결합 시약(예를 들어, CD147 항체 및 CD47 항체)의 수는 상이한 실시에서 상이할 수 있다. 2개 이상의 상이한 유형의 세포 성분 결합 시약을 갖는 조성물은 본 명세서에서 세포 성분 결합 시약 콕테일(예를 들어, 샘플 인덱싱 조성물 콕테일)로 지칭될 수 있다. 콕테일 내의 상이한 유형의 세포 성분 결합 시약의 수는 변동될 수 있다. 일부 구현예에서, 콕테일 내의 상이한 유형의 세포 성분 결합 시약의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 콕테일 내의 상이한 유형의 세포 성분 결합 시약의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 또는 100000개일 수 있다. 상이한 유형의 세포 성분 결합 시약은 동일하거나 상이한 서열(예를 들어, 샘플 인덱싱 서열)을 갖는 결합 시약 올리고뉴클레오타이드에 접합될 수 있다.

[0500] 세포 성분 표적의 정량적 분석 방법

[0501] 일부 구현예에서, 본 명세서에 개시된 방법은 또한 본 명세서에 개시된 조성물, 및 바코드 서열(예를 들어, 분자 표지 서열)을 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)의 올리고뉴클레오타이드에 회합시킬 수 있는 올리고뉴클레오타이드 프로브를 사용하여 샘플 내의 복수의 세포 성분 표적(예를 들어, 단백질 표적)의 정량적 분석에 사용될 수 있다. 세포 성분 결합 시약의 올리고뉴클레오타이드는 항체 올리고뉴클레오타이드, 샘플 인덱싱 올리고뉴클레오타이드, 세포 식별 올리고뉴클레오타이드, 제어 입자 올리고뉴클레오타이드, 제어 올리고뉴클레오타이드, 상호작용 결정 올리고뉴클레오타이드 등일 수 있거나, 이를 포함할 수 있다. 일부 구현예에서, 샘플은 단일 세포, 복수의 세포, 조직 샘플, 종양 샘플, 혈액 샘플 등일 수 있다. 일부 구현예에서, 샘플은 정상 세포, 종양 세포, 혈액 세포, B 세포, T 세포, 모체 세포, 태아 세포 등과 같은 세포 유형의 혼합물, 또는 상이한 대상체로부터의

세포의 혼합물을 포함할 수 있다.

- [0502] 일부 구현예에서, 샘플은 개별 구획, 예컨대 마이크로웰 어레이 내의 마이크로웰 내로 분리되는 복수의 단일 세포를 포함할 수 있다.
- [0503] 일부 구현예에서, 복수의 세포 성분 표적(즉, 세포 성분 표적)의 결합 표적은 탄수화물, 지질, 단백질, 세포의 단백질, 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 주요 조직적합성 복합체, 종양 항원, 수용체, 인테그린, 세포내 단백질, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 일부 구현예에서, 세포 성분 표적은 단백질 표적이다. 일부 구현예에서, 복수의 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 세포 성분을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분은 유기체 내의 모든 인코딩된 세포 성분의 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 6%, 적어도 7%, 적어도 8%, 적어도 9%, 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 적어도 98%, 적어도 99%, 또는 그 이상일 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 10, 적어도 20, 적어도 30, 적어도 40, 적어도 50, 적어도 100, 적어도 1000, 적어도 10000 개, 또는 그 이상의 상이한 세포 성분 표적을 포함할 수 있다.
- [0504] 일부 구현예에서, 복수의 세포 성분 결합 시약은 복수의 세포 성분 표적과의 특이적 결합을 위하여 샘플과 접촉된다. 결합되지 않은 세포 성분 결합 시약은, 예를 들어 세척에 의해 제거될 수 있다. 샘플이 세포를 포함하는 구현예에서, 세포에 특이적으로 결합되지 않은 어떠한 세포 성분 결합 시약도 제거될 수 있다.
- [0505] 일부 경우에, 세포의 개체군으로부터의 세포가 본 개시 내용의 기재의 웰 내로 분리(예를 들어, 분리)될 수 있다. 세포의 개체군은 분리 전에 희석될 수 있다. 세포의 개체군은 기재의 웰의 적어도 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 단일 세포를 수용하도록 희석될 수 있다. 세포의 개체군은 기재의 웰의 최대 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 단일 세포를 수용하도록 희석될 수 있다. 세포의 개체군은 희석된 개체군 내의 세포의 수가 기재 상의 웰의 수의 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 또는 100%, 또는 적어도 이들 값이 되도록 희석될 수 있다. 세포의 개체군은 희석된 개체군 내의 세포의 수가 기재 상의 웰의 수의 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 또는 100%, 또는 적어도 이들 값이 되도록 희석될 수 있다. 일부 경우에, 세포의 개체군은 세포의 수가 기재 내의 웰의 수의 약 10%가 되도록 희석된다.
- [0506] 기재의 웰 내로의 단일 세포의 분포는 푸와송 분포를 따를 수 있다. 예를 들어, 기재의 웰이 하나 초과인 세포를 가질 확률은 적어도 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 또는 10%, 또는 그 이상일 수 있다. 기재의 웰이 하나 초과인 세포를 가질 확률은 적어도 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 또는 10%, 또는 그 이상일 수 있다. 기재의 웰 내로의 단일 세포의 분포는 랜덤일 수 있다. 기재의 웰 내로의 단일 세포의 분포는 비-랜덤일 수 있다. 세포는 기재의 웰이 하나의 세포만 수용하도록 분리될 수 있다.
- [0507] 일부 구현예에서, 세포 성분 결합 시약은 형광 분자와 추가로 접합되어 개별 구획들 내로의 세포들의 유동 분류를 가능하게 할 수 있다.
- [0508] 일부 구현예에서, 본 명세서에 개시된 방법은 복수의 조성물을 복수의 세포 성분 표적과의 특이적 결합을 위하여 샘플과 접촉시키는 단계를 제공한다. 사용되는 조건은 세포 성분 결합 시약, 예를 들어 항체를 세포 성분 표적에 특이적으로 결합시킬 수 있음이 이해될 것이다. 접촉 단계 후에, 결합되지 않은 조성물은 제거될 수 있다. 예를 들어, 샘플이 세포를 포함하고, 조성물이 세포 성분 표적, 즉 세포-표면 세포 성분, 예컨대 세포-표면 단백질에 특이적으로 결합하는 구현예에서, 세포를 완충액으로 세척함으로써 결합되지 않은 조성물을 제거하여, 세포 성분 표적에 특이적으로 결합하는 조성물만이 세포와 함께 남아 있도록 할 수 있다.
- [0509] 일부 구현예에서, 본 명세서에 개시된 방법은 바코드 서열(예컨대, 분자 표지), 세포 표지, 샘플 표지 등, 또는 이들의 임의의 조합을 포함하는 올리고뉴클레오타이드(예를 들어, 바코드 또는 추계적 바코드)를 세포 성분 결합 시약과 회합된 복수의 올리고뉴클레오타이드에 회합시키는 단계를 포함할 수 있다. 예를 들어, 바코드를 포함하는 복수의 올리고뉴클레오타이드 프로브가 본 조성물의 복수의 올리고뉴클레오타이드에 대한 혼성화에 사용될 수 있다.
- [0510] 일부 구현예에서, 복수의 올리고뉴클레오타이드 프로브는 고체 지지체 상에 고정화될 수 있다. 고체 지지체는 용

액 중에서 자유롭게 부유하는, 예를 들어 비드일 수 있다. 고체 지지체는 반고체 또는 고체 어레이 내에 매립될 수 있다. 일부 구현예에서, 복수의 올리고뉴클레오타이드 프로브는 고체 지지체 상에 고정화되지 않을 수 있다. 복수의 올리고뉴클레오타이드 프로브가 세포 성분 결합 시약의 복수의 회합된 올리고뉴클레오타이드에 근접해 있는 경우, 세포 성분 결합 시약의 복수의 올리고뉴클레오타이드는 올리고뉴클레오타이드 프로브에 혼성화될 수 있다. 올리고뉴클레오타이드 프로브는 세포 성분 결합 시약의 각각의 별개의 올리고뉴클레오타이드가 본 개시 내용의 상이한 바코드 서열(예를 들어, 분자 표지)을 갖는 올리고뉴클레오타이드 프로브와 회합될 수 있도록 비-고갈성 비로 접촉될 수 있다.

[0511] 일부 구현예에서, 본 명세서에 개시된 방법은 세포 성분 표적에 특이적으로 결합된 세포 성분 결합 시약으로부터 올리고뉴클레오타이드를 탈착시키는 단계를 제공한다. 탈착은 세포 성분 결합 시약으로부터 화학 기를 분리하기 위한 다양한 방법, 예컨대 UV 광절단, 화학적 처리(예를 들어, 디티오트레이톨 처리), 가열, 효소 처리, 또는 이들의 임의의 조합으로 수행될 수 있다. 세포 성분 결합 시약으로부터 올리고뉴클레오타이드를 탈착시키는 단계는 복수의 올리고뉴클레오타이드 프로브를 본 조성물의 복수의 올리고뉴클레오타이드에 혼성화하는 단계 전에, 후에 또는 동안 중 어느 하나에서 수행될 수 있다.

[0512] 세포 성분 표적과 핵산 표적의 동시적인 정량적 분석 방법

[0513] 일부 구현예에서, 본 명세서에 개시된 방법은 또한 본 명세서에 개시된 조성물, 및 바코드 서열(예를 들어, 분자 표지 서열)을 세포 성분 결합 시약의 올리고뉴클레오타이드 및 핵산 표적 분자 둘 모두에 회합시킬 수 있는 올리고뉴클레오타이드 프로브를 사용하여 샘플 내의 복수의 세포 성분 표적(예를 들어, 단백질 표적)과 복수의 핵산 표적 분자의 동시적인 정량적 분석에 사용될 수 있다. 복수의 세포 성분 표적과 복수의 핵산 표적 분자의 동시적인 정량적 분석의 기타 다른 방법이 2017년 9월 25일에 출원된 미국 특허 출원 15/715028에 기재되어 있으며; 이의 내용은 전체가 본 명세서에 참고로 포함된다. 일부 구현예에서, 샘플은 단일 세포, 복수의 세포, 조직 샘플, 종양 샘플, 혈액 샘플 등일 수 있다. 일부 구현예에서, 샘플은 정상 세포, 종양 세포, 혈액 세포, B 세포, T 세포, 모체 세포, 태아 세포와 같은 세포 유형의 혼합물, 또는 상이한 대상체로부터의 세포의 혼합물을 포함할 수 있다.

[0514] 일부 구현예에서, 샘플은 개별 구획, 예컨대 마이크로웰 어레이 내의 마이크로웰 내로 분리되는 복수의 단일 세포를 포함할 수 있다.

[0515] 일부 구현예에서, 복수의 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 세포 성분을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분은 유기체 또는 유기체의 하나 이상의 세포 내의 모든 세포 성분, 예컨대 발현된 단백질의 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 세포 성분은 유기체 또는 유기체의 하나 이상의 세포 내의 모든 세포 성분, 예컨대 발현될 수 있는 단백질의 적어도, 또는 최대 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 또는 99%일 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있거나, 약 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 또는 10000개의 상이한 세포 성분 표적을 포함할 수 있다.

[0516] 일부 구현예에서, 복수의 세포 성분 결합 시약은 복수의 세포 성분 표적과의 특이적 결합을 위해 샘플과 접촉된다. 미결합 세포 성분 결합 시약은, 예를 들어 세척에 의해 제거될 수 있다. 샘플이 세포를 포함하는 구현예에서, 세포에 특이적으로 결합되지 않은 임의의 세포 성분 결합 시약은 제거될 수 있다.

[0517] 일부 경우에서, 세포의 개체군으로부터의 세포는 본 개시 내용의 기재의 웰 내로 분리(예를 들어, 단리)될 수 있다. 세포의 개체군은 분리 전에 희석될 수 있다. 세포의 개체군은, 기재의 웰의 적어도 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 단일 세포를 수용하도록 희석될 수 있다. 세포의 개체군은, 기재의 웰의 최대 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 단일 세포를 수용하도록 희석될 수

있다. 세포의 개체군은, 희석된 개체군에서 세포의 수가 기재 상의 웰의 수의 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 되거나, 적어도 이들 값이 되도록 희석될 수 있다. 세포의 개체군은, 희석된 개체군에서 세포의 수가 기재 상의 웰의 수의 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%가 되거나, 적어도 이들 값이 되도록 희석될 수 있다. 일부 경우에서, 세포의 개체군은 세포의 수가 기재 내 웰의 수의 약 10%가 되도록 희석된다.

[0518] 단일 세포가 기재의 웰 내로 분포되는 것은 푸와송 분포를 따를 수 있다. 예를 들어, 기재의 웰이 하나 초과인 세포를 갖는 확률은 적어도 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% 또는 10%, 또는 그 이상일 수 있다. 기재의 웰이 하나 초과인 세포를 갖는 확률은 적어도 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% 또는 10%, 또는 그 이상일 수 있다. 기재의 웰 내로의 단일 세포의 분포는 랜덤일 수 있다. 기재의 웰 내로의 단일 세포의 분포는 비-랜덤일 수 있다. 세포는 기재의 웰이 하나의 세포만 수용하도록 분리될 수 있다.

[0519] 일부 구현예에서, 세포 성분 결합 시약은 형광 분자와 추가적으로 접합되어 개별적인 구획으로의 세포의 유동 분류를 가능하게 할 수 있다.

[0520] 일부 구현예에서, 본 명세서에 개시된 방법은 복수의 조성물들을 복수의 세포 성분 표적과 특이적으로 결합시키기 위해 샘플과 접촉시키는 단계를 제공한다. 사용되는 조건은 세포 성분 결합 시약, 예를 들어 항체를 세포 성분 표적에 특이적으로 결합시킬 수 있음이 이해될 것이다. 접촉 단계 후에, 결합되지 않은 조성물은 제거될 수 있다. 예를 들어, 샘플이 세포를 포함하고, 조성물이 세포 성분 표적, 즉 세포-표면 세포 성분, 예컨대 세포-표면 단백질에 특이적으로 결합하는 구현예에서, 세포를 완충액으로 세척함으로써 결합되지 않은 조성물을 제거하여, 세포 성분 표적에 특이적으로 결합하는 조성물만이 세포와 함께 남아 있도록 할 수 있다.

[0521] 일부 구현예에서, 본 명세서에 개시된 방법은 샘플, 예를 들어 세포로부터의 복수의 핵산 표적 분자를 방출하는 단계를 제공할 수 있다. 예를 들어, 세포는 용해되어 복수의 핵산 표적 분자를 방출할 수 있다. 세포 용해는 다양한 수단 중 임의의 것, 예를 들어 화학적 처리, 삼투압 충격, 열 처리, 기계적 처리, 광학 처리 또는 이들의 임의의 조합에 의해 달성될 수 있다. 세포는 세제(예를 들어, SDS, Li 도데실 설페이트, 트리톤 X-100, 트윈-20, 또는 NP-40), 유기 용매(예를 들어, 메탄올 또는 아세톤), 또는 소화 효소(예를 들어, 프로테이나제 K, 펩신, 또는 트립신), 또는 이들의 임의의 조합을 포함하는 세포 용해 완충액의 첨가에 의해 용해될 수 있다.

[0522] 복수의 핵산 분자가 다양한 핵산 분자를 포함할 수 있음은 본 기술분야의 당업자는 이해할 것이다. 일부 구현예에서, 복수의 핵산 분자는 DNA 분자, RNA 분자, 유전체 DNA 분자, mRNA 분자, rRNA 분자, siRNA 분자, 또는 이들의 조합을 포함할 수 있고, 이중 가닥 또는 단일 가닥일 수 있다. 일부 구현예에서, 복수의 핵산 분자는 100, 1000, 10000, 20000, 30000, 40000, 50000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 종을 포함하거나, 약 100, 1000, 10000, 20000, 30000, 40000, 50000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 종을 포함한다. 일부 구현예에서, 복수의 핵산 분자는 적어도, 또는 최대 100, 1000, 10000, 20000, 30000, 40000, 50000, 100000, 또는 1000000개의 종을 포함한다. 일부 구현예에서, 복수의 핵산 분자는 샘플, 예컨대 단일 세포, 또는 복수의 세포로부터의 것일 수 있다. 일부 구현예에서, 복수의 핵산 분자는 복수의 샘플, 예컨대 복수의 단일 세포로부터 풀링될 수 있다.

[0523] 일부 구현예에서, 본 명세서에서 개시된 방법은 바코드 서열(예컨대, 분자 표지), 세포 표지, 샘플 표지 등 또는 이들의 임의의 조합을 포함할 수 있는 바코드(예를 들어, 추계적 바코드)를, 복수의 핵산 표적 분자, 및 세포 성분 결합 시약의 복수의 올리고뉴클레오타이드에 회합하는 단계를 포함할 수 있다. 예를 들어, 추계적 바코드를 포함하는 복수의 올리고뉴클레오타이드 프로브는, 복수의 핵산 표적 분자 및 조성물의 복수의 올리고뉴클레오타이드에 혼성화하는 데 사용될 수 있다.

[0524] 일부 구현예에서, 복수의 올리고뉴클레오타이드 프로브는 고체 지지체 상에 고정될 수 있다. 고체 지지체는 용액 내에서 자유롭게 부유하는, 예를 들어 비드일 수 있다. 고체 지지체는 반-고체 또는 고체 어레이에 매립될 수 있다. 일부 구현예에서, 복수의 올리고뉴클레오타이드 프로브는 고체 지지체 상에 고정화되지 않을 수 있다. 복수의 올리고뉴클레오타이드 프로브가 복수의 핵산 표적 분자 및 세포 성분 결합 시약의 복수의 올리고뉴클레오타이드와 근접하여 있는 경우, 복수의 핵산 표적 분자 및 세포 성분 결합 시약의 복수의 올리고뉴클레오타이드는 올리고뉴클레오타이드 프로브에 혼성화될 수 있다. 올리고뉴클레오타이드 프로브는 각각의 별개의 핵산 표적 분자 및 세포 성분 결합 시약의 올리고뉴클레오타이드가 본 개시 내용의 상이한 바코드 서열(예를 들어, 분자 표지)을 갖는 올리고뉴클레오타이드 프로브와 회합될 수 있도록 비-고갈성 비로 접촉될 수 있다.

[0525] 일부 구현예에서, 본 명세서에 개시된 방법은 세포 성분 표적에 특이적으로 결합된 세포 성분 결합 시약으로부터 올리고뉴클레오타이드를 탈착시키는 단계를 제공한다. 탈착은 세포 성분 결합 시약으로부터 화학 기를 분리하기 위한 다양한 방법, 예컨대 UV 광절단, 화학적 처리(예를 들어, 디티오프레이틀 처리), 가열, 효소 처리, 또는 이들의 임의의 조합으로 수행될 수 있다. 세포 성분 결합 시약으로부터 올리고뉴클레오타이드를 탈착시키는 단계는 복수의 올리고뉴클레오타이드 프로브를 본 조성물의 복수의 핵산 표적 분자 및 복수의 올리고뉴클레오타이드에 혼성화하는 단계 전에, 후에 또는 동안 중 어느 하나에서 수행될 수 있다.

[0526] 단백질 표적과 핵산 표적의 동시적인 정량적 분석

[0527] 일부 구현예에서, 본 명세서에 개시된 방법은 또한 다수의 유형의 표적 분자, 예를 들어 단백질 표적과 핵산 표적의 동시적인 정량적 분석에 사용될 수 있다. 예를 들어, 표적 분자는 세포 성분일 수 있거나, 이를 포함할 수 있다. 도 6은 단일 세포 내의 핵산 표적과 기타 다른 세포 성분 표적(예를 들어, 단백질) 둘 모두의 동시적인 정량적 분석의 예시적인 방법의 개략도를 나타낸다. 일부 구현예에서, 각각 세포 성분 결합 시약, 예컨대 항체를 포함하는 복수의 조성물(605, 605b, 605c 등)이 제공된다. 상이한 세포 성분 표적에 결합하는 상이한 세포 성분 결합 시약, 예컨대 항체가 상이한 고유 식별자와 접합된다. 다음으로, 세포 성분 결합 시약은 복수의 세포(610)를 함유하는 샘플과 인큐베이션될 수 있다. 상이한 세포 성분 결합 시약은 세포 표면 상의 세포 성분, 예컨대 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합에 특이적으로 결합할 수 있다. 결합되지 않은 세포 성분 결합 시약은, 예를 들어 세포를 완충액으로 세척함으로써 제거될 수 있다. 이어서, 세포 성분 결합 시약을 갖는 세포는 복수의 구획, 예컨대 마이크로웰 어레이 내로 분리될 수 있으며, 여기서 단일 구획(615)은 단일 세포 및 단일 비드(620)에 맞도록 크기조정된다. 각각의 비드는 복수의 올리고뉴클레오타이드 프로브(이것은 비드 상의 모든 올리고뉴클레오타이드 프로브에 공통된 세포 표지를 포함할 수 있음), 및 바코드 서열(예를 들어, 분자 표지 서열)을 포함할 수 있다. 일부 구현예에서, 각각의 올리고뉴클레오타이드 프로브는 표적-결합 영역, 예를 들어 폴리(dT) 서열을 포함할 수 있다. 세포 성분 결합 시약에 접합된 올리고뉴클레오타이드(625)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 세포 성분 결합 시약으로부터 탈착될 수 있다. 세포는 용해되어(635) 세포 내의 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(630)를 방출할 수 있다. 세포성 mRNA(630), 올리고뉴클레오타이드(625) 또는 둘 모두는, 예를 들어 폴리(dT) 서열에 혼성화함으로써, 비드(620) 상의 올리고뉴클레오타이드 프로브에 의해 포획될 수 있다. 세포성 mRNA(630) 및 올리고뉴클레오타이드(625)를 주형으로서 사용하여 세포성 mRNA(630) 및 올리고뉴클레오타이드(625)에 혼성화된 올리고뉴클레오타이드 프로브를 연장하기 위해 역전사효소가 사용될 수 있다. 역전사효소에 의해 생성된 연장 산물은 증폭되고 서열분석될 수 있다. 서열분석 판독물은 세포 표지, 바코드(예를 들어, 분자 표지), 유전자, 세포 성분 결합 시약 특이적 올리고뉴클레오타이드(예를 들어, 항체 특이적 올리고뉴클레오타이드) 등의 서열 또는 식별의 역다중화를 거칠 수 있으며, 이는 샘플 내의 각각의 단일 세포의 세포 성분 및 유전자 발현의 디지털 표현을 생성할 수 있다.

[0528] 바코드의 회합

[0529] 세포 성분 결합 시약(예를 들어, 항원 결합 시약 또는 단백질 결합 시약)과 회합된 올리고뉴클레오타이드 및/또는 핵산 분자는 올리고뉴클레오타이드 프로브와 랜덤하게 회합될 수 있다. 본 명세서에서 결합 시약 올리고뉴클레오타이드로 지칭되는, 세포 성분 결합 시약과 회합된 올리고뉴클레오타이드는 본 개시 내용의 올리고뉴클레오타이드, 예컨대 항체 올리고뉴클레오타이드, 샘플 인덱싱 올리고뉴클레오타이드, 세포 식별 올리고뉴클레오타이드, 제어 입자 올리고뉴클레오타이드, 제어 올리고뉴클레오타이드, 상호작용 결정 올리고뉴클레오타이드 등일 수 있거나, 이를 포함할 수 있다. 회합은, 예를 들어 올리고뉴클레오타이드 프로브의 표적-결합 영역을, 표적 핵산 분자의 상보적 부분 및/또는 단백질 결합 시약의 올리고뉴클레오타이드로 혼성화하는 것을 포함할 수 있다. 예를 들어, 바코드(예를 들어, 추적적 바코드)의 올리고(dT) 영역은 표적 핵산 분자의 폴리(A) 테일 및/또는 단백질 결합 시약의 올리고뉴클레오타이드의 폴리(A) 테일과 상호작용할 수 있다. 혼성화에 사용되는 분석 조건(예를 들어, 완충액 pH, 이온 강도, 온도 등)은 특정한 안정적인 혼성화물의 형성을 촉진하도록 선택될 수 있다.

[0530] 본 개시 내용은 분자 표지를, 역전사를 사용하여 표적 핵산 및/또는 세포 성분 결합 시약과 회합된 올리고뉴클레오타이드와 회합시키는 방법을 제공한다. 역전사효소는 주형으로서 RNA 및 DNA를 둘 모두 사용할 수 있다. 예를 들어, 세포 성분 결합 시약 상에 원래 접합된 올리고뉴클레오타이드는 RNA 또는 DNA 염기 중 어느 하나, 또는 둘 모두일 수 있다. 결합 시약 올리고뉴클레오타이드가 복제되고, 결합 시약 서열의 서열 또는 이의 일부분에 더하여, 세포 표지 및 바코드 서열(예를 들어, 분자 표지)에 연결될(예를 들어, 공유적으로 연결될) 수 있다. 또 다른 예로서, mRNA 분자가 복제되고, mRNA 분자의 서열 또는 이의 일부분에 더하여, 세포 표지 및 바코드 서

열(예를 들어, 분자 표지)에 연결될(예를 들어, 공유적으로 연결될) 수 있다.

[0531] 일부 구현예에서, 분자 표지는 올리고뉴클레오타이드 프로브 표적-결합 영역과, 표적 핵산 분자 및/또는 세포 성분 결합 시약과 회합된(예를 들어, 현재 또는 이전에 회합된) 올리고뉴클레오타이드의 일부분의 결합에 의해 부가될 수 있다. 예를 들어, 표적-결합 영역은 제한 부위 오버행(예를 들어, EcoRI 점착성-말단 오버행)에 특이적 혼성화가 가능할 수 있는 핵산 서열을 포함할 수 있다. 방법은 표적 핵산 및/또는 세포 성분 결합 시약과 회합된 올리고뉴클레오타이드를 제한 효소(예를 들어, EcoRI)로 처리하여 제한 부위 오버행을 생성하는 단계를 추가로 포함할 수 있다. 이들 2개의 단편을 연결하기 위해 리가제(예를 들어, T4 DNA 리가제)가 사용될 수 있다.

[0532] 고유 분자 표지 서열의 수 또는 존재의 결정

[0533] 일부 구현예에서, 본 명세서에 개시된 방법은 각각의 고유 식별자, 각각의 핵산 표적 분자, 및/또는 각각의 결합 시약 올리고뉴클레오타이드(예를 들어, 항체 올리고뉴클레오타이드)에 대한 고유 분자 표지 서열의 수 또는 존재를 결정하는 단계를 포함한다. 예를 들어, 서열분석 판독물은 각각의 고유 식별자, 각각의 핵산 표적 분자, 및/또는 각각의 결합 시약 올리고뉴클레오타이드에 대한 고유 분자 표지 서열의 수를 결정하는 데 사용될 수 있다. 또 다른 예로서, 서열분석 판독물은 분자 표지 서열(예컨대, 서열분석 판독물에서 표적, 결합 시약 올리고뉴클레오타이드, 항체 올리고뉴클레오타이드, 샘플 인덱싱 올리고뉴클레오타이드, 세포 식별 올리고뉴클레오타이드, 제어 입자 올리고뉴클레오타이드, 제어 올리고뉴클레오타이드, 상호작용 결정 올리고뉴클레오타이드 등과 회합된 분자 표지 서열 등)의 존재 또는 부재를 결정하는 데 사용될 수 있다.

[0534] 일부 구현예에서, 각각의 고유 식별자, 각각의 핵산 표적 분자, 및/또는 각각의 결합 시약 올리고뉴클레오타이드에 대한 고유 분자 표지 서열의 수는 샘플 내의 각각의 세포 성분 표적(예를 들어, 항원 표적 또는 단백질 표적) 및/또는 각각의 핵산 표적 분자의 양을 나타낸다. 일부 구현예에서, 세포 성분 표적의 양 및 그의 상응하는 핵산 표적 분자, 예를 들어 mRNA 분자의 양은 서로 비교될 수 있다. 일부 구현예에서, 세포 성분 표적의 양과 그의 상응하는 핵산 표적 분자, 예를 들어 mRNA 분자의 양의 비가 계산될 수 있다. 세포 성분은, 예를 들어 세포 표면 단백질 마커일 수 있다. 일부 구현예에서, 세포 표면 단백질 마커의 단백질 수준과 세포 표면 단백질 마커의 mRNA의 수준 사이의 비는 낮다.

[0535] 본 명세서에 개시된 방법은 다양한 적용 분야에 사용될 수 있다. 예를 들어, 본 명세서에 개시된 방법은 샘플의 단백질체 및/또는 전사체 분석에 사용될 수 있다. 일부 구현예에서, 본 명세서에 개시된 방법은 샘플에서 세포 성분 표적 및/또는 핵산 표적, 즉 바이오마커를 식별하는 데 사용될 수 있다. 일부 구현예에서, 세포 성분 표적 및 핵산 표적은 서로 상응하며, 즉 핵산 표적은 세포 성분 표적을 인코딩한다. 일부 구현예에서, 본 명세서에 개시된 방법은 샘플, 예를 들어 mRNA 분자 내에서 세포 성분 표적의 양 및 그의 상응하는 핵산 표적 분자의 양 사이의 원하는 비를 갖는 세포 성분 표적을 식별하는 데 사용될 수 있다. 일부 구현예에서, 그 비는 0.001, 0.01, 0.1, 1, 10, 100, 1000, 또는 상기 값들 중 임의의 두 값 사이의 어떤 수 또는 범위이거나, 약 0.001, 0.01, 0.1, 1, 10, 100, 1000, 또는 상기 값들 중 임의의 두 값 사이의 어떤 수 또는 범위이다. 일부 구현예에서, 그 비는 적어도, 또는 최대 0.001, 0.01, 0.1, 1, 10, 100, 또는 1000이다. 일부 구현예에서, 본 명세서에 개시된 방법은, 샘플 내의 상응하는 핵산 표적 분자의 양이 1000, 100, 10, 5, 2, 1, 0, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위이거나, 약 1000, 100, 10, 5, 2, 1, 0, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위인 샘플에서 세포 성분 표적을 식별하는 데 사용될 수 있다. 일부 구현예에서, 본 명세서에 개시된 방법은, 샘플 내의 상응하는 핵산 표적 분자의 양이 1000, 100, 10, 5, 2, 1, 또는 0 초과 또는 미만인 샘플에서 세포 성분 표적을 식별하는 데 사용될 수 있다.

[0536] 조성물 및 키트

[0537] 본 명세서에 개시된 일부 구현예는 샘플 내의 복수의 세포 성분(예를 들어, 단백질) 및/또는 복수의 핵산 표적 분자의 동시적인 정량적 분석을 위한 키트 및 조성물을 제공한다. 일부 구현예에서, 키트 및 조성물은, 각각 올리고뉴클레오타이드와 집합된 복수의 세포 성분 결합 시약(예를 들어, 복수의 단백질 결합 시약)(여기서, 올리고뉴클레오타이드는 세포 성분 결합 시약에 대한 고유 식별자를 포함함), 및 각각 표적-결합 영역, 바코드 서열(예를 들어, 분자 표지 서열)을 포함하는 복수의 올리고뉴클레오타이드 프로브(여기서, 바코드 서열은 다양한 고유 바코드 서열 세트로부터 유래됨)를 포함할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 분자 표지, 세포 표지, 샘플 표지, 또는 이들의 임의의 조합을 포함할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 링커를 포함할 수 있다. 일부 구현예에서, 올리고뉴클레오타이드 각각은 올리고뉴클레오타이드 프로브에 대한 결합 부위, 예컨대 폴리(A) 테일을 포함할 수 있다. 예를 들어, 폴리(A) 테일은, 예를 들어 올리고dA₁₈(고체 지지

체에 고착되지 않음) 또는 올리고 $A_{18}V$ (고체 지지체에 고착됨)일 수 있다. 올리고뉴클레오타이드는 DNA 잔기, RNA 잔기, 또는 둘 모두를 포함할 수 있다.

[0538] 본 명세서의 개시 내용은 복수의 샘플 인덱싱 조성물을 포함한다. 복수의 샘플 인덱싱 조성물 각각은 2개 이상의 세포 성분 결합 시약을 포함할 수 있다. 2개 이상의 세포 성분 결합 시약 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 2개 이상의 세포 성분 결합 시약 중 적어도 하나는 적어도 하나의 세포 성분 표적에 특이적으로 결합할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드는 샘플의 하나 이상의 세포의 샘플 기원을 식별하기 위한 샘플 인덱싱 서열을 포함할 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함할 수 있다.

[0539] 본 명세서의 개시 내용은 제어 입자 조성물을 포함하는 키트를 포함한다. 일부 구현예에서, 제어 입자 조성물은 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함하고, 복수의 제어 입자 올리고뉴클레오타이드 각각은 제어 바코드 서열 및 폴리(dA) 영역을 포함한다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2개는 상이한 제어 바코드 서열을 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 분자 표지 서열을 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 범용 프라이머를 위한 결합 부위를 포함할 수 있다. 또한, 본 명세서의 개시 내용은 서열분석 제어를 위한 키트를 포함한다. 일부 구현예에서, 키트는 제어 입자 조성물을 포함하며, 상기 제어 입자 조성물은 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함하고, 복수의 제어 입자 올리고뉴클레오타이드 각각은 제어 바코드 서열 및 폴리(dA) 영역을 포함한다.

[0540] 본 명세서의 개시 내용은 서열분석 제어를 위한 복수의 제어 조성물을 포함하는 키트를 포함한다. 일부 구현예에서, 복수의 제어 조성물 각각은 제어 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 복수의 결합 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 제어 올리고뉴클레오타이드는 제어 바코드 서열, 및 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적 영역을 포함한다. 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다.

[0541] 본 명세서의 개시 내용은 세포 식별을 위한 샘플 인덱싱 조성물을 포함하는 키트를 포함한다. 일부 구현예에서, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적(예를 들어, 하나 이상의 단백질 표적) 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함한다.

[0542] 본 명세서의 개시 내용은 세포 식별을 위한 키트를 포함한다. 일부 구현예에서, 키트는 2개 이상의 샘플 인덱싱 조성물을 포함한다. 2개 이상의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약(예를 들어, 항원 결합 시약)을 포함할 수 있고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함한다. 본 명세서의 개시 내용은 멀티플렉스 식별을 위한 키트를 포함한다. 일부 구현예에서, 키트는 2개의 샘플 인덱싱 조성물을 포함한다. 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약(예를 들어, 항원 결합 시약)을 포함할 수 있고, 항원 결합 시약은 하나 이상의 세포 성분 표적(예를 들어, 항원 표적) 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함한다.

[0543] 본 명세서의 개시 내용은 세포 성분 사이의 상호작용, 예를 들어 단백질-단백질 상호작용을 식별하기 위한 키트를 포함한다. 일부 구현예에서, 키트는 제1 쌍의 상호작용 결정 조성물을 포함하며, 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 제1 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약은 제1 단백질 표적에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약은 제2 단백질 표적에 특이적으로 결합할 수 있고, 상호작용 결정 올리고뉴클레오타이드는 상호작용 결정 서열 및 가교 올리고뉴클레오타이드 혼성화 영역을 포함하고, 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열; 및 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 각각 포함하는 복수의 가교 올리고뉴클레오타이드를 포함한다.

[0544] 고유 식별자(또는 세포 성분 결합 시약과 회합된 올리고뉴클레오타이드, 예컨대 결합 시약 올리고뉴클레오타이드,

항체 올리고뉴클레오타이드, 샘플 인덱싱 올리고뉴클레오타이드, 세포 식별 올리고뉴클레오타이드, 제어 입자 올리고뉴클레오타이드, 제어 올리고뉴클레오타이드, 또는 상호작용 결정 올리고뉴클레오타이드)는 임의의 적합한 길이, 예를 들어 약 25개의 뉴클레오타이드 내지 약 45개의 뉴클레오타이드 길이를 가질 수 있다. 일부 구현예에서, 고유 식별자는 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드, 15개의 뉴클레오타이드, 20개의 뉴클레오타이드, 25개의 뉴클레오타이드, 30개의 뉴클레오타이드, 35개의 뉴클레오타이드, 40개의 뉴클레오타이드, 45개의 뉴클레오타이드, 50개의 뉴클레오타이드, 55개의 뉴클레오타이드, 60개의 뉴클레오타이드, 70개의 뉴클레오타이드, 80개의 뉴클레오타이드, 90개의 뉴클레오타이드, 100개의 뉴클레오타이드, 200개의 뉴클레오타이드이거나, 대략 이들 값이거나, 이들 값 미만이거나, 이들 값 초과이거나, 또는 상기 값들 중 임의의 두 값 사이의 범위의 길이를 가질 수 있다.

[0545] 일부 구현예에서, 고유 식별자는 고유 식별자의 다양한 세트로부터 선택된다. 고유 식별자의 다양한 세트는 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 고유 식별자를 포함할 수 있거나, 약 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 고유 식별자를 포함할 수 있다. 고유 식별자의 다양한 세트는 적어도, 또는 최대 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 또는 5000개의 상이한 고유 식별자를 포함할 수 있다. 일부 구현예에서, 고유 식별자의 세트는 분석될 샘플의 DNA 또는 RNA 서열과의 최소 서열 상동성을 갖도록 설계된다. 일부 구현예에서, 고유 식별자의 세트의 서열은 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼, 또는 약 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 10개의 뉴클레오타이드만큼, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위만큼 서로 상이하거나, 이들의 상보체이다. 일부 구현예에서, 고유 식별자의 세트의 서열은 적어도, 또는 최대 1개의 뉴클레오타이드, 2개의 뉴클레오타이드, 3개의 뉴클레오타이드, 4개의 뉴클레오타이드, 5개의 뉴클레오타이드, 6개의 뉴클레오타이드, 7개의 뉴클레오타이드, 8개의 뉴클레오타이드, 9개의 뉴클레오타이드, 또는 10개의 뉴클레오타이드만큼 서로 상이하거나, 이들의 상보체이다.

[0546] 일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머를 위한 결합 부위를 포함할 수 있다. 일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머를 위한 적어도 2개의 결합 부위를 포함할 수 있다. 일부 구현예에서, 고유 식별자는 프라이머, 예컨대 범용 프라이머를 위한 적어도 3개의 결합 부위를 포함할 수 있다. 프라이머는, 예를 들어 PCR 증폭에 의한 고유 식별자의 증폭을 위해 사용될 수 있다. 일부 구현예에서, 프라이머는 네스티드 PCR 반응에 사용될 수 있다.

[0547] 임의의 적합한 세포 성분 결합 시약, 예컨대 임의의 단백질 결합 시약(예를 들어, 항체 또는 이들의 단편, 압타머, 소분자, 리간드, 펩티드, 올리고뉴클레오타이드 등 또는 이의 임의의 조합이 본 개시 내용에서 고려된다. 일부 구현예에서, 세포 성분 결합 시약은 다클론성 항체, 단클론성 항체, 재조합 항체, 단일-사슬 항체(scAb), 또는 이들의 단편, 예컨대 Fab, Fv 등일 수 있다. 일부 구현예에서, 복수의 단백질 결합 시약은 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 단백질 결합 시약을 포함할 수 있거나, 약 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 5000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 단백질 결합 시약을 포함할 수 있다. 일부 구현예에서, 복수의 단백질 결합 시약은 적어도, 또는 최대 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 또는 5000개의 상이한 단백질 결합 시약을 포함할 수 있다.

[0548] 일부 구현예에서, 올리고뉴클레오타이드는 링커를 통해 세포 성분 결합 시약과 접합된다. 일부 구현예에서, 올리고뉴클레오타이드는 단백질 결합 시약과 공유결합적으로 접합될 수 있다. 일부 구현예에서, 올리고뉴클레오타이드는 단백질 결합 시약과 비-공유결합적으로 접합될 수 있다. 일부 구현예에서, 링커는 올리고뉴클레오타이드를 단백질 결합 시약에 가역적으로 또는 비가역적으로 부착시키는 화학 기를 포함할 수 있다. 화학 기는, 예를 들어 아민 기를 통하여, 링커에 접합될 수 있다. 일부 구현예에서, 링커는 단백질 결합 시약에 접합되는 또 다른 화학 기와 안정적인 결합을 형성하는 화학 기를 포함할 수 있다. 예를 들어, 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민 등일 수 있다. 일부 구현예에서, 화학 기는 아미노산, 예컨대 리신 상에서 1차 아민, 또는 N-말단을 통해 단백질 결합 시약에 접합될 수 있다. 올리고뉴클레오타이드는, 단백질 결합 시약과 그의

단백질 표적 사이의 특이적 결합을 방해하지 않는 한, 단백질 결합 시약의 임의의 적합한 부위에 접합될 수 있다. 단백질 결합 시약이 항체인 구현예에서, 올리고뉴클레오타이드는 항원-결합 부위 외의 임의의 위치, 예를 들어, Fc 영역, C_H1 도메인, C_H2 도메인, C_H3 도메인, C_L 도메인 등에서 항체에 접합될 수 있다. 일부 구현예에서, 각각의 단백질 결합 시약은 단일 올리고뉴클레오타이드 분자와 접합될 수 있다. 일부 구현예에서, 각각의 단백질 결합 시약은 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 올리고뉴클레오타이드 분자와 접합될 수 있거나, 약 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 올리고뉴클레오타이드 분자와 접합될 수 있으며, 여기서 각각의 올리고뉴클레오타이드 분자는 동일한 고유 식별자를 포함한다. 일부 구현예에서, 각각의 단백질 결합 시약은 하나 초과와 올리고뉴클레오타이드 분자, 예를 들어, 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 또는 1000개의 올리고뉴클레오타이드 분자와 접합될 수 있으며, 여기서 올리고뉴클레오타이드 분자 각각은 동일한 고유 식별자를 포함한다.

[0549] 일부 구현예에서, 복수의 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)은 샘플 내의 복수의 세포 성분 표적(예를 들어, 단백질 표적)에 특이적으로 결합할 수 있다. 샘플은 단일 세포, 복수의 세포, 조직 샘플, 종양 샘플, 혈액 샘플 등일 수 있거나, 이를 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 세포-표면 단백질, 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합을 포함한다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 단백질을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 세포내 단백질을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 유기체 내의 모든 세포 성분 표적(예를 들어, 발현되거나 발현될 수 있는 단백질)의 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 유기체 내의 모든 세포 성분 표적(예를 들어, 발현되거나 발현될 수 있는 단백질)의 적어도, 또는 최대 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 또는 99%일 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있거나, 약 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 상이한 세포 성분 표적을 포함할 수 있다. 일부 구현예에서, 복수의 세포 성분 표적은 적어도, 또는 최대 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 1000, 또는 10000개의 상이한 세포 성분 표적을 포함할 수 있다.

[0550] 올리고뉴클레오타이드-접합된 세포 성분 결합 시약을 사용하는 샘플 인덱싱

[0551] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적 바코딩)하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다.

[0552] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

- [0553] 항체와 접합된 올리고뉴클레오타이드, 항체와 접합하기 위한 올리고뉴클레오타이드, 또는 항체와 미리 접합된 올리고뉴클레오타이드는 본 명세서에서 항체 올리고뉴클레오타이드("Ab올리고")로 지칭된다. 샘플 인텍싱과 관련하여 항체 올리고뉴클레오타이드는 본 명세서에서 샘플 인텍싱 올리고뉴클레오타이드로 지칭된다. 항체 올리고뉴클레오타이드와 접합된 항체는 본 명세서에서 고온 항체 또는 올리고뉴클레오타이드 항체로 지칭된다. 항체 올리고뉴클레오타이드와 접합되지 않은 항체는 본 명세서에서 저온 항체 또는 올리고뉴클레오타이드-부재 항체로 지칭된다. 결합 시약(예를 들어, 단백질 결합 시약)과 접합된 올리고뉴클레오타이드, 결합 시약과의 접합을 위한 올리고뉴클레오타이드, 또는 결합 시약과 미리 접합된 올리고뉴클레오타이드는 본 명세서에서 시약 올리고뉴클레오타이드로 지칭된다. 샘플 인텍싱과 관련하여 시약 올리고뉴클레오타이드는 본 명세서에서 샘플 인텍싱 올리고뉴클레오타이드로 지칭된다. 항체 올리고뉴클레오타이드와 접합된 결합 시약은 본 명세서에서 고온 결합 시약 또는 올리고뉴클레오타이드 결합 시약으로 지칭된다. 항체 올리고뉴클레오타이드와 접합되지 않은 결합 시약은 본 명세서에서 저온 결합 시약 또는 올리고뉴클레오타이드-부재 결합 시약으로 지칭된다.
- [0554] 도 7은 샘플 인텍싱을 위한 올리고뉴클레오타이드-접합된 세포 성분 결합 시약을 사용하는 예시적인 작업 흐름의 개략도를 나타낸다. 일부 구현예에서, 각각 결합 시약을 포함하는 복수의 조성물(705a, 705b 등)이 제공된다. 결합 시약은 단백질 결합 시약, 예컨대 항체일 수 있다. 세포 성분 결합 시약은 항체, 테트라머, 압타머, 단백질 스캐폴드, 또는 이들의 조합을 포함할 수 있다. 복수의 조성물(705a, 705b)의 결합 시약은 동일한 세포 성분 표적에 결합할 수 있다. 예를 들어, 복수의 조성물(705, 705b)의 결합 시약은 (결합 시약과 회합된 샘플 인텍싱 올리고뉴클레오타이드를 제외하고) 동일할 수 있다.
- [0555] 상이한 조성물은 상이한 샘플 인텍싱 서열을 갖는 샘플 인텍싱 올리고뉴클레오타이드와 접합된 결합 시약을 포함할 수 있다. 상이한 조성물의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 상이한 조성물의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 조성물의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 또는 10000개일 수 있다.
- [0556] 일부 구현예에서, 하나의 조성물 내의 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드는 동일한 샘플 인텍싱 서열을 포함할 수 있다. 하나의 조성물 내의 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드는 동일하지 않을 수 있다. 일부 구현예에서, 동일한 샘플 인텍싱 서열을 갖는, 하나의 조성물 내의 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드의 백분율은 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.9%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.9%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 샘플 인텍싱 서열을 갖는, 하나의 조성물 내의 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드의 백분율은 적어도, 또는 최대 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 99.9%일 수 있다.
- [0557] 조성물(705a, 705b)은 상이한 샘플 중의 샘플을 표지하는 데 사용될 수 있다. 예를 들어, 조성물(705a) 내의 세포 성분 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드는 하나의 샘플 인텍싱 서열을 가질 수 있고, 샘플(707a), 예컨대 환자의 샘플 내의, 흑색원으로 표시된 세포(710a)를 표지하는 데 사용될 수 있다. 조성물(705b) 내의 세포 성분 결합 시약의 샘플 인텍싱 올리고뉴클레오타이드는 또 다른 샘플 인텍싱 서열을 가질 수 있고, 샘플(707b), 예컨대 또 다른 환자의 샘플 또는 동일한 환자의 또 다른 샘플 내의, 빗금친 원으로 표시된 세포(710b)를 표지하는 데 사용될 수 있다. 세포 성분 결합 시약은 세포 표면 상의 세포 성분 표적 또는 단백질, 예컨대 세포 마커, B-세포 수용체, T-세포 수용체, 항체, 주요 조직적합성 복합체, 종양 항원, 수용체, 또는 이들의 임의의 조합에 특이적으로 결합할 수 있다. 결합되지 않은 세포 성분 결합 시약은, 예를 들어 세포를 완충액으로 세척함으로써 제거될 수 있다.

- [0558] 이어서, 세포 성분 결합 시약을 갖는 세포는 복수의 구획, 예컨대 마이크로웰 어레이 내로 분리될 수 있으며, 여기서 단일 구획(715a, 715b)은 단일 세포(710a) 및 단일 비드(720a) 또는 단일 세포(710b) 및 단일 비드(720b)에 맞도록 크기조정된다. 각각의 비드(720a, 720b)는 복수의 올리고뉴클레오타이드 프로브(이것은 비드 상의 모든 올리고뉴클레오타이드 프로브에 공통된 세포 표지를 포함할 수 있음), 및 분자 표지 서열을 포함할 수 있다. 일부 구현예에서, 각각의 올리고뉴클레오타이드 프로브는 표적-결합 영역, 예를 들어 폴리(dT) 서열을 포함할 수 있다. 조성물(705a)의 세포 성분 결합 시약에 접합된 샘플 인덱싱 올리고뉴클레오타이드(725a)는 세포 성분 결합 시약으로부터 탈착 가능하거나 탈착 불가능하도록 구성될 수 있다. 조성물(705a)의 세포 성분 결합 시약에 접합된 샘플 인덱싱 올리고뉴클레오타이드(725a)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 세포 성분 결합 시약으로부터 탈착될 수 있다. 조성물(705b)의 세포 성분 결합 시약에 접합된 샘플 인덱싱 올리고뉴클레오타이드(725b)는 세포 성분 결합 시약으로부터 탈착 가능하거나 탈착 불가능하도록 구성될 수 있다. 조성물(705b)의 세포 성분 결합 시약에 접합된 샘플 인덱싱 올리고뉴클레오타이드(725b)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 세포 성분 결합 시약으로부터 탈착될 수 있다.
- [0559] 세포(710a)는 용해되어 세포(710a) 내의 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(730a)를 방출할 수 있다. 용해된 세포(735a)는 점선으로 표시한 원으로 나타나 있다. 세포성 mRNA(730a), 샘플 인덱싱 올리고뉴클레오타이드(725a), 또는 둘 모두는, 예를 들어 폴리(dT) 서열에 혼성화함으로써, 비드(720a) 상의 올리고뉴클레오타이드 프로브에 의해 포획될 수 있다. 세포성 mRNA(730a) 및 올리고뉴클레오타이드(725a)를 주형으로서 사용하여 세포성 mRNA(730a) 및 올리고뉴클레오타이드(725a)에 혼성화된 올리고뉴클레오타이드 프로브를 연장하기 위해 역전사효소가 사용될 수 있다. 역전사효소에 의해 생성된 연장 산물은 증폭되고 서열분석될 수 있다.
- [0560] 유사하게, 세포(710b)는 용해되어 세포(710b) 내의 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(730b)를 방출할 수 있다. 용해된 세포(735b)는 점선으로 표시한 원으로 나타나 있다. 세포성 mRNA(730b), 샘플 인덱싱 올리고뉴클레오타이드(725b), 또는 둘 모두는, 예를 들어 폴리(dT) 서열에 혼성화함으로써, 비드(720b) 상의 올리고뉴클레오타이드 프로브에 의해 포획될 수 있다. 세포성 mRNA(730b) 및 올리고뉴클레오타이드(725b)를 주형으로서 사용하여 세포성 mRNA(730b) 및 올리고뉴클레오타이드(725b)에 혼성화된 올리고뉴클레오타이드 프로브를 연장하기 위해 역전사효소가 사용될 수 있다. 역전사효소에 의해 생성된 연장 산물은 증폭되고 서열분석될 수 있다.
- [0561] 서열분석 판독물은 (예를 들어, 샘플 인덱싱 올리고뉴클레오타이드(725a, 725b)의 샘플 인덱싱 서열의 측면에서) 세포 표지, 분자 표지, 유전자 정체, 및 샘플 정체의 역다중화를 거칠 수 있다. 세포 표지, 분자 표지, 및 유전자 정체의 역다중화는 샘플 내의 각각의 단일 세포의 유전자 발현의 디지털 표현을 생성할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열을 사용하는, 세포 표지, 분자 표지, 및 샘플 정체의 역다중화는 샘플 기원을 결정하는 데 사용될 수 있다.
- [0562] 일부 구현예에서, 세포 표면 상의 세포 성분 결합 시약에 대한 세포 성분 결합 시약은 고유 샘플 인덱싱 올리고뉴클레오타이드의 라이브러리에 접합되어 세포가 샘플 정체를 보유할 수 있게 할 수 있다. 예를 들어, 세포 표면 마커에 대한 항체는 고유 샘플 인덱싱 올리고뉴클레오타이드의 라이브러리에 접합되어 세포가 샘플 정체를 보유할 수 있게 할 수 있다. 이것은, 샘플 소스에 관한 정보가 라이브러리 제조 및 서열분석 전체에 걸쳐 보유되기 때문에, 다수의 샘플이 동일한 Rhapsody™ 카트리지에 상에 로딩될 수 있게 할 것이다. 샘플 인덱싱은 다수의 샘플이 단일 실험에서 함께 실시될 수 있게 할 수 있어서, 간소화 및 실험 시간 단축, 그리고 배치 효과의 제거를 가능하게 할 수 있다.
- [0563] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함한다. 방법은 복수의 바코드(예를 들어, 추계적 바코드)를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적으로 바코딩)하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함할 수 있다.

- [0564] 일부 구현예에서, 샘플 식별을 위한 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시키는 단계로서, 하나 이상의 세포 각각은 하나 이상의 세포 성분 표적을 포함하고, 복수의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계; 및 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 하나 이상의 세포 중 적어도 하나의 세포의 샘플 기원을 식별하는 단계를 포함한다.
- [0565] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는, 복수의 바코드(예를 들어, 추계적 바코드)를 사용하여 복수의 샘플 인덱싱 조성물의 샘플 인덱싱 올리고뉴클레오타이드를 바코딩(추계적으로 바코딩)하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드 중 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함한다. 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.
- [0566] 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 복수의 샘플 인덱싱 조성물의 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열의 존재 또는 부재를 식별하는 단계를 포함할 수 있다. 샘플 인덱싱 서열의 존재 또는 부재를 식별하는 단계는, 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계; 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 서열분석 데이터에서 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 상응하는 복수의 샘플 인덱싱 올리고뉴클레오타이드 중 복제된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 세포의 샘플 기원을 식별하는 단계를 포함할 수 있다.
- [0567] 일부 구현예에서, 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는, 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하기 전에, 복제 어댑터를 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 결합하는 단계를 포함한다. 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드에 결합된 복제 어댑터를 사용하여 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0568] 일부 구현예에서, 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는, 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 복제하기 전에, 포획 프로브를 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 포획 프로브를 연장하여 포획 프로브와 회합된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 적어도 하나의 샘플 인덱싱 올리고뉴클레오타이드를 복제하는 단계는 포획 프로브와 회합된 샘플 인덱싱 올리고뉴클레오타이드를 복제하여 복수의 복제된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0569] 페이지-체이닝 올리고뉴클레오타이드
- [0570] 본 명세서의 개시 내용은 바코딩된 올리고뉴클레오타이드(예컨대, 샘플 인덱싱 올리고뉴클레오타이드, 제어 입자 올리고뉴클레오타이드, 제어 올리고뉴클레오타이드, 또는 단백질 발현을 검출하기 위한 항체에 접합된 올리고뉴클레오타이드)를, 긴 오버행(예를 들어, 160개 이상의 뉴클레오타이드 길이)을 갖는 프라이머를 사용하는 바코딩된 올리고뉴클레오타이드의 PCR 증폭에 의해 연장하기 위한 시스템, 방법, 키트, 및 조성물을 포함한다. 그러한 프라이머는, 예를 들어 200개 이상의 뉴클레오타이드 길이, 예컨대 Ultramer[®] DNA 올리고뉴클레오타이드, Megamer[®] 단일 가닥 DNA 단편, 또는 gBlocks[®] 유전자 단편의 단일 가닥 DNA 단편(Integrated DNA Technologies, Inc.(미국 아이오와주 코럴빌 소재))일 수 있다. 바코딩된 올리고뉴클레오타이드는 PCR 동안 긴 프라이머가 결합하고 연장하기 위한 공통 서열(예를 들어, 40개의 뉴클레오타이드 길이인 공통 서열)을 함유할 수 있다. 일부

구현예에서, 긴 오버행은 2차 바코드 서열(예를 들어, 분자 표지 서열)을 포함하여 샘플에 복잡성을 부가할 수 있다. 따라서, 바코딩된 올리고뉴클레오타이드의 길이는, 예를 들어 200 내지 300개의 뉴클레오타이드부터 400개 초과 뉴클레오타이드까지 증가될 수 있다. 그러한 더 긴 바코딩된 올리고뉴클레오타이드는, 예를 들어 고체상 가역적 고정화(Solid Phase Reversible Immobilization, SPRI)_자성 비드(Applied Biological Materials Inc.(캐나다 브리티시 컬럼비아 리치몬드 소재))를 사용하는 DNA 정제 동안 증가하는 크기 선택을 언제나 견딜 수 있다.

[0571] 일부 구현예에서, 세포 성분 결합 시약(예를 들어, 항체 분자)과 회합된(예를 들어, 이것에 부착된) 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터의 해리 후에 바코딩(예를 들어, 추계적으로 바코딩)되어, 바코딩된 올리고뉴클레오타이드를 생성할 수 있다. 바코딩된 올리고뉴클레오타이드는 긴 오버행을 갖는 프라이머를 사용하여 추가로 신장될 수 있다. 따라서, DNA 합성 기술 또는 실험 설계 고려사항에 의해 제한될 수 있는 세포 성분 결합 시약과 회합된 올리고뉴클레오타이드의 길이 및 바코딩된 올리고뉴클레오타이드의 길이는 후속 작업 흐름(예를 들어, DNA 정제 동안의 크기 선택)을 제한하지 않는다. 하나의 비제한적이고 예시적인 실험 설계 고려사항은 세포 성분 시약의 결합 특이성 및/또는 효능에 대한 세포 성분 결합 시약에 부착된 긴 올리고뉴클레오타이드의 효과이다. 바코딩된 올리고뉴클레오타이드의 PCR 데이징-체이닝은 항체 특이성 및/또는 효능을 유지하면서 DNA 합성 기술의 제한을 극복하거나 그에 대처할 수 있다. 긴 오버행을 갖는 프라이머를 사용하여 공통 서열을 갖는 바코딩된 올리고뉴클레오타이드를 신장하기 위한 본 개시 내용의 방법은 본 명세서에서, 바코딩된 올리고뉴클레오타이드를 데이징-체이닝 증폭 프라이머를 사용하여 데이징 체이닝하여 데이징-체이닝 신장된 앰플리콘을 생성하는 것으로 지칭된다.

[0572] 본 명세서의 개시 내용은 샘플 인덱싱 올리고뉴클레오타이드(또는 기타 다른 바코딩된 올리고뉴클레오타이드, 예컨대 제어 입자 올리고뉴클레오타이드)를 연장하기 위한 방법을 포함한다. 일부 구현예에서, 복수의 샘플 각각으로부터의 하나 이상의 세포를 복수의 샘플 인덱싱 조성물 중의 샘플 인덱싱 조성물과 접촉시킨 후에, 샘플 인덱싱 올리고뉴클레오타이드는 복수의 바코드를 사용하여 바코딩되어 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성할 수 있다(도 7 참조). 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드는 긴 오버행을 갖는 복수의 프라이머(본 명세서에서 데이징-체이닝 증폭 프라이머로 지칭됨)를 사용하여 증폭되어 복수의 더 긴 앰플리콘(본 명세서에서 데이징-체이닝 신장된 앰플리콘으로 지칭됨)을 생성될 수 있다. 그러한 더 긴 앰플리콘은, 예를 들어 바코딩된 샘플 인덱싱 올리고뉴클레오타이드보다 실질적으로 더 길(예를 들어, 160개의 뉴클레오타이드 길이만큼 더 길) 수 있다.

[0573] 도 8a 및 도 8b는 바코딩된 올리고뉴클레오타이드를 데이징-체이닝하는 예시적인 작업 흐름의 개략도를 나타낸다. 일부 구현예에서, 샘플 인덱싱 조성물(805)은 세포 성분 결합 시약, 예컨대 항체를 포함할 수 있으며, 이는 샘플 인덱싱 올리고뉴클레오타이드(825)와 회합되어 있다. 샘플 인덱싱 올리고뉴클레오타이드(825)는 샘플 인덱싱 서열(825s), 데이징-체이닝 증폭 프라이머에 결합하기 위한 공통 영역(825d)(본 명세서에서 데이징-체이닝 증폭 프라이머-결합 영역 또는 서열(825d)로 지칭됨), 및 폴리(A) 테일(825a)을 포함할 수 있다. 일부 구현예에서, 데이징-체이닝 증폭 프라이머에 결합하기 위한 공통 영역(825d)은 샘플 인덱싱 올리고뉴클레오타이드(825)의 전부 또는 일부에 대해 동일할 수 있다. 일부 구현예에서, 상이한 샘플 인덱싱 올리고뉴클레오타이드(825)의 데이징-체이닝 증폭 프라이머-결합 영역(825d)은 상이하다. 샘플 인덱싱 올리고뉴클레오타이드(825)는 mRNA 모방체일 수 있다. 복수의 샘플 인덱싱 조성물 중 적어도 2개의 샘플 인덱싱 조성물(805)의 샘플 인덱싱 서열(825s)은 상이한 서열을 포함할 수 있다.

[0574] 데이징-체이닝 증폭 프라이머-결합 영역(825d)은 10 내지 200개의 뉴클레오타이드 길이의 범위로 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 데이징-체이닝 증폭 프라이머-결합 영역(825d)은 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 데이징-체이닝 증폭 프

라이머-결합 영역(825d)은 적어도, 또는 최대 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오타이드 길이일 수 있다.

[0575] 샘플 인덱싱 올리고뉴클레오타이드(825)는 복수의 바코드(815)(예를 들어, 입자, 예컨대 비드(81)와 회합된 바코드(815))를 사용하여 바코딩되어 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)를 생성할 수 있다. 일부 구현예에서, 바코드(815)는 샘플 인덱싱 올리고뉴클레오타이드(825)에 결합하기 위한 폴리(dT) 영역(815t), 선택적으로 분자 표지(815m)(예를 들어, 샘플 인덱싱 올리고뉴클레오타이드의 출현의 수를 결정하기 위함), 세포 표지(815c), 및 범용 표지(815u)를 포함할 수 있다. 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 데이지-체이닝 증폭 프라이머-결합 영역(825d)의 역상보체(825d_rt)를 포함할 수 있다.

[0576] 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 복수의 데이지-체이닝 증폭 프라이머(845)를 포함한 순방향 및 역방향 프라이머(845, 850)를 사용하여 증폭되어 복수의 데이지-체이닝 신장된 앰플리콘(860)을 생성할 수 있다. 데이지-체이닝 증폭 프라이머(845)는 데이지-체이닝 증폭 프라이머-결합 영역(825d)(본 명세서에서 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)으로 지칭됨)의 서열(845d), 이의 상보체, 이의 역상보체, 또는 이들의 조합을 포함할 수 있다. 데이지-체이닝 증폭 프라이머-결합 영역(825d)과 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 동일한 길이 또는 상이한 길이일 수 있다. 데이지-체이닝 증폭 프라이머(845)의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840) 상의 이의 역상보적 서열(825d_rt)에 결합할 수 있다. 데이지-체이닝 증폭 프라이머(845)는 오버행 영역(845o)을 포함할 수 있다(오버행 영역(845o)은 데이지-체이닝 증폭 프라이머(845)가 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(850)에 결합할 때 데이지-체이닝 증폭 프라이머(845)의 오버행임). 복수의 데이지-체이닝 신장된 앰플리콘(860) 또는 이의 일부는 샘플 식별(또는 단백질 발현 프로파일링, 서열분석 제어 등)을 위해 서열분석될 수 있다. 도 8aa 및 도 8ab에 예시된 바와 같이, 샘플 인덱싱 올리고뉴클레오타이드(825)는 200개의 뉴클레오타이드 길이일 수 있고, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 240개의 뉴클레오타이드 길이일 수 있고, 데이지-체이닝 신장된 앰플리콘(860)은 400개의 뉴클레오타이드 길이일 수 있다.

[0577] 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 예를 들어 10 내지 200개의 뉴클레오타이드 길이의 범위로 상이한 실시예에서 상이한 길이를 가질 수 있다. 일부 구현예에서, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 적어도, 또는 최대 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오타이드 길이일 수 있다.

[0578] 일부 구현예에서, 적어도 2개의 데이지-체이닝 증폭 프라이머(845)는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)을 포함할 수 있다. 2개 이상의 데이지-체이닝 증폭 프라이머(845)는 상이한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)을 포함할 수 있다. 데이지-체이닝 증폭 프라이머(845) 중 일부 또는 전부는 동일한 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)을 포함할 수 있다. 바코딩된 샘플 인덱싱 올리고뉴클레오타이드-결합 영역(845d)은 데이지-체이닝 증폭 프라이머 결합 서열(825d), 이의 상보체, 이의 역상보체, 또는 이들의 조합을 포함할 수 있다. 적어도 2개의

샘플 인덱싱 조성물(805)의 데이지-체이닝 증폭 프라이머 결합 서열(825d)은 동일한 서열을 포함할 수 있다. 적어도 2개의 샘플 인덱싱 조성물(805)의 데이지-체이닝 증폭 프라이머 결합 서열(825d)은 상이한 서열을 포함할 수 있다.

[0579] 일부 구현예에서, 오버행 영역(845o)은 바코드 서열(본 명세서에서 데이지-체이닝 증폭 프라이머 바코드 서열로 지칭됨), 예컨대 분자 표지를 포함할 수 있다. 예를 들어, 2개의 데이지-체이닝 증폭 프라이머(845o)는 동일한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역(845o)을 포함할 수 있다. 또 다른 예로서, 2개의 데이지-체이닝 증폭 프라이머(845)는 2개의 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 갖는 오버행 영역(845o)을 포함할 수 있다. 데이지-체이닝 증폭 프라이머(845)의 오버행 영역(845o)은 상이한 데이지-체이닝 증폭 프라이머 바코드 서열을 포함할 수 있다.

[0580] 오버행 영역(845)은 예를 들어 50 내지 1000개의 뉴클레오타이드 길이의 범위로 상이한 실시에서 상이한 길이를 가질 수 있다. 일부 구현예에서, 오버행 영역(845)은 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위 뉴클레오타이드 길이일 수 있거나, 약 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 오버행 영역(845)은 적어도, 또는 최대 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 또는 1000일 수 있다.

[0581] 데이지-체이닝 신장된 앰플리콘(860)은, 예를 들어 50 내지 1000개의 뉴클레오타이드 길이의 범위로 상이한 실시에서 상이한 길이를 가질 수 있다. 일부 구현예에서, 데이지-체이닝 신장된 앰플리콘(860)은 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있거나, 약 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이일 수 있다. 일부 구현예에서, 데이지-체이닝 신장된 앰플리콘(860)은 적어도, 또는 최대 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 또는 1000개일 수 있다.

[0582] 도 8ba 및 도 8bb는 바코딩된 올리고뉴클레오타이드를 데이지-체이닝하는 또 다른 예시적인 작업 흐름의 개략도를 나타낸다. 이러한 작업 흐름에서는, 스위치 올리고뉴클레오타이드(860)의 역상보적 서열이 바코딩된 샘플 인덱싱

올리고뉴클레오타이드(840)에 선택적으로 부가될 수 있다. 역전사 동안, 샘플 인덱싱 올리고뉴클레오타이드(824)의 5' 말단에 도달할 때, 효소(예를 들어, 역전사효소, 예컨대 몰로니 뮤린 백혈병 바이러스(MMLV))의 종결 전사효소 활성은 몇몇 추가의 뉴클레오타이드(대개 테옥시시티딘)를 새로 합성된 가닥(840)의 3' 말단에 부가시킨다. 이들 염기는 주형 스위치(TS) 올리고뉴클레오타이드(860)의 고착 부위로서 기능할 수 있다. 주형 스위치 올리고뉴클레오타이드(860)와 첨부된 테옥시시티딘 스트레치 사이의 염기쌍 형성 시에, 효소는 주형 가닥을 샘플 인덱싱 올리고뉴클레오타이드(825)로부터 주형 스위치 올리고뉴클레오타이드(860)로 "스위칭"하고, TS 올리고뉴클레오타이드(860)의 5' 말단으로 복제를 계속한다. 따라서, 생성된 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 샘플 인덱싱 올리고뉴클레오타이드(825) 및 주형 스위치 올리고뉴클레오타이드(860)의 역상보체를 함유한다.

[0583] 주형 스위치 올리고뉴클레오타이드(860)의 역상보체 서열을 갖는 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 순방향 및 역방향 프라이머(845, 850)를 사용하여 증폭되어 주형 스위치 올리고뉴클레오타이드(860)의 역상보체를 갖는 데이지-체이닝 신장된 앰플리콘(860)을 생성할 수 있다. 2개의 프라이머 중 하나는 도 8aa 및 도 8ab에서의 범용 표지(845u)와 동등할 수 있는, 증폭을 위한 범용 서열인 판독물 1 서열(845r1)에 결합될 수 있다. 데이지-체이닝 증폭 프라이머(845)인 제2 프라이머는 데이지-체이닝 증폭 프라이머-결합 영역(825d)의 역상보체 서열(825d_rt)에 결합할 수 있다. 따라서, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(860)에 긴 테일이 부가될 수 있다. 효소적 단편화, 말단 수복, A-테일링, 및 선택적으로 샘플 인덱싱 PCR 후에, 생성된 앰플리콘은 P5 서열(865p5), P7 서열(865p7), 판독물 1 서열(815r1), 및 (예를 들어, 가교 증폭을 사용한) 차세대 서열분석을 위한 판독물 2 서열(865r2)을 포함할 수 있다. 도 8aa 및 도 8ab에 예시된 바와 같이, 샘플 인덱싱 올리고뉴클레오타이드(825)는 200개의 뉴클레오타이드 길이일 수 있고, 바코딩된 샘플 인덱싱 올리고뉴클레오타이드(840)는 대략 260개의 뉴클레오타이드 길이일 수 있고, 데이지-체이닝 신장된 앰플리콘(860)은 406개 초과 뉴클레오타이드 길이일 수 있다.

[0584] 단일 세포 서열분석 제어 입자

[0585] 본 명세서의 개시 내용은, 예를 들어 단일 세포 서열분석 제어에 사용될 수 있는 제어 입자 조성물을 포함한다. 일부 구현예에서, 제어 입자 조성물은 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함한다. 복수의 제어 입자 올리고뉴클레오타이드와 회합된 제어 입자는 또한 작용화된 제어 입자로 본 명세서에서 지칭된다. 도 9a는 복수의 올리고뉴클레오타이드로 작용화된 입자의 비제한적이고 예시적인 개략도이다. 도 9a는 제어 입자와 회합된 제어 입자 올리고뉴클레오타이드가 제어 바코드 서열 및 폴리(dA) 영역을 포함하여, mRNA 폴리(A) 테일을 모방할 수 있음을 나타낸다. 제어 입자 올리고뉴클레오타이드는 분자 표지 서열, 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다. 제어 입자와 회합된 제어 입자 올리고뉴클레오타이드는 서로 동일하거나 상이할 수 있다. 일부 구현예에서, 제어 입자와 회합된 적어도 2개의 제어 입자 올리고뉴클레오타이드는 상이한 제어 바코드 서열을 갖는다. 일부 구현예에서, 복수의 제1 제어 입자 올리고뉴클레오타이드 및 복수의 제2 제어 올리고뉴클레오타이드는 제어 입자와 회합되고, 제1 입자 올리고뉴클레오타이드 및 제2 입자 올리고뉴클레오타이드는 상이한 제어 바코드 서열을 갖는다.

[0586] 비드, 예컨대 CompBead™ Plus(BD(미국 뉴저지주 프랭클린 레이크 소재))는 올리고뉴클레오타이드와 접합된 항체로 작용화될 수 있다. CompBeads Plus는 약 7.5 미크론 크기로, 면역 세포의 크기와 유사하다. 올리고뉴클레오타이드와 접합된 항체로 작용화된 경우, CompBead Plus는 단일 세포 작업 흐름을 위한 제어 세포 또는 제어 입자로서 사용될 수 있다. AbO 작용화된 비드는 단일 세포 서열분석 제어로서 임의의 단일 세포 작업 흐름으로 사용될 수 있다.

[0587] 제어 입자 올리고뉴클레오타이드

[0588] 제어 바코드 서열의 길이는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 제어 바코드 서열은 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이거나, 약 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440,

450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이다. 일부 구현예에서, 제어 바코드 서열은 적어도, 또는 최대 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 또는 1000개의 뉴클레오타이드 길이이다.

[0589] 제어 입자 올리고뉴클레오타이드의 길이는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이거나, 약 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이다. 일부 구현예에서, 제어 입자 올리고뉴클레오타이드는 적어도, 또는 최대 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 또는 1000개의 뉴클레오타이드 길이이다.

[0590] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000, 1000000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000, 또는 1000000000개일 수 있다.

[0591] 복수의 제어 입자 올리고뉴클레오타이드는 동일하거나 상이한 제어 바코드 서열을 포함할 수 있다. 예를 들어, 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 2개는 상이한 제어 바코드 서열을 포함할 수 있다. 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드의 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000, 또는 1000000000개의 제어 바코드 서열은 동일할 수 있다. 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000, 또는 1000000000개, 또는 이들 중 임의의 두 값들 사이의 어떤 수 또는 범위, 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40,

50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 1000000, 10000000, 100000000개, 또는 이들 중 임의의 두 값들 사이의 어떤 수 또는 범위의 제어 바코드 서열은 동일할 수 있다.

[0592] 복수의 제어 입자 올리고뉴클레오타이드의 적어도, 또는 최대 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값들 사이의 어떤 수 또는 범위의 제어 바코드 서열은 동일할 수 있다. 복수의 제어 입자 올리고뉴클레오타이드의 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값들 사이의 어떤 수 또는 범위, 또는 약 0.000000001%, 0.00000001%, 0.0000001%, 0.000001%, 0.00001%, 0.0001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 또는 이들 중 임의의 두 값들 사이의 어떤 수 또는 범위의 제어 바코드 서열은 동일할 수 있다.

[0593] 일부 구현예에서, 제어 바코드 서열은 종의 유전체 서열과 상동성이 아니다. 제어 바코드 서열은 종의 유전체 서열과 상동성일 수 있다. 종은 비-포유동물 종일 수 있다. 비-포유동물 종은 파지 종일 수 있다. 파지 종은 T7 파지, PhiX 파지, 또는 이들의 조합일 수 있다.

[0594] 제어 입자

[0595] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 통해 제어 입자와 회합된다. 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0596] 제어 입자의 직경은 1 내지 1000 마이크로미터, 예컨대 10 내지 100 마이크로미터 또는 7.5 마이크로미터일 수 있거나, 약 1 내지 1000 마이크로미터, 예컨대 10 내지 100 마이크로미터 또는 7.5 마이크로미터일 수 있다. 일부 구현예에서, 제어 입자의 직경은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 마이크로미터, 또는 이들 중 임의의 두 값들 사이의 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 마이크로미터, 또는 이들 중 임의의 두 값들 사이의 수 또는 범위일 수 있다. 일부 구현예에서, 제어 입자의 직경은 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000 마이크로미터일 수 있다.

[0597] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 고정화될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 상에 부분적으로 고정화될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 봉입될 수 있다. 복수의 제어 입자 올리고뉴클레오타이드는 제어 입자 내에 부분적으로 봉입될 수 있다. 제어 입자는 붕괴성일 수 있다.

[0598] 일부 구현예에서, 제어 입자는 비드일 수 있다. 비드는 세파로스 비드, 스트렙타비딘 비드, 아가로스 비드, 자성 비드, 접합된 비드, 단백질 A 접합된 비드, 단백질 G 접합된 비드, 단백질 A/G 접합된 비드, 단백질 L 접합

된 비드, 올리고(dT) 접합된 비드, 실리카 비드, 실리카-유사 비드, 항-비오틴 마이크로비드, 항-형광색소 마이크로비드, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 제어 입자는 폴리디메틸실록산(PDMS), 폴리스티렌, 유리, 폴리프로필렌, 아가로스, 젤라틴, 하이드로겔, 상자성체, 세라믹, 플라스틱, 유리, 메틸스티렌, 아크릴 중합체, 티타늄, 라텍스, 세파로스, 셀룰로스, 나일론, 실리콘, 또는 이들의 임의의 조합의 재료를 포함할 수 있다. 제어 입자는 붕괴성 하이드로겔 입자를 포함할 수 있다.

[0599] 단백질 결합 시약

[0600] 일부 구현예에서, 제어 입자는 복수의 제1 단백질 결합 시약에 회합되고, 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 하나와 회합된다. 도 9b는 올리고뉴클레오타이드로 작용화된 다수의 항체로 코팅된 비제한적이고 예시적인 입자를 나타낸다. 제1 단백질 결합 시약은 제1 항체(예를 들어, 1차 항체, 또는 2차 항체)를 포함할 수 있다. 제어 입자 올리고뉴클레오타이드는 링커를 통해 제1 단백질 결합 시약에 접합될 수 있다. 제1 제어 입자 올리고뉴클레오타이드는 링커를 포함할 수 있다. 링커는 화학 기를 포함할 수 있다. 화학 기는 제1 단백질 결합 시약에 가역적으로 또는 비가역적으로 부착될 수 있다. 화학 기는 UV 광절단성 기, 이황화 결합, 스트렙타비딘, 비오틴, 아민, 이황화 연결, 또는 이들의 임의의 조합을 포함할 수 있다.

[0601] 일부 구현예에서, 제어 입자는 복수의 제2 단백질 결합 시약에 회합된다. 복수의 제2 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 하나와 회합될 수 있다. 도 9c는 올리고뉴클레오타이드로 작용화된 복수의 제1 항체 및 올리고뉴클레오타이드로 작용화되지 않은 복수의 제2 항체들로 코팅된 비제한적이고 예시적인 입자를 나타낸다. 제어 입자 상의 항체는 (예를 들어, 제어 입자 올리고뉴클레오타이드와 회합된) 고온 항체 및 (예를 들어, 제어 입자 올리고뉴클레오타이드와 회합되지 않은) 저온 항체의 비로 적정되어 제어 입자로부터 수득된 서열분석 관독물의 양을 변경시킬 수 있다. 제1 항체 및 제2 항체는 동일하거나 상이할 수 있다.

[0602] 도 9d는 상대적으로 높은 풍부도, 중간 풍부도 및 낮은 풍부도로, 복수의 제1 제어 입자 올리고뉴클레오타이드, 복수의 제2 항체에 접합된 복수의 제2 제어 입자 올리고뉴클레오타이드 및 제3 제어 입자 올리고뉴클레오타이드로 작용화된 입자의 비제한적이고 예시적인 개략도이다. 복수의 제1 제어 입자 올리고뉴클레오타이드는 복수의 제1 단백질 결합 시약에 접합될 수 있다. 복수의 제2 제어 입자 올리고뉴클레오타이드는 복수의 제2 단백질 결합 시약에 접합될 수 있다. 복수의 제3 제어 입자 올리고뉴클레오타이드는 복수의 제3 단백질 결합 시약에 접합될 수 있다.

[0603] 제1, 제2 및 제3의 제어 입자 올리고뉴클레오타이드의 상대적 풍부도는 상이한 발현 수준을 갖는 mRNA를 모방할 수 있다. 제1 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오타이드와 제2 단백질 결합 시약과 회합된 제어 입자 올리고뉴클레오타이드는 상이한 제어 바코드 서열을 포함할 수 있다. 제어 입자 상에서 상이한 단백질 결합 시약, 예컨대 항체, 및 상이한 제어 입자 올리고뉴클레오타이드는 적정되어 표준 곡선을 생성할 수 있다. 제1 단백질 결합 시약, 제2 단백질 결합 시약, 또는 제3 단백질 결합 시약은 동일하거나 상이한 단백질 결합 시약일 수 있다.

[0604] 일부 구현예에서, 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 (또는 제3) 제어 입자 올리고뉴클레오타이드의 수의 비는 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 또는 이들 중 임의의 두 수 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 또는 이들 중 임의의 두 수 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 (또는 제3) 제어 입자 올리고뉴클레오타이드의 수의 비는 적어도, 또는 최대 1:1, 1:1.1,

1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 또는 1:100일 수 있다.

[0605] 일부 구현예에서, 제1 단백질 결합 시약은 파트너 결합 시약(예를 들어, 2차 항체)과 회합될 수 있고, 제1 단백질 결합 시약은 파트너 결합 시약을 사용하여 제어 입자와 회합된다. 파트너 결합 시약은 파트너 항체를 포함할 수 있다. 파트너 항체는 항-고양이 항체, 항-닭 항체, 항-소 항체, 항-개 항체, 항-당나귀 항체, 항-염소 항체, 항-기니피그 항체, 항-햄스터 항체, 항-말 항체, 항-인간 항체, 항-라마 항체, 항-원숭이 항체, 항-마우스 항체, 항-돼지 항체, 항-토끼 항체, 항-랫 항체, 항-양 항체, 또는 이들의 조합을 포함할 수 있다. 파트너 항체는 면역글로불린 G(IgG), F(ab') 단편, F(ab')₂ 단편, 이들의 조합, 또는 이들의 단편을 포함할 수 있다.

[0606] 일부 구현예에서, 제1 단백질 결합 시약은 동일한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드 중 2개 이상과 회합된다. 제1 단백질 결합 시약은 상이한 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드 중 2개 이상과 회합될 수 있다. 일부 구현예에서, 복수의 제1 단백질 결합 시약 중 적어도 하나는 복수의 제어 입자 올리고뉴클레오타이드 중 어느 것과도 회합되지 않는다. 제어 입자 올리고뉴클레오타이드와 회합된 제1 단백질 결합 시약 및 임의의 제어 입자 올리고뉴클레오타이드와 회합되지 않은 제1 단백질 결합 시약은 동일한 단백질 결합 시약일 수 있다.

[0607] 제어 바코드의 다양성

[0608] 하나의 제어 입자와 회합된 복수의 제어 입자 올리고뉴클레오타이드는 상이한 제어 바코드 서열을 갖는 다수의 제어 입자 올리고뉴클레오타이드를 포함할 수 있다. 제어 입자 올리고뉴클레오타이드가 갖는 제어 바코드 서열의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 제어 입자 올리고뉴클레오타이드가 갖는 제어 바코드 서열의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 제어 입자 올리고뉴클레오타이드가 갖는 제어 바코드 서열의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 또는 1000000개일 수 있다.

[0609] 일부 구현예에서, 동일한 제어 입자 올리고뉴클레오타이드 서열을 갖는 제어 입자 올리고뉴클레오타이드의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 제어 입자 올리고뉴클레오타이드 서열을 갖는 제어 입자 올리고뉴클레오타이드의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 또는 1000000개일 수 있다.

[0610] 일부 구현예에서, 복수의 제어 입자 올리고뉴클레오타이드는 제1 제어 바코드 서열을 각각 포함하는 복수의 제1 제어 입자 올리고뉴클레오타이드, 및 제2 제어 바코드 서열을 각각 포함하는 복수의 제2 제어 입자 올리고뉴클레오타이드를 포함하고, 제1 제어 바코드 서열 및 제2 제어 바코드 서열은 상이한 서열을 갖는다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수 및 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 대략 동일할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수 및 복수의 제2 제어 입자 올리고뉴클레오타이드의 수는 상이할 수 있다. 복수의 제1 제어 입자 올리고뉴클레오타이드의 수는 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 적어도 2배, 10배, 100배, 또는 그 이상일 수 있다. 일부 구현예에서, 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 비는 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31,

1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 제1 제어 입자 올리고뉴클레오타이드의 수와 복수의 제2 제어 입자 올리고뉴클레오타이드의 수의 비는 적어도, 또는 최대 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 또는 1:10000일 수 있다.

[0611] 검출가능한 모이어티

[0612] 일부 구현예에서, 제어 입자는 검출가능한 모이어티, 예를 들어, 광학 모이어티, 예컨대 형광단 또는 발색단과 회합된다. 제어 입자 올리고뉴클레오타이드는 검출가능한 모이어티, 예를 들어, 광학 모이어티와 회합될 수 있다. 일부 구현예에서, 제1 단백질 결합 시약은 광학 모이어티와 회합될 수 있다(도 9e). 제2 단백질 결합 시약은 광학 모이어티와 회합될 수 있다. 광학 모이어티와 회합된 제어 입자(예를 들어, 형광 태그된 비드)는 또한 이미징 및 유세포분석에 사용될 수 있다.

[0613] 검출가능한 모이어티는 스펙트럼상으로 구별가능한 검출가능한 모이어티의 군으로부터 선택될 수 있다. 스펙트럼상으로 구별가능한 검출가능한 모이어티는, 그의 방출 스펙트럼이 중복될 수 있다 하더라도, 구별가능한 방출 스펙트럼을 갖는 검출가능한 모이어티를 포함한다. 검출가능한 모이어티의 비제한적인 예는 잔틴 유도체: 플루오레세인, 로다민, 오레곤 그린, 에오신, 및 텍사스 레드; 시아닌 유도체: 시아닌, 인도카르보시아닌, 옥사카르보시아닌, 티아카르보시아닌, 및 메로시아닌; 세타(Seta), 세타우(SeTau) 및 스퀘어(Square) 염료를 포함하는 스쿠아레인 유도체 및 고리-치환된 스쿠아레인; 나프탈렌 유도체(단일 및 프로단 유도체); 쿠마린 유도체; 옥사디아졸 유도체: 피리딜옥사졸, 니트로벤즈옥사디아졸 및 벤즈옥사디아졸; 안트라센유도체: DRAQ5, DRAQ7 및 CyTRAK 오렌지를 포함하는 안트라퀴논; 피렌 유도체: 캐스캐이드 블루; 옥사진 유도체: 나일 레드, 나일 블루, 크레실 바이올렛, 옥사진 170; 아크리딘 유도체: 프로플라빈, 아크리딘 오렌지, 아크리딘 옐로우; 아틸메틴 유도체: 아우라민, 크리스탈 바이올렛, 말라카이트 그린; 및 테트라피롤 유도체: 포르핀, 프탈로시아닌, 빌리루빈을 포함한다. 검출가능한 모이어티의 기타 다른 비제한적인 예는 히드록시쿠마린, 아미노쿠마린, 메톡시쿠마린, 캐스캐이드 블루, 퍼시픽 블루(Pacific Blue), 퍼시픽 오렌지, 루시퍼(Lucifer) 옐로우, NBD, R-피코에리트린(PE), PE-Cy5 접합체, PE-Cy7 접합체, 레드 613, PerCP, TruRed, FluorX, 플루오레세인, BODIPY-FL, Cy2, Cy3, Cy3B, Cy3.5, Cy5, Cy5.5, Cy7, TRITC, X-로다민, 리사민 로다민 B, 텍사스 레드, 알로피코시아닌(APC), APC-Cy7 접합체, 웨스트(Hoechst) 33342, DAPI, 웨스트 33258, 시토스(SYTOX) 블루, 크로모마이신 A3, 미트라마이신, YOYO-1, 에티디움 브로마이드, 아크리딘 오렌지, SYTOX 그린, TOTO-1, TO-PRO-1, TO-PRO: 시아닌 모노머, 티아졸 오렌지, CyTRAK 오렌지, 프로피디움 요오드드(PI), LDS 751, 7-AAD, SYTOX 오렌지, TOTO-3, TO-PRO-3, DRAQ5, DRAQ7, Indo-1, Fluo-3, Fluo-4, DCFH, DHR, 및 SNARF를 포함한다.

[0614] 검출가능한 모이어티의 여기 파장은 다양할 수 있으며, 예를 들어 10, 20, 30, 40, 50, 60, 70, 80, 90, 100,

110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 나노미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 나노미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 검출가능한 모이어티의 방출 파장 또한 다양할 수 있으며, 예를 들어 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 나노미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 나노미터, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다.

[0615] 검출가능한 모이어티의 분자량은 다양할 수 있으며, 예를 들어 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 달톤(Da), 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 달톤(Da), 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 검출가능한 모이어티의 분자량은 또한 다양할 수 있으며, 예를 들어 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 킬로달톤(kDa), 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 킬로달톤(kDa), 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다.

[0616] 스펙트럼상으로 구별가능한 검출가능한 모이어티의 군은, 예를 들어 5개의 상이한 형광단, 5개의 상이한 발색단, 5개의 형광단 및 발색단의 조합, 4개의 상이한 형광단 및 하나의 비-형광단의 조합, 4개의 발색단 및 하나의 비-발색단의 조합, 또는 4개의 형광단 및 발색단 및 하나의 비-형광단 비-발색단의 조합을 포함한다. 일

부 구현예에서, 모이어티 중 검출가능한 모이어티는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개이거나, 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 스펙트럼상으로 구별가능한 검출가능한 모이어티일 수 있다.

[0617] 제어 입자 작업 흐름

[0618] AbO 작용화된 비드는 단일 세포 서열분석 제어로서 임의의 단일 세포 작업 흐름과 함께 사용될 수 있다. 단일 세포 작업 흐름은 마이크로웰 어레이 또는 마이크로웰 카트리지(예를 들어, BD Rhapsody™) 또는 마이크로유체학적 디바이스(예를 들어, 10X Genomics(미국 캘리포니아주 샌프란시스코 소재), Drop-seq(McCarroll Lab, Harvard Medical School(미국 매사추세츠주 캠브리지 소재); 그 내용이 전체가 본 명세서에 참고로 포함된 문헌 [Macosko et al., Cell, 2015 May 21 16;5:1202]), 또는 Abseq(Mission Bio(미국 캘리포니아주 샌프란시스코 소재)); 그 내용이 전체가 본 명세서에 참고로 포함된 문헌 [Shahi et al., Sci Rep. 2017 Mar 14;7:44447])를, 추계적 바코드(예를 들어, BD Rhapsody, 또는 Drop-seq) 또는 방출가능한 추계적 바코드를 봉입하고 있는 붕괴성 하이드로겔 입자(예를 들어, 10X Genomics, 또는 Abseq)와 회합된 고체 또는 반고체 입자와 함께 사용할 수 있다. 작용화된 비드는 단일 세포 작업 흐름의 효율을 결정하기 위한 제어일 수 있으며, 이는 벌크 RNAseq 또는 마이크로어레이 연구에 사용되는 외부 RNA 제어 컨소시엄(ERCCs)과 유사하다.

[0619] 본 명세서의 개시 내용은 복수의 제어 입자 올리고뉴클레오타이드를 사용하여 표적의 수를 결정하기 위한 방법을 포함한다. 표적(예를 들어, 유전자 발현)의 수를 결정하기 위한 방법은 본 명세서에 개시된 기타 다른 방법과 함께 사용될 수 있다. 예를 들어, 작업 흐름은 복수의 제어 입자 올리고뉴클레오타이드를 사용하여 단백질 발현 및 유전자 발현을 결정하는 데 사용될 수 있다. 일부 구현예에서, 방법은, 복수의 바코드(예를 들어, 추계적 바코드)를 사용하여 복수의 세포 중 하나의 세포의 복수의 표적 및 복수의 제어 입자 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적으로 바코딩)하여 복수의 바코딩된 표적(예를 들어, 추계적으로 바코딩된 표적) 및 복수의 바코딩된 제어 입자 올리고뉴클레오타이드(예를 들어, 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드)를 생성하는 단계를 포함하며, 여기서 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 상이한 바코드 서열(예를 들어, 분자 표지)을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하고, 제어 입자 조성물은 복수의 제어 입자 올리고뉴클레오타이드와 회합된 제어 입자를 포함하고, 복수의 제어 입자 올리고뉴클레오타이드 각각은 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 제어 바코드 서열 및 의사-표적 영역을 포함한다. 방법은, 복수의 바코딩된 표적 및 복수의 바코딩된 제어 입자 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 서열분석 데이터에서 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수를 계수하는 단계를 포함할 수 있다. 방법은, 복수의 표적 중 적어도 하나의 표적에 대하여, 서열분석 데이터에서 표적과 회합된 별개의 서열을 갖는 바코드 서열의 수를 계수하는 단계; 및 표적의 수를 예측하는 단계로서, 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수와 상관되는 단계를 포함할 수 있다. 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다. 의사-표적 영역은 표적의 하위서열을 포함할 수 있다.

[0620] 일부 구현예에서, 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수, 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수, 및 제어 바코드 서열을 포함하는 복수의 제어 입자 올리고뉴클레오타이드의 수와 상관될 수 있다. 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수, 및 제어 바코드 서열을 포함하는 복수의 제어 입자 올리고뉴클레오타이드의 수와 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수의 비와 상관될 수 있다.

[0621] 예를 들어, 제어 입자가 특정 제어 바코드 서열을 갖는 100개의 제어 입자 올리고뉴클레오타이드를 갖고, 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수(예를 들어, 라이브러리 제조 과정을 건디는 제어 바코드 서열을 갖는 제어 입자 올리고뉴클레오타이드의 수)가 80개인 경우, 라이브러리 제조(예를 들어, 역전사, 증폭 등)의 효율은 80%이다. 따라서, 상이한 라이브러리 제조로부터의 데이터는 라이브러리 제조 효율을 사용하여 정규화함으로써 비교될 수 있다.

[0622] 또 다른 예로서, 제어 입자는 저발현 유전자를 모방하는 특정 제어 바코드 서열을 갖는 5개의 제어 입자 올리고뉴클레오타이드를 포함할 수 있다. 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자

표지)의 수가 5개이고, 저발현 유전자가 검출되지 않는 경우, 저발현 유전자는 발현되지 않는다(또는 세포가 유전자의 5개 미만의 mRNA를 갖는다)고 결론지을 수 있다. 그러나, 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수가 0이고, 저발현 유전자가 검출되지 않는 경우, 저발현 유전자는 발현되지 않는다고 결론지을 수 없다.

[0623] 포획 효율은 상이한 풍부도를 갖는 제어 입자 올리고뉴클레오타이드에 대해 결정될 수 있다. 정규화는 2개 이상의 제어 바코드 서열을 갖는 제어 입자 올리고뉴클레오타이드의 포획 효율을 기반으로 수행될 수 있다. 일부 구현예에서, 서열분석 데이터에서 제어 바코드 서열을 갖는 복수의 제어 입자 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수의 계수는, 서열분석 데이터에서 제1 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수를 계수하는 단계; 및 서열분석 데이터에서 제2 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수를 계수하는 단계를 포함한다. 예측되는 표적의 수는 계수된 표적과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수, 제1 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수, 및 제2 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지)의 수와 상관될 수 있다.

[0624] 일부 구현예에서, 방법은 복수의 표적과 제어 입자 및 복수의 제어 입자 올리고뉴클레오타이드를 추계적으로 바코딩하기 전에, 제어 입자로부터 복수의 제어 입자 올리고뉴클레오타이드 중 적어도 하나를 방출하는 단계를 포함한다.

[0625] 일부 구현예에서, 복수의 바코드(예를 들어, 추계적 바코드)를 사용하여 복수의 표적 및 복수의 제어 입자 올리고뉴클레오타이드를 바코딩(예를 들어, 추계적으로 바코딩)하는 단계는, 복수의 바코드를 복수의 표적 중의 표적 및 복수의 제어 입자 올리고뉴클레오타이드 중의 제어 입자 올리고뉴클레오타이드와 접촉시켜 표적 및 제어 입자 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 표적 및 제어 입자 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 표적 및 복수의 바코딩된 제어 입자 올리고뉴클레오타이드(예를 들어, 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드)를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소, 역전사효소, 또는 이들의 조합을 사용하여 바코드를 연장하는 단계를 포함할 수 있다.

[0626] 일부 구현예에서, 방법은 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 추계적으로 바코딩된 표적 및 복수의 추계적으로 바코딩된 제어 입자 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 바코드 서열(예를 들어, 분자 표지)의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분 또는 바코드 서열(예를 들어, 분자 표지)의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 바코드 서열(예를 들어, 분자 표지)의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분, 또는 바코드 서열(예를 들어, 분자 표지)의 적어도 일부분 및 제어 입자 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0627] 마이크로웰 카트리지 또는 어레이 작업 흐름

[0628] 도 10은 단일 세포 서열분석 제어를 위해 올리고뉴클레오타이드로 작용화된 입자의 사용의 예시적인 작업 흐름의 개략도이다. 일부 구현예에서, 제어 입자 조성물은 제어 입자(1004)와 회합된 복수의 제어 입자 올리고뉴클레오타이드를 포함한다. 예를 들어, 제어 입자(1040)는 제어 입자(1040)에 결합된 항체(1005)에 접합된 제어 입자 올리고뉴클레오타이드(1025)와 회합될 수 있다. 제어 입자 올리고뉴클레오타이드(1025)로 작용화된 복수의 제어 입자(1040)는 복수의 세포 내에, 예를 들어 5%로 스파이크될 수 있다. 제어 입자(1040)는 이후의 작업 흐름에서 "세포"로서 처리될 수 있다. 제어 입자(1040)는 또한 제어 세포 또는 제어 세포 입자로 지칭될 수 있다. 이어서, 세포(1010) 및 제어 입자(1040)는 복수의 구획, 예컨대 마이크로웰 어레이의 웰 내로 분리될 수 있으며, 여기서 단일 구획(1015a, 1015b)은 단일 세포 또는 제어 입자 및 단일 비드(1020a, 1020b)에 맞추어 크기조정된다. 비드(1020a, 1020b)는 구획(1015a, 1015b) 내로 로딩될 수 있다.

[0629] 각각의 비드는, 비드 상에서 모든 올리고뉴클레오타이드 프로브에 공통적인 세포 표지를 포함할 수 있는 복수의 올리고뉴클레오타이드 프로브, 및 바코드 서열(예를 들어, 분자 표지 서열)을 포함할 수 있다. 일부 구현예에서, 각각의 올리고뉴클레오타이드 프로브는 표적-결합 영역, 예를 들어 폴리(dT) 서열을 포함할 수 있다. 항체(1005)에 접합된 올리고뉴클레오타이드(1025)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 항체(1005)로부터 탈착될 수 있다. 세포(1010)는 용해되어(1035) 세포 내에 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(1030)를 방출

할 수 있다. 세포성 mRNA(1030) 및 제어 입자 올리고뉴클레오타이드(1025)는 각각 비드(1020a, 1020b) 상에서 올리고뉴클레오타이드 프로브에 의해, 예를 들어 폴리(dT) 서열에 혼성화함으로써 포획될 수 있다. 비드는 회수될 수 있고, 포획된 세포성 mRNA(1030) 및 제어 입자 올리고뉴클레오타이드(1025)(예를 들어, 총 약 5000개의 세포에 상응함)는 풀링될 수 있다.

[0630] 역전사효소를 사용하여 세포성 mRNA(1030) 및 제어 입자 올리고뉴클레오타이드(1025)를 주형으로서 사용하여 세포성 mRNA(1030) 및 제어 입자 올리고뉴클레오타이드(1025)에 혼성화된 올리고뉴클레오타이드 프로브를 연장할 수 있다. 역전사효소에 의해 생성된 연장 산물은 증폭 및 서열분석될 수 있다. 서열분석 판독물은 세포 표지, 바코드 서열(예를 들어, 분자 표지), 유전자 정제, 제어 입자 올리고뉴클레오타이드 서열 등의 역다중화 처리될 수 있으며, 단일 세포 유전자 발현 프로파일 및 작업 흐름의 전체 또는 일부의 정량 효율(예를 들어, 세포 포획 효율)을 결정하는 데 사용될 수 있다. 예를 들어, 포획된 제어 입자의 수는 데이터 내에서 제어 바코드 서열에 회합된 세포 표지의 수에 기반하여 결정될 수 있다. 작업 흐름의 효율은 포획된 제어 입자의 수와 스파이크된 제어 입자의 비율 수 있다.

[0631] 마이크로유체학적 작업 흐름

[0632] 도 10은 단일 세포 서열분석 제어를 위해 올리고뉴클레오타이드로 작용화된 입자를 사용하는 또 다른 예시적인 작업 흐름의 개략도이다. 제어 입자 올리고뉴클레오타이드(1025)로 작용화된 복수의 제어 입자(1040)는 복수의 세포 내에 예를 들어, 5%로 스파이크될 수 있다. 제어 입자(1040)는 이후의 작업 흐름에서 "세포"로서 처리될 수 있다. 제어 입자(1040)는 또한 제어 세포 또는 제어 세포 입자로 지칭될 수 있다. 이어서, 세포(1010) 및 제어 입자(1040)는 마이크로유체학적 디바이스를 사용하여 복수의 구획, 예컨대 액적(1045a, 1045b)으로 분리될 수 있다. 각각의 액적(1045a, 1045b)은 하나의 세포(1010) 또는 하나의 제어 입자(1040) 및 하이드로겔 비드(1020a, 1020b)를 포함할 수 있다.

[0633] 각각의 비드(1020a, 1020b)는, 비드 상에서 모든 올리고뉴클레오타이드 프로브에 일반적인 세포 표지를 포함할 수 있는 복수의 올리고뉴클레오타이드 프로브, 및 바코드 서열(예를 들어, 분자 표지 서열)을 포함할 수 있다. 일부 구현예에서, 각각의 올리고뉴클레오타이드 프로브는 표적-결합 영역, 예를 들어 폴리(dT) 서열을 포함할 수 있다. 비드(1020a, 1020b)는 이후의 작업 흐름(예를 들어, 역전사)을 위한 시약을 포함할 수 있다. 항체(1005)에 접합된 올리고뉴클레오타이드(1025)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 항체(1005)로부터 탈착될 수 있다. 세포(1010)는 용해되어(1035) 세포 내에 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(1030)를 방출할 수 있다. 세포성 mRNA(530) 및 제어 입자 올리고뉴클레오타이드(1025)는 각각 비드(1020a, 1020b)로부터 방출된 올리고뉴클레오타이드 프로브에 의해, 예를 들어 폴리(dT) 서열에 혼성화함으로써 포획될 수 있다. 역전사효소는 세포성 mRNA(1030) 및 올리고뉴클레오타이드(1025)를 주형으로서 사용하여 세포성 mRNA(1030) 및 올리고뉴클레오타이드(1025)에 혼성화된 올리고뉴클레오타이드 프로브를 연장하는 데 사용될 수 있다.

[0634] 액적(1045a, 1045b)이 파괴된 후, 역전사효소에 의해 생성된 연장 산물은 풀링될 수 있고, 증폭 및 서열분석될 수 있다. 서열분석 판독물은 세포 표지, 분자 표지, 유전자 정제, 제어 입자 올리고뉴클레오타이드 서열 등의 역다중화 처리되어 단일 세포 유전자 발현 프로파일 및 작업 흐름의 전체 또는 일부의 정량 효율을 결정할 수 있다.

[0635] 단일 세포 서열분석 효율을 결정하기 위한 제어 올리고뉴클레오타이드

[0636] 또한, 본 명세서의 개시 내용은 단일 세포 서열분석 효율을 결정하기 위한 방법, 키트 및 시스템을 포함한다. 그러한 방법, 키트 및 시스템은 본 명세서에 개시된 임의의 적합한 방법, 키트 및 시스템, 예컨대 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 사용하여 세포 성분 발현 수준(예를 들어, 단백질 발현 수준)을 측정하기 위한 방법, 키트 및 시스템에, 또는 이들과 조합하여 사용될 수 있다.

[0637] 일부 구현예에서, 단일 세포를 올리고뉴클레오타이드와 접합된 항체로(예를 들어, 범용 항체 또는 바이오마커 항체로) 표지하고, 그와 함께 차세대 서열분석 라이브러리를 생성함으로써, NGS 판독물 내 올리고뉴클레오타이드로부터의 신호는 단일 세포 NGS 효율을 결정하는 데 사용될 수 있다. 이후 이는 상이한 단일 세포 서열분석 플랫폼에 대한 효율에 대한 평가 도구 또는 QC 단계로서 사용될 수 있다. 예를 들어, 제어 올리고뉴클레오타이드는 본 명세서에 개시된 임의의 적합한 방법, 키트 및 시스템, 예를 들어 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 사용하여 세포 성분 발현 수준(예를 들어, 단백질 발현 수준)을 측정하기 위한 방법, 키트 및 시스템에서 사용될 수 있다.

[0638] 올리고뉴클레오타이드와 접합된 항체(본 명세서에서 "Ab0"로 지칭됨)는 단일 세포 서열분석 제어로서 임의의 단일

세포 작업 흐름과 함께 사용될 수 있다. 단일 세포 작업 흐름은 마이크로웰 어레이 또는 마이크로웰 카트리지(예를 들어, BD Rhapsody™) 또는 마이크로유체학적 디바이스(예를 들어, 10X Genomics(미국 캘리포니아주 샌프란시스코 소재), Drop-seq(McCarroll Lab, Harvard Medical School(미국 매사추세츠주 캠프리지 소재); 그 내용이 전체가 본 명세서에 참고로 포함된 문헌(Macosko et al., Cell, 2015 May 21 16;5:1202)), 또는 Abseq(Mission Bio(미국 캘리포니아주 샌프란시스코 소재); 그 내용이 전체가 본 명세서에 참고로 포함된 문헌[Shahi et al., Sci Rep. 2017 Mar 14;7:44447])를, 바코드, 예컨대 추계적 바코드(예를 들어, BD Rhapsody, 또는 Drop-seq) 또는 방출가능한 바코드, 예컨대 추계적 바코드(예를 들어, 10X Genomics, 또는 Abseq)를 봉입하고 있는 봉괴성 하이드로겔 입자와 회합된 고체 또는 반고체 입자와 함께 사용할 수 있다. Ab0는 단일 세포 작업 흐름의 효율을 결정하기 위한 제어일 수 있다. 예를 들어, 10X Genomics로부터의 단일 세포 서열분석 플랫폼은 액적 내에 단일 세포를 캡슐화하기 위하여 에멀션을 사용하여 단일 세포 포획을 수행한다. 이들 액적은 쉽게 가시화될 수 없기 때문에, 단일 세포의 포획 효율은 쉽게 결정될 수 없다. 그러한 단일 세포 서열분석 작업 흐름의 업스트림에서 Ab0의 사용은, 사용자가 서열분석 후 단일 세포 포획 효율 및 더블렛 형성 속도를 평가할 수 있게 한다.

[0639] 도 11은 단일 세포 서열분석 효율의 결정을 위한 제어 올리고뉴클레오타이드-접합된 항체의 사용의 예시적인 작업 흐름의 개략도를 나타낸다. 일부 구현예에서, 하나 이상의 세포(예를 들어, 5000개)(1100)는 마이크로웰 카트리지 또는 어레이의 마이크로웰(1115) 상에 로딩되기 전에 제어 올리고뉴클레오타이드(1125)와 접합된 항체(1105)로 염색될 수 있다. 이어서, 세포(1110)는 복수의 구획, 예컨대 마이크로웰 어레이의 웰 내로 분리될 수 있으며, 단일 구획(1115)은 단일 세포 및 단일 비드(1120)에 맞도록 크기조정된다.

[0640] 비드는, 예를 들어 복수의 올리고뉴클레오타이드 프로브를 포함할 수 있으며, 이는 비드 상에서 모든 올리고뉴클레오타이드 프로브에 공통적인 세포 표지, 및 바코드 서열(예를 들어, 분자 표지 서열)을 포함할 수 있다. 일부 구현예에서, 각각의 올리고뉴클레오타이드 프로브는 표적-결합 영역, 예를 들어 폴리(dT) 서열을 포함할 수 있다. 항체(1105)에 접합된 올리고뉴클레오타이드(1125)는 화학적, 광학적 또는 기타 다른 수단을 사용하여 항체(1105)로부터 탈착될 수 있다. 세포(1110)는 용해되어(1135) 세포 내에 핵산, 예컨대 유전체 DNA 또는 세포성 mRNA(1130)를 방출할 수 있다. 세포성 mRNA(1130) 및 제어 올리고뉴클레오타이드(1125)는 비드(1120) 상에서 올리고뉴클레오타이드 프로브에 의해, 예를 들어 폴리(dT) 서열에 혼성화함으로써 포획될 수 있다. 비드는 회수될 수 있고, 포획된 세포성 mRNA(1130)(예를 들어, 총 약 5000개의 세포에 상응함)는 풀링될 수 있다.

[0641] 역전사효소는 세포성 mRNA(1130) 및 올리고뉴클레오타이드(1125)를 주형으로서 사용하여 세포성 mRNA(1130) 및 올리고뉴클레오타이드(1125)에 혼성화된 올리고뉴클레오타이드 프로브를 연장하는 데 사용될 수 있다. 역전사효소에 의해 생성된 연장 산물은 증폭 및 서열분석될 수 있다. 서열분석 관독물은 세포 표지, 바코드 서열(예를 들어, 분자 표지), 유전자 정체, 제어 올리고뉴클레오타이드 서열 등의 역다중화 처리되어 단일 세포 유전자 발현 프로파일 및 작업 흐름의 전체 또는 일부의 정량 효율(예를 들어, 세포 포획 효율)을 결정할 수 있다. 포획되고 라이브러리 제조를 성공적으로 통과한 세포의 수(예를 들어, 5000개 미만의 세포)가 결정될 수 있다.

[0642] 도 12는 단일 세포 서열분석 효율을 결정하기 위한 제어 올리고뉴클레오타이드-접합된 항체의 사용의 예시적인 작업 흐름의 또 다른 개략도를 나타낸다. 도 12에서, 단일 세포(1210a, 1210b) 및 단일 입자(1220a, 1220b)를 함유하는 액적(1245a, 1245b)은 마이크로유체 디바이스를 사용하여 형성될 수 있다. 단일 세포(1210a, 1210b)는 제어 올리고뉴클레오타이드(1225a, 1225b)와 접합된 항체(1205a, 1205b)에 결합될 수 있다. 액적(1245a, 1245b) 내 세포 용해 및 역전사 후, 액적은 파괴될 수 있으며, 내용물은 라이브러리 제조를 위해 풀링될 수 있다. 포획되고 라이브러리 제조를 성공적으로 통과한 세포의 수가 결정될 수 있다.

[0643] 도 13a 내지 도 13c는, 제어 올리고뉴클레오타이드가 세포 계수에 사용될 수 있음을 나타내는 플롯이다. 도 13a 및 도 13b는, Ab0의 제어 올리고뉴클레오타이드가 세포 계수를 위한 대조군으로서 사용될 수 있음을 나타낸다. mRNA 계수 플롯 및 제어 올리고뉴클레오타이드 계수 플롯의 하락 지점은 100% 포획 및 라이브러리 제조 효율이 달성된 경우와 일치할 수 있다. 도 13c는 종래의 mRNA-만의 세포 표지 호출(calling)의 사용이 전사적으로 높은 세포의 안개 속에서 전사적으로 낮은 세포를 놓칠 수 있음을 나타낸다. 이 방법은 n개의 세포 표지에서 컷오프를 호출할 수 있다. 이것은 활성화된 T 세포가 RNA 전사에서 몇 배로 더 높을 수 있는 활성화된 T 세포의 큰 개체군 내에서 휴지기 T 세포일 경우 일어날 수 있다. 이것은 또한 표적화된 패널(예를 들어, 암 패널)에서, 표적화된 유전자의 발현이 낮은 비-표적화된 세포(비-암 세포)가 감소되는 경우에 일어날 수 있다. 그러나, 단백질 발현이 훨씬 더 높기 때문에, 전사적으로 낮은 세포는 호출될 기회가 여전히 더 높다. 이러한 방법은 $m > n$ 인 경우, m개의 세포 표지에서 컷오프를 호출할 수 있다.

- [0644] 본 명세서의 개시 내용은 서열분석 제어를 위한 방법을 포함한다. 예를 들어, 방법은 단일 세포 서열분석 효율을 결정하는 데 사용될 수 있다. 단일 세포 서열분석 효율을 결정하기 위한 방법은 본 명세서에서 개시된 기타 다른 방법과 함께 사용될 수 있다. 예를 들어, 단일 세포 서열분석 효율에 사용된 방법은 단백질 발현을 결정하기 위한 방법과 함께 사용될 수 있다. 또 다른 예로서, 작업 흐름을 사용하여 단일 세포 서열분석 효율, 단백질 발현, 및/또는 유전자 발현을 결정할 수 있다.
- [0645] 일부 구현예에서, 방법은, 복수의 세포 중 하나 이상의 세포를 복수의 제어 조성물 중 하나의 제어 조성물과 접촉시키는 단계로서, 복수의 세포 중 하나의 세포는 복수의 표적 및 복수의 단백질 표적을 포함하고, 복수의 제어 조성물 각각은 제어 올리고뉴클레오타이드와 회합된 단백질 결합 시약을 포함하고, 단백질 결합 시약은 복수의 단백질 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 제어 올리고뉴클레오타이드는 제어 바코드 서열, 및 복수의 바코드 중 적어도 하나의 표적-결합 영역에 실질적으로 상보적인 서열을 포함하는 의사-표적 영역을 포함하는, 단계; 복수의 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지 서열), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 바코드 서열(예를 들어, 분자 표지 서열)은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 제어 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 적어도 하나의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징(예를 들어, 포획되고 라이브러리 제조를 성공적으로 통과한 세포의 수)을 결정하는 단계를 포함한다. 일부 구현예에서, 의사-표적 영역은 폴리(dA) 영역을 포함한다.
- [0646] 일부 구현예에서, 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열을 결정하는 단계; 및 복수의 바코딩된 제어 올리고뉴클레오타이드와 회합된 별개의 서열을 갖는 세포 표지 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 방법은, 결정된 하나 이상의 세포의 수에 기초하여, 단일 세포 포획 효율을 결정하는 단계를 포함할 수 있다. 방법은, 결정된 하나 이상의 세포의 수와 복수의 세포의 수의 비에 기초하여 단일 세포 포획 효율을 결정하는 것을 포함하는 단계를 포함할 수 있다.
- [0647] 일부 구현예에서, 서열분석 데이터에서 복수의 바코딩된 제어 올리고뉴클레오타이드의 특징을 사용하여 하나 이상의 세포의 적어도 하나의 특징을 결정하는 단계는, 서열분석 데이터에서 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열(예를 들어, 분자 표지 서열)의 수를 결정하는 단계; 및 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계를 포함한다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열의 수를 결정하는 단계는, 서열분석 데이터에서 각각의 세포 표지에 대하여, 세포 표지 및 제어 바코드 서열과 회합된 최대 수의 별개의 서열을 갖는 바코드 서열의 수를 결정하는 단계를 포함할 수 있다. 세포 표지 및 제어 바코드 서열과 회합된 별개의 서열을 갖는 바코드 서열의 수를 사용하여 하나 이상의 세포의 수를 결정하는 단계는, 최대 수의 별개의 서열을 갖는 바코드 서열의 수와, 최대 수의 별개의 서열을 갖는 바코드 서열의 수와 연관된 서열분석 데이터에서 세포 표지의 수의 플롯을 생성하는 단계; 및 플롯에서의 컷오프를 하나 이상의 세포의 수로서 결정하는 단계를 포함할 수 있다.
- [0648] 일부 구현예에서, 방법은 제어 올리고뉴클레오타이드를 바코딩하기 전에, 세포 성분 결합 시약(예를 들어, 단백질 결합 시약)으로부터 제어 올리고뉴클레오타이드를 방출하는 단계를 포함한다. 일부 구현예에서, 방법은 복수의 제어 조성물 중 결합되지 않은 제어 조성물을 제거하는 단계를 포함한다. 결합되지 않은 제어 조성물을 제거하는 단계는 복수의 세포 중 하나 이상의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 제어 조성물을 제거하는 단계는 제어 조성물 중의 적어도 하나의 세포 성분 결합 시약에 결합된 세포를 유세포분석을 사용하여 선택하는 단계를 포함할 수 있다.
- [0649] 일부 구현예에서, 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 추계적 바코드를 사용하여 제어 올리고뉴클레오타이드를 바코딩하여 복수의 추계적으로 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 일부 구현예에서, 복수의 바코드를 사용하여 복수의 제어 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 복수의 제어 조성물의 제어 올리고뉴클레오타이드와 접촉시켜 제어 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 제어 올리고뉴클레오타이드에 혼성화된 추계적 바코드를 연장하여 복수의 바코딩된 제어 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 증합효소, 역전사효소, 또는 이들의 조합을 사용하여 바코드를 연장하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은 복수의 바코딩된 제어 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바

코딩된 제어 올리고뉴클레오타이드를 증폭시키는 단계는 증합효소 연쇄 반응(PCR)을 사용하여, 추계적 바코드 서열(예를 들어, 분자 표지 서열)의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함한다. 서열분석 데이터를 획득하는 단계는 분자 표지 서열의 적어도 일부분 및 제어 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다.

[0650] 세포 오버로딩(overloading) 및 멀티플렛 식별

[0651] 본 명세서의 개시 내용은 또한 세포 오버로딩 및 멀티플렛을 식별하기 위한 방법, 키트 및 시스템을 포함한다. 그러한 방법, 키트 및 시스템은 본 명세서에 개시된 임의의 적합한 방법, 키트 및 시스템, 예를 들어 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 사용하여 세포 성분 발현 수준(예컨대, 단백질 발현 수준)을 측정하기 위한 방법, 키트 및 시스템에서, 또는 이와 조합되어 사용될 수 있다.

[0652] 현재 세포-로딩 기술을 사용하여, 약 20000개 세포가 약 60000개의 마이크로웰을 갖는 마이크로웰 카트리지를 또는 어레이 내로 로딩되는 경우, 2개 이상의 세포(더블렛 또는 멀티플렛으로 지칭됨)를 갖는 마이크로웰 또는 액적의 수는 최소일 수 있다. 그러나, 로딩된 세포의 수가 증가하는 경우, 다수의 세포를 갖는 마이크로웰 또는 액적의 수는 현저히 증가할 수 있다. 예를 들어, 약 50000개의 세포가 마이크로웰 카트리지를 또는 어레이의 약 60000개의 마이크로웰 내로 로딩되는 경우, 다수의 세포를 갖는 마이크로웰의 백분율은 예컨대 11 내지 14%로 상당히 높을 수 있다. 많은 수의 세포를 마이크로웰 내로 그렇게 로딩하는 것을 세포 오버로딩으로 지칭할 수 있다. 그러나, 세포가 여러 군(예를 들어, 5개)으로 나누어지고, 각각의 군 내의 세포가 별개의 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드로 표지되는 경우, 2개 이상의 샘플 인덱싱 서열과 회합된 세포 표지(예를 들어, 바코드, 예컨대 추계적 바코드의 세포 표지)는 서열분석 데이터 내에서 식별되고 후속 가공으로부터 제거될 수 있다. 일부 구현예에서, 세포는 다수의 군(예를 들어, 10000개)으로 나누어지고, 각각의 군 내의 세포가 별개의 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드로 표지되고, 2개 이상의 샘플 인덱싱 서열과 회합된 샘플 표지는 서열분석 데이터 내에서 식별되고 후속 가공으로부터 제거될 수 있다. 일부 구현예에서, 상이한 세포는 별개의 세포 식별 서열을 갖는 세포 식별 올리고뉴클레오타이드로 표지되고, 2개 이상의 세포 식별 올리고뉴클레오타이드와 회합된 세포 식별 서열은 서열분석 데이터에서 식별되고 후속 가공으로부터 제거될 수 있다. 마이크로웰 카트리지를 또는 어레이 내의 마이크로웰의 수에 비하여 그러한 더 많은 수의 세포가 마이크로웰 내로 로딩될 수 있다.

[0653] 본 명세서의 개시 내용은 샘플 식별을 위한 방법을 포함한다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 세포 성분을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지 서열), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 바코드 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열을 식별하는 단계; 및 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석(예를 들어, 단일 세포 mRNA 프로파일링, 또는 전체 전사체 분석)에서 제외하는 단계를 포함한다. 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 바코드 서열(예를 들어, 분자 표지 서열), 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함한다.

[0654] 예를 들어, 방법은 샘플 인덱싱을 사용하여 50000개 이상의 세포(10000 내지 20000개의 세포와 대비됨)를 로딩하는 데 사용될 수 있다. 샘플 인덱싱은 올리고뉴클레오타이드-접합된 세포 성분 결합 시약(예를 들어, 항체) 또는 세포 성분(예를 들어, 범용 단백질 마커)에 대한 세포 성분 결합 시약을 사용하여, 고유 샘플 인덱스로 상이한 샘플로부터의 세포를 표지할 수 있다. 상이한 샘플로부터의 2개 이상의 세포, 하나의 샘플의 세포의 상이한 개체군으로부터의 2개 이상의 세포, 또는 하나의 샘플의 2개 이상의 세포가 동일한 마이크로웰 또는 액적 내에 포획될 때, 조합된 "세포"(또는 2개 이상의 세포의 내용물)는 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올

리고뉴클레오타이드(또는 상이한 세포 식별 서열을 갖는 세포 식별 올리고뉴클레오타이드)와 회합될 수 있다. 상이한 세포 개체군의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 상이한 개체군의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 개체군의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100개일 수 있다. 각각의 개체군 내의 세포의 수 또는 평균수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 각각의 개체군 내의 세포의 수 또는 평균수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 각각의 개체군 내의 세포의 수 또는 평균수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100개일 수 있다. 각각의 개체군 내의 세포의 수 또는 평균수가 충분히 작은(예를 들어, 개체군당 세포수가 50, 25, 10, 5, 4, 3, 2, 또는 1개 이하) 경우, 세포 오버로딩 및 멀티플렛 식별을 위한 샘플 인덱싱 조성물은 세포 식별 조성물로 지칭될 수 있다.

[0655] 샘플의 세포는 샘플의 세포를 다수의 개체군으로 분류함으로써 다수의 개체군으로 나누어질 수 있다. 서열분석 데이터에서 하나 초과인 샘플 인덱싱 서열과 회합된 "세포"는 서열분석 데이터에서 하나의 세포 표지 서열(예를 들어, 바코드, 예컨대 추계적 바코드의 세포 표지 서열)과 회합된 2개 이상의 샘플 인덱싱 서열에 기초하여 "멀티플렛"으로서 식별될 수 있다. 조합된 "세포"의 서열분석 데이터가 또한 본 명세서에서 멀티플렛으로 지칭된다. 멀티플렛은 더블렛, 트리플렛, 쿼텟, 퀸텟, 섀스텟, 셉텟, 옥텟, 노넷, 또는 이들의 임의의 조합일 수 있다. 멀티플렛은 임의의 n -플렛일 수 있다. 일부 구현예에서, n 은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 이들 중 임의의 두 값 사이의 범위이거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 이들 중 임의의 두 값 사이의 범위이다. 일부 구현예에서, n 은 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20이다.

[0656] 단일 세포의 발현 프로파일을 결정할 때, 2개의 세포가 하나의 세포로서 식별될 수 있고, 2개의 세포의 발현 프로파일들은 하나의 세포에 대한 발현 프로파일(더블렛 발현 프로파일로 지칭됨)로서 식별될 수 있다. 예를 들어, 바코딩(예를 들어, 추계적 바코딩)을 사용하여 2개의 세포의 발현 프로파일들을 결정할 때, 2개의 세포의 mRNA 분자는 동일한 세포 표지를 갖는 바코드와 회합될 수 있다. 또 다른 예로서, 2개의 세포는 하나의 입자(예를 들어, 비드)와 회합될 수 있다. 입자는 동일한 세포 표지를 갖는 바코드를 포함할 수 있다. 세포를 용해시킨 후에, 2개의 세포 내의 mRNA 분자는 입자의 바코드와 회합될 수 있으며, 이에 따라 동일한 세포 표지와 회합될 수 있다. 더블렛 발현 프로파일은 발현 프로파일의 해석을 왜곡할 수 있다.

[0657] 더블렛은 상이한 샘플 인덱싱 서열을 갖는 2개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 더블렛은 또한 2개의 샘플 인덱싱 서열을 갖는 2개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 더블렛은 상이한 서열의 2개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된 2개의 세포(또는 2개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된 2개 이상의 세포)가 동일한 마이크로웰 또는 액적(droplet) 내에 포획되는 경우 발생할 수 있으며, 이때 조합된 "세포"는 상이한 샘플 인덱싱 서열을 갖는 2개의 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다. 트리플렛은, 모두 상이한 샘플 인덱싱 서열을 갖는 3개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 3개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 쿼텟은, 모두 상이한 샘플 인덱싱 서열을 갖는 4개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 4개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 퀸텟은, 모두 상이한 샘플 인덱싱 서열을 갖는 5개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 5개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 섀스텟은, 모두 상이한 샘플 인덱싱 서열을 갖는 6개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 6개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 셉텟은, 모두 상이한 샘플 인덱싱 서열을 갖는 7개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 7개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 옥텟은, 모두 상이한 샘플 인덱싱 서열을 갖는 8개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 8개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 노넷은, 모두 상이한 샘플 인덱싱 서열을 갖는 9개의 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포", 또는 9개의 상이한 샘플 인덱싱

서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된, 조합된 "세포"를 지칭할 수 있다. 멀티플렛은 상이한 서열의 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드와 회합된 2개 이상의 세포(또는 2개 이상의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합된 2개 이상의 세포)가 동일한 마이크로웰 또는 액적 내에 포획되는 경우 발생할 수 있으며, 이때 조합된 "세포"는 2개 이상의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드와 회합될 수 있다.

[0658] 또 다른 예에서, 방법은 샘플 오버로딩의 맥락에서 또는 마이크로웰 어레이의 마이크로웰 상에 세포를 로딩하거나 세포를 함유하는 액적을 생성하는 맥락에서의 여부와 관계 없이, 멀티플렛 식별을 위해 사용될 수 있다. 2개 이상의 세포가 하나의 마이크로웰 내에 로딩되는 경우, 조합된 "세포"(또는 2개 이상의 세포의 내용물)로부터의 결과 데이터는 이상 유전자 발현 프로파일을 갖는 멀티플렛이다. 샘플 인덱싱을 사용함으로써, 상이한 샘플 인덱싱 서열을 갖는 2개 이상의 샘플 인덱싱 올리고뉴클레오타이드(또는 2개 이상의 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드)와 각각 회합되거나 이에 배정된 세포 표지를 찾음으로써 이들 멀티플렛 중 일부를 인식할 수 있다. 샘플 인덱싱 서열을 사용하여, 본 명세서에 개시된 방법은 (샘플 오버로딩 여부와 관련해서든 마이크로웰 어레이의 마이크로웰 상에 세포를 로딩하거나 세포를 함유하는 액적을 생성하는 것과 관련해서든 어느 것이든 간에) 멀티플렛 식별에 사용될 수 있다. 일부 구현예에서, 방법은, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계로서, 제1 복수의 세포 각각 및 제2 복수의 세포 각각은 하나 이상의 세포 성분을 포함하고, 2개의 샘플 인덱싱 조성물 각각은 샘플 인덱싱 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 중 적어도 하나에 특이적으로 결합할 수 있고, 샘플 인덱싱 올리고뉴클레오타이드는 샘플 인덱싱 서열을 포함하고, 2개의 샘플 인덱싱 조성물의 샘플 인덱싱 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지 서열), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 바코드 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 샘플 인덱싱 서열과 각각 회합된 하나 이상의 다중 세포 표지 서열을 식별하는 단계를 포함한다.

[0659] 마이크로웰 카트리지의 마이크로 웰 상으로 또는 미세유체 장치를 사용하여 생성된 액적 내로 로딩될 수 있는 세포의 수는 멀티플렛 비율(rate)에 의해 제한될 수 있다. 더 많은 세포의 로딩은 더 많은 멀티플렛을 생성할 수 있으며, 이는 단일 세포 데이터에서 노이즈를 식별하고 생성하는 것이 어려울 수 있다. 샘플 인덱싱을 사용하여, 방법은 멀티플렛을 더 정확하게 표지하거나 식별하고, 서열분석 데이터 또는 후속 분석에서 멀티플렛을 제거하는 데 사용될 수 있다. 더 높은 신뢰도로 멀티플렛을 식별할 수 있다면, 멀티플렛 비율에 대한 사용자 허용치가 증가하고, 각각의 마이크로웰 카트리지 상에 더 많은 세포를 로딩하거나 적어도 하나의 세포를 각각 갖는 액적을 생성할 수 있다.

[0660] 일부 구현예에서, 제1 복수의 세포 및 제2 복수의 세포를 2개의 샘플 인덱싱 조성물과 각각 접촉시키는 단계는, 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제1 샘플 인덱싱 조성물과 접촉시키는 단계; 및 제1 복수의 세포를 2개의 샘플 인덱싱 조성물 중 제2 샘플 인덱싱 조성물과 접촉시키는 단계를 포함한다. 복수의 세포의 수와 복수의 샘플 인덱싱 조성물의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 복수의 세포 및/또는 샘플 인덱싱 조성물의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 복수의 세포 및/또는 샘플 인덱싱 조성물의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개일 수 있다. 세포의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 세포의 수 또는 평균수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 수 또는 평균수, 또는 세포는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000,

10000, 100000, 또는 1000000개일 수 있다.

- [0661] 일부 구현예에서, 방법은, 2개의 샘플 인덱싱 조성물 중 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계를 포함한다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 제1 복수의 세포 및 제2 복수의 세포 중의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 샘플 인덱싱 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 샘플 인덱싱 조성물의 적어도 하나의 세포 성분 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터의 하나 이상의 세포를 용해시키는 단계를 포함한다.
- [0662] 일부 구현예에서, 샘플 인덱싱 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 샘플 인덱싱 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 샘플 인덱싱 올리고뉴클레오타이드를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다.
- [0663] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하는 단계는 복수의 바코드를 샘플 인덱싱 올리고뉴클레오타이드와 접촉시켜 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 샘플 인덱싱 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0664] 일부 구현예에서, 방법은, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 바코드 서열(예를 들어, 분자 표지 서열)의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 바코드 서열의 적어도 일부분 및 샘플 인덱싱 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함한다.
- [0665] 일부 구현예에서, 복수의 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 샘플 인덱싱 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 샘플 인덱싱 올리고뉴클레오타이드를 생성하는 단계를 포함한다.
- [0666] 일부 구현예에서, 방법은, 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 표적의 복제물을 바코드의 표적-결합 영역과 접촉시키는 단계; 및 복수의 바코드를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다.
- [0667] 일부 구현예에서, 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함한다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.
- [0668] 일부 구현예에서, 세포 식별을 위한 방법은, 제1 복수의 하나 이상의 세포 및 제2 복수의 하나 이상의 세포를 2개의 세포 식별 조성물과 각각 접촉시키는 단계로서, 제1 복수의 하나 이상의 세포 각각 및 제2 복수의 하나 이상의 세포 각각은 하나 이상의 세포 성분을 포함하고, 2개의 세포 식별 조성물 각각은 세포 식별 올리고뉴클레

오티드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 중 적어도 하나에 특이적으로 결합할 수 있고, 세포 식별 올리고뉴클레오티드는 세포 식별 서열을 포함하고, 2개의 세포 식별 조성물의 세포 식별 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 세포 식별 올리고뉴클레오티드를 바코딩하여 복수의 바코딩된 세포 식별 올리고뉴클레오티드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지 서열), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 바코드 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 세포 식별 올리고뉴클레오티드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 세포 식별 서열과 각각 회합된 하나 이상의 세포 표지 서열을 식별하는 단계; 및 2개 이상의 세포 식별 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 획득된 서열분석 데이터로부터 제거하는 단계 및/또는 2개 이상의 세포 식별 서열과 각각 회합된 하나 이상의 세포 표지 서열과 연관된 서열분석 데이터를 후속 분석(예를 들어, 단일 세포 mRNA 프로파일링, 또는 전체 전사체 분석)에서 제외하는 단계를 포함한다. 일부 구현예에서, 세포 식별 올리고뉴클레오티드는 바코드 서열(예를 들어, 분자 표지 서열), 범용 프라이머를 위한 결합 부위, 또는 이들의 조합을 포함할 수 있다.

[0669] 멀티플렛(예를 들어, 더블렛, 트리플렛 등)은 상이한 서열의 2개 이상의 세포 식별 올리고뉴클레오티드와 회합된 2개 이상의 세포(또는 2개 이상의 상이한 세포 식별 서열을 갖는 세포 식별 올리고뉴클레오티드와 회합된 2개 이상의 세포)가 동일한 마이크로웰 또는 액적 내에 포획되는 경우 발생할 수 있으며, 이때 조합된 "세포"는 2개 이상의 상이한 세포 식별 서열을 갖는 세포 식별 올리고뉴클레오티드와 회합될 수 있다.

[0670] 세포 식별 조성물은 세포 오버로딩과 관련해서든 마이크로웰 어레이의 마이크로웰 상에 세포를 로딩하거나 세포를 함유하는 액적을 생성하는 것과 관련해서든 어느 것이든 간에, 멀티플렛 식별을 위해 사용될 수 있다. 2개 이상의 세포가 하나의 마이크로웰 내에 로딩되는 경우, 조합된 "세포"(또는 2개 이상의 세포의 내용물)로부터의 결과 데이터는 이상 유전자 발현 프로파일을 갖는 멀티플렛이다. 세포 식별을 사용함으로써, 상이한 세포 식별 서열을 갖는 2개 이상의 세포 식별 올리고뉴클레오티드(또는 2개 이상의 세포 식별 서열을 갖는 세포 식별 올리고뉴클레오티드)와 각각 회합되거나 이에 배정된 세포 표지(예를 들어, 바코드, 예컨대 체계적 바코드의 세포 표지)를 찾음으로써 이들 멀티플렛 중 일부를 인식할 수 있다. 본 명세서에 개시된 방법은 세포 식별 서열을 이용하여, (샘플 오버로딩 여부의 맥락에서, 또는 마이크로웰 어레이의 마이크로웰 상에 세포를 로딩하거나 세포를 함유하는 액적의 생성의 맥락에서의 여부에 관계없이) 멀티플렛 식별에 사용될 수 있다. 일부 구현예에서, 방법은, 제1 복수의 하나 이상의 세포 및 제2 복수의 하나 이상의 세포를 2개의 세포 식별 조성물과 각각 접촉시키는 단계로서, 제1 복수의 하나 이상의 세포 각각 및 제2 복수의 하나 이상의 세포 각각은 하나 이상의 세포 성분을 포함하고, 2개의 세포 식별 조성물 각각은 세포 식별 올리고뉴클레오티드와 회합된 세포 성분 결합 시약을 포함하고, 세포 성분 결합 시약은 하나 이상의 세포 성분 표적 중 적어도 하나에 특이적으로 결합할 수 있고, 세포 식별 올리고뉴클레오티드는 세포 식별 서열을 포함하고, 2개의 세포 식별 조성물의 세포 식별 서열은 상이한 서열을 포함하는, 단계; 복수의 바코드를 사용하여 세포 식별 올리고뉴클레오티드를 바코딩하여 복수의 바코딩된 세포 식별 올리고뉴클레오티드를 생성하는 단계로서, 복수의 바코드 각각은 세포 표지 서열, 바코드 서열(예를 들어, 분자 표지 서열), 및 표적-결합 영역을 포함하고, 복수의 바코드 중 적어도 2개의 바코드의 바코드 서열은 상이한 서열을 포함하고, 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열을 포함하는, 단계; 복수의 바코딩된 세포 식별 올리고뉴클레오티드의 서열분석 데이터를 획득하는 단계; 및 획득된 서열분석 데이터에서 2개 이상의 세포 식별 서열과 각각 회합된 하나 이상의 다중 세포 표지 서열을 식별하는 단계를 포함한다.

[0671] 마이크로웰 카트리지의 마이크로 웰 상으로 또는 마이크로유체학적 디바이스를 사용하여 생성된 액적 내로 로딩될 수 있는 세포의 수는 멀티플렛 비율에 의해 제한될 수 있다. 더 많은 세포의 로딩은 더 많은 멀티플렛을 생성할 수 있으며, 단일 세포 데이터에서 노이즈를 식별하고 생성하는 것이 어려울 수 있다. 방법은 세포 식별을 이용하여, 멀티플렛을 더 정확하게 표지하거나 식별하고, 서열분석 데이터 또는 후속 분석에서 멀티플렛을 제거하는 데 사용될 수 있다. 더 높은 신뢰도로 멀티플렛을 식별할 수 있다면, 멀티플렛 비율에 대한 사용자 허용치가 증가하고, 각각의 마이크로웰 카트리지 상에 더 많은 세포를 로딩하거나 적어도 하나의 세포를 각각 가진 액적을 생성할 수 있다.

[0672] 일부 구현예에서, 제1 복수의 하나 이상의 세포 및 제2 복수의 하나 이상의 세포를 2개의 세포 식별 조성물과 각각 접촉시키는 단계는, 제1 복수의 하나 이상의 세포를 2개의 세포 식별 조성물 중 제1 세포 식별 조성물과 접촉시키는 단계; 및 제1 복수의 하나 이상의 세포를 2개의 세포 식별 조성물 중 제2 세포 식별 조성물과 접촉

시키는 단계를 포함한다. 복수의 세포 식별 조성물의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 세포 식별 조성물의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 2개의 값들 사이의 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 2개의 값들 사이의 수 또는 범위일 수 있다. 일부 구현예에서, 세포 식별 조성물의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 또는 1000000개일 수 있다. 각각의 복수의 하나 이상의 세포 내의 세포의 수 또는 평균수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 각각의 복수의 하나 이상의 세포 내의 세포의 수 또는 평균수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 100000, 1000000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 각각의 복수의 하나 이상의 세포 내의 세포의 수 또는 평균수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 10000, 또는 1000000개일 수 있다.

[0673] 일부 구현예에서, 방법은, 2개의 세포 식별 조성물 중 결합되지 않은 세포 식별 조성물을 제거하는 단계를 포함한다. 결합되지 않은 세포 식별 조성물을 제거하는 단계는 제1 복수의 하나 이상의 세포 및 제2 복수의 하나 이상의 세포의 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않은 세포 식별 조성물을 제거하는 단계는 유세포분석을 사용하여 2개의 세포 식별 조성물 중 적어도 하나의 세포 성분 결합 시약에 결합된 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 복수의 샘플 각각으로부터 하나 이상의 세포를 용해시키는 단계를 포함한다.

[0674] 일부 구현예에서, 세포 식별 올리고뉴클레오타이드는 세포 성분 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은 세포 성분 결합 시약으로부터 세포 식별 올리고뉴클레오타이드를 탈착시키는 단계를 포함할 수 있다. 세포 식별 올리고뉴클레오타이드를 탈착시키는 단계는 UV 광절단, 화학적 처리(예를 들어, 환원성 시약, 예컨대 디티오프레이톨을 사용함), 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 세포 성분 결합 시약으로부터 세포 식별 올리고뉴클레오타이드를 탈착시키는 단계를 포함할 수 있다.

[0675] 일부 구현예에서, 복수의 바코드를 사용하여 세포 식별 올리고뉴클레오타이드를 바코딩하는 단계는, 복수의 바코드를 세포 식별 올리고뉴클레오타이드와 접촉시켜 세포 식별 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 세포 식별 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 생성하는 단계를 포함한다. 바코드를 연장하는 단계는 DNA 중합효소를 사용하여 바코드를 연장하여 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다. 바코드를 연장하는 단계는 역전사효소를 사용하여 바코드를 연장하여 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.

[0676] 일부 구현예에서, 방법은, 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함한다. 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 바코드 서열(예를 들어, 분자 표지 서열)의 적어도 일부분 및 세포 식별 올리고뉴클레오타이드의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 세포 식별 올리고뉴클레오타이드의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 바코드 서열의 적어도 일부분 및 세포 식별 올리고뉴클레오타이드의 적어도 일부분을 서열분석하는 단계를 포함한다. 일부 구현예에서, 적어도 하나의 세포의 샘플 기원을 식별하는 단계는 적어도 하나의 바코딩된 세포 식별 올리고뉴클레오타이드의 세포 식별 서열에 기초하여 복수의 바코딩된 표적의 샘플 기원을 식별하는 단계를 포함한다.

[0677] 일부 구현예에서, 복수의 바코드를 사용하여 세포 식별 올리고뉴클레오타이드를 바코딩하여 복수의 바코딩된 세포 식별 올리고뉴클레오타이드를 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포 식별 올리고뉴클레오타이드를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 세포 식별 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0678] 세포 성분-세포 성분 상호작용의 결정

- [0679] 본 명세서의 개시 내용은 세포 성분 사이의 상호작용을 결정하기 위한 방법, 예를 들어 다중화된 근접 결찰 방법을 포함한다. 그러한 방법, 키트 및 시스템은 본 명세서에 개시된 임의의 적합한 방법, 키트 및 시스템, 예를 들어 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 사용하여 세포 성분 발현 수준(예컨대, 단백질 발현 수준)을 측정하기 위한 방법, 키트 및 시스템에, 또는 이들과 조합하여 사용될 수 있다.
- [0680] 본 명세서에 개시된 방법은 하나 또는 다수의 세포 성분(예를 들어, 단백질) 사이의 상호작용을 검출, 식별, 결정, 및/또는 측정하는 데 사용될 수 있다. 예를 들어, 방법은 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 200, 300개, 또는 그 이상의 쌍의 세포 성분 사이의 상호작용을 검출, 식별, 결정, 및/또는 측정하는 데 사용될 수 있다. 하나의 쌍의 세포 성분 표적 내의 세포 성분 표적은 다른 쌍의 세포 성분 표적 내의 세포 성분 표적과 중복될 수 있다. 예를 들어, 제1 쌍의 세포 성분 표적은 단백질 A 및 단백질 B를 포함할 수 있고, 제2 쌍의 세포 성분 표적은 단백질 A 및 단백질 C를 포함할 수 있다.
- [0681] 일부 구현예에서, 방법은 상호작용 결정 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 사용하여 세포 성분 결합 시약의 표적(본 명세서에서 세포 성분 표적으로 지칭됨) 사이의 상호작용을 결정한다. 예를 들어, 방법은 항체-올리고뉴클레오타이드 접합체를 사용하여 단일 세포 내의 단백질-단백질 상호작용을 검출할 수 있다. 항체는 동종이량체 또는 이종이량체일 수 있다. 각각의 세포 성분 결합 시약은 고유 상호작용 결정 서열을 함유하는 상호작용 결정 올리고뉴클레오타이드와 회합됨(예를 들어, 그에 부착되거나 결합할) 수 있다. 예를 들어, 각각의 항체는 고유 바코드 서열을 함유하는 올리고뉴클레오타이드에 접합될 수 있다. 함께 결찰된 상호작용 결정 올리고뉴클레오타이드의 쌍을 결정함으로써, 서로 상호작용하는 세포 성분 표적을 밝혀낼 수 있다. 예를 들어, 함께 결찰된 바코드 쌍을 결정함으로써, 서로 상호작용하는 단백질을 밝혀낼 수 있다. 이 방법은 또한 단일 세포 mRNA 서열분석 방법, 예컨대 웰-기반 및 액적-기반 단일 세포 mRNA 서열분석 방법과 함께 사용될 수 있다.
- [0682] 일부 구현예에서, 방법은 선형 증폭을 이용하여 서열분석을 위한 앰플리콘을 생성할 수 있다. 방법은 원형 주형을 생성하지 않을 수 있으며, 회전 환 증폭(rolling circle amplification)을 사용하여 형광 프로브 또는 기타 다른 방법에 의한 하류 검출을 위한 충분한 주형을 생성할 수 있다. 방법은 단일 세포 내의 세포 성분 표적 사이의 상호작용을 결정하는 데 사용될 수 있다. 예를 들어, 방법은 수천 내지 수만 개의 단일 세포 전체에 걸쳐 수백 개의 세포 성분-세포 성분 상호작용(예를 들어, 단백질-단백질 상호작용)을 검출하는 데 사용될 수 있으며, 판독치로서 서열분석을 사용한다. 동일한 단백질 또는 세포 성분 표적 사이의 상호작용이 또한 결정될 수 있다. 방법은, 입력으로서 벌크 RNA를 필요로 하는 형광 및 정량적 PCR(qPCR)을 판독치로서 이용하지 않으며, 단지 적은 수의 세포(예를 들어, 100개 미만의 세포)만을 한번에 처리할 수 있으며, 제한된 양의 세포 성분-세포 성분 상호작용을 한번에 검출할 수 있다. 다중화된 방법은 항체와 회합되거나 그에 접합된 바코딩된 올리고뉴클레오타이드를 이용할 수 있으며, 세포 성분 결합 시약은 다수의 단일 세포 내의 세포 성분-세포 성분 상호작용을 고처리량 방식으로 결정하는 데 사용될 수 있다.
- [0683] 더욱이, 방법은 또한 기타 다른 단일 세포 RNA 서열분석 방법과 함께 사용될 수 있다. 예를 들어, mRNA 발현 수준, 단백질 발현 수준(또는 세포 성분 표적의 수준), 및/또는 단백질-단백질 상호작용(또는 세포 성분 사이의 상호작용)을 결정하기 위하여 하나의 단일 작업 흐름이 사용될 수 있다. 결과적으로, 단일 실험에서, 단일 세포(또는 다수의 단일 세포) 내로부터의 mRNA 발현, 단백질 발현, 및 단백질-단백질 상호작용에 관한 데이터가 획득될 수 있다.
- [0684] 일부 구현예에서, 방법은 2개의 세포 사이의 세포 성분-세포 성분 상호작용 및 세포내 세포 성분-세포 성분 상호작용을 검출하는 데 사용될 수 있다. 예를 들어, 방법은 2개의 세포 사이의 단백질-단백질 상호작용뿐만 아니라, 세포내 단백질-단백질 상호작용을 검출하는 데 사용될 수 있다. 방법은 항체 염색 또는 인큐베이션(예를 들어, 세포내 단백질 또는 표적에 대한 항체의 결합을 가능하게 하는 세포 고정) 전에 추가의 세포 제조를 필요로 할 수 있다.
- [0685] 도 14a 내지 도 14f는 한 쌍의 상호작용 결정 조성물을 사용하여 세포 성분(예를 들어, 단백질) 사이의 상호작용을 결정하는 예시적인 작업 흐름의 개략도를 나타낸다. 방법은 도 14a에 예시된 2가지 유형의 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)(예를 들어, 항체-올리고뉴클레오타이드 접합체)을 이용할 수 있다. 2가지 유형의 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)은 2가지 유형의 올리고뉴클레오타이드(1412a, 1412b)와 회합된 2개의 동일하거나 상이한 항체(1408a, 1408b)를 포함할 수 있다. 회합(예를 들어, 접합)은 가역적 또는 비가역적일(예를 들어, 절단 가능 또는 절단 불가능할) 수 있다. 한 가지 유형의 올리고뉴클레오타이드(1412a)(본 명세서에서 유형 1 올리고뉴클레오타이드(1412a)로 지칭됨)는 범용 서열분석 영역(1416)을 가질 수 있다. 일부 구현예에서, 범용 서열분석 영역(1416)은 범용 프라이머에 의한 증폭을 위하여 모든 항체-올리고뉴클레오타이드 조성물

에 공통될 수 있다. 기타 다른 유형의 올리고뉴클레오타이드(1412b)(본 명세서에서 유형 2 올리고뉴클레오타이드(1412b)로 지칭됨)는 바코드, 예컨대 추계적 바코드에 대한 혼성화를 위한 폴리(dA) 서열(1420)을 가질 수 있다. 예를 들어, 유형 2 올리고뉴클레오타이드(1412b)의 폴리(dA) 서열(1420)은 비드와 회합된(예를 들어, 비드 상에 고정화된, 비드 상에 부분적으로 고정화된, 방출 가능하게 비드 내에 봉입된, 비드 내에 부분적으로 봉입된, 또는 이들이 임의로 조합된) 추계적 바코드에 혼성화될 수 있다. 올리고뉴클레오타이드 유형 둘 모두(1412a, 1412b)는 또한 각각의 항체(1404a, 1404b)에 고유한 바코드 서열(1424a, 1424b)을 포함한다. 2가지 올리고뉴클레오타이드 유형(1412a, 1412b)은 도 14b에 예시된 가교 올리고(1432)에 대한 혼성화를 위한 상이한 서열(1428a, 1428b)을 포함할 수 있다. 유형 1 올리고뉴클레오타이드(1412a) 및 유형 2 올리고뉴클레오타이드(1412b)는 5'에서 3' 방향으로 및 3'에서 5' 방향으로 항체(1408a, 1408b)와 회합되는 것으로 나타나 있지만, 이들은 단지 예시적이며 제한적이고자 하지 않는다. 일부 구현예에서, 유형 1 올리고뉴클레오타이드(1412a) 및 유형 2 올리고뉴클레오타이드(1412b)는 3'에서 5' 방향으로 및 5'에서 3' 방향으로 항체(1408a, 1408b)와 회합될 수 있다.

[0686] 도 14b를 참조하면, 가교 올리고뉴클레오타이드(1432)는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)에 결합하여 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)을 결합을 위하여 함께 합칠 수 있다. 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)의 서열은 각각의 조성물(1404a, 1404b)에 대해 상이할 수 있다. 일부 구현예에서, 상이한 가교 올리고뉴클레오타이드(1432)가 각각의 항체 쌍에 대해 사용될 수 있다.

[0687] 방법은 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 쌍을 함께 합쳐서, 회합된 올리고뉴클레오타이드(1412a, 1412b)가 서로 함께 결합되거나 결합될 수 있도록 하고, 추후에 서열분석을 위하여 비드와 회합된 바코드(예를 들어, 추계적 바코드)에 의해 포획될 수 있도록 하는 것에 기초할 수 있다. 올리고뉴클레오타이드(1412a, 1412b)는 항체(1408a, 1408b)의 표적(1436a, 1436b)이 서로의 특성의 역치 거리 또는 범위 이내, 예컨대 30 내지 40 nm에 있다면 서로 결합될 수 있다.

[0688] 다중화된 방법은 항체 또는 세포 성분 결합 시약과 회합되거나 그에 접합된 바코딩된 올리고뉴클레오타이드를 이용하여, 다수의 단일 세포 내의 단백질-단백질 상호작용 또는 세포 성분 사이의 상호작용을 고처리량 방식으로 결정할 수 있다. 일부 구현예에서, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위, 또는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 세포 내의 세포 성분 사이의 상호작용이 세포 작업 흐름에서 결정될 수 있다. 일부 구현예에서, 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 10^5 , 10^6 , 10^7 , 10^8 , 10^9 개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 세포 내의 세포 성분 사이의 상호작용이 세포 작업 흐름에서 결정될 수 있다.

[0689] 염색

[0690] 본 검정을 수행하기 위하여, 먼저 항체-올리고뉴클레오타이드 조성물(1404a, 1404b) 또는 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 쌍(또는 세포 성분 결합 시약-올리고뉴클레오타이드 조성물)을 이용하여 세포가 염색될 수 있다. 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)(또는 세포 성분 결합 시약)의 수는 상이한 실시예에서 상이할 수 있다. 일부 구현예에서, 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다.

[0691] 상이한 조성물(1404a, 1404b)의 항체(1408a, 1408b)(또는 세포 성분 결합 시약)는 상이하거나 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합할 수 있

다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 백분율은 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 백분율은 적어도, 또는 최대 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0692]

항체(1408a, 1408b)(또는 세포 성분 결합 시약) 중 일부 또는 전부가 동일한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합할 수 있거나, 이들 중 어느 것도 그럴 수 없다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 항체(1408a, 1408b)의 수는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 항체(1408a, 1408b)의 수는 적어도, 또는 최대 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 항체(1408a, 1408b)의 백분율은 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 항체(1408a, 1408b)의 백분율은 적어도, 또는 최대 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0693]

항체(1408a)(또는 세포 성분 결합 시약) 중 일부 또는 전부가 동일한 서열 또는 상이한 서열을 갖는 범용 서열 분석 영역을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합할 수 있거나, 이들 중 어느 것도 그럴 수 없다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열 분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 항체(1408a)의 수는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열 분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 항체(1408a)의 수는 적어도, 또는 최대 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90,

100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 항체(1408a)의 백분율은 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 항체(1408a)의 수는 적어도, 또는 최대 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0694] 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)(세포 성분 결합 시약-올리고뉴클레오타이드 조성물)의 쌍의 수는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 쌍의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 항체-올리고뉴클레오타이드 조성물(1404a, 1404b)의 쌍의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다.

[0695] 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)(또는 세포 성분 결합 시약)는 상이하거나 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합할 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 적어도, 또는 최대 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 백분율은 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 상이한 바코드 서열 또는 동일한 바코드 서열(1424a, 1424b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합하거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 백분율은 적어도, 또는 최대 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0696] 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)(또는 세포 성분 결합 시약) 중 일부 또는 전부가 동일한 가교 올리고뉴클레오타이드 혼성화 영역을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합할 수 있거나, 이들 중 어느 것도 그럴 수 없다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 수는 적어도, 또는 최대 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200,

300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 %는 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합되거나 그것에 결합하는 상이한 쌍의 조성물(1404a, 1404b)의 항체(1408a, 1408b)의 백분율은 적어도, 또는 최대 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0697]

상이한 조성물(1404a)의 항체(1408a)(또는 세포 성분 결합 시약) 중 일부 또는 전부가 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합할 수 있거나, 이들 중 어느 것도 그럴 수 없다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 상이한 조성물(1404a)의 항체(1408a)의 수는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 상이한 조성물(1404a)의 항체(1408a)의 수는 적어도, 또는 최대 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 상이한 조성물(1404a)의 항체(1408a)의 백분율은 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 100%, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일한 서열 또는 상이한 서열을 갖는 범용 서열분석 영역(1416a)을 갖는 상호작용 결정 올리고뉴클레오타이드(1412a)와 회합되거나 그것에 결합하는 상이한 조성물(1404a)의 항체(1408a)의 백분율은 적어도, 또는 최대 0%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 33%, 40%, 50%, 60%, 70%, 80%, 90%, 99.9%, 99.99%, 또는 100%일 수 있다.

[0698]

표적(1436a, 1436b)의 수는, 예컨대 2개 내지 적어도 50 내지 100개의 단백질 표적으로 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 표적(1436a, 1436b)의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 표적(1436a, 1436b)의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000개일 수 있다.

[0699]

결찰

[0700]

결합되지 않은 조성물(1404a, 1404b)을 세척한 후에, 가교 올리고뉴클레오타이드(1432)가 첨가될 수 있다. 유형 1 조성물과 유형 2 조성물(1404a, 1404b)로 이루어진 쌍은 가교 올리고뉴클레오타이드 혼성화 영역에 혼성화할 수 있는 가교 올리고뉴클레오타이드(1432)를 통해 함께 연결될 수 있다. 예를 들어, 유형 1 조성물과 유형 2 조성물(1404a, 1404b)로 이루어진 쌍은, 상응하는 단백질 표적(1436a, 1436b)이 역치 거리 이내, 예컨대 30 내지 40 nm에 있다면, 가교 올리고뉴클레오타이드 혼성화 영역에 혼성화할 수 있는 가교 올리고뉴클레오타이드(1432)를 통해 함께 연결될 수 있다. 이어서, DNA 리가제를 첨가하여 2개의 올리고뉴클레오타이드(1412a, 1412b)를 결합하여, 결합된 올리고뉴클레오타이드(1412)를 형성할 수 있으며, 그 결과 조성물(1436a, 1436b)은 함께 결합된다. 예를 들어, 가교 올리고뉴클레오타이드(1432)는 결찰을 촉진시키기 위하여 결찰 영역이 이중 가닥을 갖게 할 수 있다. 가교 올리고뉴클레오타이드(1432)는 각각의 올리고뉴클레오타이드(1412a, 1412b) 상의 혼성화 영역(1428a, 1428b)에

결합할 수 있으며, 이는 DNA 리가제가 2개의 올리고뉴클레오타이드(1412a, 1412b)를 함께 결합할 수 있게 한다.

[0701] 역치 거리는 상이한 실시에서 상이할 수 있다. 일부 구현예에서, 역치 거리는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 nm, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 nm, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 예를 들어, 거리는 30 내지 40 nm의 범위일 수 있다. 일부 구현예에서, 역치 거리는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 또는 1000 nm일 수 있다. 역치 거리는 올리고뉴클레오타이드(1412a, 1412b)의 길이, 항체(1408a, 1408b)에 대한 올리고뉴클레오타이드(1412a, 1412b)의 회합 또는 부착의 위치, 표적(1436a, 1436b)의 크기, 또는 이들의 임의의 조합에 의해 영향을 받을 수 있다.

[0702] 세포 용해 및 결합된 올리고뉴클레오타이드의 바코드에 대한 혼성화

[0703] 이어서, 세포를 용해시켜, 도 14c에 예시된 바와 같이 결합된 올리고뉴클레오타이드(1412)의 포획을 가능하게 할 수 있다. 예를 들어, 폴리(dT) 영역(1444)을 갖는 바코드(1440)(예를 들어, 추계적 바코드)를 갖는 비드(1438)는 폴리(dA) 영역을 통해 결합된 올리고뉴클레오타이드(1412)를 포획할 수 있다. 비드(1438)는 표지, 예컨대 세포 표지 및/또는 분자 표지(1448) 및 하위서열 증폭을 위한 범용 서열(1452)을 추가로 포함할 수 있다.

[0704] 역전사

[0705] 이어서, 역전사를 수행하여 도 14d에 나타난 결합된 올리고뉴클레오타이드(1412)의 상보적 서열(1456)을 생성할 수 있다. 예를 들어, 몰로니 뮤린 백혈병 바이러스(MMLV)-기반 또는 Taq-기반 역전사효소가 결합된 올리고뉴클레오타이드(1412)의 상보적 서열을 생성하는 데 사용될 수 있다. MMLV-기반 역전사효소는 상보적 가닥의 생성 동안 가교 올리고뉴클레오타이드(1432)를 변위시킬 가닥-변위 메커니즘을 갖는다. Taq-기반 역전사효소는 가교 올리고뉴클레오타이드(1432)를 제거할 5'에서 3'으로의 엑소뉴클레아제 활성을 갖는다.

[0706] 라이브러리 제조

[0707] 역전사 후에는, 예를 들어 단일 세포 mRNA 서열분석에 대한 표준 프로토콜에 따라 서열분석 라이브러리가 제조될 수 있다. 도 14e를 참조하면, 상보적 서열(1456)(또는 이의 역상보체(1456r))은 바코드(1440) 상의 범용 서열(1452)(예를 들어, 범용 P5 서열 또는 부분 범용 P5 서열) 및 유형 1 항체-올리고뉴클레오타이드 조성물(1404a)의 범용 서열분석 영역(1416)(예를 들어, 범용 N1 서열)을 통해 증폭될 수 있다. 위치자 r은 역상보체를 나타낸다. 제1 라운드의 PCR(본 명세서에서 PCR1로 지칭됨)은 범용 프라이머(예를 들어, P5 및 N1 서열, 또는 이들의 역상보체를 갖는 프라이머)를 사용하여, 하나 이상의 표지 및 범용 서열을 갖는 결합된 올리고뉴클레오타이드(1456r)에 상응하는 전장(또는 전장에 가까운) 앰플리콘을 생성할 수 있다.

[0708] 동일한 프라이머가 제2 라운드의 PCR 증폭(본 명세서에서 PCR2로 지칭됨)에 사용될 수 있거나, N1 프라이머는 도 14f에 예시된 바와 같이 추가의 특이성을 위하여 가교 올리고뉴클레오타이드 혼성화 영역 둘 모두(1428a, 1428b)에 결합하는 N2 프라이머에 의해 대체될 수 있다. 이러한 N2 프라이머가 사용되는 경우, 범용 서열분석 영역(1416) 및 범용 서열(1452) 둘 모두를 갖는 대부분의 결합된 올리고뉴클레오타이드(1456r)는 증폭될 수 있다. 일부 구현예에서, 이러한 N2 프라이머가 제2 라운드의 PCR에 사용되는 경우, 범용 서열분석 영역(1416) 및 범용 서열(1452) 둘 모두를 갖는 대부분의 결합된 올리고뉴클레오타이드(1456r)는 증폭될 수 있다.

[0709] 세포 성분 사이의 상호작용을 결정하기 위한 상호작용 결정 조성물의 쌍의 사용

[0710] 본 명세서의 개시 내용은 단백질-단백질 상호작용을 결정하기 위한 시스템, 방법, 및 키트를 포함한다. 일부 구현예에서, 방법은, 세포를 도 14a에 예시된 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)과 접촉시키는 단계를 포함한다. 세포는 제1 단백질 표적(1436a) 및 제2 단백질 표적(1436b)을 포함할 수 있다. 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 각각은 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합된 단백질 결합 시약(1408a, 1408b)을 포함한다. 제1 쌍의 상호작용 결정 조성물 중 하나의 단백질 결합 시약(1408a, 1408b)은 제1 단백질 표적(1436a)에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 단백질 결합 시약(1408b)은 제2 단백질 표적(1436b)에 특이적으로 결합할 수 있다. 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 상호작용 결정 서열(1424a, 1434b) 및 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 포함한다. 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 중의 상호작용 결정 서열(1424a, 1424b)은 상이한 서열을 포함할 수 있다.

- [0711] 도 14b를 참조하면, 방법은 가교 올리고뉴클레오타이드(1432)를 사용하여 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드(1412)를 생성하는 단계를 포함할 수 있다. 가교 올리고뉴클레오타이드(1432)는 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 포함할 수 있다.
- [0712] 도 14c를 참조하면, 방법은 복수의 바코드(1440)를 사용하여 결합된 상호작용 결정 올리고뉴클레오타이드(1412)를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계를 포함할 수 있다. 복수의 바코드(1440) 각각은 바코드 서열(1448) 및 포획 서열(1444)을 포함할 수 있다. 방법은 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 방법은 획득된 서열분석 데이터에서 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 상호작용 결정 서열(1424a, 1424b)의 회합에 기초하여 제1 단백질 표적과 제2 단백질 표적(1436a, 1436b) 사이의 상호작용을 결정하는 단계를 포함할 수 있다.
- [0713] 본 명세서의 개시 내용은 세포 성분 표적 사이의 상호작용을 결정하기 위한 시스템, 방법, 및 키트를 포함한다. 일부 구현예에서, 방법은, 세포를 도 14a에 예시된 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)과 접촉시키는 단계를 포함한다. 세포는 제1 세포 성분 표적 및 제2 세포 성분 표적(예를 들어, 제1 단백질 표적(1436a) 및 제2 단백질 표적(1436b))을 포함할 수 있다. 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합된 세포 성분 결합 시약(예를 들어, 항체(1408a, 1408b))을 포함할 수 있다. 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 중 하나의 세포 성분 결합 시약은 제1 세포 성분 표적(예를 들어, 단백질 표적(1436a))에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나(1404b)의 세포 성분 결합 시약은 제2 세포 성분 표적(예를 들어, 단백질 표적(1436b))에 특이적으로 결합할 수 있다. 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 상호작용 결정 서열(1424a, 1424b) 및 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)을 포함할 수 있다. 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 중의 상호작용 결정 서열(1424a, 1424b)은 상이한 서열을 포함할 수 있다.
- [0714] 도 14b를 참조하면, 방법은 가교 올리고뉴클레오타이드(1432)를 사용하여 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드(1412)를 생성하는 단계를 포함할 수 있다. 가교 올리고뉴클레오타이드(1432)는 제1 쌍의 상호작용 결정 조성물들(1404a, 1404b)의 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 포함할 수 있다.
- [0715] 도 14c를 참조하면, 방법은 복수의 바코드(1440)를 사용하여 결합된 상호작용 결정 올리고뉴클레오타이드(1412)를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계를 포함할 수 있다. 복수의 바코드(1440) 각각은 바코드 서열(1448) 및 포획 서열(1444)을 포함할 수 있다. 방법은 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 방법은 획득된 서열분석 데이터에서 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 상호작용 결정 서열(1424a, 1424b)의 회합에 기초하여 제1 세포 성분 표적과 제2 세포 성분 표적(1436a, 1436b) 사이의 상호작용을 결정하는 단계를 포함할 수 있다. 일부 구현예에서, 2개의 세포 성분 결합 시약 중 적어도 하나는 단백질 결합 시약을 포함할 수 있다. 단백질 결합 시약은 2개의 상호작용 결정 올리고뉴클레오타이드 중 하나와 회합될 수 있다. 하나 이상의 세포 성분 표적은 적어도 하나의 단백질 표적을 포함할 수 있다.
- [0716] 일부 구현예에서, 세포를 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)과 접촉시키는 단계는, 세포를 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 각각과 순차적으로 또는 동시에 접촉시키는 단계를 포함한다. 제1 단백질 표적(1436a)은 제2 단백질 표적(1436b)과 동일할 수 있다. 제1 단백질 표적(1436a)은 제2 단백질 표적(1436b)과 상이할 수 있다. 제1 세포 성분 표적과 제2 세포 성분 표적은 동일하거나 상이할 수 있다.
- [0717] 상호작용 결정 올리고뉴클레오타이드(1412a, 1412)는 상이한 실시에서 상이한 길이를 가질 수 있다. 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990,

1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이거나, 약 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위의 뉴클레오타이드 길이이다. 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 적어도, 또는 최대 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 128, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 또는 1000개의 뉴클레오타이드 길이이다.

[0718] 일부 구현예에서, 방법은, 세포를 제2 쌍의 상호작용 결정 조성물과 접촉시키는 단계를 포함한다. 세포는 제3 세포 성분 표적 및 제4 세포 성분 표적(예를 들어, 제3 단백질 표적 및 제4 단백질 표적)을 포함할 수 있다. 제2 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함할 수 있다. 제2 쌍의 상호작용 결정 조성물 중 하나의 세포 성분 결합 시약은 제3 세포 성분 표적에 특이적으로 결합할 수 있고, 제2 쌍의 상호작용 결정 조성물 중 다른 하나의 세포 성분 결합 시약은 제4 세포 성분 표적에 특이적으로 결합할 수 있다. 일부 구현예에서, 제3 세포 성분 표적 및 제4 세포 성분 표적 중 적어도 하나는 제1 세포 성분 표적 및 제2 세포 성분 표적 중 하나와 상이할 수 있다. 일부 구현예에서, 제3 세포 성분 표적 및 제4 세포 성분 표적 중 적어도 하나와 제1 세포 성분 표적 및 제2 세포 성분 표적 중 적어도 하나는 동일할 수 있다.

[0719] 일부 구현예에서, 방법은, 세포를 3개 이상의 쌍의 상호작용 결정 조성물과 접촉시키는 단계를 포함한다. 복수의 쌍의 상호작용 결정 조성물 중 소정 수의 상호작용 결정 조성물의 상호작용 결정 서열은 동일하거나 상이한 서열을 포함할 수 있다. 일부 구현예에서, 동일하거나 상이한 서열을 갖는 상호작용 결정 서열을 갖는 상호작용 결정 조성물의 수는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있거나, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000개, 또는 이들 중 임의의 두 값 사이의 어떤 수 또는 범위일 수 있다. 일부 구현예에서, 동일하거나 상이한 서열을 갖는 상호작용 결정 서열을 갖는 상호작용 결정 조성물의 수는 적어도, 또는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 또는 10000개일 수 있다.

[0720] 일부 구현예에서, 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b)은 상이한 서열을 포함한다. 가교 올리고뉴클레오타이드 혼성화 영역(1428a, 1428b) 중 적어도 하나는 가교 올리고뉴클레오타이드(1432)의 2개의 혼성화 영역 중 적어도 하나에 상보적일 수 있다.

[0721] 일부 구현예에서, 가교 올리고뉴클레오타이드(1432)를 사용하여 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)의 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)를 결합하는 단계는, 가교 올리고뉴클레오타이드(1432)의 제1 혼성화 영역을 상호작용 결정 올리고뉴클레오타이드의 가교 올리고뉴클레오타이드 혼성화 영역 중 제1 가교 올리고뉴클레오타이드 혼성화 영역(1428a)과 혼성화하는 단계; 가교 올리고뉴클레오타이드의 제2 혼성화 영역을 상호작용 결정 올리고뉴클레오타이드의 가교 올리고뉴클레오타이드 혼성화 영역 중 제2 가교 올리고뉴클레오타이드 혼성화 영역(1428b)과 혼성화하는 단계; 및 가교 올리고뉴클레오타이드(1432)에 혼성화된 상호작용 결정 올리고뉴클레오타이드(1428a, 1428b)를 결합하여 결합된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함한다.

[0722] 일부 구현예에서, 하나 이상의 세포 성분 표적(예를 들어, 단백질 표적(1436a, 1436b)) 중 적어도 하나는 세포 표면 상에 있다. 일부 구현예에서, 방법은, 세포를 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)과 접촉시키기 전에, 세포를 고정시키는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 제1 쌍의 상호작용 결정 조성물들(1404a, 1404b) 중 결합되지 않은 상호작용 결정 조성물을 제거하는 단계를 포함할 수 있다. 결합되지 않은 상호작용 결정 조성물을 제거하는 단계는 세포를 세척 완충액으로 세척하는 단계를 포함할 수 있다. 결합되지 않

은 상호작용 결정 조성물을 제거하는 단계는 유세포분석을 사용하여 세포를 선택하는 단계를 포함할 수 있다. 일부 구현예에서, 방법은, 세포를 용해시키는 단계를 포함할 수 있다.

[0723] 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드는 단백질 결합 시약으로부터 탈착 가능 또는 탈착 불가능하도록 구성된다. 방법은, 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b) 또는 결합된 상호작용 결정 올리고뉴클레오타이드(1412)를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)를 탈착시키는 단계는 UV 광절단, 화학적 처리, 가열, 효소 처리, 또는 이들의 임의의 조합에 의해 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)를 세포 성분 결합 시약으로부터 탈착시키는 단계를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 세포의 유전체 서열과 상동성이 아닐 수 있다. 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 종의 유전체 서열과 상동성일 수 있다.

[0724] 일부 구현예에서, 제1 쌍의 상호작용 결정 조성물들(1404a, 1404b) 중 하나의 상호작용 결정 올리고뉴클레오타이드(1412b)는 포획 서열(1444)에 상보적인 서열(1420)을 포함한다. 포획 서열(1444)은 폴리(dT) 영역을 포함할 수 있다. 포획 서열(1444)에 상보적인 상호작용 결정 올리고뉴클레오타이드의 서열(1420)은 폴리(dA) 영역을 포함할 수 있다. 일부 구현예에서, 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)는 제2 바코드 서열을 포함한다. 제1 쌍의 상호작용 결정 조성물(1404a, 1404b) 중 다른 하나의 상호작용 결정 올리고뉴클레오타이드(1412a)는 범용 프라이머를 위한 결합 부위(1416)를 포함할 수 있다. 상호작용 결정 올리고뉴클레오타이드(1404a, 1404b)는 검출가능한 모이어티와 회합될 수 있다.

[0725] 일부 구현예에서, 세포 성분 결합 시약은 동일하거나 상이한 상호작용 결정 서열(1428a, 1428b)을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합될 수 있다. 일부 구현예에서, 복수의 상호작용 결정 조성물들(1404a, 1404b) 중 하나는 상호작용 결정 올리고뉴클레오타이드와 회합되지 않은 제2 단백질 결합 시약을 포함한다. 세포 성분 결합 시약과 제2 세포 성분 결합 시약은 동일할 수 있다. 세포 성분 결합 시약은 검출가능한 모이어티와 회합될 수 있다. 세포 성분 결합 시약(1404a, 1404b)은 동일한 서열을 갖는 2개 이상의 상호작용 결정 올리고뉴클레오타이드(1412a, 1412b)와 회합될 수 있다.

[0726] 일부 구현예에서, 방법은, 2개 이상의 세포를 제1 쌍의 상호작용 결정 조성물(1404a, 1404b)과 접촉시키는 단계를 포함한다. 2개 이상의 세포 각각은 제1 세포 성분 표적 및 제2 세포 성분 표적(예를 들어, 제1 단백질 표적 및 제2 단백질 표적(1436a, 1436b))을 포함할 수 있다. 2개 이상의 세포 중 적어도 하나는 단일 세포를 포함한다.

[0727] 도 14b를 참조하면, 바코드(1440)는 세포 표지 서열(1448), 범용 프라이머를 위한 결합 부위, 또는 이들의 임의의 조합을 포함한다. 복수의 바코드 중 적어도 2개의 바코드는 동일한 세포 표지 서열(1448)을 포함한다.

[0728] 일부 구현예에서, 세포 성분 표적은 세포의 단백질, 세포내 단백질, 또는 이들의 임의의 조합이거나, 이들을 포함한다. 세포 성분 표적은 지질, 탄수화물, 또는 이들의 임의의 조합일 수 있거나, 이들을 포함할 수 있다. 세포 성분 표적은 10 내지 100개의 상이한 세포 성분 표적을 포함하는 군으로부터 선택될 수 있다.

[0729] 일부 구현예에서, 복수의 바코드는 입자(예를 들어, 비드(1438))와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 부분적으로 고정화될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 봉입될 수 있다. 복수의 바코드 중 적어도 하나의 바코드는 입자 내에 부분적으로 봉입될 수 있다. 입자는 붕괴성일 수 있다. 일부 구현예에서, 입자의 바코드는 적어도 1000, 10000개, 또는 그 이상의 상이한 바코드 서열로부터 선택되는 바코드 서열을 포함한다.

[0730] 도 14c를 참조하면, 복수의 바코드(1440)를 사용하여 상호작용 결정 올리고뉴클레오타이드(1412)를 바코딩하는 단계는, 복수의 바코드(1440)를 상호작용 결정 올리고뉴클레오타이드(1412)와 접촉시켜 상호작용 결정 올리고뉴클레오타이드에 혼성화된 바코드를 생성하는 단계; 및 상호작용 결정 올리고뉴클레오타이드에 혼성화된 바코드를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계를 포함한다. 바코드(1440)를 연장하는 단계는 DNA 중합효소를 사용하여 바코드(1440)를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계를 포함할 수 있다. 바코드(1440)를 연장하는 단계는 역전사효소를 사용하여 바코드(1440)를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(14)를 생성하는 단계를 포함할 수 있다. 바코드(14)를 연장하는 단계는 물로니 뮤린 백혈병 바이러스(M-MLV) 역전사효소 또는 Taq DNA 중합효소를 사용하여 바코드(14)를 연장하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계를 포함할 수 있다. 바코드(14)를 연장하는 단계는 결합된 상호작용 결정 올리고뉴클레오타이드(1412)부터 가교 올리고뉴클레오타이드(1432)를 변위시키는 것을 포함할 수 있다.

- [0731] 도 14d 및 도 14e를 참조하면, 방법은, 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 증폭시켜 복수의 앰플리콘을 생성하는 단계를 포함할 수 있다. 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 증폭시키는 단계는 중합효소 연쇄 반응(PCR)을 사용하여, 바코드 서열(1448)의 적어도 일부분 및 상호작용 결정 올리고뉴클레오타이드(1428a, 1428b)의 적어도 일부분을 증폭시키는 단계를 포함할 수 있다. 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)의 서열분석 데이터를 획득하는 단계는 복수의 앰플리콘의 서열분석 데이터를 획득하는 단계를 포함할 수 있다. 서열분석 데이터를 획득하는 단계는 바코드 서열(1448)의 적어도 일부분 및 상호작용 결정 올리고뉴클레오타이드(1428a, 1428b)의 적어도 일부분을 서열분석하는 단계를 포함할 수 있다. 일부 구현예에서, 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)의 서열분석 데이터를 획득하는 단계는 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)의 부분적인 및/또는 완전한 서열(또는 이들의 역서열, 상보체, 역상보체(1456r), 또는 이들의 임의의 조합)을 획득하는 단계를 포함한다.
- [0732] 도 14c를 참조하면, 복수의 바코드(1440)은 복수의 추계적 바코드를 포함할 수 있다. 복수의 추계적 바코드 각각의 바코드 서열은 분자 표지 서열(1448)을 포함할 수 있다. 복수의 추계적 바코드 중 적어도 2개의 추계적 바코드의 분자 표지 서열(1448)은 상이한 서열을 포함할 수 있다. 복수의 바코드(1440)를 사용하여 상호작용 결정 올리고뉴클레오타이드(1412)를 바코딩하여 복수의 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 생성하는 단계는 복수의 추계적 바코드를 사용하여 상호작용 결정 올리고뉴클레오타이드(1412)를 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 상호작용 결정 올리고뉴클레오타이드를 생성하는 단계를 포함할 수 있다.
- [0733] 방법은 또한 기타 다른 단일 세포 RNA 서열분석 방법과 함께 사용될 수 있다. 예를 들어, mRNA 발현 수준, 단백질 발현 수준(또는 세포 성분 표지의 발현 수준), 및/또는 단백질-단백질 상호작용(또는 세포 성분 사이의 상호작용)을 결정하기 위하여 하나의 단일 작업 흐름이 사용될 수 있다. 결과적으로, 단일 실험에서, 단일 세포(또는 다수의 단일 세포) 내로부터의 mRNA 발현, 단백질 발현, 및 단백질-단백질 상호작용에 관한 데이터가 획득될 수 있다. 일부 구현예에서, 방법은, 복수의 바코드(1440)를 사용하여 세포의 복수의 표적(예를 들어, 관심 대상의 mRNA 종)을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계; 및 바코딩된 표적의 서열분석 데이터를 획득하는 단계를 포함한다. 복수의 바코드(1440)를 사용하여 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는, 표적의 복제물을 바코드(1440)의 표적-결합 영역(예를 들어, 폴리(dT) 영역(1444))과 접촉시키는 단계; 및 복수의 바코드(1440)를 사용하여 복수의 표적을 역전사하여 복수의 역전사된 표적을 생성하는 단계를 포함할 수 있다. 방법은, 복수의 바코딩된 표적의 서열분석 데이터를 획득하기 전에, (도 14e 및 도 14f에 예시된 바코딩된 상호작용 결정 올리고뉴클레오타이드(1456)를 증폭시키는 단계와 유사한) 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계를 포함할 수 있다. 바코딩된 표적을 증폭시켜 복수의 증폭된 바코딩된 표적을 생성하는 단계는, 중합효소 연쇄 반응(PCR)에 의해 바코딩된 표적을 증폭시키는 단계를 포함할 수 있다. 복수의 바코드(1440)를 사용하여 세포의 복수의 표적을 바코딩하여 복수의 바코딩된 표적을 생성하는 단계는 복수의 추계적 바코드를 사용하여 세포의 복수의 표적을 추계적으로 바코딩하여 복수의 추계적으로 바코딩된 표적을 생성하는 단계를 포함할 수 있다.
- [0734] 본 명세서에서 개시된 구현예는 또한 세포 성분-세포 성분 상호작용(예를 들어, 단백질-단백질 상호작용)을 식별하기 위한 키트를 포함한다. 일부 구현예에서, 키트는, 제1 쌍의 상호작용 결정 조성물을 포함하며, 여기서 제1 쌍의 상호작용 결정 조성물 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 제1 쌍의 상호작용 결정 조성물 중 하나의 세포 성분 결합 시약은 제1 세포 성분 표적에 특이적으로 결합할 수 있고, 제1 쌍의 상호작용 결정 조성물 중 다른 하나의 세포 성분 결합 시약은 제2 세포 성분 표적에 특이적으로 결합할 수 있고, 상호작용 결정 올리고뉴클레오타이드는 상호작용 결정 서열 및 가교 올리고뉴클레오타이드 혼성화 영역을 포함하고, 제1 쌍의 상호작용 결정 조성물의 상호작용 결정 서열은 상이한 서열을 포함하고; 제1 쌍의 상호작용 결정 조성물의 가교 올리고뉴클레오타이드 혼성화 영역에 특이적으로 결합할 수 있는 2개의 혼성화 영역을 각각 포함하는 복수의 가교 올리고뉴클레오타이드를 포함한다.
- [0735] 일부 구현예에서, 키트는, 제2 쌍의 상호작용 결정 조성물을 포함하며, 여기서 제2 쌍의 상호작용 결정 조성물의 각각은 상호작용 결정 올리고뉴클레오타이드와 회합된 세포 성분 결합 시약을 포함하고, 제2 쌍의 상호작용 결정 조성물 중 하나의 세포 성분 결합 시약은 제3 세포 성분 표적에 특이적으로 결합할 수 있고, 제2 쌍의 상호작용 결정 조성물 중 다른 하나의 세포 성분 결합 시약은 제4 세포 성분 표적에 특이적으로 결합할 수 있다. 일부 구현예에서, 키트는, 상호작용 결정 조성물의 3개 이상의 쌍을 포함한다.
- [0736] 일부 구현예에서, 키트는, 복수의 바코드를 포함하고, 복수의 바코드 각각은 바코드 서열 및 포획 서열을 포함한다. 복수의 바코드는 복수의 추계적 바코드를 포함할 수 있으며, 여기서 복수의 추계적 바코드 각각의 바코드 서열은 바코드 서열(예를 들어, 분자 표지 서열)을 포함하고, 복수의 추계적 바코드 중 적어도 2개의 추계적 바

코드의 바코드 서열은 상이한 서열을 포함한다. 일부 구현예에서, 복수의 바코드는 입자와 회합된다. 복수의 바코드 중 적어도 하나의 바코드는 입자 상에 고정화되거나, 입자 상에 부분적으로 고정화되거나, 입자 내에 봉입되거나, 입자 내에 부분적으로 봉입되거나, 이들의 임의의 조합일 수 있다. 입자는 붕괴성일 수 있다. 입자는 비드를 포함할 수 있다.

[0737] 일부 구현예에서, 키트는, DNA 중합효소를 포함한다. 키트는 역전사효소를 포함할 수 있다. 키트는, 몰로니 뮤린 백혈병 바이러스(M-MLV) 역전사효소 또는 Taq DNA 중합효소를 포함할 수 있다. 일부 구현예에서, 방법은 고정제(예를 들어, 포르말린, 파라포름알데히드, 글루타르알데히드/오스뮴 테트록시드, 알코올성 고정제, 헤페스-글루탐산 완충액-매개 유기 용매 보호 효과(HOPE), 부인 용액, 또는 이들의 임의의 조합)를 포함한다.

[0738] 실시예

[0739] 상기 논의된 구현예의 일부 양태들은 하기 실시예에서 더 상세히 개시되며, 이는 본 발명의 개시 내용의 범주를 어떤 방식으로든 제한하고자 하지 않는다.

[0740] 실시예 1

[0741] 단백질 결합 시약과의 접합을 위한 올리고뉴클레오티드

[0742] 본 실시예는 단백질 결합 시약과 접합될 수 있는 올리고뉴클레오티드의 설계를 보여준다. 올리고뉴클레오티드는 단백질 발현 및 유전자 발현을 동시에 결정하는 데 사용될 수 있다. 이러한 올리고뉴클레오티드는 또한 동일하거나 상이한 샘플의 세포를 결정하기 위하여 샘플 인덱싱에 사용될 수 있다.

[0743] 95량체 올리고뉴클레오티드 설계

[0744] 하기 방법을 사용하여 단백질 발현과 유전자 발현의 동시 결정 또는 샘플 인덱싱을 위한 후보 올리고뉴클레오티드 서열 및 상응하는 프라이머 서열을 생성하였다.

[0745] 1. 서열 생성 및 제거

[0746] 하기 공정을 사용하여 단백질 발현 및 유전자 발현의 동시 결정 또는 샘플 인덱싱을 위한 후보 올리고뉴클레오티드를 생성하였다.

[0747] 단계 1a. 원하는 길이(45 bp)를 갖는 다수의 후보 서열(50000개의 서열)을 랜덤하게 생성시킨다.

[0748] 단계 1b. 전사 조절제 LSRR 서열을 생성된 서열의 5' 말단에 첨부시키고, 폴리(A) 서열을(25 bp) 생성된 서열의 3' 말단에 첨부시킨다.

[0749] 단계 1c. 40% 내지 50%의 범위에 있는 GC 함량을 갖지 않는 생성 및 첨부된 서열을 제거한다.

[0750] 단계 1d. 각각 하나 이상의 헤어핀 구조를 갖는 잔류 서열을 제거한다.

[0751] 잔류하는 후보 올리고뉴클레오티드 서열의 수는 423개이었다.

[0752] 2. 프라이머 설계

[0753] 하기 방법을 사용하여 잔류하는 423개의 후보 올리고뉴클레오티드 서열에 대한 프라이머를 설계하였다.

[0754] 2.1 N1 프라이머: N1 프라이머로서 범용 N1 서열인 5'-GTTGTCAAGATGCTACCGTTCAGAG-3'(LSRR 서열: 서열번호 5)를 사용한다.

[0755] 2.2 N2 프라이머(특이적 샘플 인덱스 올리고뉴클레오티드를 증폭시키기 위함; 예를 들어, 도 15b 내지 도 15d에 서의 N2 프라이머):

[0756] 2.2a. N1 서열의 하류에서 출발하지 않는 후보 N2 프라이머를 제거한다.

[0757] 2.2b. 후보 올리고뉴클레오티드 서열의 최종 35 bp에서 중복되는 후보 N2 프라이머를 제거한다.

[0758] 2.2c. 올리고뉴클레오티드를 사용하여 연구되는 세포의 종의 전사체(예를 들어, 인간 전사체 또는 마우스 전사체)에 대해 정렬된 프라이머 후보를 제거한다.

[0759] 2.2d. 디폴트 제어(default control)로서 ILR2 서열(ACACGACGCTCTCCGATCT; 서열 번호 8)을 사용하여 프라이머-프라이머 상호작용을 최소화하거나 피한다.

- [0760] 423개의 후보 올리고뉴클레오타이드 서열 중, 390개의 후보에 대한 N2 프라이머를 설계하였다.
- [0761] 3. 필터링
- [0762] 하기 공정을 사용하여 잔류하는 390개의 후보 프라이머 서열을 필터링하였다.
- [0763] 3a. 복수의 A로 종결되는 랜덤 서열을 갖는 임의의 후보 올리고뉴클레오타이드 서열을 제거하여(즉, 폴리(A) 서열의 유효 길이는 25 bp를 초과함), 폴리(A) 테일을 모든 바코드에 대해 동일한 길이로 유지한다.
- [0764] 3b. 복수의 G의 연속된 올리고 합성에서의 추가 비용 및 잠재적으로 더 낮은 수율로 인해, 4개 이상의 연속 G(3개 초과 G)를 갖는 임의의 후보 올리고뉴클레오타이드 서열을 제거한다.
- [0765] 도 15a는 상기 방법을 사용하여 생성된 비제한적이고 예시적인 후보 올리고뉴클레오타이드 서열을 나타낸다.
- [0766] 200량체 올리고뉴클레오타이드 설계
- [0767] 하기 방법을 사용하여 단백질 발현과 유전자 발현의 동시 결정 및 샘플 인덱싱을 위한 후보 올리고뉴클레오타이드 서열 및 상응하는 프라이머 서열을 생성하였다.
- [0768] 1. 서열 생성 및 제거
- [0769] 하기를 사용하여 단백질 발현과 유전자 발현의 동시 결정 및 샘플 인덱싱을 위한 후보 올리고뉴클레오타이드 서열을 생성하였다.
- [0770] 1a. 원하는 길이(128 bp)를 갖는 다수의 후보 서열(100000개의 서열)을 랜덤하게 생성한다.
- [0771] 1b. 전사 조절제 LSRR 서열 및 비-인간, 비-마우스인 추가의 고착 서열(anchor sequence)을 생성된 서열의 5' 말단에 첨부시키고, 폴리(A) 서열(25 bp)을 생성된 서열의 3' 말단에 첨부시킨다.
- [0772] 1c. GC 함량이 40% 내지 50%의 범위에 있는 않는 생성 및 첨부된 서열을 제거한다.
- [0773] 1d. 헤어핀 구조 점수에 기초하여 잔류하는 후보 올리고뉴클레오타이드 서열을 분류한다.
- [0774] 1e. 최저 헤어핀 점수를 갖는 1000개의 잔류하는 후보 올리고뉴클레오타이드 서열을 선택한다.
- [0775] 2. 프라이머 설계
- [0776] 하기 방법을 사용하여 최저 헤어핀 점수를 갖는 400개의 후보 올리고뉴클레오타이드 서열에 대한 프라이머를 설계하였다.
- [0777] 2.1 N1 프라이머: N1 프라이머로서 범용 N1 서열인 5'-GTTGTCAAGATGCTACCGTTCAGAG-3'(LSRR 서열; 서열 번호 5)를 사용한다.
- [0778] 2.2 N2 프라이머(특이적 샘플 인덱스 올리고뉴클레오타이드를 증폭시키기 위함; 예를 들어, 도 15b 및 도 15c에서의 N2 프라이머):
- [0779] 2.2a. N1 서열의 23 nt 하류에서 출발하지 않는 후보 N2 프라이머를 제거한다(고착 서열은 모든 후보 올리고뉴클레오타이드 서열에 걸쳐 범용이었다).
- [0780] 2.2b. 표적 서열의 마지막 100 bp에서 중복되는 후보 N2 프라이머를 제거한다. 생성된 프라이머 후보는 표적 서열의 48번째 뉴클레오타이드와 100번째 뉴클레오타이드 사이일 수 있다.
- [0781] 2.2c. 올리고뉴클레오타이드를 사용하여 연구되는 세포의 종의 전사체(예를 들어, 인간 전사체 또는 마우스 전사체)에 대해 정렬된 프라이머 후보를 제거한다.
- [0782] 2.2d. 디플트 제어로서 ILR2 서열(5'-ACACGACGCTCTCCGATCT-3'; 서열 번호 8)을 사용하여 프라이머-프라이머 상호작용을 최소화하거나 피한다.
- [0783] 2.2e. 표적 서열의 마지막 100 bp에서 중복되는 N2 프라이머 후보를 제거한다.
- [0784] 400개의 후보 올리고뉴클레오타이드 서열 중, 392개의 후보에 대한 N2 프라이머를 설계하였다.
- [0785] 3. 필터링
- [0786] 하기를 사용하여 잔류하는 392개의 후보 프라이머 서열을 필터링하였다.

- [0787] 3a. 복수의 A로 종결되는 랜덤 서열을 갖는 임의의 후보 올리고뉴클레오타이드 서열을 제거하여(즉, 폴리(A) 서열의 유효 길이는 25 bp를 초과함), 폴리(A) 테일을 모든 바코드에 대해 동일한 길이로 유지한다.
- [0788] 3b. 복수의 G의 연속된 올리고 합성에서의 추가 비용 및 잠재적으로 더 낮은 수율로 인해, 4개 이상의 연속 G(3개 초과 G)를 갖는 임의의 후보 올리고뉴클레오타이드 서열을 제거한다.
- [0789] 도 15b는 상기 방법을 사용하여 생성된 비제한적이고 예시적인 후보 올리고뉴클레오타이드 서열을 나타낸다. 도 15b에 나타난 네스티드 N2 프라이머는 표적화된 증폭을 위해 항체 또는 샘플 특이적 서열에 결합할 수 있다. 도 15c는 표적화된 증폭을 위한 고정 서열에 상응하는 네스티드 범용 N2 프라이머를 이용하여 동일한 비제한적이고 예시적인 후보 올리고뉴클레오타이드 서열을 나타낸다. 도 15d는 1 단계로 표적화된 증폭을 위한 N2 프라이머를 이용하여 동일한 비제한적이고 예시적인 후보 올리고뉴클레오타이드 서열을 나타낸다.
- [0790] 종합하면, 이들 데이터는 상이한 길이의 올리고뉴클레오타이드 서열이 단백질 발현 및 유전자 발현 또는 샘플 인덱싱의 동시 결정을 위해 설계될 수 있음을 나타낸다. 올리고뉴클레오타이드 서열은 범용 프라이머 서열, 항체 특이적 올리고뉴클레오타이드 서열 또는 샘플 인덱싱 서열, 및 폴리(A) 서열을 포함할 수 있다.
- [0791] 실시예 2
- [0792] 상이한 항체:올리고뉴클레오타이드 비를 이용한 검출 감도의 비교
- [0793] 본 실시예는 1, 2, 또는 3개 올리고뉴클레오타이드와 접합된 항-CD4 항체를 사용하여 CD4 단백질의 검출 감도를 보여준다.
- [0794] 대상체의 냉동 말초혈액 단핵 세포(PBMC)를 해동시켰다. 해동된 PBMC를 실온에서 20분 동안 0.06 $\mu\text{g}/100 \mu\text{l}$ (올리고뉴클레오타이드-접합된 항체 스톱의 1:333 희석)에서 항-CD4 항체의 3가지 유형으로 염색하였다. 항-CD4 항체 유형의 각각의 유형을 1, 2, 또는 3개의 올리고뉴클레오타이드("항체 올리고뉴클레오타이드")와 접합하였다. 항체 올리고뉴클레오타이드의 서열은 도 16에 나타난다. 세포를 세척하여 결합되지 않은 항-CD4 항체를 제거하였다. 살아있는 세포를 수득하기 위하여 유세포분석을 이용하여 분류하기 위해 Calcein AM(BD(미국 뉴저지주, 프랭클린레이크 소재)) 및 Draq7TM(Abcam(영국, 캄브리지 소재))로 세포를 염색하였다. 세포를 세척하여 과량의 Calcein AM 및 Draq7TM를 제거하였다. Calcein AM으로 염색되고(살아있는 세포), Draq7TM로는 염색되지 않은(죽지 않거나 투과화되지 않은 세포) 단일 세포를 유세포분석을 사용하여, BD RhapsodyTM 카트리지 내로 분류하였다.
- [0795] 단일 세포 및 비드를 함유하는 웰 중, 웰 내의 3500개의 단일 세포를 5 mM DTT를 갖는 용해 완충액 내에서 용해시켰다. BD RhapsodyTM 비드를 사용하여 CD4 mRNA 발현 프로파일을 결정하였다. CD4 단백질 발현 프로파일을 BD RhapsodyTM 비드 및 항체 올리고뉴클레오타이드를 사용하여 결정하였다. 간략하게는, 세포 용해 후에 mRNA 분자를 방출하였다. RhapsodyTM 비드를 분자 표지, 세포 표지, 및 폴리T 영역을 각각 함유하는 추계적 바코드와 회합시켰다. 용해된 세포로부터 방출된 mRNA 분자의 폴리(A) 영역을 추계적 바코드의 폴리T 영역에 혼성화하였다. 올리고뉴클레오타이드의 폴리(A) 영역을 추계적 바코드의 폴리T 영역에 혼성화하였다. mRNA 분자를 추계적 바코드를 사용하여 역전사하였다. 항체 올리고뉴클레오타이드를 추계적 바코드를 사용하여 복제하였다. 역전사 및 복제는 하나의 샘플 분취액 내에서 동시에 발생하였다.
- [0796] 혈액 패널 N1 프라이머를 사용하여 488개 혈액 패널 유전자의 mRNA 발현 프로파일, 및 항체 올리고뉴클레오타이드 N1 프라이머("PCR 1")를 사용하여 CD4 단백질의 발현 프로파일을 결정하기 위한 프라이머를 사용하여, 60도의 어닐링 온도에서 15회의 사이클 동안 역전사된 산물 및 복제된 산물을 PCR 증폭시켰다. 과량의 추계적 바코드를 암퇘어 클린업을 이용하여 제거하였다. PCR1로부터의 산물을 2개의 분취액으로 나누었으며, 하나는 혈액 패널 N2 프라이머를 사용하여, 488개의 혈액 패널 유전자의 mRNA 발현 프로파일을 결정하기 위한 분취액이고, 하나는 항체 올리고뉴클레오타이드 N2 프라이머("PCR 2")를 사용하여, CD4 단백질의 발현 프로파일을 결정하기 위한 분취액이다. 두 분취액 모두를 60도의 어닐링 온도에서 15회의 사이클 동안 PCR 증폭시켰다. 용해된 세포 내 CD4 단백질의 발현을 도 16에 예시된 바와 같은 항체 올리고뉴클레오타이드에 기초하여 결정하였다("PCR 2"). 서열분석 어댑터 결합("PCR 3") 후, 서열분석 데이터를 수득 및 분석하였다. 세포 유형을 488개의 혈액 패널 유전자의 발현 프로파일에 기초하여 결정하였다.
- [0797] 도 17a 내지 도 17f는 CD4 단백질 발현 및 유전자 발현을 동시에 고처리량 방식으로 측정하기 위하여 올리고뉴클레오타이드-접합된 항체 사용의 결과를 나타내는, 비제한적이고 예시적인 t-분포화 추계적 네이버 임베딩

(Distributed Stochastic Neighbor Embedding: tSNE) 투영 플롯이다. 1, 2 또는 3개의 항체 올리고뉴클레오타이드에 접합된 항-CD4 항체를 이용하여 CD4 발현 세포 유형(예를 들어, CD4 T 세포) 내에서 명백하고 확실하게 CD4 단백질 발현을 검출하였다(각각 도 17b, 도 17d, 및 도 17f). 도 18a 내지 도 18f는 CD4 mRNA 및 단백질의 CD4 T 세포에서(높은 CD4 발현), CD8 T 세포(최소 CD4 발현), 및 골수 세포(일부 CD4 발현)에서의 발현을 나타내는 비제한적이고 예시적인 막대 차트이다. 유사한 서열분석 깊이를 이용하여, CD4 단백질 수준에 대한 검출 감도는 더 높은 항체:올리고뉴클레오타이드 비에서 증가되었으며, 1:3 비가 1:1 및 1:2 비보다 더 양호하게 수행되었다(도 19). 유세포분석을 사용하여 분류된 세포의 세포 표면 상에서의 CD4 단백질의 발현을, 도 20a 내지 도 20d에 나타낸 바와 같이 FlowJo(FlowJo(미국 오리건주 애쉬랜드 소재))를 사용하여 확인하였다. 도 21a 내지 도 21f는 두 샘플의 CD4 T 세포, CD8 T 세포, 및 골수 세포에서 CD4 mRNA 및 CD4 단백질의 발현을 나타내는 비제한적이고 예시적인 막대 차트이다. 2개의 상이한 샘플 제조 프로토콜을 사용하여 제2 샘플을 제조하였다. 도 22는 1:3의 항체:올리고뉴클레오타이드 비를 갖는 상이한 샘플 제조 프로토콜을 사용하여 결정된 CD4 단백질 수준에 대한 검출 감도를 나타내는 비제한적이고 예시적인 막대 차트이다.

[0798] 전체적으로, 이들 데이터는 항-CD4 항체와 접합된 올리고뉴클레오타이드에 기초하여 CD4 단백질 발현이 명백하고 강력하게 검출될 수 있음을 나타낸다. CD4 단백질 수준에 대한 검출 감도는 더 높은 항체:올리고뉴클레오타이드 비와 함께 증가할 수 있다.

[0799] 실시예 3

[0800] 샘플 인택싱

[0801] 본 실시예는 높은 표지화 효율 및 낮은 비특이적 표지화 또는 스푼 오버(spill over)로 샘플 인택싱을 사용하여 상이한 샘플의 세포를 식별하는 것을 보여준다. 본 실시예는 또한 샘플 인택싱에 대한 샘플 인택싱 올리고뉴클레오타이드의 길이 및 샘플 인택싱 올리고뉴클레오타이드의 절단성의 효과를 보여준다.

[0802] 도 23은 샘플 인택싱을 수행하고, 샘플 인택싱에 대한 샘플 인택싱 올리고뉴클레오타이드의 길이 및 샘플 인택싱 올리고뉴클레오타이드의 절단성의 효과를 결정하기 위한 비제한적이고 예시적인 실험 설계를 나타낸다. 도 23에서의 좌측 컬럼은, 주르카트 세포를 상이한 샘플 인택싱 올리고뉴클레오타이드를 갖는 항-Cbcbcb체로 표지하여, 샘플 인택싱에 대한 샘플 인택싱 올리고뉴클레오타이드의 길이 및 샘플 인택싱 올리고뉴클레오타이드의 절단성의 효과를 결정하였음을 나타낸다. 항-CD147 항체와 접합된 올리고뉴클레오타이드는 항체로부터 절단 가능한 200개의 뉴클레오타이드 길이(T-링커), 절단 가능한 95개의 뉴클레오타이드 길이(T-링커), 및 절단 불가능한 95개의 뉴클레오타이드 길이였다. 추가의 샘플 제조 후에, RhapsodyTM 플로우셀(도 23에 "플로우셀 1"로 표지되어 있고, "FC1"로 약기됨)을 사용하여, 샘플 인택싱에 대한 샘플 인택싱 올리고뉴클레오타이드의 길이 및 샘플 인택싱 올리고뉴클레오타이드의 절단성의 효과를 결정하였다. 샘플 제조는 Calcein AM 및 Draq7TM 염색, 유세포분석을 사용한 세포 분류, 및 체계적 바코딩을 포함하였다.

[0803] 도 23에서의 가운데 컬럼은 절단 가능한 3개의 상이한 샘플 인택싱 올리고뉴클레오타이드(도 23에서 "단길이 1," "단길이 2," 및 "단길이 3"으로 표지되어 있는 샘플 인택싱 올리고뉴클레오타이드는 S1, S2, 및 S3으로 약기될 수 있음)와 접합된 항-CD147 항체로 세포를 표지하였음을 나타낸다. 라모스 세포를 절단 가능한 상이한 샘플 인택싱 올리고뉴클레오타이드(도 23에서 샘플 인택싱 올리고뉴클레오타이드는 "단길이 1" 및 "단길이 2"로 표지됨)와 접합된 2가지 유형의 항-CD147 항체로 표지하였다. 주르카트 세포를 절단 가능한 상이한 샘플 인택싱 올리고뉴클레오타이드(도 23에서 샘플 인택싱 올리고뉴클레오타이드는 "단길이 1" 및 "단길이 3"으로 표지됨)와 접합된 2가지 유형의 항-CD147 항체로 표지하였다. 표지된 세포를 1:1 비로 혼합한 후, 추가의 샘플 제조 및 RhapsodyTM 플로우셀(도 23에서 "플로우셀 2"로 표지되어 있고, "FC2"로 약기됨)을 사용한 샘플 인택싱의 효율 및 특이성의 결정을 수행하였다. 도 23에서의 우측 컬럼은 대조 실험을 나타내는데, 여기서는 표지되지 않은 주르카트 세포와 라모스 세포를 1:1 비로 혼합한 후, 추가의 샘플 제조 및 RhapsodyTM 플로우셀(도 23에서 "플로우셀 3"으로 표지되어 있고, "FC3"으로 약기됨)을 사용한 분석을 수행하였다.

[0804] 세포를 실온에서 20분 동안 1:3 항체:올리고뉴클레오타이드 비의 샘플 인택싱 올리고뉴클레오타이드와 접합된 항체로 염색하였다. 올리고뉴클레오타이드-접합된 항체 스톱으로부터 pH 7.5 희석제를 사용하여 100 μ l 중 1:20 희석률로 항체를 희석시켰다. 세포를 세척하여 결합되지 않은 항-CD147 항체를 제거하였다. 살아 있는 세포를 수득하기 위하여 유세포분석을 이용하여 분류하기 위해 Calcein AM(살아 있는 세포를 위한 염색) 및 Draq7TM(죽었거나 투과화된 세포를 위한 염색)로 세포를 염색하였다. 세포를 세척하여 과량의 Calcein AM 및 Draq7TM을 제거하

었다. Calcein AM으로 염색되고 Draq7™로는 염색되지 않은 단일 세포를, 유세포분석을 사용하여, 플로우셀을 함
유하는 BD Rhapsody™ 카트리지 내로 분류하였다.

[0805] 단일 세포 및 비드를 함유하는 웰 중, 웰 내의 1000개의 단일 세포를 용해 완충액 중에서 용해시켰다. BD Rhapsody™ 비드를 사용하여 CD147 mRNA 발현 프로파일을 결정하였다. 간략하게는, 세포 용해 후에 mRNA 분자를 방출하였다. Rhapsody™ 비드를 분자 표지, 세포 표지, 및 폴리T 영역을 각각 함유하는 추계적 바코드와 회합시켰다. 용해된 세포로부터 방출된 mRNA 분자의 폴리(A) 영역을 추계적 바코드의 폴리T 영역에 혼성화하였다. 절단 가능한 경우, 항체로부터 절단된 샘플 인덱싱 올리고뉴클레오타이드의 폴리(A) 영역은 추계적 바코드의 폴리T 영역에 혼성화하였다. mRNA 분자를 추계적 바코드를 사용하여 역전사하였다. 항체 올리고뉴클레오타이드를 추계적 바코드를 사용하여 복제하였다. 역전사 및 복제는 60도 어닐링 온도에서 15회의 사이클 동안 하나의 샘플 분취물 내에서 동시에 일어났다.

[0806] 혈액 패널 N1 프라이머를 사용하여 488개 혈액 패널 유전자의 mRNA 발현 프로파일, 및 샘플 인덱싱 올리고뉴클레오타이드 N1 프라이머("PCR 1")를 사용하여 CD147 단백질의 발현을 결정하기 위한 프라이머를 사용하여, 60도의 어닐링 온도에서 15회의 사이클 동안 역전사된 산물 및 복제된 산물을 PCR 증폭시켰다. 과량의 프라이머는 앰플리클론을 이용하여 제거하였다. PCR1로부터의 산물을 혈액 패널 N2 프라이머 및 어댑터 결찰을 위한 측접 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드 N1 프라이머를 사용하여 60도의 어닐링 온도에서 15회의 사이클 동안 추가로 PCR 증폭시켰다("PCR 2"). 서열분석 어댑터 결찰 후 서열분석 데이터를 획득하고 분석하였다("PCR 3"). 488개의 혈액 패널 유전자의 발현 프로파일에 기초하여 세포 유형을 결정하였다.

[0807] 표 1은 도 23에 예시된 실험 설계를 사용하여 획득된 서열분석 데이터의 메트릭스(metrics)를 나타낸다. 샘플 인덱싱 올리고뉴클레오타이드의 3가지 유형(도 23에서 좌측 컬럼에 나타난 절단 가능한 95량체, 절단 불가능한 95량체, 및 절단 가능한 200량체) 중 어느 하나와 접합된 3가지 유형의 항-CD147 항체는 샘플 인덱싱에 성공적으로 사용되었다. 서열분석 데이터에서 총 관독물 수의 11% 초과가 샘플 인덱싱 올리고뉴클레오타이드에 기인하였다.

표 1

[0808] 샘플 인덱싱 올리고뉴클레오타이드의 서열분석 메트릭스

샘플	FC1(올리고)			FC2(샘플 인덱싱)		
	주르카트(J)			J & R	라모스(R)	J
올리고 유형	절단 가능한 95량체	절단 불가능한 95량체	절단 가능한 200량체	절단 가능한 95량체		
총 관독물의 %	11.4	22.4	11.8	16.7	12.9	11
총 분자수의 %	29	41.4	7.5	28.1	20.3	16.3
세포당 평균 몰수	1043	1679	303	NA	NA	NA
세포당 중앙(median) 몰수	895	1536	256	NA	NA	NA
올리고 RSEC 서열 깊이	3.9	4.6	11.54	1.74	1.87	2
혈액 패널 RSEC 서열 깊이	10.72	10.72	10.72	4.1	4.1	4.1

[0809] 절단 가능한 95량체(도 23에서 가운데 컬럼에서 "단길이 1," "단길이 2," 및 "단길이 3"으로 표시됨) 중 어느 하나와 접합된 3가지 유형의 항-CD147 항체는 샘플 인덱싱에 성공적으로 사용되어 상이한 샘플의 라모스 및 주르카트 세포를 구별하였다. 서열분석 데이터에서 총 관독물 수의 11% 초과가 샘플 인덱싱 올리고뉴클레오타이드에 기인하였다. 백분율은 주르카트 및 라모스 세포 둘 모두를 함유하는 샘플의 경우에 가장 높았다(16.7%).

[0810] 도 24a 내지 도 24c는 상이한 샘플 인덱싱 올리고뉴클레오타이드(절단 가능한 95량체, 절단 불가능한 95량체, 및 절단 가능한 200량체)와 접합된 3가지 유형의 항-CD147 항체가 CD147의 단백질 발현 수준을 결정하는 데 사용될 수 있음을 나타내는 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 플롯이다. 세포의 발현 프로파일의 tSNE 투영 플롯 상에서의, 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 결정된 CD147 발현의 오버레이는 상이한 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 결정된 CD147 발현 패턴이 유사하였음을 나타낸다. 3개의 샘플 인덱싱 올리고뉴클레오타이드의 검출은 100%였다. 도 25는 세포당 GAPDH 발현의 오버레이를 갖는 비제한적이고 예시적인 tSNE 플롯이다. 플롯의 우측 상단 모서리에 있는 "세포수"는 GAPDH가 낮았는데, 이는 이것이 실제 세포수

가 아님을 시사한다. 이들 "세포수"는 샘플 인덱싱 올리고뉴클레오타이드에서 높았다.

- [0811] 도 26a 내지 도 26c는 3가지 유형의 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 검출된 샘플 인덱싱 올리고뉴클레오타이드의 분자수를 나타내는 비제한적이고 예시적인 히스토그램이다. 절단 불가능한 샘플 인덱싱 올리고뉴클레오타이드가 검출된 분자 수의 관점에서 가장 감도가 높았다.
- [0812] 도 27a 내지 도 27c는 CD147 발현이 분열성 세포에서 더 높았음(도 23에서의 "플로우셀 1"을 사용하여 결정됨)을 나타내는 플롯 및 막대 차트이다. CD147 발현은 주르카트 세포의 샘플 내의 세포 전체에 걸쳐 균일하지 않았다. 세포는 세포 사이클 유전자 AURKB에 기초하여 분류될 수 있다. AURKB 유전자의 mRNA 발현은 샘플 인덱싱을 사용하여 결정된 CD147 단백질 발현과 상관되었다. 도 27c는 AURKB+ 및 AURKB- 세포에서 3개의 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 결정된, CD147 발현의 비교를 나타낸다.
- [0813] 도 28a 내지 도 28c는 샘플 인덱싱이 상이한 샘플의 세포를 식별하는 데 사용될 수 있음을 나타내는, 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다. 488개의 혈액 패널 유전자의 발현 프로파일을 함유하는, 서열분석 데이터에서의 샘플 인덱싱 올리고뉴클레오타이드의 판독물을 필터링하였다. 필터링된 서열분석 데이터를 사용하여, 도 28a 내지 도 28c에서의 tSNE 투영 플롯을 생성하였다. 주르카트 세포에 상응하는 클러스터와 라모스 세포에 상응하는 클러스터(도 23에서 "단길이 1"로 표시된 샘플 인덱싱 올리고뉴클레오타이드의 풍부도에 기초하여 결정됨)는 도 28a의 tSNE 투영 플롯에서 명백히 분리되었다(도 28b 및 도 28c에 대해서도 비슷함).
- [0814] 샘플 인덱싱은 세포 유형과 올바르게 매칭되었다. 예를 들어, 도 28b는, 도 23에서 "단길이 2"로 표시된 샘플 인덱싱 올리고뉴클레오타이드로 표시된 라모스 세포가 CD147의 높은 발현을 갖는 반면, 표시되지 않은 주르카트 세포가 CD147의 최소한의 발현을 가졌음을 나타낸다. 도 28c는, 도 23에서 "단길이 3"으로 표시된 샘플 인덱싱 올리고뉴클레오타이드로 표시된 주르카트 세포가 CD147의 높은 발현을 갖는 반면, 표시되지 않은 라모스 세포는 CD147의 낮은 발현을 가졌음을 나타낸다. 3.6% 빈도의 작은 더블렛 클러스터는 웰의 점유의 이미지 분석으로부터 예상된 더블렛 비율인 3.92%와 근접하게 매칭되었다.
- [0815] 도 29a 내지 도 29c는 결정된 샘플 인덱싱 올리고뉴클레오타이드의 분자수에 기초한 세포당 샘플 인덱싱 서열의 비제한적이고 예시적인 히스토그램이다. 서열분석 데이터를 250의 컷오프를 사용하여 분포 기반 오류 정정(distribution based error correction, DBEC)으로 정정하였다. DBEC는 2017년 5월 25일에 출원된 미국 특허 출원 15/605874에 기재되어 있으며, 이의 내용은 전체가 본 명세서에 참고로 포함된다. DBEC 후에, 1000s에서의 "신호" 피크와 구별가능한 100s에서의 명백한 "노이즈" 피크가 있는 것으로 나타났다(도 29b 및 도 29c). 도 29b 및 도 29c는 또한 샘플 인덱싱 올리고뉴클레오타이드의 검출이 250의 노이즈 감산 후에 높은 100s에 있을 것임을 나타낸다.
- [0816] 도 30a 내지 d는 CD3D(주르카트 세포의 경우) 및 JCHAIN(라모스 세포의 경우)의 mRNA 발현 및 주르카트 및 라모스 세포의 샘플 인덱싱에 기초하여 결정된 세포 유형의 주석을 비교하는 비제한적이고 예시적인 플롯이다. 도 30a는 JCHAIN 발현의 신호 피크와 세포당 샘플 인덱싱 올리고뉴클레오타이드의 분자 수가 적은 노이즈 신호의 명백한 분리를 나타내는 비제한적이고 예시적인 히스토그램이다. 도 30b는 라모스 발현의 신호 피크와 세포당 샘플 인덱싱 올리고뉴클레오타이드의 분자 수가 적은 노이즈 신호의 명백한 분리를 나타내는 비제한적이고 예시적인 히스토그램이다. 더블렛은 250 초과와 총 수를 갖는 하나 초과와 샘플 인덱싱 서열을 갖는 세포로서 식별되었다. 도 30c에서, 라모스 세포는 JCHAIN의 mRNA 발현을 사용하여 식별되었고, 주르카트 세포는 CD3D의 mRNA 발현을 사용하여 식별되었다.
- [0817] 라모스 세포는 절단 가능한 상이한 샘플 인덱싱 올리고뉴클레오타이드(도 23에서 샘플 인덱싱 올리고뉴클레오타이드는 "단길이 1" 및 "단길이 2"로 표시됨)와 접합된 2가지 유형의 항-CD147 항체로 표시하였다. 주르카트 세포는 절단 가능한 상이한 샘플 인덱싱 올리고뉴클레오타이드(도 23에서 샘플 인덱싱 올리고뉴클레오타이드는 "단길이 1" 및 "단길이 3"으로 표시됨)와 접합된 2가지 유형의 항-CD147 항체로 표시하였다. 도 30d에서, 라모스 세포의 정체는 서열분석 데이터에서 "단길이 2" 샘플 인덱싱 올리고뉴클레오타이드의 총 수(250 초과)의 수에 기초하여 식별되었고, 주르카트 세포의 정체는 서열분석 데이터에서 "단길이 3" 샘플 인덱싱 올리고뉴클레오타이드의 총 수(250 초과)의 수에 기초하여 식별되었다.
- [0818] 도 31a 내지 도 31c는 CD3D 및 JCHAIN(도 31a), JCHAIN(도 31b), 및 CD3D(도 31c)의 mRNA 발현의 오버레이를 갖는 주르카트 및 라모스 세포의 mRNA 발현 프로파일의 비제한적인 tSNE 투영 플롯이다. JCHAIN의 mRNA 발현은 예상된 바와 같이 라모스 세포에서 더 높았고(도 31b), 주르카트 세포에서는 최소한이었다(도 31c). CD3D의 mRNA 발현은 주르카트 세포에서 더 높았고(도 31c), 라모스 세포에서는 최소한이었다(도 31b).

[0819] 도 28b 및 도 28c와 도 31b 및 도 31c의 비교는 상이한 샘플의 세포를 구별하는 데 있어서의 샘플 인텍싱의 높은 수행을 밝혀주었다. 샘플 기원을 결정하는 데 있어서 샘플 인텍싱의 수행은 도 32에 나타나 있다. 도 32는 250의 DBEC 컷오프로 샘플 인텍싱을 사용하여 결정된 세포 유형의 오버레이를 갖는 주르카트 및 라모스 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다. 표 2는 도 32에 나타난 샘플 인텍싱을 사용하여 결정된 세포 유형의 요약이다. 표 3은 샘플 기원을 결정하는 데 있어서의 샘플 인텍싱의 감도 및 특이성을 나타내는데, 이때 표 2에서의 더블렛 및 불명확한 세포는 제외하였다.

표 2

[0820] 샘플 인텍싱을 사용한 샘플 기원 결정의 요약.

	S2	S2+S3	S3	표지되지 않음
더블렛	0	36	1	0
주르카트	0	5	658	12
라모스	833	4	0	0
불명	30	0	17	0

표 3

[0821] 샘플 기원을 결정하는 데 있어서 샘플 인텍싱의 특이성 및 감도

	특이성	감도
S2(라모스)	99.5% (833/(833+4))	100% (833/(833+0))
S3(주르카트)	99.25% (658/(658+4))	98.2% (658/(658+12))

[0822] 도 33a 내지 도 33c는 "단길이 3" 샘플 인텍싱 올리고뉴클레오타이드, "단길이 2" 및 "단길이 3" 샘플 인텍싱 올리고뉴클레오타이드, 및 "단길이 2" 샘플 인텍싱 올리고뉴클레오타이드로 표지되지 않거나 표지된 라모스 및 주르카트 세포(도 33), 라모스 세포(도 33b), 및 주르카트 세포(도 33c)에 대한 세포당 샘플 인텍싱 올리고뉴클레오타이드의 분자수의 비제한적이고 예시적인 막대 차트이다. 1% 미만의 단일 세포가 "단길이 2" 및 "단길이 3" 샘플 인텍싱 올리고뉴클레오타이드 둘 모두로 표지되었다(도 34a 내지 도 34c).

[0823] 도 35a 내지 도 35c는 도 23에 개략적으로 설명된 바와 같이 상이한 플로우셀을 사용하여 제조된 샘플간의 주르카트 및 라모스 세포의 발현 프로파일에 대한 배치 효과를 나타내는 비제한적이고 예시적인 tSNE 플롯이다. 도 35a는 세포의 상이한 샘플 조제물의 오버레이를 갖는 주르카트 및 라모스 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯을 나타낸다. 상이한 샘플 조제물 사이의 488개의 혈액 패널 유전자(도 35b) 및 CD3D 유전자(도 35c)에서 mRNA 발현 프로파일의 변동은 상이한 플로우셀을 사용하여 제조된 샘플의 상이한 서열분석 깊이로 인한 배치 효과의 결과일 수 있다.

[0824] 종합하면, 이들 데이터는 샘플 인텍싱 올리고뉴클레오타이드의 3가지 모든 버전(절단 가능한 95량체, 절단 불가능한 95량체, 및 절단 가능한 200량체)이 샘플 인텍싱에 사용될 수 있으며, 이때 절단 불가능한 95량체가 최고의 효능을 가짐을 나타낸다. 샘플 기원을 결정하기 위한 샘플 인텍싱은 높은 특이성 및 감도를 가질 수 있다.

[0825] 실시예 4

[0826] 고온:저온 항체 역가

[0827] 본 실시예는, 항체 올리고뉴클레오타이드가 서열분석 데이터에서 총 판독물의 원하는 백분율(예를 들어, 2%)을 차지하도록 하는, 올리고뉴클레오타이드-접합된 항체("고온 항체")와 올리고뉴클레오타이드와 접합되지 않은 항체("저온 항체")의 비의 결정을 보여준다.

[0828] PE 완충액을 이용하여 1:20, 1:100, 1:200, 1:300, 1:400, 1:600, 1:800, 및 1:1000의 희석물로 항-CD147 항체 스톱을 희석하였다. 100 μ l의 염색 완충액(FBS)(BD(미국 뉴저지주 프랭클린 레이크 소재)) 내 약 150000 주르카트 세포를 다양한 항체 희석물로, 실온에서 20분 동안 염색하였다. 염색 후, 세포를 500 μ l의 염색 완충액으로 1회 세척하고, 형광 강도의 측정을 위해 200 μ l에 재현탁시켰다. Muse™ Autophagy LC3-항체(EMD Millipore(미국 매사추세츠주 빌레리카 소재))를 사용하여 주르카트 세포에 결합된 항-CD147 항체를 검출하였다. 다양한 항-

CD147 항체 희석물로 염색된 세포로부터의 형광 강도 또는 염색되지 않은 세포를 결정하고, 비교하여 항체에 대한 최적 희석을 결정하였다(도 36aa 내지 도 36ai, 도 36b 및 도 36c). 형광 강도는 희석이 더 높아짐에 따라 감소되었다. 99% 초과인 세포를 1:800의 희석 비율로 염색하였다. 형광 신호는 1:800에서 떨어지기 시작했다. 세포를 희석 비 1:200 이하까지 포화 염색시켰다. 세포를 희석 비 1:400 이하까지 포화에 가깝게 염색시켰다.

[0829] 도 37은 항체 올리고뉴클레오타이드가 서열분석 데이터에서 총 판독물의 원하는 백분율을 차지하도록 하는 올리고뉴클레오타이드-접합된 항체의 염색 농도 결정을 위한 비제한적이고 예시적인 실험 설계를 나타낸다. 항-CD147 항체는 절단 가능한 95량체 항체 올리고뉴클레오타이드와 1:3의 항체:올리고뉴클레오타이드 비("고온 항체")로 접합되었다. 고온 항체를 pH 7.5 희석액을 사용하여 1:100 비 또는 1:800 비로 희석하였다. 10% 고온 항체:90% 저온 항체의 혼합물을 9 μ l의 저온 항-CD147 항체 및 1 μ l의 고온 항체를 사용하여 제조하였다. 1% 고온 항체:90% 저온 항체의 혼합물을 9 μ l의 저온 항-CD147 항체 및 1 μ l의 10% 고온 항체:90% 저온 항체의 혼합물을 사용하여 제조하였다.

[0830] 약 5십만 개의 세포를 갖는 해동된 말초혈액 단핵 세포(PBMC)를, 100% 고온 항체(1%의 스톱 고온 항체), 10% 고온 항체:90% 저온 항체의 혼합물(0.1%의 스톱 고온 항체), 1% 고온 항체:99% 저온 항체의 혼합물(0.01%의 스톱 고온 항체)로 1:100 희석된 스톱, 및 100% 고온 항체로 1:800 희석된 스톱(0.0125%의 스톱 고온 항체)를 이용하여, 100 μ l의 염색 완충액(FBS) 중에서 염색하였다. 염색 후, 세포를 세척하여 결합되지 않은 항체 분자를 제거하였다. 살아있는 세포를 수득하기 위하여 유세포분석을 이용하여 분류하기 위해 세포를 Calcein AM 및 Draq7TM로 염색하였다. 세포를 세척하여 과량의 Calcein AM 및 Draq7TM을 제거하였다. Calcein AM 으로 염색되고(살아있는 세포), Draq7TM로는 염색되지 않은(죽지 않거나 투과화되지 않은 세포) 단일 세포를 유세포분석을 사용하여, BD RhapsodyTM 카트리지를 내로 분류하였다.

[0831] 단일 세포 및 비드를 함유하는 웰 중, 웰 내의 단일 세포 중 1000개를 용해 완충액 내에 용해하였다. 각각의 단일 세포의 경우, mRNA 분자를 역전사하였으며, 항체 올리고뉴클레오타이드를 그 세포에 대한 비드와 접합된 추적적 바코드를 사용하여 복제하였다. 혈액 패널 N1 프라이머를 사용하여 488개 혈액 패널 유전자의 mRNA 발현 프로파일을, 및 샘플 인덱싱 올리고뉴클레오타이드 N1 프라이머("PCR 1")를 사용하여 CD147 단백질의 발현을 결정하기 위한 프라이머를 사용하여, 60도의 어닐링 온도에서 15회의 사이클 동안 역전사 및 복제 후 샘플을 PCR 증폭시켰다. 과량의 프라이머는 앰플어 클리닝을 이용하여 제거하였다. PCR1로부터의 산물을 혈액 패널 N2 프라이머 및 어댑터 결합을 위한 측점 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드 N1 프라이머를 사용하여 60도의 어닐링 온도에서 15회의 사이클 동안 추가로 PCR 증폭시켰다("PCR 2"). 서열분석 어댑터 결합 후 서열분석 데이터를 획득하고 분석하였다("PCR 3").

[0832] 도 38a 내지 도 38d는 항체 올리고뉴클레오타이드의 예측된 크기와 일치하는 피크(화살표로 표시됨)를 나타내는 비제한적이고 예시적인 생체분석자 추적선이다. 항체 올리고뉴클레오타이드 피크는, 고온 항체가 저온 항체로 적정됨에 따라 감소되었다.

[0833] 표 4는 서열분석 데이터 매트릭스의 요약이다. 1:100 희석된 스톱을 사용하여 제조된 1% 고온 항체:99% 저온 항체의 혼합물로 세포를 염색함으로써, 항체 올리고뉴클레오타이드는 서열분석 데이터에서 총 원래 판독물의 2.4%를 차지하였다. 그러나, 도 39aa 내지 39ac 및 도 39ba 내지 도 39bc 및 도 40b, 도 41b, 및 도 42b에 나타난 바와 같이, 재귀 치환 오류 정정(RSEC) 또는 분포-기반 오류 정정(DBEC) 후 검출된 항체 올리고뉴클레오타이드의 분자 수의 분포 히스토그램은, 세포가 1:100 희석된 스톱을 사용하여 제조된 1% 고온 항체:99% 저온 항체의 혼합물을 이용하여 염색된 경우, 명확한 신호 피크를 포함하지 않았다. RSEC는 2017년 5월 25일에 출원된 미국 특허 출원 15/605874에 기재되어 있으며, 이의 내용은 전체가 본 명세서에 참고로 포함된다.

[0834] 도 40a 내지 도 40c는 올리고뉴클레오타이드-접합된 항-CD147 항체 분자가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형을 혈액 패널 내 488개의 유전자의 발현 프로파일을 사용하여 결정하였다(도 40a). 1:100 희석된 스톱을 사용하여 제조된 10% 고온 항체:90% 저온 항체의 혼합물로 세포를 염색하였으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 명확한 신호를 생성하였다(도 40b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 40c에 나타난다.

[0835] 도 41a 내지 도 41c는 올리고뉴클레오타이드-접합된 항-CD147 항체가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형을 혈액 패널에서 488개의 유전자의 발현 프로파일

을 사용하여 결정하였다(도 41a). 1:100 희석된 스톱을 사용하여 제조된 1% 고온 항체:99% 저온 항체의 혼합물로 세포를 염색하였으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 명확한 신호를 생성하지 않았다(도 41b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 41c에 나타난다.

[0836] 도 42a 내지 도 42c는 올리고뉴클레오타이드-접합된 항-CD147 항체 분자가 다양한 세포 유형을 표지하는 데 사용될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 세포 유형을 혈액 패널에서 488개의 유전자의 발현 프로파일을 사용하여 결정하였다(도 42a). 1:800의 희석된 스톱을 이용하여 세포를 염색하였으며, 이는 검출된 항체 올리고뉴클레오타이드의 분자 수를 나타내는 히스토그램에서 명확한 신호를 생성하였다(도 42b). 항체 올리고뉴클레오타이드에 의한 다양한 세포 유형의 표지는 도 42c에 나타난다.

표 4

[0837] 서열분석 데이터 매트릭스의 요약.

샘플	FC4 - 1:800 희석률, 저온 항체 없음	FC3 - 1:100 희석률, 1:100 저온 항체	FC2 - 1:100 희석률, 1:10 저온 항체
총 미가공 판독물	31.3M	27.3M	29.2M
올리고에 배정된 총 미가공 판독물	9161642 (29.2%)	660044 (2.4%)	4013438 (13.7%)
검출된 세포수	1010	983	907
RSEC 올리고 MI	577110	20054	170742
DBEC 올리고 MI	477216	9319	110629
% Q30	76.45	71.89	73.45
세포 표지에 정렬된 %	84.63	79.47	82.58
앰플리콘에 고유하게 배정된 %	73.19	65.4	69.58
평균 미가공 서열 길이	5.58	6.52	6.4
평균 RSEC 서열 길이	8.69	10.41	10.25
평균 DBEC 서열 길이	15.96	23.29	22.48
Ab올리고 RSEC 서열 길이	8.08	12.9	11.2
Ab올리고 DBEC 서열 길이	12.4	30.9	21.2
세포당 평균 판독물	11257	14414	15150
세포당 평균 분자수	553.8	608.3	647.1
세포당 중앙 몰수	246.5	279	278
패널 내의 유전자의 수	489	489	489
검출된 총 유전자수	438	439	441
세포당 평균 유전자수	68.39	73.6	73.12

[0838] 종합하면, 이들 데이터는, 올리고뉴클레오타이드-접합된 항체("고온 항체") 및 올리고뉴클레오타이드와 접합되지 않은 항체("저온 항체")의 비가, 항체 올리고뉴클레오타이드가 서열분석 데이터에서의 총 판독물의 원하는 백분율을 차지하고, 신호 항체 올리고뉴클레오타이드를 나타내는 데이터가 노이즈 항체 올리고뉴클레오타이드를 나타내는 데이터로부터 명확히 분리되도록 조정될 수 있음을 나타낸다.

[0839] 실시예 5

[0840] 정규화

[0841] 본 실시예는 올리고뉴클레오타이드-접합된 항체("고온 항체") 및 올리고뉴클레오타이드와 접합되지 않은 항체("저온 항체")의 혼합물을 이용하여, 어떻게 정규화가 항체의 단백질 표적의 풍부도에 관계 없이, 원하는 커버리지로 서열분석 데이터에서 총 판독물의 원하는 백분율을 차지하는 항체 올리고뉴클레오타이드를 생성할 수 있는지를 보여준다.

[0842] 표 5는 고온 항체를 사용하여 10000개의 B 세포에서 가변하는 풍부도의 3개의 세포 표면 마커 중 정량의 비교를 나타낸다. 3개의 세포 표면 마커의 상대적 발현 수준을 해결하는 데 요구되는 총 판독물의 수는 4752만 판독물이었다.

표 5

[0843] 고온 항체를 사용한 단백질 정량화의 예.

항원	세포당 분자수	상대 풍부도	고온 항체:저온 항체 비	서열분석에 의해 검출된 세포당 분자수	서열분석 깊이가 4로 주어진 경우의 관독물
CD21	210000	105	1:0	840	33.6M
HLA-DR	85000	42.5	1:0	340	13.6M
CD40	2000	1	1:0	8	0.32M

표 6

[0844] 고온 및 저온 항체를 사용한 단백질 정량화의 예.

항원	세포당 분자수	상대 풍부도	고온 항체:저온 항체 비	서열분석에 의해 검출된 세포당 분자수	서열분석 깊이가 4로 주어진 경우의 관독물 수	항체 비에 기초한 예상된 분자수	서열분석에 의한 상대 풍부도
CD21	210000	105	1:100	8.3	0.33M	8.3 X 100=830	103.75
HLA-DR	85000	42.5	1:40	8.5	0.34M	8.5 X 40=340	42.5
CD40	2000	1	1:0	8	0.32M	8 X 1=8	1

[0845] 표 6은 3개 세포 표면 마커의 상대 발현 수준을 해결하는 데 요구된 관독물의 총 수는, 고온 항체:저온 항체의 혼합물을 사용하여 백만 관독물이었음을 나타낸다. 또한, 표 5에 나타난 정량 결과(1 백만 관독물 대 4752만 관독물)와 비교하여, 3개의 세포 표면 마커의 발현 수준을 정량하는 데 고온 항체:저온 항체의 혼합물이 사용된 경우, 3개의 모든 단백질 마커 중 최적 커버리지(예를 들어, 서열분석 깊이 4)를 달성하는 데 관독물의 수의 단지 2%만이 필요하다. 더 높은 백분율의 저온 항체를 갖는 혼합물을 사용하여, 고발현 단백질 분자의 정규화는 낮은 풍부도의 단백질의 검출, 파라미터 수, 및 서열분석 비용 간의 상충을 감소시켜서, 분석을 더 매력적인 도구로 만든다.

[0846] 종합하면, 이들 데이터는 원하는 커버리지를 갖는 서열분석 데이터에서 총 관독물의 원하는 수가 고온 항체:저온 항체의 혼합물을 사용하여 상이한 풍부도의 단백질 표적(예를 들어, 항원)을 위해 달성될 수 있음을 나타낸다.

[0847] 실시예 6

[0848] 제어 입자

[0849] 본 실시예는 상이한 서열을 갖는 제어 입자 올리고뉴클레오타이드를 포함하는 제어 입자의 생성 및 포획 효율을 결정하기 위한 작용화된 제어 입자의 용도를 보여준다.

[0850] 재료

[0851] BD CompBead Plus 항-마우스 Ig(7.5 um) 입자 세트(51-9006274)

[0852] BD 염색 완충액(FBS)

[0853] 절차

[0854] 1. BD CompBead Plus를 사용 전에 철저히 와동시킨다(적어도 1분).

[0855] 2. 800 uL의 염색 완충액을 튜브에 첨가한다(표 7은 상이한 풍부도로 5개의 서열을 갖는 올리고뉴클레오타이드에 접합된 CD147을 갖는 염색 완충액의 조성을 나타냄. 도 43a는 염색 완충액의 조성을 나타내는 플롯임.).

[0856] 3. CompBead Plus의 항-마우스의 5개의 충분한 액적(대략 300 uL)을 첨가한다.

[0857] 4. 20 uL의 하기 염색 각테일을 튜브에 첨가한다. 와동시킨다.

[0858] 5. 광을 피하여 실온에서 30분 인큐베이션한다.

- [0859] 6. 200 g에서 10분 동안 비드를 회전시킨다.
- [0860] 7. 상청액을 조심히 제거하고, 1 mL 염색 완충액으로 재현탁시킨다.
- [0861] 8. 200 g에서 10분 동안 비드를 회전시킨다.
- [0862] 9. 상청액을 조심히 제거하고, 1 mL 염색 완충액으로 재현탁시켜 작용화된 CompBead 스톱 용액을 생성한다.
- [0863] 10. 비드를 계수한다.

표 7

[0864] 염색 카테일의 조성

항체	염색 완충액 중 최종 %	이전 희석률	염색 용액(μl)
CD147-LZ15	1	1:1	1
CD147-LZ16	0.2	1:5	1
CD147-LZ17	0.1	1:10	1
CD147-LZ18	0.02	1:50	1
CD147-LZ19	0.01	1:100	1
염색 완충액			95

[0865] 결과

[0866] 도 44a 및 도 44b는 혈구계에서 세포의 명시야 이미지(도 44a, 백색 원) 및 제어 입자(도 44b, 흑색 원)이다. 도 45a 및 도 45b는 형광단과 회합된 올리고뉴클레오티드-접합된 항체에 결합된 제어 입자의 위상차(도 45a, 10 X) 및 형광(도 45b, 10 X) 이미지이다. 형광 현미경을 사용하여 5 ul의 작용화된 CompBead 스톱 용액은 약 400000개의 작용화된 제조된 CompBeads와 함께 약 2000개의 세포(4%의 총 투입)를 함유하였음을 결정하였다.

[0867] 도 46은 카트리지의 마이크로웰 내로 로딩되는 세포 및 입자를 나타내는 제어 입자의 이미지이다. CompBeads는 규칙적인 Rhapsody™ 실험과 함께 사용될 수 있다. 522개의 작용화된 CompBeads를 복수의 세포 내로 첨가하였다. 20000개의 세포(제어 입자 포함) 서열 중, 156개는 20개 초과인 모든 제어 입자 올리고뉴클레오티드의 합계를 가졌다. 이에 따라, 156개의 제어 입자는 서열이었다. 도 43b는 염색 완충액 내 이들의 풍부도와 상관된 5개의 상이한 제어 바코드 서열(LZ15 내지 LZ19)을 갖는 제어 입자 올리고뉴클레오티드의 수를 나타내는 플롯이다.

[0868] 종합하면, 이들 데이터는 입자(예를 들어, CompBead Plus)가 올리고뉴클레오티드(예를 들어, 제어 입자 올리고뉴클레오티드)로 작용화될 수 있음을 나타낸다. 작용화된 입자는 단일 세포 서열분석 작업 흐름을 사용하여 포획되고 서열분석된 입자의 수를 결정할 수 있다.

[0869] 실시예 7

[0870] 샘플 인덱싱을 위한 항체 카테일

[0871] 본 실시예는 샘플 인덱싱을 위한 항체 카테일을 사용함으로써 표지화 감도를 증가시킬 수 있음을 보여준다.

[0872] 샘플 인덱싱은 샘플을 표지하기 위하여 다수의 단백질 표적에 대한 올리고뉴클레오티드-접합된 항체를 이용할 수 있다. 본 실시예에서는, 단일 항체를 사용하는 대신에, 2개의 항체의 카테일을 대신 사용하였다. 항체의 카테일을 동일한 샘플 인덱싱 올리고뉴클레오티드로 표지하였다. 도 47a 내지 도 47c는 샘플 인덱싱을 위한 항체 카테일을 사용함으로써 표지화 감도를 증가시킬 수 있음을 나타내는 플롯이다. 도 47a 내지 도 47c는 CD147 항체, CD47 항체, 및 CD147 및 CD47 항체 둘 모두에 의한 PBMC의 표지화 감도를 나타낸다. 도 47c는 CD147 및 CD47 항체 둘 모두가 사용되었을 때 더 많은 세포가 표지되었고, 더 우수한 신호 대 노이즈 분리가 달성되었음을 나타낸다. 표 8은 CD147 및 CD47 항체 둘 모두가 사용되었을 때 증가된 표지화 감도를 나타낸다. CD147 및 CD47 항체 둘 모두의 사용은 이중성 샘플 유형을 표지화하는 데 있어서 더 높은 감도를 초래하였다.

표 8

[0873] CD147 항체, CD47 항체, 또는 CD147 및 CD47 항체 둘 모두에 의한 표지화 감도.

항체	감도
CD147	86.2%
CD47	82.5%
CD147+CD47	92.6%

[0874] 종합하면, 이 데이터는 항체 콕테일의 사용이 표지화 감도를 증가시킬 수 있음을 나타낸다. 이는 세포 유형간 및 세포 상태간에 단백질 발현이 다양할 수 있어서, 시험되는 모든 세포 유형을 표지하는 범용 항체를 찾을 수 있기 때문일 수 있다. 항체 콕테일의 사용은 더 많은 샘플 유형의 더 감도 높고 효율적인 표지화를 가능하게 할 수 있다. 항체 콕테일은 또한 더 넓은 범위의 항체를 포함할 수 있다. 상이한 샘플 유형을 잘 표지하는 항체는 함께 풀링되어 모든 세포 유형 또는 관심 대상의 세포 유형을 충분히 표지하는 콕테일을 생성할 수 있다.

[0875] 실시예 8

[0876] 데이터세트에서의 멀티플렛의 존재

[0877] 본 실시예는 서열분석 데이터에서의 멀티플렛이 상이한 샘플 인덱싱 올리고뉴클레오타이드로 세포를 태깅함으로써 식별되고 제거될 수 있음을 설명한다.

[0878] 도 48은 멀티플렛이 샘플 인덱싱을 사용하여 식별되고 서열분석 데이터로부터 제거될 수 있음을 나타내는 비제한적이고 예시적인 플롯이다. 도 48에 나타난 바와 같이, 20000개의 단일 세포를 갖는 샘플이 액적-기반 플랫폼을 사용하여 액적 내로 분배될 때, 더블렛 및 단일 세포 발현 데이터에서의 멀티플렛의 비율은 8%에 근접할 수 있다. 20000개의 단일 세포를 갖는 샘플이 Rhapsody™ (Becton, Dickinson and Company(미국 뉴저지주 프랭클린 레이크스 소재))의 웰 내로 분배될 때, 마이크로웰 어레이 및 단일 세포 발현 데이터에서의 멀티플렛의 비율은 훨씬 더 낮을 수 있다(예를 들어, 단지 4%에 근접함). 서열분석 데이터에서의 멀티플렛은 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별되고 제거될 수 있다. 예를 들어, 20000개의 단일 세포를 갖는 샘플이 2개의 하위개체군으로 나누어지고, 각각의 하위개체군 내의 세포는 동일한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드로 표지되고, 상이한 하위개체군 내의 세포는 2개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드로 표지될 때, 멀티플렛과 연관된 서열 데이터는 식별되고 제거될 수 있는데, 이때 생성된 서열분석 데이터는 약 2%의 멀티플렛을 함유하였다. 8개의 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드의 경우, 20000개의 단일 세포의 서열분석 데이터는 단지 약 0.5%의 멀티플렛을 포함할 수 있다. 더 많은 수의 샘플 인덱싱 서열의 경우, 서열분석 데이터에서의 멀티플렛의 비율은 훨씬 더 낮을 수 있다.

[0879] 따라서, 단일 세포 서열분석 데이터에 잔류하는 멀티플렛의 비율은 상이한 샘플 인덱싱 서열을 갖는 샘플 인덱싱 올리고뉴클레오타이드로 세포를 표지함으로써 실질적으로 낮아지거나 제거될 수 있다.

[0880] 실시예 9

[0881] 멀티플렛을 식별하기 위한 샘플 인덱싱

[0882] 본 실시예는 샘플 태그를 사용하여 서열분석 데이터에서 멀티플렛을 식별하고 제거하는 것을 보여준다.

[0883] 2마리의 마우스로부터 6개의 조직(골수, 지방(성선 백색 지방 조직(gWAT)), 결장, 간, 폐, 및 비장)을 입수하였다. 이들 단리된 조직으로부터 CD45+ 단일 세포를 단리하고, 형광-활성화 세포 분류(FACS)를 사용하여 분류하여 12개의 샘플을 생성하였다. 12개의 샘플의 CD45+ 단일 세포를 RNA-서열분석을 위한 BD™ 단일-세포 다중화 키트(Single-Cell Multiplexing Kit)를 사용하여 12개의 상이한 샘플 태그(본 명세서에서 샘플 인덱싱 조성물로 지칭됨)로 태깅하고, Rhapsody™ 카트리지 상에 로딩하였다. 샘플 태그는, 3' RNA-서열 검정(예컨대, BD Rhapsody™ 검정)에서 포획되고 증폭될 수 있는, 올리고뉴클레오타이드(본 명세서에서 샘플 인덱싱 올리고뉴클레오타이드로 지칭됨)에 접합된 항체였다. 그러한 샘플 태그는 높은 특이성 및 감도(99% 초과)를 가졌다. 세포로부터의 mRNA 분자 및 샘플 인덱싱 올리고뉴클레오타이드를 Rhapsody™ 자성 비드를 사용하여 포획하고, Rhapsody™ 면역 반응 패널-마우스(Immune Response Panel-Mouse)(마우스)를 사용하여 바코딩하고 증폭시켰다. 세포의 발현 프로파일을 합성 멀티플렛 발현 프로파일을 사용하여 싱글렛 발현 프로파일 및 멀티플렛 발현 프로파일로서 결

정하고 식별하였다.

- [0884] 도 49는 샘플 태그를 사용하여 싱글렛 또는 멀티플렛으로서 식별된 2마리의 마우스로부터의 6개의 조직의 12개의 샘플로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다. 도 49는 다수의 클러스터로서의 멀티플렛 발현 프로파일을 나타낸다. 상이한 조직을 태깅하기 위한 샘플 태그 서열(본 명세서에서 샘플 인덱싱 올리고뉴클레오타이드의 샘플 인덱싱 인덱스로 지칭됨)을 사용하여, 2가지 이상의 세포 유형 또는 아형의 세포의 발현 프로파일을 각각 포함하는 멀티플렛 발현 프로파일을 식별하였다. 2개의 상이한 조직으로부터의 CD45+ 세포의 멀티플렛을 이들 세포로부터 획득된 서열 데이터에서의 샘플 태그 서열의 존재에 기초하여 식별하였다. 핵산 샘플 태그를 사용하여 샘플을 인덱싱하고 멀티플렛을 식별하는 수행은 합성 멀티플렛 발현 프로파일을 사용하여 멀티플렛을 식별하는 수행과 비견되었다. 도 49b는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별된 멀티플렛이 제거된 2마리의 마우스로부터의 6개의 조직의 12개의 샘플로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다.
- [0885] 도 50a는 샘플 인덱싱 올리고뉴클레오타이드를 사용하여 식별된 멀티플렛이 제거된 2마리의 마우스로부터의 CD45+ 단일 세포의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다. 멀티플렛 발현 프로파일이 제거된 상태에서, 생물학적 복제물인 2마리의 마우스는 도 50a 내지 도 50c에 나타나 있는 바와 같이 유사한 발현 프로파일을 나타내었다. 멀티플렛 발현 프로파일을 제거한 후의 6개의 상이한 조직의 면역 프로파일링의 결과가 도 51a 내지 도 51f에 나타나 있다. 도 51a 내지 도 51f에 나타나 있는 6개의 샘플에서의 상이한 세포 유형(예를 들어, B 세포, 대식세포 등)을 세포의 발현 프로파일에 기초하여 식별하였다.
- [0886] 6개의 상이한 조직으로부터의 대식세포, T 세포, 및 B 세포의 예시적인 발현 프로파일이 도 52a 내지 도 52c에 나타나 있다. 도 52a 내지 도 52c는, 상이한 조직에서의 동일한 세포 유형이 고유 발현 프로파일을 나타내었고, 결장 면역 개체군이 6개의 조직 중에서 가장 뚜렷하였음을 나타낸다. 도 53a 내지 도 53d는 결장과 비장에서의 대식세포를 비교하는 비제한적이고 예시적인 플롯이다. 도 53a는 결장 및 비장에서의 대식세포의 발현 프로파일을 비교하는 비제한적이고 예시적인 tSNE 투영 플롯으로서, 이는 2개의 조직에서의 대식세포가 별개의 발현 프로파일을 가졌음을 나타낸다. 도 53b는 결장 및 비장에서 대식세포의 발현 프로파일이 낮은 상관관계를 가졌음을 나타내는 비제한적이고 예시적인 플롯이다. 낮은 상관관계는 2개의 조직에서의 대식세포가 도 53a에 나타난 별개의 발현 프로파일을 가졌다는 것을 추가로 뒷받침하였다. 도 53c 및 도 53d는 결장 및 비장에서의 대식세포 내의 Areg 유전자 및 Rora 유전자의 발현 프로파일의 비제한적이고 예시적인 tSNE 투영 플롯이다. 이들 2개의 유전자는 결장 및 비장에서 대식세포의 5.2% 및 9.5%로 발현되었다. 도 53a, 도 53c, 및 도 53d는 이들 2개의 유전자가 단지 결장에서의 대식세포에서만 발현되었음을 나타낸다. 도 52a 내지 도 52c 및 도 53a 내지 도 53d에 나타난 바와 같은 6개의 상이한 조직으로부터의 28000개 초과 마우스 면역 세포의 단일-세포 프로파일링은 주요 면역 세포 유형에 대한 조직 특이적 유전자 발현 특징을 밝혀주었다.
- [0887] 종합하면, 이들 데이터는 샘플 태깅이 동일한 단일-세포 실험 상으로 다수의 샘플의 풀링을 가능하게 할 수 있음을 나타낸다. 샘플 태깅은 샘플 처리량을 증가시키고, 기술적 오류로 인한 샘플간의 배치 효과를 제거하고, 멀티플렛의 검출을 가능하게 할 수 있다.
- [0888] 실시예 10
- [0889] 샘플 인덱싱 및 발현 프로파일
- [0890] 본 실시예는 세포의 샘플 태깅의 사용이 세포의 mRNA 발현 프로파일을 변경시키지 않는다는 것을 보여준다.
- [0891] 도 54는 샘플 인덱싱 조성물로의 세포의 태깅 또는 염색이 말초 혈액 단핵 세포(PBMC)에서의 mRNA 발현 프로파일을 변경시키지 않았음을 나타내는 비제한적이고 예시적인 플롯이다. 도 54는 샘플 태깅을 갖거나 갖지 않는 PBMC의 발현 프로파일들의 tSNE 투영 플롯의 오버레이를 나타낸다. 이러한 오버랩은 샘플 태깅이 PBMC에서의 mRNA 발현 프로파일을 변경시키지 않았음을 나타낸다.
- [0892] 종합하면, 이 데이터는 올리고뉴클레오타이드와 접합된 항체에 의한 샘플 태깅 또는 염색이 세포의 발현 프로파일을 변경시키지 않음을 나타낸다. 따라서, 샘플 태깅은 태깅된 세포의 발현 프로파일을 변경시키지 않고서 다수의 샘플의 풀링을 가능하게 하고/하거나, 샘플 처리량을 증가시키고/증가시키거나, 샘플간의 배치 효과를 제거하고/제거하거나, 멀티플렛의 검출을 가능하게 할 수 있다.
- [0893] 용어
- [0894] 앞서 기재된 구현예 중 적어도 일부에서, 구현예에서 사용된 하나 이상의 요소는 또 다른 구현예에서 상호교환

가능하게 사용될 수 있는데, 단, 그러한 대체는 기술적으로 실현가능해야 한다. 청구된 발명 주제의 범주로부터 벗어나지 않고서 상기 기재된 방법 및 구조에 다양한 기타 다른 생략, 추가 및 변형이 이루어질 수 있음이 당업자에 의해 이해될 것이다. 모든 그러한 변형 및 변경은 첨부된 청구범위에 의해 정의된 바와 같은 발명 주제의 범주 내에 속하는 것이므로 한다.

[0895] 본 명세서에서 실질적으로 임의의 복수 및/또는 단수 용어의 사용에 대하여, 당업자는, 문맥 및/또는 적용에 적절한 바에 따라 복수에서 단수 및/또는 단수에서 복수로 번역할 수 있다. 다양한 단수/복수 치환은 명확성을 위해 본 명세서에 명시적으로 설명될 수 있다. 본 명세서 및 첨부된 청구범위에서 사용되는 바와 같이, 단수형("a", "an" 및 "the")은 맥락이 명확하게 달리 지시하지 않으면 복수형 지시 대상을 포함한다. 본 명세서에서 "또는"에 대한 임의의 언급은 달리 기재되지 않는 한 "및/또는"을 포괄하고자 한다.

[0896] 일반적으로, 본 명세서 및 특히 첨부된 청구범위(예를 들어, 첨부된 청구범위의 본문)에서 사용된 용어는 일반적으로 "개방형" 용어(예를 들어, 용어 "포함하는"은 "포함하지만, 이에 제한되지 않는"으로서 해석되어야 하고, 용어 "갖는"은 "적어도 갖는"으로 해석되어야 하며, 용어 "포함하다"는 "포함하지만, 이에 제한되지 않는" 등으로서 해석되어야 함)이고자 함은 당업자에 의해 이해될 것이다. 특정 수의 도입된 청구항 언급이 의도된 경우, 그러한 의도는 청구항에서 명시적으로 언급될 것이며, 그러한 언급 부재가 부재하면 그러한 의도는 없는 것으로 당업자는 추가로 이해할 것이다. 예를 들어, 이해에 대한 보조로서, 다음의 첨부된 청구범위는 도입 구절 "적어도 하나" 및 "하나 이상의"의 사용을 포함하여 청구범위 언급을 도입할 수 있다. 그러나, 그러한 구절의 사용은, 동일한 청구항이 도입 구절인 "하나 이상의" 또는 "적어도 하나" 및 부정관사(예컨대 "a" 또는 "an")(예를 들어, "a" 및/또는 "an"은 "적어도 하나" 또는 "하나 이상의"를 의미하는 것으로 해석되어야 함)를 포함하는 경우라도, 부정관사("a" 또는 "an")에 의한 청구항 언급의 도입이, 그러한 도입된 청구항 언급을 포함하는 임의의 특정 청구항을 단지 하나의 그러한 언급만을 포함하는 구현으로 제한하는 것을 의미하는 것으로 해석되어서는 안 되며; 청구항 언급을 도입하는 데 사용된 정관사의 사용에 대해서도 동일하게 적용된다. 추가적으로, 도입된 청구항 언급의 특정 수가 명시적으로 언급된 경우라도, 당업자는 그러한 언급이 적어도 언급된 수(예를 들어, 다른 수식어 없이 "2개의 언급"의 언급은 적어도 2개의 언급, 또는 2개 이상의 언급을 의미함)를 의미하는 것으로 해석되어야 함을 인식할 것이다. 나아가, "A, B, 및 C 등 중 적어도 하나"와 유사한 관례가 사용되는 경우에, 일반적으로 그러한 구조는 당업자가 그 관례를 이해하는 그러한 의미로 의도된다(예를 들어, "A, B, 및 C 등 중 적어도 하나를 갖는 시스템"은 A 단독, B 단독, C 단독, A와 B를 함께, A와 C를 함께, B와 C를 함께 및/또는 A, B 및 C를 함께 갖는 시스템을 포함하지만, 이에 제한되지 않을 것임). "A, B, 또는 C 등 중 적어도 하나"와 유사한 관례가 사용되는 그러한 경우, 일반적으로 그러한 구조는 당업자가 그 관례를 이해하는 의미이고자 한다(예를 들어, "A, B, 또는 C 등 중 적어도 하나를 갖는 시스템"은 A 단독, B 단독, C 단독, A와 B를 함께, A와 C를 함께, B와 C를 함께 및/또는 A, B 및 C를 함께 갖는 시스템을 포함하지만, 이에 제한되지 않을 것임). 실질적으로 임의의 선언적 단어 및/또는 2개 이상의 대안적 용어를 나타내는 구절은, 설명, 청구범위 또는 도면 어디에 위치하든, 그 용어들 중 하나, 용어들 중 어느 하나 또는 두 용어 모두를 포함하는 가능성을 고려하는 것으로 이해되어야 함이 추가로 이해되어야 할 것이다. 예를 들어, 구절 "A 또는 B"는 "A" 또는 "B" 또는 "A 및 B"의 가능성을 포함하는 것으로 이해될 것이다.

[0897] 추가적으로, 본 개시 내용의 특성 또는 양태가 마쿠쉬 그룹 방식으로 기재된 경우, 당업자는 본 개시 내용이 또한 그에 따라 그 마쿠쉬 그룹의 구성원 중 임의의 개별 구성원 또는 하위군에 관해서도 기재되는 것임을 이해할 것이다.

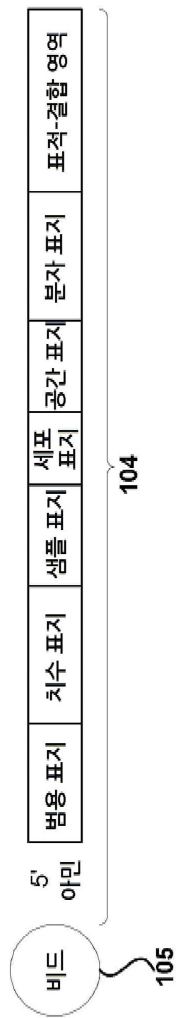
[0898] 당업자에 의해 이해될 바와 같이, 임의의 그리고 모든 목적을 위하여, 예컨대 기재된 설명의 제공과 관련하여, 본 명세서에 개시된 모든 범위는 또한, 임의의 및 모든 가능한 하위 범위 및 그의 하위 범위의 조합을 포괄한다. 임의의 열거된 범위는 적어도 균등하게 1/2, 1/3, 1/4, 1/5, 1/10 등으로 분해되는 동일 범위를 충분히 기재 및 가능하게 하는 것으로서 쉽게 인식될 수 있다. 비제한적인 예로서, 본 명세서에서 논의된 각각의 범위는 아래 1/3, 중간 1/3 및 윗부분 1/3로 용이하게 분해될 수 있다. 또한 당업자에 의해 이해될 바와 같이, 모든 언어, 예컨대 "이하", "적어도", "초과", "미만" 등은 언급된 수를 포함하고, 상기 논의된 것과 같이 이후에 하위 범위로 분해될 수 있는 범위를 지칭한다. 마지막으로, 당업자에 의해 이해될 바와 같이, 범위는 각각의 개별적인 구성원을 포함한다. 따라서, 예를 들어, 1 내지 3개의 물품을 갖는 군은 1, 2, 또는 3 물품을 갖는 군을 지칭한다. 유사하게, 1 내지 5개의 물품을 갖는 군은 1, 2, 3, 4, 또는 5개의 물품을 갖는 군 등을 지칭한다.

[0899] 다양한 양태 및 구현예가 본 명세서에 개시되어 있지만, 다른 양태 및 구현예가 당업자에게 명백할 것이다. 본 명세서에 개시된 다양한 양태 및 구현예는 예시 목적을 위한 것이고 제한적이고자 하지 않으며, 진정한 범주 및

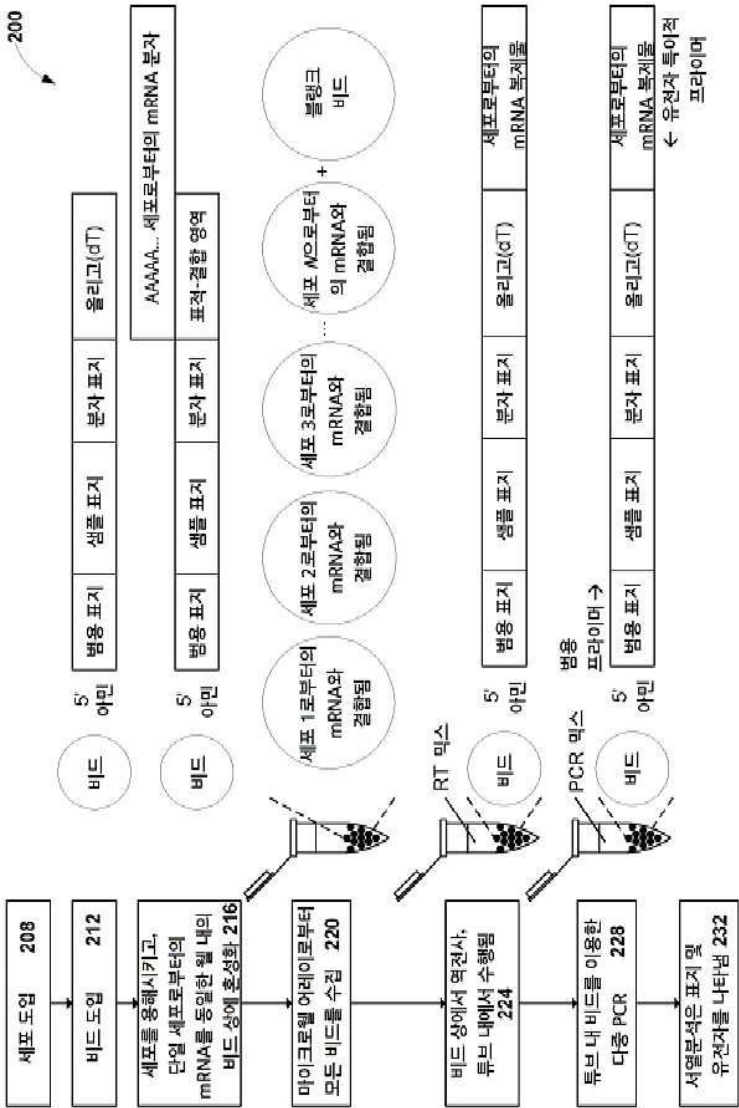
사상은 하기의 청구범위에 의해 지시되어 있다.

도면

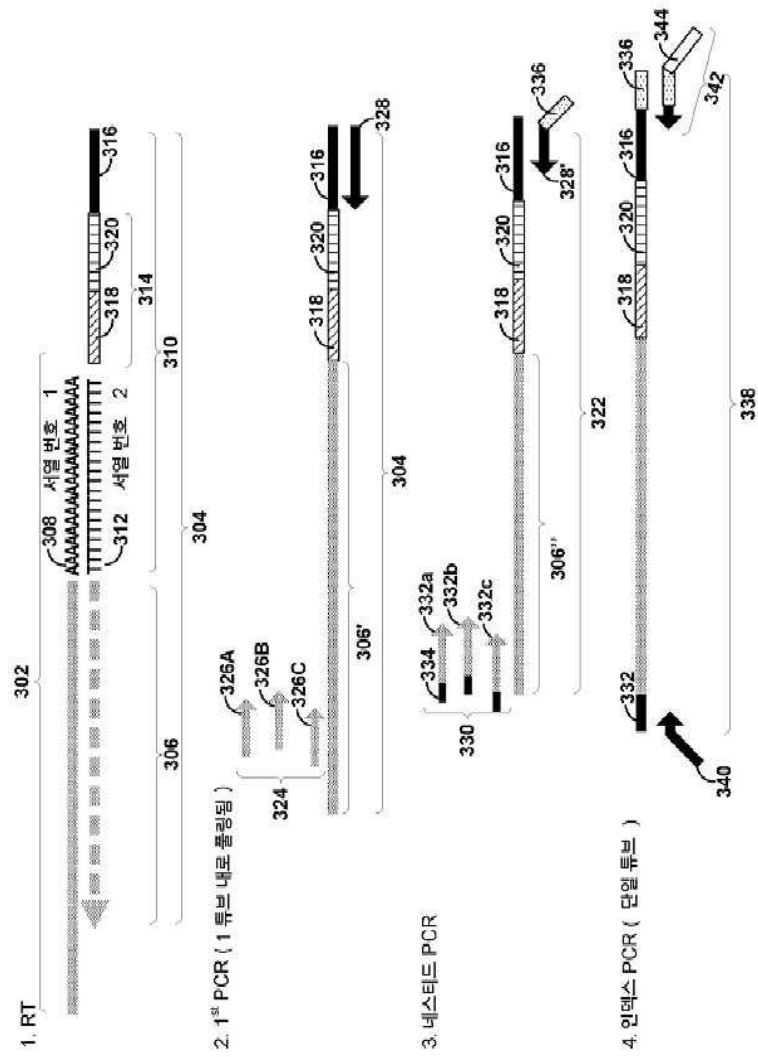
도면1



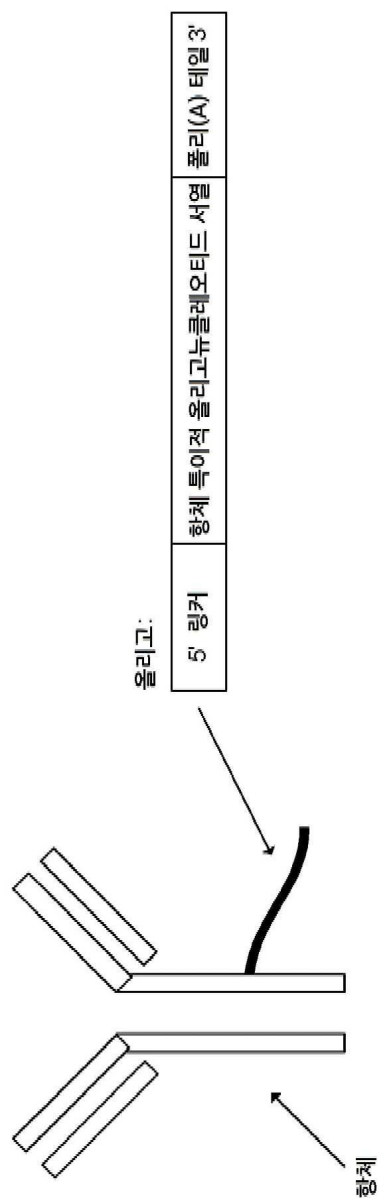
도면2



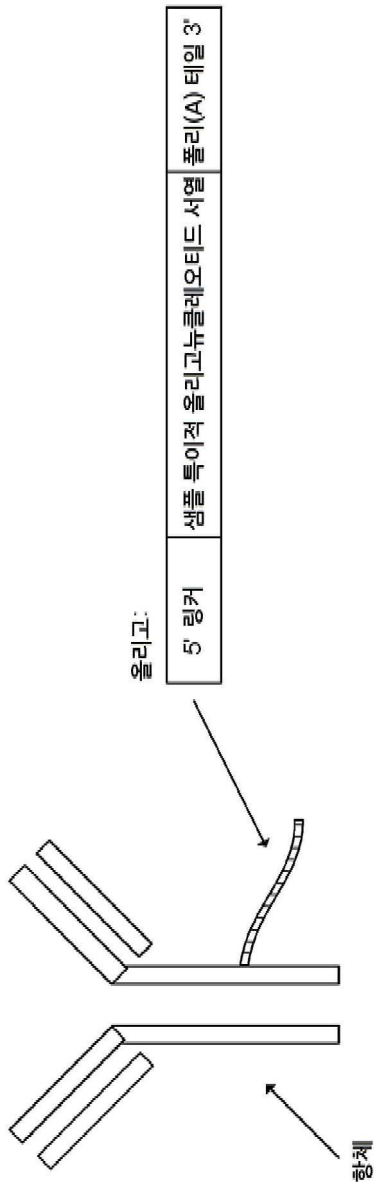
도면3



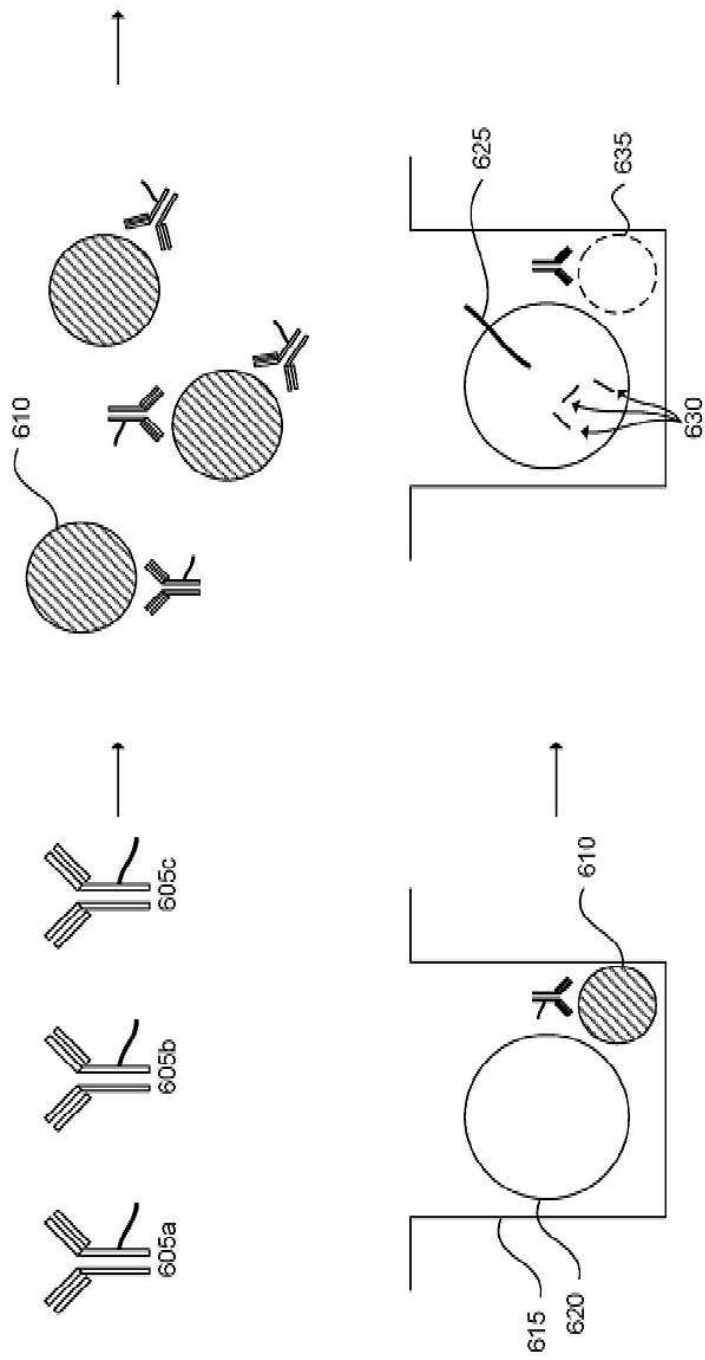
도면4



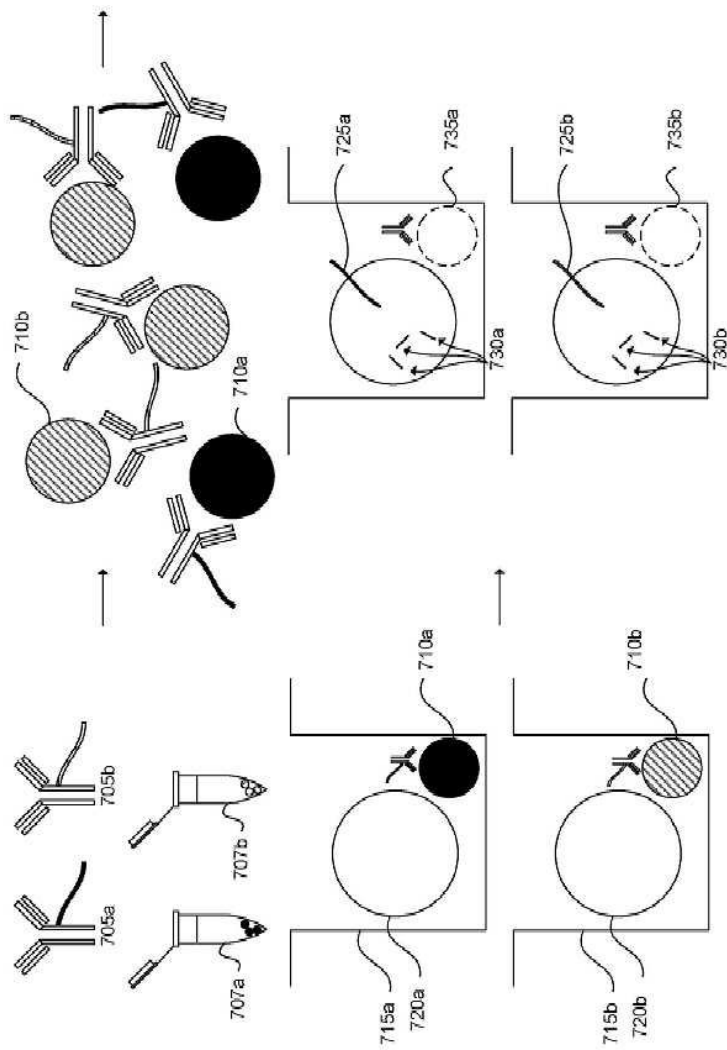
도면5



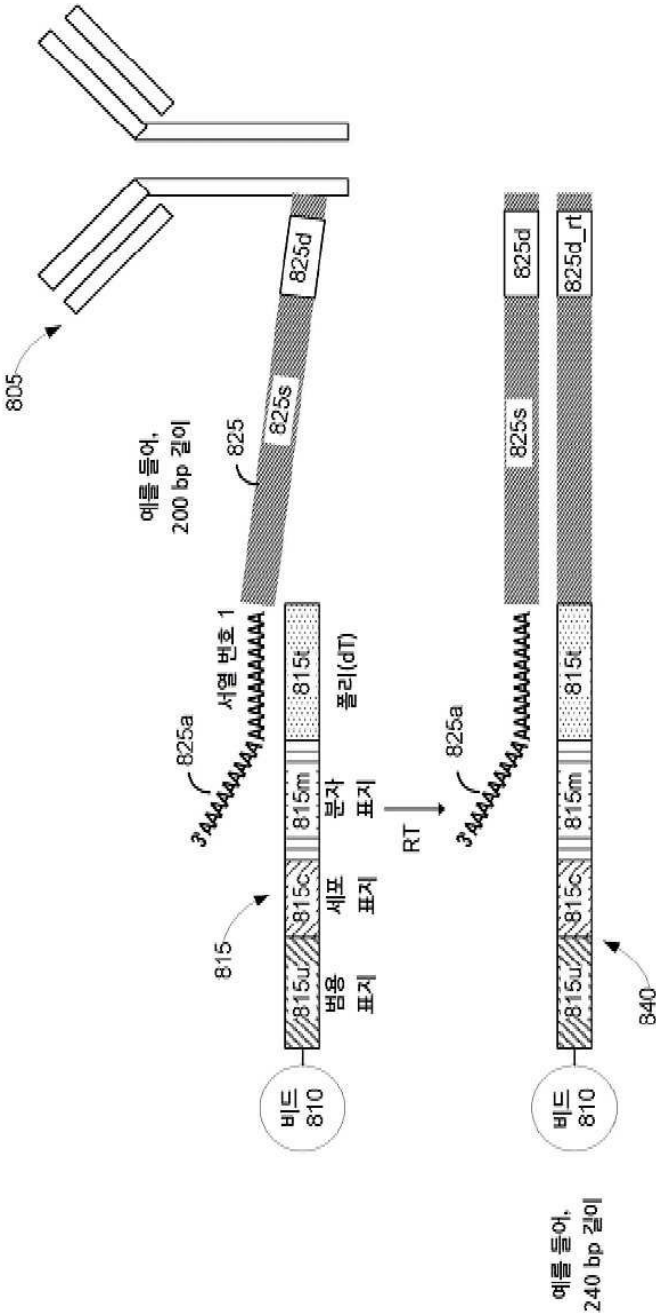
도면6



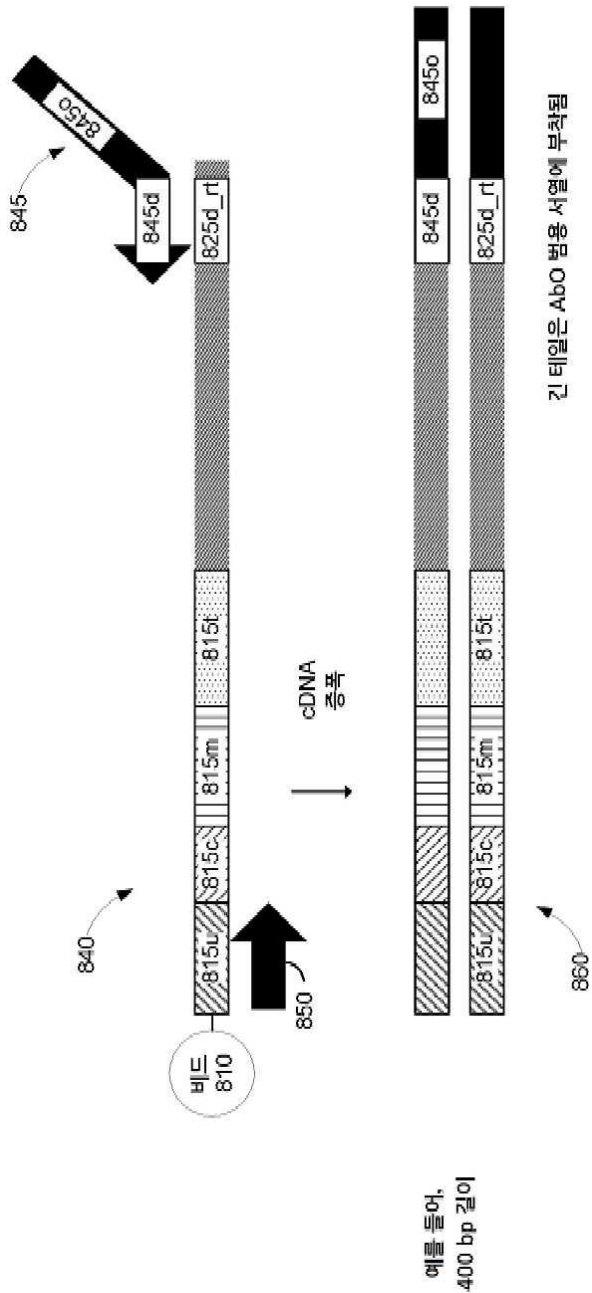
도면7



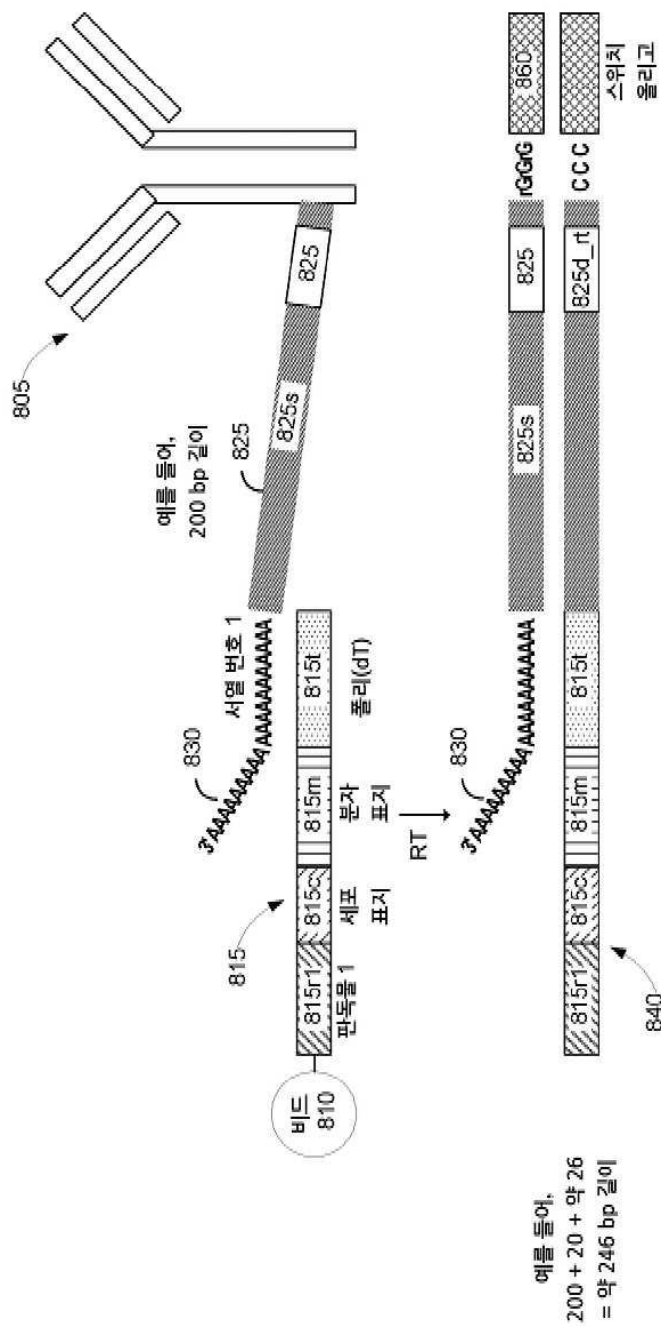
도면 8aa



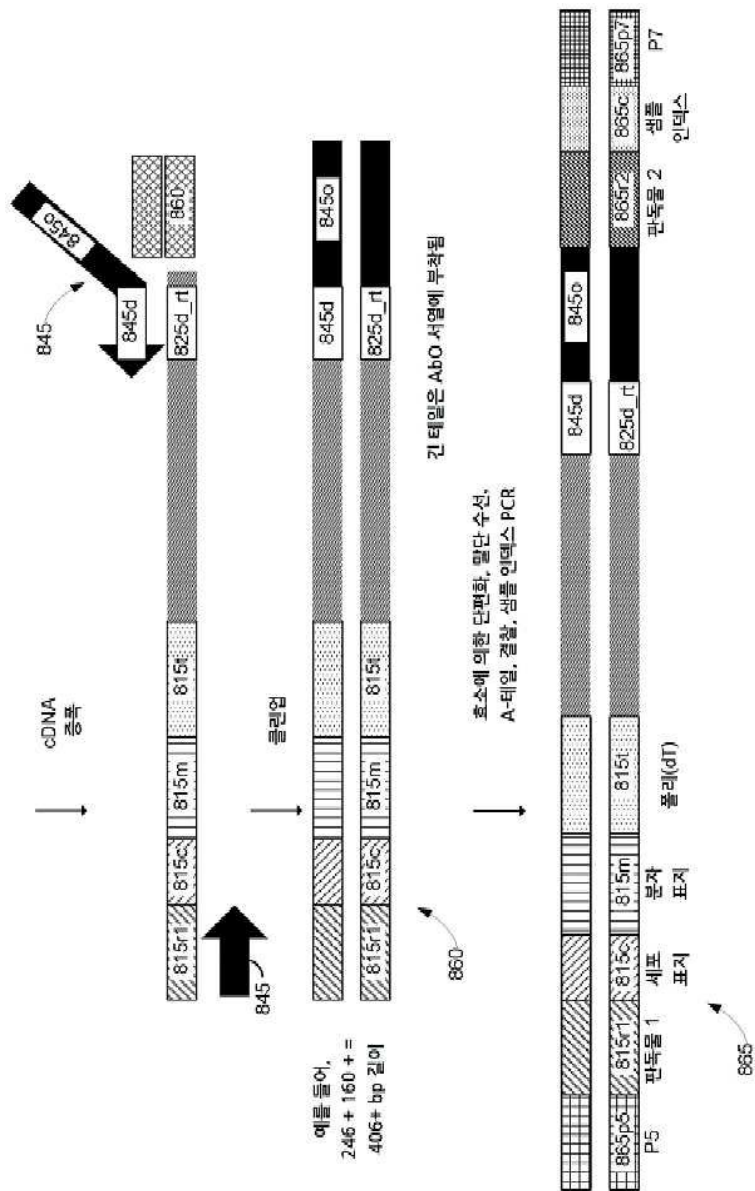
도면8ab



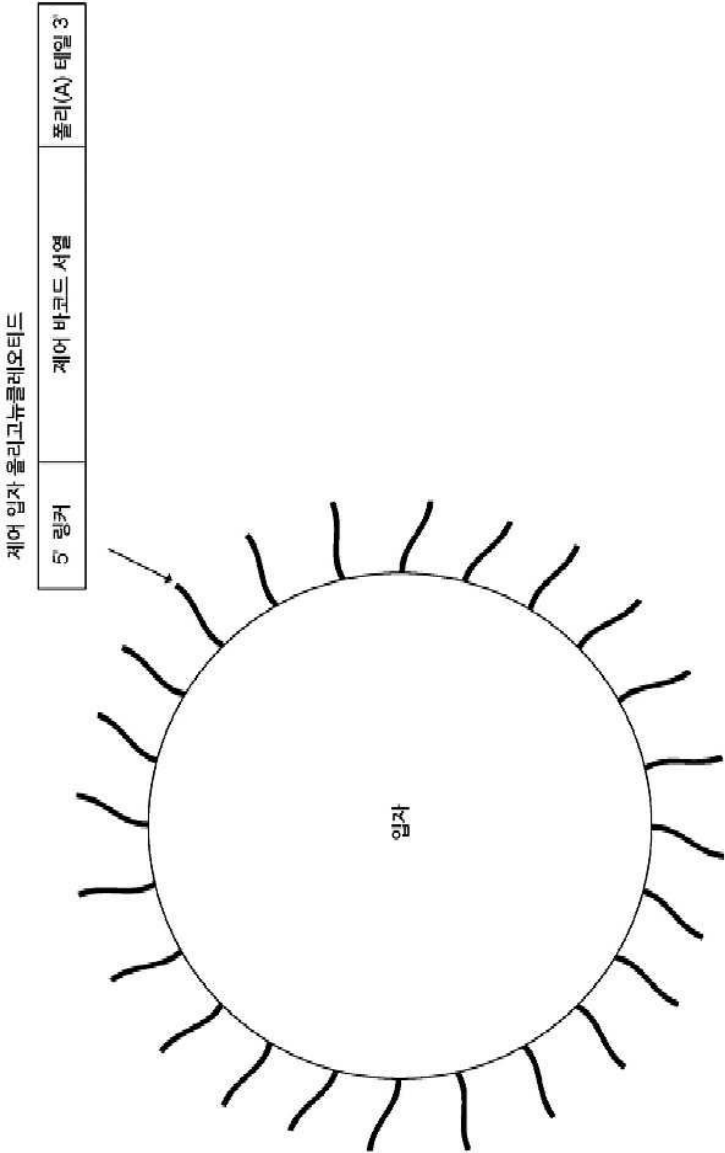
도면 8ba



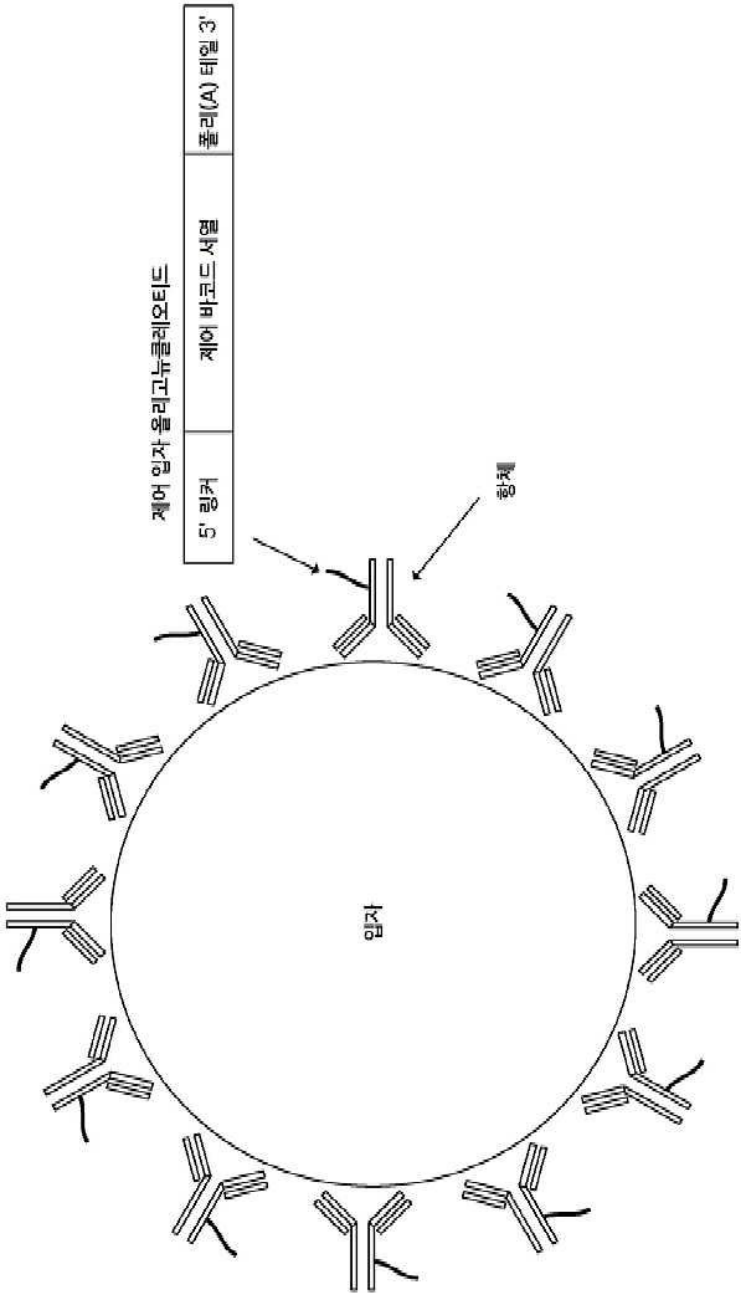
도면 8bb



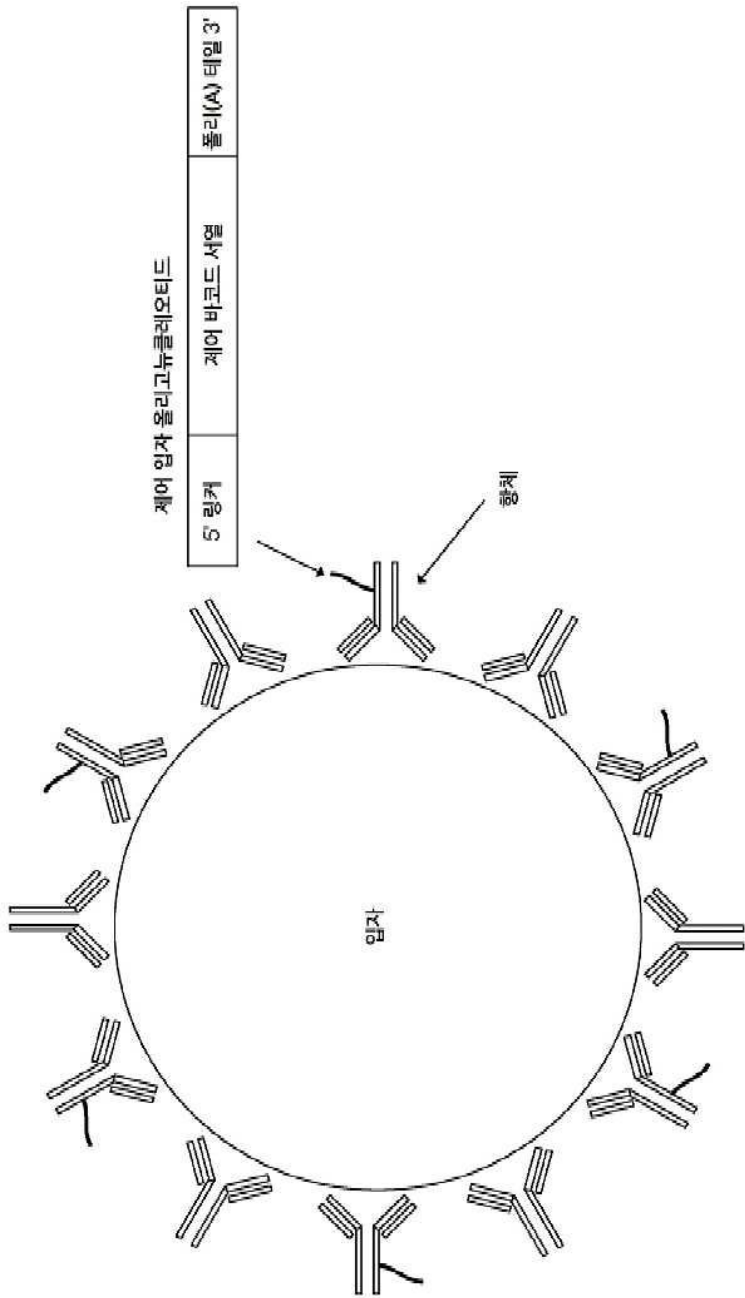
도면9a



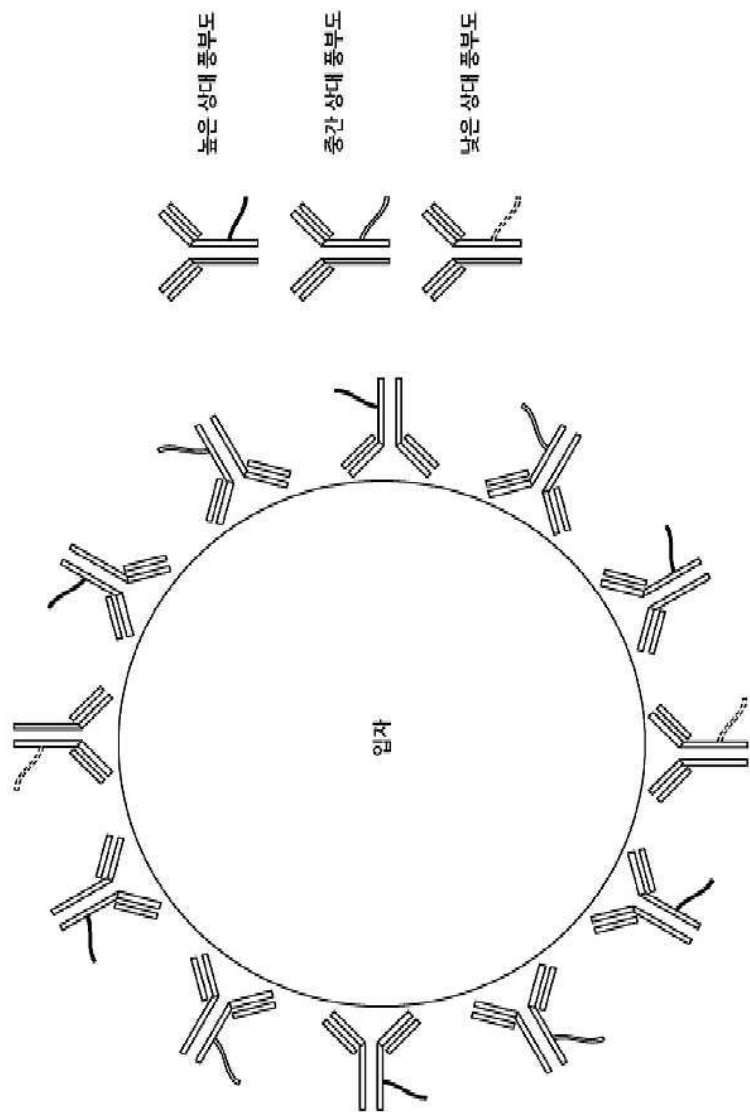
도면9b



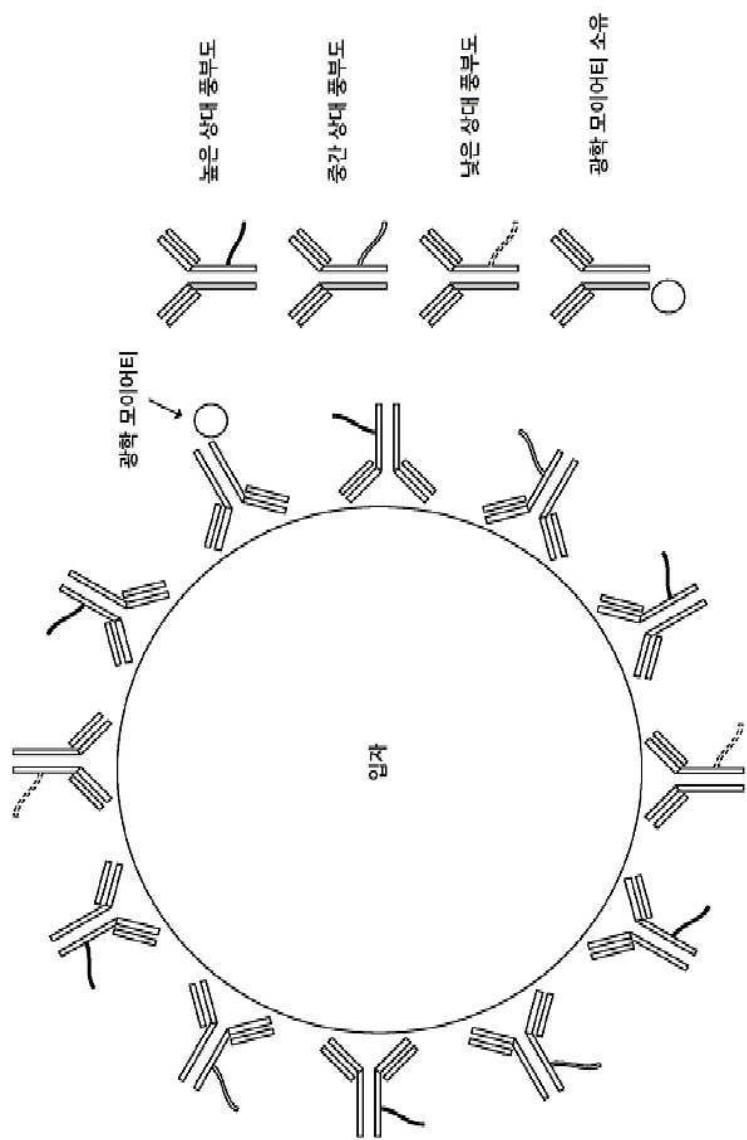
도면9c



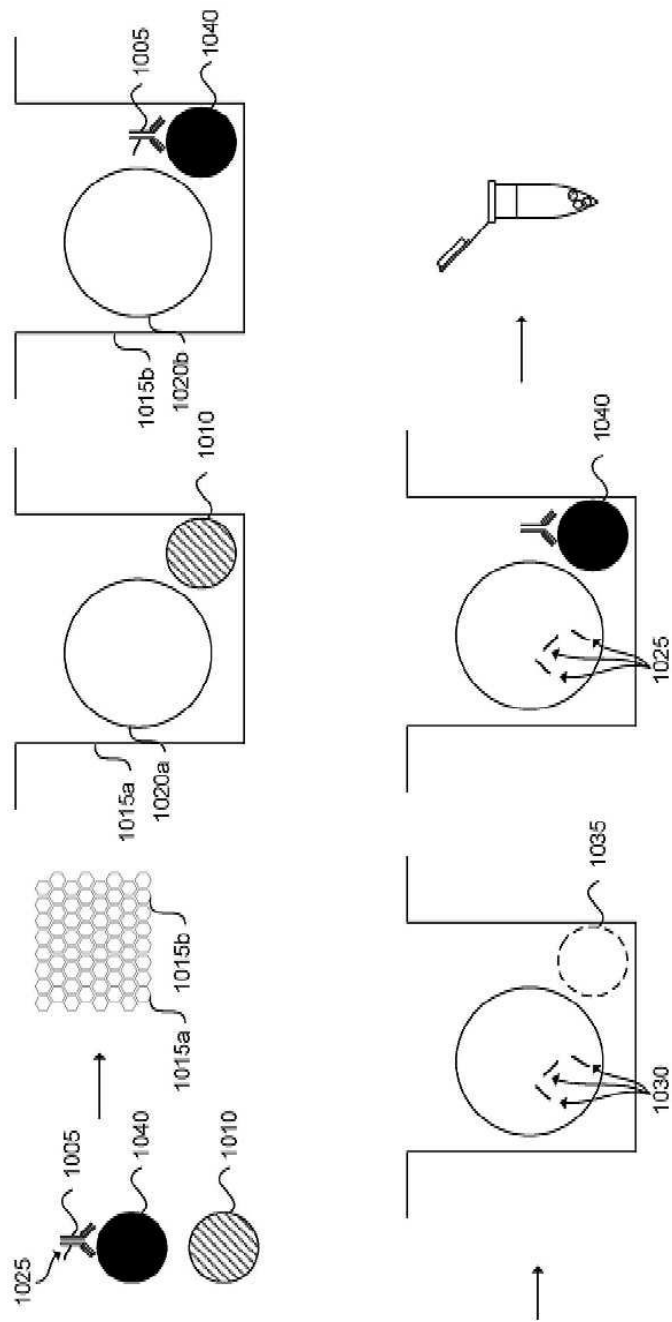
도면9d



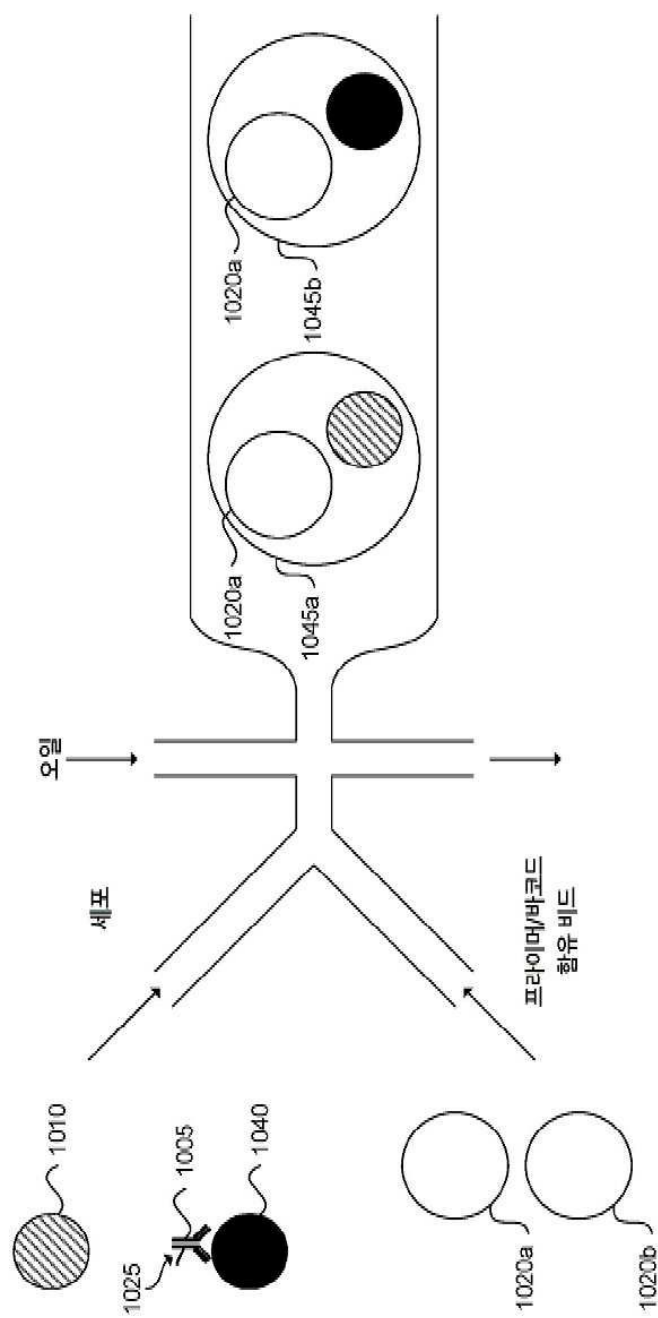
도면9e



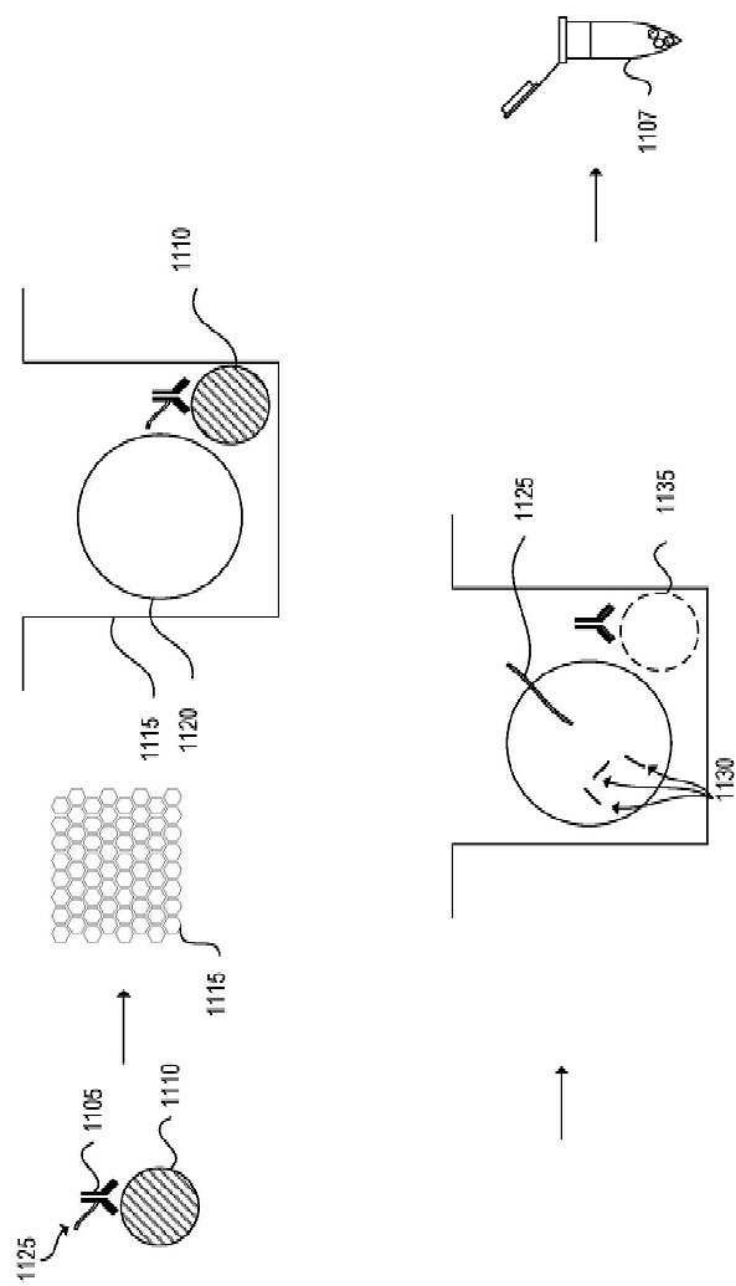
도면10a



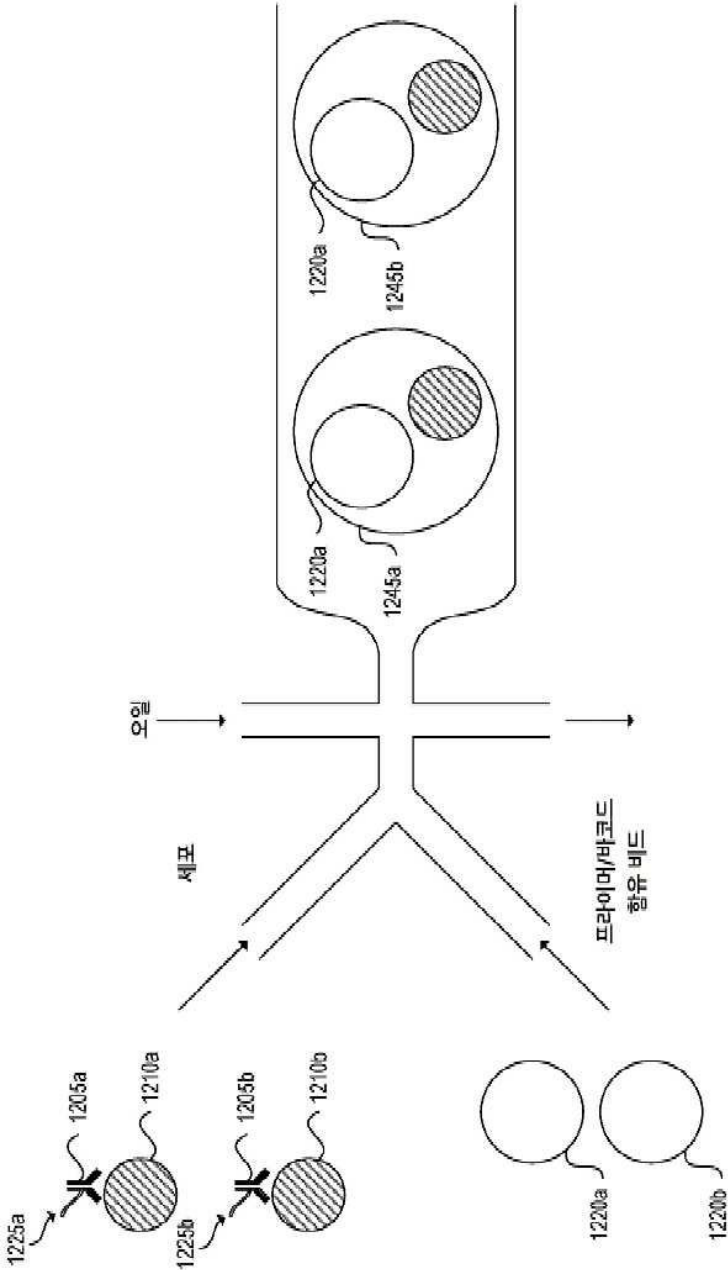
도면10b



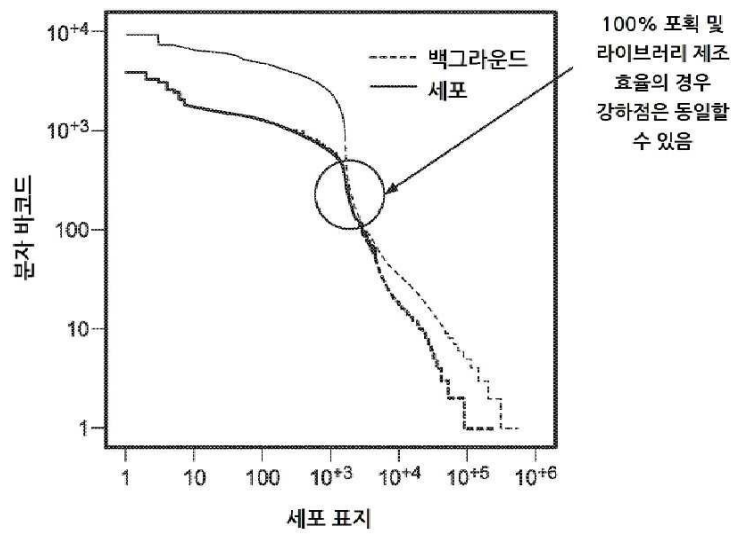
도면11



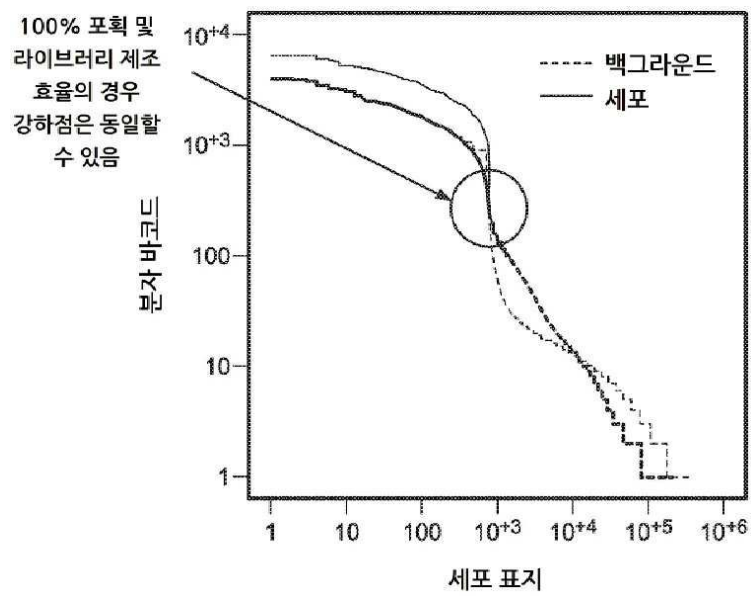
도면12



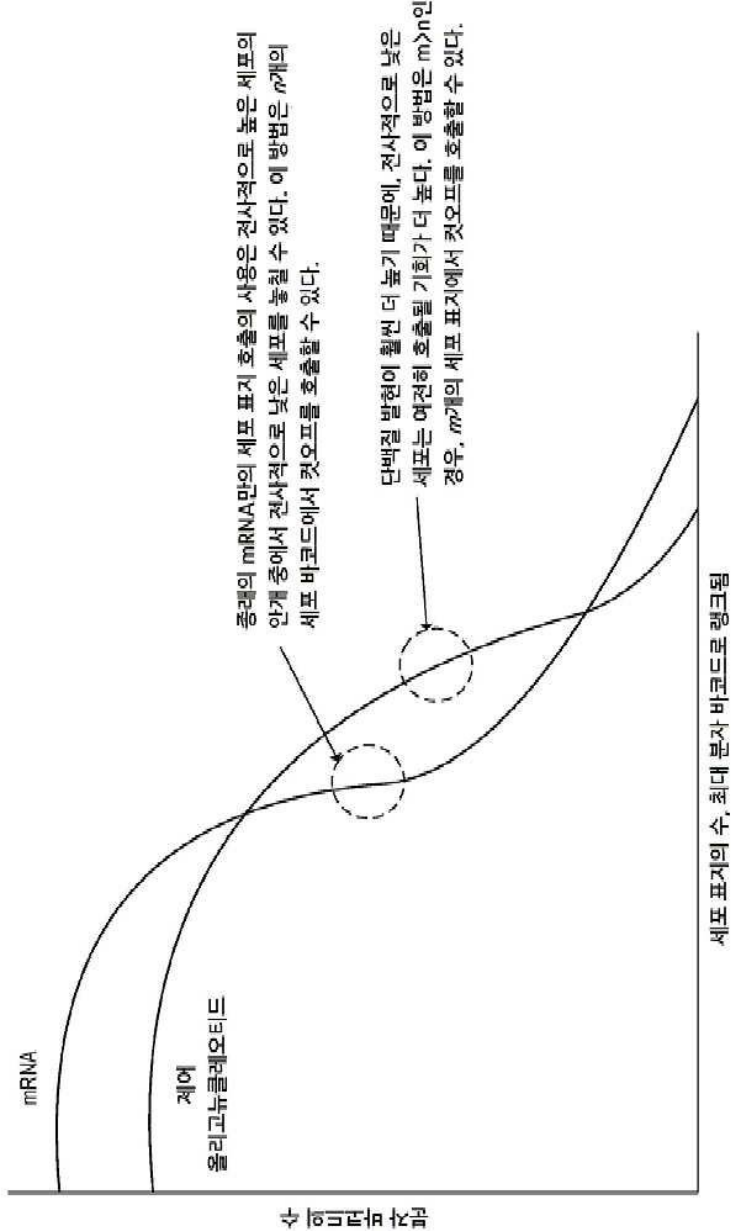
도면13a



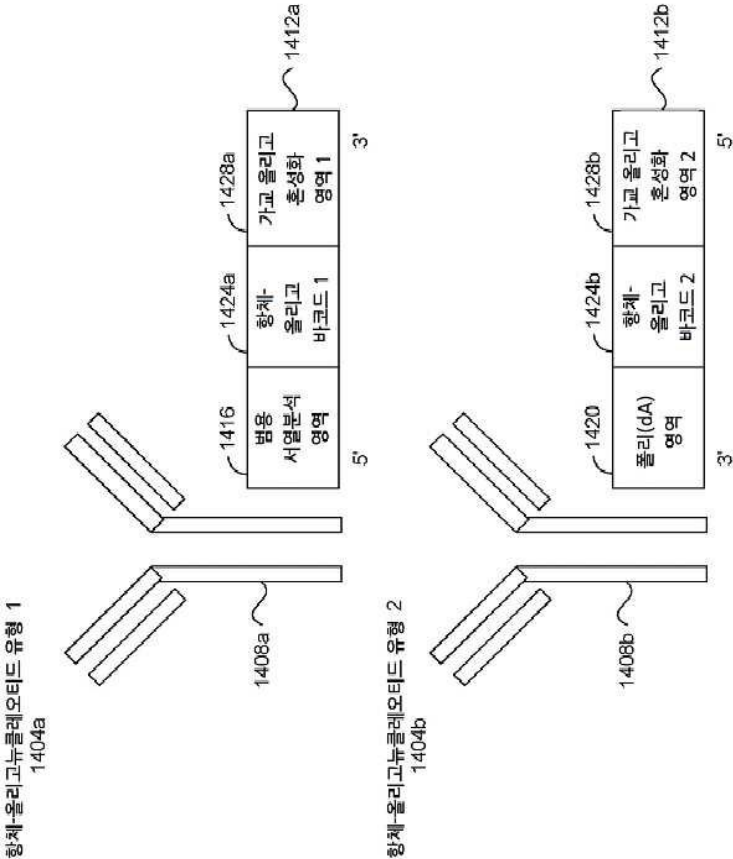
도면13b



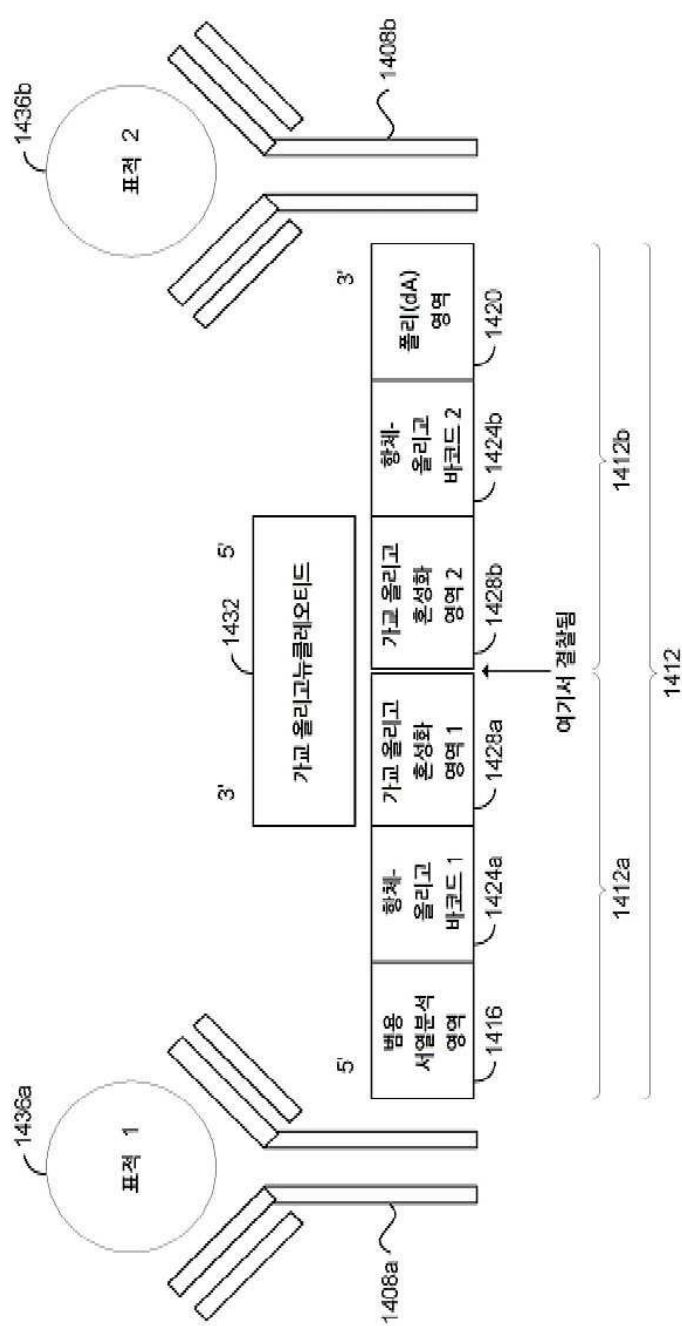
도면13c



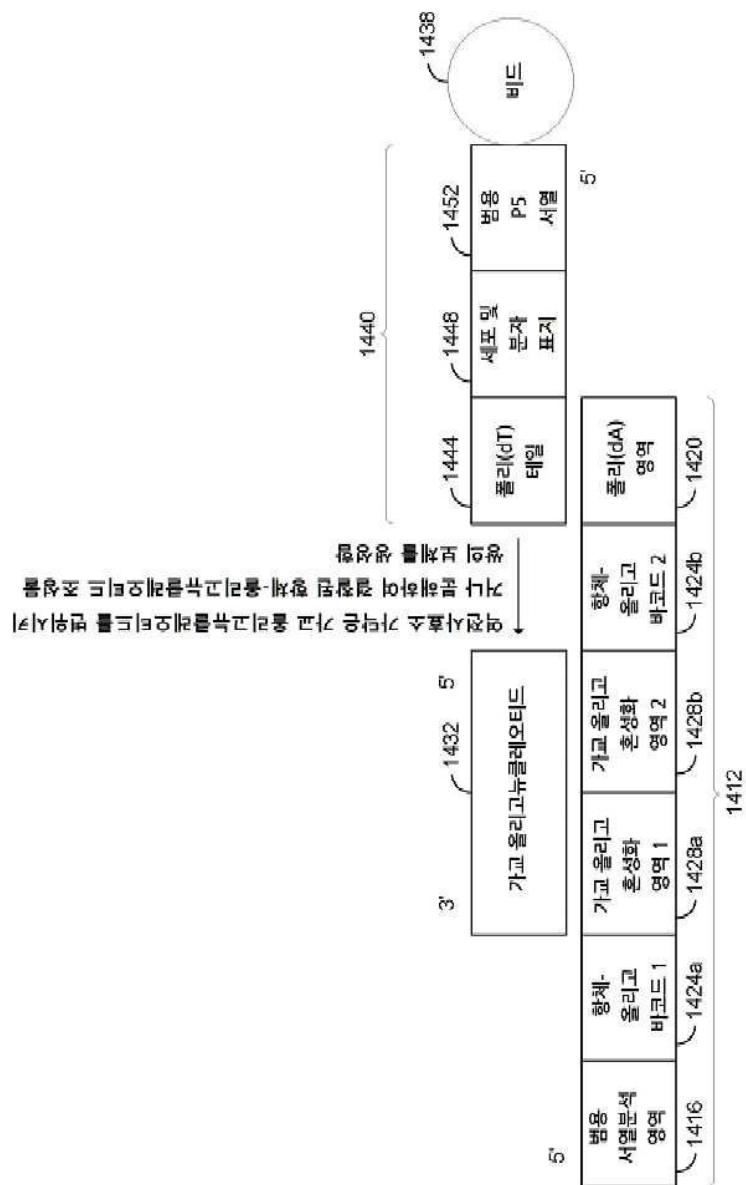
도면14a



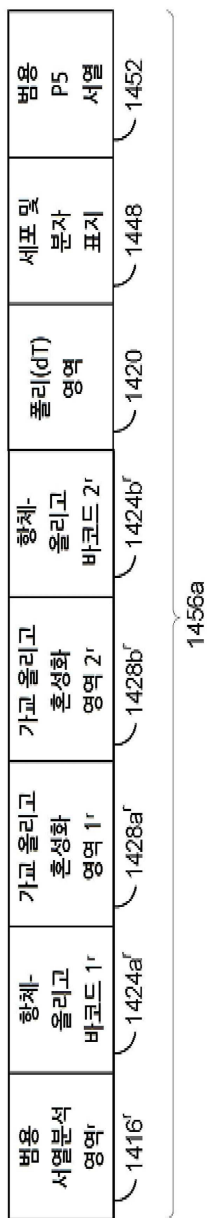
도면14b



도면 14c

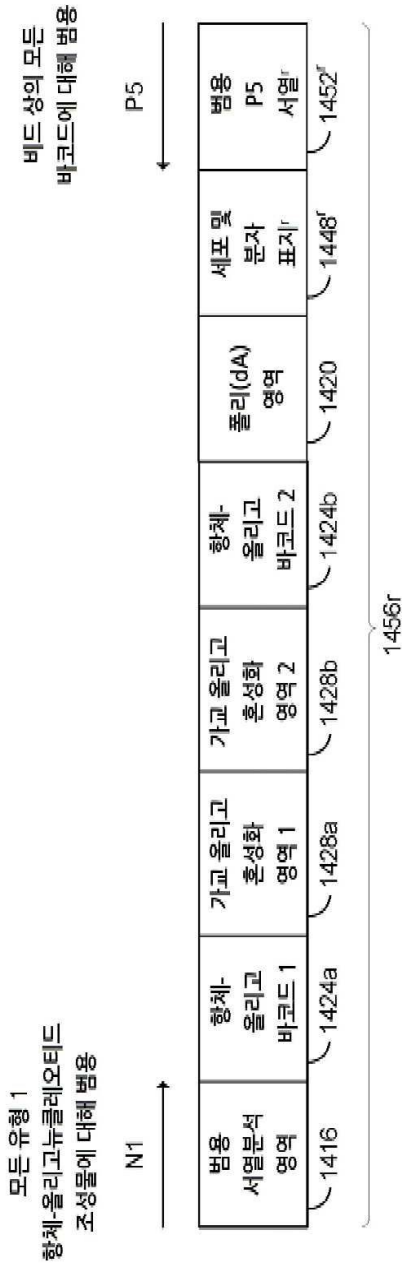


도면14d

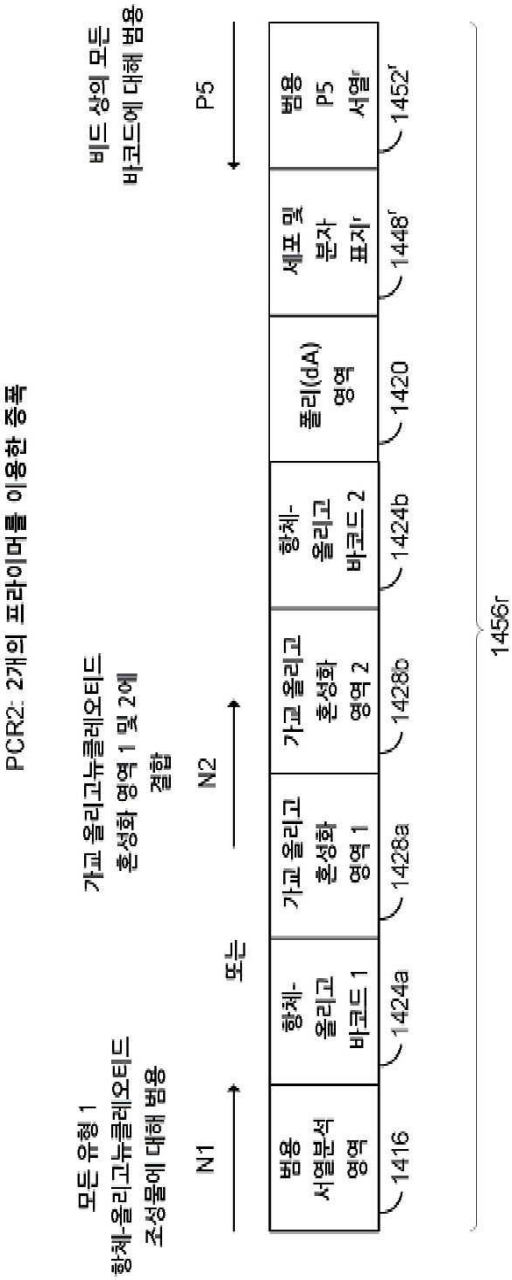


도면14e

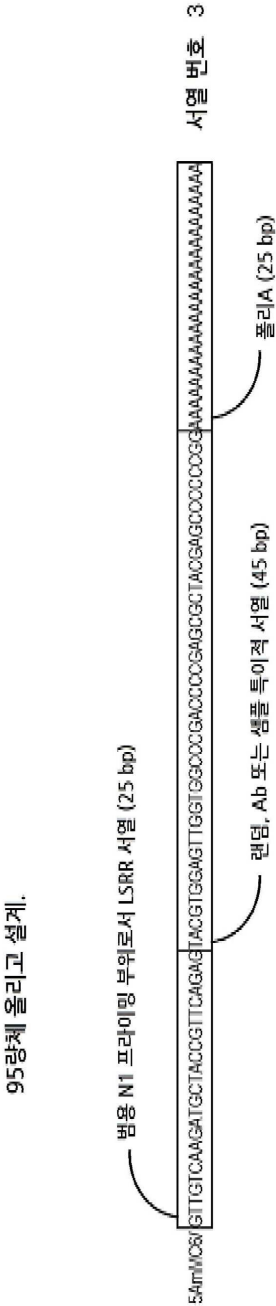
PCR1: 2개의 프라이머를 이용한 증폭



도면14f

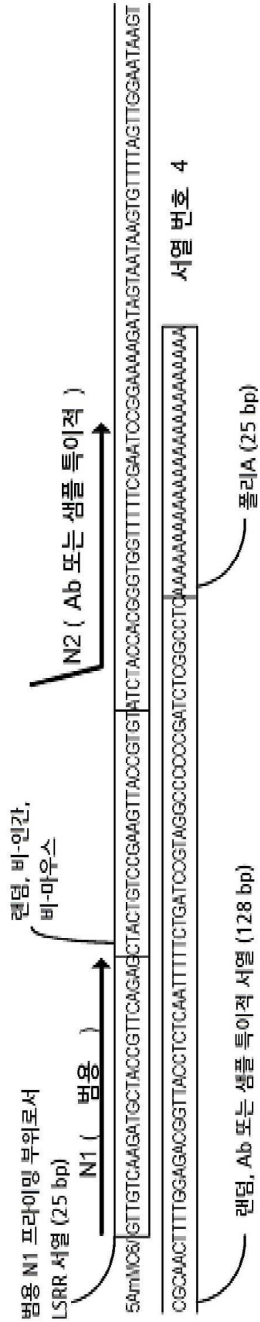


도면15a



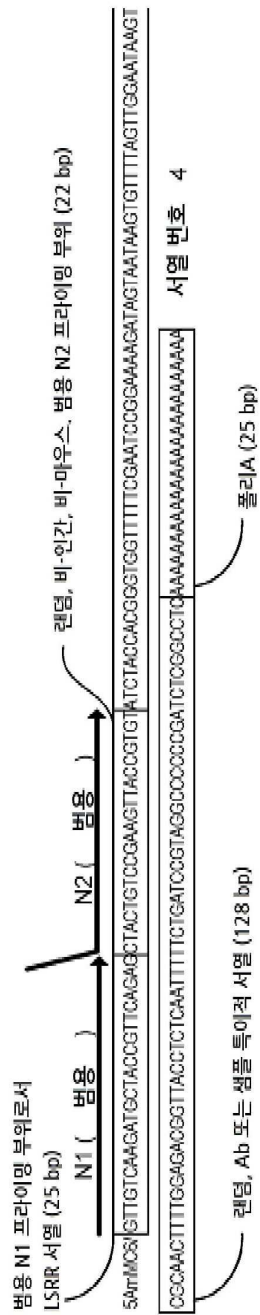
도면15b

200량체 올리고 설계. 표적화된 증폭 도식(네스티드 특이적)



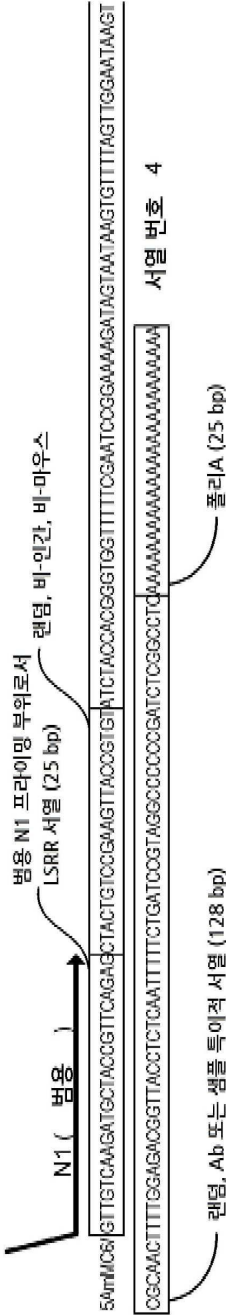
도면 15c

200량체 올리고 설탕계, 표적화된 증폭 도식(네스티드 벡터)

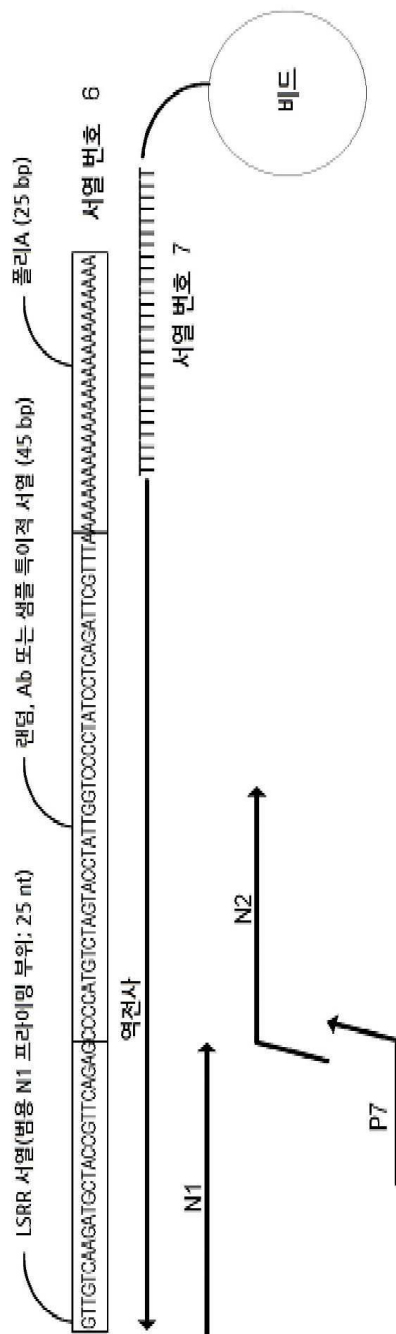


도면15d

200량체 올리고 설계. 표적화된 증폭 도식(1 단계)

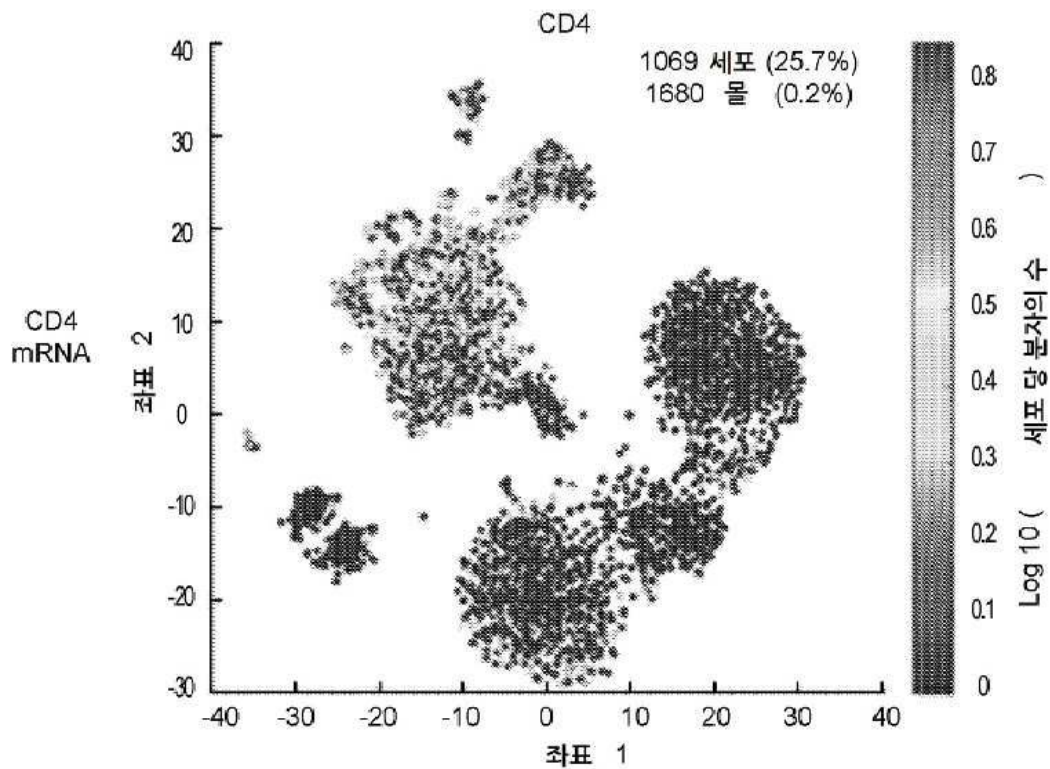


도면 16



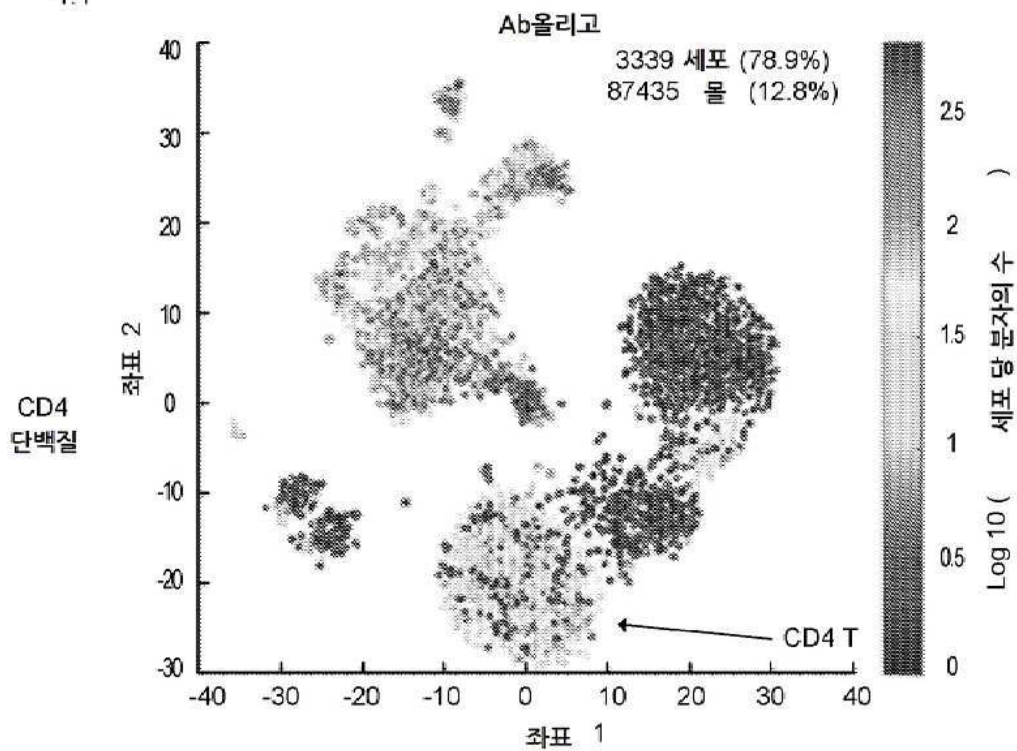
도면17a

1:1

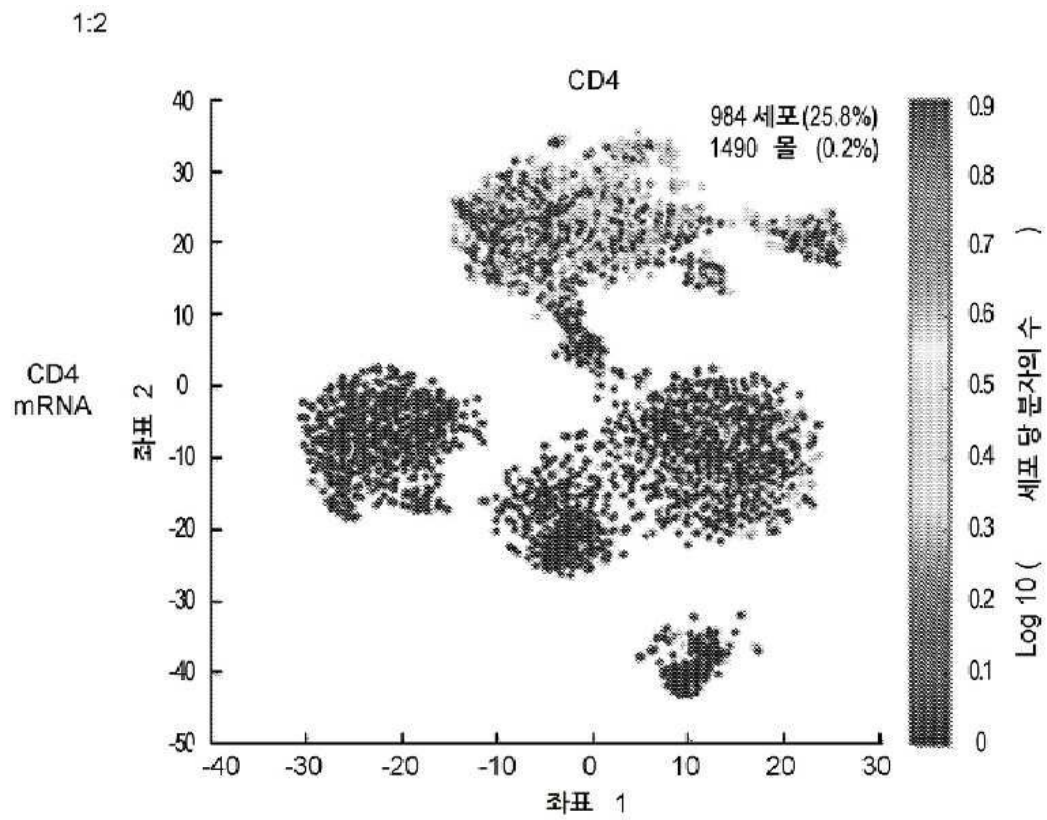


도면17b

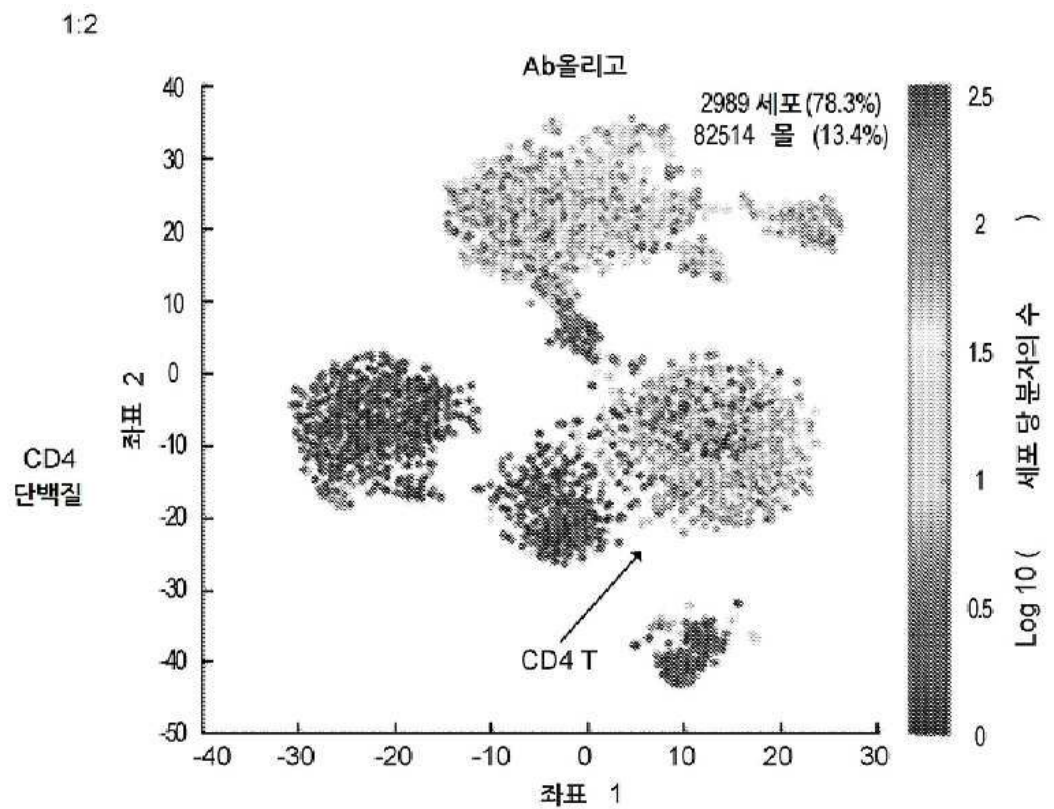
1:1



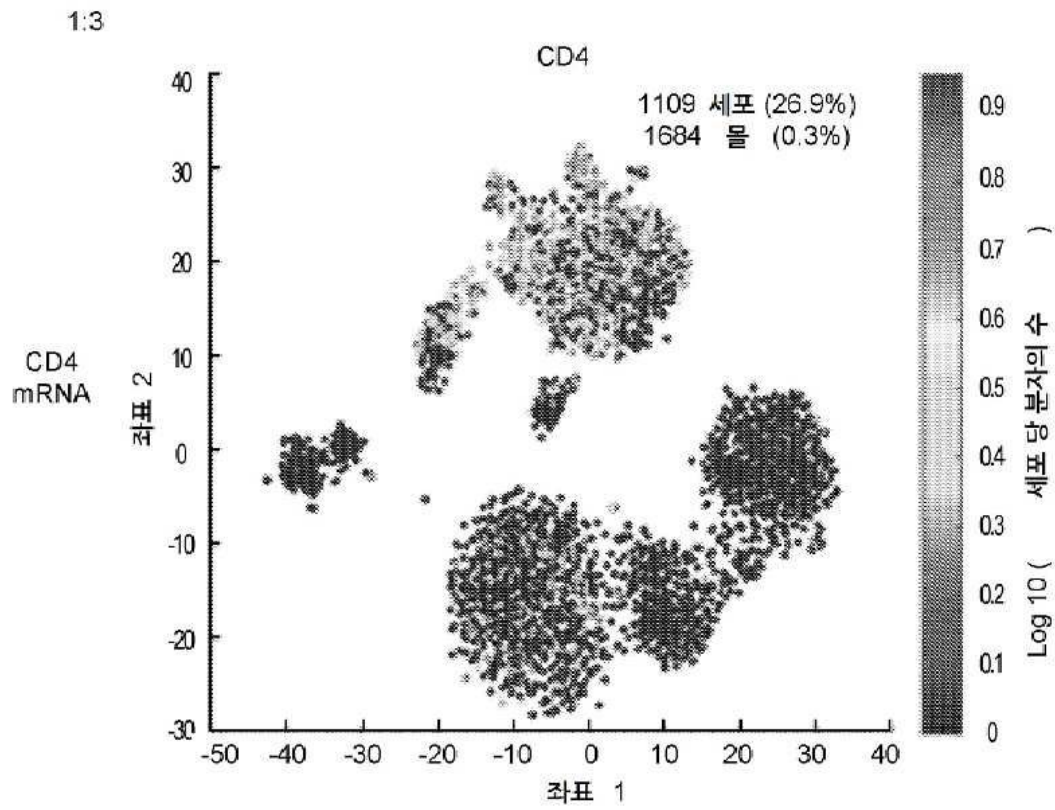
도면17c



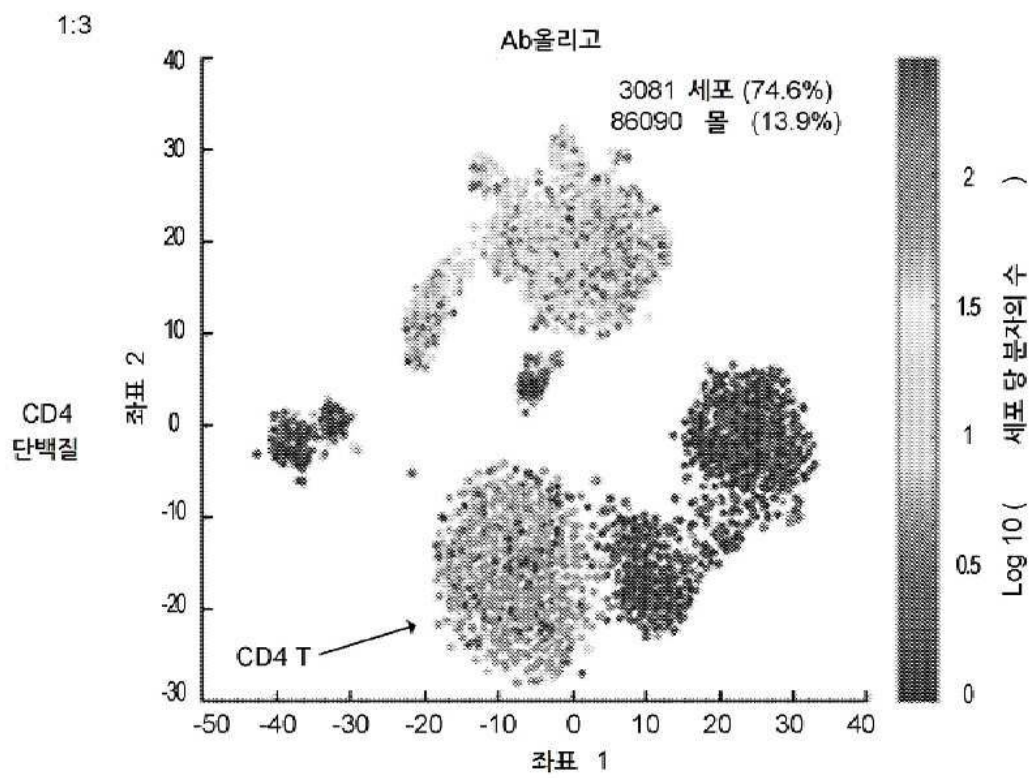
도면17d



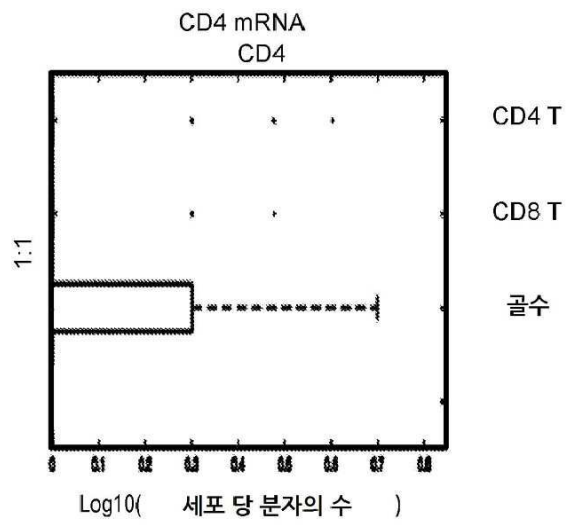
도면17e



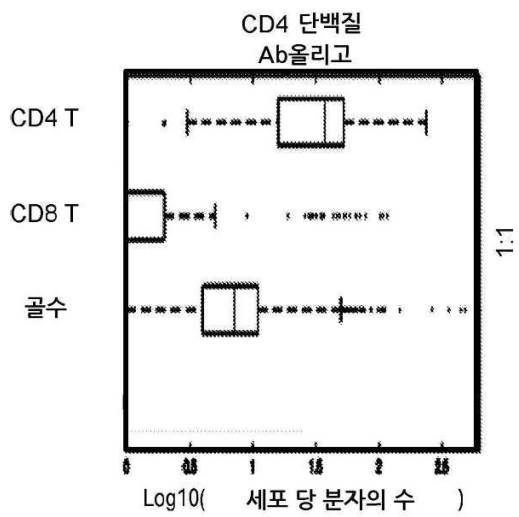
도면17f



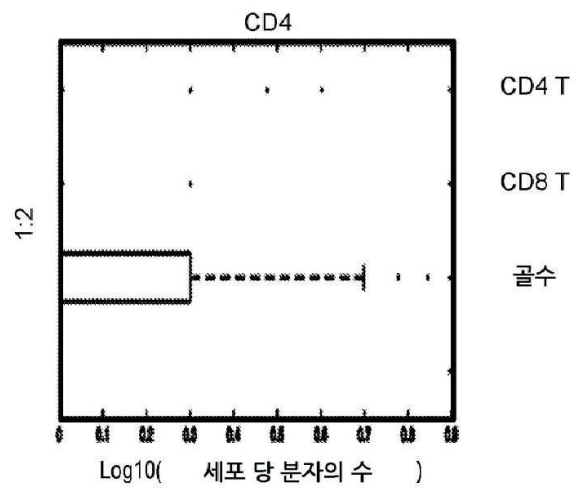
도면18a



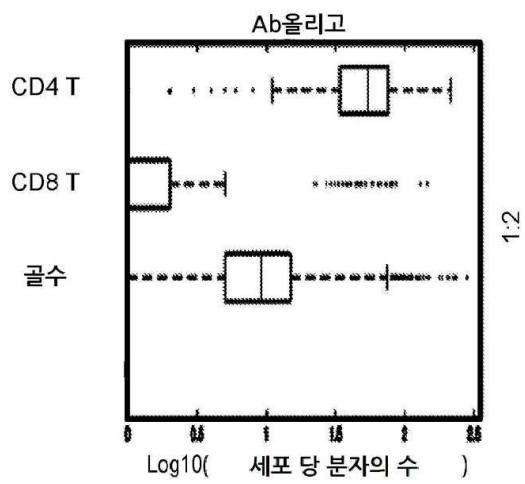
도면18b



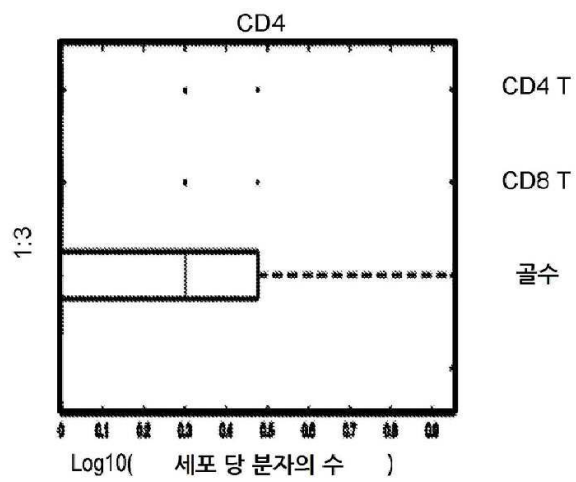
도면18c



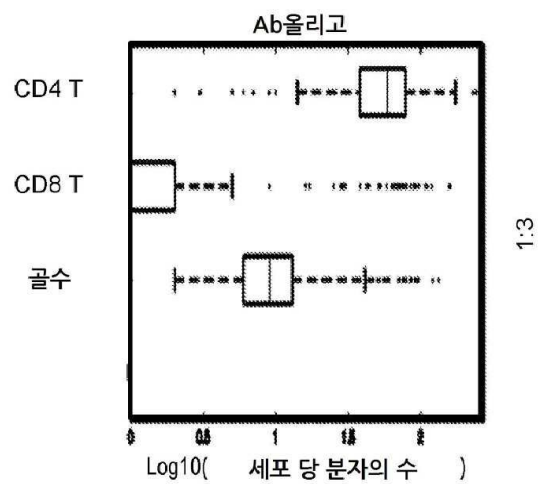
도면18d



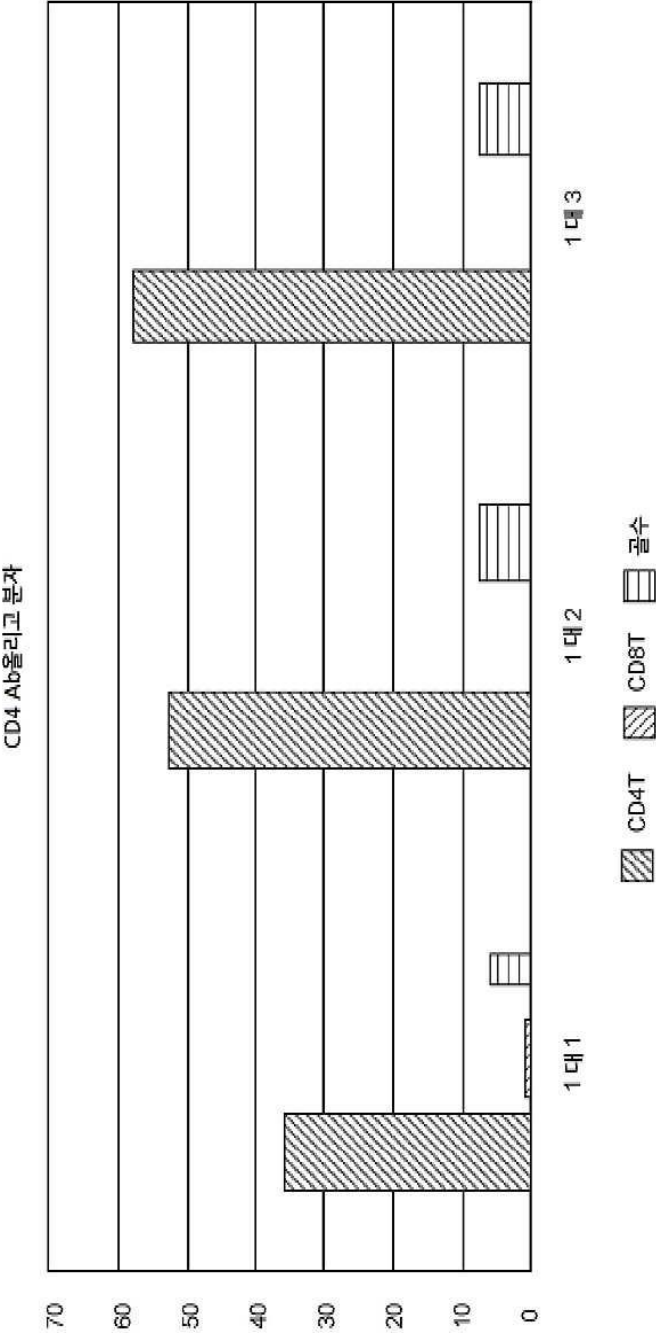
도면18e



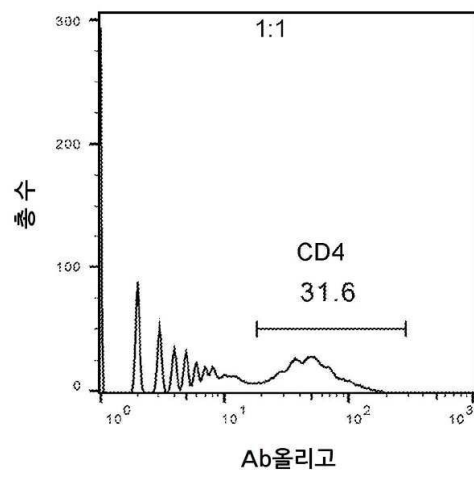
도면18f



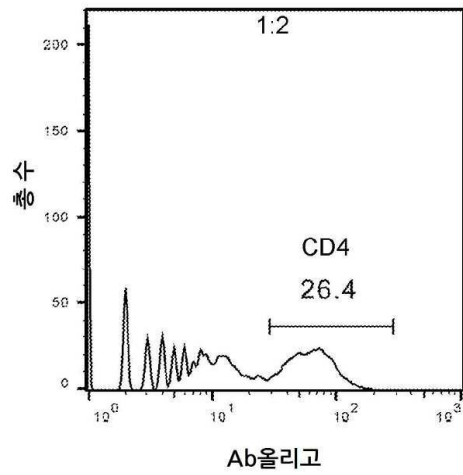
도면19



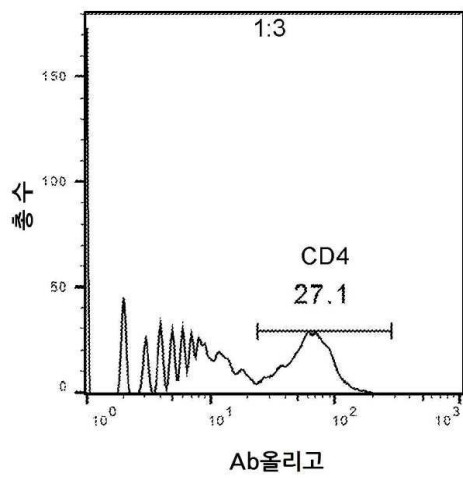
도면20a



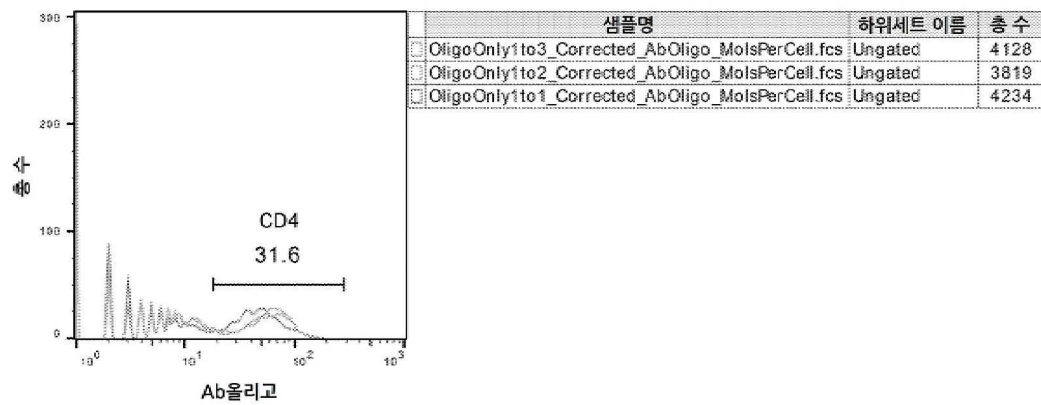
도면20b



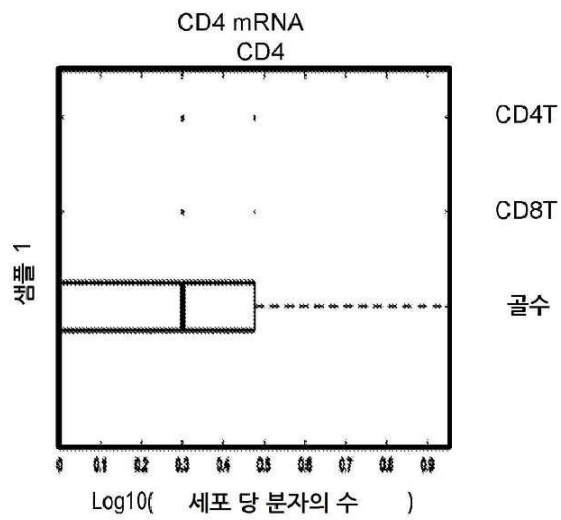
도면20c



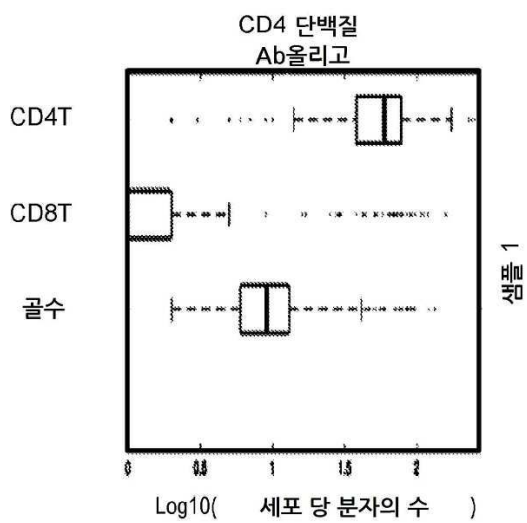
도면20d



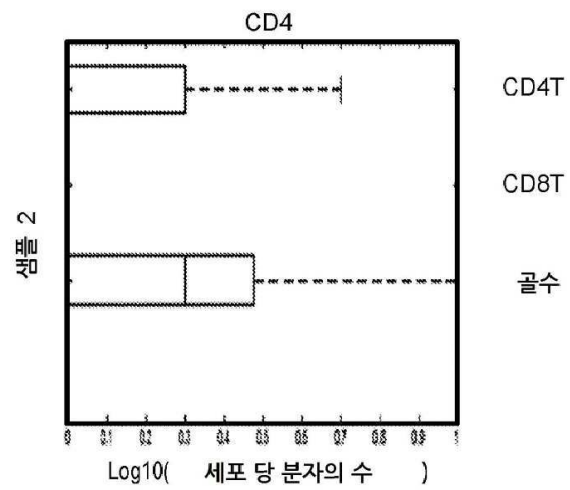
도면21a



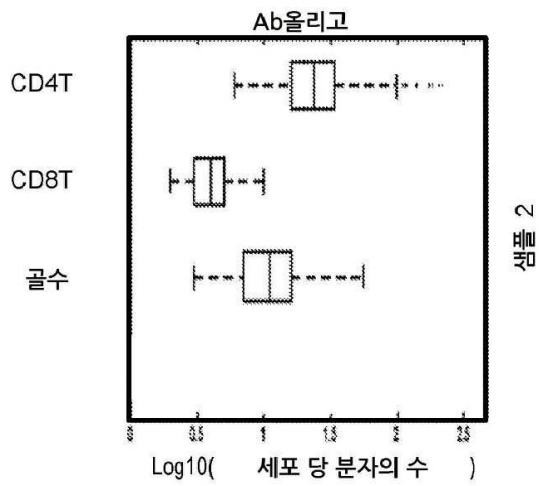
도면21b



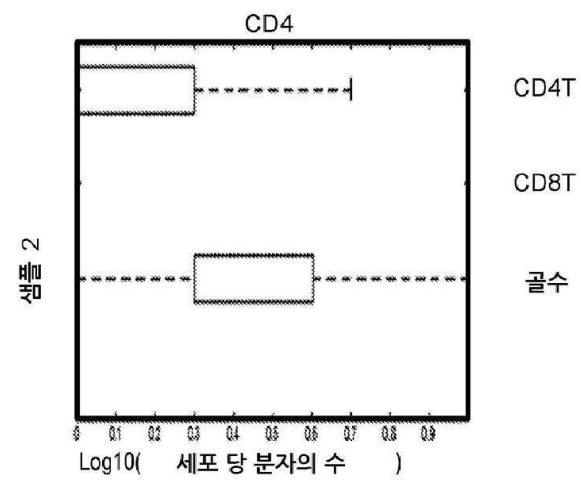
도면21c



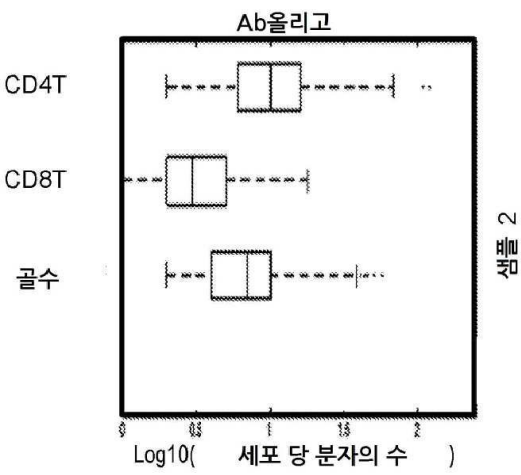
도면21d



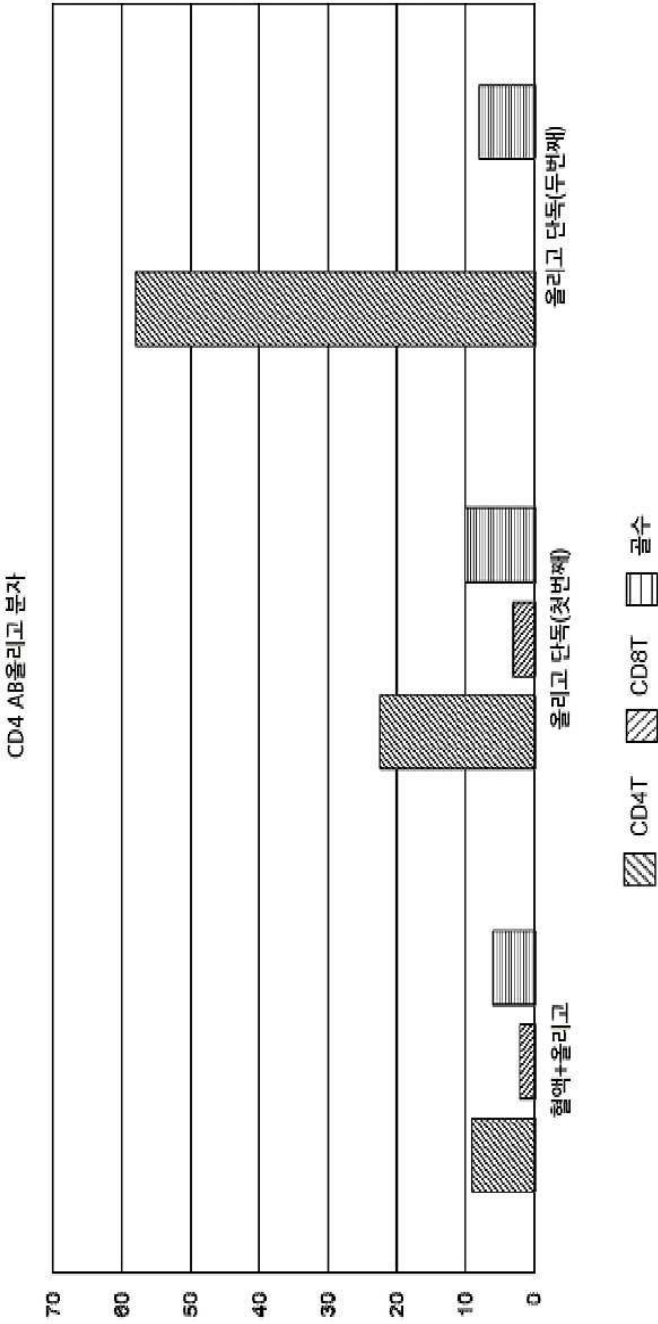
도면21e



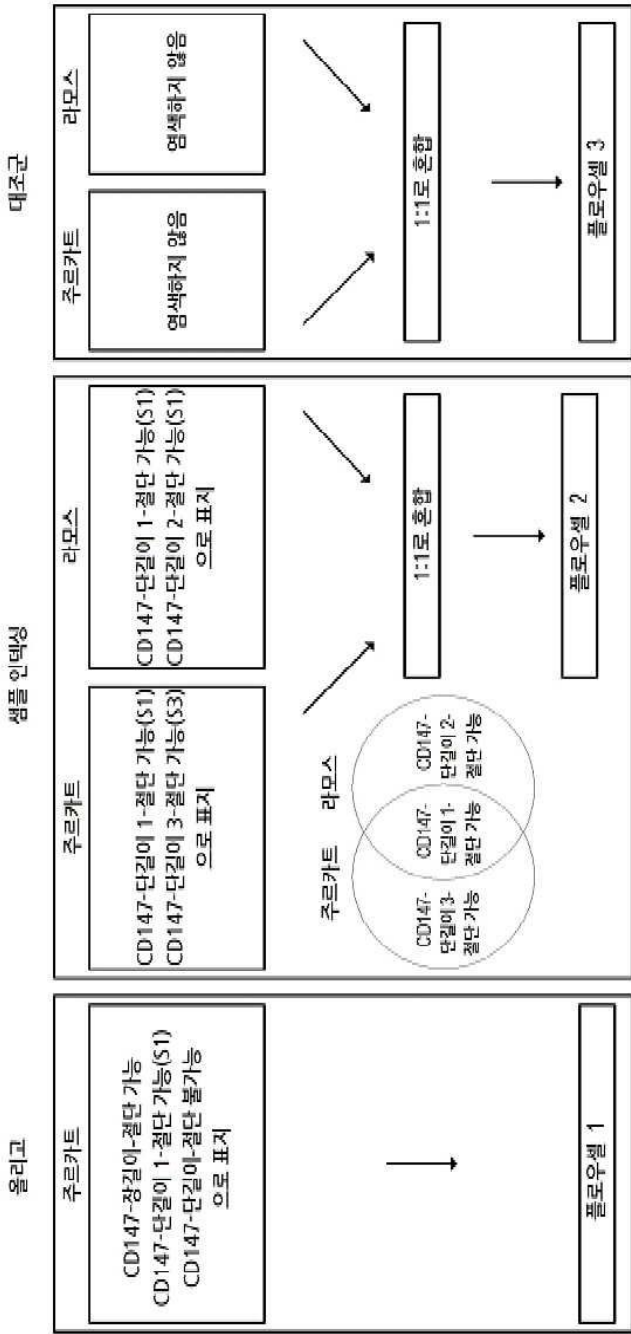
도면21f



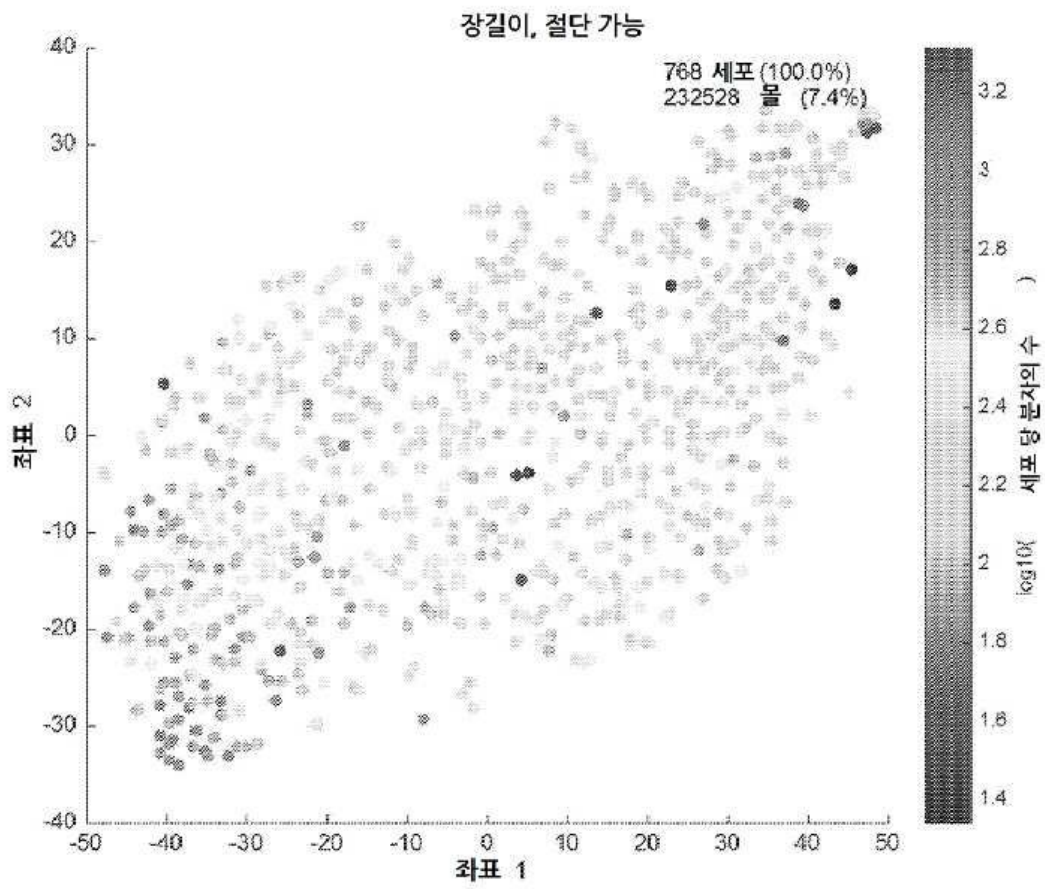
도면22



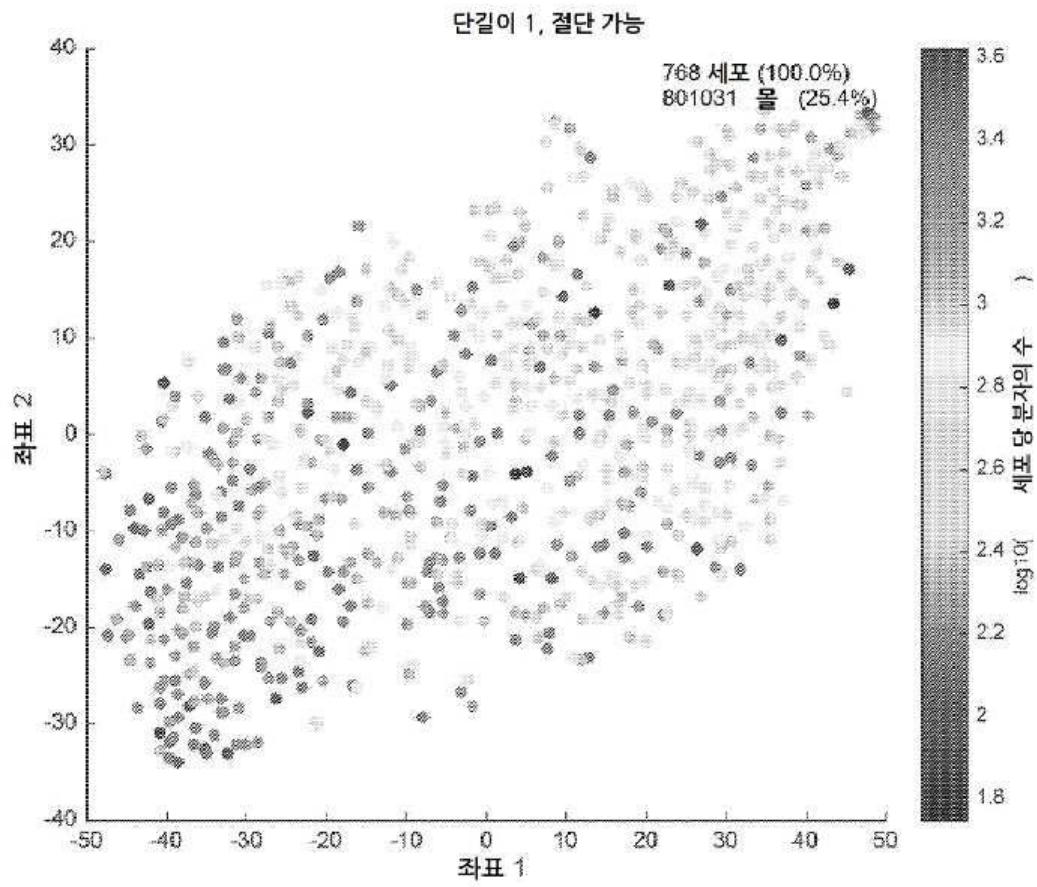
도면23



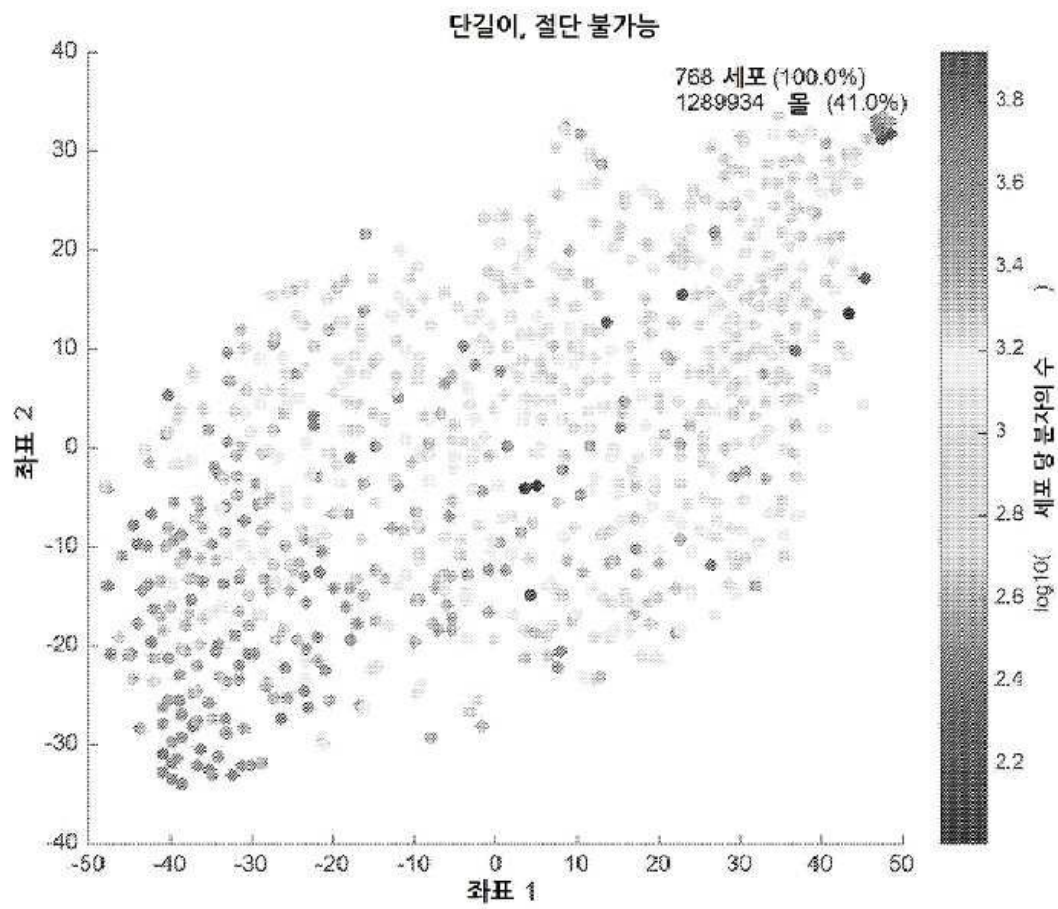
도면24a



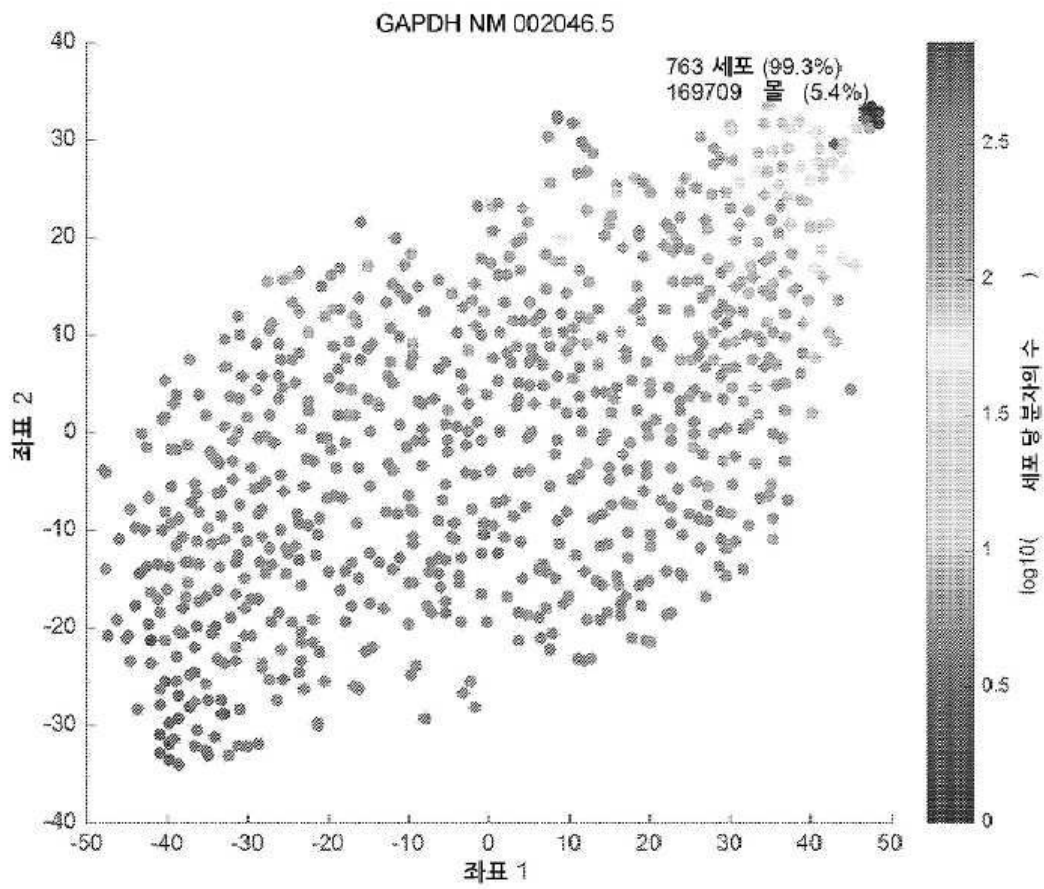
도면24b



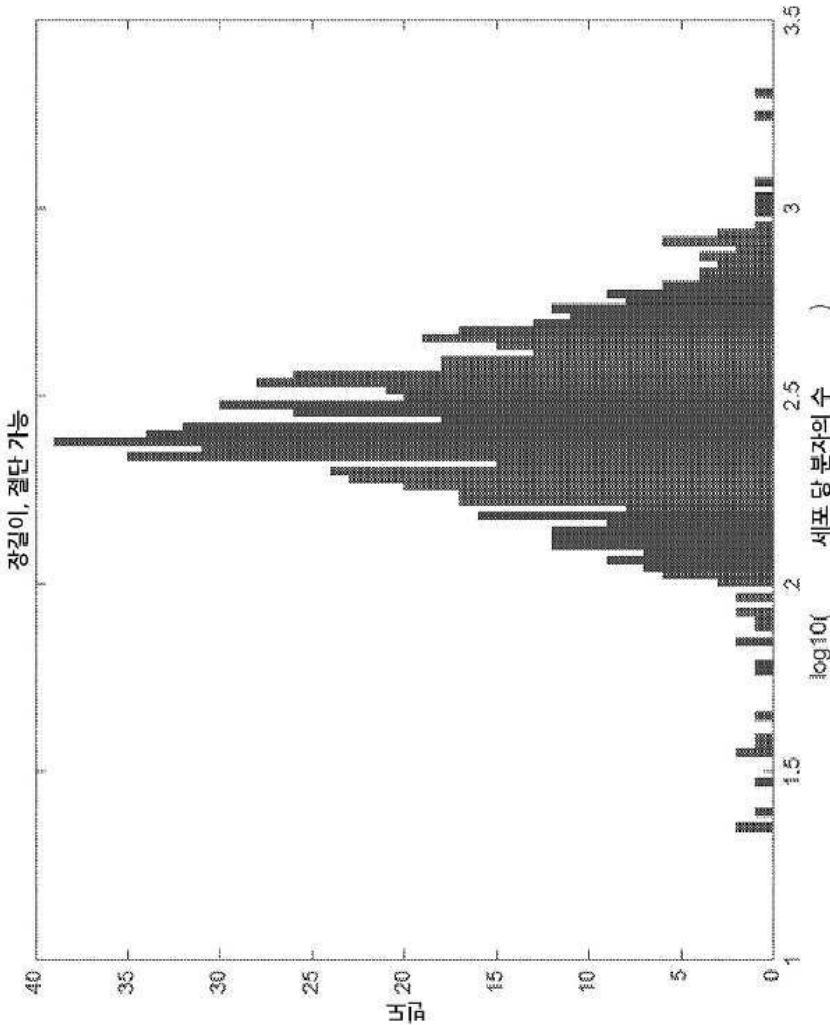
도면24c



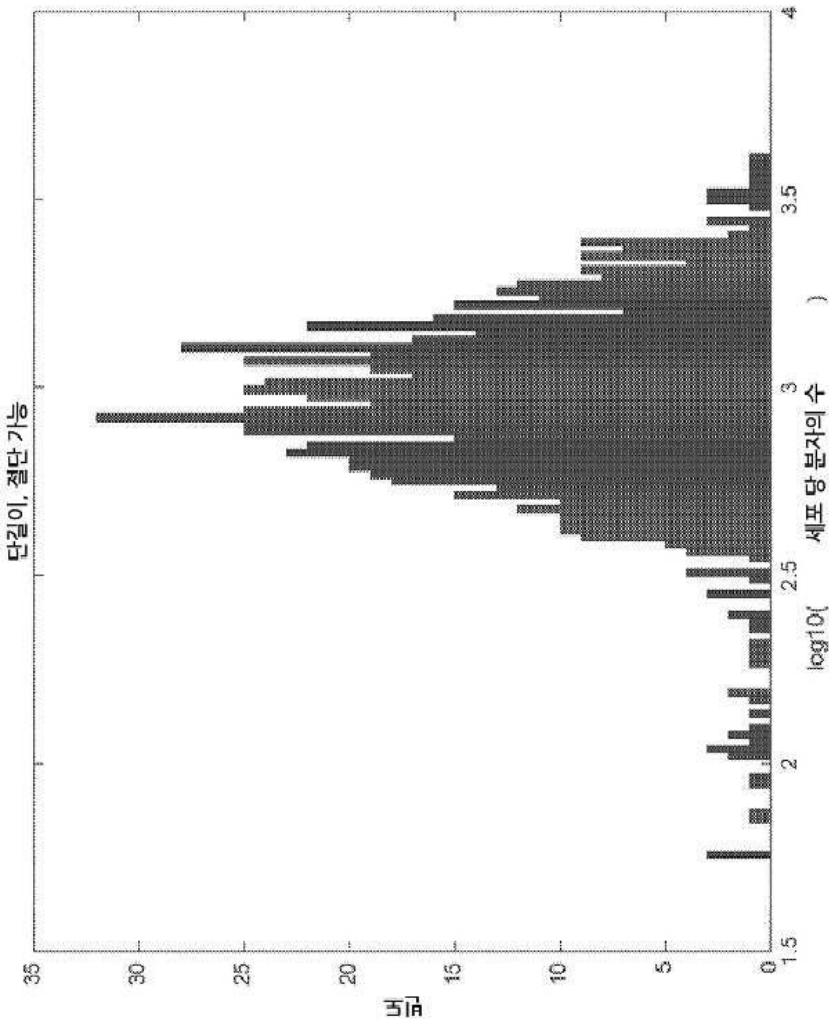
도면25



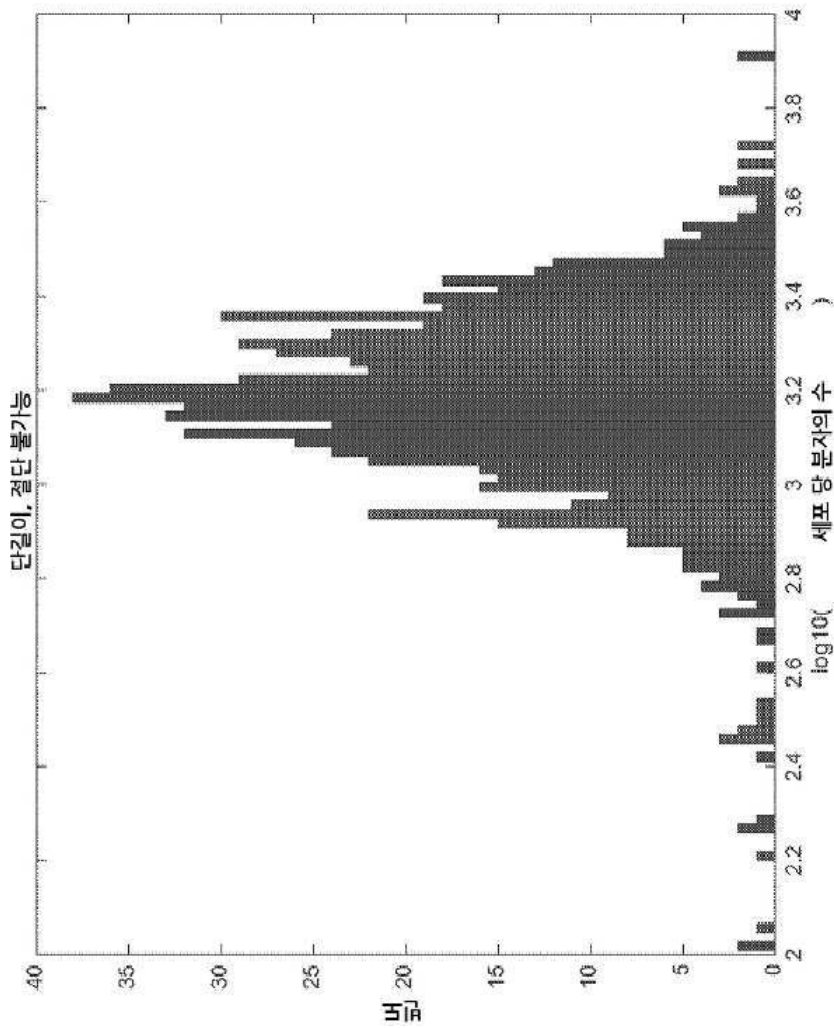
도면26a



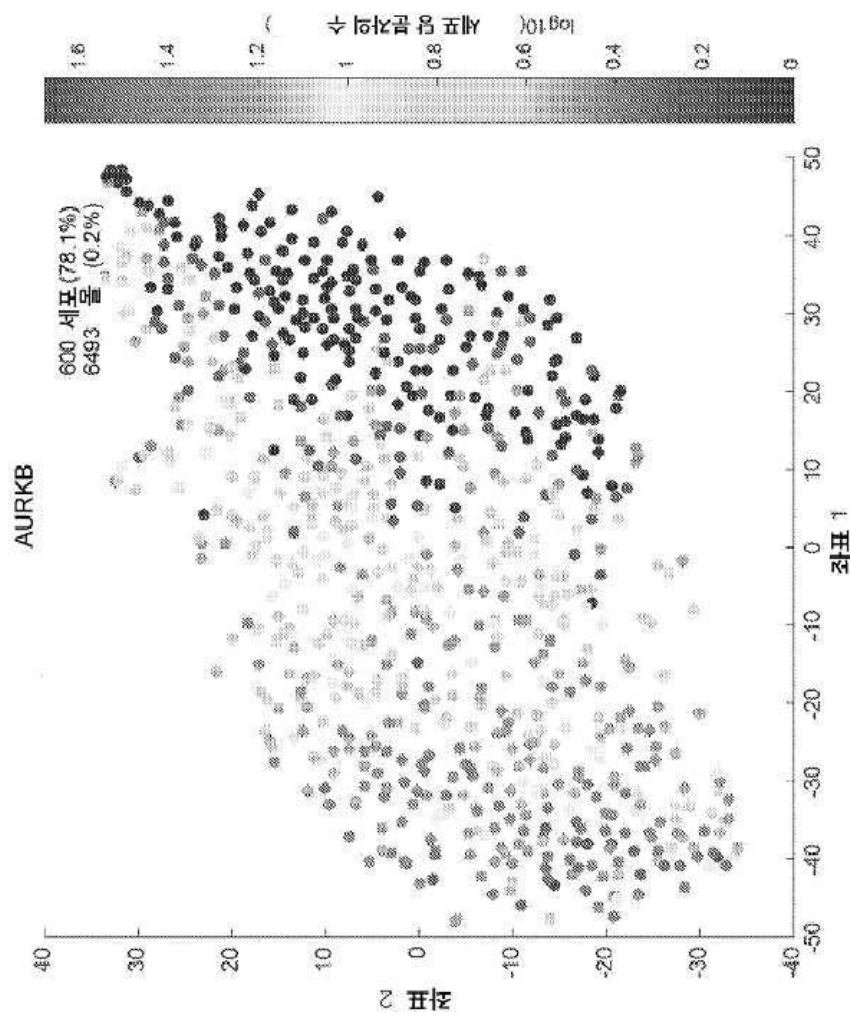
도면26b



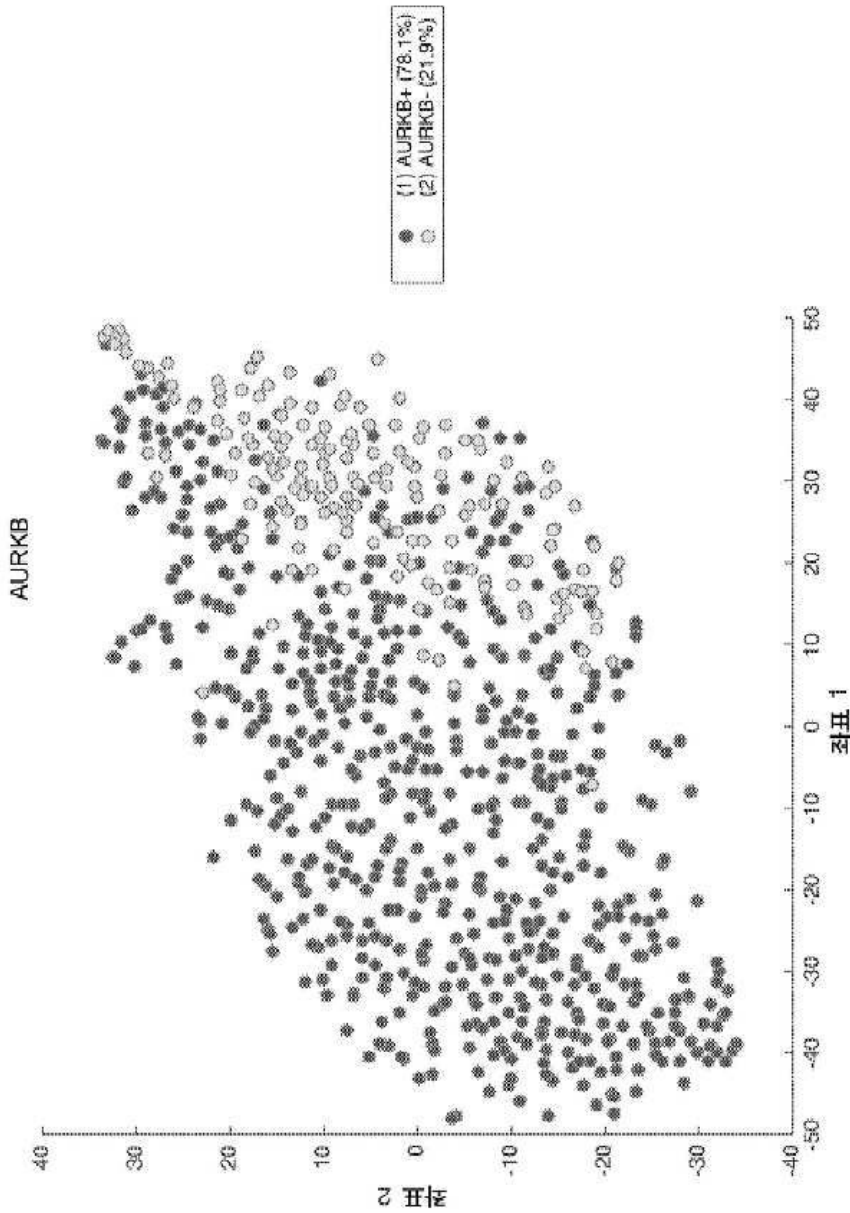
도면26c



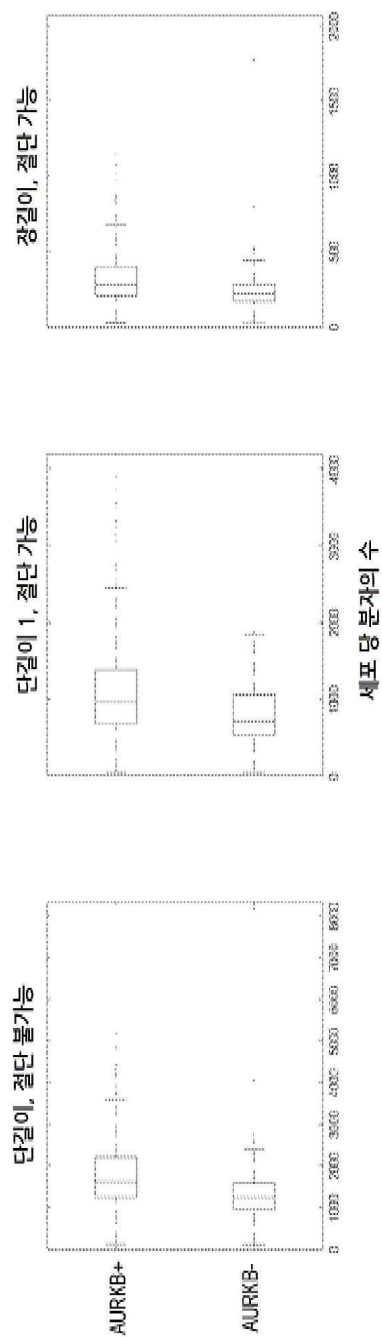
도면27a



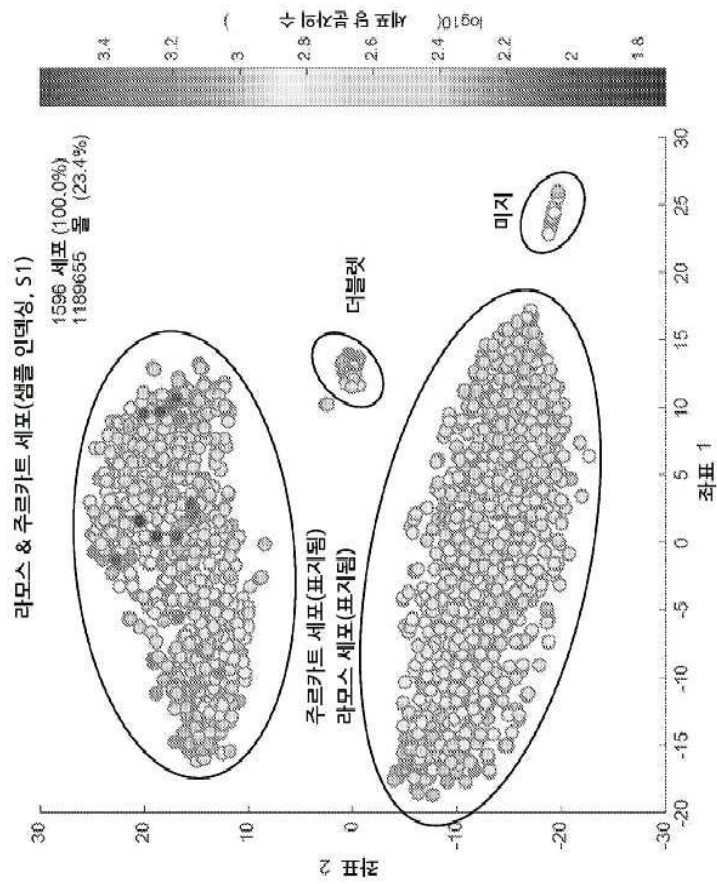
도면27b



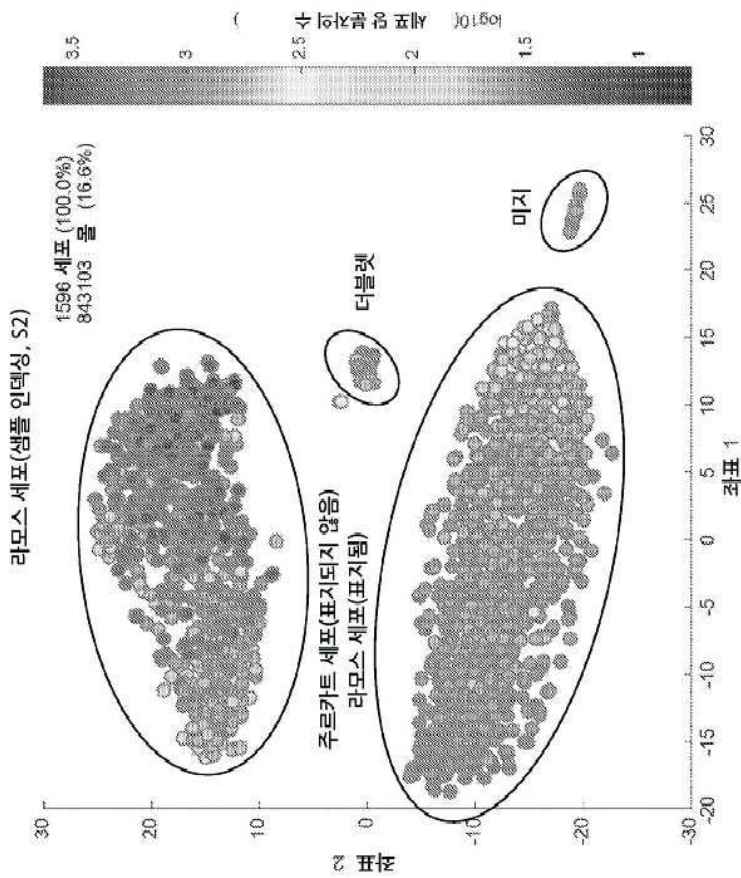
도면27c



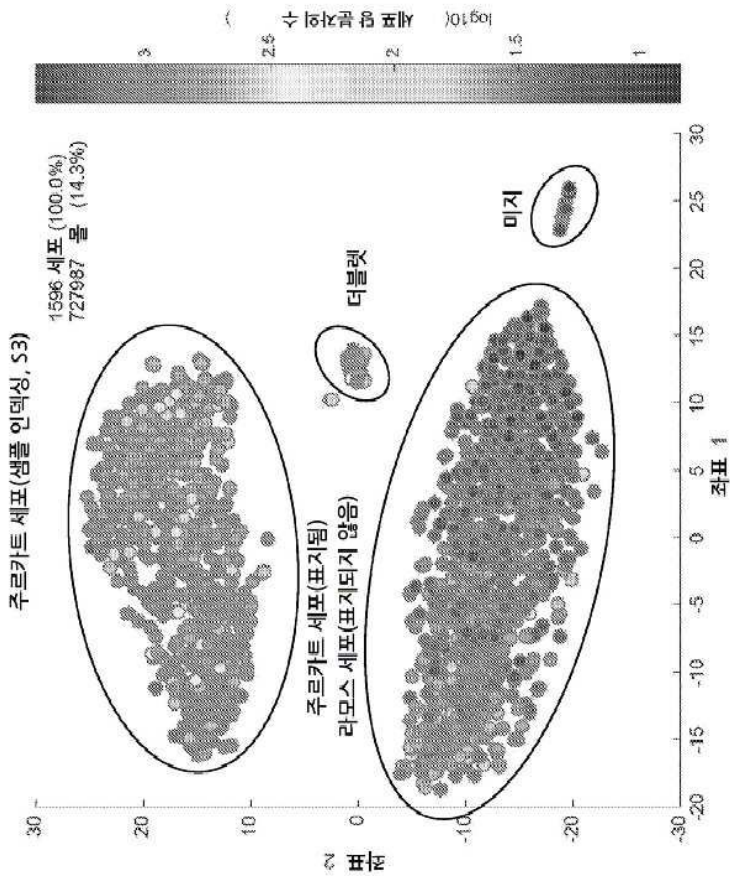
도면28a



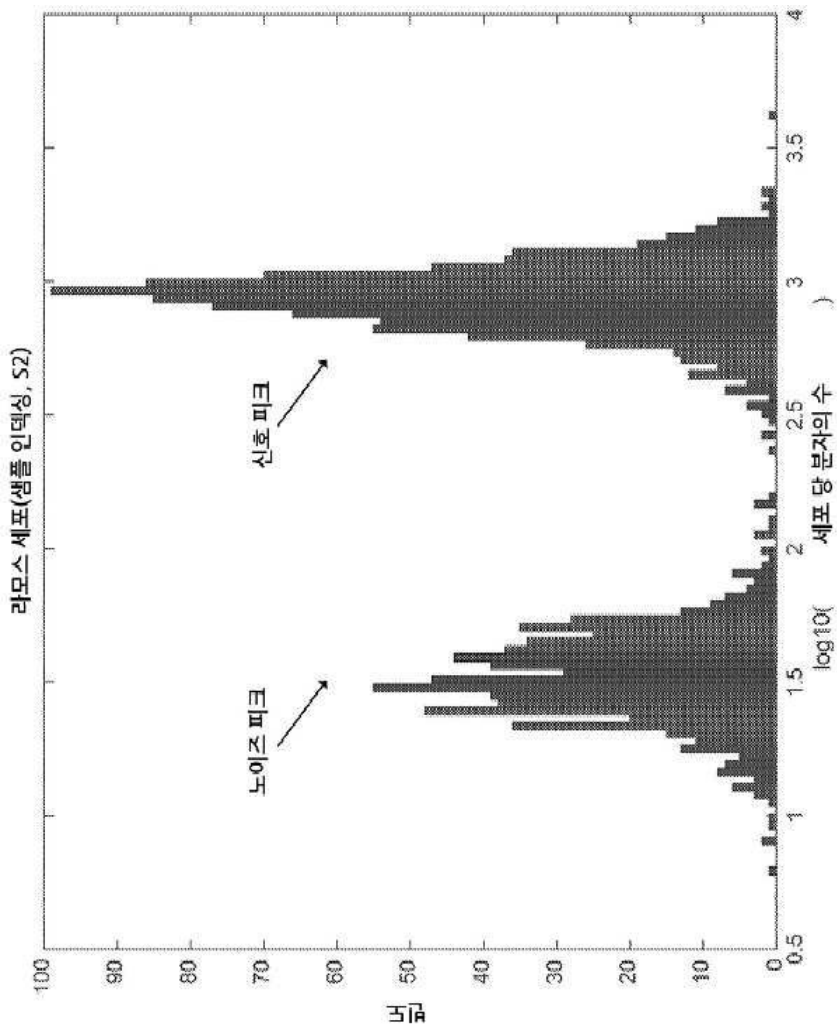
도면28b



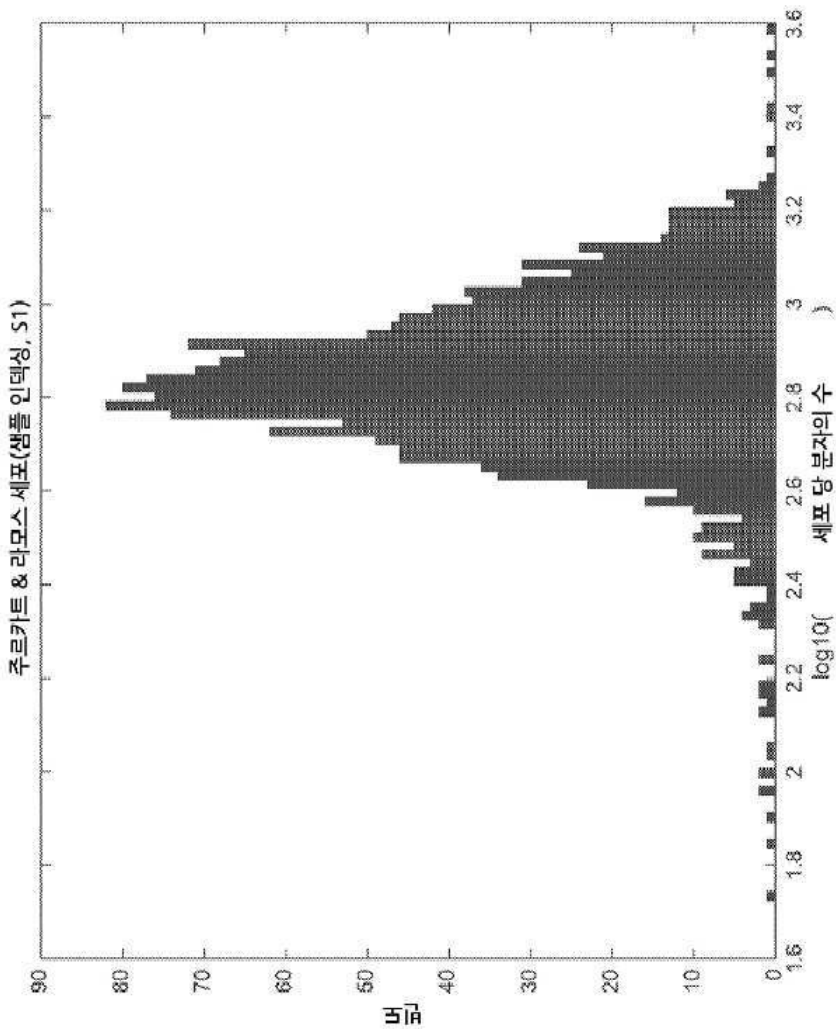
도면28c



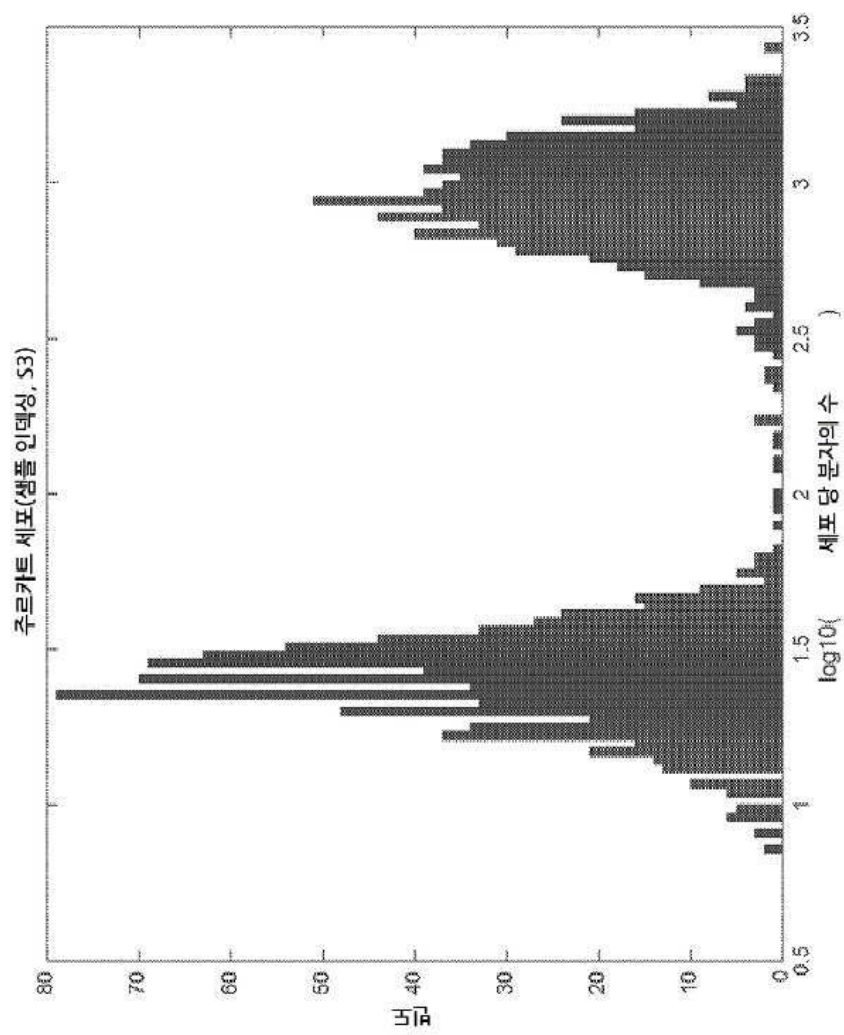
도면29a



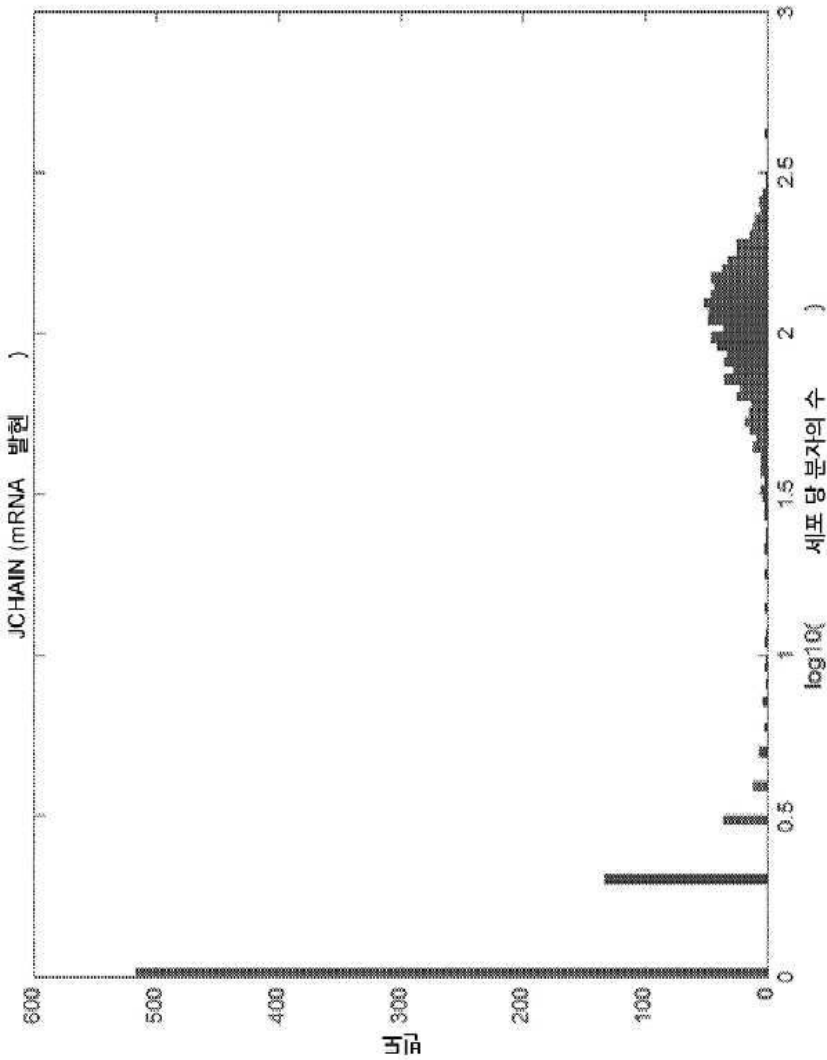
도면29b



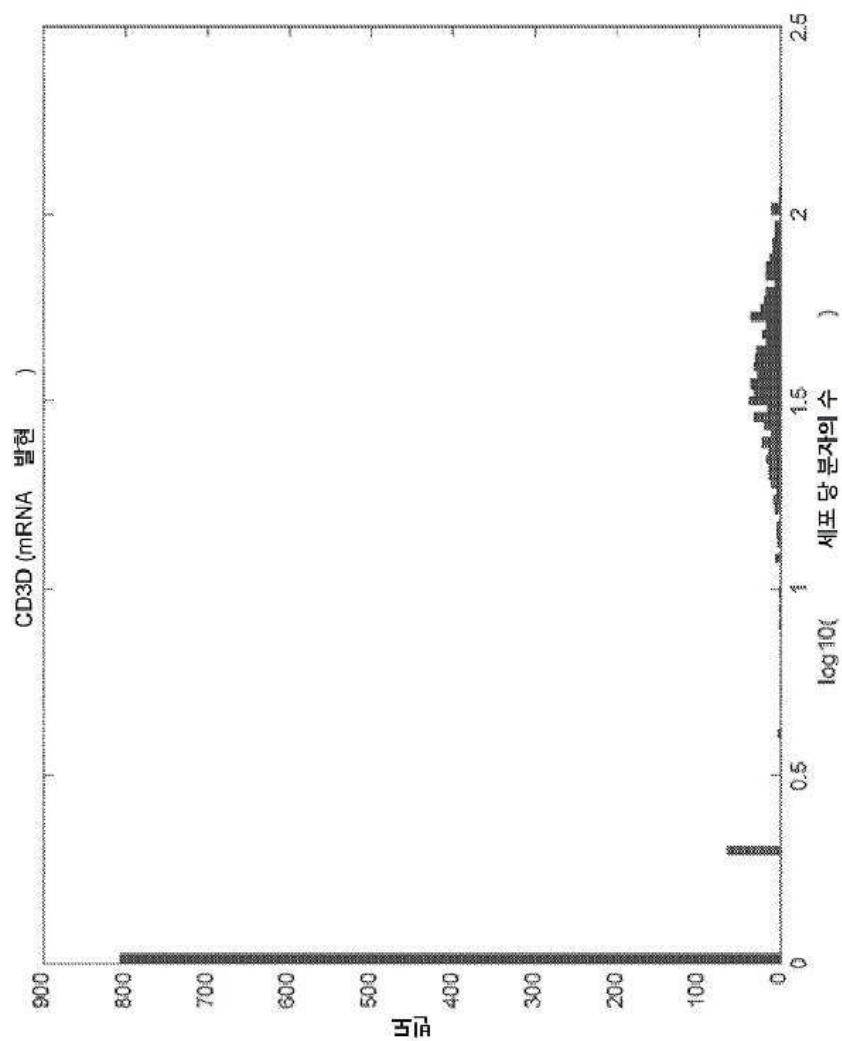
도면29c



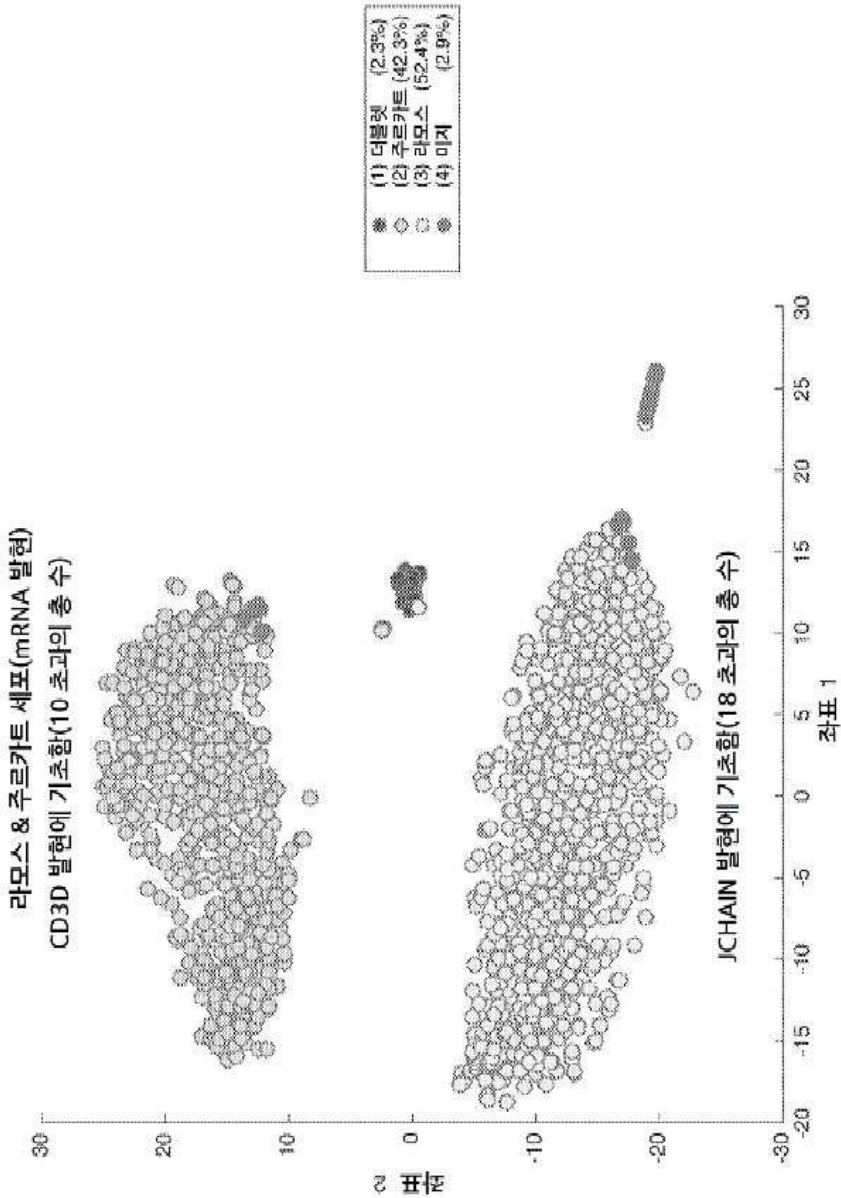
도면30a



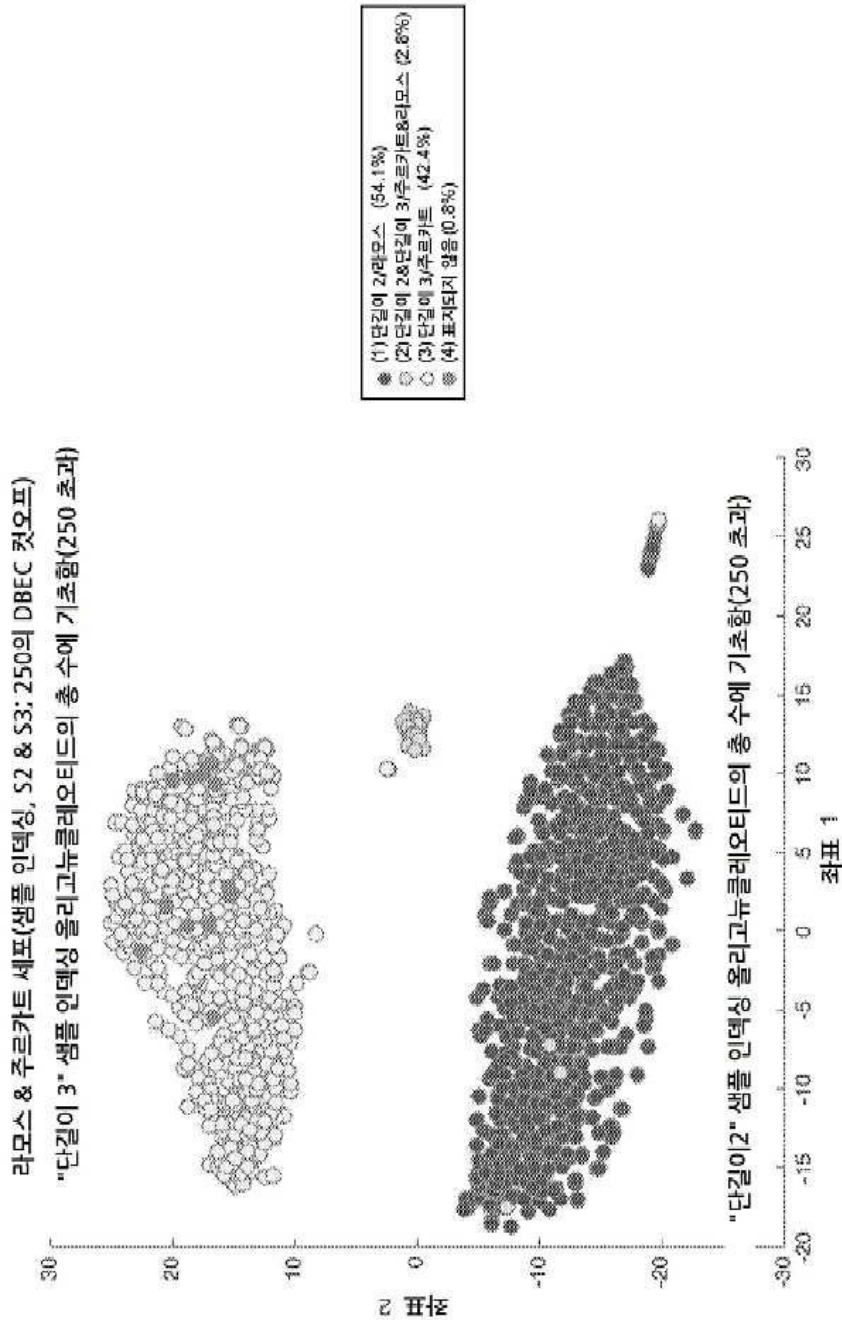
도면30b



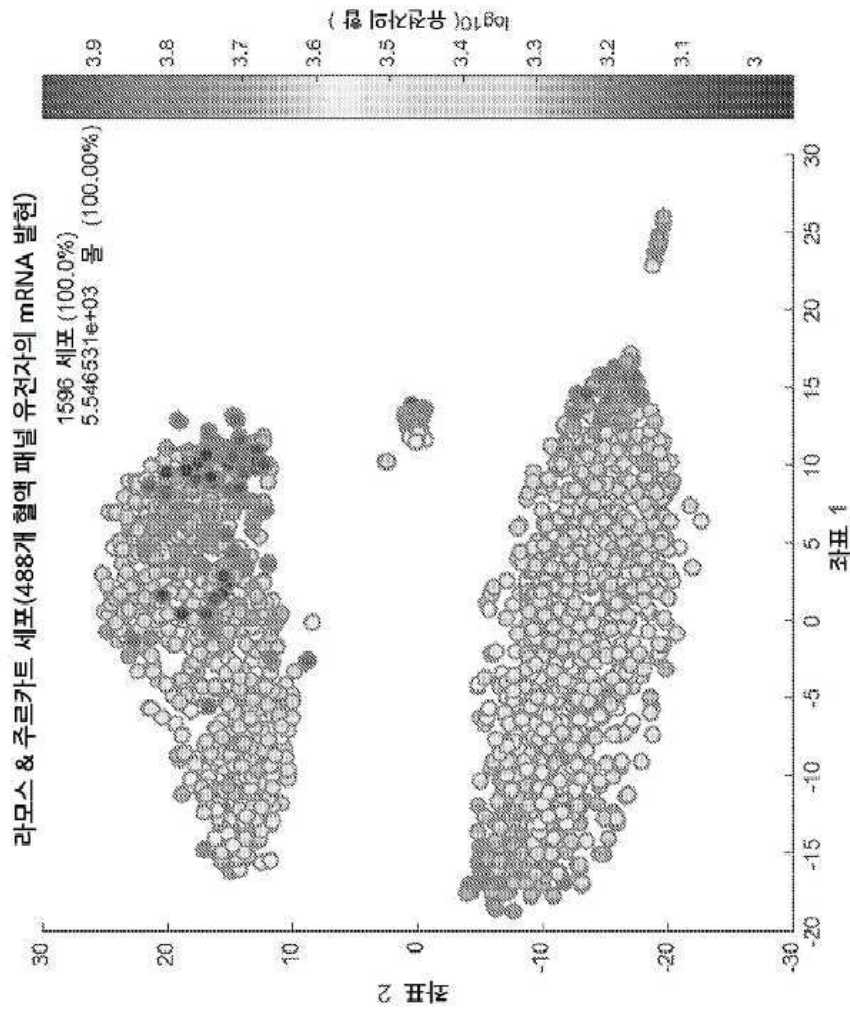
도면30c



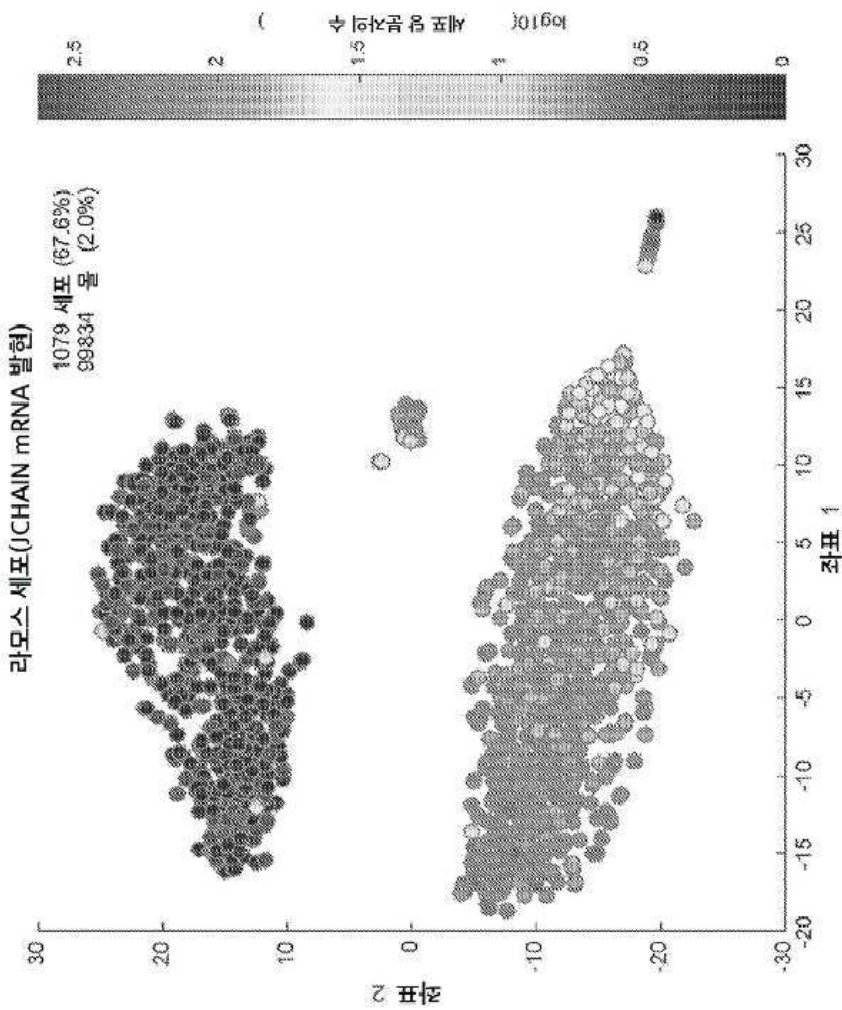
도면30d



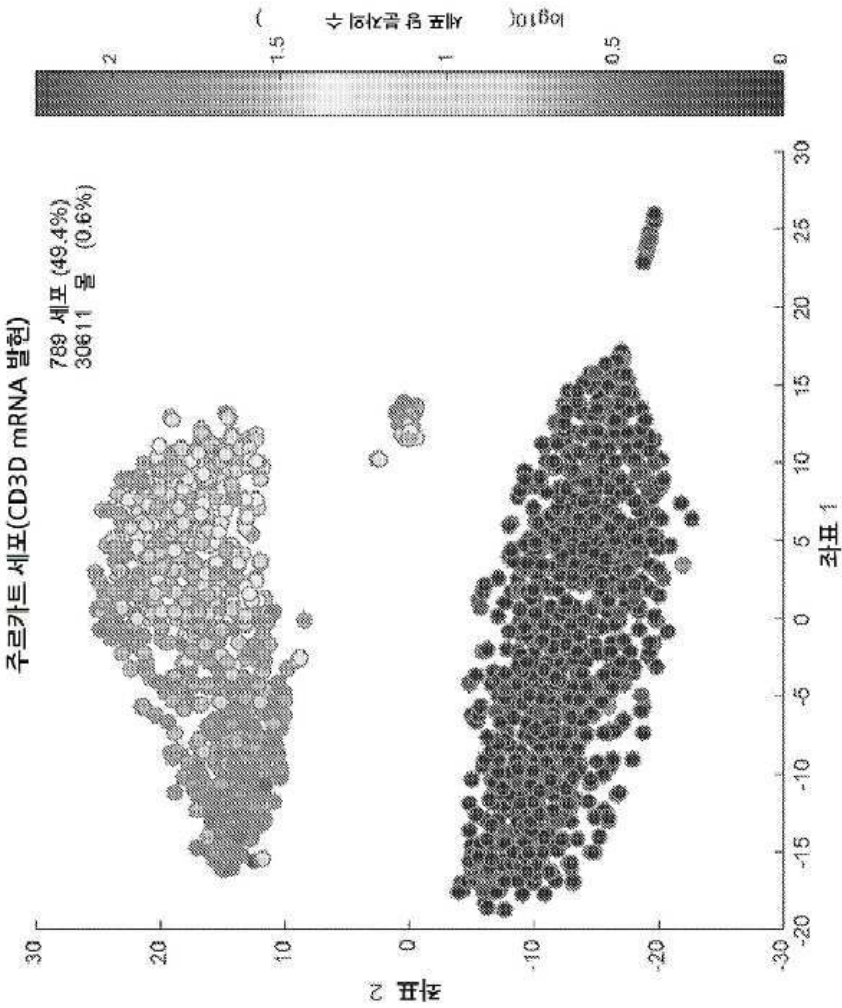
도면31a



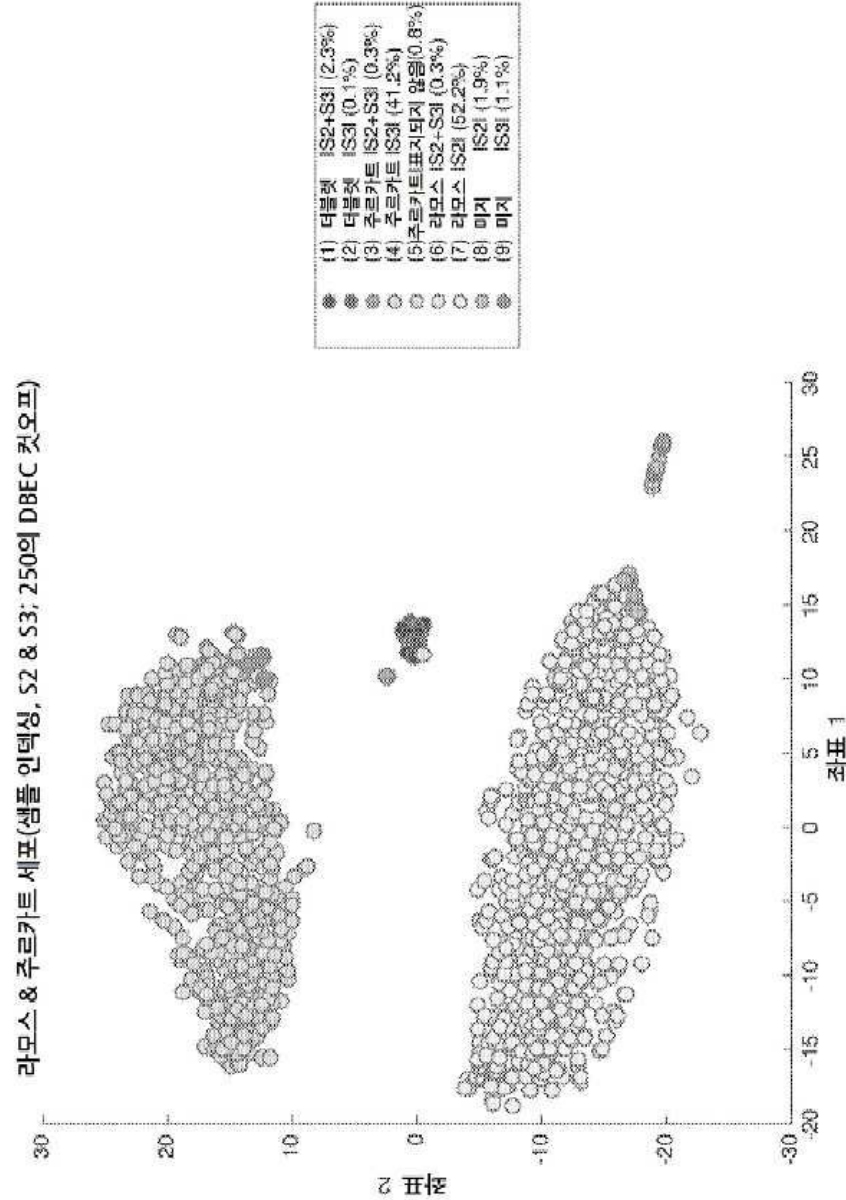
도면31b



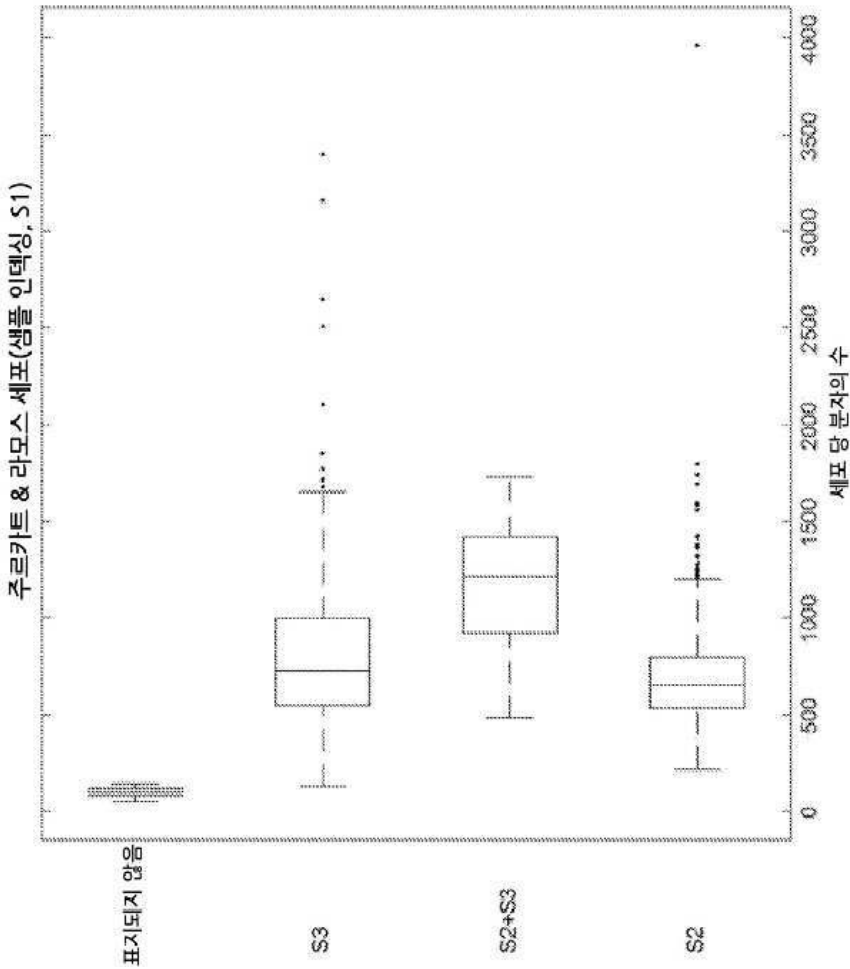
도면31c



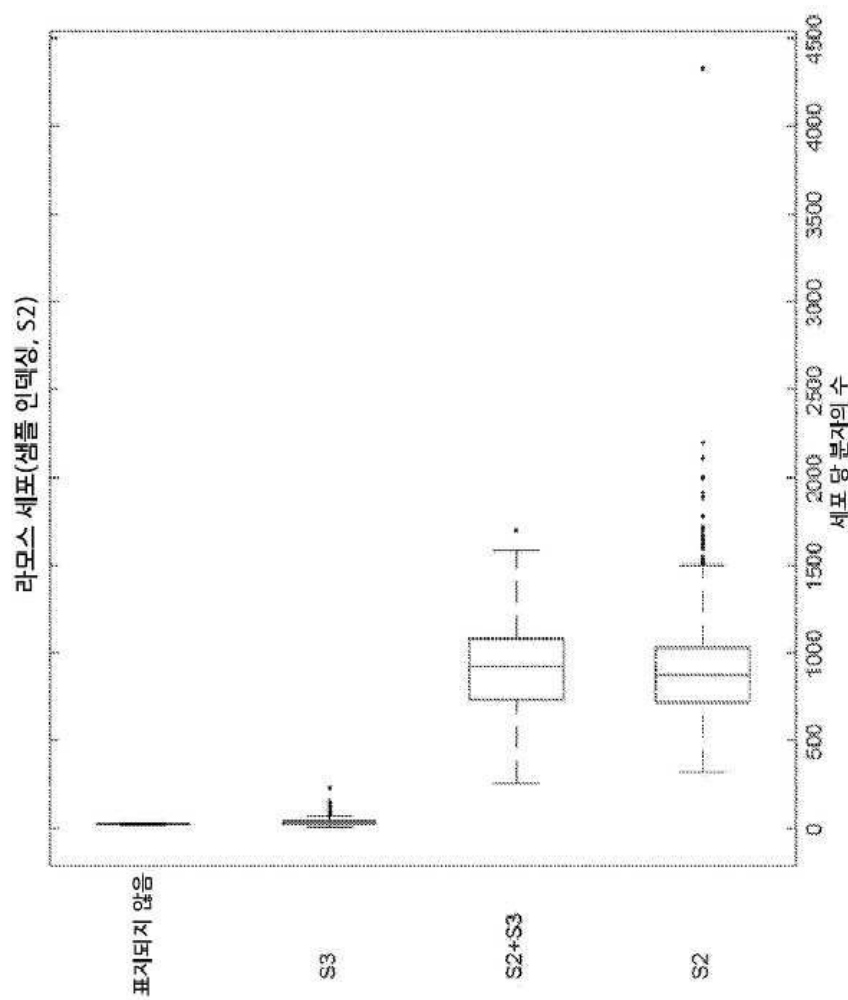
도면32



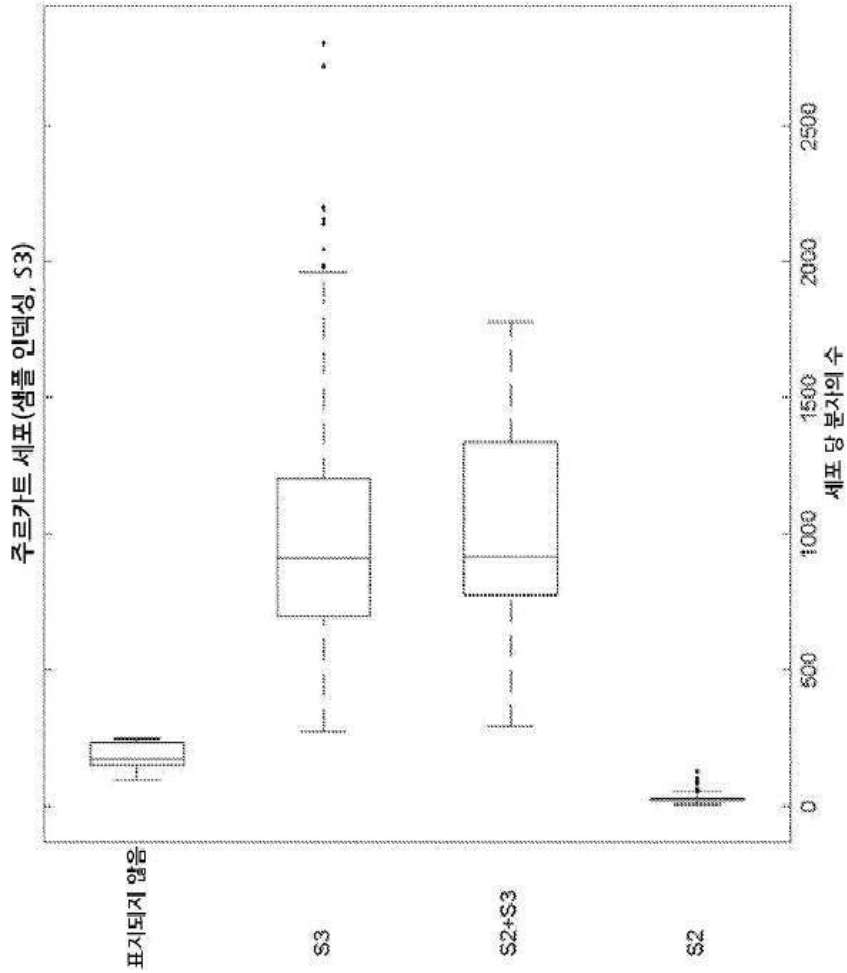
도면33a



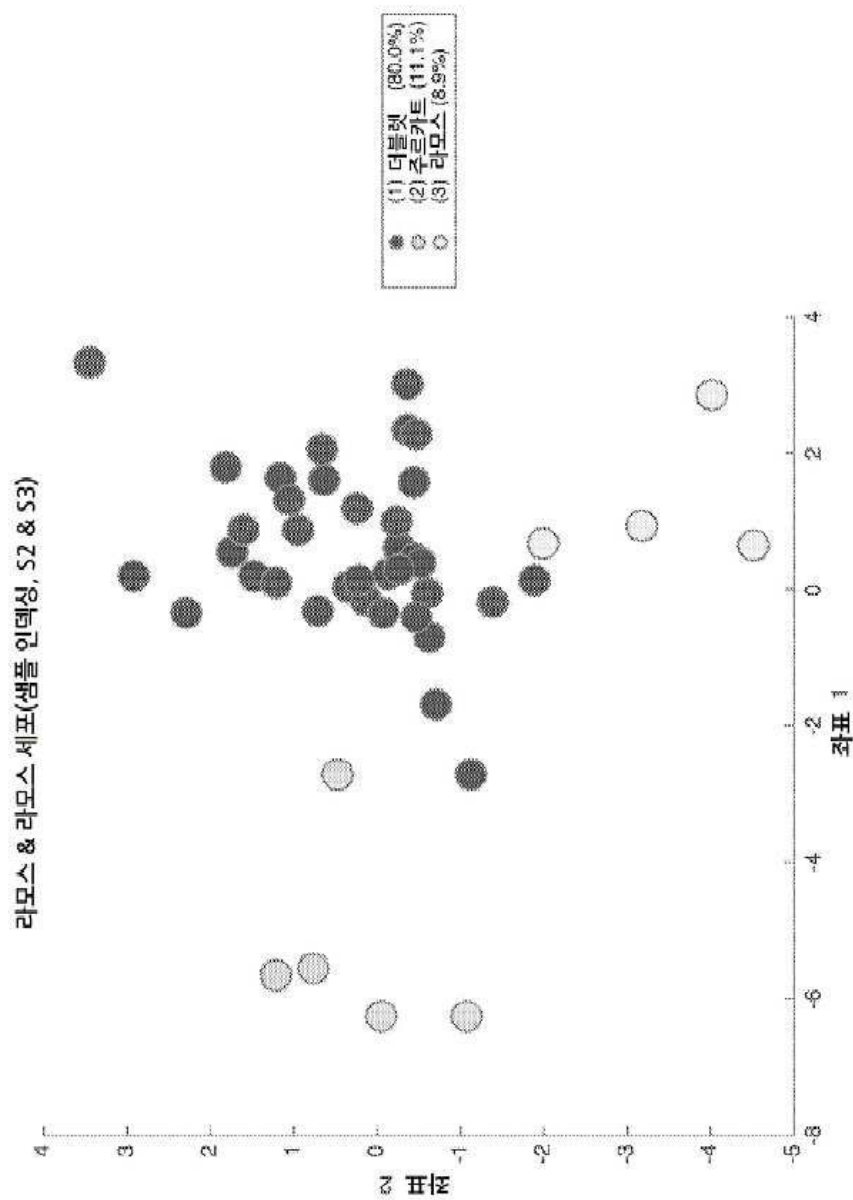
도면33b



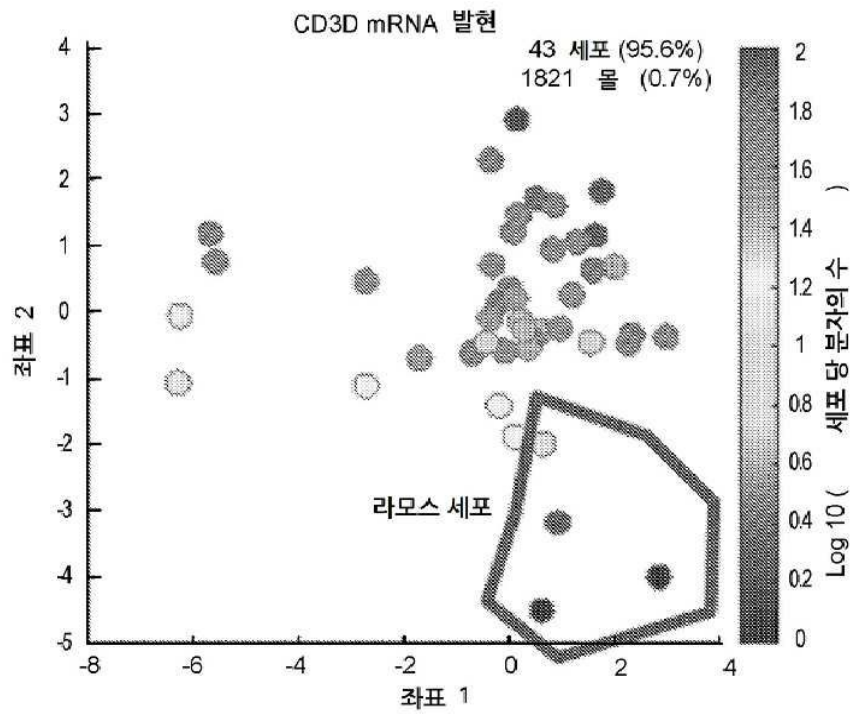
도면33c



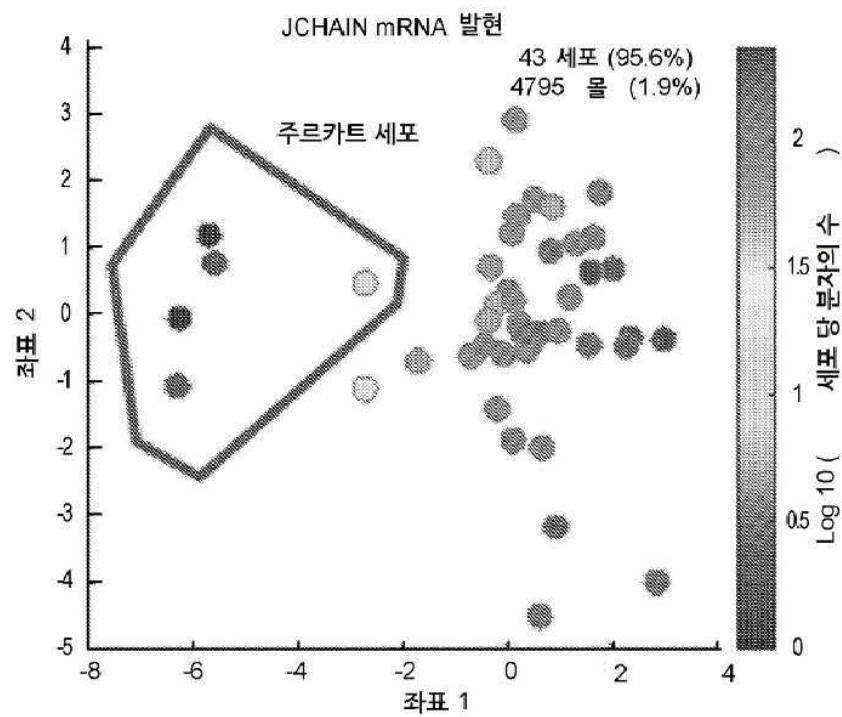
도면34a



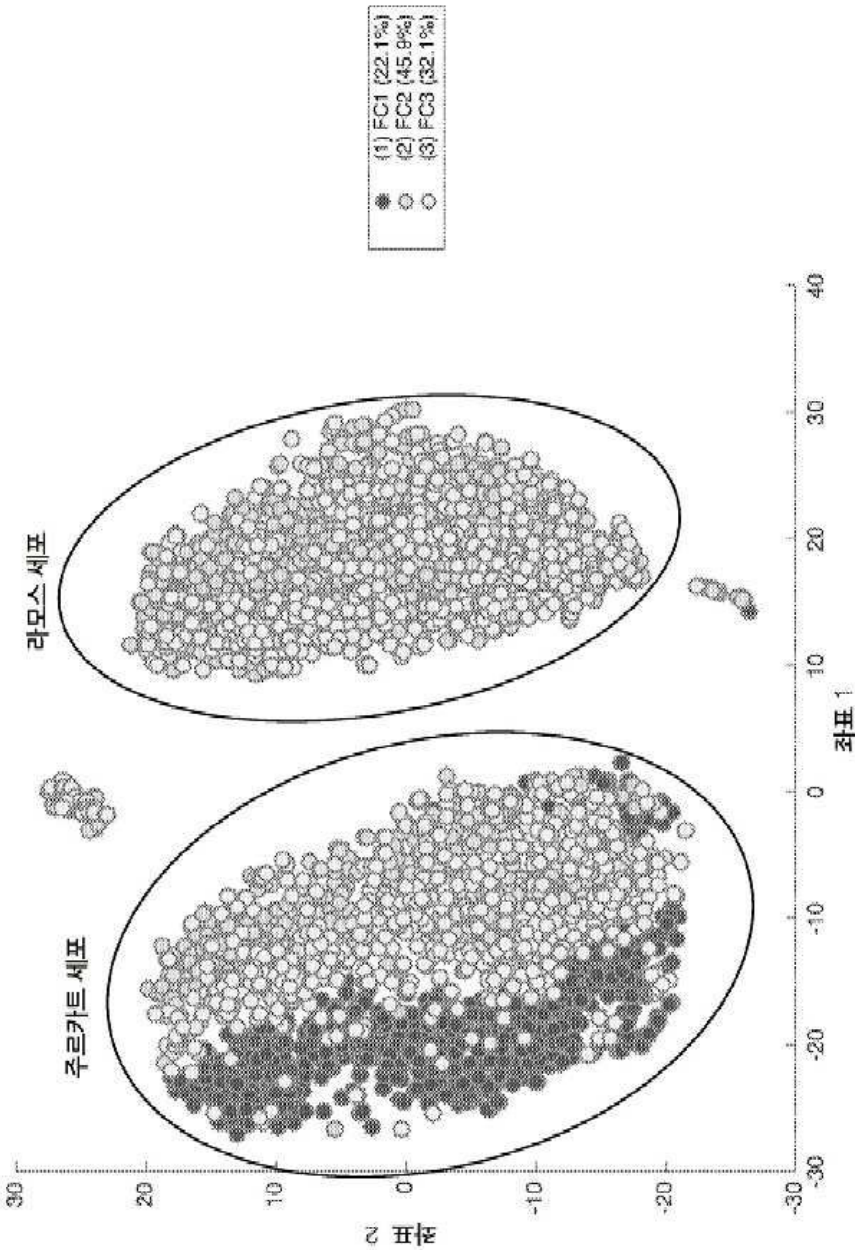
도면34b



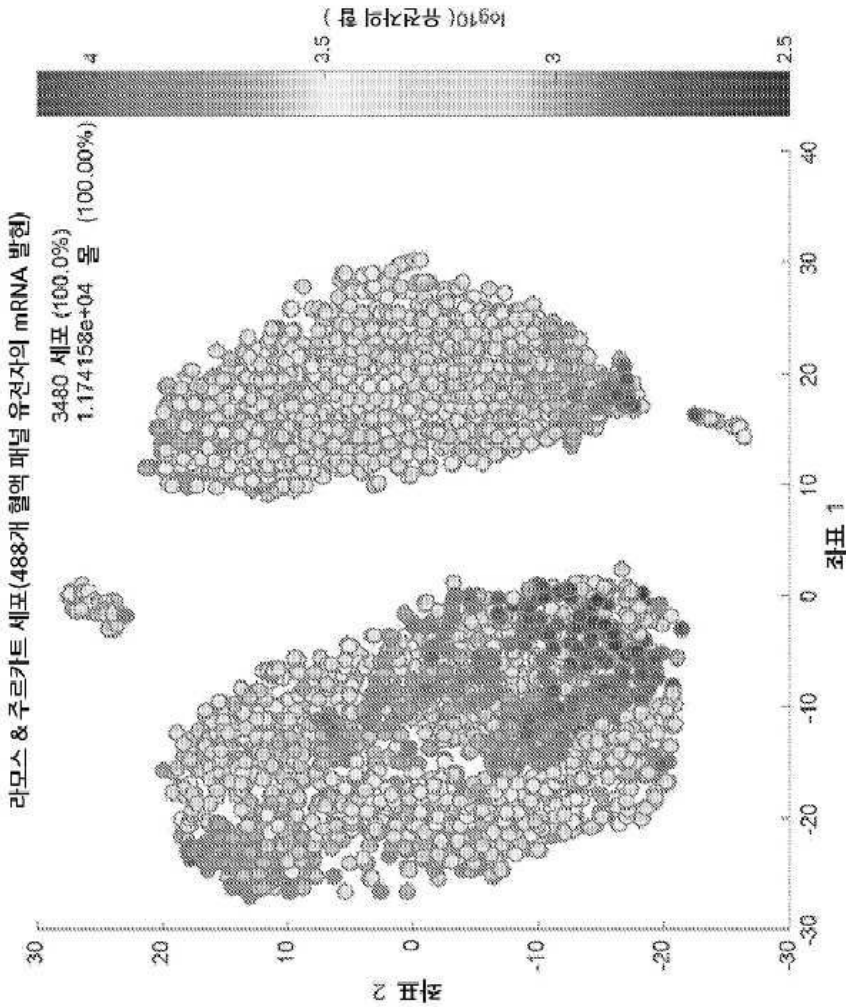
도면34c



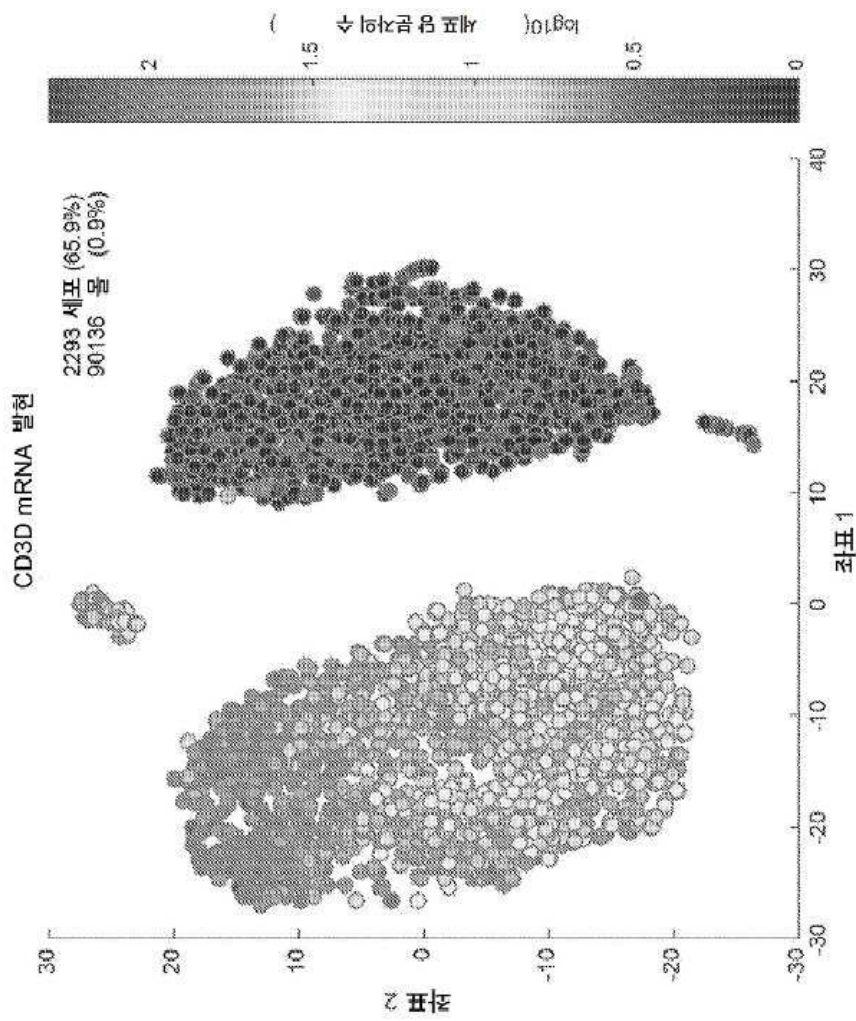
도면35a



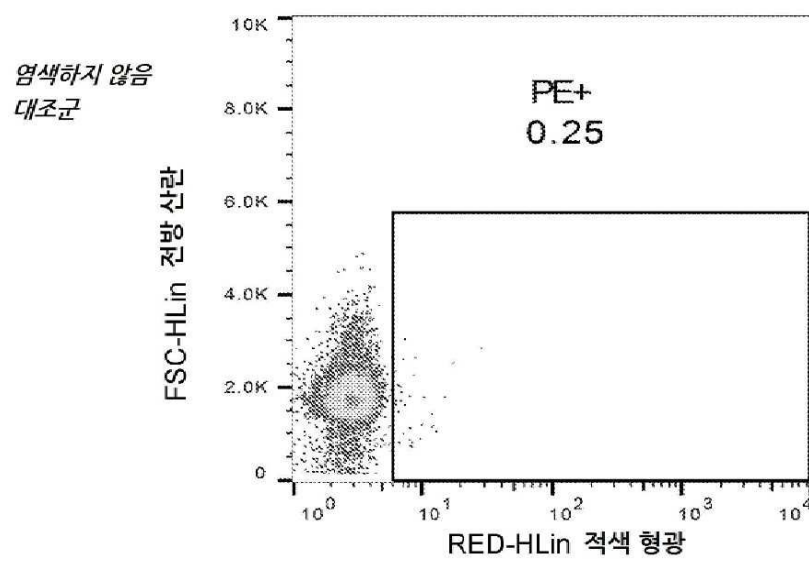
도면35b



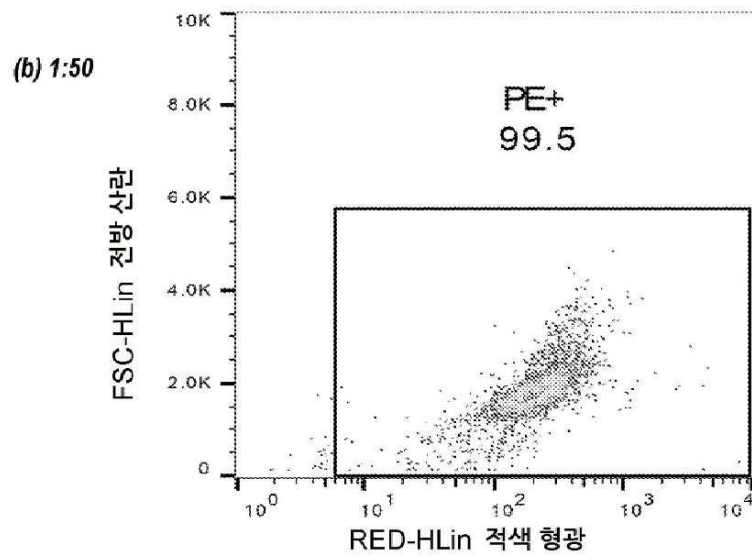
도면35c



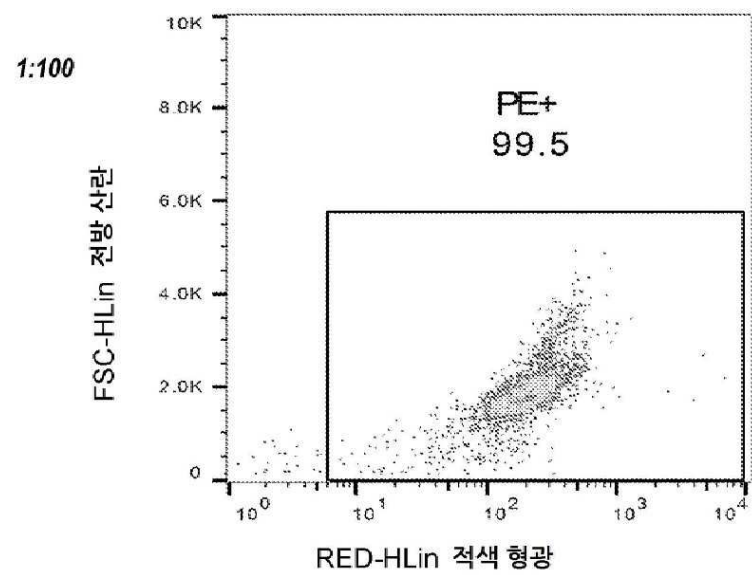
도면36aa



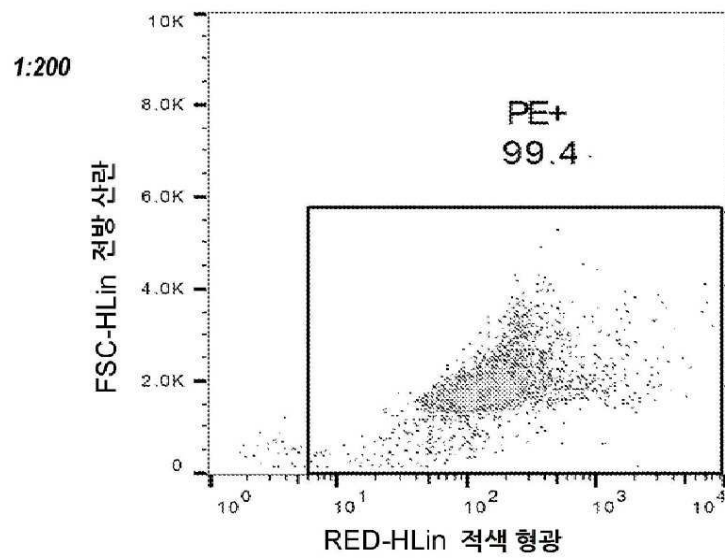
도면36ab



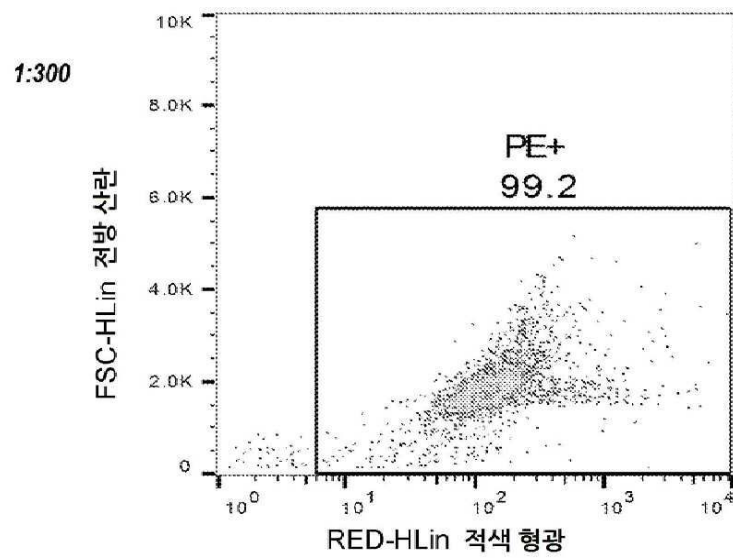
도면36ac



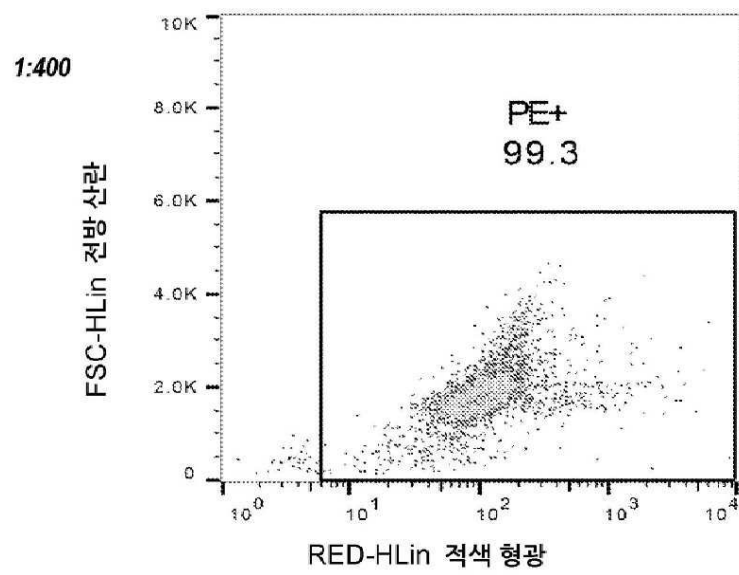
도면36ad



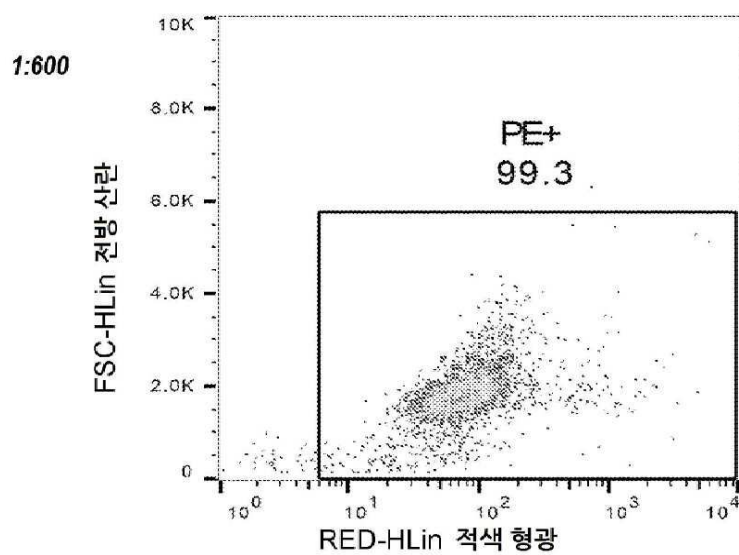
도면36ae



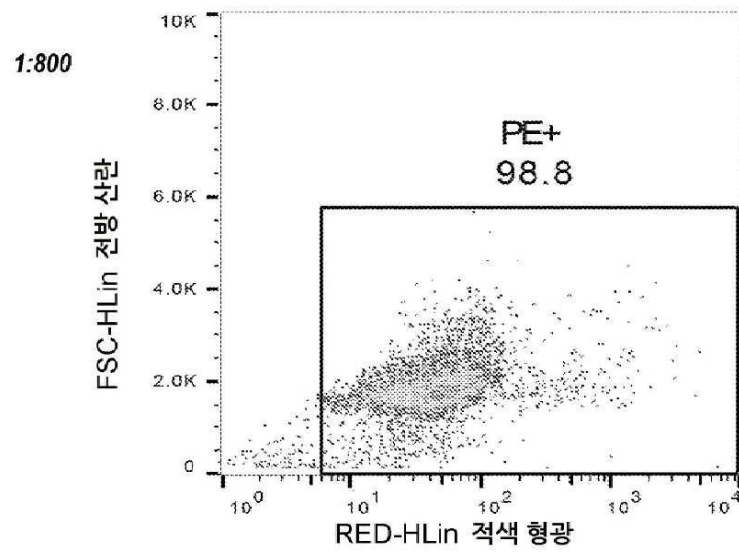
도면36af



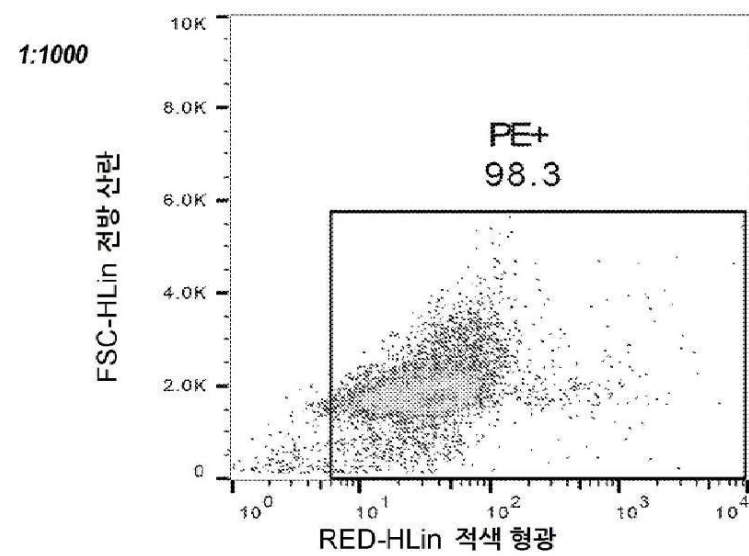
도면36ag



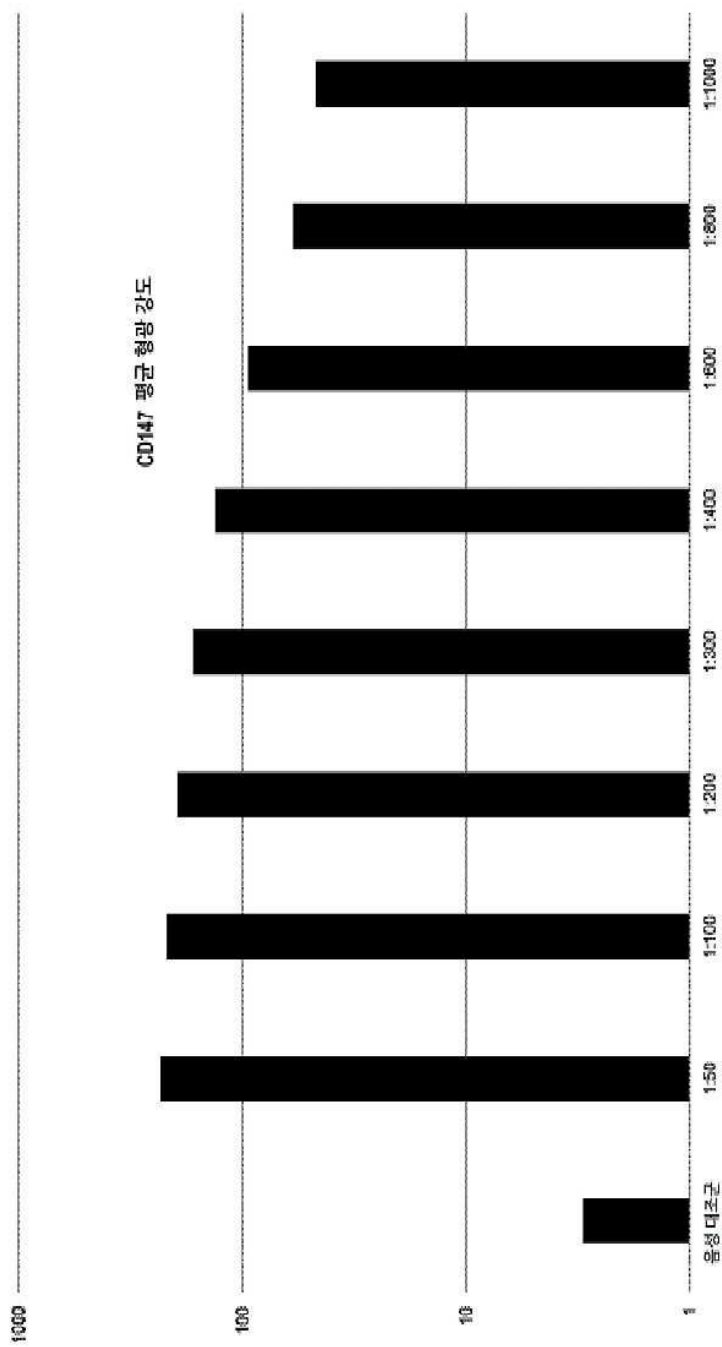
도면36ah



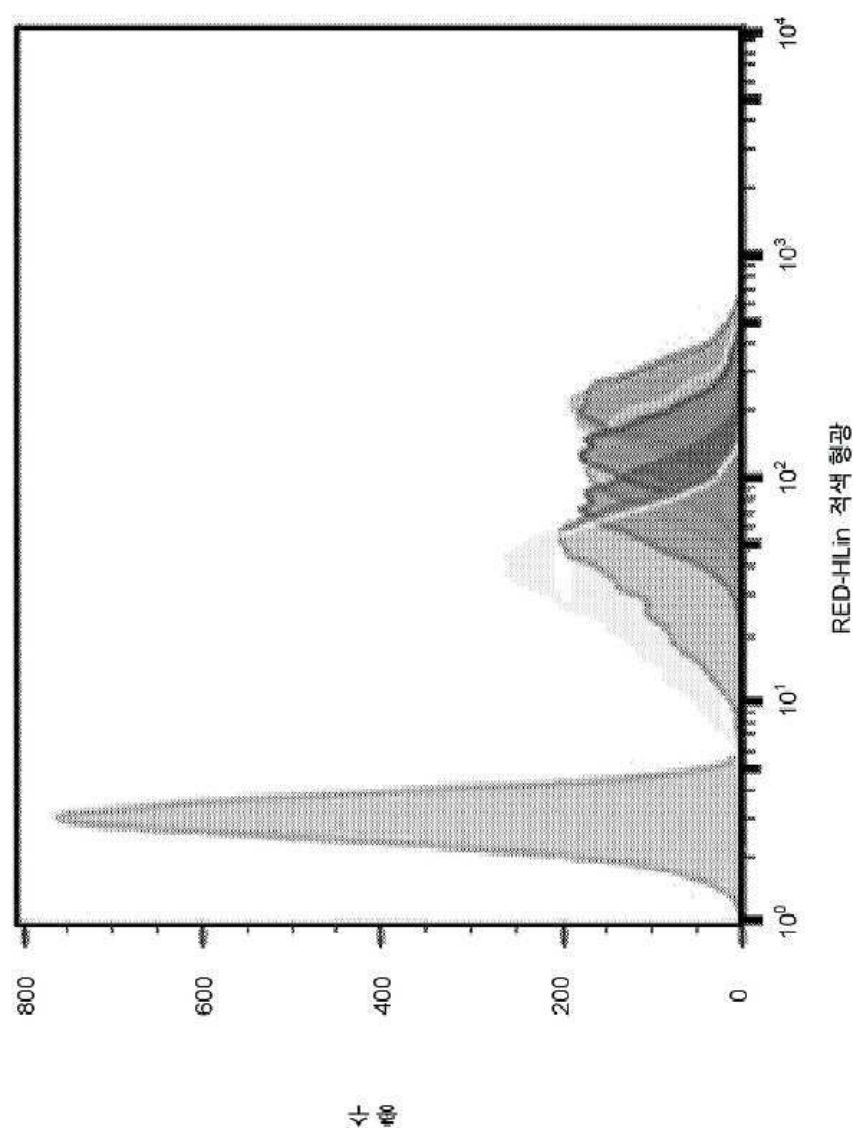
도면36ai



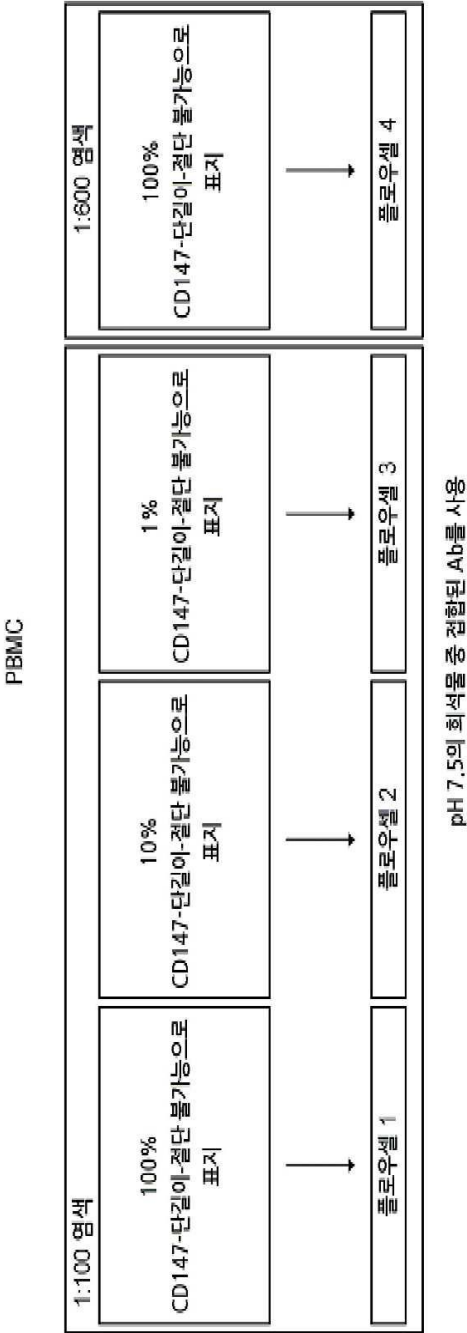
도면36b



도면36c

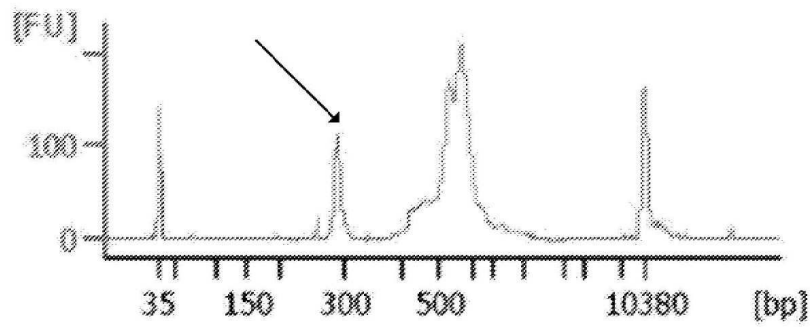


도면37



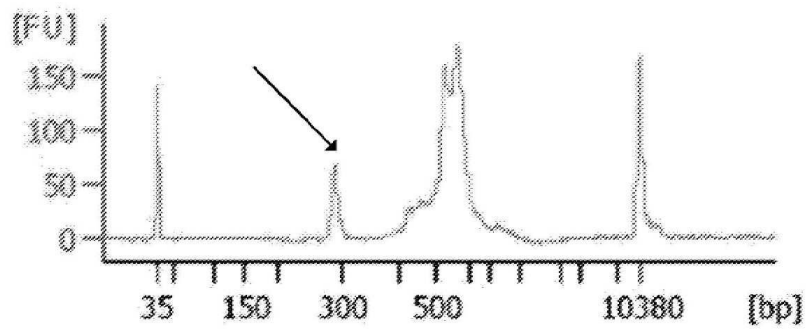
도면38a

FC1 - 1:100 희석 , 100% 고온



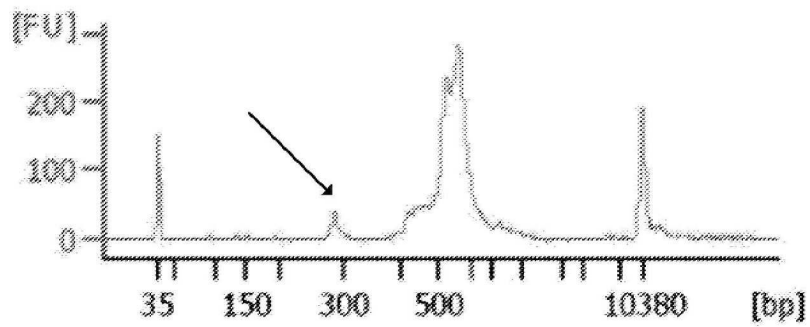
도면38b

FC1 - 1:800 희석 , 100% 고온

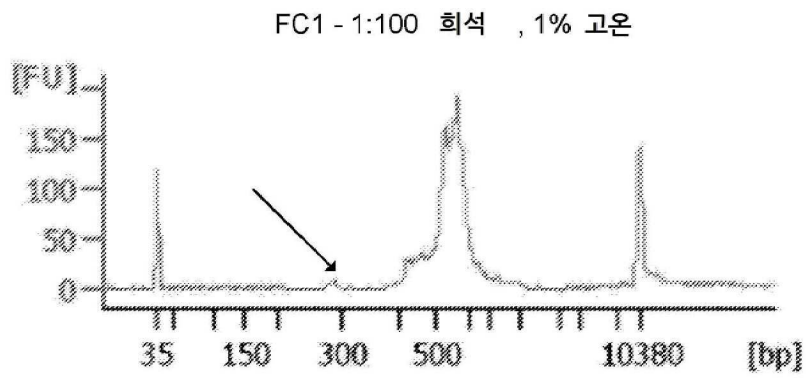


도면38c

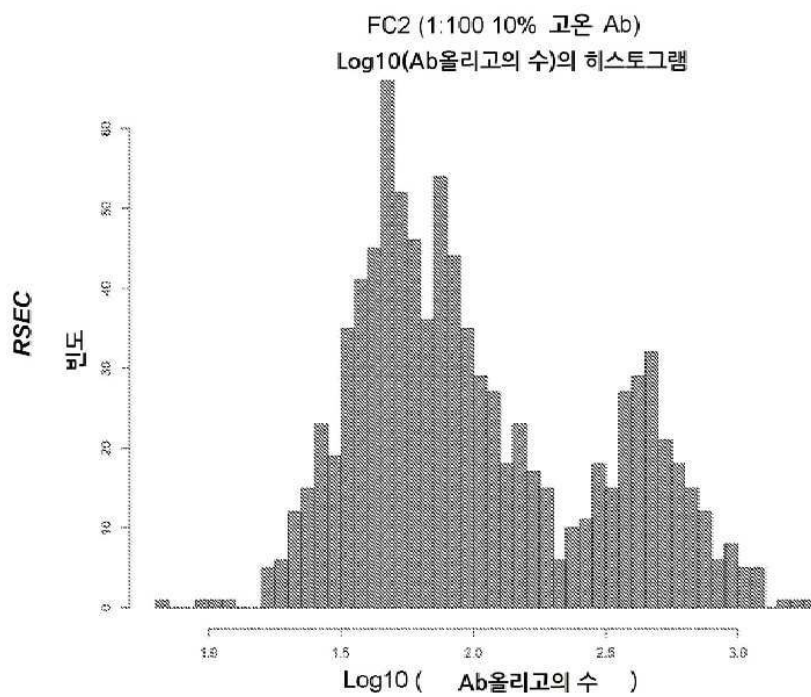
FC1 - 1:100 희석 , 10% 고온



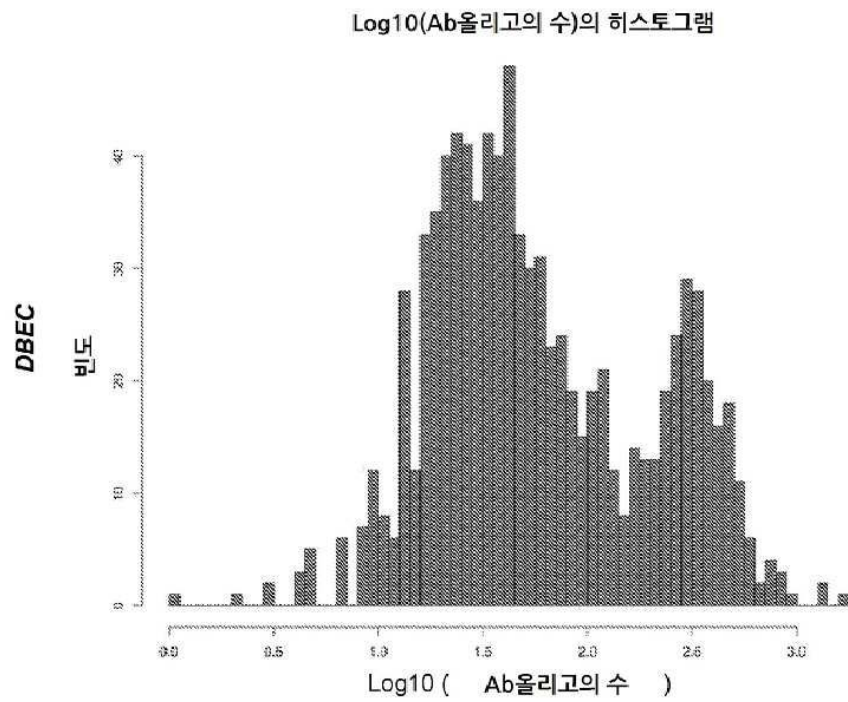
도면38d



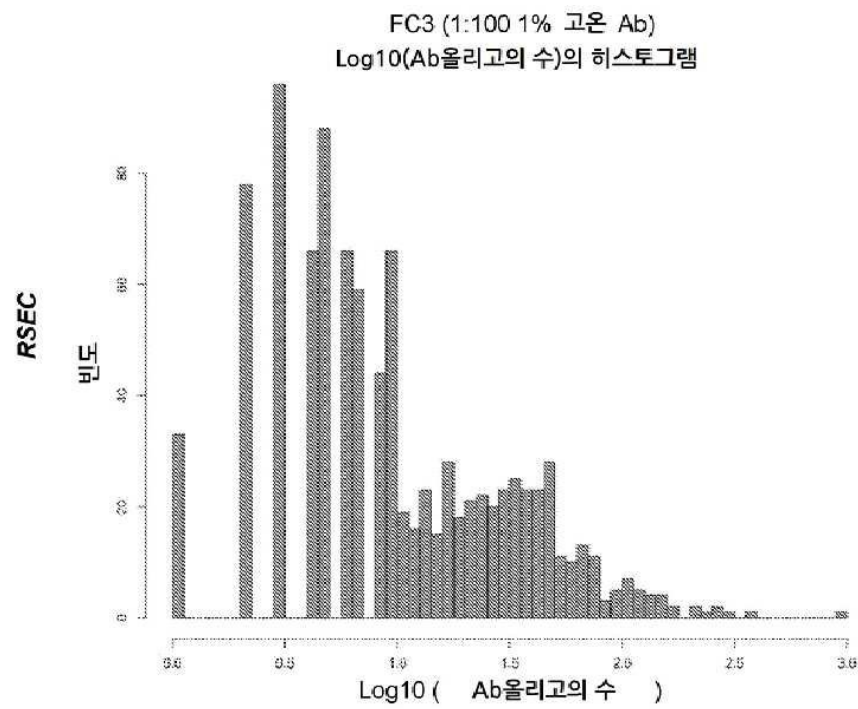
도면39aa



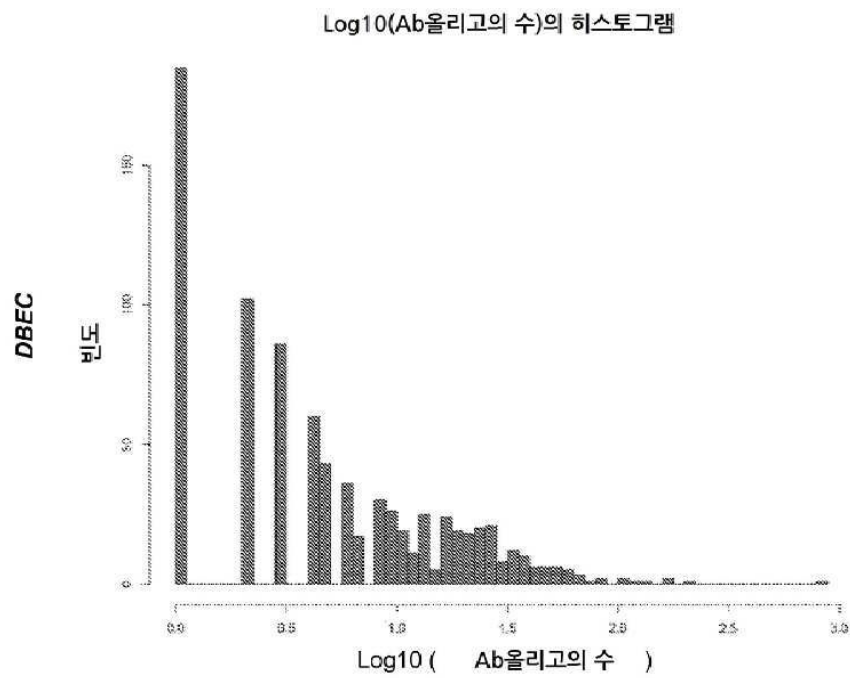
도면39ba



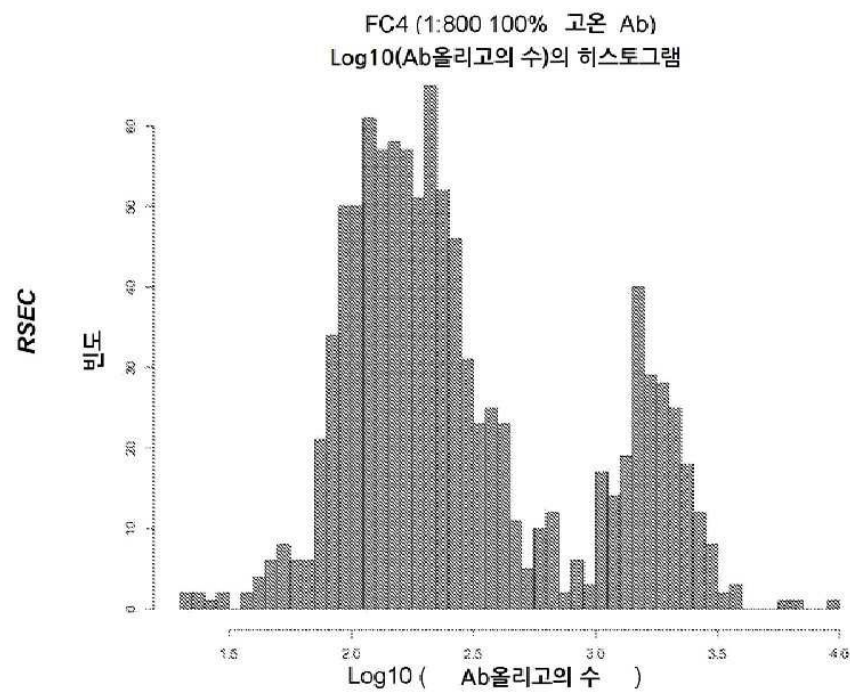
도면39ab



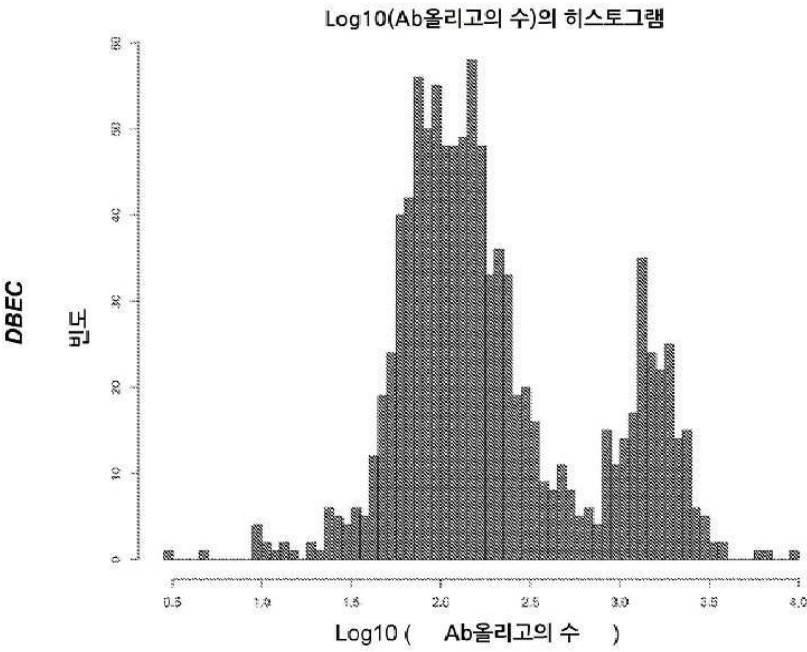
도면39bb



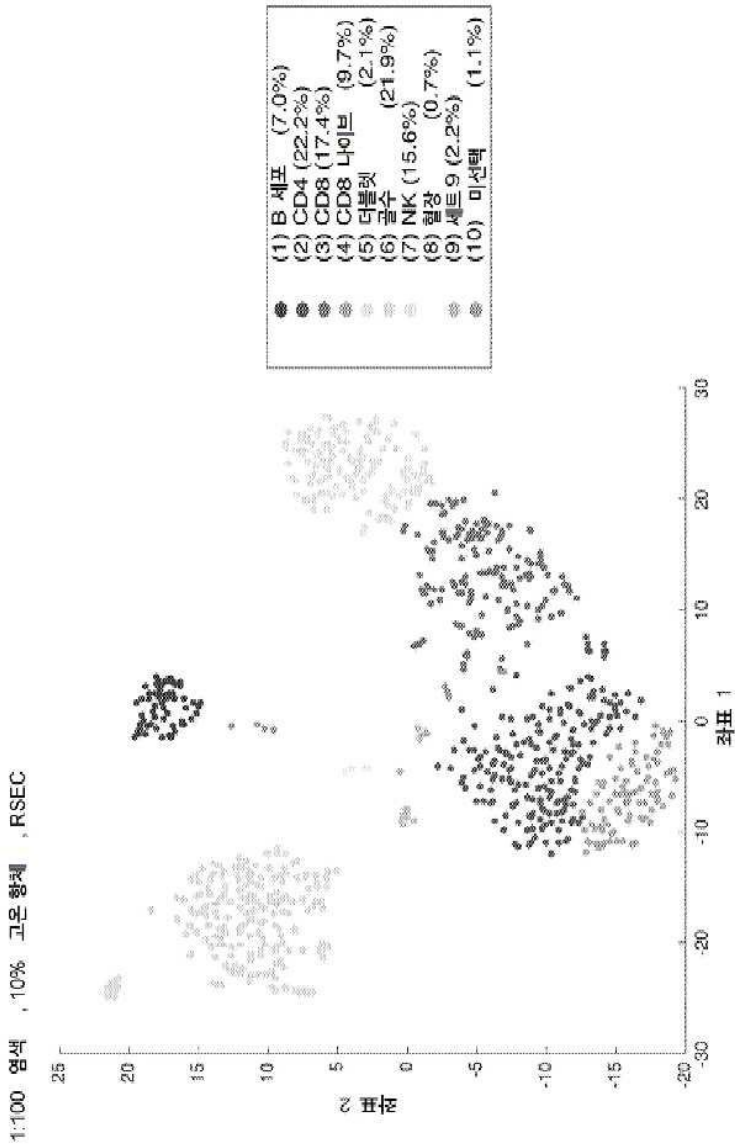
도면39ac



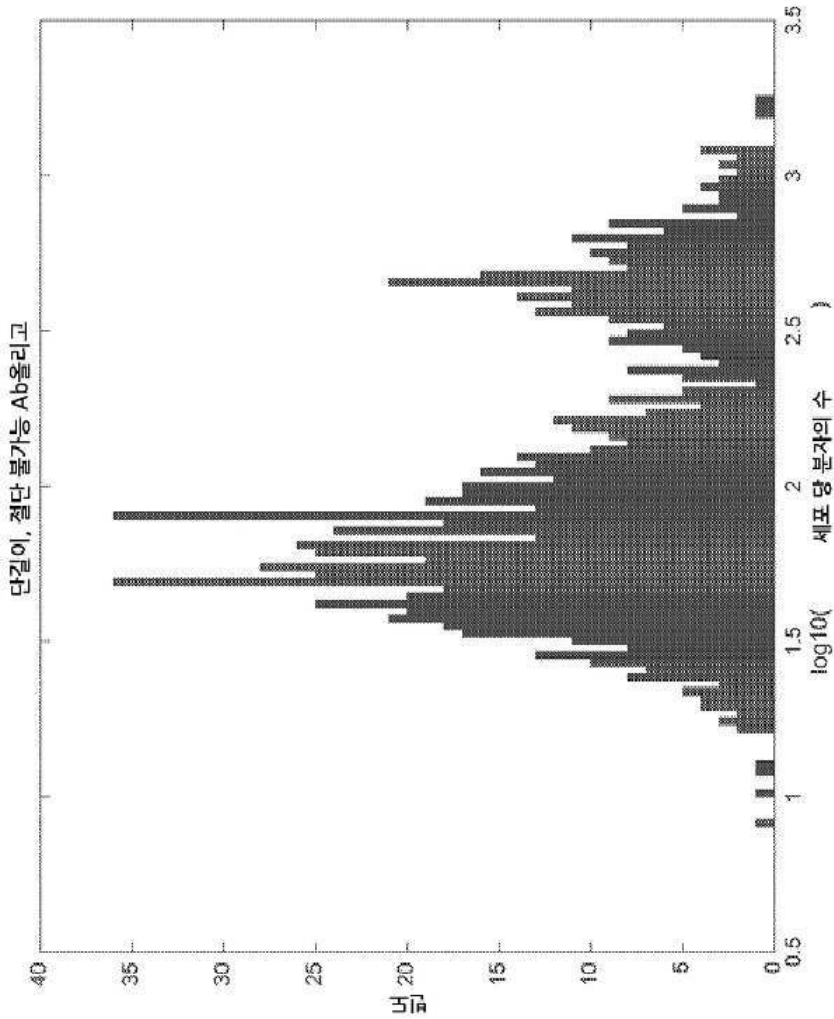
도면39bc



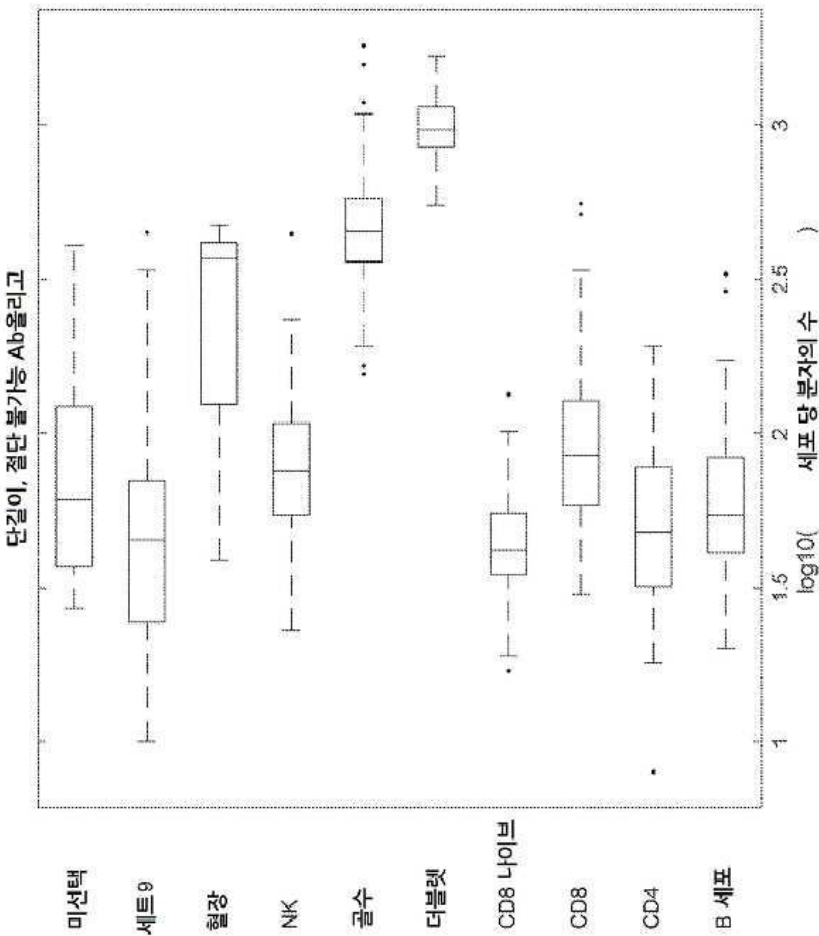
도면40a



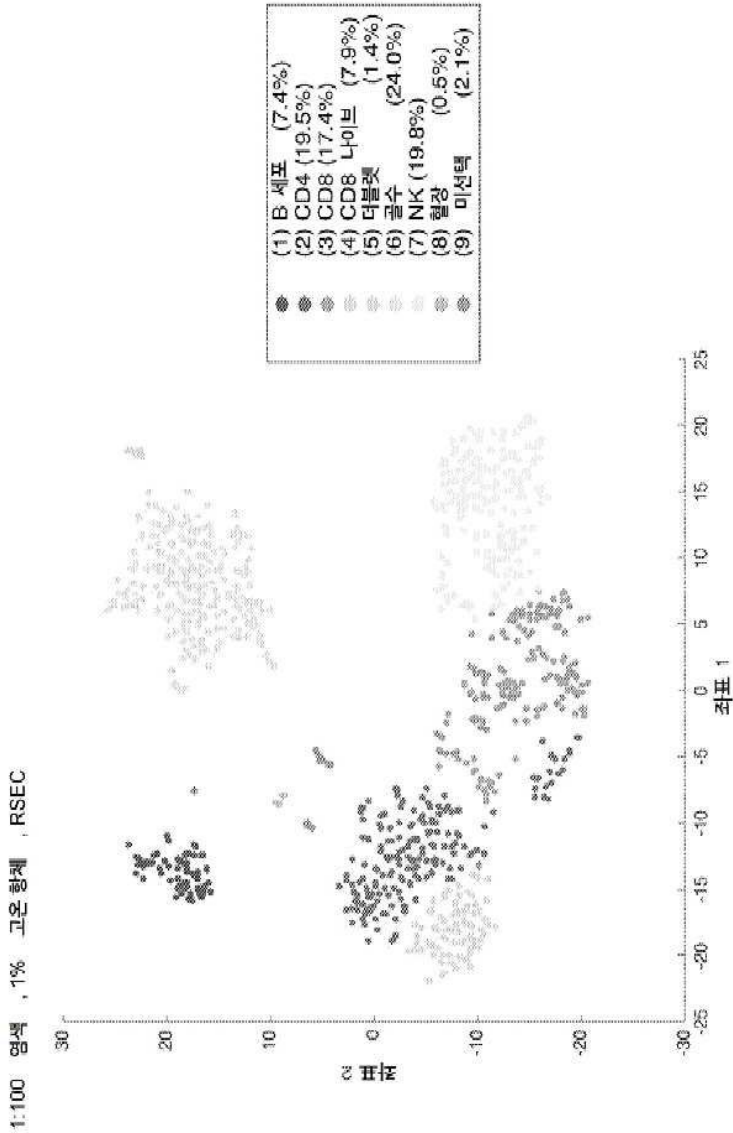
도면40b



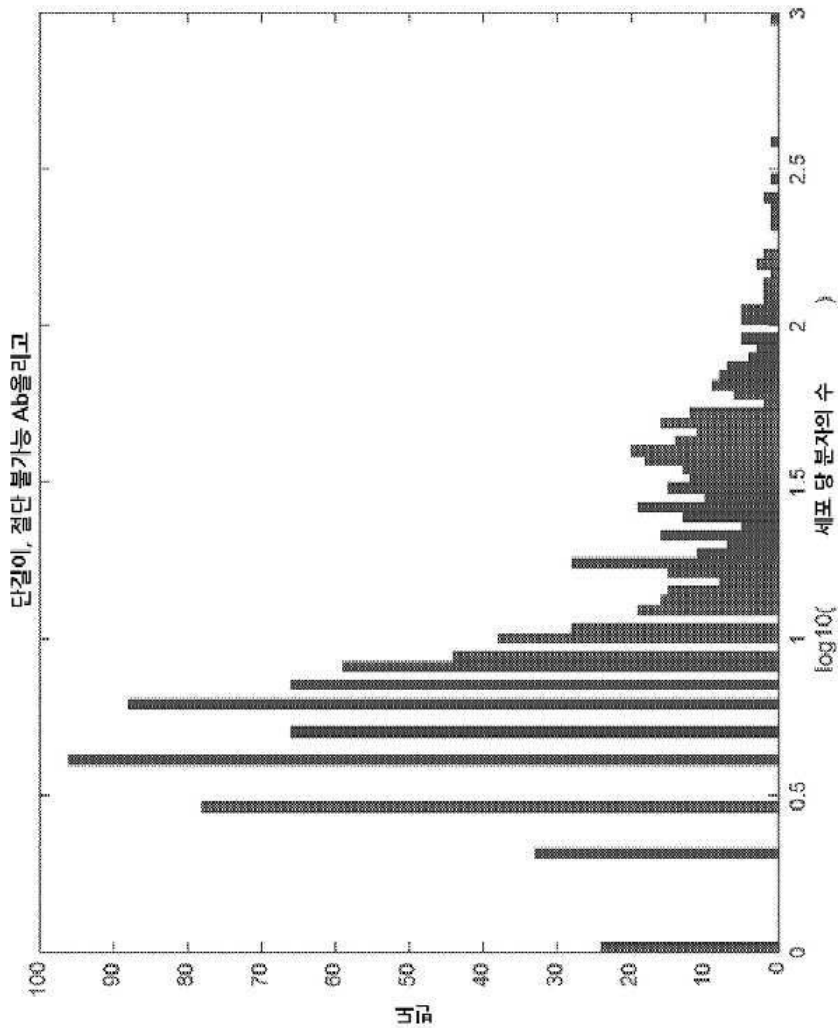
도면40c



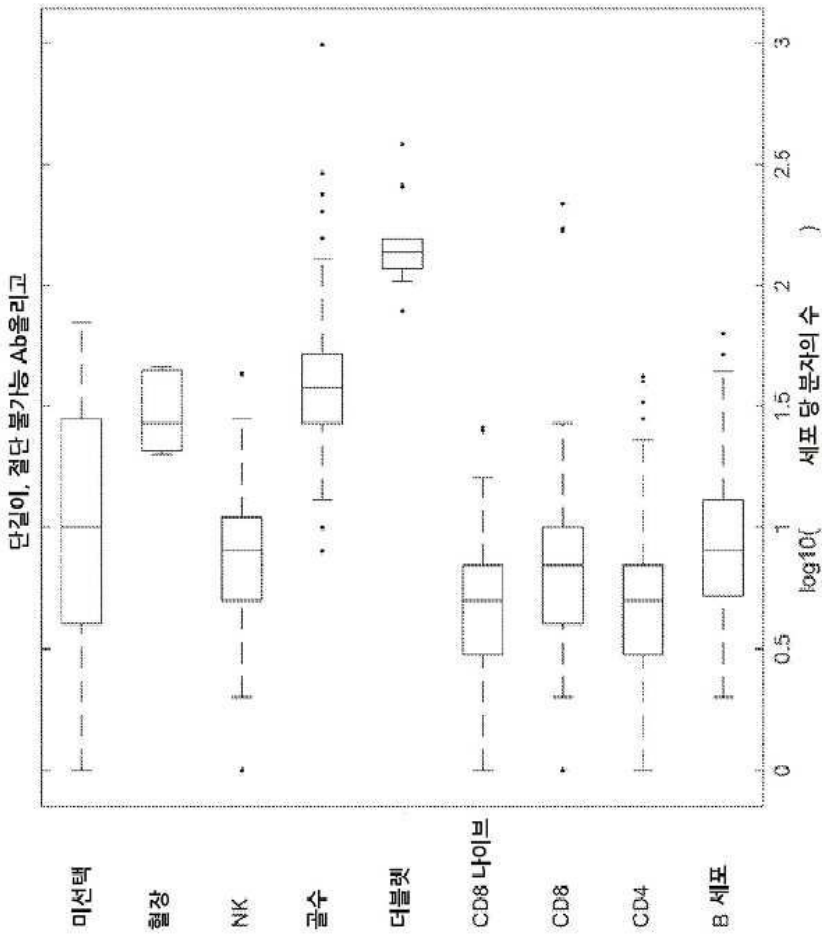
도면41a



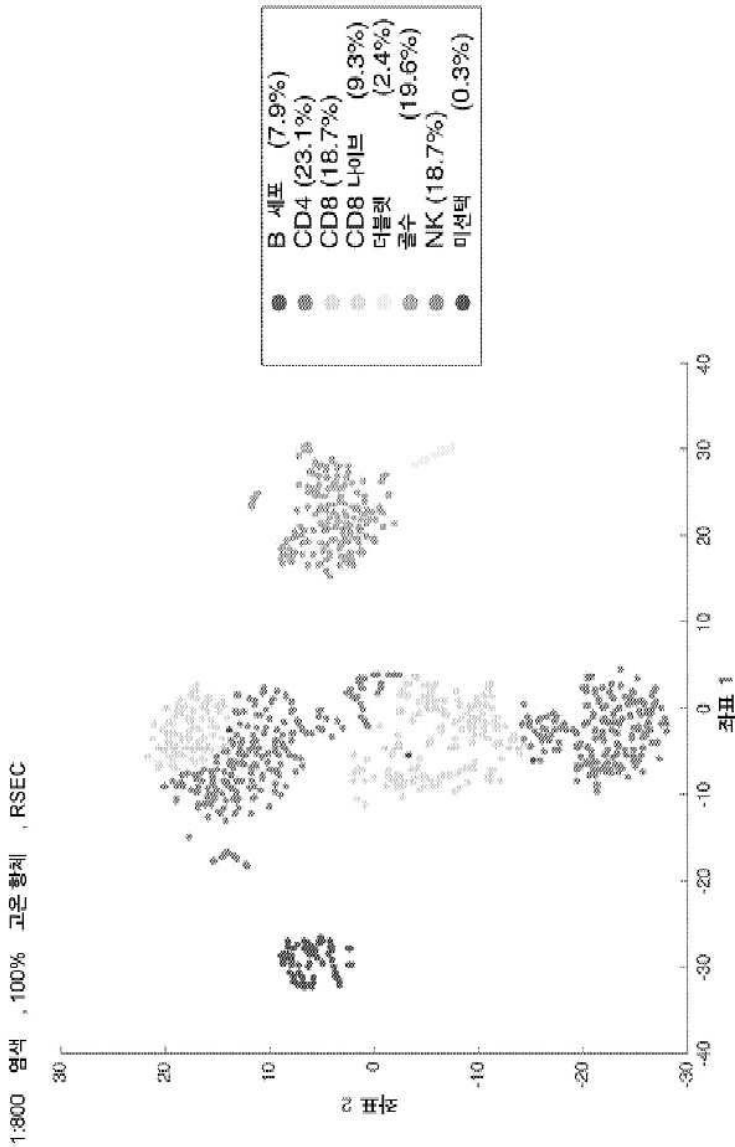
도면41b



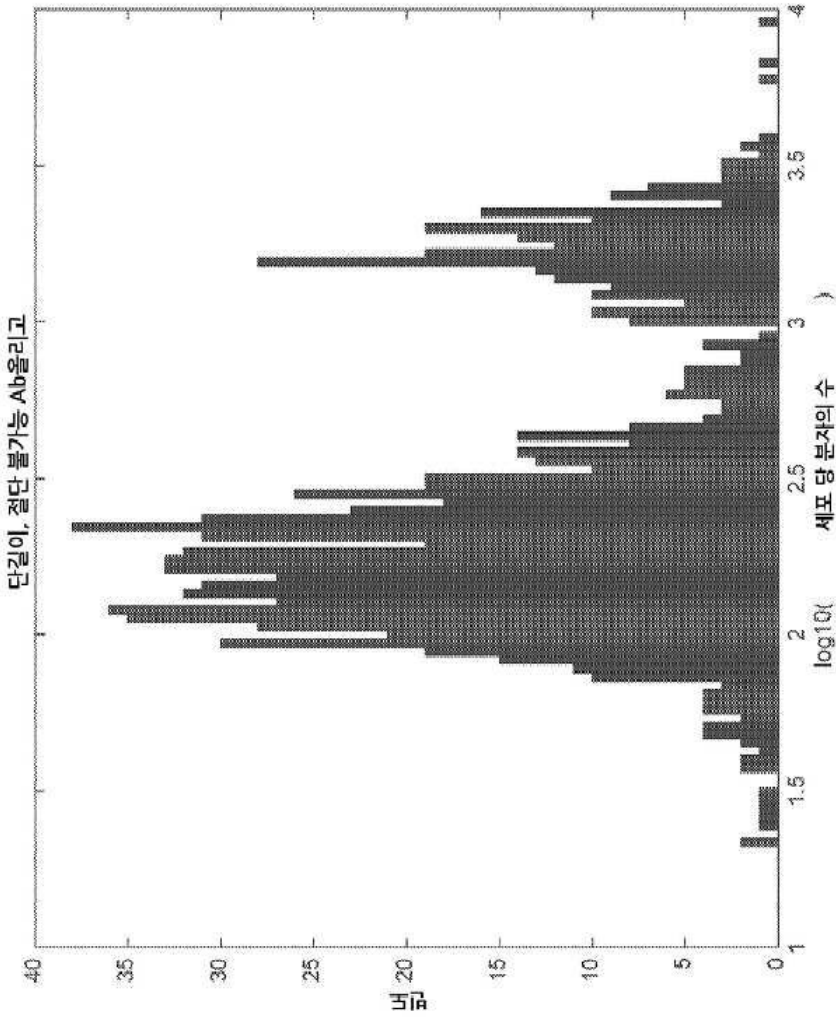
도면41c



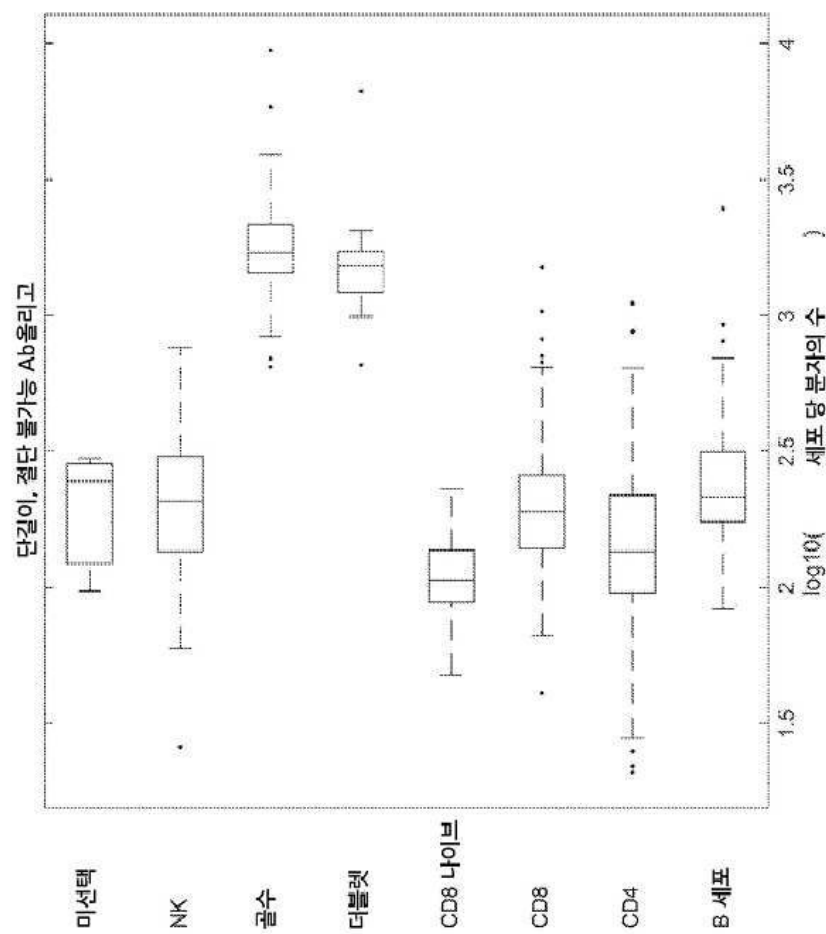
도면42a



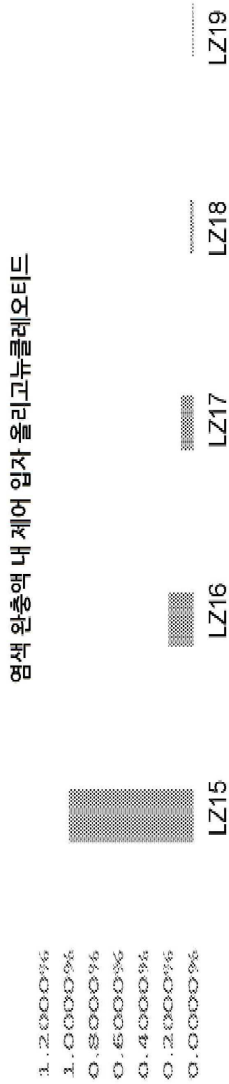
도면42b



도면42c

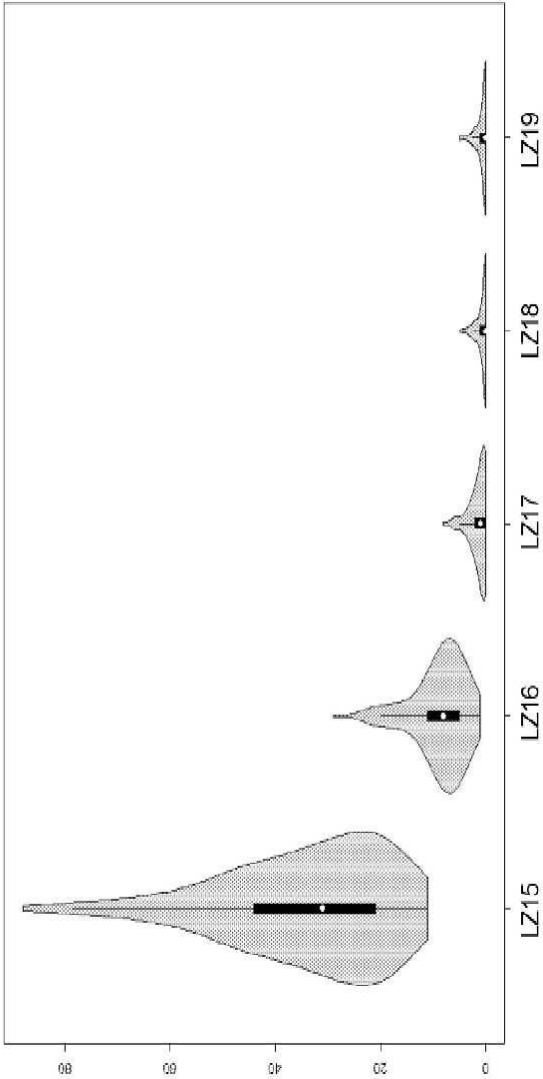


도면43a

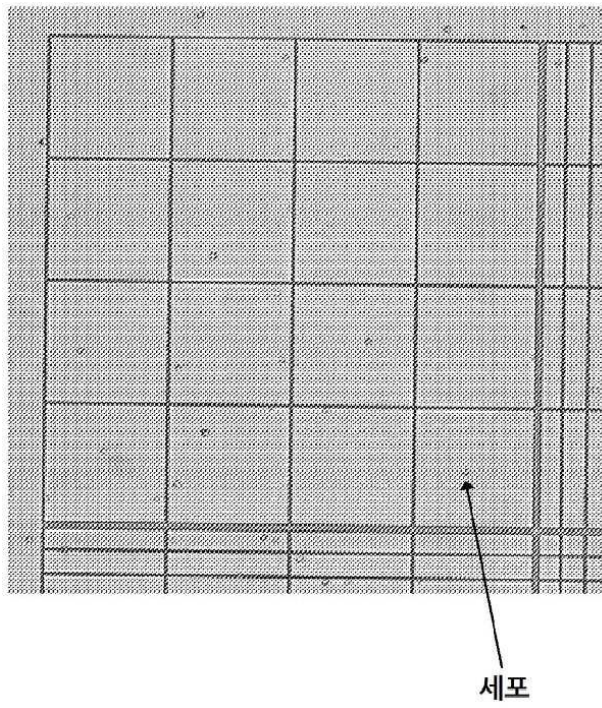


도면43b

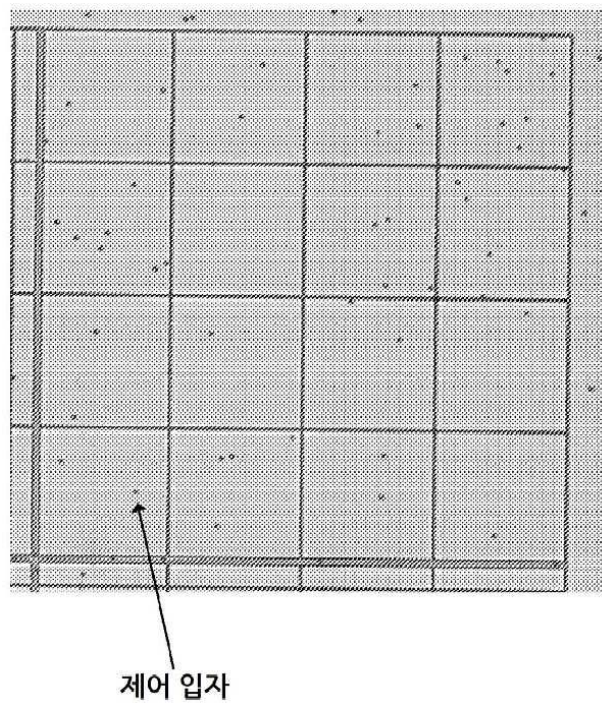
검출된 제어 입자 올리고뉴클레오타이드(올리고뉴클레오타이드의 함> 20)



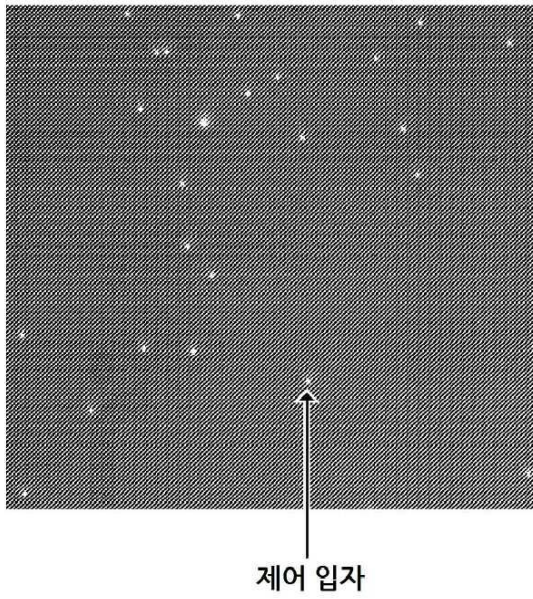
도면44a



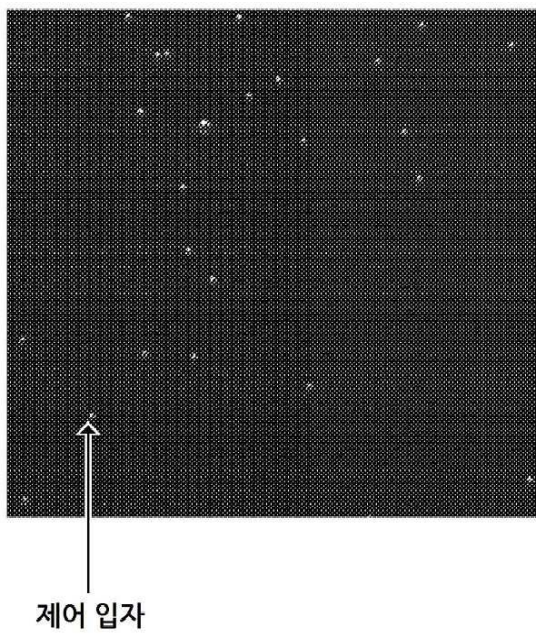
도면44b



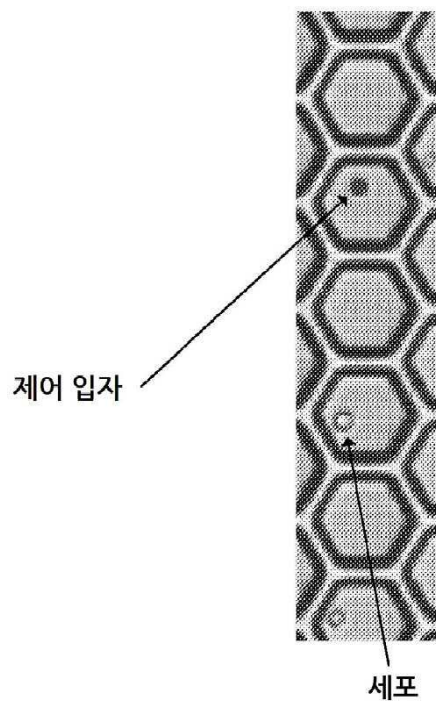
도면45a



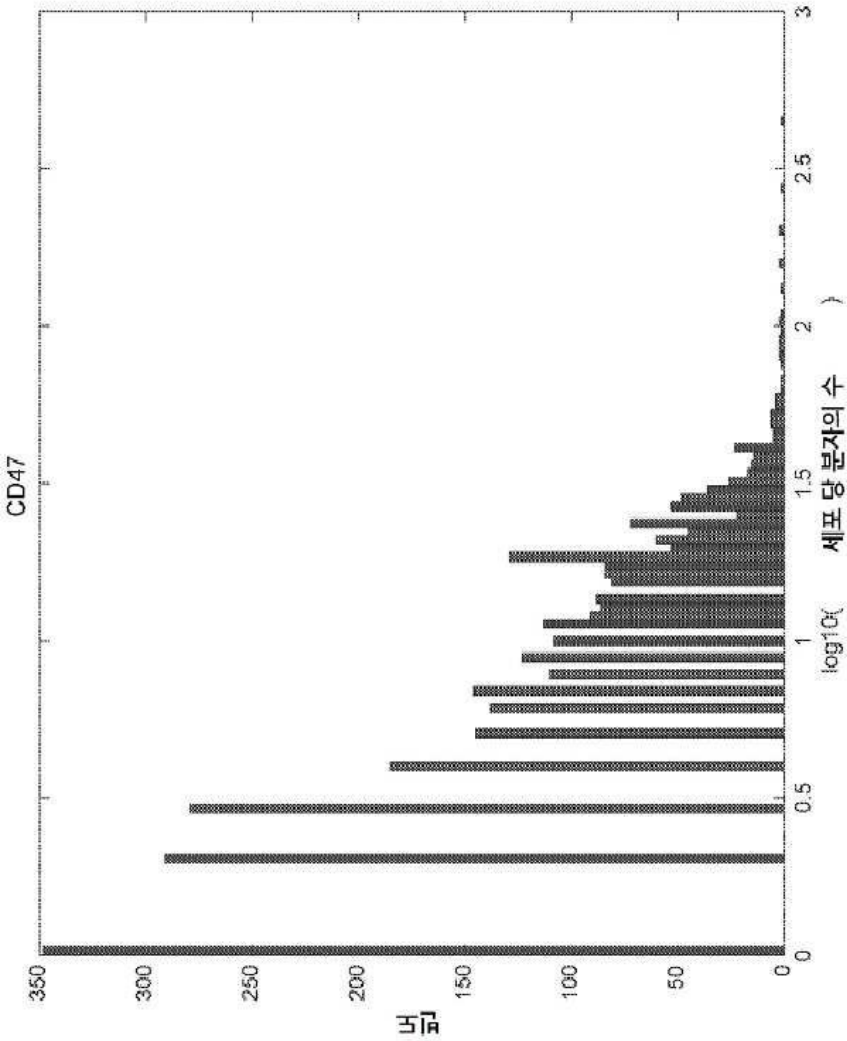
도면45b



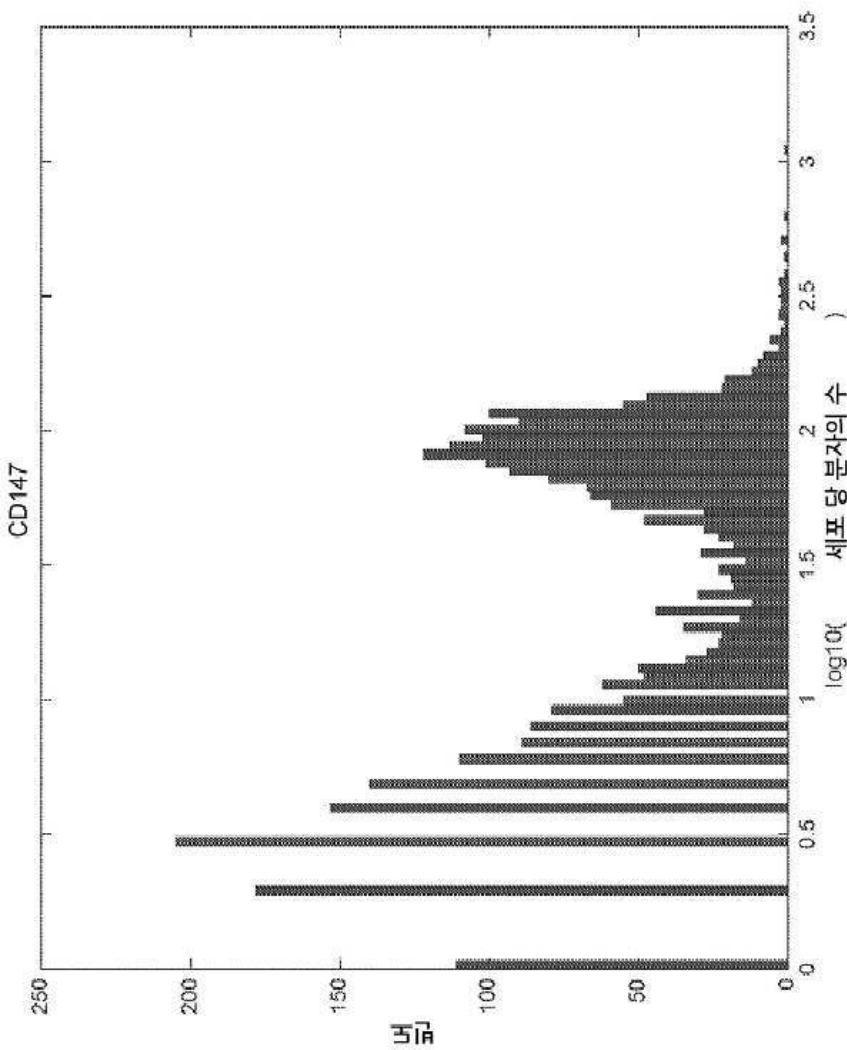
도면46



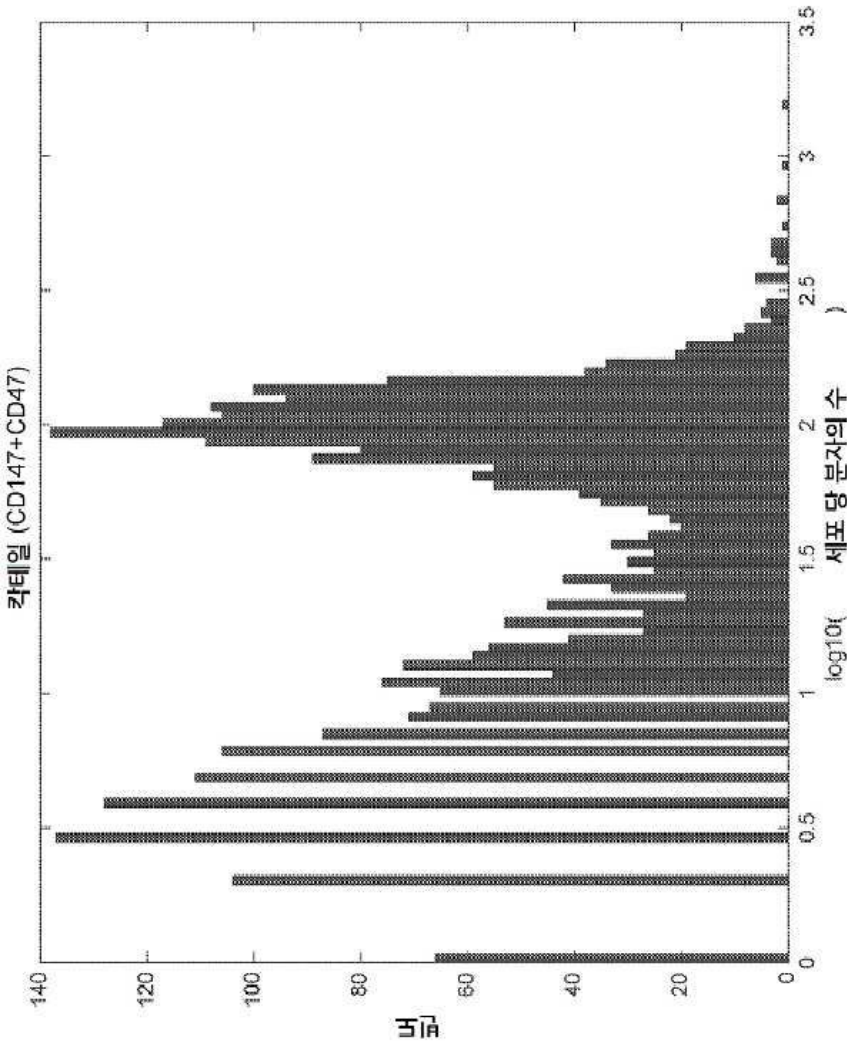
도면47a



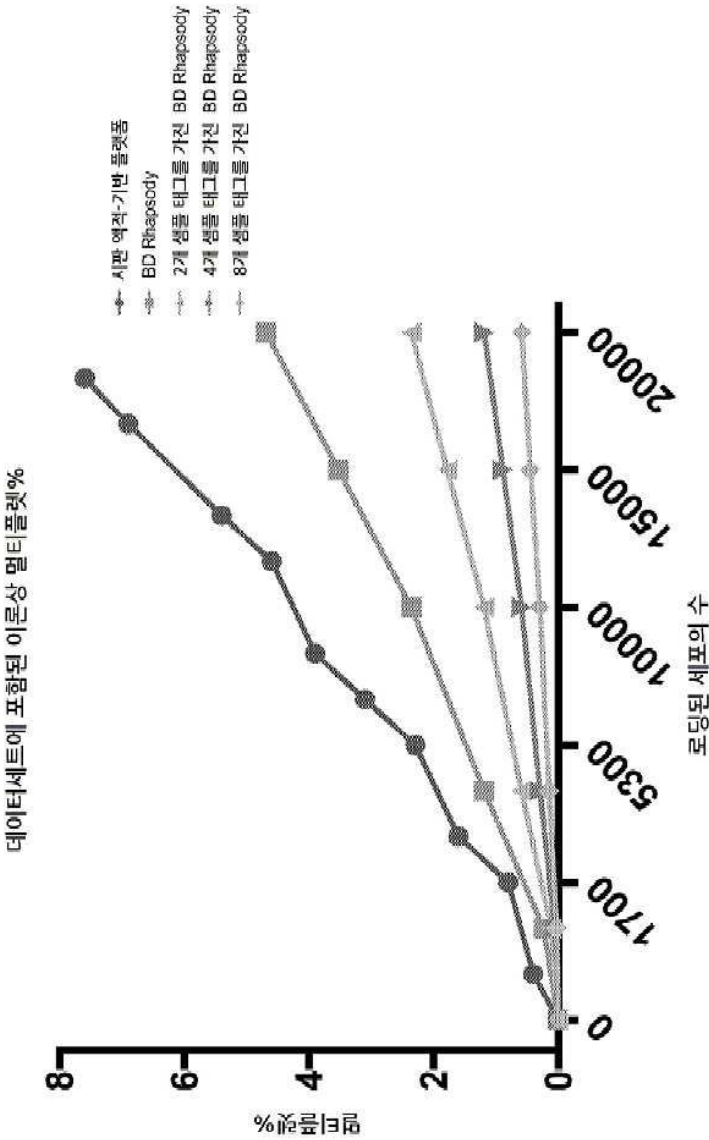
도면47b



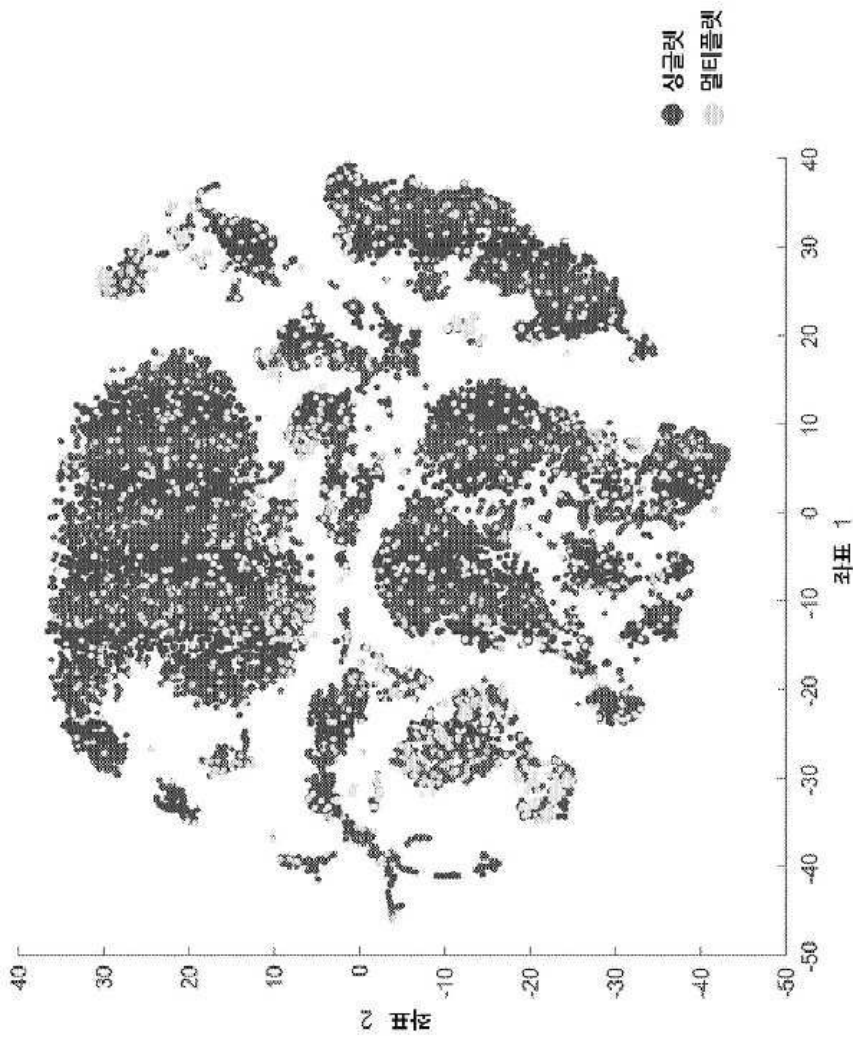
도면47c



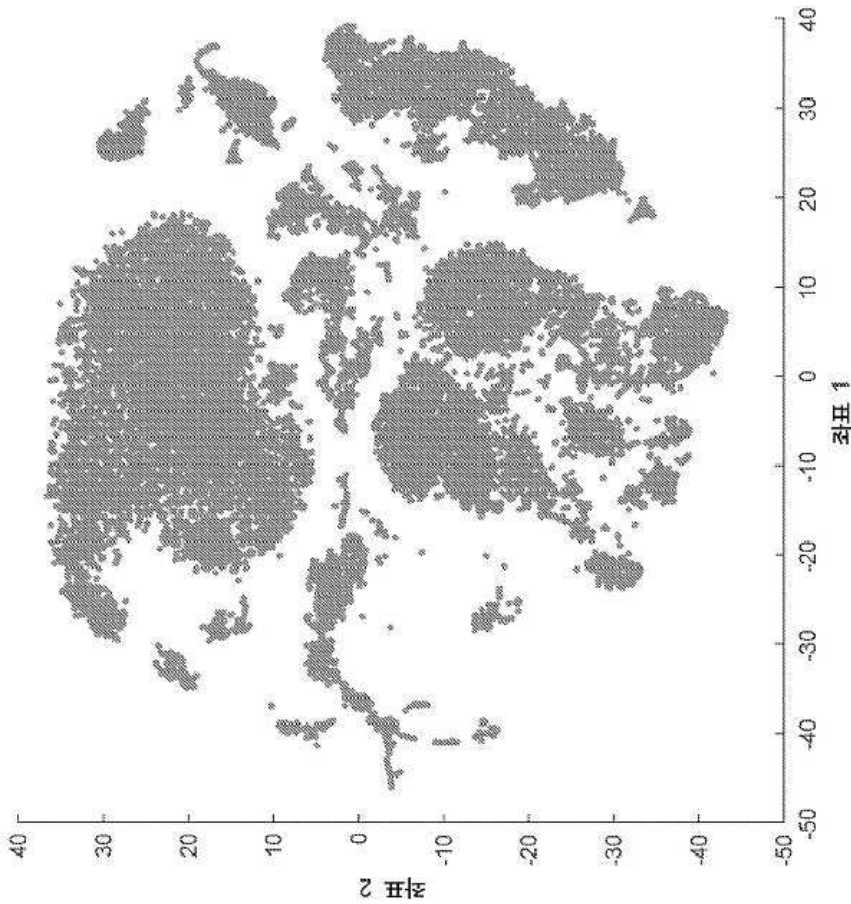
도면48



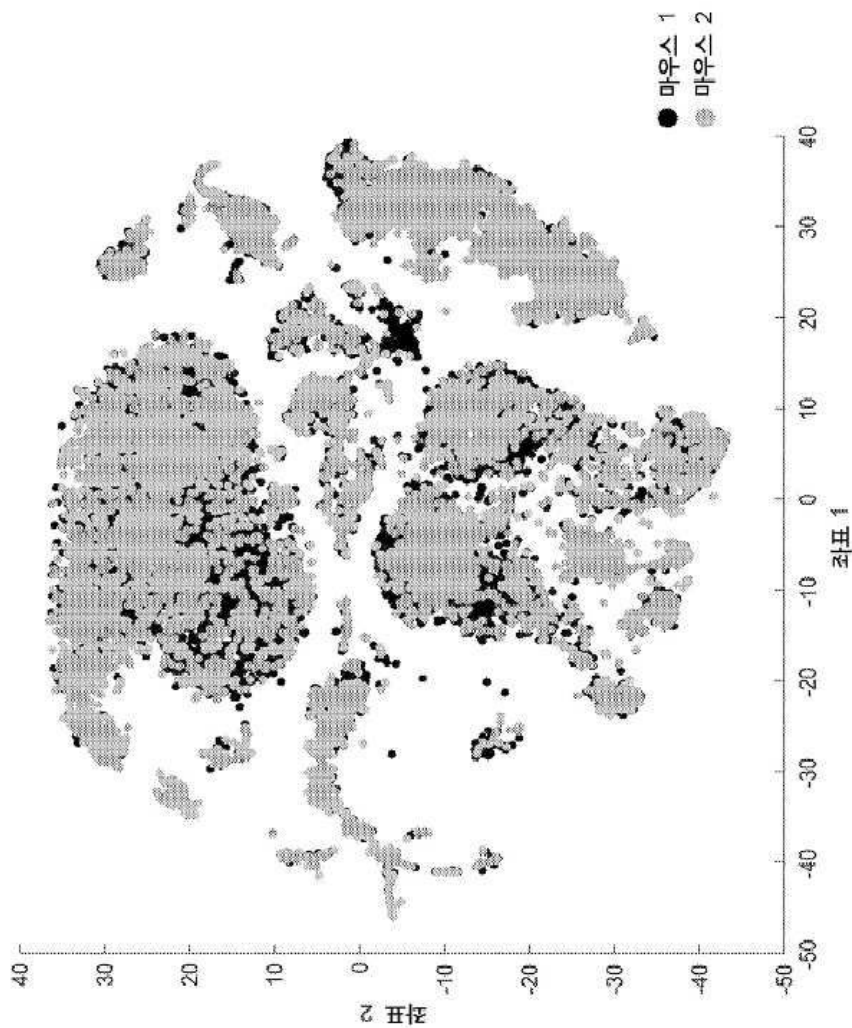
도면49a



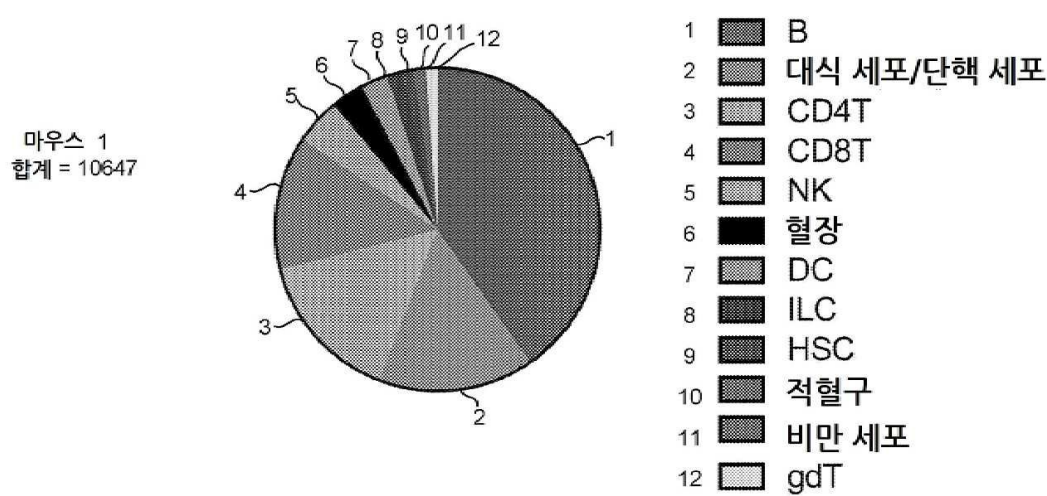
도면49b



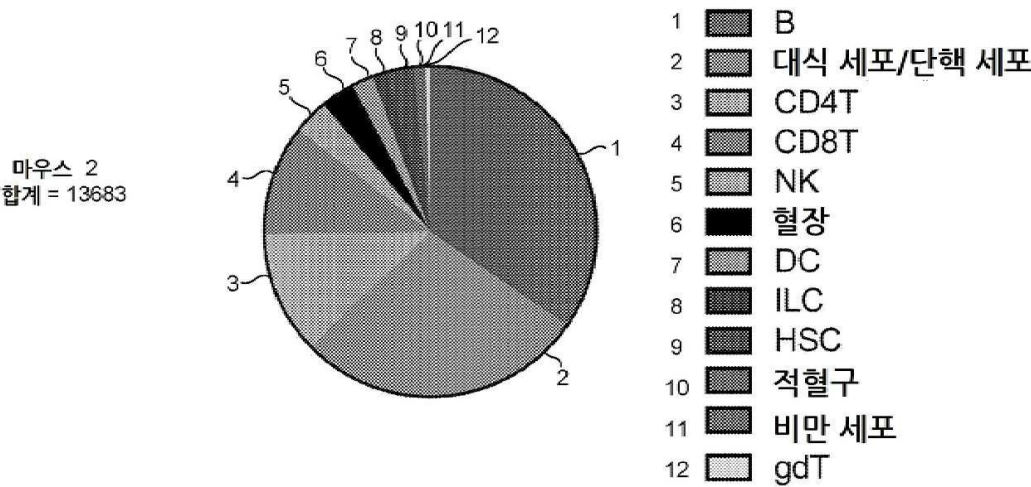
도면50a



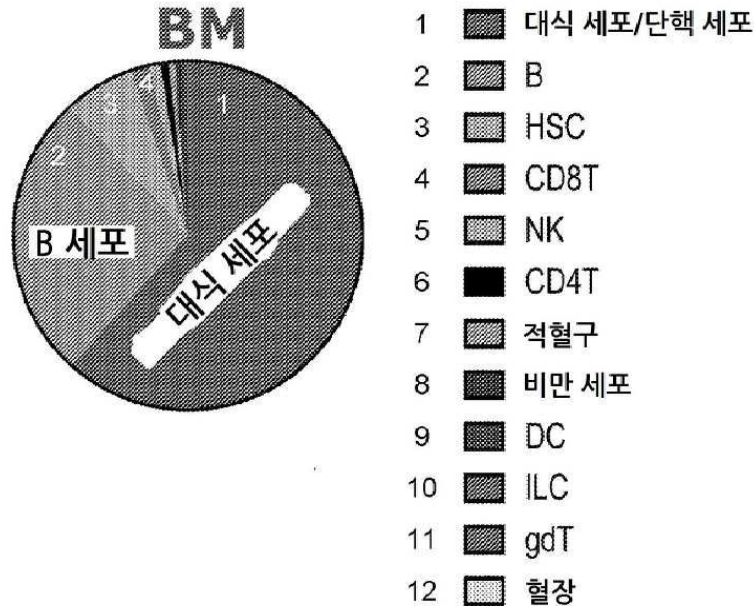
도면50b



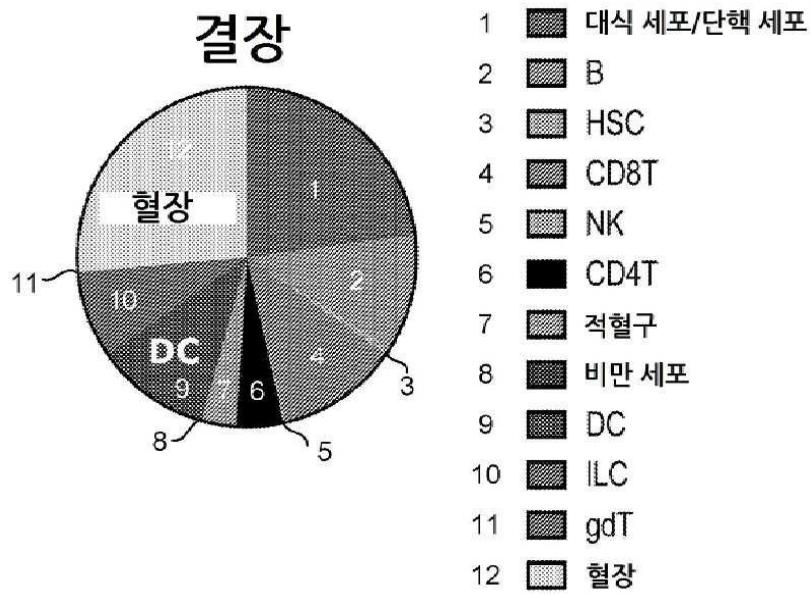
도면50c



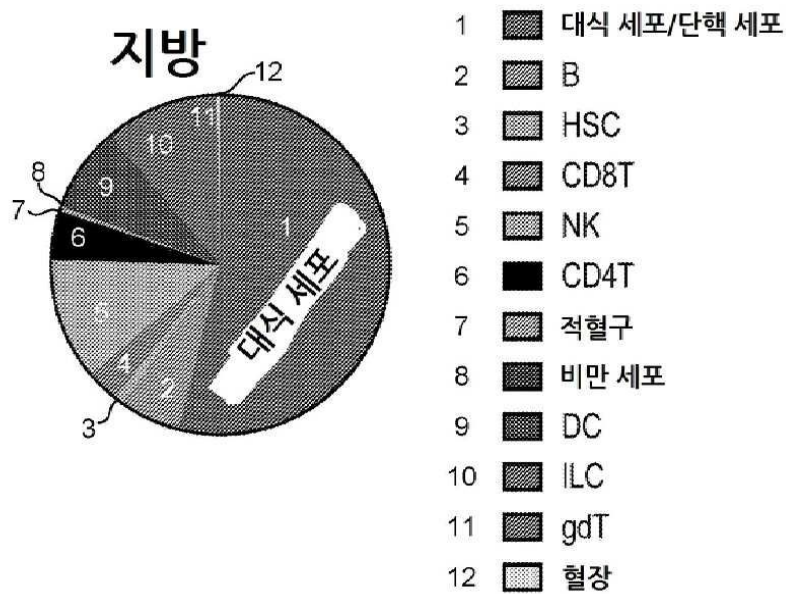
도면51a



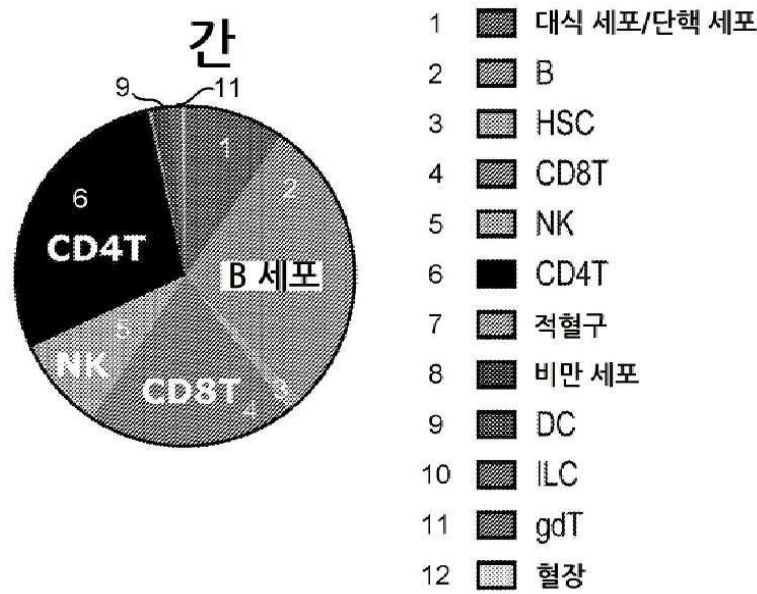
도면51b



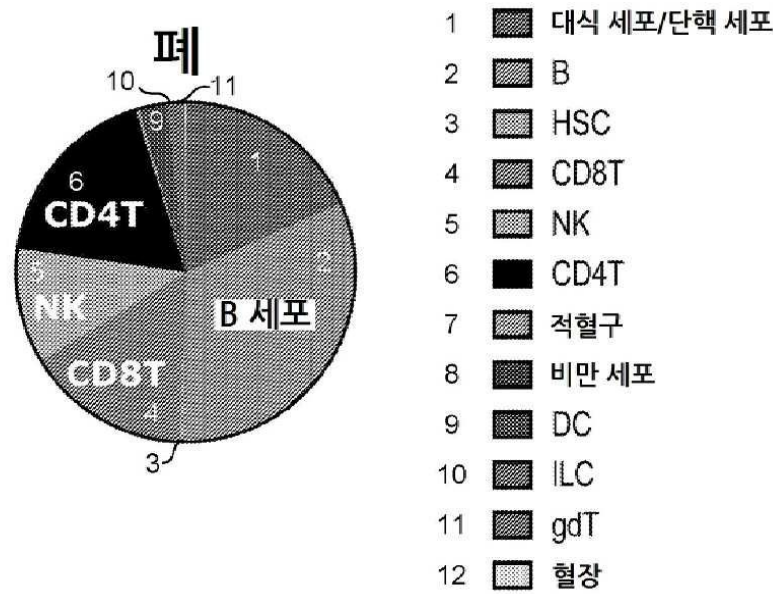
도면51c



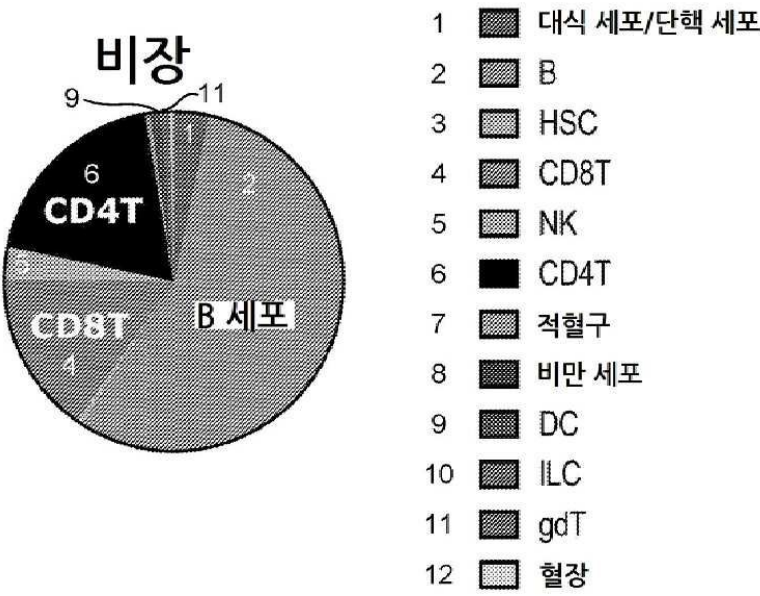
도면51d



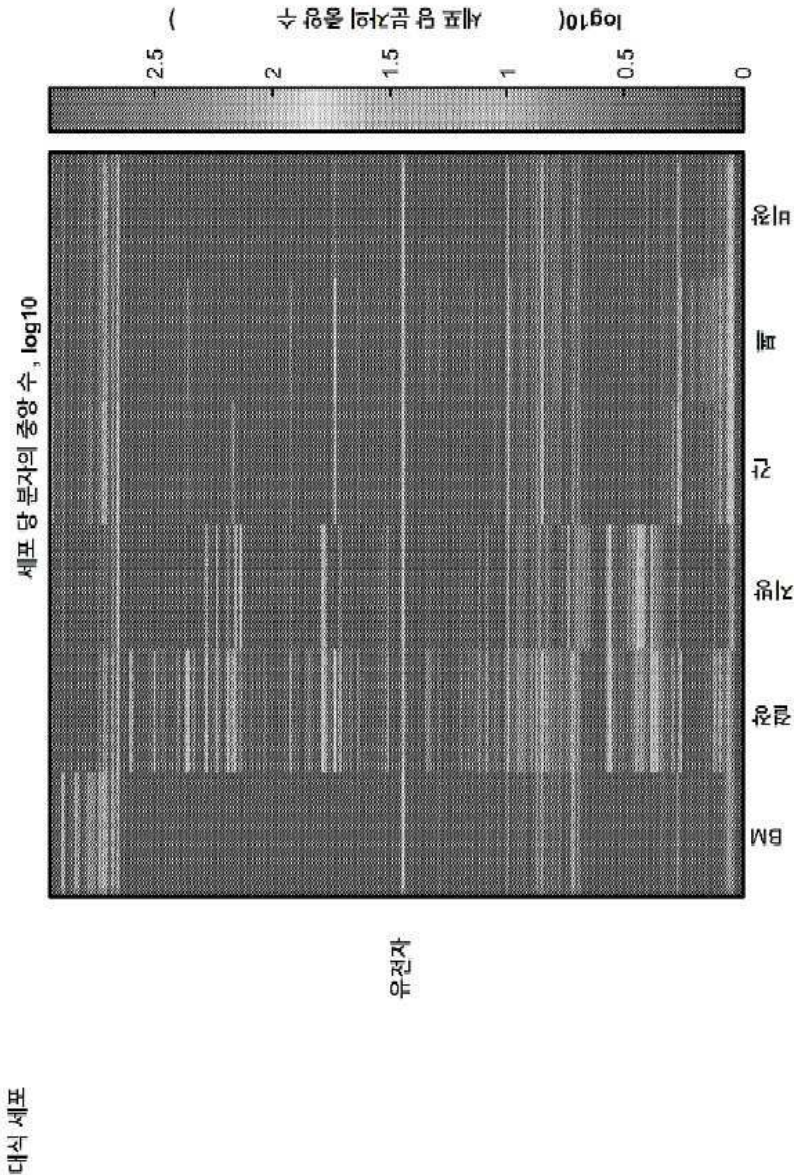
도면51e



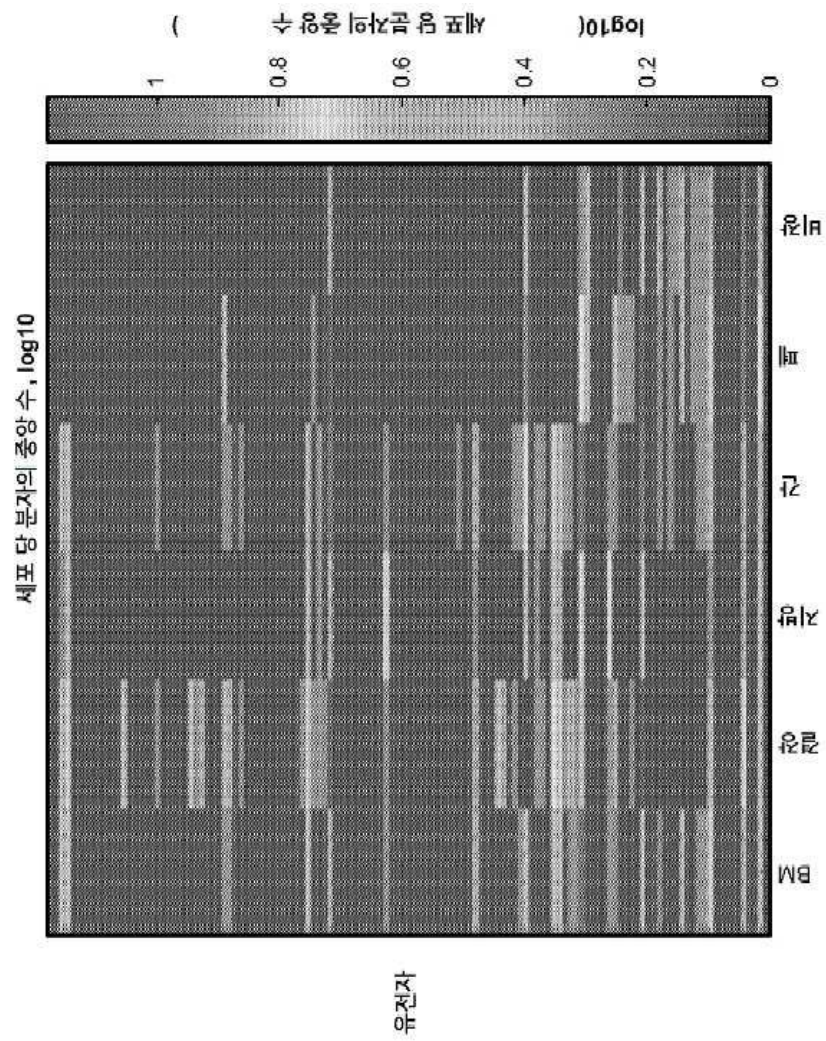
도면51f



도면52a

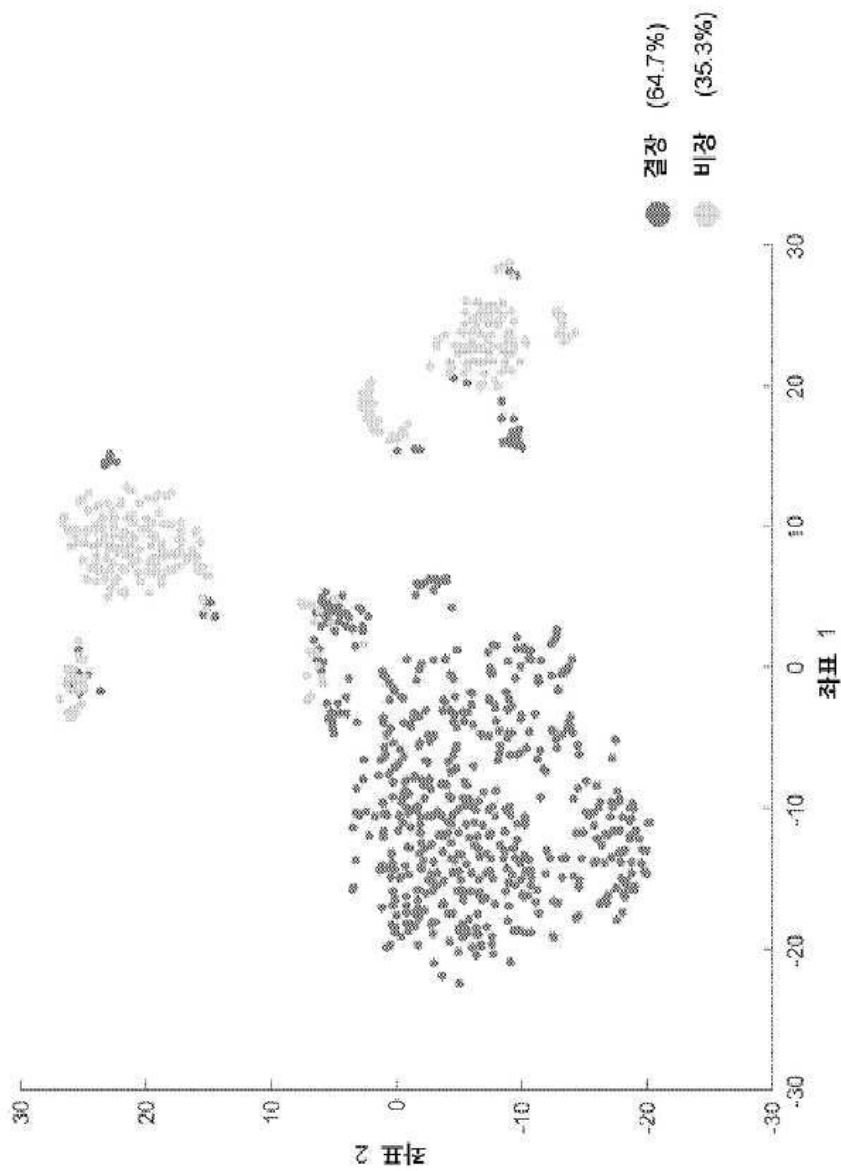


도면 52b

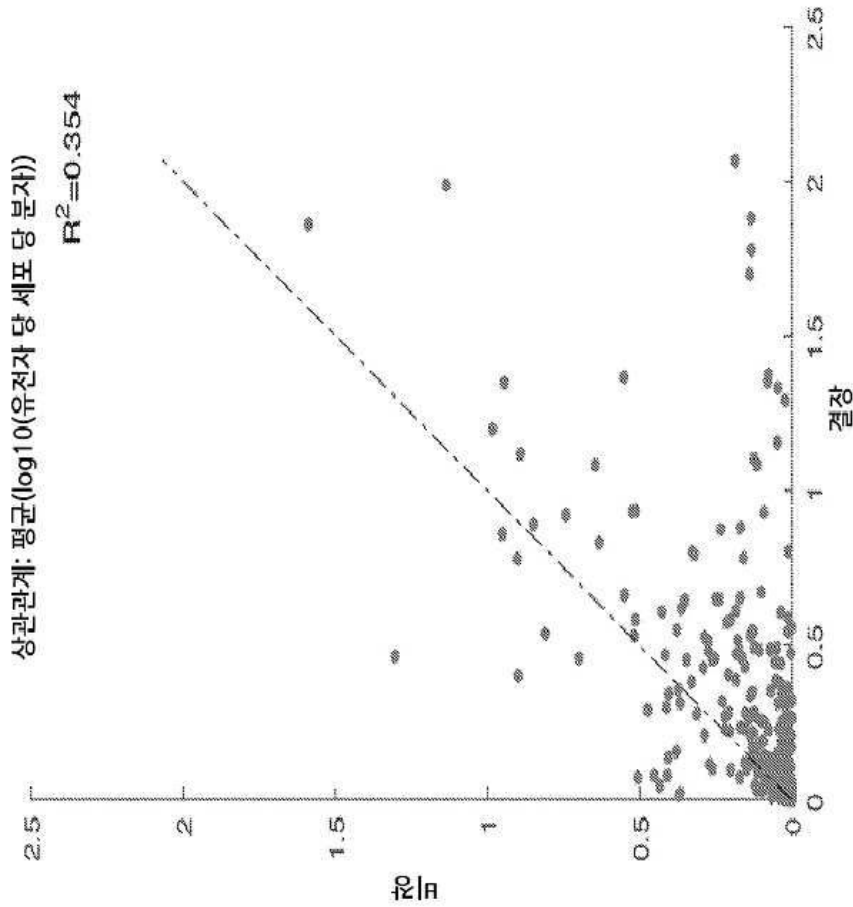


1. 서론

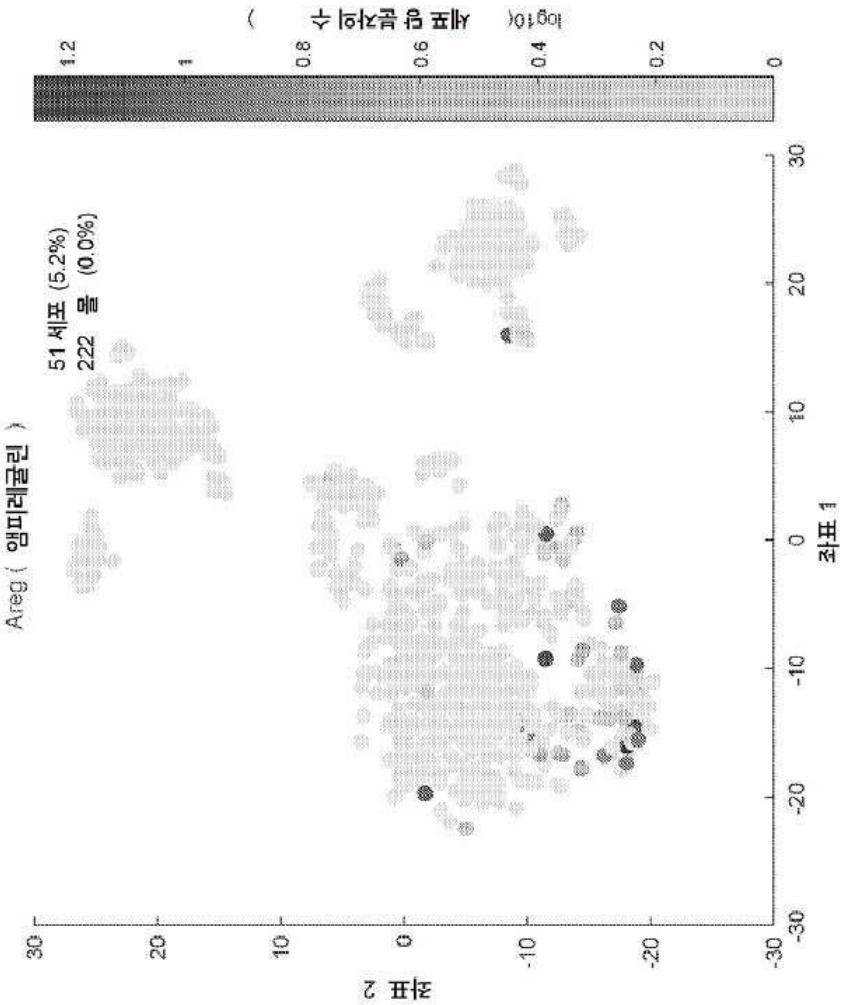
도면53a



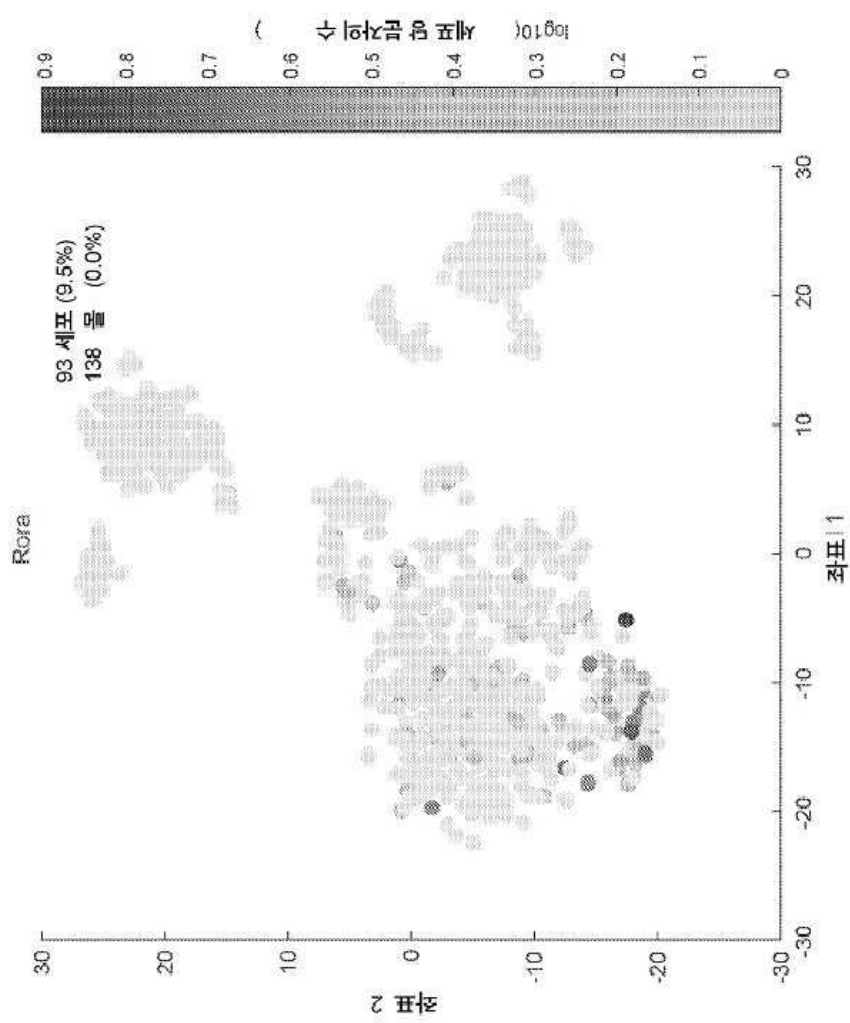
도면53b



도면53c



도면53d



도면54



서열 목록

SEQUENCE LISTING

<110> Cellular Research, Inc.
Chang, Christina
Fan, Christina
Shum, Eleen
Martin, Jody
Bansal, Nidhanjali
Ghadiali, James
Lazaruk, Katherine
Lam, Gretchen

<120> SAMPLE INDEXING FOR SINGLE CELLS
<130> BDCRI.033WO
<150> 62/515,285
<151> 2017-06-05

<150> 62/532,905

<151> 2017-07-14

<150> 62/532,949

<151> 2017-07-14

<150> 62/532,971

<151> 2017-07-14

<150> 62/554,425

<151> 2017-09-05

<150> 62/578,957

<151> 2017-10-30

<150> 62/645,703

<151> 2018-03-20

<160> 8

<170> Patent In version 3.5

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> note="Description of Artificial Sequence: Synthetic
Oligonucleotide"

<400> 1

aaaaaaaaa aaaaaaaaaa

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> note="Description of Artificial Sequence: Synthetic
Oligonucleotide"

<400> 2

tttttttttt tttttttttt

20

<210> 3

<211> 95

<212> DNA

<213> Artificial Sequence
 <220><223> note="Description of Artificial Sequence: Synthetic
 Oligonucleotide"
 <220><221> 5AmMC6
 <222> (1)..(1)
 <223> 5' Amino Modifier C6
 <400> 3
 gttgtcaaga tgctaccgtt cagagtacgt ggagttggtg gcccgacccc gagcgctacg 60
 agccccccgg aaaaaaaaaa aaaaaaaaaa aaaaa 95
 <210> 4
 <211> 200
 <212> DNA

<213> Artificial Sequence
 <220><223> note="Description of Artificial Sequence: Synthetic
 Oligonucleotide"
 <220><221> 5AmMC6
 <222> (1)..(1)
 <223> 5' Amino Modifier C6
 <400> 4
 gttgtcaaga tgctaccgtt cagagctact gtccgaagtt accgtgtatc taccacgggt 60
 ggtttttcga atccggaaaa gatagtaata agtgtttttag ttggaataag tcgcaacttt 120
 tggagacggg tacctctcaa tttttctgat ccgtaggccc cccgatctcg gcctcaaaaa 180
 aaaaaaaaaa aaaaaaaaaa 200
 <210>
 > 5
 <211> 25
 <212> DNA

<213> Artificial Sequence
 <220><223> note="Description of Artificial Sequence: Synthetic
 Oligonucleotide"
 <400> 5
 gttgtcaaga tgctaccgtt cagag 25
 <210> 6
 <211> 95

<212> DNA

<213> Artificial Sequence

<220><223> note="Description of Artificial Sequence: Synthetic
Oligonucleotide"

<400> 6

gttgtcaaga tgctaccgtt cagagcccca tgtctagtag ctattgggtcc cctatcctca 60

gattcgttta aaaaaaaaaa aaaaaaaaaa aaaaa 95

<210> 7

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> note="Description of Artificial Sequence: Synthetic
Oligonucleotide"

<400> 7

tttttttttt tttttttttt tttttt 26

<210> 8

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> note="Description of Artificial Sequence: Synthetic
Oligonucleotide"

<400> 8

acacgacgct cttccgatct 20