



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0050411  
(43) 공개일자 2018년05월14일

- (51) 국제특허분류(Int. Cl.)  
*C12N 15/10* (2017.01) *C12Q 1/6876* (2018.01)  
*C12Q 1/6886* (2018.01) *G01N 33/50* (2017.01)
- (52) CPC특허분류  
*C12N 15/1079* (2013.01)  
*C12N 15/1093* (2013.01)
- (21) 출원번호 10-2018-7010998
- (22) 출원일자(국제) 2016년09월16일  
 심사청구일자 없음
- (85) 번역문제출일자 2018년04월18일
- (86) 국제출원번호 PCT/US2016/052336
- (87) 국제공개번호 WO 2017/049231  
 국제공개일자 2017년03월23일
- (30) 우선권주장  
 62/220,879 2015년09월18일 미국(US)  
 (뒷면에 계속)

- (71) 출원인  
 트위스트 바이오사이언스 코포레이션  
 미국, 캘리포니아 94158, 샌프란시스코, 스위트 545, 사우쓰, 미션 베이 불러바드 455
- (72) 발명자  
 콕스 앤서니  
 미국 94040 캘리포니아주 마운틴 뷰 플라자 코트 1768  
 트로이쉬 세바스티안  
 미국 94134 캘리포니아주 샌프란시스코 울지 스트리트 933  
 천 쓰위엔  
 미국 94402 캘리포니아주 샌 마테오 애버딘 드라이브 1302
- (74) 대리인  
 김진희, 김태홍

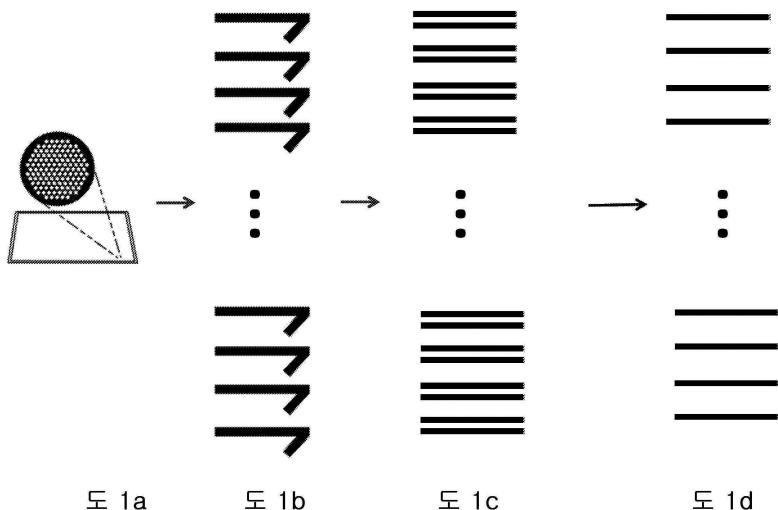
전체 청구항 수 : 총 35 항

(54) 발명의 명칭 올리고핵산 변이체 라이브러리 및 그의 합성

### (57) 요약

본원에서는 핵산 서열의 소정의 변이체를 코딩하는 고도로 정확한 올리고핵산 라이브러리를 생성하는 방법을 개시한다. 변이 정도는 완전할 수 있고, 이로써 포화된 변이체 라이브러리를 수득할 수 있거나, 또는 덜 완전할 수 있고, 이로써, 변이체의 선택적 라이브러리를 수득할 수 있다. 본원에 기술된 변이체 올리고핵산 라이브러리는 전사 또는 번역에 의해 추가 프로세싱을 위해 디자인될 수 있다. 본원에 기술된 변이체 올리고핵산 라이브러리는 변이체 RNA, DNA 및/또는 단백질 집단 생성을 위해 디자인될 수 있다. 본원에서는 생물학적 기능 조절, 및 질환 치료용 또는 감소용 치료제 디자인에서 응용되는, 활성이 증가 또는 감소된 변이체 종을 확인하는 방법을 추가로 제공한다.

### 대 표 도



(52) CPC특허분류

*C12N 15/113* (2013.01)  
*C12Q 1/6876* (2018.05)  
*C12Q 1/6886* (2018.05)  
*G01N 33/502* (2013.01)  
*G01N 33/5029* (2013.01)  
*G01N 33/5041* (2013.01)  
*C12N 2320/30* (2013.01)  
*C12N 2330/31* (2013.01)  
*C12Q 2600/158* (2013.01)

---

(30) 우선권주장

62/263,548 2015년12월04일 미국(US)  
62/354,034 2016년06월23일 미국(US)

## 명세서

### 청구범위

#### 청구항 1

- a) 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 변이체 코돈 서열을 코딩하는 것인 단계;
- b) 표면을 갖는 구조물을 제공하는 단계;
- c) 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계;
- d) 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계;
- e) 변이체 핵산의 라이브러리를 세포로 전달하고, 복수의 변이체 단백질을 발현시키는 단계; 및
- f) 복수의 변이체 단백질들 중 한 변이체 단백질과 연관된 활성을 확인하는 단계로서, 상기 활성을 단일 기준 서열에 의해 코딩된 단백질과 비교하여 조절되는 것인 단계

를 포함하는, 단백질 활성을 조절하는 방법.

#### 청구항 2

제1항에 있어서, 활성이 세포의 생식, 성장, 부착, 사멸, 이동, 에너지 생산, 산소 이용률, 대사 활성, 세포 신호전달, 노화, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합을 포함하는 것인 방법.

#### 청구항 3

제1항에 있어서, 세포가 진핵 세포 또는 원핵 세포인 방법.

#### 청구항 4

제1항에 있어서, 세포가 박테리아, 식물, 마우스 또는 영장류의 세포인 방법.

#### 청구항 5

제1항에 있어서, 변이체 핵산의 라이브러리가 변이체 유전자 또는 그의 단편에 대한 서열을 코딩하는 것인 방법.

#### 청구항 6

제1항에 있어서, 변이체 단백질이 항체, 효소 또는 웹티드인 방법.

#### 청구항 7

제1항에 있어서, 변이체 단백질이 또 다른 분자에 대하여 증진되거나 감소된 결합 친화도를 갖는 것인 방법.

#### 청구항 8

제1항에 있어서, 변이체 단백질은 증진되거나 감소된 효소 활성을 갖는 것인 방법.

#### 청구항 9

제1항에 있어서, 변이체 핵산의 라이브러리에 의해 코딩된 단일 변이체 핵산을, 그를 필요로 하는 피험체에게 투여하는 단계를 추가로 포함하는 방법.

#### 청구항 10

제1항에 있어서, 약 30,000 이상의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한

소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 방법.

### 청구항 11

- a) 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 소정의 변이체 서열을 코딩하는 것인 단계;
  - b) 표면을 갖는 구조물을 제공하는 단계;
  - c) 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되고, 약 30,000 이상의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 단계;
  - d) 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계;
  - e) 변이체 핵산의 라이브러리를 제1 세포 세트로 전달하는 단계; 및
  - f) 세포 활성의 변화를 측정하는 단계로서, 세포 활성은 제1 세포 세트 또는 제2 세포 세트에 대하여 측정하고, 제2 세포 세트는 제1 세포 세트로부터 단리된 적어도 하나의 발현 생성물로 처리된 것인 단계
- 를 포함하는, 세포 활성을 조절하는 방법.

### 청구항 12

제11항에 있어서, 세포 활성이 생식, 성장, 부착, 사멸, 이동, 에너지 생산, 산소 이용률, 대사 활성, 세포 신호전달, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합을 포함하는 것인 방법.

### 청구항 13

제11항에 있어서, 제1 세포 세트 또는 제2 세포 세트가 진핵 세포 또는 원핵 세포인 방법.

### 청구항 14

제11항에 있어서, 제1 세포 세트 또는 제2 세포 세트가 박테리아, 식물, 마우스 또는 영장류의 세포인 방법.

### 청구항 15

제11항에 있어서, 변이체 핵산의 라이브러리가 변이체 유전자 또는 그의 단편에 대한 서열을 코딩하는 것인 방법.

### 청구항 16

제15항에 있어서, 각각의 변이체 유전자를 변역하여 단백질을 형성하는 단계를 추가로 포함하는 방법.

### 청구항 17

제16항에 있어서, 단백질이 항체, 효소 또는 펩티드인 방법.

### 청구항 18

제17항에 있어서, 변이체 핵산의 라이브러리가 항체의 가변 영역 또는 불변 영역의 적어도 일부를 코딩하는 것인 방법.

### 청구항 19

제18항에 있어서, 변이체 핵산의 라이브러리가 항체의 적어도 하나의 CDR 영역을 코딩하는 것인 방법.

### 청구항 20

제16항에 있어서, 단백질이 또 다른 분자에 대하여 증진되거나 감소된 결합 친화도를 갖거나, 또는 단백질이 증진되거나 감소된 효소 활성을 갖는 것인 방법.

**청구항 21**

제11항에 있어서, 변이체 핵산의 라이브러리가 전사 조절 서열을 코딩하는 것인 방법.

**청구항 22**

제21항에 있어서, 전사 조절 서열이 프로모터, UTR 또는 터미네이터 서열인 방법.

**청구항 23**

제11항에 있어서, 변이체 핵산의 라이브러리는, 전사될 때 mRNA, miRNA 또는 shRNA를 코딩하는 서열을 코딩하는 것인 방법.

**청구항 24**

- a) 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 변이체 코돈 서열을 코딩하는 것인 단계;
  - b) 표면을 갖는 구조물을 제공하는 단계;
  - c) 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계;
  - d) 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계;
  - e) 변이체 핵산의 라이브러리를 피험체로부터 수득된 세포로 전달하는 단계;
  - f) 변이체 핵산의 라이브러리에 의해 코딩된 변이체 핵산과 연관된 유해 활성의 감소를 확인하는 단계; 및
  - g) 변이체 핵산의 라이브러리에 의해 코딩된 변이체 핵산을, 그를 필요로 하는 피험체에게 투여하여, 질환 상태를 치료하거나 경감하는 단계
- 를 포함하는, 질환 상태를 치료하거나 경감하는 방법.

**청구항 25**

제24항에 있어서, 질환 상태가 세포 증식성 장애, 자가면역 장애, 바이러스성 장애 또는 박테리아성 장애와 연관된 것인 방법.

**청구항 26**

제24항에 있어서, 약 30,000 이상의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 방법.

**청구항 27**

- a) 복수의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 동일하지 않은 올리고핵산 각각의 길이는 20개 이상의 염기이고, 상기 복수의 동일하지 않은 올리고핵산은, 적어도 하나의 서열에 대한 3개 이상의 코돈 각각에 대해 최대 19종의 변이체를 코딩하며, 상기 복수의 동일하지 않은 올리고핵산은 1 이상의 유전자 및 그의 변이체를 집합적으로 코딩하는 것인 단계;
  - b) 표면을 갖는 구조물을 제공하는 단계;
  - c) 복수의 동일하지 않은 올리고핵산을 합성하는 단계로서, 상기 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계; 및
  - d) 복수의 동일하지 않은 올리고핵산으로부터 변이체 핵산의 라이브러리를 조립하는 단계
- 를 포함하는 핵산 합성 방법.

**청구항 28**

제27항에 있어서, 복수의 동일하지 않은 올리고핵산이 75,000 이상의 동일하지 않은 올리고핵산을 포함하는 것인 방법.

#### 청구항 29

제27항에 있어서, 단일 유전자 및 그의 변이체를 집합적으로 코딩하는 복수의 동일하지 않은 올리고핵산의 서브세트가 상기 구조물 표면 상의 단일 클러스터 내에 위치하는 것인 방법.

#### 청구항 30

제29항에 있어서, 구조물 표면이 6,000개 이상의 단일 클러스터를 포함하는 것인 방법.

#### 청구항 31

제30항에 있어서, 각각의 클러스터는 직경이 약 0.5 내지 2 mm인 채널 내에 위치하는 것인 방법.

#### 청구항 32

제29항에 있어서, 단일 클러스터가 핵산 신장을 위한 50 내지 500개의 좌위를 포함하는 것인 방법.

#### 청구항 33

제27항에 있어서, 복수의 동일하지 않은 올리고핵산이 1종 초과의 유전자의 변이체를 집합적으로 코딩하는 것인 방법.

#### 청구항 34

제27항에 있어서, 복수의 동일하지 않은 올리고핵산이 5,000종 이상의 유전자의 변이체를 집합적으로 코딩하는 것인 방법.

#### 청구항 35

제27항에 있어서, 변이체 핵산의 라이브러리가 효소, 펩티드 또는 항체의 적어도 일부를 코딩하는 것인 방법.

### 발명의 설명

#### 기술 분야

[0001]

#### 상호 참조

[0002]

본 출원은 2015년 9월 18일에 제출된 미국 가출원 제62/220,879호, 2015년 12월 4일에 제출된 미국 가출원 제62/263,548호, 및 2016년 6월 23일에 제출된 미국 가출원 제62/354,034호를 우선권으로 주장하며, 상기 가출원들은 모두 본원에 참고로 인용되어 있다.

[0003]

#### 서열 목록

[0004]

본 출원은 ASCII 포맷으로 전자 제출된 서열 목록을 포함하고, 이 서열 목록은 그 전체가 본원에 참고로 인용되어 있다. 2016년 9월 12일에 생성된 상기 ASCII 사본의 명칭은 44854\_718\_601\_SL.txt이고, 그 크기는 7,841 바이트이다.

### 배경 기술

[0005]

합성 생물학의 기초는, 맞춤식 경로 및 유기체의 신속하고 적절한 생성 및 최적화를 위해 접근가능화되어야 하는 디자인, 구축, 및 시험 프로세스(DNA를 필요로 하는 반복 프로세스)이다. 디자인 단계에서, DNA를 구성하는 A, C, T 및 G 뉴클레오티드는 관심의 대상의 되는 좌위 또는 경로를 구성하게 되는 다양한 유전자 서열로 제조되고, 여기서 각 서열 변이체는 시험하고자 하는 구체적인 가설을 나타낸다. 이들 변이체 유전자 서열은 진화 생물학에서 기원이 된 개념인 서열 공간의 서브세트이며, 이는 유전자, 게놈, 트랜스크립트 및 프로테옴을 구성하는 서열 전체에 관한 것이다.

[0006]

다수의 상이한 변이체는 전형적으로는 서열 공간의 샘플링이 적절히 이루어질 수 있도록 하고, 최적화된 디자인의 가능성을 최대화시키기 위하여 각각의 디자인-구축-시험 사이클을 위해 디자인된다. 비록 개념은 간단하기는

하지만, 종래 합성 방법의 속도, 처리량 및 품질에 관한 프로세스상의 장애가 상기 사이클 진행 속도를 감소시키고, 이로 인해 개발 시간은 연장된다. 매우 정확한 DNA의 비용은 높고, 현행 합성 기술의 처리량은 한정되어 있기 때문에, 서열 공간을 충분하게 분석하지 못한다는 점이 속도 제한 단계로 여전히 남아있다.

[0007] 구축 단계로 시작되는 두 프로세스: 올리고핵산 합성 및 유전자 합성은 주목할 만하다. 과거에 상이한 유전자 변이체의 합성은 분자 클로닝을 통해 달성되었다. 이 접근법은 강건하기는 하지만, 확장이 불가능하다. 초기의 화학적 유전자 합성 노력은 중복 서열 상동성을 갖는 다수의 올리고핵산을 제조하는 것에 주력하였다. 이어서, 풀링하고, 다회에 걸쳐 폴리머라제 연쇄 반응(PCR: polymerase chain reaction)을 수행하여 중복 올리고핵산을 전장의 이중 가닥 유전자로 연쇄시킬 수 있다. 장시간 소요, 노동 집약적인 구성, 다량의 포스포라미다이트 요구, 고가인 원료, 및 하류 단계에 필요한 양보다 유의적으로 더 적은, 나노몰 정도인 양의 최종 생성물의 제조를 비롯한 다수의 인자가 상기 방법을 방해하고, 다수의 별개의 올리고핵산은 유전자 1개의 합성을 준비하는 데 한 96 웰 플레이트를 필요로 하였다.

[0008] 마이크로어레이 상에서의 올리고핵산 합성이 유전자 합성 처리량을 유의적으로 증가시켰다. 다수의 올리고핵산은 마이크로어레이 표면 상에서 합성된 후, 절단되고, 함께 풀링될 수 있다. 특정 유전자가 되는 각 올리고핵산은, 올리고뉴클레오티드의 특정 하위집단이 테풀링되도록 하고, 관심의 대상이 되는 유전자로 조립될 수 있도록 하는 독특한 바코드 서열을 함유한다. 본 프로세스 중 상기 단계에서, 각각의 하위풀은 96 웰 플레이트 중의 한 웰로 전달되고, 이를 통해 96개의 유전자로의 처리량은 증가된다. 이러한 방법의 처리량은 고전적 방법보다 100 배 정도 더 높지만, 비용 효율이 부족하고, 왕복 소요 시간은 느리기 때문에, 한 번에 수천 개의 서열을 필요로 하는 디자인, 구축, 시험 사이클을 적절히 지원하지 못한다.

### 발명의 내용

[0009] 본원에서는 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 변이체 코돈 서열을 코딩하는 것인 단계; 표면을 갖는 구조물을 제공하는 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계; 변이체 핵산의 라이브러리를 세포로 전달하고, 복수의 변이체 단백질을 발현시키는 단계; 및 복수의 변이체 단백질들 중 한 변이체 단백질과 연관된 활성을 확인하는 단계로서, 상기 활성을 단일 기준 서열에 의해 코딩된 단백질과 비교하여 조절되는 것인 단계를 포함하는, 단백질 활성을 조절하는 방법을 제공한다. 본원에서는 활성이 세포의 생식, 성장, 부착, 사멸, 이동, 에너지 생산, 산소 이용률, 대사 활성, 세포 신호전달, 노화, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합을 포함하는 것인 방법을 추가로 제공한다. 본원에서는 세포가 전핵 세포 또는 원핵 세포인 방법을 추가로 제공한다. 본원에서는 세포가 박테리아, 식물, 마우스 또는 영장류의 세포인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 변이체 유전자 또는 그의 단편에 대한 서열을 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 변이체 단백질이 항체, 효소 또는 웹티드인 방법을 추가로 제공한다. 본원에서는 변이체 단백질이 또 다른 분자에 대하여 증진되거나 감소된 결합 친화도를 갖는 것인 방법을 추가로 제공한다. 본원에서는 변이체 단백질이 증진되거나 감소된 효소 활성을 갖는 것인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리에 의해 코딩된 단일 변이체를 그를 필요로 하는 괴형체에게 투여하는 것인 방법을 추가로 제공한다. 본원에서는 약 30,000 이상의 동일하지 않은 올리고핵산이 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 방법을 추가로 제공한다.

[0010] 본원에서는 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 소정의 변이체 서열을 코딩하는 것인 단계; 표면을 갖는 구조물을 제공하는 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되고, 약 30,000 이상의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계; 변이체 핵산의 라이브러리를 제1 세포 세트로 전달하는 단계; 및 세포 활성의 변화를 측정하는 단계로서, 여기서, 세포 활성은 제1 세포 세트 또는 제2 세포 세트에 대하여 측정되고, 여기서, 제2 세포 세트는 제1 세포 세트로부터 단리된 적어도 하나의 발현 생성물로 처리된 것인 단계를 포함하는, 세포 활성을 조절하는 방법을 제공한다. 본원에서는 세포 활성이 생식, 성장, 부

착, 사멸, 이동, 에너지 생산, 산소 이용률, 대사 활성, 세포 신호전달, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합을 포함하는 것인 방법을 추가로 제공한다. 본원에서는 제1 세포 세트 또는 제2 세포 세트가 진핵 세포 또는 원핵 세포인 방법을 추가로 제공한다. 본원에서는 제1 세포 세트 또는 제2 세포 세트가 박테리아, 식물, 마우스 또는 영장류의 세포인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 변이체 유전자 또는 그의 단편에 대한 서열을 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 각각의 변이체 유전자를 번역하여 단백질을 형성하는 단계를 포함하는 방법을 추가로 제공한다. 본원에서는 단백질이 항체, 효소 또는 웹티드인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 항체의 가변 영역 또는 불변 영역의 적어도 일부를 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 항체의 적어도 하나의 CDR 영역을 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 단백질이 또 다른 분자에 대하여 증진되거나 감소된 결합 친화도를 갖거나, 또는 단백질이 증진되거나 감소된 효소 활성을 갖는 것인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 전사 조절 서열을 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 전사 조절 서열이 프로모터, UTR 또는 터미네이터 서열인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 전사될 때, mRNA, miRNA, 또는 shRNA를 코딩하는 서열을 코딩하는 것인 방법을 추가로 제공한다.

[0011] 본원에서는 약 30,000 이상의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각은 단일 기준 서열과 비교하여 변이체 코돈 서열을 코딩하는 것인 단계; 표면을 갖는 구조물을 제공하는 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 합성하는 단계로서, 약 30,000 이상의 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계; 약 30,000 이상의 동일하지 않은 올리고핵산을 DNA 폴리머라제 및 단일 기준 서열과 혼합하여 변이체 핵산의 라이브러리를 형성하는 단계; 변이체 핵산의 라이브러리를 피험체로부터 수득된 세포로 전달하는 단계; 변이체 핵산의 라이브러리에 의해 코딩된 변이체 핵산과 연관된 유해 활성을 감소를 확인하는 단계; 및 변이체 핵산의 라이브러리에 의해 코딩된 변이체 핵산을, 그를 필요로 하는 피험체에게 투여하여 질환 상태를 치료하거나 경감하는 단계를 포함하는, 질환 상태를 치료하거나 경감하는 방법을 제공한다. 본원에서는 질환 상태가 세포 증식성, 자가면역, 바이러스성, 또는 박테리아성 장애와 연관된 것인 방법을 추가로 제공한다. 본원에서는 약 30,000 이상의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 방법을 추가로 제공한다.

[0012] 본원에서는 복수의 동일하지 않은 올리고핵산을 코딩하는 소정의 서열을 제공하는 단계로서, 동일하지 않은 올리고핵산 각각의 길이는 20개 이상의 염기이고, 여기서, 복수의 동일하지 않은 올리고핵산은, 적어도 하나의 서열에 대한 3개 이상의 코돈 각각에 대해 약 19종의 변이체를 코딩하고, 여기서, 복수의 동일하지 않은 올리고핵산은 1종 이상의 유전자 및 그의 변이체를 집합적으로 코딩하는 것인 단계; 표면을 갖는 구조물을 제공하는 단계; 복수의 동일하지 않은 올리고핵산을 합성하는 단계로서, 상기 동일하지 않은 올리고핵산 각각이 상기 표면으로부터 신장되는 것인 단계; 및 복수의 동일하지 않은 올리고핵산으로부터 변이체 핵산의 라이브러리를 조립하는 단계를 포함하는, 핵산 합성 방법을 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 75,000 이상의 동일하지 않은 올리고핵산을 포함하는 것인 방법을 추가로 제공한다. 본원에서는 단일 유전자 및 그의 변이체를 집합적으로 코딩하는 복수의 동일하지 않은 올리고핵산의 서브세트가 상기 구조물 표면 상의 단일 클러스터 내에 위치하는 것인 방법을 추가로 제공한다. 본원에서는 구조물 표면이 6,000개 이상의 단일 클러스터를 포함하는 것인 방법을 추가로 제공한다. 본원에서는 각각의 클러스터가 직경이 약 0.5 내지 2 mm인 채널 내에 위치하는 것인 방법을 추가로 제공한다. 본원에서는 단일 클러스터가 핵산 신장을 위한 50 내지 500개의 좌위를 포함하는 것인 방법을 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 1종 초과의 유전자의 변이체를 집합적으로 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 5,000종 이상의 유전자의 변이체를 집합적으로 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 변이체 핵산의 라이브러리가 효소, 웹티드 또는 항체의 적어도 일부를 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 각각의 변이체 핵산이 유전자의 적어도 85%에 대한 서열을 코딩하는 것인 방법을 추가로 제공한다. 본원에서는 단백질 라이브러리가, 집합적으로 본원에 기술된 동일하지 않은 올리고핵산을 포함하는 복수의 발현 구축물의 발현에 의해 생성된 것인 방법을 추가로 제공한다.

[0013] 본원에서는 복수의 동일하지 않은 올리고핵산을 포함하는 올리고핵산 라이브러리로서, 여기서, 동일하지 않은 올리고핵산 각각의 길이는 적어도 12개의 염기이고, 여기서, 복수의 동일하지 않은 올리고핵산은 3개 이상의 코돈 각각에 대해 약 19종의 변이체를 코딩하고, 여기서, 복수의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 올리고핵산 라이브러리를 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 복수의 동일하지 않은 올리고핵산에 대

한 소정의 서열과 비교하여 총 오류율이 1,500개의 염기 중 1개 미만인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 2,000개의 염기 중 1개 미만인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 각각의 복수의 동일하지 않은 올리고핵산의 길이가 적어도 30개의 염기인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 각각의 복수의 동일하지 않은 올리고핵산의 길이가 적어도 50개의 염기인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 각각의 복수의 동일하지 않은 올리고핵산의 길이가 12 내지 100개의 염기인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 집합적으로 유전자의 적어도 85%에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 집합적으로 유전자 중 복수의 엑손에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 유전자가 항체, 효소 또는 펩티드의 적어도 일부를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 항체의 가변 영역 또는 불변 영역을 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 항체의 적어도 하나의 CDR 영역을 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 집합적으로 발현 카세트의 하나 이상의 세그먼트에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 발현 카세트가 적어도 하나의 프로모터 영역을 포함하고, 복수의 동일하지 않은 올리고핵산이 프로모터 영역의 적어도 일부를 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 발현 카세트가 2개의 프로모터 영역을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈이 연속적으로 존재하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈이 연속적으로 존재하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈 중 적어도 2개가 서로 적어도 하나의 코돈 위치에 의해 이격되어 있는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 라이브러리가 4, 5, 6, 7, 8, 9 또는 10개의 코돈 중 코돈 변이체를 코딩하는 동일하지 않은 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 적어도 4개의 코돈 중 모든 가능한 코돈 변이체를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 동일하지 않은 올리고핵산 중 어느 것도 3개 초과의 히스티딘 잔기에 대한 코돈을 코딩하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 동일하지 않은 올리고핵산 중 어느 것도 4개 초과의 히스티딘 잔기에 대한 코돈을 코딩하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다.

[0014] 본원에서는 75,000 이상의 동일하지 않은 올리고핵산을 포함하는 올리고핵산 라이브러리로서, 여기서, 각각의 75,000 이상의 동일하지 않은 올리고핵산의 길이가 적어도 30개의 염기이고, 여기서, 75,000 이상의 동일하지 않은 올리고핵산이 적어도 하나의 서열에 대한 3개 이상의 코돈 각각에 대해 적어도 3종의 변이체를 코딩하고, 여기서, 75,000 이상의 동일하지 않은 올리고핵산이 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 올리고핵산 라이브러리를 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 75,000 이상의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,500개의 염기 중 1개 미만인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 각각의 75,000 이상의 동일하지 않은 올리고핵산이 적어도 50개의 염기를 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 집합적으로 유전자의 적어도 85%에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 집합적으로 동일 유전자 중 복수의 엑손에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 유전자가 항체, 효소, 또는 어댑터 단백질의 적어도 일부를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 항체의 가변 영역 또는 불변 영역을 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 항체의 적어도 하나의 상보성 결정 영역(CDR: complementarity-determining region)을 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 집합적으로 발현 카세트의 하나 이상의 세그먼트에 대한 서열을 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 발현 카세트가 적어도 하나의 프로모터 영역을 포함하고, 75,000 이상의 동일하지 않은 올리고핵산이 프로모터 영역의 적어도 일부를 코딩하는 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 발현 카세트가 2개의 프로모터 영역을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈이 연속적으로 존재하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈이 연속적으로 존재하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 3개 이상의 코돈 중 적어도 2개가 서로 적어도 하나의 코돈

위치에 의해 이격되어 있는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 75,000 이상의 동일하지 않은 올리고핵산이 3개 이상의 코돈 중 모든 가능한 코돈 변이체를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 동일하지 않은 올리고핵산 중 어느 것도 3개 초과의 히스티딘 잔기에 대한 코돈을 코딩하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 동일하지 않은 올리고핵산 중 어느 것도 4개 초과의 히스티딘 잔기에 대한 코돈을 코딩하지 않는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 라이브러리가 적어도 100,000개의 동일하지 않은 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 라이브러리가 적어도 700,000개의 동일하지 않은 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 라이브러리가 적어도 1,000,000개의 동일하지 않은 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다.

[0015] 본원에서는 복수의 동일하지 않은 올리고핵산을 포함하는 올리고핵산 라이브러리로서, 여기서, 각 동일하지 않은 올리고핵산의 길이는 약 20 내지 130개의 염기이고, 여기서, 복수의 동일하지 않은 올리고핵산이 집합적으로 적어도 하나의 서열에 대한 3개 이상의 코돈 각각에 대해 약 19종의 변이체를 코딩하고, 여기서, 복수의 동일하지 않은 올리고핵산은, 상기 복수의 동일하지 않은 올리고핵산에 대한 소정의 서열과 비교하여 총 오류율이 1,000개의 염기 중 1개 미만인 올리고핵산 라이브러리를 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 50 내지 500개의 동일하지 않은 올리고핵산을 포함하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 복수의 동일하지 않은 올리고핵산이 집합적으로 적어도 50종의 변이체 유전자를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다. 본원에서는 50 내지 500개의 동일하지 않은 올리고핵산이 집합적으로 적어도 50종의 변이체 유전자를 코딩하는 것인 올리고핵산 라이브러리를 추가로 제공한다.

#### 참고 문헌 인용

[0017] 본 명세서에서 언급된 모든 공개문헌, 특히, 및 특히 출원은, 마치 각각의 개별 공개문헌, 특히, 또는 특히 출원이 참조로 인용되는 것으로 구체적으로 및 개별적으로 명시된 것과 같은 정도로 본원에서 참조로 인용된다.

#### 도면의 간단한 설명

[0018] 도 1a-1d는 PCR 돌연변이유발 단계를 도입한 변이체 생물학적 문자 합성을 위한 프로세스 작업 흐름을 도시한 것이다.

도 2a-2d는 단일의 소정의 코돈 부위에서 참조 올리고핵산 서열과 상이한 핵산 서열을 포함하는 올리고핵산 생성을 위한 프로세스 작업 흐름을 도시한 것이다.

도 3a-3f는 각각의 변이체가 단일 코돈 위치에 상이한 핵산 서열을 포함하는 것인 올리고핵산 변이체 세트를 주형 올리고핵산으로 생성하기 위한 대안적 작업 흐름을 도시한 것이다. 각각의 변이체 올리고핵산은 그의 단일 코돈 위치에서 상이한 아미노산을 코딩하고, 상이한 코돈은 X, Y, 및 Z로 표시된다.

도 4a-4e는 각각의 잔기가 단일 동그라미로 표시된, 다수의 아미노산을 갖는 참조 아미노산 서열(도 4a), 및 본원에 기술된 방법을 사용하여 생성된 변이체 아미노산 서열(도 4b, 4c, 4d, & 4e)을 도시한 것이다. 참조 아미노산 서열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 5a-5b는 참조 아미노산 서열(도 5a, 서열 번호: 24), 및 각각의 변이체가 ("X"로 표시되는) 단일의 잔기 변이체를 포함하는 것인 변이체 아미노산 서열의 라이브러리(도 5b, 출현 순서대로 각각 서열 번호: 25-31)를 도시한 것이다. 참조 아미노산 서열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 6a-6b는 참조 아미노산 서열(도 6a), 및 각각의 변이체가 두 부위에 단일 위치 변이체를 포함하는 것인 변이체 아미노산 서열의 라이브러리(도 6b)를 도시한 것이다. 각각의 변이체는 다르게 패턴화된 동그라미로 표시되어 있다. 참조 아미노산 서열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 7a-7b는 참조 아미노산 서열(도 7a), 및 각 스트레치가, 서열이 참조 아미노산 서열과 상이한 (히스티딘을 코딩하는) 3개 부위의 위치 변이체를 갖는 것인 (동그라미 주위에 테두리 상자로 표시된) 아미노산 스트레치를 각각의 변이체가 포함하는 것인 변이체 아미노산 서열의 라이브러리(도 7b)를 도시한 것이다. 참조 아미노산 서

열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 8a-8b는 참조 아미노산 서열(도 8a), 및 각 스트레치가, 서열이 참조 아미노산 서열과 상이한 (패턴화된 동그라미로 표시된) 1개 부위의 단일 위치 변이체를 갖는 것인 (동그라미 주위에 테두리 상자로 표시된) 2개의 아미노산 스트레치를 각각의 변이체가 포함하는 것인 변이체 아미노산 서열의 라이브러리(도 8b)를 도시한 것이다. 참조 아미노산 서열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 9a-9b는 참조 아미노산 서열(도 9a), 및 각 스트레치가, 서열이 참조 아미노산 서열과 상이한 단일 부위의 다중 위치 변이체를 갖는 것인 (패턴화된 동그라미로 표시된) 아미노산 스트레치를 각각의 변이체가 포함하는 것인 아미노산 서열 변이체의 라이브러리(도 9b)를 도시한 것이다. 본 도면에서, 5개 위치는 상이하며, 여기서, 1번 위치는 50/50 K/R 비를 가지고; 2번 위치는 50/25/25 V/L/S 비를 가지고, 3번 위치는 50/25/25 Y/R/D 비를 가지고, 4번 위치는 모든 아미노산에 대하여 동일한 비를 가지고, 5번 위치는 G/P에 대해 75/25 비를 가진다. 참조 아미노산 서열 및 변이체 서열은 본원에 기술된 프로세스에 의해 생성된 핵산 및 그의 변이체에 의해 코딩된 것이다.

도 10은 CDR1, CDR2, 및 CDR3 영역을 갖는 항체를 코딩하는 주형 올리고핵산을 도시하는 것으로서, 여기서, 각 CDR 영역은 변이를 위한 다중 부위를 포함하고, (별표로 표시된) 각 단일 부위는, 주형 올리고핵산 서열과 상이한 임의의 코돈 서열과 상호교환이 가능한 단일 위치 및/또는 다중, 연속 위치의 스트레치를 포함하는 것이다.

도 11은 발현 카세트의 변이체 라이브러리를 생성하기 위해 2개의 발현 카세트 섹션(예컨대, 프로모터, 오픈 리딩 프레임, 및 터미네이터)을 교환함으로써 제조된 예시적인 개수의 변이체를 도시한 것이다.

도 12는 본원에 개시된 바와 같은 유전자 합성을 위한 예시적인 프로세스 작업 흐름을 나타내는 단계 다이어그램을 보여주는 것이다.

도 13은 컴퓨터 시스템의 예를 도시한 것이다.

도 14는 컴퓨터 시스템의 컴퓨터 아키텍처를 도시한 블록 다이어그램이다.

도 15는 복수의 컴퓨터 시스템, 복수의 휴대 전화 및 개인 정보 단말기, 및 네트워크 부착 저장 장치(NAS: Network Attached Storage)를 도입하도록 구성된 네트워크를 나타내는 다이어그램이다.

도 16은 공유 가상 주소 메모리 공간을 사용하는 멀티프로세서 컴퓨터 시스템의 블록 다이어그램이다.

도 17은 젤 전기영동에 의해 분리된 PCR 반응 생성물의 바이오애널라이저(BioAnalyzer) 플롯을 도시한 것이다.

도 18은 각 PCR 생성물 세트가 단일 코돈 위치에서 야생형 주형 핵산과 서열이 상이하고, 여기서, 각 세트의 단일 코돈 위치는 야생형 주형 핵산 서열 중 상이한 부위에 위치하는 것인, 96개의 PCR 생성물 세트를 보여주는 전기영동도를 도시한 것이다. 각 PCR 생성물 세트는, 각각의 변이체가 그의 단일 코돈 위치에서 상이한 아미노산을 코딩하는 것인 19종의 변이체 올리고핵산을 포함한다.

### 발명을 실시하기 위한 구체적인 내용

[0019]

달리 명시되지 않는 한, 본 개시내용은 당업계의 기술 범위 내에 포함되어 있는 종래 문자 생물학 기술을 이용한다. 달리 정의되지 않는 한, 본원에서 사용된 모든 기술 용어 및 과학 용어는 당업계의 숙련가가 보편적으로 이해하는 것과 동일한 의미를 가진다.

[0020]

### 정의

[0021]

본 개시내용 전역에 걸쳐, 다양한 실시양태가 범위 형식으로 제시된다. 범위 형식의 설명은 단지 편의 및 간결함을 위한 것이며, 임의의 실시양태의 범주에 대한 변경불가능한 제한으로서 해석되어서는 안 된다는 것을 이해하여야 한다. 따라서, 문맥상 달리 분명하게 명시되지 않은 한, 범위에 관한 설명은 모든 가능한 하위범위 뿐만 아니라, 그 범위 내의 개별 수치 값도 하한 단위의 10분의 1까지 구체적으로 개시한 것으로서 간주되어야 한다. 예를 들어, 예컨대, 1 내지 6이라는 범위에 관한 설명은 하위범위, 예컨대, 1 내지 3, 1 내지 4, 1 내지 5, 2 내지 4, 2 내지 6, 3 내지 6 등 뿐만 아니라, 그 범위 내의 개별 값, 예를 들어, 1.1, 2, 2.3, 5, 5.9도 구체적으로 개시한 것으로서 간주되어야 한다. 이는 범위 폭과 상관없이 적용된다. 이러한 개체 범위의 상한 및 하한은 독립적으로 더 작은 범위 내에 포함될 수 있고, 이 또한 언급된 범위에서 임의의 구체적으로 배제된 한계를 제외하고 본 개시내용 내에 포함된다. 문맥상 달리 분명하게 명시되지 않은 한, 언급된 범위가 한계 중 하나

또는 둘 모두를 포함하는 경우, 포함된 한계 중 하나 또는 둘 모두를 배제한 범위 또한 본 개시내용 내에 포함된다.

[0022] 본원에서 사용된 용어는 단지 특정 실시양태를 기술하기 위한 것이며, 어느 실시양태도 제한하는 것으로 의도되지 않는다. 본원에서 사용되는 바, 문맥상 달리 명시하지 않은 한, "하나"("a," "an") 및 "그"라는 단수 형태는 복수 형태도 포함하는 것으로 의도된다. 본 명세서에서 사용될 때, "포함하다" 및 "~을 포함하는"이라는 용어는 언급된 특징, 정수, 단계, 작업, 요소 및/또는 성분이 존재함을 명시하지만, 하나 이상의 다른 특징, 정수, 단계, 작업, 요소, 성분, 및/또는 그의 군의 존재 또는 부가를 배제시키는 것은 아님을 추가로 이해할 것이다. 본원에서 사용되는 바, "및/또는"이라는 용어는 연관된 열거 목록 중 하나 이상의 것의 임의의 및 모든 조합을 포함한다.

[0023] 문맥상 구체적으로 언급되지 않거나, 또는 문맥으로부터 명백해지지 않는 한, 수치 또는 수치 범위와 관련하여 본원에서 사용되는 "약"이라는 용어는 언급된 수치 및 수치 +/- 그의 10%, 또는 범위로 열거된 값인 경우, 열거된 하한 아래로 10% 내지 열거된 상한 위로 10%를 의미하는 것으로 이해된다.

#### [0024] 변이체 라이브러리 합성

[0025] 본원에 기술된 방법은 각각의 것이 적어도 하나의 소정의 참조 핵산 서열의 소정의 변이체를 코딩하는 것인 올리고핵산의 라이브러리의 합성을 제공한다. 일부 경우에서, 소정의 기준 서열은 단백질을 코딩하는 핵산 서열이고, 변이체 라이브러리는, 합성된 핵산에 의해 코딩된 후속 단백질 중의 단일의 잔기의 복수의 상이한 변이체가 표준 번역 프로세스에 의해 생성되도록 적어도 한 단일 코돈의 변이를 코딩하는 서열을 포함한다. 뉴클레오티드 변이를 중복 또는 둔단 올리고핵산 프라이머 내로 도입함으로써 핵산 서열 중 합성된 특정 변경을 도입할 수 있다. 대안적으로, 올리고핵산 집단은 집합적으로 긴 핵산(예컨대, 유전자) 및 그의 변이체를 코딩할 수 있다. 이러한 배열로 올리고핵산 핵산은 하이브리드화될 수 있고, 표준 문자 생물학 기술에 의해 긴 핵산(예컨대, 유전자) 및 그의 변이체를 형성할 수 있다. 긴 핵산(예컨대, 유전자) 및 그의 변이체가 세포에서 발현될 때, 변이체 단백질 라이브러리가 생성된다. 유사하게, 본원에서는 RNA 서열(예컨대, miRNA, shRNA, 및 mRNA) 또는 DNA 서열(예컨대, 인핸서, 프로모터, UTR, 및 터미네이터 영역)을 코딩하는 변이체 라이브러리를 합성하는 방법을 제공한다. 또한, 본원에서는 본원에 기술된 방법을 사용하여 합성된 라이브러리 중 선택된 변이체에 대한 하류 적용을 제공한다. 하류 적용은 생물학적으로 관련된 기능, 예컨대, 생화학적 친화도, 효소 활성, 세포 활성 변화가 증진된 변이체 핵산 또는 단백질 서열, 및 질환 상태 치료용 또는 예방용의 변이체 핵산 또는 단백질 서열을 확인하는 것을 포함한다.

#### [0026] PCR 돌연변이유발이 후속되는 합성

[0027] 올리고핵산의 변이체 라이브러리의 합성을 위한 제1 프로세스는 PCR 돌연변이유발 방법을 위한 것이다. 본 작업 흐름에서, 각 올리고핵산이, 참조 핵산 서열의 소정의 변이체인 소정의 서열을 코딩하는 것인, 복수의 올리고핵산이 합성된다. **도 1a-1d**에 도시된 예시적인 작업 흐름에 관한 도면을 참조하면, 여기서, 올리고핵산은 표면 상에서 생성된다. **도 1a**는 121개의 좌위를 포함하는 표면의 단일 클러스터의 확대도를 도시한 것이다. **도 1b**에 도시된 각각의 올리고핵산은 **도 1c**의 변이체 긴 핵산의 라이브러리를 제조하기 위하여 참조 핵산 서열로부터의 증폭을 위해 사용될 수 있는 프라이머이다. 이어서, 변이체 긴 핵산의 라이브러리를 임의적으로 전사 및/또는 번역시켜 **도 1d**의 변이체 RNA 또는 단백질 라이브러리를 생성한다. 본 예시도에서, 표면이 실질적으로 평면인 장치는 **도 1a**에 도시된 올리고핵산의 드 노보 합성을 위해 사용된다. 일부 경우에서, 장치는 좌위의 클러스터를 포함하고, 여기서, 각 좌위는 올리고핵산 신장을 위한 부위이다. 일부 경우에서, 단일 클러스터는 원하는 변이체 서열 라이브러리를 생성하는 데 필요한 모든 올리고핵산 변이체를 포함한다. 대안적 방식에서, 플레이트는 클러스터로 분리되지 않은 좌위의 필드를 포함한다.

[0028] 본원에 기술된 드 노보 합성된 올리고핵산 라이브러리는, 각각이 제1 위치인 위치 "x\*"에 적어도 하나의 변이체 서열을 포함하는 것인 복수의 올리고핵산을 포함할 수 있고, 각각의 변이체 올리고핵산은 제1 신장 생성물을 생성하기 위한 제1 라운드의 PCR에서 프라이머로서 사용된다. 이러한 예에서, 제1 올리고핵산(220) 중 위치 "x"는 기준 서열로부터 변이체 코돈 서열, 즉, 19개의 가능한 변이체 중 하나를 코딩한다. **도 2a**를 참조한다. 제1 올리고핵산의 것과 중복된 서열을 포함하는 제2 올리고핵산(225) 또한 제2 신장 생성물을 생성하기 위한 별개인 다른 라운드의 PCR에서 프라이머로서 사용된다. 추가로, 외부 프라이머((215), (230))는 긴 핵산 서열로부터의 단편의 증폭을 위해 사용될 수 있다. 생성된 증폭 생성물은 긴 핵산 서열((235), (240))의 단편이다. **도 2b**를 참조한다. 이어서, 긴 핵산 서열((235), (240))의 단편을 하이브리드화하고, 신장 반응에 가하여 긴 핵산(245)의 변이체를 형성한다. **도 2c**를 참조한다. 제1 및 제2 신장 생성물의 중복 단부가 제2 라운드의 PCR의 프라이머

로서 작용할 수 있고, 이로써, 변이체를 함유하는 제3 신장 생성물이 생성될 수 있다(도 2d). 수율을 증가시키기 위해, 긴 핵산의 변이체는 DNA 폴리머라제, 중폭 시약, 및 외부 프라이머((215), (230))를 포함하는 반응에서 증폭된다. 일부 경우에서, 제2 올리고핵산은 제한하는 것은 아니지만, 변이체 부위에 인접한 서열을 포함한다. 대안적 방식에서, 제2 올리고핵산과 중복되는 영역을 갖는 제1 올리고핵산이 생성된다. 이러한 시나리오에서, 최대 19종의 변이체를 위하여 단일 코돈에 변이를 갖는 제1 올리고핵산이 합성된다. 제2 올리고핵산은 변이체 서열을 포함하지 않는다. 임의적으로, 제1 집단은 제1 올리고핵산 변이체, 및 상이한 코돈 부위에 변이체를 코딩하는 추가의 올리고핵산을 포함한다. 대안적으로, 제1 올리고핵산 및 제2 올리고핵산은 둘다 라이제이션을 위해 디자인될 수 있다.

[0029] 대안적으로, 돌연변이유발 PCR 방법이 도 3a-3f에 도시되어 있다. 상기 프로세스에서, 제1 및 제2 가닥((305), (310))을 포함하는 주형 핵산 분자(300)를 제1 프라이머(315) 및 제2 프라이머(320)를 함유하는 PCR 반응에서 증폭시킨다(도 3a). 증폭 반응은 뉴클레오티드 시약으로서 우라실을 포함한다. 우라실 표지된 신장 생성물(325)(도 3b)을 생성하고, 임의적으로 정제하고, 제1 올리고핵산(335) 및 복수의 제2 올리고핵산(330)을 사용하여 후속 PCR 반응을 위한 주형으로서의 역할을 하고, 이로써, 제1 신장 생성물((340) 및 (345))이 생성된다(도 3c-3d). 본 프로세스에서, 복수의 제2 올리고핵산(330)은 (도 3c에서 X, Y, 및 Z로 표시된) 변이체 서열을 코딩하는 올리고핵산을 포함한다. 우라실 표지된 주형 핵산을 우라실 특이 절제 시약, 예컨대, 뉴 잉글랜드 바이오랩스(New England Biolabs)로부터 상업적으로 이용가능한 USER 다이제스트(USER digest)에 의해 분해시킨다. 변이체(335) 및 변이체 X, Y, 및 Z를 포함하는 상이한 코돈(330)을 부가하고, 제한된 PCR 단계를 수행하여 도 3D를 생성한다. 우라실 함유 주형 분해 후, 신장 생성물의 중복 단부는 제1 외부 프라이머(350) 및 제2 외부 프라이머(355)와 함께 조합하여 프라이머로서 작용하는 제1 신장 생성물((340) 및 (345))을 이용하여 PCR 반응을 프라이밍하는 역할을 하고, 이로써, 도 3f의 변이체 부위에 복수의 변이체 X, Y, 및 Z를 함유하는 핵산 분자(360)의 라이브러리가 생성된다.

#### 긴 핵산의 변이체 및 비변이체 부분을 포함하는 집단의 드 노보 합성

[0030] [0031] 변이체 라이브러리 합성을 위한 제2 프로세스에서, 표면은 긴 핵산의 다중 단편의 드 노보 합성을 위해 사용되며, 여기서, 단편 중 적어도 하나는, 각 베전이 상이한 변이체 서열의 것인 다중 베전으로 합성된다. 상기 방식에서, 변이체 긴 범위의 핵산의 라이브러리를 조립하는 데 필요한 단편 모두 드 노보 합성된다. 합성된 단편은 합성 후, 단편 라이브러리가 하이브리드화될 수 있도록 중복 서열을 가질 수 있다. 하이브리드화 후, 임의의 상보적인 캡을 충전시키기 위해 신장 반응을 수행할 수 있다.

[0032] 대안적으로, 합성된 단편을 프라이머를 이용하여 증폭시킨 후, 그에 대해 둔단 라이제이션 또는 중복 하이브리드화를 수행할 수 있다. 일부 경우에서, 장치는 좌위의 클러스터를 포함하고, 여기서, 각 좌위는 올리고핵산 신장을 위한 부위이다. 일부 경우에서, 단일 클러스터는 원하는 변이체 서열 라이브러리를 생성하기 위해 모든 올리고핵산 변이체 및 소정의 긴 핵산의 다른 단편 서열을 포함한다. 클러스터는 약 50 내지 500개의 좌위를 포함할 수 있다. 일부 방식에서, 클러스터는 500개 초과의 좌위를 포함할 수 있다.

[0033] 제1 올리고핵산 집단 중의 각각의 개별 올리고핵산은 클러스터의 별개의, 개별적으로 주소지정가능한 좌위 상에서 생성될 수 있다. 한 올리고핵산 변이체는 복수의 개별적으로 주소지정가능한 좌위에 의해 제시될 수 있다. 제1 올리고핵산 집단 중 각각의 변이체는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10회 이상 제시될 수 있다. 일부 경우에서, 제1 올리고핵산 집단 중 각각의 변이체는 3개 이하의 좌위에서 제시된다. 일부 경우에서, 제1 올리고핵산 집단 중 각각의 변이체는 2개의 좌위에서 제시된다. 일부 경우에서, 제1 올리고핵산 집단 중 각각의 변이체는 오직 단 하나의 단일 좌위에서 제시된다.

[0034] 본원에서는 중복이 감소된 핵산 라이브러리를 생성하는 방법을 제공한다. 일부 경우에서, 원하는 변이체 올리고뉴클레오티드를 수득하기 위해 1회 초과로 변이체 올리고뉴클레오티드를 합성할 필요없이 변이체 올리고뉴클레오티드를 생성할 수 있다. 일부 경우에서, 본 개시내용은 원하는 변이체 올리고뉴클레오티드를 생성하기 위해 1, 2, 3, 4, 5회, 6, 7, 8, 9, 10회 이상 초과로 변이체 올리고뉴클레오티드를 합성할 필요없이 변이체 올리고뉴클레오티드를 생성하는 방법을 제공한다.

[0035] 원하는 변이체 올리고뉴클레오티드를 수득하기 위해 1개 초과의 별개의 부위에서 변이체 올리고뉴클레오티드를 합성할 필요없이 변이체 올리고뉴클레오티드를 생성할 수 있다. 본 개시내용은 원하는 변이체 올리고뉴클레오티드를 생성하기 위해 1개 초과의 부위, 2개 초과의 부위, 3개 초과의 부위, 4개 초과의 부위, 5개 초과의 부위, 6개 초과의 부위, 7개 초과의 부위, 8개 초과의 부위, 9개 초과의 부위, 또는 10개 초과의 부위에서 변이체 올리고뉴클레오티드를 합성할 필요없이 변이체 올리고뉴클레오티드를 생성하는 방법을 제공한다. 일부 경우에서,

올리고뉴클레오티드는 최대 6, 5, 4, 3, 2, 또는 1개의 별개의 부위에서 합성된다. 동일한 올리고뉴클레오티드는 표면 상의 1, 2, 또는 3개의 별개의 좌위에서 합성될 수 있다.

[0036] 일부 경우에서, 단일 변이체 올리고뉴클레오티드를 나타내는 좌위의 양은 하류 프로세싱, 예컨대, 증폭 반응 또는 세포 검정법에서 필요한 핵산 물질의 양의 함수이다. 일부 경우에서, 단일 변이체 올리고뉴클레오티드를 나타내는 좌위의 양은 단일 클러스터에서 이용가능한 좌위의 함수이다.

[0037] 본원에서는 참조 핵산 중 복수의 부위에서 상이한 변이체 올리고핵산을 포함하는 올리고핵산의 라이브러리를 생성하는 방법을 제공한다. 상기 경우에서, 각각의 변이체 라이브러리는 좌위의 클러스터 내의 개별적으로 주소지정가능한 좌위 상에서 생성된다. 올리고핵산 라이브러리에 의해 제시되는 변이체 부위의 개수는 클러스터 중 개별적으로 주소지정가능한 좌위의 개수 및 각 부위에서의 원하는 변이체의 개수에 의해 결정될 것이라는 것을 이해할 것이다. 일부 경우에서, 각각의 클러스터는 약 50 내지 500개의 좌위를 포함한다. 일부 경우에서, 각각의 클러스터는 100 내지 150개의 좌위를 포함한다.

[0038] 예시적인 방식에서, 19종의 변이체는 19개의 가능한 변이체 아미노산 각각의 것을 코딩하는 코돈에 상응하는 변이체 부위에서 제시된다. 또 다른 예시적인 경우에서, 61종의 변이체는 19개의 가능한 변이체 아미노산 각각의 것을 코딩하는 트리플릿에 상응하는 변이체 부위에서 제시된다. 비제한적인 예에서, 클러스터는 121개의 개별적으로 주소지정가능한 좌위를 포함한다. 상기 예에서, 올리고핵산 집단은 단일 부위 변이체 각각의 것의 6개의 복제물(6개의 복제물 x 1종의 변이체 부위 x 19종의 변이체 = 114개의 좌위), 이중 부위 변이체 각각의 것의 3개의 복제물(3개의 복제물 x 2종의 변이체 부위 x 19종의 변이체 = 114개의 좌위), 또는 삼중 부위 변이체 각각의 것의 2개의 복제물(2개의 복제물 x 3종의 변이체 부위 x 19종의 변이체 = 114개의 좌위)을 포함한다. 일부 경우에서, 올리고핵산 집단은 4, 5, 6개, 또는 6개 초과의 변이체 부위에 변이체를 포함한다.

[0039] 코돈 변이

[0040] 본원에 기술된 변이체 올리고핵산 라이브러리는, 각각의 올리고핵산이 참조 핵산 서열과 비교하여 변이체 코돈 서열을 코딩하는 것인 복수의 올리고핵산을 포함할 수 있다. 일부 경우에서, 제1 올리고핵산 집단의 각각의 올리고핵산은 단일 변이체 부위에 변이체를 함유한다. 일부 경우에서, 제1 올리고핵산 집단은 단일 변이체 부위에 복수의 변이체를 함유하고, 이로써, 제1 올리고핵산 집단은 동일한 변이체 부위에 1개 초과의 변이체를 함유한다. 제1 올리고핵산 집단은 동일한 변이체 부위에 다중 코돈 변이체를 집합적으로 코딩하는 올리고핵산을 포함할 수 있다. 제1 올리고핵산 집단은 동일한 위치에서 최대 19개 이상의 코돈을 집합적으로 코딩하는 올리고핵산을 포함할 수 있다. 제1 올리고핵산 집단은 동일한 위치에서 최대 60종의 변이체 트리플릿을 집합적으로 코딩하는 올리고핵산을 포함할 수 있거나, 또는 제1 올리고핵산 집단은 동일한 위치에서 최대 61개의 상이한 코돈 트리플릿을 집합적으로 코딩하는 올리고핵산을 포함할 수 있다. 각각의 변이체는 번역 동안 상이한 아미노산을 생성하는 코돈을 코딩할 수 있다. 하기 표 1은 변이체 부위에 대하여 가능한 각각의 코돈 (및 대표적인 아미노산)에 관한 목록을 제공한다.

표 1

## 코돈 및 아미노산 목록

아미노산	1문자 코드	3문자 코드	코돈					
알라닌	A	Ala	GCA	GCC	GCG	GCT		
시스테인	C	Cys	TGC	TGT				
아스파르트산	D	Asp	GAC	GAT				
글루탐산	E	Glu	GAA	GAG				
페닐알라닌	F	Phe	TTC	TTT				
글리신	G	Gly	GGA	GGC	GGG	GGT		
히스티딘	H	His	CAC	CAT				
이소류신	I	Iso	ATA	ATC	ATT			
리신	K	Lys	AAA	AAG				
류신	L	Leu	TTA	TTG	CTA	CTC	CTG	CTT
메티오닌	M	Met	ATG					
아스파라긴	N	Asn	AAC	AAT				
프롤린	P	Pro	CCA	CCC	CCG	CCT		
글루타민	Q	Gln	CAA	CAG				
아르기닌	R	Arg	AGA	AGG	CGA	CGC	CGG	CGT
세린	S	Ser	AGC	AGT	TCA	TCC	TCG	TCT
트레오닌	T	Thr	ACA	ACC	ACG	ACT		
발린	V	Val	GTA	GTC	GTG	GTT		
트립토판	W	Trp	TGG					
티로신	Y	Tyr	TAC	TAT				

[0041]

[0042]

올리고핵산 집단은 다중 위치에서 최대 20개의 코돈을 집합적으로 코딩하는 다양한 올리고핵산을 포함할 수 있다. 상기 경우에서, 집단 중 각각의 올리고핵산은 동일한 올리고핵산 중 1개 초과의 위치에서 코돈에 대한 변이를 포함한다. 일부 경우에서, 집단 중 각각의 올리고핵산은 단일 올리고핵산 중 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 이상의 코돈에서 코돈에 대한 변이를 포함한다. 일부 경우에서, 각각의 변이체 긴 핵산은 단일의 긴 핵산 중 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 이상의 코돈에서 코돈에 대한 변이를 포함한다. 일부 경우에서, 변이체 올리고핵산 집단은 단일 올리고핵산 중 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 이상의 코돈에서 코돈에 대한 변이를 포함한다. 일부 경우에서, 변이체 올리고핵산 집단은 단일의 긴 핵산 중 적어도 약 10, 20, 30, 40, 50,

60, 70, 80, 90, 100개 이상의 코돈에서 코돈에 대한 변이를 포함한다.

[0043] 본원에서는 제2 올리고핵산 집단이, 복수의 개별적으로 주소지정가능한 좌위를 함유하는 제2 클러스터 상에서 생성되는 것인 프로세스를 제공한다. 제2 올리고핵산 집단은, 각 코돈 위치에 대하여 불변인 (즉, 각 위치에서 동일한 아미노산을 코딩하는) 복수의 제2 올리고핵산을 포함할 수 있다. 제2 올리고핵산은 제1 올리고핵산의 적어도 일부와 중복될 수 있다. 일부 경우에서, 제2 올리고핵산은 제1 올리고핵산에 있는 변이체 부위를 함유하지 않는다. 대안적으로, 제2 올리고핵산 집단은 하나 이상의 코돈 위치에 대한 변이체를 포함하는 복수의 제2 올리고핵산을 포함할 수 있다.

[0044] 본원에서는 다중 코돈 위치에 변이체를 포함하는 단일의 올리고핵산 집단을 생성하는 것인, 올리고핵산의 라이브러리를 합성하는 방법을 제공한다. 제1 올리고핵산 집단은 복수의 개별적으로 주소지정가능한 좌위를 함유하는 제1 클러스터 상에서 생성될 수 있다. 상기 경우에서, 제1 올리고핵산 집단은 상이한 코돈 위치에 변이체를 포함한다. 일부 경우에서, 상이한 부위는 연속적으로 존재한다(즉, 연속 아미노산을 코딩하는 것이다). 제1 올리고뉴클레오티드 산 집단은 동일한, 또는 추가의 변이체 부위에서 최대 19의 코돈 변이체를 집합적으로 코딩하는 다양한 올리고핵산을 포함할 수 있다. 제1 올리고뉴클레오티드 산 집단은 위치  $x$ 에 최대 19종의 변이체, 위치  $y$ 에 최대 19종의 변이체, 및 위치  $z$ 에 최대 19종의 변이체를 함유하는 복수의 제1 올리고핵산을 포함할 수 있다. 상기 방식에서, 각각의 변이체는 상이한 아미노산을 코딩하고, 이로써, 각각의 상이한 변이체 부위에서 최대 19의 아미노산 변이체가 코딩된다. 추가 경우에서, 제2 올리고핵산 집단은 복수의 개별적으로 주소지정가능한 좌위를 함유하는 제2 클러스터 상에서 생성된다. 제2 올리고핵산 집단은, 각 코돈 위치에 대하여 불변인 (즉, 각 위치에서 동일한 아미노산을 코딩하는) 복수의 제2 올리고핵산을 포함할 수 있다. 제2 올리고핵산은 제1 올리고핵산의 적어도 일부와 중복될 수 있다. 일부 경우에서, 제2 올리고핵산은 제1 올리고핵산에 있는 변이체 부위를 함유하지 않을 수 있다.

[0045] 본원에 기술된 프로세스에 의해 생성된 변이체 핵산 라이브러리는 변이체 단백질 라이브러리를 생성한다. 제1의 예시적인 방식에서, 주형 올리고핵산은, 전사 및 번역되었을 때, 단일 동그라미로 표시된, 다수의 코돈 위치를 갖는 참조 아미노산 서열(도 4a)을 생성하는 서열을 코딩한다. 주형의 올리고핵산 변이체는 본원에 기술된 방법을 사용하여 생성될 수 있다. 일부 경우에서, 단일 변이체가 올리고핵산 중에 존재하고, 이로써, 단일 아미노산 서열이 생성된다(도 4b). 일부 경우에서, 하나 이상의 코돈에 의해 이격되어 있는 것인 1개 초파의 변이체가 올리고핵산 중에 존재하고, 이로써, 변이체 잔기 사이가 이격되어 있는 단백질이 생성된다(도 4c). 일부 경우에서, 변이체가 일련으로 연속되고, 서로 인접하거나, 또는 연속되어 있는 것인 1개 초파의 변이체가 올리고핵산 중에 존재하고, 이로써, 잔기로 이루어진, 이격화된 변이체 스트레치가 생성된다(도 4d). 일부 경우에서, 각각의 변이체 스트레치가 일련으로 연속되고, 인접한 또는 연속되어 있는 변이체를 포함하는 것인 2종의 변이체 스트레치가 올리고핵산 중에 존재한다(도 4e).

[0046] 본원에서는 각각의 변이체가 단일 위치 코돈 변이체를 포함하는 것인, 올리고핵산 변이체의 라이브러리를 생성하는 방법을 제공한다. 한 경우에서, 주형 올리고핵산은 도 5a에서 예시적인 아미노산 잔기가 그의 각 1문자 코드 단백질 코돈이 기재된 동그라미로 표시되어 있는 다수의 코돈 위치를 가진다. 도 5b는 변이체 핵산의 라이브러리에 의해 코딩된 아미노산 변이체의 라이브러리를 도시하는 것이며, 여기서, 각각의 변이체는 상이한 단일 부위에 위치하는, "X"로 표시된 단일 위치 변이체를 포함한다. 제1 위치 변이체는 알라닌을 대체하는 임의의 코돈을 가지고, 제2 변이체는 트립토판을 대체하는 변이체 핵산의 라이브러리에 의해 코딩되는 임의의 코돈을 가지고, 제3 변이체는 이소류신을 대체하는 임의의 코돈을 가지고, 제4 변이체는 리신을 대체하는 임의의 코돈을 가지고, 제5 변이체는 아르기닌을 대체하는 임의의 코돈을 가지고, 제6 변이체는 글루탐산을 대체하는 임의의 코돈을 가지고, 제7 변이체는 글루타민을 대체하는 임의의 코돈을 가진다. 모든 코돈 변이체 또는 모든 코돈 변이체보다 적은 개수의 코돈 변이체가 변이체 핵산 라이브러리에 의해 코딩되었을 때, 생성된, 상응하는 아미노산 서열 변이체의 집단은 단백질 발현(즉, DNA 전사, 이어서, 번역 및 프로세싱 이벤트가 수행되는 표준 세포 이벤트) 후에 생성된다.

[0047] 일부 방식에서, 다중 부위의 단일 위치 변이체를 포함하는 라이브러리가 생성된다. 도 6a에 도시된 바와 같은 야생형 주형을 제공한다. 도 6b는 2개 부위의 단일 위치 코돈 변이체를 포함하는 생성된 아미노산 서열을 도시한 것이며, 여기서, 상이한 아미노산을 코딩하는 각각의 코돈 변이체는 다르게 패턴화된 동그라미로 표시되어 있다.

[0048] 본원에서는 다중 부위, 단일 위치 변이체의 스트레치를 갖는 라이브러리를 생성하는 방법을 제공한다. 각각의 올리고핵산 스트레치는 1, 2, 3, 4, 5개 이상의 변이체를 가질 수 있다. 각각의 올리고핵산 스트레치 적어도 1

종의 변이체를 가질 수 있다. 각각의 올리고핵산 스트레치는 적어도 2종의 변이체를 가질 수 있다. 각각의 올리고핵산 스트레치는 적어도 3종의 변이체를 가질 수 있다. 예를 들어, 5개의 올리고핵산 스트레치는 1종의 변이체를 가질 수 있다. 5개의 올리고핵산 스트레치는 2종의 변이체를 가질 수 있다. 5개의 올리고핵산 스트레치는 3종의 변이체를 가질 수 있다. 5개의 올리고핵산 스트레치는 4종의 변이체를 가질 수 있다. 예를 들어, 4개의 올리고핵산 스트레치는 1종의 변이체를 가질 수 있다. 4개의 올리고핵산 스트레치는 2종의 변이체를 가질 수 있다. 4개의 올리고핵산 스트레치는 3종의 변이체를 가질 수 있다. 4개의 올리고핵산 스트레치는 4종의 변이체를 가질 수 있다.

[0049] 일부 경우에서, 단일 위치 변이체는 모두 동일한 아미노산, 예컨대, 히스티딘을 코딩할 수 있다. 도 7a에 도시된 바와 같은 참조 아미노산 서열을 제공한다. 상기 방식에서, 올리고핵산 스트레치는 다중 부위의 단일 위치 변이체를 코딩하고, 도 7b에서와 같이, 발현되었을 때, 모두가 히스티딘을 코딩하는 것인 단일 위치 변이체를 갖는 아미노산 서열이 생성된다. 일부 실시양태에서, 본원에 기술된 방법에 의해 합성된 변이체 라이브러리는 생성된 아미노산 서열 중 4개 초과의 히스티딘 잔기를 코딩하지 않는다.

[0050] 일부 경우에서, 본원에 기술된 방법에 의해 생성된 핵산의 변이체 라이브러리는 아미노산 서열이 별개의 변이 스트레치를 갖는 것인 아미노산 서열의 발현을 제공한다. 주형 아미노산 서열은 도 8a에 도시되어 있다. 올리고핵산 스트레치는 2개의 스트레치 중에 단 하나의 변이체 코돈을 가질 수 있고, 발현되었을 때, 도 8b에 도시되어 있는 아미노산 서열이 생성될 수 있다. 변이체는 아미노산 변이가 단일 스트레치 중 상이한 위치에 존재한다는 것을 나타내기 위해 도 8b에서 다르게 패턴화된 동그라미로 표시되어 있다.

[0051] 본원에서는 각 부위에 대한 변이체가 선택적으로 조절되는 것인, 1, 2, 3개 이상의 코돈 변이체를 포함하는 올리고핵산 라이브러리를 합성하는 방법 및 장치를 제공한다. 단일 부위 변이체에 대한 두 아미노산의 비는 약 1:100, 1:50, 1:10, 1:5, 1:3, 1:2, 1:1일 수 있다. 단일 부위 변이체에 대한 3개의 아미노산의 비는 약 1:1:100, 1:1:50, 1:1:20, 1:1:10, 1:1:5, 1:1:3, 1:1:2, 1:1:1, 1:10:10, 1:5:5, 1:3:3, 또는 1:2:2일 수 있다. 도 9a는 야생형 핵산 서열에 의해 코딩된 야생형 참조 아미노산 서열을 도시한 것이다. 도 9b는 각 위치가 생성된 변이체 단백질 라이브러리 중 특정 비로 아미노산을 가질 수 있는 것인 (패턴화된 동그라미로 표시된) 서열 스트레치를 각각의 변이체가 포함하는 것인 아미노산 변이체의 라이브러리를 도시한 것이다. 생성된 변이체 단백질 라이브러리는 본원에 기술된 방법에 의해 생성된 변이체 핵산 라이브러리에 의해 코딩된다. 이러한 예시에서, 5개의 위치는 달라진다: 제1 위치(900)는 50/50 K/R 비를 가지고; 제2 위치(910)는 50/25/25 V/L/S 비를 가지고, 제3 위치(920)는 50/25/25 Y/R/D 비를 가지고, 제4 위치(930)는 20개의 아미노산 모두에 대하여 동일한 비를 가지고, 제5 위치(940)는 G/P에 대하여 75/25 비를 가진다. 본원에 기술된 비는 단지 예시일 뿐이다.

[0052] 일부 경우에서, 궁극적으로는 단백질의 아미노산 서열로 번역되는 핵산 서열을 코딩하는 합성된 변이체 라이브러리가 생성된다. 예시적인 아미노산 서열은 작은 웹티드 뿐만 아니라, 큰 웹티드, 예컨대, 항체 서열의 적어도 일부를 코딩하는 것을 포함한다. 일부 경우에서, 합성된 각각의 올리고핵산은 항체 서열의 일부 중의 변이체 코돈을 코딩한다. 합성된 변이체 올리고핵산의 일부가 코딩하는 예시적인 항체 서열은 그의 항원 결합 또는 가변 영역, 또는 그의 단편을 포함한다. 본원에 기술된 올리고핵산이 그의 일부를 코딩하는 것인 항체 단편의 예로는 제한 없이, Fab, Fab', F(ab')2 및 Fv 단편, 디아바디, 선형 항체, 단일 쇄 항체 분자, 및 항체 단편으로부터 형성된 다중특이적 항체를 포함한다. 본원에 기술된 올리고핵산이 그의 일부를 코딩하는 것인 항체 영역의 예로는 제한 없이, Fc 영역, Fab 영역, Fab 영역의 가변 영역, Fab 영역의 불변 영역, 중쇄 또는 경쇄( $V_H$  또는  $V_L$ )의 가변 도메인, 또는  $V_H$  또는  $V_L$ 의 특이적인 상보성 결정 영역(CDR)을 포함한다. 본원에 개시된 방법에 의해 생성된 변이체 라이브러리를 통해 본원에 기술된 항체 영역 중 하나 이상의 것을 변이시킬 수 있다. 한 예시적인 프로세스에서, 수개의 CDR을 코딩하는 핵산에 대한 변이체 라이브러리가 생성된다. 도 10을 참조한다. 각 CDR 영역이 변이를 위한 다중 부위를 포함하는 것인 CDR1(1010), CDR2(1020), 및 CDR3(1030) 영역을 갖는 항체를 코딩하는 주형 핵산은 본원에 기술된 방법에 의해 변형된다. 중쇄 또는 경쇄의 단일 가변 도메인 중 3개의 CDR 각각에 대한 변이 (1015), (1025), 및 (1035)가 생성된다. 별표로 표시된 각 부위는, 주형 올리고핵산 서열과 상이한 임의의 코돈 서열과 상호교환이 가능한 단일 위치, 다중, 연속 위치의 스트레치, 또는 그 둘 모두를 포함할 수 있다. 본원에 사용되는 방법을 사용함으로써 변이체 라이브러리의 다양성은 최대  $\sim 10^{10}$  이상의 다양도로 까지 크게 증가될 수 있다.

[0053] 발현 카세트의 변이

[0054]

일부 경우에서, 발현 구축물의 일부를 코딩하는 합성된 변이체 라이브러리가 생성된다. 발현 구축물의 일부의 예로는 프로모터, 오픈 리딩 프레임, 및 종결 종결 영역을 포함한다. 일부 경우에서, 발현 구축물은 1, 2, 3개 이상의 발현 카세트를 코딩한다. 도 11에 도시된 바와 같이, 발현 구축물 카세트의 일부를 구성하는 단일 부위 또는 다중 부위의 별개 영역에서의 코돈 변이를 코딩하는 올리고뉴클레오티드 라이브러리가 생성될 수 있다. 2 개의 구축물을 발현하는 한 카세트를 생성하기 위해 제1 프로모터(1110), 제1 오픈 리딩 프레임(1120), 제1 터미네이터(1130), 제2 프로모터(1140), 제2 오픈 리딩 프레임(1150), 또는 제2 터미네이터 서열(1160)로 이루어진 변이체 서열의 적어도 일부를 코딩하는 변이체 올리고핵산을 합성하였다. 이전 실시예에서 기술된 바와 같이, 증폭 라운드 후, 1,024개의 발현 구축물로 이루어진 라이브러리가 생성되었다. 도 11은 단지 일례의 방식을 제공하는 것이다. 일부 경우에서, 추가의 조절 서열, 예컨대, 비번역 조절 영역(UTR: untranslated regulatory region) 또는 인핸서 영역 또한 본원에서 언급된 발현 카세트에 포함된다. 발현 카세트는, 그에 대한 변이체 서열이 본원에 기술된 방법에 의해 생성되는 것인 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 구성 요소를 포함할 수 있다. 일부 경우에서, 발현 구축물은 다중시스트론성 벡터 중 1종 초과의 유전자를 포함한다. 한 예에서, 합성된 DNA 올리고핵산을 바이러스 벡터(예컨대, 렌티바이러스) 내로 삽입한 후, 세포로의 형질도입을 위해 패키징하거나, 또는 세포 내로의 전달을 위해 비바이러스 벡터로 삽입한 후, 스크리닝 및 분석을 수행한다.

[0055]

본원에 개시된 핵산을 삽입하기 위한 발현 벡터는 진핵성 벡터(예컨대, 박테리아 및 진균) 및 원핵성 벡터(예컨대, 포유동물, 식물 및 곤충 발현 벡터)를 포함한다. 예시적인 발현 벡터로는 제한 없이, 포유동물 발현 벡터: pSF-CMV-NEO-NH2-PPT-3XFLAG, pSF-CMV-NEO-COOH-3XFLAG, pSF-CMV-PURO-NH2-GST-TEV, pSF-OXB20-COOH-TEV-FLAG(R)-6His, pCEP4 pDEST27, pSF-CMV-Ub-KrYFP, pSF-CMV-FMDV-daGFP, pEF1a-m체리-N1 벡터(pEF1a-mCherry-N1 Vector), pEF1a-td토마토 벡터(pEF1a-tdTomato Vector), pSF-CMV-FMDV-하이그로(pSF-CMV-FMDV-Hygro), pSF-CMV-PGK-푸로(pSF-CMV-PGK-Puro), pMCP-tag(m), 및 pSF-CMV-PURO-NH2-CMYC; 박테리아 발현 벡터: pSF-OXB20-베타Gal(pSF-OXB20-BetaGal), pSF-OXB20-Fluc, pSF-OXB20, 및 pSF-Tac; 식물 발현 벡터: pRI 101-AN DNA 및 pCambia2301(pCambia2301); 및 효모 발현 벡터: pTYB21 및 pKLAC2, 및 곤충 벡터: pAc5.1/V5-His A 및 pDEST8을 포함한다. 예시적인 세포로는 제한 없이, 원핵 세포 및 진핵 세포를 포함한다. 예시적인 진핵 세포로는 제한 없이, 동물 세포, 식물 세포, 및 진균 세포를 포함한다. 예시적인 동물 세포로는 제한 없이, 곤충 세포, 어류 세포, 포유동물 세포를 포함한다. 예시적인 포유동물 세포로는 마우스 세포, 인간 세포 및 영장류의 세포를 포함한다. 본원에 기술된 방법에 의해 합성된 핵산은 제한하는 것은 아니지만, 형질감염, 형질도입, 및 전기천공을 비롯한, 당업계에 공지된 각종 방법에 의해 수행됨으로써 세포 내로 전달될 수 있다. 시험되는 세포 기능의 예로는 제한 없이, 세포 증식, 이동/부착, 대사 활성, 및 세포 신호전달 활성 변화를 포함한다.

[0056]

### 고도의 병렬식 핵산 합성

[0057]

본원에서는 혁신적인 합성 플랫폼을 생성하기 위하여 소형화, 병렬화, 및 규소 상의 나노웰 내에서 올리고핵산 합성에서부터 유전자 조립까지의 종단 간 프로세스의 수직적 통합을 이용하는 플랫폼 접근법을 제공한다. 96 웰 플레이트와 동일한 풋프린트를 갖는 본원에 기술된 장치는, 단일의 고도로 병렬화된 작업에서 최대 대략 1,000,000개 이상의 올리고핵산, 또는 10,000종 이상의 유전자를 생성하는 종래의 합성 방법과 비교하여 최대 1,000배 이상만큼 처리량을 증가시킬 수 있는 규소 합성 플랫폼을 제공한다.

[0058]

차세대 서열분석이 출현함에 따라, 고해상도 게놈 데이터는 정상적인 생물학 및 질환 발병기전, 둘 모두에서 다양한 유전자의 생물학적 역할을 철저하게 조사하는 연구를 위해 중요한 인자가 되었다. 분자 생물학의 중심 원리 및 "순차적인 정보의 잔기 하나씩 하나씩 이루어지는 전달"이라는 개념이 이러한 연구의 핵심이다. DNA에서 코딩되는 게놈 정보는 메세지로 전사되고, 이는 주어진 생물학적 경로 내에서 활성 생성물인 단백질로 번역된다.

[0059]

또 다른 흥미로운 연구 영역은 고도로 특이적인 세포 표면에 초점을 맞춘 치료학적 분자의 발견, 개발, 및 제조에 관한 것이다. 고도로 다양한 DNA 서열 라이브러리가 표적화된 치료제의 개발 경로의 핵심이 된다. 유전자 돌연변이체는, 이상적으로 그의 치료학적 표적에 대하여 높은 친화도를 가진 단백질의 고도한 발현을 위해 최적화된 유전자로 종결되는 디자인, 구축, 및 시험 단백질 조작 사이클에서 단백질을 발현시키는 데 사용된다. 한 예로서, 수용체의 결합 포켓을 고려한다. 결합 포켓 내의 모든 잔기의 모든 서열 순열을 시험할 수 있는 능력을 통해 철저히 조사할 수 있고, 성공의 기회를 증가시킬 수 있다. 연구원이 수용체 내의 특정 부위에서 모든 가능한 돌연변이를 생성하기 위해 시도하는 포화 돌연변이유발은 이러한 개발 도전에 대한 한 접근법을 나타낸다. 비록 고가이고, 시간 및 노동 집약적이기는 하지만, 이를 통해 각각의 변이체는 각 위치 내로 도입될 수 있다. 그에 반해, DNA의 짧은 스트레치 또는 소수의 선택된 위치가 광범위하게 변형될 수 있는 조합 돌연변이유발은

편향성을 보이는 불완전한 변이체 레퍼토리를 생성한다.

[0060] 약물 개발 경로를 가속화시키기 위해, 시험에 이용될 수 있는 알맞은 위치에서 의도된 빈도로 이용가능한 원하는 변이체를 포함하는 라이브러리-다시 말해, 정밀 라이브러리를 통해서 비용을 절감할 수 있을 뿐만 아니라, 스크리닝 작업을 완료해서 회송하는 데 걸리는 시간을 단축시킬 수 있다. 본원에서는 각각의 의도된 변이체를 원하는 빈도로 정밀하게 도입하는 올리고핵산 합성 변이체 라이브러리를 합성하는 방법을 제공한다. 최종 사용자에게 이는 비용은 절감하고, 스크리닝 시간은 단축시키면서, 서열 공간을 철저하게 샘플링할 뿐만 아니라, 효율적인 방식으로 상기와 같은 가설을 질의할 수 있는 능력으로 해석된다. 게놈와이드 편집이 각각의 변이체 및 서열 순열이 최적의 기능을 위해 시험될 수 있고, 수천 개의 유전자가 전체 경로 및 게놈을 재구축시켜 약물 발견을 위한 생물학적 시스템을 재조작하는 데 사용될 수 있는 라이브러리에서 중요한 경로를 설명할 수 있다.

[0061] 제1 예에서, 약물 그 자체는 본원에 기술된 방법을 사용함으로써 최적화될 수 있다. 예를 들어, 명시된 항체의 기능을 개선시키기 위해, 상기 항체의 일부를 코딩하는 변이체 올리고핵산 라이브러리를 디자인하고, 합성한다. 이어서, 본원에 기술된 프로세스(예컨대, PCR 돌연변이유발, 이어서, 백터 내로의 삽입)에 의해 항체에 대한 변이체 핵산 라이브러리를 생성할 수 있다. 이어서, 생산 세포주에서 항체를 발현시키고, 증강된 활성에 대해 스크리닝한다. 예시적인 스크린으로는 항원에의 결합 친화도, 안정성, 또는 효과기 기능(예컨대, ADCC, 보체, 또는 아폽토시스) 조절을 조사하는 것을 포함한다. 항체를 최적화시키는 영역의 예로는 제한 없이, Fc 영역, Fab 영역, Fab 영역의 가변 영역, Fab 영역의 불변 영역, 중쇄 또는 경쇄( $V_H$  또는  $V_L$ ),  $V_H$  또는  $V_L$ 의 특정 상보성 결정 영역(CDR)을 포함한다.

[0062] 대안적으로, 최적화되는 분자는 활성화제 또는 경쟁적 억제제로서 사용하기 위한 수용체 결합 에피토프이다. 핵산의 변이체 라이브러리의 합성 후, 핵산의 변이체 라이브러리를 백터 내로 삽입한 후, 세포에서 발현될 수 있다. 수용체 항원을 세포(예컨대, 곤충, 포유동물, 또는 박테리아)에서 발현시킨 후, 정제할 수 있거나, 또는 세포(예컨대, 포유동물)에서 발현시켜 서열 변이로부터 초래되는 기능적 결과를 조사할 수 있다. 기능적 결과로는 제한 없이, 단백질 발현, 결합 친화도 및 안정성 변화를 포함한다. 세포 기능적 결과로는 제한 없이, 생식, 성장, 부착, 사멸, 이동, 에너지 생산, 산소 이용률, 대사 활성, 세포 신호전달, 노화, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합의 변화를 포함한다. 일부 실시양태에서, 최적화를 위해 선택되는 단백질 유형은 효소, 수송체 단백질s, G 단백질 커플링된 수용체, 전암 개폐 이온 채널, 전사 인자, 폴리머라제, 어댑터 단백질(2개의 다른 단백질을 함께 모으는 역할을 하는, 효소 활성이 없는 단백질), 및 세포골격 단백질을 포함한다. 예시적인 효소 유형으로는 제한 없이, 신호전달 효소(예컨대, 단백질 키나제, 단백질 포스파타제, 포스포디에스테라제, 히스톤 데아세틸라제, 및 GTP아제)를 포함한다.

[0063] 본원에서는 전체 경로 또는 전체 게놈에 관여하는 분자에 대한 변이체를 포함하는 변이체 핵산 라이브러리를 제공한다. 예시적인 경로로는 제한 없이, 대사, 세포 사멸, 세포 주기 진행, 면역 세포 활성화, 염증 반응, 혈관신생, 림프구신생, 저산소증 및 산화적 스트레스 반응, 또는 세포 부착/이동 경로를 포함한다. 세포 사멸 경로 중의 단백질의 예로는 제한 없이, Fas, Cadd, 카스파제 3, 카스파제 6, 카스파제 8, 카스파제 9, 카스파제 10, IAP, TNFR1, TNF, TNFR2, NF-kB, TRAFs, ASK, BAD, 및 Akt를 포함한다. 세포 주기 경로 중의 단백질의 예로는 제한 없이, NFKB, E2F, Rb, p53, p21, 사이클린 A, 사이클린 B, 사이클린 D, 사이클린 E, 및 cdc 25를 포함한다. 세포 이동 경로 중의 단백질의 예로는 제한 없이, Ras, Raf, PLC, 코필린, MEK, ERK, MLP, LIMK, ROCK, RhoA, Src, Rac, 미오신 II, ARP2/3, MAPK, PIP2, 인테그린, 탈린, 칸들린, 미그필린 및 필라민을 포함한다.

[0064] 본원에 기술된 방법에 의해 합성된 핵산 라이브러리는 각종 세포 유형에서 발현될 수 있다. 예시적인 세포 유형으로는 원핵생물(예컨대, 박테리아 및 진균) 및 진핵생물(예컨대, 식물 및 동물)을 포함한다. 예시적인 동물로는 제한 없이, 마우스, 토끼, 영장류, 어류 및 곤충을 포함한다. 예시적인 식물로는 제한 없이, 단자엽 식물 및 쌍자엽 식물을 포함한다. 예시적인 식물로는 또한 제한 없이, 미세조류, 켈프, 시아노박테리아, 및 녹조류, 갈조류 및 홍조류, 밀, 담배, 및 옥수수, 벼, 목화, 채소류 및 과실류도 포함한다.

[0065] 본원에 기술된 방법에 의해 합성된 핵산 라이브러리는 질환 상태와 연관된 각종 세포에서 발현될 수 있다. 질환 상태와 연관된 세포로는 세포주, 조직 샘플, 피험체로부터의 1차 세포, 피험체로부터 확장된 배양된 세포, 또는 모델 시스템 중의 세포를 포함한다. 예시적인 모델 시스템으로는 제한 없이, 질환 상태의 식물 및 동물 모델을 포함한다.

[0066] 본원에 기술된 방법에 의해 합성된 핵산 라이브러리는 세포 활성 변화를 사정하는 각종 세포 유형에서 발현될 수 있다. 예시적인 세포 활성으로는 제한 없이, 증식, 사이클 진행, 세포 사멸, 부착, 이동, 생식, 세포 신호전달, 에너지 생산, 산소 이용률, 대사 활성, 및 노화, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합

을 포함한다.

[0067]

질환 상태의 예방, 감소, 또는 치료와 연관된 변이체 분자를 확인하기 위해, 본원에 기술된 변이체 핵산 라이브러리는 질환 상태와 연관된 세포, 또는 질환 상태가 유도될 수 있는 것인 세포에서 발현된다. 일부 경우에서, 작용체는 세포에서 질환 상태를 유도하는 데 사용된다. 질환 상태 유도를 위한 도구의 예로는 제한 없이, Cre/Lox 재조합 시스템, LPS 염증 유도, 및 저혈당증을 유도하는 스트렙토조토신을 포함한다. 질환 상태와 연관된 세포는 특정 질환 상태를 앓는 피험체로부터의 세포 뿐만 아니라, 모델 시스템으로부터의 세포 또는 배양된 세포일 수 있다. 예시적인 질환 상태로는 박테리아성, 진균성, 바이러스성, 자가면역, 또는 증식성 장애(예컨대, 암)를 포함한다. 일부 경우에서, 변이체 핵산 라이브러리는 모델 시스템, 세포주, 또는 피험체로부터 유래된 1차 세포에서 발현되고, 적어도 하나의 세포 활성 변화에 대하여 스크리닝된다. 예시적인 세포 활성으로는 제한 없이, 증식, 사이클 진행, 세포 사멸, 부착, 이동, 생식, 세포 신호전달, 에너지 생산, 산소 이용률, 대사 활성, 및 노화, 자유 라디칼 손상에 대한 반응, 또는 이들의 임의의 조합을 포함한다.

[0068]

## 기판

[0069]

본원에서는 각각의 클러스터가 올리고핵산의 부착 및 합성을 지원하는 것인 복수의 좌위를 포함하는 것인 복수의 클러스터를 포함하는 기판을 제공한다. 본원에서 사용되는 바, "좌위"라는 용어는 단일의 소정의 서열을 코딩하는 올리고뉴클레오티드가 표면으로부터 연장되는 것을 지원하는 구조 상의 별개의 영역을 지칭한다. 일부 경우에서, 좌위는 2차원 표면, 예컨대, 실질적으로 평면 상에 위치한다. 일부 경우에서, 좌위란, 예컨대, 웰, 마이크로웰, 채널, 또는 포스트와 같은 표면 상의 별개의 돌출된 또는 함몰된 부위를 지칭한다. 일부 경우에서, 좌위의 표면은 올리고핵산 합성을 위해 적어도 하나의 뉴클레오티드에, 또는 바람직하게, 올리고핵산 집단의 합성을 위해 동일한 뉴클레오티드로 이루어진 집단에 부착되도록 능동적으로 관능화된 물질을 포함한다. 일부 경우에서, 올리고핵산은 동일한 핵산 서열을 코딩하는 올리고핵산 집단을 지칭한다. 일부 경우에서, 장치의 표면은 기판의 하나 또는 복수의 표면을 포함한다.

[0070]

제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 평균 오류율은 1/1,000 미만, 1/1,250 미만, 1/1,500 미만, 1/2,000 미만, 1/3,000 미만, 또는 그보다 덜 빈번할 수 있다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 평균 오류율은 1/500, 1/600, 1/700, 1/800, 1/900, 1/1,000, 1/1,100, 1/1,200, 1/1,250, 1/1,300, 1/1,400, 1/1,500, 1/1,600, 1/1,700, 1/1,800, 1/1,900, 1/2,000, 1/3,000 미만, 또는 그 미만이다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 평균 오류율은 1/1,000 미만이다.

[0071]

일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 총 오류율은 소정의 서열과 비교하여 1/500, 1/600, 1/700, 1/800, 1/900, 1/1,000, 1/1,100, 1/1,200, 1/1,250, 1/1,300, 1/1,400, 1/1,500, 1/1,600, 1/1,700, 1/1,800, 1/1,900, 1/2,000, 1/3,000 미만, 또는 그 미만일 수 있다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 총 오류율은 1/500, 1/600, 1/700, 1/800, 1/900, 또는 1/1,000 미만이다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산에 대한 총 오류율은 1/1,000 미만이다.

[0072]

일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산을 위해 오류 보정 효소가 사용될 수 있다. 일부 경우에서, 오류 보정된 올리고핵산의 총 오류율은 소정의 서열과 비교하여 1/500, 1/600, 1/700, 1/800, 1/900, 1/1,000, 1/1,100, 1/1,200, 1/1,300, 1/1,400, 1/1,500, 1/1,600, 1/1,700, 1/1,800, 1/1,900, 1/2,000, 1/3,000 미만, 또는 그 미만일 수 있다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산의 경우, 오류 보정된 올리고핵산의 총 오류율은 1/500, 1/600, 1/700, 1/800, 1/900, 또는 1/1,000 미만일 수 있다. 일부 경우에서, 제공된 시스템 및 방법을 사용하여 라이브러리 내에서 합성된 올리고핵산의 경우, 오류 보정된 올리고핵산의 총 오류율은 1/1,000 미만일 수 있다.

[0073]

오류율이 유전자 변이체의 라이브러리의 제조를 위한 유전자 합성의 값을 제한할 수 있다. 오류율이 1/300인 경우, 1,500개의 염기쌍을 가진 유전자에서 클론 중 약 0.7%가 정확할 것이다. 올리고뉴클레오티드 합성으로부터의 오류 대부분이 프레임 쉬프트 돌연변이를 초래하기 때문에, 상기 라이브러리에서 99% 초파의 클론이 전장의 단백질을 생성하지 못할 것이다. 오류율을 75%만큼 감소시키면, 정확한 클론의 분율이 40배만큼 증가하게 될 것이다. 개선된 합성 품질, 및 대량 병렬식 및 시간 효율적 방식으로 가능해지는 오류 보정 방법의 적용가능성, 이 둘 모두에 기인하여, 본 개시내용의 방법 및 조성물은 통상적으로 관찰된 유전자 합성 방법보다 더 낮은 오류율로 큰 올리고뉴클레오티드 및 유전자 라이브러리를 빠르게 드 노보 합성할 수 있다. 따라서, 라이브러리는 라이브러리에 걸쳐, 또는 라이브러리의 80%, 85%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, 99.9%,

99.95%, 99.98%, 99.99% 초과, 또는 그 초과인 것에 걸쳐 1/300, 1/400, 1/500, 1/600, 1/700, 1/800, 1/900, 1/1,000, 1/1,250, 1/1,500, 1/2,000, 1/2,500, 1/3,000, 1/4,000, 1/5,000, 1/6,000, 1/7,000, 1/8,000, 1/9,000, 1/10,000, 1/12,000, 1/15,000, 1/20,000, 1/25,000, 1/30,000, 1/40,000, 1/50,000, 1/60,000, 1/70,000, 1/80,000, 1/90,000, 1/100,000, 1/125,000, 1/150,000, 1/200,000, 1/300,000, 1/400,000, 1/500,000, 1/600,000, 1/700,000, 1/800,000, 1/900,000, 1/1,000,000 미만, 또는 그 미만인 염기 삽입, 결실, 치환 또는 총 오류율로 합성될 수 있다. 본 개시내용의 방법 및 조성물은 추가로 라이브러리의 적어도 서브세트 내의 올리고뉴클레오티드 또는 유전자의 적어도 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, 99.9%, 99.95%, 99.98%, 99.99% 이상이 소정의/미리 선택된 서열과 비교하여 오류 부재 서열에 관한 것과 연관된 낮은 오류율을 가진 큰 합성 올리고뉴클레오티드 및 유전자 라이브러리에 관한 것이다. 일부 경우에서, 라이브러리 내에서 단리된 부피 내의 올리고뉴클레오티드 또는 유전자의 적어도 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.95%, 99.98%, 99.99% 이상이 동일한 서열을 가진다. 일부 경우에서, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% 초과, 또는 그 초과의 유사성 또는 동일성과 관련된 임의의 올리고뉴클레오티드 또는 유전자의 적어도 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, 99.9%, 99.95%, 99.98%, 99.99% 이상이 동일한 서열을 가진다. 일부 경우에서, 올리고뉴클레오티드 또는 유전자 상의 특정된 좌우와 관련된 오류율이 최적화된다. 따라서, 큰 라이브러리의 일부로서 하나 이상의 올리고뉴클레오티드 또는 유전자의 주어진 좌우 또는 복수의 선택된 좌우들은 각각 1/300, 1/400, 1/500, 1/600, 1/700, 1/800, 1/900, 1/1,000, 1/1,250, 1/1,500, 1/2,000, 1/2,500, 1/3,000, 1/4,000, 1/5,000, 1/6,000, 1/7,000, 1/8,000, 1/9,000, 1/10,000, 1/12,000, 1/15,000, 1/20,000, 1/25,000, 1/30,000, 1/40,000, 1/50,000, 1/60,000, 1/70,000, 1/80,000, 1/90,000, 1/100,000, 1/125,000, 1/150,000, 1/200,000, 1/300,000, 1/400,000, 1/500,000, 1/600,000, 1/700,000, 1/800,000, 1/900,000, 1/1,000,000 미만, 또는 그 미만인 오류율을 가질 수 있다. 다양한 경우에서, 상기 오류 최적화된 좌우는 적어도 1, 2, 3, 4, 5, 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, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 30,000, 50,000, 75,000, 100,000, 500,000, 1,000,000, 2,000,000, 3,000,000종 이상의 좌우를 포함할 수 있다. 오류 최적화된 좌우는 적어도 1, 2, 3, 4, 5, 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, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,500, 2,000, 2,500, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 30,000, 75,000, 100,000, 500,000, 1,000,000, 2,000,000, 3,000,000개 이상의 올리고뉴클레오티드 또는 유전자에 분포될 수 있다.

[0074] 오류율은 오류 보정하에 또는 오류 보정 없이 달성될 수 있다. 오류율은 라이브러리에 걸쳐, 또는 라이브러리의 80%, 85%, 90%, 93%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.8%, 99.9%, 99.95%, 99.98%, 99.99% 초과, 또는 그 초과인 것에 걸쳐 달성될 수 있다.

[0075] 본원에서는 통상의 지지체 상의 주소지정가능한 위치에 상이한 소정의 서열을 갖는, 복수의 올리고핵산의 합성을 지원하는 표면을 포함할 수 있는 구조물을 제공한다. 일부 경우에서, 장치는 2,000; 5,000; 10,000; 20,000; 30,000; 50,000; 75,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000; 2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; 10,000,000개 초과, 또는 그 초과의 동일하지 않은 올리고핵산의 합성을 지원한다. 일부 경우에서, 장치는 상이한 서열을 코딩하는, 2,000; 5,000; 10,000; 20,000; 30,000; 50,000; 75,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000; 2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; 10,000,000개 초과, 또는 그 초과의 올리고핵산의 합성을 지원한다. 일부 경우에서, 올리고핵산 중 적어도 일부는 동일한 서열을 갖거나, 또는 동일한 서열로 합성되도록 구성된다.

[0076] 본원에서는 길이가 약 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 또는 2,000개의 염기인 올리고핵산의 제조 및 성장을 위한 방법 및 장치를 제공한다. 일부 경우에서, 형성된 올리고핵산의 길이는 약 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 또는 225개의 염기이다. 올리고핵산의 길이는 적어도 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100개의 염기일 수 있다. 올리고핵산의 길이는 10 내지 225개의 염기, 12 내지 100개의 염기, 20

내지 150개의 염기, 20 내지 130개의 염기, 또는 30 내지 100개의 염기일 수 있다.

[0077] 일부 경우에서, 올리고핵산은, 각 좌위가 올리고핵산 집단의 합성을 지원하는 것인, 기판의 상이한 좌위 상에서 합성된다. 일부 경우에서, 각 좌위는 또 다른 좌위 상에서 성장된 올리고핵산 집단과는 상이한 서열을 갖는 올리고핵산 집단의 합성을 지원한다. 일부 경우에서, 장치의 좌위는 복수의 클러스터 내에 위치한다. 일부 경우에서, 장치는 적어도 10, 500, 1,000, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 11,000, 12,000, 13,000, 14,000, 15,000, 20,000, 30,000, 40,000, 50,000개 이상의 클러스터를 포함한다. 일부 경우에서, 장치는 2,000; 5,000; 10,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,300,000; 1,400,000; 1,500,000; 1,600,000; 1,700,000; 1,800,000; 1,900,000; 2,000,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000; 2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; 또는 10,000,000개 초과, 또는 그 초과의 상이한 좌위를 포함한다. 일부 경우에서, 장치는 약 10,000개의 상이한 좌위를 포함한다. 단일 클러스터 내의 좌위의 양은 다른 경우에는 달라진다. 일부 경우에서, 각각의 클러스터는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 130, 150, 200, 300, 400, 500종 이상의 좌위를 포함한다. 일부 경우에서, 각각의 클러스터는 약 50-500개의 좌위를 포함한다. 일부 경우에서, 각각의 클러스터는 약 100-200개의 좌위. 일부 경우에서, 각각의 클러스터는 약 100-150개의 좌위를 포함한다. 일부 경우에서, 각각의 클러스터는 약 109, 121, 130 또는 137개의 좌위를 포함한다. 일부 경우에서, 각각의 클러스터는 약 19, 20, 61, 64종 이상의 좌위를 포함한다.

[0078] 장치 상에서 합성된 상이한 올리고핵산의 개수는 기판에서 이용가능한 상이한 좌위의 개수에 의존할 수 있다. 일부 경우에서, 장치의 클러스터 내의 좌위의 밀도는 적어도 또는 약 1개의 좌위/ $\text{mm}^2$ , 10개의 좌위/ $\text{mm}^2$ , 25개의 좌위/ $\text{mm}^2$ , 50개의 좌위/ $\text{mm}^2$ , 65개의 좌위/ $\text{mm}^2$ , 75개의 좌위/ $\text{mm}^2$ , 100개의 좌위/ $\text{mm}^2$ , 130개의 좌위/ $\text{mm}^2$ , 150개의 좌위/ $\text{mm}^2$ , 175개의 좌위/ $\text{mm}^2$ , 200개의 좌위/ $\text{mm}^2$ , 300개의 좌위/ $\text{mm}^2$ , 400개의 좌위/ $\text{mm}^2$ , 500개의 좌위/ $\text{mm}^2$ , 1,000개의 좌위/ $\text{mm}^2$  이상이다. 일부 경우에서, 장치는 약 10개의 좌위/ $\text{mm}^2$  내지 약 10개의 좌위/약 500  $\text{mm}^2$ , 약 25개의 좌위/ $\text{mm}^2$  내지 약 25개의 좌위/약 400  $\text{mm}^2$ , 약 50개의 좌위/ $\text{mm}^2$  내지 약 50개의 좌위/약 500  $\text{mm}^2$ , 약 100개의 좌위/ $\text{mm}^2$  내지 약 100개의 좌위/약 500  $\text{mm}^2$ , 약 150개의 좌위/ $\text{mm}^2$  내지 약 150개의 좌위/약 500  $\text{mm}^2$ , 약 10개의 좌위/ $\text{mm}^2$  내지 약 10개의 좌위/약 250  $\text{mm}^2$ , 약 50개의 좌위/ $\text{mm}^2$  내지 약 50개의 좌위/약 250  $\text{mm}^2$ , 약 10개의 좌위/ $\text{mm}^2$  내지 약 10개의 좌위/약 200  $\text{mm}^2$ , 또는 약 50개의 좌위/ $\text{mm}^2$  내지 약 50개의 좌위/약 200  $\text{mm}^2$ 를 포함한다. 일부 경우에서, 클러스터 내의 2개의 인접한 좌위의 중심부 사이의 거리는 약 10  $\text{um}$  내지 약 500  $\text{um}$ , 약 10  $\text{um}$  내지 약 200  $\text{um}$ , 또는 약 10  $\text{um}$  내지 약 100  $\text{um}$ 이다. 일부 경우에서, 인접한 좌위의 두 중심부 사이의 거리는 약 10  $\text{um}$ , 20  $\text{um}$ , 30  $\text{um}$ , 40  $\text{um}$ , 50  $\text{um}$ , 60  $\text{um}$ , 70  $\text{um}$ , 80  $\text{um}$ , 90  $\text{um}$  또는 100  $\text{um}$  초과이다. 일부 경우에서, 2개의 인접한 좌위의 중심부 사이의 거리는 약 200  $\text{um}$ , 150  $\text{um}$ , 100  $\text{um}$ , 80  $\text{um}$ , 70  $\text{um}$ , 60  $\text{um}$ , 50  $\text{um}$ , 40  $\text{um}$ , 30  $\text{um}$ , 20  $\text{um}$  또는 10  $\text{um}$  미만이다. 일부 경우에서, 각 좌위의 너비는 약 0.5  $\text{um}$ , 1  $\text{um}$ , 2  $\text{um}$ , 3  $\text{um}$ , 4  $\text{um}$ , 5  $\text{um}$ , 6  $\text{um}$ , 7  $\text{um}$ , 8  $\text{um}$ , 9  $\text{um}$ , 10  $\text{um}$ , 20  $\text{um}$ , 30  $\text{um}$ , 40  $\text{um}$ , 50  $\text{um}$ , 60  $\text{um}$ , 70  $\text{um}$ , 80  $\text{um}$ , 90  $\text{um}$  또는 100  $\text{um}$ 이다. 일부 경우에서, 각 좌위의 너비는 약 0.5  $\text{um}$  내지 100  $\text{um}$ , 약 0.5  $\text{um}$  내지 50  $\text{um}$ , 약 10  $\text{um}$  내지 75  $\text{um}$ , 또는 약 0.5  $\text{um}$  내지 50  $\text{um}$ 이다.

[0079] 일부 경우에서, 장치내 클러스터의 밀도는 적어도 또는 약 1개의 클러스터/100  $\text{mm}^2$ , 1개의 클러스터/10  $\text{mm}^2$ , 1개의 클러스터/5  $\text{mm}^2$ , 1개의 클러스터/4  $\text{mm}^2$ , 1개의 클러스터/3  $\text{mm}^2$ , 1개의 클러스터/2  $\text{mm}^2$ , 1개의 클러스터/1  $\text{mm}^2$ , 2개의 클러스터/1  $\text{mm}^2$ , 3개의 클러스터/1  $\text{mm}^2$ , 4개의 클러스터/1  $\text{mm}^2$ , 5개의 클러스터/1  $\text{mm}^2$ , 10개의 클러스터/1  $\text{mm}^2$ , 50개의 클러스터/1  $\text{mm}^2$  이상이다. 일부 경우에서, 장치는 약 1개의 클러스터/10  $\text{mm}^2$  내지 약 10개의 클러스터/1  $\text{mm}^2$ 를 포함한다. 일부 경우에서, 2개의 인접한 클러스터의 중심부 사이의 거리는 약 50  $\text{um}$ , 100  $\text{um}$ , 200  $\text{um}$ , 500  $\text{um}$ , 1,000  $\text{um}$ , 또는 2000  $\text{um}$ , 또는 5000  $\text{um}$  미만이다. 일부 경우에서, 2개의 인접한 클러스터의 중심부 사이의 거리는 약 50  $\text{um}$  내지 약 100  $\text{um}$ , 약 50  $\text{um}$  내지 약 200  $\text{um}$ , 약 50  $\text{um}$  내지 약 300  $\text{um}$ , 약 50  $\text{um}$  내지 약 500  $\text{um}$ , 및 약 100  $\text{um}$  내지 약 2,000  $\text{um}$ 이다. 일부 경우에서, 2개의 인접한 클러스터의 중심부 사이의 거리는 0.05  $\text{mm}$  내지 약 50  $\text{mm}$ , 약 0.05  $\text{mm}$  내지 약 10  $\text{mm}$ , 약 0.05  $\text{mm}$  내지 약 5  $\text{mm}$ , 약 0.05  $\text{mm}$  내지 약 4  $\text{mm}$ , 약 0.05  $\text{mm}$  내지 약 3  $\text{mm}$ , 약 0.05  $\text{mm}$  내지 약 2  $\text{mm}$ , 약 0.1  $\text{mm}$  내지 10  $\text{mm}$ , 약 0.2  $\text{mm}$  내지 10  $\text{mm}$ , 약 0.3  $\text{mm}$  내지 약 10  $\text{mm}$ , 약 0.4  $\text{mm}$  내지 약 10  $\text{mm}$ , 약 0.5  $\text{mm}$  내지 약 10  $\text{mm}$ , 약 0.5  $\text{mm}$  내지 약 5  $\text{mm}$ , 또는 약 0.5  $\text{mm}$  내지 약 2  $\text{mm}$ 이다. 일부 경우에서, 1차원 선을 따라 각각의 클러스터의 직경 또는 너비는 약 0.5 내지 2  $\text{mm}$ , 약 0.5 내지 1  $\text{mm}$ , 또는 약 1 내지 2  $\text{mm}$ 이다. 일부 경우에서, 1차원 선을 따라 각각의 클러스터의 직경 또는 너비는 약 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 또는 2  $\text{mm}$ 이다. 일부 경우에서, 1차원 선을 따라 각각의 클러스터의 내부 직경 또는 너비는 약 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.15, 1.2, 1.3,

1.4, 1.5, 1.6, 1.7, 1.8, 1.9 또는 2 mm이다.

[0080] 장치의 크기는 대략 표준 96 웰 플레이트의 크기, 예를 들어, 약 100 내지 200 mm x 약 50 내지 150 mm일 수 있다. 일부 경우에서, 장치의 직경은 약 1,000 mm, 500 mm, 450 mm, 400 mm, 300 mm, 250 mm, 200 mm, 150 mm, 100 mm 또는 50 mm 이하이다. 일부 경우에서, 장치의 직경은 약 25 mm 내지 1,000 mm, 약 25 mm 내지 약 800 mm, 약 25 mm 내지 약 600 mm, 약 25 mm 내지 약 500 mm, 약 25 mm 내지 약 400 mm, 약 25 mm 내지 약 300 mm, 또는 약 25 mm 내지 약 200 mm이다. 장치 크기의 비제한적인 예로는 약 300 mm, 200 mm, 150 mm, 130 mm, 100 mm, 76 mm, 51 mm 및 25 mm를 포함한다. 일부 경우에서, 장치의 평면적은 적어도 약 100 mm<sup>2</sup>; 200 mm<sup>2</sup>; 500 mm<sup>2</sup>; 1,000 mm<sup>2</sup>; 2,000 mm<sup>2</sup>; 5,000 mm<sup>2</sup>; 10,000 mm<sup>2</sup>; 12,000 mm<sup>2</sup>; 15,000 mm<sup>2</sup>; 20,000 mm<sup>2</sup>; 30,000 mm<sup>2</sup>; 40,000 mm<sup>2</sup>; 50,000 mm<sup>2</sup> 이상이다. 일부 경우에서, 장치의 두께는 약 50 mm 내지 약 2,000 mm, 약 50 mm 내지 약 1,000 mm, 약 100 mm 내지 약 1,000 mm, 약 200 mm 내지 약 1,000 mm, 또는 약 250 mm 내지 약 1,000 mm이다. 장치 두께의 비제한적인 예로는 275 mm, 375 mm, 525 mm, 625 mm, 675 mm, 725 mm, 775 mm 및 925 mm를 포함한다. 일부 경우에서, 장치의 두께는 직경에 따라 달라지며, 기판 조정에 의존한다. 예를 들어, 규소 이외의 다른 물질을 포함하는 장치는 직경이 동일한 규소 장치와 다른 두께를 가진다. 장치 두께는 사용되는 물질의 기계적 강도에 의해 측정될 수 있고, 장치는 취급하는 동안 균열 없이 그 자신의 중량을 지탱할 수 있을 정도로 충분히 두꺼워야 한다. 일부 경우에서, 구조물은 본원에 기술된 복수의 장치를 포함한다.

### 표면 재료

[0082] 본원에서는 본원에 기술된 방법 및 조성물에 적합한 임의의 다양한 물질로부터 제작된 기판, 장치 및 반응기를 제공한다. 특정 경우에서, 장치 재료는 낮은 수준의 뉴클레오티드 결합을 보이도록 제작된다. 일부 경우에서, 장치 재료는 높은 수준의 뉴클레오티드 결합을 보이는 상이한 표면을 생성하도록 개질된다. 일부 경우에서, 장치 재료는 가시광 및/또는 UV 광에 대하여 투과성을 나타낸다. 일부 경우에서, 장치 재료는 충분한 전도성을 띠고, 예컨대, 기판 전체 또는 그 일부에 걸쳐 균일한 전지강을 형성할 수 있다. 일부 경우에서, 전도성 재료는 전기 접지에 연결된다. 일부 경우에서, 장치는 열 전도성이거나, 또는 절연 처리가 된 것이다. 일부 경우에서, 재료는 화학적 또는 생화학적 반응, 예를 들어, 올리고핵산 합성 반응 프로세스를 견딜 수 있을 정도로 내화학성 및 내열성을 띤다. 일부 경우에서, 장치는 가요성 재료를 포함한다. 가요성 재료로는 제한 없이, 개질된 나일론, 비개질된 나일론, 니트로셀룰로스, 폴리프로필렌 등을 포함한다. 일부 경우에서, 장치는 강성 재료를 포함한다. 강성 재료로는 제한 없이, 유리, 훈연 실리카, 규소, 이산화규소, 질화규소, 플라스틱(예를 들어, 폴리테트라플루오로에틸렌, 폴리프로필렌, 폴리스티렌, 폴리카보네이트, 및 그의 브랜드 등), 및 금속(예를 들어, 금, 백금 등)을 포함한다. 일부 경우에서, 장치는 규소, 폴리스티렌, 아가로스, 텍스트란, 셀룰로스계 중합체, 폴리아크릴아미드, 폴리디메틸실록산(PDMS), 유리, 또는 이들의 임의의 조합을 포함하는 재료로부터 제작된다. 일부 경우에서, 장치는 본원에 열거된 재료, 또는 당업계에 공지된 임의의 다른 적합한 재료의 조합을 이용하여 제작된다.

### 표면 아키텍처

[0084] 본원에서는 돌출된 및/또는 함몰된 특징부를 포함하는 장치를 제공한다. 상기와 같은 특징부를 갖는 것의 한가지 이점은 올리고핵산 합성을 지원할 수 있는 표면적이 증가한다는 점이다. 일부 경우에서, 돌출된 및/또는 함몰된 특징부를 갖는 장치는 3차원 기판으로 지칭된다. 일부 경우에서, 3차원 장치는 하나 이상의 채널을 포함한다. 일부 경우에서, 하나 이상의 좌위는 채널을 포함한다. 일부 경우에서, 채널은 증착 장치, 예컨대, 올리고핵산 합성기를 통한 시약 증착에 접근가능하다. 일부 경우에서, 시약 및/또는 유체는 유체 소통하는 하나 이상의 채널 중 더욱 큰 웰 중에 수집된다. 예를 들어, 장치는 클러스터를 포함하는 복수의 좌위에 상응하는 복수의 채널을 포함하고, 복수의 채널은 클러스터의 한 웰과 유체 소통한다. 일부 방법에서, 올리고핵산 라이브러리는 클러스터의 복수의 좌위에서 합성된다.

[0085] 일부 경우에서, 구조물은 표면 상에서의 올리고핵산 합성을 위해 조절된 방식의 유동 및 물질 전달 경로를 허용하도록 구성된다. 일부 경우에서, 장치의 구성은 올리고핵산 합성 동안의 조절된 방식 및 고른 분포의 물질 전달 경로, 화학적 노출 시간, 및/또는 세척 효능을 허용한다. 일부 경우에서, 장치의 구성은 예를 들어, 성장하는 올리고핵산을 위해 충분한 부피를 제공함으로써 성장하는 올리고핵산에 의해 배제된 부피가, 올리고핵산을 성장시키는 데 이용가능하거나, 또는 그에 적합한 초기의 이용가능한 부피의 50, 45, 40, 35, 30, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1% 이하보다 더 큰 부피를 차지하지 않도록 함으로써 스윕 효율(sweep efficiency)을 증가시킬 수 있다. 일부 경우에서, 3차원 구조물은 관리된 방식의 유체 유동이 화학적 노출을 빠르게 교체할 수 있게 허용한다.

- [0086] 본원에서는 1 fM, 5 fM, 10 fM, 25 fM, 50 fM, 75 fM, 100 fM, 200 fM, 300 fM, 400 fM, 500 fM, 600 fM, 700 fM, 800 fM, 900 fM, 1 pM, 5 pM, 10 pM, 25 pM, 50 pM, 75 pM, 100 pM, 200 pM, 300 pM, 400 pM, 500 pM, 600 pM, 700 pM, 800 pM, 900 pM 이상의 DNA 양을 합성하는 방법을 제공한다. 일부 경우에서, 올리고뉴클레오티드 라이브러리는 유전자의 약 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 또는 100%의 길이를 포괄할 수 있다. 유전자는 최대 약 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 100%까지로 달라질 수 있다.
- [0087] 동일하지 않은 올리고핵산은 집합적으로 유전자의 적어도 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 100%를 코딩할 수 있다. 일부 경우에서, 올리고핵산은 유전자의 50%, 60%, 70%, 80%, 85%, 90%, 95% 이상의 서열을 코딩할 수 있다. 일부 경우에서, 올리고핵산은 유전자의 80%, 85%, 90%, 95% 이상의 서열을 코딩할 수 있다.
- [0088] 일부 경우에서, 분리는 물리적 구조에 의해 달성된다. 일부 경우에서, 분리는 표면을 차별적으로 관능화하여 올리고핵산 합성을 위한 능동적 및 수동적 영역을 생성함으로써 달성된다. 차별적 관능화는 또한 기판 표면에 걸쳐 교대로 소수성을 띠게 만들어 중착된 시약의 비딩(beading) 또는 습윤화를 유발하는 물 접촉각 효과를 일으킴으로써 달성된다. 더욱 큰 구조물을 사용하면, 이웃 스콧의 시약에 의한 상이한 올리고핵산 합성 위치의 튀김(splashing) 또는 교차 오염을 감소시킬 수 있다. 일부 경우에서, 장치, 예컨대, 올리고핵산 합성기는 상이한 올리고핵산 합성 위치에 시약을 중착시키는 데 사용된다. 3차원 특징부를 갖는 기판은 낮은 오류율(예컨대, 약 1:500, 1:1000, 1:1500, 1:2,000; 1:3,000; 1:5,000; 또는 1:10,000 미만)로 다수의 올리고핵산(예컨대, 약 10,000개 초과)을 합성할 수 있도록 하는 방식으로 구성된다. 일부 경우에서, 장치는 1 mm<sup>2</sup>당 약 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 300, 400 또는 500개의 특징부, 또는 그 초과의 밀도로 특징부를 포함한다.
- [0089] 장치의 한 웰은 기판의 또 다른 웰과 동일하거나, 또는 상이한 너비, 높이, 및/또는 부피를 가질 수 있다. 장치의 한 채널은 기판의 또 다른 채널과 동일하거나, 또는 상이한 너비, 높이, 및/또는 부피를 가질 수 있다. 일부 경우에서, 클러스터의 너비는 약 0.05 mm 내지 약 50 mm, 약 0.05 mm 내지 약 10 mm, 약 0.05 mm 내지 약 5 mm, 약 0.05 mm 내지 약 4 mm, 약 0.05 mm 내지 약 3 mm, 약 0.05 mm 내지 약 2 mm, 약 0.05 mm 내지 약 1 mm, 약 0.05 mm 내지 약 0.5 mm, 약 0.05 mm 내지 약 0.1 mm, 약 0.1 mm 내지 10 mm, 약 0.2 mm 내지 10 mm, 약 0.3 mm 내지 약 10 mm, 약 0.4 mm 내지 약 10 mm, 약 0.5 mm 내지 10 mm, 약 0.5 mm 내지 약 5 mm, 또는 약 0.5 mm 내지 약 2 mm이다. 일부 경우에서, 클러스터를 포함하는 웰의 너비는 약 0.05 mm 내지 약 50 mm, 약 0.05 mm 내지 약 10 mm, 약 0.05 mm 내지 약 5 mm, 약 0.05 mm 내지 약 4 mm, 약 0.05 mm 내지 약 3 mm, 약 0.05 mm 내지 약 2 mm, 약 0.05 mm 내지 약 1 mm, 약 0.05 mm 내지 약 0.5 mm, 약 0.05 mm 내지 약 0.1 mm, 약 0.1 mm 내지 10 mm, 약 0.2 mm 내지 10 mm, 약 0.3 mm 내지 약 10 mm, 약 0.4 mm 내지 약 10 mm, 약 0.5 mm 내지 약 10 mm, 약 0.5 mm 내지 약 5 mm, 또는 약 0.5 mm 내지 약 2 mm이다. 일부 경우에서, 클러스터의 너비는 약 5 mm, 4 mm, 3 mm, 2 mm, 1 mm, 0.5 mm, 0.1 mm, 0.09 mm, 0.08 mm, 0.07 mm, 0.06 mm 또는 0.05 mm 이하이다. 일부 경우에서, 클러스터의 너비는 약 1.0 내지 1.3 mm이다. 일부 경우에서, 클러스터의 너비는 약 1.150 mm이다. 일부 경우에서, 웰의 너비는 약 5 mm, 4 mm, 3 mm, 2 mm, 1 mm, 0.5 mm, 0.1 mm, 0.09 mm, 0.08 mm, 0.07 mm, 0.06 mm 또는 0.05 mm 이하이다. 일부 경우에서, 웰의 너비는 약 1.0 내지 1.3 mm이다. 일부 경우에서, 웰의 너비는 약 1.150 mm이다. 일부 경우에서, 클러스터의 너비는 약 0.08 mm이다. 일부 경우에서, 웰의 너비는 약 0.08 mm이다. 클러스터의 너비는 2차원 또는 3차원 기판 내의 클러스터를 지칭할 수 있다.
- [0090] 일부 경우에서, 웰의 높이는 약 20  $\mu$ m 내지 약 1,000  $\mu$ m, 약 50  $\mu$ m 내지 약 1,000  $\mu$ m, 약 100  $\mu$ m 내지 약 1,000  $\mu$ m, 약 200  $\mu$ m 내지 약 1,000  $\mu$ m, 약 300  $\mu$ m 내지 약 1,000  $\mu$ m, 약 400  $\mu$ m 내지 약 1,000  $\mu$ m, 또는 약 500  $\mu$ m 내지 약 1,000  $\mu$ m이다. 일부 경우에서, 웰의 높이는 약 1,000  $\mu$ m 미만, 약 900  $\mu$ m 미만, 약 800  $\mu$ m 미만, 약 700  $\mu$ m 미만, 또는 약 600  $\mu$ m 미만이다.
- [0091] 일부 경우에서, 장치는 클러스터 내의 복수의 좌위에 상응하는 복수의 채널을 포함하고, 여기서, 채널의 높이 또는 깊이는 약 5  $\mu$ m 내지 약 500  $\mu$ m, 약 5  $\mu$ m 내지 약 400  $\mu$ m, 약 5  $\mu$ m 내지 약 300  $\mu$ m, 약 5  $\mu$ m 내지 약 200  $\mu$ m, 약 5  $\mu$ m 내지 약 100  $\mu$ m, 약 5  $\mu$ m 내지 약 50  $\mu$ m, 또는 약 10  $\mu$ m 내지 약 50  $\mu$ m이다. 일부 경우에서, 채널의 높이는 100  $\mu$ m 미만, 80  $\mu$ m 미만, 60  $\mu$ m 미만, 40  $\mu$ m 미만, 또는 20  $\mu$ m 미만이다.
- [0092] 일부 경우에서, 채널, 좌위(예컨대, 실질적으로 평평한 기판에서), 또는 채널 및 좌위 둘 모두(예컨대, 좌위가 채널에 상응하는 것인 3차원 장치에서)의 직경은 약 1  $\mu$ m 내지 약 1,000  $\mu$ m, 약 1  $\mu$ m 내지 약 500  $\mu$ m, 약 1  $\mu$ m 내지 약 200  $\mu$ m, 약 1  $\mu$ m 내지 약 100  $\mu$ m, 약 5  $\mu$ m 내지 약 100  $\mu$ m, 또는 약 10  $\mu$ m 내지 약 100  $\mu$ m, 예를

들어, 약 90  $\mu\text{m}$ , 80  $\mu\text{m}$ , 70  $\mu\text{m}$ , 60  $\mu\text{m}$ , 50  $\mu\text{m}$ , 40  $\mu\text{m}$ , 30  $\mu\text{m}$ , 20  $\mu\text{m}$ , 또는 10  $\mu\text{m}$ 이다. 일부 경우에서, 채널, 좌위 또는 채널 및 좌위 둘 모두의 직경은 약 100  $\mu\text{m}$ , 90  $\mu\text{m}$ , 80  $\mu\text{m}$ , 70  $\mu\text{m}$ , 60  $\mu\text{m}$ , 50  $\mu\text{m}$ , 40  $\mu\text{m}$ , 30  $\mu\text{m}$ , 20  $\mu\text{m}$ , 또는 10  $\mu\text{m}$  미만이다. 일부 경우에서, 2개의 인접한 채널, 좌위, 또는 채널 및 좌위의 중심부 사이의 거리는 약 1  $\mu\text{m}$  내지 약 500  $\mu\text{m}$ , 약 1  $\mu\text{m}$  내지 약 200  $\mu\text{m}$ , 약 1  $\mu\text{m}$  내지 약 100  $\mu\text{m}$ , 약 5  $\mu\text{m}$  내지 약 200  $\mu\text{m}$ , 약 5  $\mu\text{m}$  내지 약 100  $\mu\text{m}$ , 약 5  $\mu\text{m}$  내지 약 50  $\mu\text{m}$ , 또는 약 5  $\mu\text{m}$  내지 약 30  $\mu\text{m}$ , 예를 들어, 약 20  $\mu\text{m}$ 이다.

### [0093] 표면 개질

다양한 경우에서, 표면 개질은 장치 표면, 또는 장치 표면의 선택된 부위 또는 영역의 하나 이상의 화학적 및/ 또는 물리적 특성을 변화시키기 위해 추가 또는 공제 프로세스에 의해 표면을 화학적으로 및/또는 물리적으로 변경시키는 데 사용된다. 예를 들어, 표면 개질으로는 제한 없이, (1) 표면의 습윤화 특성을 변화시키는 것, (2) 표면을 관능화하는 것, 즉, 표면 작용기를 제공하거나, 개질시키거나, 또는 치환하는 것, (3) 표면을 탈관능화하는 것, 즉, 표면 작용기를 제거하는 것, (4) 다르게는, 예컨대, 에칭을 통해 표면의 화학적 조성을 변경시키는 것, (5) 표면 조도를 증가 또는 감소시키는 것, (6) 표면 상에 코팅, 예컨대, 표면의 습윤화 특성과 상이한 습윤화 특성을 나타내는 코팅을 제공하는 것, 및/또는 (7) 표면 상에 미립자를 침착시키는 것을 포함한다.

일부 경우에서, 표면 위에 (부착 프로모터로 지정되는) 화학층을 추가하는 것이 기판의 표면 상에의 좌위의 구조화된 패턴화를 촉진시킨다. 부착 프로모션 적용을 위한 예시적인 표면으로는 제한 없이, 유리, 규소, 이산화규소 및 질화규소를 포함한다. 일부 경우에서, 부착 프로모터는 표면 에너지가 높은 화학물질이다. 일부 경우에서, 제2 화학층을 기판의 표면 상에 증착시킨다. 일부 경우에서, 제2 화학층의 표면 에너지는 낮다. 일부 경우에서, 표면 상에 코팅된 화학층의 표면 에너지는 표면 상의 액적의 국재화를 지원한다. 선택된 패턴화 배열에 의존하여, 좌위의 근접성 및/또는 좌위에서의 유체 접촉 면적은 변경가능하다.

일부 경우에서, 예컨대, 올리고핵산 합성을 위해 그 위에 핵산 또는 다른 모이어티가 증착되어 있는 것인 장치 표면, 또는 분해된 좌위는 평활하거나, 또는 실질적으로 평평하거나(예컨대, 2차원), 또는 요철, 예컨대, 돌출된 또는 함몰된 특징부(예컨대, 3차원 특징부)를 가진다. 일부 경우에서, 장치 표면은 하나 이상의 상이한 화합물 층으로 개질된다. 관심의 대상이 되는 상기와 같은 개질 층으로는 제한 없이, 무기 및 유기 층, 예컨대, 금속, 금속 산화물, 중합체, 소형 유기 분자 등을 포함한다. 비제한적인 중합체 층은 웹티드, 단백질, 핵산 또는 그의 모방체(예컨대, 웹티드 핵산 등), 다당류, 인지질, 폴리우레탄, 폴리에스테르, 폴리카보네이트, 폴리우레아, 폴리아미드, 폴리에틸렌아민, 폴리아릴렌 술피드, 폴리실록산, 폴리이미드, 폴리아세테이트, 및 본원에 기술되거나, 또는 다르게는 당업계에 공지된 임의의 다른 적합한 화합물을 포함한다. 일부 경우에서, 중합체는 이 종중합체이다. 일부 경우에서, 중합체는 동종중합체이다. 일부 경우에서, 중합체는 작용성 모이어티를 포함하거나, 또는 접합된 것이다.

일부 경우에서, 장치의 분해된 좌위는 표면 에너지를 증가 및/또는 감소시키는 하나 이상의 모이어티로 관능화된다. 일부 경우에서, 모이어티는 화학적으로 불활성이다. 일부 경우에서, 모이어티는 원하는 화학 반응, 반응, 예를 들어, 올리고핵산 합성 반응 중 하나 이상의 프로세스를 지원하도록 구성된다. 표면의 표면 에너지, 또는 소수성은 표면 상에 부착되는 뉴클레오티드의 친화도를 결정짓는 인자이다. 일부 경우에서, 장치 관능화 방법은 (a) 이산화규소를 포함하는 표면을 갖는 장치를 제공하는 단계; 및 (b) 본원에 기술되어 있거나, 또는 다르게는 당업계에 공지된 적합한 실란화제, 예를 들어, 유기작용성 알콕시실란 분자를 사용하여 표면을 실란화시키는 단계를 포함할 수 있다.

일부 경우에서, 유기작용성 알콕시실란 분자로는 디메틸클로로-옥토데실-실란, 메틸디클로로-옥토데실-실란, 트리클로로-옥토데실-실란, 트리메틸-옥토데실-실란, 트리에틸-옥토데실-실란, 또는 이들의 임의의 조합을 포함한다. 일부 경우에서, 장치 표면은 폴리에틸렌/폴리프로필렌(감마 방사선 조사 또는 크롬산 산화, 및 하이드록시알킬 표면으로의 환원에 의해 관능화됨), 고도로 가교결합된 폴리스티렌-디비닐벤젠(클로로메틸화에 의해 유도체화되고, 벤질아민 작용성 표면으로 아민화됨), 나일론(말단 아미노핵실 기는 직접적인 반응성을 나타냄)으로 관능화, 또는 환원된 폴리테트라플루오로에틸렌으로 에칭된 것을 포함한다. 다른 방법 및 관능화제는 미국 특허 번호 5474796(상기 특허는 그 전문이 본원에서 참조로 포함된다)에 기술되어 있다.

일부 경우에서, 장치 표면은 전형적으로 장치 표면 상에 존재하는 반응성 친수성 모이어티를 통해 실란을 장치 표면에 커플링시키는 데에 효과적인 반응 조건하에서 실란의 혼합물을 함유하는 유도체화 조성물과의 접촉에 의해 관능화된다. 실란화는 일반적으로 유기작용성 알콕시실란 분자와의 자가 조립을 통해 표면을 커버한다.

다양한 실록산 관능화 시약은 현재 당업계에 공지되어 있는 바와 같이, 예컨대, 표면 에너지를 낮추거나, 또는

증가시키기 위해 추가로 사용될 수 있다. 유기작용성 알콕시실란은 그의 유기 작용기에 따라 분류될 수 있다.

[0101] 본원에서는 뉴클레오시드에 커플링할 수 있는 작용제의 패턴화를 함유할 수 있는 장치를 제공한다. 일부 경우에서, 장치는 활성제로 코팅될 수 있다. 일부 경우에서, 장치는 부동화제로 코팅될 수 있다. 본원에 기술된 코팅 물질에 포함되는 예시적인 활성제로는 제한 없이, N-(3-트리에톡시실릴프로필)-4-하이드록시부티르아미드(HAPS), 11-아세톡시운데실트리에톡시실란, n-데실트리에톡시실란, (3-아미노프로필)트리메톡시실란, (3-아미노프로필)트리에톡시실란, 3-글리시독시프로필트리메톡시실란(GOPS), 3-아이오도-프로필트리메톡시실란, 부틸-알데하이드-트리메톡시실란, 이량체 2차 아미노알킬 실록산, (3-아미노프로필)-디에톡시-메틸실란, (3-아미노프로필)-디메틸-에톡시실란, 및 (3-아미노프로필)-트리메톡시실란, (3-글리시독시프로필)-디메틸-에톡시실란, 글리시독시-트리메톡시실란, (3-메르캅토프로필)-트리메톡시실란, 3-4 에폭시시클로헥실-에틸트리메톡시실란, 및 (3-메르캅토프로필)-메틸-디메톡시실란, 알릴 트리클로로클로로실란, 7-옥트-1-에닐 트리클로로클로로실란, 또는 비스(3-트리메톡시실릴프로필) 아민을 포함한다.

[0102] 본원에 기술된 코팅 물질에 포함되는 예시적인 부동화제로는 제한 없이, 퍼플루오로옥틸트리클로로실란; 트리데카플루오로-1,1,2,2-테트라하이드로옥틸)트리클로로실란; 1H,1H,2H,2H-플루오로옥틸트리에톡시실란(FOS); 트리클로로(1H,1H,2H,2H-퍼플루오로옥틸)실란; tert-부틸-[5-플루오로-4-(4,4,5,5-테트라메틸-1,3,2-디옥사보로란-2-일)인돌-1-일]-디메틸-실란; CYTOP™; 플루오리너트(Fluorinert)™; 퍼플루오록틸트리클로로실란(PFOTCS); 퍼플루오로옥틸디메틸클로로실란(PFODCS); 퍼플루오로데실트리에톡시실란(PFDTES); 펜타플루오로페닐-디메틸프로필클로로-실란(PFPTE); 퍼플루오로옥틸트리에톡시실란; 퍼플루오로옥틸트리메톡시실란; 옥틸클로로실란; 디메틸클로로-옥토데실-실란; 메틸디클로로-옥토데실-실란; 트리클로로-옥토데실-실란; 트리메틸-옥토데실-실란; 트리에틸-옥토데실-실란; 또는 옥타데실트리클로로실란을 포함한다.

[0103] 일부 경우에서, 관능화제로는 탄화수소 실란, 예컨대, 옥타데실트리클로로실란을 포함한다. 일부 경우에서, 관능화제로는 11-아세톡시운데실트리에톡시실란, n-데실트리에톡시실란, (3-아미노프로필)트리메톡시실란, (3-아미노프로필)트리에톡시실란, 글리시딜옥시프로필/트리메톡시실란 및 N-(3-트리에톡시실릴프로필)-4-하이드록시부티르아미드를 포함한다.

#### 올리고뉴클레오티드 합성

[0105] 본 개시내용의 올리고핵산 합성 방법은 포스포라미다이트 화학법을 포함하는 프로세스를 포함할 수 있다. 일부 경우에서, 올리고핵산 합성은 염기를 포스포라미다이트와 커플링하는 것을 포함한다. 올리고핵산 합성은 커플링 조건하에서 포스포라미다이트를 증착시켜 염기를 커플링하는 것을 포함할 수 있고, 여기서, 동일한 염기가 임의적으로 1회 초파로 포스포라미다이트로 증착되고, 즉, 이중으로 커플링된다. 올리고핵산 합성은 비반응 부위의 캡핑을 포함할 수 있다. 일부 경우에서, 캡핑은 임의적이다. 올리고핵산 합성은 또한 산화 또는 산화 단계 또는 산화 단계들을 포함할 수 있다. 올리고핵산 합성은 탈차단, 탈트리틸화, 및 황화를 포함할 수 있다. 일부 경우에서, 올리고핵산 합성은 산화 또는 황화 중 어느 하나를 포함한다. 일부 경우에서, 올리고핵산 합성 반응 동안 단계 또는 각 단계 동안 예를 들어, 테트라졸 또는 아세토니트릴을 사용하여 장치를 세척한다. 포스포라미다이트 합성 방법에서 어느 한 단계 동안의 시간은 약 2 min, 1 min, 50 sec, 40 sec, 30 sec, 20 sec 및 10 sec 미만일 수 있다.

[0106] 포스포라미다이트를 사용하는 올리고핵산 합성 방법은 포스파이트 트리에스테르 결합 형성을 위해 성장하는 올리고핵산 쇄에의 포스포라미다이트 빌딩 블록(예컨대, 뉴클레오시드 포스포라미다이트)의 잇따른 부가를 포함할 수 있다. 포스포라미다이트 올리고핵산 합성은 3'에서 5' 방향으로 진행된다. 포스포라미다이트 올리고핵산 합성을 통해 합성 사이클에 따라 하나의 뉴클레오티드를 성장하는 핵산 쇄에 조절된 방식으로 부가할 수 있다. 일부 경우에서, 각 합성 사이클은 커플링 단계를 포함한다. 포스포라미다이트 커플링은 활성화된 뉴클레오시드 포스포라미다이트와, 예를 들어, 링커를 통해 기판에 결합된 뉴클레오시드 사이의 포스파이트 트리에스테르 결합의 형성을 포함한다. 일부 경우에서, 뉴클레오시드 포스포라미다이트를 활성화된 장치에 제공한다. 일부 경우에서, 뉴클레오시드 포스포라미다이트를 활성화제와 함께 장치에 제공한다. 일부 경우에서, 뉴클레오시드 포스포라미다이트를 기판에 결합된 뉴클레오시드 대비 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100배 이상의 과량으로 장치에 제공한다. 일부 경우에서, 뉴클레오시드 포스포라미다이트를 부가하는 것은 무수 환경에서, 예를 들어, 무수 아세토니트릴 중에서 수행된다. 뉴클레오시드 포스포라미다이트 부가 후, 장치를 임의적으로 세척한다. 일부 경우에서, 커플링 단계는 1회 이상의 추가 회차로 반복되고, 임의적으로, 세척 단계는 기판에 뉴클레오시드 포스포라미다이트를 부가하는 단계들 사이에 이루어진다. 일부 경우에서, 본원에서 사용되는 올리고핵산 합성 방법은 1, 2, 3회 이상의 순차

적인 커플링 단계를 포함한다. 다수 경우에서, 커플링 이전에, 장치에 결합된 뉴클레오시드는 보호기 제거에 의해 탈보호화되고, 여기서, 보호기는 중합화를 막는 작용을 한다. 통상의 보호기는 4,4'-디메톡시트리틸(DMT)이다.

[0107] 커플링 후, 포스포라미다이트 올리고핵산 합성 방법은 임의적으로 캡핑 단계를 포함한다. 캡핑 단계에서, 성장하는 올리고핵산을 캡핑제로 처리한다. 캡핑 단계는, 내부 염기가 결실된 올리고핵산이 형성되지 못하게 막으면서, 커플링 이후의 비반응 기관 결합 5'-OH 기의 추가의 쇄 신장을 차단하는 데 유용하다. 추가로, 1H-테트라졸로 활성화된 포스포라미다이트는 작게 구아노신의 06 위치와 반응할 수 있다. 이론으로 제한하지 않으면서, I<sub>2</sub>/물에 의한 산화시, 상기 부산물은 가능하게는 06-N7 이동을 통해 탈퓨린화될 수 있다. 퓨린 결여(apurinic) 부위는 올리고뉴클레오티드의 최종 탈보호화의 과정에서 결단된 상태로 종결되어, 전장의 생성물의 수율을 감소시킬 수 있다. 06 변형은 I<sub>2</sub>/물을 사용한 산화 이전에 캡핑 시약을 사용한 처리에 의해 제거될 수 있다. 일부 경우에서, 올리고핵산 합성 동안 캡핑 단계를 포함하는 것이 캡핑을 포함하지 않는 합성과 비교하여 오류율을 감소시킨다. 한 예로서, 캡핑 단계는 아세트산 무수물과 1-메틸이미다졸의 혼합물로 기관 결합 올리고핵산을 처리하는 것을 포함한다. 캡핑 단계 후, 장치를 임의적으로 세척한다.

[0108] 일부 경우에서, 뉴클레오시드 포스포라미다이트의 부가 후, 및 임의적으로 캡핑 및 1회 이상의 세척 단계 이후에, 장치에 결합된, 성장하는 핵산을 산화시킨다. 산화 단계는 포스파이트 트리에스테르가, 자연적으로 발생된 포스페이트 디에스테르 뉴클레오시드간 결합의 보호된 전구체인 4 배위 포스페이트 트리에스테르로 산화되는 것을 포함한다. 일부 경우에서, 성장하는 올리고핵산의 산화는 임의적으로 약염기(예컨대, 피리딘, 루티딘 또는 콜리딘)의 존재하에서 아이오딘 및 물로 처리함으로써 달성된다. 산화는 무수 조건하에서, 예컨대, tert-부틸 하이드로페온시드 또는 (1S)-(+)-(10-캄포술포닐)-옥사지리딘(CSO)을 사용하여 수행될 수 있다. 일부 방법에서, 캡핑 단계는 산화 이후에 수행된다. 지속될 수 있는, 산화로부터의 잔류성 물이 후속 커플링을 억제시킬 수 있는 바, 제2 캡핑 단계는 장치를 건조시킬 수 있다. 산화 후, 장치 및 성장하는 올리고핵산을 임의적으로 세척한다. 일부 경우에서, 올리고뉴클레오티드 포스포로티오에이트를 수득하기 위해 산화 단계를 황화 단계로 치환하고, 여기서, 임의의 캡핑 단계는 황화 이후에 수행될 수 있다. 3-(디메틸아미노메틸리텐)아미노)-3H-1,2,4-디티아졸-3-티온, DDTT, 보카주(Beaucage) 시약으로도 공지되어 있는 3H-1,2-벤조디티올-3-온 1,1-디옥시드, 및 N,N,N'-테트라에틸티우람 디솔피드(TETD)를 포함하나, 이에 제한되지 않는 다수의 시약이 황을 효율적으로 전달할 수 있다.

[0109] 뉴클레오시드를 도입하는 후속 사이클이 커플링을 통해 이루어지도록 하기 위해, 장치에 결합된, 성장하는 올리고핵산의 보호된 5' 단부를 제거하면, 이로써, 1차 하이드록실 기는 다음 뉴클레오시드 포스포라미다이트와 반응한다. 일부 경우에서, 보호기는 DMT이고, 탈차단은 디클로로메탄 중 트리클로로아세트산으로 이루어진다. 연장된 시간 동안, 또는 권장된 산 용액보다 더욱 강한 산을 이용하여 탈트리틸화를 수행하면, 고체 지지체에 결합된 올리고뉴클레오티드의 탈퓨린화는 증가될 수 있고, 이로써 원하는 전장의 생성물의 수율은 감소된다. 본원에 기술된 본 개시내용의 방법 및 조성물은 원하지 않는 탈퓨린화 반응을 제한하는 조절된 탈차단 조건을 제공한다. 일부 경우에서, 탈차단 이후에 장치에 결합된 올리고핵산을 세척한다. 일부 경우에서, 탈차단 이후의 효율적인 세척이 합성된 올리고핵산이 낮은 오류율을 갖는 데 기여한다.

[0110] 올리고핵산 합성 방법은 전형적으로 하기 단계 순서를 반복하는 것을 포함한다: 보호된 단량체를 능동적으로 관능화된 표면(예컨대, 좌위)에 적용시켜 활성화된 표면과, 링커 또는 사전에 탈보호화된 단량체와 연결시키는 단계; 적용된 단량체를 탈보호화하여 후속하여 적용되는 보호된 단량체와 반응을 하도록 하는 단계; 및 연결을 위해 또 다른 보호된 단량체를 적용시키는 단계. 하나 이상의 중가 단계는 산화 또는 황화를 포함한다. 일부 경우에서, 1회 이상의 세척 단계가 상기 단계 중 한 단계 또는 그들 모두 이전에 또는 그 이후에 진행된다.

[0111] 포스포라미다이트 기반의 올리고핵산 합성 방법은 일련의 화학적 단계를 포함한다. 일부 경우에서, 합성 방법 중 하나 이상의 단계는 시약 사이클링을 포함하고, 여기서, 방법 중 하나 이상의 단계는 단계에 유용한 시약을 장치에 적용시키는 것을 포함한다. 예를 들어, 시약은 일련의 액체 중착 및 진공 건조 단계에 의해 사이클링된다. 3차원 특징부를 포함하는 기관, 예컨대, 웰, 마이크로웰, 채널 등의 경우, 시약이 임의적으로 웰 및/또는 채널을 통해 장치의 하나 이상의 영역을 통해 통과하도록 한다.

[0112] 본원에 기술된 방법 및 시스템은 올리고뉴클레오티드 합성을 위한 올리고뉴클레오티드 합성 장치에 관한 것이다. 합성은 동시에 진행될 수 있다. 예를 들어 적어도 또는 적어도 약 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 1,000, 10,000, 50,000, 75,000, 100,000개

이상의 올리고뉴클레오티드가 동시에 합성될 수 있다. 동시에 합성될 수 있는 올리고핵산의 총 개수는 2-100,000, 3-50,000, 4-10,000, 5-1,000, 6-900, 7-850, 8-800, 9-750, 10-700, 11-650, 12-600, 13-550, 14-500, 15-450, 16-400, 17-350, 18-300, 19-250, 20-200, 21-150, 22-100, 23-50, 24-45, 25-40, 30-35개일 수 있다. 동시에 합성된 올리고뉴클레오티드의 총 개수는 상기 값 중 임의의 것에 의해 제한된 임의 범위, 예를 들어, 25-100 내에 포함될 수 있다는 것을 당업자는 이해한다. 동시에 합성된 올리고뉴클레오티드의 총 개수는 상기 범위의 종점으로서 작용한 상기 값 중 임의의 것에 의해 정의된 임의 범위 내에 포함될 수 있다. 장치 내에서 합성된 올리고뉴클레오티드의 총 몰 질량, 또는 각 올리고뉴클레오티드의 몰 질량은 적어도 또는 적어도 약 10, 20, 30, 40, 50, 100, 250, 500, 750, 1,000, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 25,000, 50,000, 75,000, 100,000 퍼코몰 이상일 수 있다. 각 올리고뉴클레오티드의 길이 또는 장치 내의 올리고뉴클레오티드의 평균 길이는 적어도 또는 적어도 약 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 150, 200, 300, 400, 500개의 뉴클레오티드 길이 그 이상일 수 있다. 각 올리고뉴클레오티드의 길이 또는 장치 내의 올리고뉴클레오티드의 평균 길이는 최대 또는 최대 약 500, 400, 300, 200, 150, 100, 50, 45, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10개의 뉴클레오티드 길이 그 이하일 수 있다. 각 올리고뉴클레오티드의 길이 또는 장치 내의 올리고뉴클레오티드의 평균 길이는 10-500, 9-400, 11-300, 12-200, 13-150, 14-100, 15-50, 16-45, 17-40, 18-35, 19-25인 것을 포함할 수 있다. 각 올리고뉴클레오티드의 길이 또는 장치 내의 올리고뉴클레오티드의 평균 길이는 상기 값 중 임의의 것에 의해 제한된 임의 범위, 예를 들어, 100-300 내에 포함될 수 있다는 것을 당업자는 이해한다. 각 올리고뉴클레오티드의 길이 또는 장치 내의 올리고뉴클레오티드의 평균 길이는 상기 범위의 종점으로서 작용한 상기 값 중 임의의 것에 의해 정의된 임의 범위 내에 포함될 수 있다.

[0113]

본원에서 제공된, 표면 상에서 올리고핵산을 합성하는 방법을 통해 빠른 속도로 합성할 수 있다. 한 예로서, 시간당 적어도 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, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 125, 150, 175, 200개의 뉴클레오티드 그 이상이 합성된다. 뉴클레오티드로는 아데닌, 구아닌, 티민, 시토신, 우리딘 빌딩 블록, 또는 그의 유사체/변형된 버전을 포함한다. 일부 경우에서, 올리고핵산 라이브러리는 기판 상에서 동시에 합성된다. 예를 들어, 약 또는 적어도 약 100; 1,000; 10,000; 30,000; 75,000; 100,000; 1,000,000; 2,000,000; 3,000,000; 4,000,000; 또는 5,000,000개의 분해된 좌위를 포함하는 장치가 적어도 동일한 개수의 상이한 올리고핵산의 합성을 지지하며, 여기서, 상이한 서열을 코딩하는 올리고핵산은 분해된 좌위 상에서 합성된다. 일부 경우에서, 올리고핵산 라이브러리는 약 3개월, 2개월, 1개월, 3주, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2일, 24시간 이하인 시간 이내에 본원에 기술된 낮은 오류율로 장치 상에서 합성된다. 일부 경우에서, 본원에 기술된 기판 및 방법을 사용하여 낮은 오류율로 합성된 올리고핵산 라이브러리로부터 조립된 더욱 큰 핵산은 약 3개월, 2개월, 1개월, 3주, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2일, 24시간 이하인 시간 이내에 제조된다.

[0114]

일부 경우에서, 본원에 기술된 방법은 복수의 코돈 부위에서 상이한 변이체 올리고핵산을 포함하는 올리고핵산 라이브러리를 생성하는 것을 제공한다. 일부 경우에서, 올리고핵산은 변이체 코돈 부위 중 1개 부위, 2개 부위, 3개 부위, 4개 부위, 5개 부위, 6개 부위, 7개 부위, 8개 부위, 9개 부위, 10개 부위, 11개 부위, 12개 부위, 13개 부위, 14개 부위, 15개 부위, 16개 부위, 17개 부위 18개 부위, 19개 부위, 20개 부위, 30개 부위, 40개 부위, 50개 부위, 또는 그 초과를 가질 수 있다.

[0115]

일부 경우에서, 변이체 코돈 부위 중 하나 이상의 부위가 인접해 있을 수 있다. 변이체 코돈 부위 중 하나 이상의 부위는 인접해 있지 않고, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 코돈에 의해 이격화되어 있을 수 있다.

[0116]

일부 경우에서, 올리고핵산은 모든 변이체 코돈 부위가 서로 인접해 있고, 이로써, 변이체 스트레치 코돈 부위를 형성하는 것인 변이체 코돈 부위의 다중 부위를 포함할 수 있다. 일부 경우에서, 올리고핵산은 변이체 코돈 부위 중 어느 것도 서로 인접해 있지 않은 것인 변이체 코돈 부위의 다중 부위를 포함할 수 있다. 일부 경우에서, 올리고핵산은 일부 변이체 코돈 부위가 서로 인접해 있고, 이로써, 변이체 스트레치 코돈 부위를 형성하고, 변이체 코돈 부위 중 일부는 서로 인접해 있지 않은 것인 변이체 코돈 부위의 다중 부위를 포함할 수 있다.

[0117]

도면을 참조하면, 도 12는 더 짧은 올리고핵산으로부터 핵산(예컨대, 유전자)을 합성하기 위한 예시적인 프로세스 작업 흐름을 도시한 것이다. 작업 흐름은 일반적으로 (1) 단일 가닥 올리고핵산 라이브러리의 드 노보 합성 단계, (2) 올리고핵산을 연결하여 더 큰 단편을 형성하는 단계, (3) 오류 보정 단계, (4) 품질 관리 단계, 및 (5) 운반 단계로 나뉜다. 드 노보 합성 이전에, 의도하는 핵산 서열 또는 핵산 서열 군을 미리 선택한다. 예를 들어, 생성을 위해 유전자 군을 미리 선택한다.

- [0118] 일단 생성을 위해 큰 올리고핵산을 선택하였으면, 드 노보 합성을 위해 소정의 올리고핵산 라이브러리를 디자인 한다. 고밀도 올리고핵산 어레이 생성을 위한 각종의 적합한 방법이 공지되어 있다. 작업 흐름 예에서, 장치 표면 층(1201)을 제공한다. 상기 예에서, 올리고핵산 합성 프로세스를 개선시키기 위해 표면의 화학적 성질을 변경시킨다. 액체에는 반발력을 가하기 위해 표면 에너지가 낮은 면적을 생성하고, 동시에, 액체에 인력을 가하기 위해 표면 에너지가 높은 면적을 생성한다. 표면 그 자체는 평면 형태일 수 있거나, 예컨대, 표면적을 증가시키는 돌출부 또는 마이크로웰과 같은 형상 변화를 함유할 수 있다. 작업 흐름 예에서, 선택된 고에너지 표면 분자는 국제 특허 출원 공개 WO/2015/021080(이 특허는 그 전문이 본원에서 참조로 포함된다)에 개시된 바와 같이, DNA 화학적 성질을 지원하는 이중적 기능을 한다.
- [0119] 올리고핵산 어레이의 반응계내 제조는 고체 지지체 상에서 생성되고, 이는 단일 뉴클레오티드 신장 프로세스를 이용하여 다중 올리고머를 동시에 신장시킨다. 물질 증착 장치, 예컨대, 올리고핵산 합성기는 다중 올리고핵산이 동시에 잔기를 한 번에 하나씩 신장시켜 소정의 핵산 서열을 갖는 올리고머를 생성하게 하는 단계적 방식으로 시약을 이형시킬 수 있도록 디자인된다(1202). 일부 경우에서, 이 단계에서 올리고핵산을 표면으로부터 절단한다. 절단은 예컨대, 암모니아 또는 메틸아민을 이용한 기체 절단을 포함한다.
- [0120] 생성된 올리고핵산 라이브러리를 반응 챔버에 배치한다. 본 예시적인 작업 흐름에서, 반응 챔버(이는 또한 "나노반응기"로도 지칭됨)는 PCR 시약을 함유하는 규소 코팅된 웰이고, 올리고핵산 라이브러리 상에 내려 놓는다(1203). 올리고핵산의 밀봉(1204) 이전 또는 이후에, 시약을 첨가하여 올리고핵산을 기판으로부터 이형시킨다. 예시적인 작업 흐름에서, 나노반응기의 밀봉(1205) 이후에, 올리고핵산을 이형시킨다. 일단 이형되고 나면, 단일 가닥 올리고핵산의 단편은 전체 긴 범위의 DNA 서열에 걸쳐 포괄하도록 하기 위해 하이브리드화한다. 각각의 합성된 올리고핵산은 풀에서 적어도 하나의 다른 올리고핵산과 작은 일부분이 중복되도록 디자인되었기 때문에 부분적 하이브리드화(1205)가 가능하다.
- [0121] 하이브리드화 후, PCA 반응을 개시한다. 폴리머라제 사이클 동안, 올리고핵산은 상보성 단편에 어닐링되고, 캡은 폴리머라제에 의해 충전된다. 각 사이클은 올리고핵산이 서로 어느 것을 찾게 되는지에 따라 무작위적으로 각종 단편의 길이를 증가시킨다. 단편 사이의 상보성을 통해서 완전한 큰 범위의 이중 가닥 DNA가 형성될 수 있다(1206).
- [0122] PCA 완료 후, 나노반응기를 장치로부터 분리시키고(1207), PCR용 프라이머가 있는 장치와의 상호작용을 위해 배치한다(1208). 밀봉 후, 나노반응기에 대해 PCR을 수행하고(1209), 더욱 큰 핵산을 증폭시킨다. PCR(1210) 후, 나노챔버를 개봉하고(1211), 오류 보정 시약을 첨가하고(1212), 챔버를 밀봉하고(1213), 오류 보정 반응을 통해 이중 가닥 PCR 증폭 생성물로부터 미스매칭된 염기쌍 및/또는 상보성이 부족한 가닥을 제거한다(1214). 나노반응기를 개봉하고, 분리시킨다(1215). 이어서, 오류 보정된 생성물에 추가 프로세싱 단계, 예컨대, PCR 및 분자바 코딩을 수행한 후, 운반(1223)을 위해 패키징한다(1222).
- [0123] 일부 경우에서, 품질 관리 조치를 취하다. 오류 보정 후, 품질 관리 단계는 오류 보정된 생성물의 증폭을 위해 서열분석 프라이머를 갖는 웨이퍼와의 상호작용(1216), 오류 보정된 증폭 생성물을 함유하는 챔버에의 웨이퍼 밀봉(1217), 및 추가 라운드의 증폭 수행(1218)을 포함한다. 나노반응기를 개봉하고(1219), 생성물을 풀링하고(1220), 서열분석한다(1221). 허용되는 품질 관리 측정 후, 패키징된 생성물(1222)을 운반(1223)을 승인한다.
- [0124] 일부 경우에서, 본원에 개시된 중복 프라이머를 이용하여 예컨대, 도 12에서의 것과 같은 작업 흐름에 의해 생성된 핵산에 대해 돌연변이유발을 수행한다. 일부 경우에서, 프라이머 라이브러리는 고체 지지체 상에서의 반응계내 제조에 의해 생성되고, 이는 단일 뉴클레오티드 신장 프로세스를 사용하여 동시에 다중 올리고머를 신장시킨다. 증착 장치, 예컨대, 올리고핵산 합성기는 다중 올리고핵산이 동시에 잔기를 한 번에 하나씩 신장시켜 소정의 핵산 서열을 갖는 올리고머를 생성하게 하는 단계적 방식으로 시약을 이형시킬 수 있도록 디자인된다(1202).
- [0125] 컴퓨터 시스템
- [0126] 본원에 기술된 시스템 중 임의의 시스템은 컴퓨터에 자동가능하게 연결될 수 있고, 근거리 방식으로 또는 원격적으로 컴퓨터를 통해 자동화될 수 있다. 다양한 경우에서, 본 개시내용의 방법 및 시스템은 컴퓨터 시스템상의 소프트웨어 프로그램 및 그의 사용을 추가로 포함할 수 있다. 따라서, 분배/진공/재충전 기능의 동시화, 예컨대, 물질 증착 장치 이동, 분배 작용 및 진공 작동의 조직화 및 동시화를 위한 전산화된 조절은 본 개시내용의 범위 내에 있다. 컴퓨터 시스템은 사용자에 의해 특정된 염기 서열과 물질 증착 장치의 위치를 접속시켜 올바른 시약을 기판의 특정된 영역에게 전달하도록 프로그래밍될 수 있다.

[0127]

도 13에 도시된 컴퓨터 시스템(1300)은 고정된 매체(1312)를 가진 서버(1309)에 임의적으로 연결될 수 있는 매체(1311) 및/또는 네트워크 포트(1305)로부터의 지시를 판독할 수 있는 논리 장치로서 이해될 수 있다. 예컨대, 도 13에 제시된 시스템은 CPU(1301), 디스크 드라이브(1303), 임의적 입력 장치, 예컨대, 키보드(1315) 및/또는 마우스(1316), 및 임의적 모니터(1307)를 포함할 수 있다. 데이터 통신은 제시된 통신 매체를 통해 근거리 또는 원거리에 위치하는 서버 쪽으로 달성될 수 있다. 통신 매체는 데이터를 전송하고/거나, 수신하는 임의의 수단을 포함할 수 있다. 예를 들어, 통신 매체는 네트워크 연결, 무선 연결 또는 인터넷 연결일 수 있다. 상기 연결은 월드 와이드 웹(World Wide Web)상에서의 통신을 제공할 수 있다. 본 개시내용과 관련된 데이터는 도 13에 도시된 바와 같이, 관계자(1322)에 의한 접수 및/또는 검토를 위해 상기와 같은 네트워크 또는 연결 상에서 전송될 수 있다는 것이 구상된다.

[0128]

도 14는 본 개시내용의 예시적 일례와 관련하여 사용될 수 있는 컴퓨터 시스템(1400)의 첫 번째 예시적 아키텍처를 도시한 블록 다이어그램이다. 도 14에 도시된 바와 같이, 예시적 컴퓨터 시스템은 지시를 프로세싱하기 위한 프로세서(1402)를 포함할 수 있다. 프로세서의 비제한적인 예로는 인텔(Intel) 제온(Xeon)<sup>TM</sup> 프로세서, AMD 옵테론(Opteron)<sup>TM</sup> 프로세서, 삼성(Samsung) 32-비트 RISC ARM 1176JZ(F)-S v10<sup>TM</sup> 프로세서, ARM 코텍스(Cortex)-A8 삼성 S5PC100<sup>TM</sup> 프로세서, ARM 코텍스-A8 애플 A4(ARM Cortex-A8 Apple A4)<sup>TM</sup> 프로세서, 마벨(Marvell) PXA 930<sup>TM</sup> 프로세서, 또는 기능적으로 등가인 프로세서를 포함한다. 다중의 실행 쓰레드가 병렬식 프로세싱을 위해 사용될 수 있다. 일부 경우에서, 다중 프로세서 또는 다중 코어를 가진 프로세서 또한 단일 컴퓨터 시스템, 클러스터, 또는 복수의 컴퓨터, 휴대 전화, 및/또는 개인 정보 단말기 장치를 포함하는, 네트워크상에 걸쳐 분산된 시스템에서 사용될 수 있다.

[0129]

도 14에 도시된 바와 같이, 고속 캐시(1404)를 프로세서(1402)에 연결하거나, 또는 그에 도입하여 프로세서(1402)에 의해 최근에 사용되었거나, 또는 빈번하게 사용되는 지시 또는 데이터를 위한 고속 메모리를 제공할 수 있다. 프로세서(1402)는 프로세서 버스(1408)에 의해 노쓰 브리지(1406)에 연결된다. 노쓰 브리지(1406)는 메모리 버스(1412)에 의해 랜덤 액세스 메모리(RAM: random access memory)(1410)에 연결되고, 프로세서(1402)에 의해 RAM(1410)에의 접근을 관리한다. 노쓰 브리지(1406)는 또한 칩셋 버스(1416)에 의해 사우쓰 브리지(1414)에도 연결된다. 결국, 사우쓰 브리지(1414)는 주변 장치 버스(1418)에 연결된다. 주변 장치 버스는 예를 들면, PCI, PCI-X, PCI 익스프레스(Express) 또는 다른 주변 장치 버스일 수 있다. 노쓰 브리지 및 사우쓰 브리지는 종종 프로세서 칩셋으로 지정되고, 프로세서, RAM, 및 주변 장치 버스(1418) 상의 주변 구성 요소 사이의 데이터 전달을 관리한다. 일부 대안적 구성에서, 별개의 노쓰 브리지 칩을 사용하는 대신, 노쓰 브리지의 기능성을 프로세서 내로 도입할 수 있다. 일부 경우에서, 시스템(1400)은 주변 장치 버스(1418)에 부착된 액셀러레이터 카드(1422)를 포함할 수 있다. 액셀러레이터는 필드 프로그램 가능 게이트 어레이(FPGA: field programmable gate array), 또는 특정 프로세싱을 가속화하기 위한 다른 하드웨어를 포함할 수 있다. 예를 들어, 액셀러레이터는 적응 데이터 재구조화를 위해, 또는 확장된 세트 프로세싱에서 사용된 대수식 평가를 위해 사용될 수 있다.

[0130]

소프트웨어 및 데이터는 외부 저장 장치(1424) 내에 저장되고, 프로세서에 의한 사용을 위해 RAM(1410) 및/또는 캐시(1404) 내에 로딩될 수 있다. 시스템(1400)은 시스템 자원을 관리하기 위한 운영 시스템을 포함하고; 운영 시스템의 비제한적인 예로는 본 개시내용의 예시적인 경우에 따라 데이터 저장 및 최적화를 관리하기 위한 운영 시스템의 상부 상에서 실행되는 응용 소프트웨어 뿐만 아니라, 리눅스(Linux), 윈도우즈(Windows)<sup>TM</sup>, MACOS<sup>TM</sup>, 블랙베리(BlackBerry) OS<sup>TM</sup>, iOS<sup>TM</sup> 및 다른 기능적으로 등가인 운영 시스템도 포함한다. 본 예에서, 시스템(1400)은 외부 저장 장치, 예컨대, 네트워크 부착 저장 장치(NAS)에의 네트워크 접속을 제공하기 위해 주변 장치 버스에 연결된 네트워크 접속 카드들(NIC: network interface card)((1420) 및 (1421)), 및 분산된 병렬식 프로세싱을 위해 이용될 수 있는 다른 컴퓨터 시스템도 포함한다.

[0131]

도 15는 복수의 컴퓨터 시스템((1502a) 및 (1502b)), 복수의 휴대 전화 및 개인 정보 단말기(1502c) 및 네트워크 부착 저장 장치(NAS)((1504a) 및 (1504b))를 포함하는 네트워크(1500)를 보여주는 다이어그램이다. 예시적인 경우에서, 시스템((1502a), (1502b) 및 (1502c))은 데이터 저장을 관리할 수 있고, 네트워크 부착 저장 장치(NAS)((1504a) 및 (1504b))에 저장된 데이터에 대한 데이터 접근을 최적화할 수 있다. 수학적 모델은 데이터를 위해 이용할 수 있고, 컴퓨터 시스템((1502a) 및 (1502b)), 및 휴대 전화 및 개인 정보 단말기 시스템(1502c)에 걸쳐 분산된 병렬식 프로세싱을 이용하여 평가될 수 있다. 스템((1502a) 및 (1502b)), 및 휴대 전화 및 개인 정보 단말기 시스템(1502c)은 또한 네트워크 부착 저장 장치((1504a) 및 (1504b))에 저장된 데이터의 적응 데이터 재구조화를 위한 병렬식 프로세싱을 제공할 수 있다. 도 15는 단지 일례만을 보여주는 것이며, 매우 다양한 다른 컴퓨터 아키텍처 및 시스템이 본 개시내용의 다양한 경우와 함께 이용될 수 있다. 예를 들어, 블레이드 서버

를 이용하여 병렬식 프로세싱을 제공할 수 있다. 후면판을 통해 프로세서 블레이드를 연결하여 병렬식 프로세싱을 제공할 수 있다. 저장 장치 또한 후면판에 연결될 수 있거나, 또는 별개의 네트워크 접속을 통해 네트워크 부착 저장 장치(NAS)로서 연결될 수 있다. 일부 예시적인 경우에서, 프로세서는 별개의 메모리 공간을 유지할 수 있고, 다른 프로세서에 의한 병렬식 프로세싱을 위해 네트워크 접속, 후면판 또는 다른 연결기를 통해 데이터를 전송할 수 있다. 다른 경우에서, 프로세서 중 일부 또는 그들 모두는 공유 가상 주소 메모리 공간을 이용할 수 있다.

[0132] **도 16**은 예시적인 경우에 따라 공유 가상 주소 메모리 공간을 이용하는 멀티프로세서 컴퓨터 시스템(1600)의 블록 다이어그램이다. 상기 시스템은 공유된 메모리 서브시스템(1604)에 접근할 수 있는 복수의 프로세서(1602a-f)를 포함한다. 상기 시스템은 복수의 프로그램 가능 하드웨어 메모리 알고리즘 프로세서(MAP: memory algorithm processor)(1606a-f)를 메모리 서브시스템(1604)에 도입한다. 각각의 MAP(1606a-f)는 메모리(1608a-f) 및 하나 이상의 필드 프로그램 가능 케이트 어레이(FPGA)(1610a-f)를 포함할 수 있다. MAP는 설정가능한 기능 단위를 제공하고, 특정 알고리즘 또는 알고리즘의 일부는 각각의 프로세서로 세심하게 조정된 프로세싱을 위해 FPGA들(1610a-f)에게 제공될 수 있다. 예를 들어, 예시적 경우에서, MAP를 이용하여 데이터 모델에 관한 대수식을 평가하고, 적응 데이터 재구조화를 수행할 수 있다. 이러한 예에서, 각각의 MAP는 본 목적을 위해 모든 프로세서에 의해 전체적으로 접근가능하다. 한 구성에서, 각각의 MAP는 직접 메모리 접근(DMA: Direct Memory Access)을 이용하여 연관된 메모리(1608a-f)에 접근함으로써, 상기 메모리가 각각의 마이크로프로세서(1602a-f)로부터 독립적으로 및 비동기적으로 과제를 실행하게 할 수 있다. 이러한 구성에서, MAP는 알고리즘의 파이프라인 및 병렬식 실행을 위해 결과를 또 다른 MAP에게 직접 공급할 수 있다.

[0133] 상기 컴퓨터 아키텍처 및 시스템은 단지 예일 뿐이고, 병용 프로세서, 보조프로세서, FPGA 및 다른 프로그램 가능 논리 장치의 임의의 조합을 이용하는 시스템, 시스템온칩(SOC: system on chip), 응용 주문형 집적 회로(ASIC: application specific integrated circuit), 및 다른 프로세싱 및 논리 소자를 비롯한, 매우 다양한 다른 컴퓨터, 휴대 전화 및 개인 정보 단말기 아키텍처 및 시스템이 예시적인 경우와 관련하여 사용될 수 있다. 일부 경우에서, 컴퓨터 시스템 모두 또는 그의 일부가 소프트웨어 또는 하드웨어에서 실행될 수 있다. 랜덤 액세스 메모리, 하드 드라이브, 플래쉬 메모리, 테이프 드라이브, 디스크 어레이, 네트워크 부착 저장 장치(NAS), 및 다른 근거리 또는 분산된 데이터 저장 장치 및 시스템을 비롯한, 임의의 다양한 데이터 저장 매체가 예시적인 경우와 관련하여 사용될 수 있다.

[0134] 예시적인 경우에서, 상기 또는 다른 컴퓨터 아키텍처 및 시스템 중 임의의 컴퓨터 아키텍처 및 시스템 상에서 실행하는 소프트웨어 모듈을 이용하여 컴퓨터 시스템을 실행할 수 있다. 다른 경우에서, 펌웨어, 프로그램 가능 논리 장치, 예컨대, **도 13**에서 참조된 필드 프로그램 가능 케이트 어레이(FPGA), 시스템온칩(SOC), 응용 주문형 집적 회로(ASIC), 또는 다른 프로세싱 및 논리 소자에서 시스템의 기능을 부분적으로 또는 완전히 실행할 수 있다. 예를 들어, 하드웨어 액셀러레이터 카드, 예컨대, **도 13**에 도시된 액셀러레이터 카드(1322)의 사용을 통한 하드웨어 가속으로 세트 프로세서(Set Processor) 및 옵티마이저(Optimizer)를 실행할 수 있다.

[0135] 하기 실시예는 당업자에게 본원에 개시된 실시양태의 원리 및 실시를 더욱 명확하게 예시하기 위해 기술된 것이며, 임의의 청구된 실시양태의 범주를 제한하는 것으로 해석되지 않아야 한다. 달리 언급되지 않는 한, 모든 부(part) 및 백분율은 중량 기준으로 한다.

### 실시예

[0137] 하기 실시예는 본 개시내용의 다양한 실시양태를 예시하기 위한 목적으로 제공되는 것이며, 어느 방식으로든 본 개시내용을 제한하는 것으로 의도되지 않는다. 본원에 기술된 방법과 함께 본 실시예는 이제 바람직한 실시양태를 나타내고, 예시적인 것이며, 본 개시내용의 범주를 제한하는 것으로 의도되지 않는다. 특히 청구범위의 범주에 의해 정의되는 바와 같이, 본 개시내용의 정신 내에 포함되는 그 안에서의 변형 및 다른 사용이 당업자에게 일어날 것이다.

### 실시예 1: 장치 표면의 관능화

[0139] 울리고핵산 라이브러리의 부착 및 합성을 지원하기 위해 장치를 관능화시켰다. 먼저, 90%  $H_2SO_4$  및 10%  $H_2O_2$ 를 포함하는 피라나 용액을 이용하여 장치 표면을 20분 동안 습식 방식으로 세정하였다. 수개의 비커에서 DI 물을 이용하여 장치를 행구고, DI 물 거위목형 수도꼭지하에서 5 min 동안 유지시키고,  $N_2$ 를 이용하여 건조시켰다. 이어서, 장치를  $NH_4OH$ (1:100; 3 mL:300 mL) 중에 5 min 동안 침지시키고, 핸드건(handgun)을 이용하여 DI 물로 행구고, 3개의 연속된 비커에서 DI 물을 이용하여 매회 1 min 동안 침지시킨 후, 다시 핸드건을 이용하여 DI 물

로 행구었다. 이어서, 장치 표면을  $O_2$ 에 노출시켜 장치를 플라즈마 세정하였다. SAMCO PC-300 기계를 사용하여 하류 모드에서 1 min 동안 250 와트에서  $O_2$  플라즈마 에칭을 실시하였다.

[0140] 하기 파라미터를 이용하여 YES-1224P 기상 증착 오븐 시스템을 사용함으로써 N-(3-트리에톡시실릴프로필)-4-하이드록시부티르아미드를 포함하는 용액으로 세정 장치 표면을 능동적으로 관능화시켰다: 0.5 내지 1 토르, 60 min, 70°C, 135°C 증발기. 브루어 사이언스(Brewer Science) 200X 스판 코터를 이용하여 장치 표면을 레지스트 코팅하였다. SPR™ 3612 포토레지스트를 40 sec 동안 2,500 rpm으로 장치 상에 스판 코팅하였다. 장치를 브루어 햄 플레이트 상에서 90°C에서 30 min 동안 미리 베이킹하였다. 칼 수스(Karl Suss) MA6 마스크 얼라이너 기계를 사용하여 장치에 대해 포토리소그래피를 실시하였다. 장치를 2.2 sec 동안 노출시키고, MSF 26A 중에서 1 min 동안 발색시켰다. 잔류 발색제는 핸드건으로 행구고, 장치를 5 min 동안 물 중에 침지시켰다. 장치를 오븐에서 100°C에서 30 min 동안 베이킹한 후, 니콘(Nikon) L200을 이용하여 리소그래피 결함에 대하여 시각적으로 검사하였다. 디스컴(descum) 프로세스를 이용함으로써 SAMCO PC-300 기계를 사용하여 잔류 레지스트를 제거하여 1 min 동안 250 와트에서  $O_2$  플라즈마 에칭을 실시하였다.

[0141] 장치 표면을 10  $\mu$ l 경질 미네랄 오일과 함께 혼합된 퍼플루오로옥틸트리클로로실란 용액 100  $\mu$ l로 수동적으로 관능화시켰다. 장치를 챔버에 배치하고, 10 min 동안 펌핑한 후, 밸브를 펌프에 대해 닫고, 10 min 동안 정지시켰다. 챔버를 대기로 환기로 시켰다. 최대 전력으로 (Crest 시스템 상에서 9) 초음파처리하면서, 70°C에서 5 min 동안 500 mL NMP 중에 2회에 걸쳐 침지시킴으로써 장치를 레지스트 스트리핑시켰다. 이어서, 최대 전력으로 초음파처리하면서, 실온에서 5 min 동안 500 mL 이소프로판을 중에 장치를 침지시켰다. 장치를 200 푸르프 에탄올 300 mL 중에 디핑하고, N<sub>2</sub>를 이용하여 드라이 건조시켰다. 관능화된 표면을 활성화시켰고, 이는 올리고핵산 합성을 위한 지지체로서 작용하였다.

#### 실시예 2: 올리고뉴클레오티드 합성 장치 상에서의 50-mer 서열 합성

[0143] 2차원 올리고뉴클레오티드 합성 장치를 플로우셀로 조립하고, 이를 플로우셀 (Applied Biosystems (ABI394 DNA Synthesizer")에 연결시켰다. 2차원 올리고뉴클레오티드 합성 장치를 N-(3-트리에톡시실릴프로필)-4-하이드록시부티르아미드 (Gelest)로 균일하게 관능화시키고, 이를 이용하여 본원에 기술된 올리고뉴클레오티드 합성 방법을 사용함으로써 50 bp의 예시적인 올리고뉴클레오티드 ("50-mer 올리고뉴클레오티드")를 합성하였다.

[0144] 50-mer의 서열은 서열 번호: 20으로 기술하였다.  
 5'AGACAATCAACCATTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTTTTTT3' (서열 번호: 20), 여기서, #는 탈보호 동안 표면으로부터 올리고를 유리시킬 수 있는 절단가능한 링커인, 티미딘-숙시닐 핵사미드 CED 포스포라미다이트(ChemGenes로부터의 CLP-2244)를 나타낸다.

[0145] 하기 표 2의 프로토콜에 따른 표문 DNA 합성 화학법(커플링, 캡핑, 산화, 및 탈차단) 및 ABI 합성기를 이용하여 합성을 수행하였다.

#### 표 2

표 2		
일반 DNA 합성 프로세스 명칭	프로세스 단계	시간 (sec)
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	23
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
DNA 염기 침가 (포스포라미다이트+)	활성화제 메니폴드 세척	2
	플로우셀로 향하는 활성화제	6

[0146]

표 2		
일반 DNA 합성 프로세스 명칭	프로세스 단계	시간 (sec)
활성화제 유동)	플로우셀로 향하는 활성화제 + 포스포라미다이트	6
	플로우셀로 향하는 활성화제	0.5
	플로우셀로 향하는 활성화제 + 포스포라미다이트	5
	플로우셀로 향하는 활성화제	0.5
	플로우셀로 향하는 활성화제 + 포스포라미다이트	5
	플로우셀로 향하는 활성화제	0.5
	플로우셀로 향하는 활성화제 + 포스포라미다이트	5
	25 sec 동안 인큐베이션	25
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	15
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
DNA 염기 침가 (포스포라미다이트+ 활성화제 유동)	활성화제 매니폴드 세척	2
	플로우셀로 향하는 활성화제	5
	플로우셀로 향하는 활성화제 + 포스포라미다이트	18
	25 sec 동안 인큐베이션	25
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	15
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
캡핑(CapA+B, 1:1, 유동)	플로우셀로 향하는 CapA+B	15
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	15
	아세토니트릴 시스템 세척	4
산화(산화제 유동)	플로우셀로 향하는 산화제	18
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	15
	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	15
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	23
	N2 시스템 세척	4

[0147]

표 2		
일반 DNA 합성 프로세스 명칭	프로세스 단계	시간 (sec)
	아세토니트릴 시스템 세척	4
탈차단(탈차단제 유동)	플로우셀로 향하는 탈차단제	36
세척(아세토니트릴 세척 유동)	아세토니트릴 시스템 세척	4
	N2 시스템 세척	4
	아세토니트릴 시스템 세척	4
	플로우셀로 향하는 아세토니트릴	18
	N2 시스템 세척	4.13
	아세토니트릴 시스템 세척	4.13
	플로우셀로 향하는 아세토니트릴	15

[0148]

플로우셀을 통한 별크 시약 전달과 유사한 방식으로 포스포라미다이트/활성화제 조합을 전달하였다. 전 기간 동

안 시약으로 인해 환경은 "습윤" 상태로 유지되는 바, 건조 단계는 수행하지 않았다.

[0150] 더욱 빠르게 유동할 수 있도록 하기 위해 ABI 394 합성기로부터 유동 제한기를 제거하였다. (유동 제한기가 있는 경우, 모든 시약에 대하여 ~50  $\mu$ L/sec인 것과 비교하여) 유동 제한기가 없는 경우, 아미다이트(ACN 중 0.1 M), 활성화제, (ACN 중 0.25 M 벤조일티오텐트라졸("BTT"; GlenResearch로부터의 30-3070-xx)), 및 Ox(20% 피리딘, 10% 물, 및 70% THF 중 0.02 M I2)에 대한 유동 속도는 대략 ~100  $\mu$ L/sec였고, 아세토니트릴("ACN") 및 캡핑 시약(CapA 및 CapB의 1:1 믹스, 여기서, CapA는 THF/피리딘 중의 아세트산 무수물이고, CapB는 THF 중의 16% 1-메틸아미디졸이다)에 대한 유동 속도는 대략 ~200  $\mu$ L/sec였고, 데블록(Deblock)(톨루엔 중 3% 디클로로아세트산)에 대한 유동 속도는 대략 ~300  $\mu$ L/sec였다. 산화제를 완전히 압출시키는 데 소요되는 시간을 관찰하였고, 그에 맞춰 화학물질 유도 시간을 위한 타이밍을 조정하였고, 추가의 ACN 세척을 상이한 화학물질 사이에 도입하였다. 올리고뉴클레오티드 합성 후, 칩을 75 psi에서 밤새도록 기체 암모니아 중에서 탈보호화하였다. 물 5방울을 표면에 가하여 올리고핵산을 회수하였다. 이어서, 회수된 올리고핵산을 바이오애널라이저 소형 RNA 칩 상에서 분석하였다(데이터는 나타내지 않음).

#### [0151] 실시예 3: 올리고뉴클레오티드 합성 장치 상에서의 100-mer 서열 합성

[0152] 50-mer 서열 합성을 위해 실시예 2에 기술된 것과 동일한 프로세스를, 첫 번째 것은 N-(3-트리에톡시실릴프로필)-4-하이드록시부티르아미드로 균일하게 관능화된 것이고, 두 번째 것은 11-아세톡시운데실트리에톡시실란 및 n-데실트리에톡시실란의 5/95 믹스로 관능화된 것인, 2개의 상이한 규소 칩 상에서 100-mer 올리고뉴클레오티드 ("100-mer 올리고뉴클레오티드"; 5' CGGGATCCTTATCGTCATCGTCGATACAGATCCCGACCCATTGCTGTCACCAGTCATGCTAGCCATACCATGATGATGATGATGATGAGAACCCGCAT##TTTTTTTTT3', 여기서, #는 티미딘-숙시닐 핵사미드 CED 포스포라미다이트(ChemGenes로부터의 CLP-2244)를 나타낸다; 서열 번호: 21)를 합성하는 데 사용하였고, 표면으로부터 추출된 올리고핵산을 바이오애널라이저 기계 상에서 분석하였다(데이터는 나타내지 않음).

[0153] 하기 열 사이클링 프로그램을 사용하여 50  $\mu$ L PCR 믹스(25  $\mu$ L NEB Q5 마스터믹스, 2.5  $\mu$ L 10  $\mu$ M 정방향 프라이머, 2.5  $\mu$ L 10  $\mu$ M 역방향 프라이머, 표면으로부터 추출된 1  $\mu$ L 올리고핵산, 및 전체가 50  $\mu$ L가 될 때까지의 물) 중에서 정방향 프라이머(5'ATGCGGGGTTCTCATCATC3'; 서열 번호: 22) 및 역방향 프라이머(5'CGGGATCCTTATCGTCATCG3'; 서열 번호: 23)를 이용함으로써 상기 두 칩으로부터의 10개의 샘플 모두 추가로 PCR 증폭시켰다:

[0154] 98 C, 30 sec

[0155] 98 C, 10 sec; 63C, 10 sec; 72C, 10 sec; 사이클 12회 반복

[0156] 72C, 2 min.

[0157] PCR 생성물을 또한 바이오애널라이저 상에서 전개시켰고(데이터는 나타내지 않음), 이를 통해 100-mer 위치에 뾰족한 피크가 존재한다는 것이 입증되었다. 이어서, PCR 증폭 샘플을 클로닝하고, 생어(Sanger) 서열분석하였다. 하기 표 3에는 칩 1로부터의 스폿 1-5로부터 채취된 샘플, 및 칩 2로부터의 스폿 6-10으로부터 채취된 샘플에 대하여 수행된 생어 서열분석의 결과가 요약되어 있다.

## 표 3

스폿	오류율	사이클 효율
1	1/763 bp	99.87%
2	1/824 bp	99.88%
3	1/780 bp	99.87%
4	1/429 bp	99.77%
5	1/1525 bp	99.93%
6	1/1615 bp	99.94%
7	1/531 bp	99.81%
8	1/1769 bp	99.94%
9	1/854 bp	99.88%
10	1/1451 bp	99.93%

[0158]

이로써, 상이한 표면 화학법을 통해 2개의 칩 상에서 고품질이고, 균일한 합성된 올리고뉴클레오티드가 반복되었다. 전체적으로, 서열분석된 262개의 100-mer 중 233개에 상응하는 89%가 오류가 없는 완변한 서열이었다.

[0159]

마지막으로, 하기 표 4에는 스폿 1-10으로부터의 올리고뉴클레오티드 샘플로부터 수득된 서열에 대한 오류 특징이 요약되어 있다.

표 4

샘플 ID/ 스폿 번호	OSA_0 046/1	OSA_0 047/2	OSA_0 048/3	OSA_0 049/4	OSA_0 050/5	OSA_0 051/6	OSA_0 052/7	OSA_0 053/8	OSA_0 054/9	OSA_00 55/10
총 서열	32	32	32	32	32	32	32	32	32	32
시열분석 품질	25 / 28	27 / 27	26 / 30	21 / 23	25 / 26	29 / 30	27 / 31	29 / 31	28 / 29	25 / 28
올리고 품질	23 / 25	25 / 27	22 / 26	18 / 21	24 / 25	25 / 29	22 / 27	28 / 29	26 / 28	20 / 25
ROI 배치 계수	2500	2698	2561	2122	2499	2666	2625	2899	2798	2348
ROI 돌연변이	2	2	1	3	1	0	2	1	2	1
ROI 다중염기 결실	0	0	0	0	0	0	0	0	0	0
ROI 작은 삼입	1	0	0	0	0	0	0	0	0	0
ROI 단일 염기 결실	0	0	0	0	0	0	0	0	0	0
큰 결실 계수	0	0	1	0	0	1	1	0	0	0
돌연변이 G>A	2	2	1	2	1	0	2	1	2	1
돌연변이 T>C	0	0	0	1	0	0	0	0	0	0
ROI 오류 계수	3	2	2	3	1	1	3	1	2	1
ROI 오류율	Err: 834 총 ~1	Err: 1350 총 ~1	Err: 1282 총 ~1	Err: 708 총 ~1	Err: 2500 총 ~10	Err: 2667 총 ~17	Err: 876 총 ~1	Err: 2900 총 ~1	Err: 1400 총 ~1	Err: 2349 총 ~1
ROI - 프라이머 오류율	MP Err: 763 총 ~1	MP Err: 824 총 ~1	MP Err: 780 총 ~1	MP Err: 429 총 ~1	MP Err: 1525 총 ~1	MP Err: 1615 총 ~1	MP Err: 531 총 ~1	MP Err: 1769 총 ~1	MP Err: 854 총 ~1	MP Err: 1451 총 ~1

[0161]

## 실시예 4: 단일 부위, 단일 위치 돌연변이유발에 의한 올리고핵산 라이브러리 생성

[0162]

일련의 PCR 반응에서 사용하기 위한 올리고핵산 프라이머를 드 노보 합성하여 주형 핵산의 올리고핵산 변이체의 라이브러리를 생성하였다(도 2a-2d 참조). 도 2a에서 4가지 유형의 프라이머가 생성되었다: 외부 5' 프라이머 (215), 외부 3' 프라이머(230), 내부 5' 프라이머(225), 및 내부 3' 프라이머(220). 일반적으로 표 2에 개략적으로 서술되어 있는 바와 같은 올리고핵산 합성 방법을 사용하여 내부 5' 프라이머/제1 올리고핵산(220) 및 내부 3' 프라이머/제2 올리고핵산(225)을 생성하였다. 내부 5' 프라이머/제1 올리고핵산(220)은 소정의 서열로 이루어진 최대 19의 프라이머로 구성된 세트를 나타내며, 여기서, 세트 내 각각의 프라이머는 서열 중의 단일 부위 중 단일 코돈에서 또 다른 것과 상이하다.

[0164]

각각의 클러스터가 121개의 개별적으로 주소지정가능한 좌위를 갖는, 적어도 2개의 클러스터가 있는 장치 상에서 올리고핵산 합성을 수행하였다.

[0165]

내부 5' 프라이머(225) 및 내부 3' 프라이머(220)를 별개의 클러스터에서 합성하였다. 내부 5' 프라이머(225)를 121회 복제하여 단일 클러스터 내의 121개의 좌위 상에서 연장시켰다. 내부 3' 프라이머(220)의 경우, 변이체 서열로 이루어지 19개의 프라이머 각각의 것을 각각 6개의 상이한 좌위에서 연장시켜 114개의 상이한 좌위 상에서 114개의 올리고핵산을 연장시켰다.

[0166]

합성된 올리고핵산을 장치 표면으로부터 절단하고, 플라스틱 바이알로 전달하였다. 도 2b에 도시된 바와 같이, 긴 핵산 서열(235), (240)의 단편을 이용하여 제1 PCR 반응을 수행하였다. 도 2c-2d에 도시된 바와 같이, 프라이머 조합 및 주형으로서 제1 PCR 반응의 생성물을 이용하여 제2 PCR 반응을 수행하였다. 바이오애널라이저 상에서 제2 PCR 생성물 분석을 수행하였고, 이는 도 17의 트레이스에 제시된 바와 같았다.

[0167] **실시예 5: 96개의 상이한 단일 위치 변이체의 세트를 포함하는 올리고핵산 라이브러리 생성**

[0168] 드 노보 올리고핵산 합성을 사용하여 일반적으로 **도 2a**에 제시되어 있고, 실시예 2에서 다루어진 것과 같은 4개의 프라이머 세트를 생성하였다. 내부 5' 프라이머(220)를 위해, 각 프라이머 세트가 주형 올리고핵산의 단일 부위 내에 배치된 상이한 단일 코돈을 표적화하는 것인 96개의 상이한 프라이머 세트를 생성하였다. 각 프라이머 세트를 위해, 각각의 변이체가 단일 부위의 상이한 아미노산을 코딩하는 코돈을 포함하는 것인 19개의 상이한 변이체를 생성하였다. 일반적으로 **도 2a-2d**에 제시되어 있고, 실시예 2에서 기술된 바와 같이, 생성된 프라이머를 사용하여 2 라운드에 걸쳐 PCR을 수행하였다. 96개의 증폭 생성물 세트를 전기영동도(**도 18**)로 시각화하였고, 이를 사용하여 100% 증폭 성공률을 산출하였다.

[0169] **실시예 6: 500개의 상이한 단일 위치 변이체의 세트를 포함하는 올리고핵산 라이브러리 생성**

[0170] 드 노보 올리고핵산 합성을 사용하여 일반적으로 **도 2a**에 제시되어 있고, 실시예 2에서 다루어진 것과 같은 4개의 프라이머 세트를 생성하였다. 내부 5' 프라이머(220)를 위해, 각 프라이머 세트가 주형 올리고핵산의 단일 부위 내에 배치된 상이한 단일 코돈을 표적화하는 것인 500개의 상이한 프라이머 세트를 생성하였다. 각 프라이머 세트를 위해, 각각의 변이체가 단일 부위의 상이한 아미노산을 코딩하는 코돈을 포함하는 것인 19개의 상이한 변이체를 생성하였다. 일반적으로 **도 2a**에 제시되어 있고, 실시예 2에서 기술된 바와 같이, 생성된 프라이머를 사용하여 2 라운드에 걸쳐 PCR을 수행하였다. 전기영동도는 상이한 단일 부위에 19종의 변이체를 포함하는 것인 핵산의 집단을 포함하는 500개의 PCR 생성물 세트 각각의 것을 나타낸다(데이터는 나타내지 않음). 라이브러리의 종합 서열분석 분석 결과, 미리 선택된 코돈 돌연변이에 걸쳐 99% 초과의 성공률을 보였다(서열 트레이스 및 분석 데이터는 나타내지 않음).

[0171] **실시예 7: 1번 위치에 대한 단일 부위 돌연변이유발 프라이머**

[0172] 코돈 변이 디자인의 한 예가 황색 형광 단백질에 대해 하기 표 3에 제시되어 있다. 이러한 경우, 서열 50-mer로부터의 단일 코돈을 19회에 걸쳐 다르게 하였다. 변이체 핵산 서열은 굵은체로 표시되어 있다. 야생형 프라이머 서열은 ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGTGGTGCCCAT(서열 번호: 1)이다. 본 경우에서, 야생형 코돈은 서열 번호: 1에서 밑줄로 표시되어 있는 발린을 코딩한다. 그러므로, 하기의 19종의 변이체는 발린을 코딩하는 코돈은 배제한다. 대안적 예에서, 모든 트리플릿이 고려되어야 한다면, 이때는 야생형 코돈에 대한 대안적 서열을 포함하는 60종의 변이체 모두가 생성될 것이다.

[0173]

[표 3]

서열 번호	변이체 번호	변이체 코돈
2	atgTTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	F
3	atgTTAAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	L
4	atgATTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	I
5	atgTCTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	S
6	atgCCTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	P
7	atgACTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	T
8	atgGCTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	A
9	atgTATAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	Y
10	atgCATAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	H
11	atgCAAAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	Q
12	atgAATAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	N
13	atgAAAAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	K
14	atgGATAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	D
15	atgGAAAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	E
16	atgTGTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	C
17	atgTGGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	W
18	atgCGTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	R
19	atgGGTAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCAT	G

[0174]

## 실시예 8: 단일 부위, 이중 위치 올리고핵산 변이체

[0175]

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행하였다. 단일 부위에서, 각 위치가 아미노산을 코딩하는 코돈인 2개의 연속된 코돈 위치에 대해 합성된 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스터를 생성하였다. 이러한 방식에서, 2개의 위치에 대해 위치당 19종의 변이체가 생성되었고, 여기서, 각각의 올리고핵산에 대해 3개의 복제물을 포함하며, 이로써, 114개의 올리고핵산이 합성되었다.

[0176]

## 실시예 9: 다중 부위, 이중 위치 올리고핵산 변이체

[0177]

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행하였다. 각 위치가 아미노산을 코딩하는 코돈인 2개의 비연속 코돈 위치에 대해 합성된 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스터를 생성하였다. 이러한 방식에서, 2개의 위치에 대해 위치당 19종의 변이체가 생성되었다.

[0178]

## 실시예 10: 단일 스트레치, 삼중 위치 올리고핵산 변이체

[0179]

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행하였다. 3개의 연속된 코돈 위치에 대해 합성된 참조 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스터를 생성하였다. 3개의 연속된 코돈 위치 방식에서, 3개의 위치에 대해 위치당 19종의 변이체가 생성되었고, 여기서, 각각의 올리고핵산에 대해 2개의 복제물을 포함하며, 이로써, 114개의 올리고핵산이 합성되었다.

[0180]

## 실시예 11: 다중 부위, 삼중 위치 올리고핵산 변이체

[0181]

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행하였다. 적어도 3개의 비연속 코돈 위치에 대해 합성된 참조 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스터를 생성하였다. 소정의 영역 내에서, 3개의 히스티딘 잔기를 코딩하는 코돈의 위치는 다양하였다.

[0182]

## 실시예 12: 다중 부위, 다중 위치 올리고핵산 변이체

[0183]

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행하였다. 1개 이상의 스트레치 중 1개 이상의 코돈 위치에 대해 합성된 참조 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스

터를 생성하였다. 라이브러리에서 5개의 위치는 다양하였다. 제1 위치는 발현 단백질 중 50/50 K/R 비의 생성물에 대한 코돈을 코딩하였고; 제2 위치는 발현 단백질 중 50/25/25 V/L/S 비의 생성물에 대한 코돈을 코딩하였고, 제3 위치는 발현 단백질 중 50/25/25 Y/R/D 비의 생성물에 대한 코돈을 코딩하였고, 제4 위치는 발현 단백질 중 모든 아미노산에 대하여 동일한 비의 생성물에 대한 코돈을 코딩하였고, 제5 위치는 발현 단백질 중 75/25 G/P 비의 생성물에 대한 코돈을 코딩하였다.

[0185] **실시예 13: 다중 변이체 부위를 갖는 CDR 중의 스트레치**

변이체가 각 위치에서 미리 선택되어 있는 것인, 단일 부위 또는 다중 부위에서의 코돈 변이를 코딩하는 올리고뉴클레오티드 라이브러리를 실시예 4-6 및 8-12에서와 같이 생성한다. 변이체 영역은 CDR의 적어도 일부를 코딩한다. 예를 들어, 도 10을 참조한다. 합성된 올리고핵산을 장치 표면으로부터 유리시키고, 프라이머로서 사용하여 핵산 라이브러리를 생성하고, 이를 세포에서 발현시켜 변이체 단백질 라이브러리를 생성한다. 변이체 항체를 에피토프에의 결합 친화도 증가에 관하여 사정한다.

[0187] **실시예 14: 다양한 펩티드를 발현하기 위한 모듈식 플라스미드 성분**

도 11에 도시된 바와 같이, 발현 구축물 카세트의 일부를 구성하는 각각의 별개 영역에 대한 단일 부위 또는 다중 부위에서의 코돈 변이를 코딩하는 올리고뉴클레오티드 라이브러리를 실시예 4-6 및 8-12에서와 같이 생성한다. 2개의 구축물을 발현하는 한 카세트를 생성하기 위해 제1 프로모터(1110), 제1 오픈 리딩 프레임(1120), 제1 터미네이터(1130), 제2 프로모터(1140), 제2 오픈 리딩 프레임(1150), 또는 제2 터미네이터 서열(1160)로 이루어진 변이체 서열의 적어도 일부를 코딩하는 변이체 올리고핵산을 합성하였다. 이전 실시예에서 기술된 바와 같이, 증폭 라운드 후, 1,024개의 발현 구축물로 이루어진 라이브러리가 생성된다.

[0189] **실시예 15: 다중 부위, 단일 위치 변이체**

핵산의 적어도 일부를 코딩하는 영역 중의 단일 부위 또는 다중 부위에서의 코돈 변이를 코딩하는 올리고뉴클레오티드 라이브러리를 실시예 4-6 및 8-12에서와 같이 생성한다. 라이브러리가 다중 부위, 단일 위치 변이체로 구성된 것인, 올리고핵산 변이체의 라이브러리가 생성된다. 예를 들어, 도 6b를 참조한다.

[0191] **실시예 16: 변이체 라이브러리 합성**

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행한다. 각각의 동일하지 않은 올리고핵산은 아미노산 서열의 상이한 코돈 변이체를 코딩하는 것인, 적어도 30,000개의 동일하지 않은 올리고핵산을 드 노보 합성한다. 적어도 30,000개의 동일하지 않은 올리고핵산 각각의 것에 대한 소정의 서열과 비교하였을 때, 합성된 적어도 30,000개의 동일하지 않은 올리고핵산은 1,000개의 염기 중 1개 미만인 총 오류율을 가진다. 라이브러리는 긴 핵산의 PCR 돌연변이유발을 위해 사용되고, 적어도 30,000개의 동일하지 않은 변이체 핵산이 형성된다.

[0193] **실시예 17: 웰에서의 변이체 라이브러리 합성**

실시예 2에 기술된 것과 유사한 조건하에서 드 노보 올리고핵산 합성을 수행한다. 2개의 코돈 위치에 대해 합성된 참조 올리고핵산의 소정의 변이체를 함유하는 것인, 장치 상의 단일 클러스터를 생성한다. 2개의 연속된 코돈 위치 방식에서, 2개의 위치에 대해 위치당 19종의 변이체가 생성되었고, 여기서, 각각의 올리고핵산에 대해 2개의 복제물을 포함하며, 이로써, 38개의 올리고핵산이 합성되었다. 각각의 변이체 서열의 길이는 40개의 염기이다. 동일한 클러스터에서, 추가의 비변이체 올리고핵산 서열이 생성되고, 여기서, 추가의 비변이체 올리고핵산 및 변이체 핵산은 집합적으로 유전자의 코딩 서열의 38종의 변이체를 코딩한다. 각 올리고핵산은 또 다른 올리고핵산에 대해 역상보성인 적어도 하나의 영역을 가진다. 클러스터 중 올리고핵산을 기체 암모니아 절단에 의해 유리시킨다. 물을 포함하는 펫을 클러스터와 접촉시키고, 올리고핵산을 취하고, 올리고핵산을 작은 바이알로 옮겨 놓는다. 바이알은 또한 중합효소 사이클링 조립(PCA: polymerase cycling assembly) 반응을 위해 DNA 폴리머라제 시약을 함유한다. 올리고핵산은 어닐링하고, 캡은 연장 반응에 의해 충전되고, 생성된 이중 가닥 DNA 분자가 형성되며, 이로써, 변이체 핵산 라이브러리가 형성된다. 변이체 핵산 라이브러리를 임의적으로 제한 효소에 가한 후, 이어서, 발현 벡터로 라이제이션시킨다.

[0195] **실시예 18: 단백질 결합 친화도 변화에 관한 변이체 핵산 라이브러리 스크리닝**

실시예 16 또는 17에 기술된 바와 같이 복수의 발현 벡터를 생성한다. 본 실시예에서, 발현 벡터는 HIS 태그부착된 박테리아 발현 벡터이다. 벡터 라이브러리를 박테리아 세포 내로 전기천공시킨 후, HIS 태그부착된 변이체 단백질의 발현 및 정제를 위해 클론을 선택한다. 표적 분자에의 결합 친화도의 변화에 대하여 변이체 단백질을

스크리닝한다.

[0197] 방법에 의해, 예컨대, 금속 이온 코팅된 수지(예컨대, IDA-아가로스 또는 NTA-아가로스)를 사용하여 HIS 태그부착된 단백질을 단리시키는 것인 금속 친화성 크로마토그래피(IMAC: metal affinity chromatography)를 사용하여 친화도를 조사한다. 히스티딘 잔기 스트링이 특이적인 완충제 조건하에서 니켈, 코발트 및 구리를 비롯한, 여러 유형의 고정화된 금속 이온에 결합하기 때문에 발현된 His 태그부착된 단백질을 정제할 수 있고, 검출할 수 있다. 예시적인 결합/세척 완충제는 10-25 mM 이미다졸을 함유하는, pH 7.2의 트리스 완충제 염수(TBS: Tris-buffer saline)를 구성한다. IMAC 칼럼으로부터 포획된 His 태그부착된 단백질의 용리 및 회수는 고농도의 이미다졸(적어도 200 mM)(용리제), 낮은 pH(예컨대, 0.1 M 글리신-HCl, pH 2.5) 또는 과량의 강한 칼레이터(예컨대, EDTA)로 수행된다.

[0198] 대안적으로, 항His-태그 항체는 His 태그부착된 단백질을 포함하는 검정 방법, 예컨대, His 태그부착된 단백질을 단리시키는 풀-다운 검정법, 또는 His 태그부착된 단백질을 검출하는 면역블롯팅 검정법에서 사용하기 위한 용도로 상업적으로 이용가능하다.

#### 실시예 19: 세포 부착 및 이동의 조절인자에 대한 활성 변화에 관한 변이체 혼합 라이브러리 스크리닝

[0200] 실시예 16 또는 17에 기술된 바와 같이 복수의 발현 벡터를 생성한다. 본 실시예에서, 발현 벡터는 GFP 태그부착된 포유동물 발현 벡터이다. 라이브러리로부터 단리된 클론을 포유동물 세포에 일시적으로 형질감염시킨다. 대안적으로, 단백질을 발현시키고, 발현 구축물을 함유하는 세포로부터 단리시킨 후, 추가 측정을 위해 단백질을 세포로 전달한다. 면역형광 검정법을 수행하여 GFP 태그부착된 변이체 발현 생성물의 세포 국재화의 변화를 사정한다. FACS 검정법을 수행하여 GFP 태그부착된 변이체 단백질 발현 생성물의 비변이체 베전과 상호작용하는 막획단 단백질의 입체형태적 상태의 변화를 사정한다. 상처 치유 검정법을 수행하여 GFP 태그부착된 변이체 단백질을 발현하는 세포의, 세포 배양 디쉬 상의 스크레이프에 의해 생성된 공간을 침입할 수 있는 능력의 변화를 사정한다. 형광 광원 및 카메라를 이용하여 GFP 태그부착된 단백질을 발현하는 세포를 확인하고, 추적한다.

#### 실시예 20: 바이러스성 장애 진행을 억제하는 웨티드에 대한 변이체 혼합 라이브러리 스크리닝

[0202] 실시예 16 또는 17에 기술된 바와 같이 복수의 발현 벡터를 생성하였다. 본 실시예에서, 발현 벡터는 FLAG 태그부착된 포유동물 발현 벡터이고, 변이체 혼합 라이브러리는 웨티드 서열을 코딩한다. 1차 포유동물 세포는 바이러스성 장애를 앓는 피험체로부터 수득한 것이다. 대안적으로, 건강한 피험체로부터의 1차 세포를 바이러스로 감염시킨다. 세포를 일련의 마이크로웰 디쉬 상에 플레이팅한다. 변이체 라이브러리로부터 단리된 클론을 세포에 일시적으로 형질감염시킨다. 대안적으로, 단백질을 발현시키고, 발현 구축물을 함유하는 세포로부터 단리시킨 후, 추가 측정을 위해 단백질을 세포로 전달한다. 세포 생존 검정법을 수행하여 변이체 웨티드와 연관된 생존 증진에 대해 감염된 세포를 사정한다. 예시적인 바이러스로는 제한 없이, 조류 독감 바이러스, 지카 바이러스, 한타바이러스, C형 간염 바이러스, 두창 바이러스를 포함한다.

[0203] 한 예시적인 검정법은, 세포에 침가되었을 때, 원형질막을 가로질러 확산되고, 경미하게 양이온성을 띠는 뉴트럴 레드의 특성에 기인하여 산성 리소좀 구획에 축적되는 것인 뉴트럴 레드 염료를 사용하여 뉴트럴 레드 세포 독성 검정법이다. 바이러스 유도성 세포 변성은 막의 단편화 및 리소좀 ATP 주동 양성자 전위 활성 손실을 유도한다. 그 결과에 따른 세포내 뉴트럴 레드의 감소는 다중웰 플레이트 포맷으로 분광 광도법을 사용하여 사정될 수 있다. 변이체 웨티드를 발현하는 세포를 신호 색상 획득(gain-of-signal color) 검정법에서 세포내 뉴트럴 레드의 증가에 의해 점수화한다. 바이러스 유도성 세포 변성을 억제시키는 웨티드에 대해 세포를 사정한다.

#### 실시예 21: 세포의 대사 활성을 증가 또는 감소시키는 변이체 단백질에 관한 스크리닝

[0205] 세포의 대사 활성을 변화시키는 발현 생성물을 확인하기 위한 목적으로 실시예 16 또는 17에 기술된 바와 같이 복수의 발현 벡터를 생성하였다. 본 실시예에서, 발현 벡터를 (예컨대, 형질감염 또는 형질도입을 통해) 일련의 마이크로웰 디쉬 상에 플레이팅된 세포로 전달한다. 이어서, 세포를 하나 이상의 대사 활성 변화에 관하여 스크리닝한다. 대안적으로, 단백질을 발현시키고, 발현 구축물을 함유하는 세포로부터 단리시킨 후, 대사 활성을 측정하기 위해 단백질을 세포로 전달한다. 임의적으로, 하나 이상의 대사 활성 변화에 관하여 스크리닝하기 이전에 대사 활성을 측정하기 위한 세포를 독소로 처리한다. 투여되는 독소의 예로는 제한 없이, 보툴리눔 독소(면역학적 유형: A, B, C1, C2, D, E, F, 및 G 포함), 스타필로코구스(*staphylococcus*) 장내 독소 B, 예르시니아 페스티스(*Yersinia pestis*), C형 간염, 머스타드 작용제, 중금속, 시아나이드, 내독소, 바실러스 안트라시스(*Bacillus anthracis*), 지카 바이러스, 조류 독감, 제초제, 살충제, 수은, 유기포스페이트, 및 리신을 포함한다.

[0206]

기초 에너지 요구량은 호기성 트리카복실산(TCA: aerobic tricarboxylic acid) 또는 크레브스(Kreb's) 회로, 또는 협기성 해당 반응을 포함하는 산화성 인산화에 의해 대사 기질, 예컨대, 글루코스의 산화로부터 도출된다. 해당 반응이 주된 에너지 공급원일 때, 세포의 대사 활성은 세포가 산성을 띠는 대사 생성물, 예컨대, 락테이트 및  $\text{CO}_2$ 를 배출하는 속도를 모니터링함으로써 추정될 수 있다. 호기성 대사인 경우, 세포의 산소 소비 및 산화성 자유 라디칼 생성이 세포의 에너지 요구량을 반영한다. 세포내 산화-환원 전위는 NADH 및  $\text{NAD}^+$ 의 자가 형광 측정값에 의해 측정될 수 있다. 예컨대, 열과 같이, 세포에 의해 방출된 에너지의 양은 정상 환경하에서 산소 소비량(예컨대, 4.8 kcal/1  $\text{O}_2$ )으로부터 예측될 수 있는, 대사 동안 생산 및/또는 소비된 물질에 대한 분석 값으로부터 도출된다. 열 생산과 산소 이용률 사이의 커플링은 독소에 의해 교란될 수 있다. 직접 미세 열량 측정법은 열적으로 격리된 샘플의 온도 상승을 측정한다. 따라서, 산소 소비 측정과 함께 조합되었을 때, 열량 측정법은 독소의 비커플링 활성을 검출하는 데 사용될 수 있다.

[0207]

대사 활성의 각종 마커의 변화를 측정하는 다양한 방법 및 장치가 당업계에 공지되어 있다. 예를 들어, 상기 방법, 장치, 및 마커는 미국 특허 번호 7,704,745(상기 특허는 그 전문이 본원에서 참조로 포함된다)에서 논의된다. 간략하면, 하기 특징들 중 임의의 것의 측정값을 각 세포 집단에 대하여 기록한다: 글루코스, 락테이트,  $\text{CO}_2$ , NADH 및  $\text{NAD}^+$  비, 열,  $\text{O}_2$  소비, 및 자유 라디칼 생산. 스크리닝된 세포로는 간세포, 대식세포 또는 신경아세포종 세포를 포함할 수 있다. 스크리닝된 세포로는 세포주, 피험체로부터의 1차 세포, 또는 모델 시스템(예컨대, 마우스 모델)으로부터의 세포를 포함할 수 있다.

[0208]

다중웰 플레이트의 챔버 내에 위치하는 단일 세포, 또는 세포 집단의 산소 소비 속도를 측정하는 데 다양한 기술이 이용가능하다. 예를 들어, 세포를 포함하는 챔버는 온도, 전류 또는 형광 변화를 기록하기 위한 센서 뿐만 아니라, 광학 시스템, 예컨대, 형광을 모니터링하기 위해 각 챔버에 결합된, 섬유 결합형 광학 시스템을 가질 수 있다. 본 실시예에서, 각 챔버는 조명원이 챔버 내부의 분자를 여기시킬 수 있도록 하기 위해 창을 가진다. 섬유 결합형 광학 시스템은 자가 형광을 검출하여 세포내 NADH/NAD 비를 측정할 수 있고, 전압 및 칼슘 감수성 염료를 검출하여 막횡단 전위 및 세포내 칼슘을 측정할 수 있다. 추가로,  $\text{CO}_2$  및/또는  $\text{O}_2$  감수성 형광 염료 신호의 변화를 검출한다.

[0209]

#### 실시예 22: 암 세포의 선택적 표적화에 관한 변이체 핵산 라이브러리 스크리닝

[0210]

실시예 16 또는 17에 기술된 바와 같이 복수의 발현 벡터를 생성한다. 본 실시예에서, 발현 벡터는 FLAG 태그부착된 포유동물 발현 벡터이고, 변이체 핵산 라이브러리는 웨티드 서열을 코딩한다. 변이체 라이브러리로부터 단리된 클론을 암 세포 및 비암 세포에 따로따로 일시적으로 형질감염시킨다. 각각이 변이체 핵산에 의해 코딩되는 변이체 웨티드를 발현하는 것인 암 세포 및 비암 세포, 둘 모두에 대하여 세포 생존 및 세포 사멸 검정법을 수행한다. 변이체 웨티드와 연관된 선택적 암 세포 사멸에 관하여 세포를 사정한다. 암 세포는 임의적으로 암 진단을 받은 피험체로부터의 암 세포주 또는 원발성 암 세포이다. 암 진단을 받은 피험체로부터의 원발성 암 세포인 경우, 스크리닝 검정법에서 확인된 변이체 웨티드는 임의적으로 피험체에게로의 투여를 위해 선택된다. 대안적으로, 단백질을 발현시키고, 단백질 발현 구축물을 함유하는 세포로부터 단리시킨 후, 추가 측정을 위해 단백질을 암 세포 및 비암 세포로 전달한다.

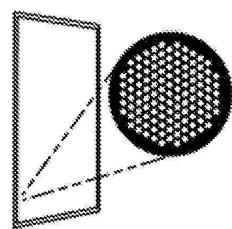
[0211]

본 개시내용의 바람직한 실시양태가 본원에 제시 및 기술되어 있지만, 상기 실시양태는 단지 예로서 제공된 것이라는 것이 자명할 것이다. 이에 본 발명을 벗어나지 않으면서 다수의 변형, 변화, 및 치환이 당업자에게 이루어질 수 있을 것이다. 본원에 기술된 본 개시내용의 실시양태에 대한 다양한 대안이 본 개시내용을 실시하는 데 사용될 수 있다는 것을 이해하여야 한다. 하기 특허청구범위는 본 개시내용의 범주를 정의하고, 이들 특허청구 범위의 내의 방법 및 구조, 및 그의 등가물이 특허청구범위에 의해 포함되는 것으로 의도된다.

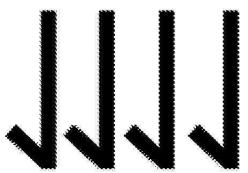
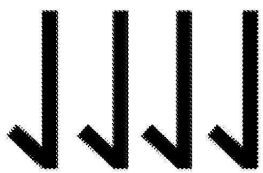
도면

도면1

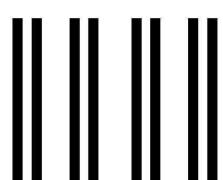
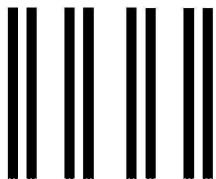
도 1a



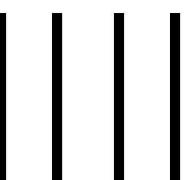
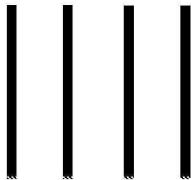
도 1b



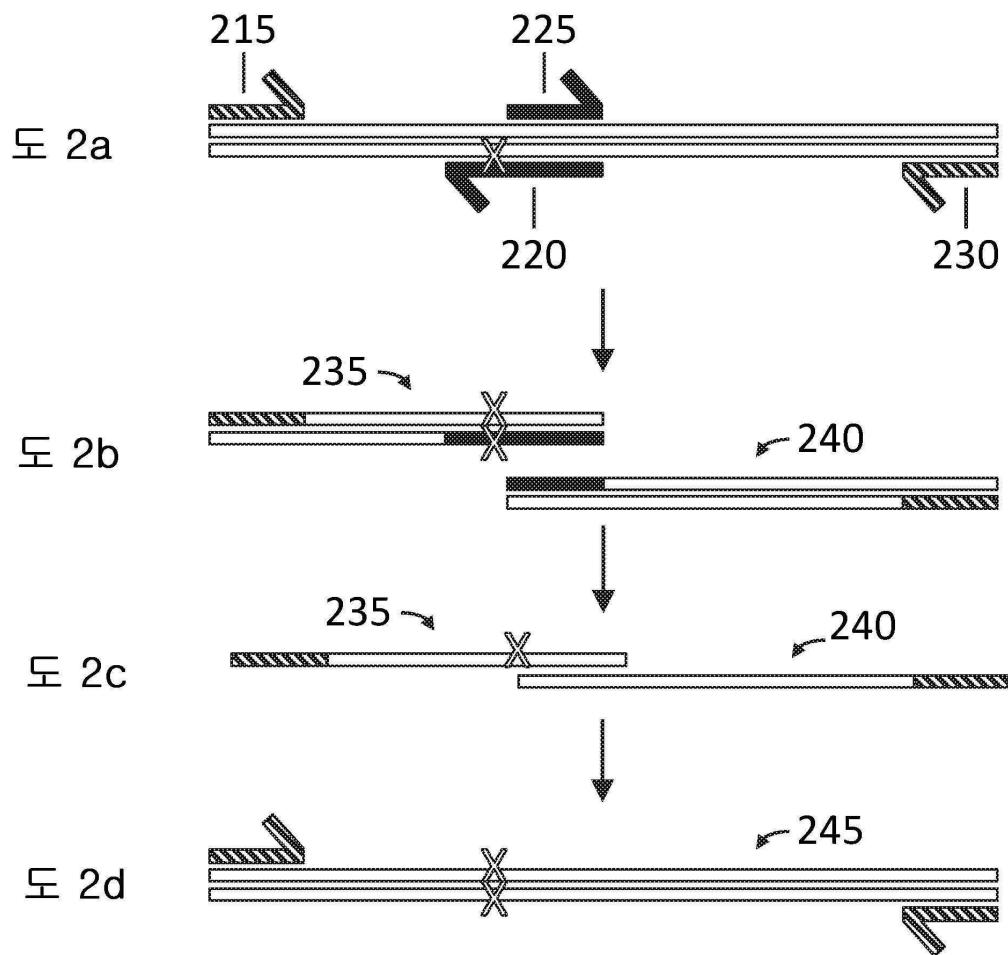
도 1c



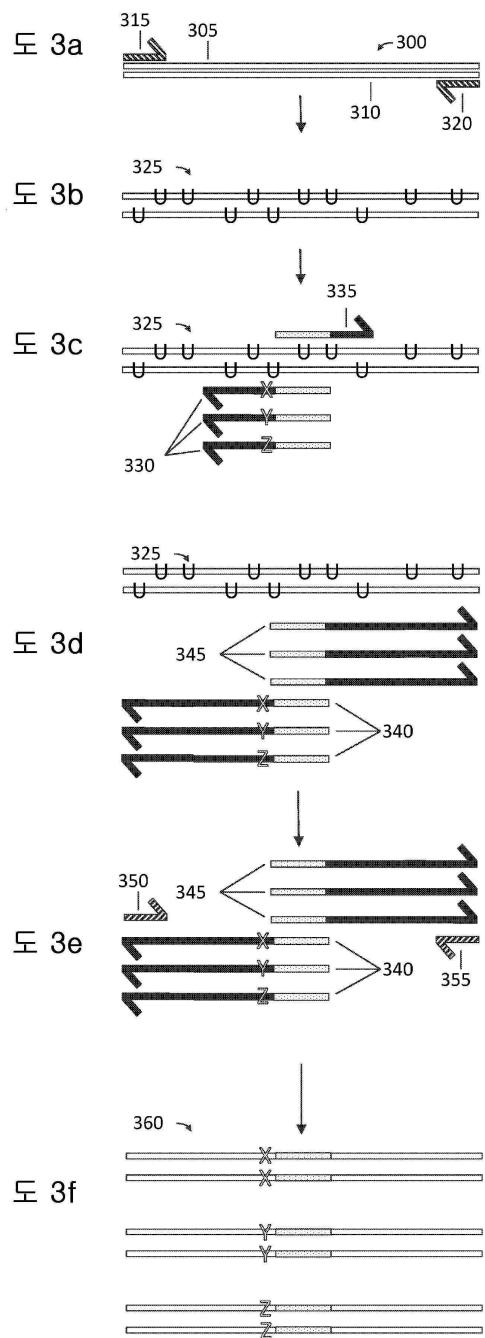
도 1d



도면2



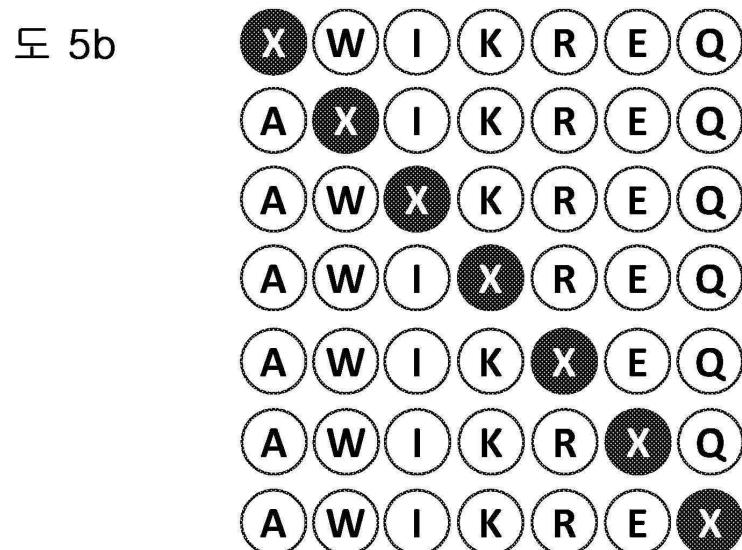
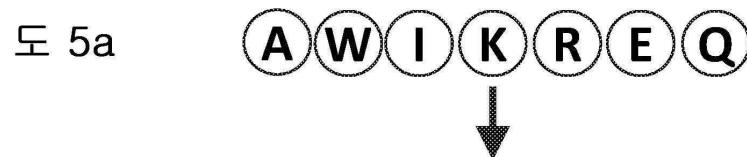
## 도면3



## 도면4

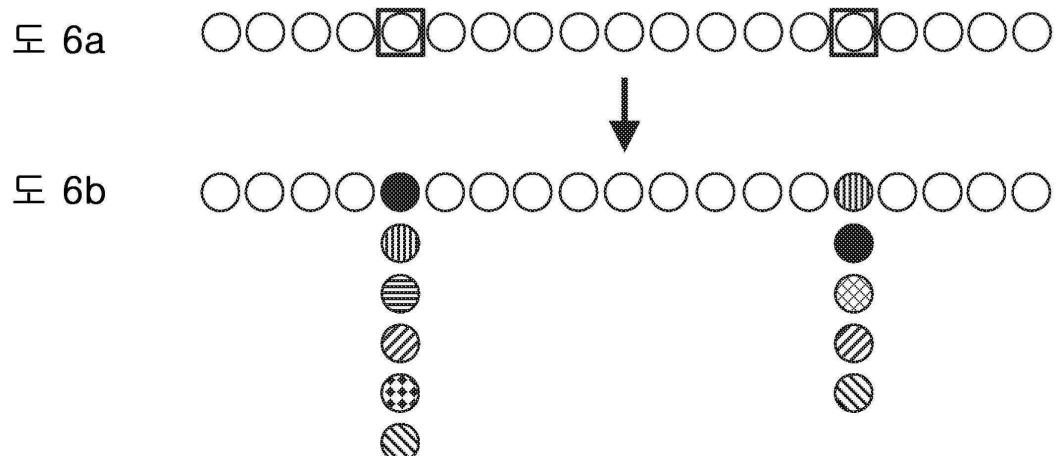


## 도면5

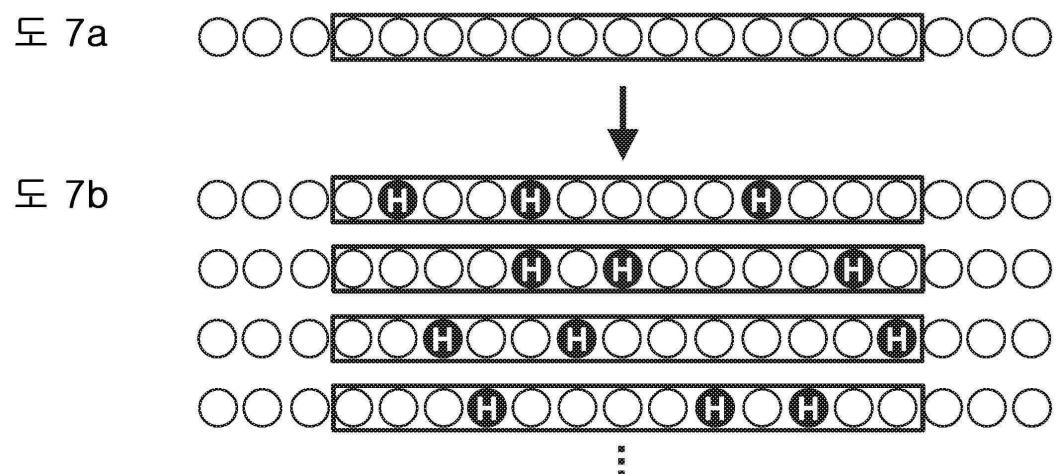


= 임의의 코돈

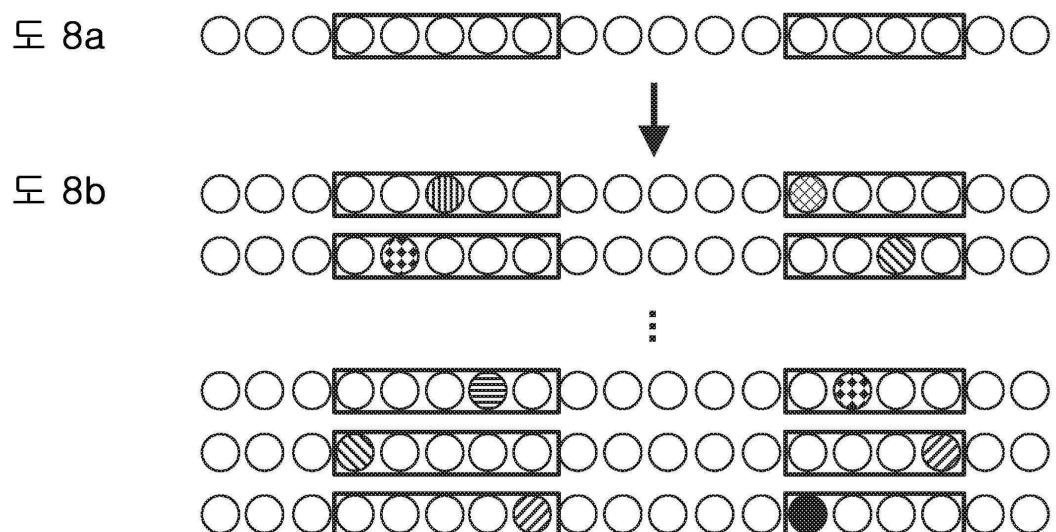
## 도면6



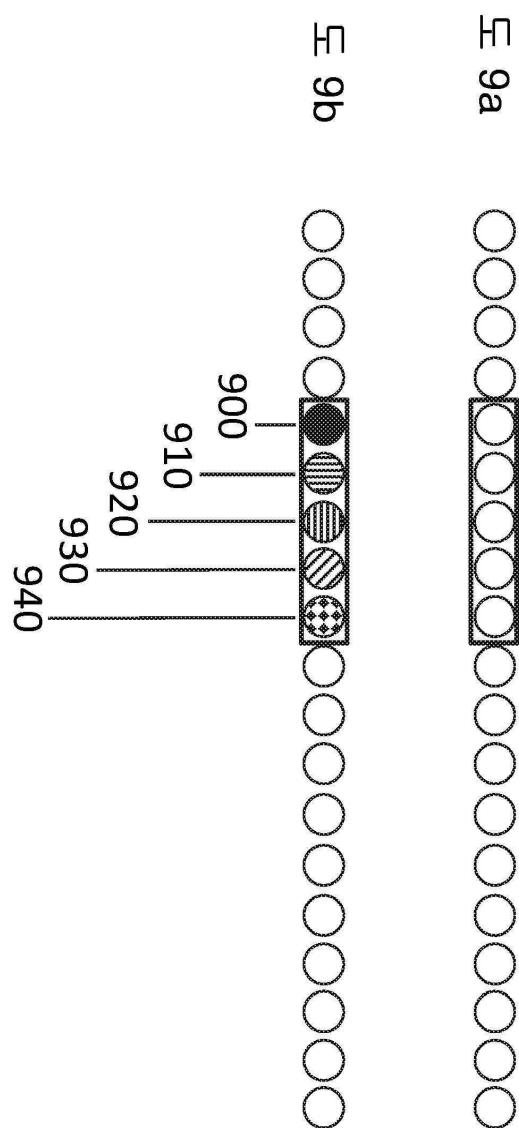
## 도면7



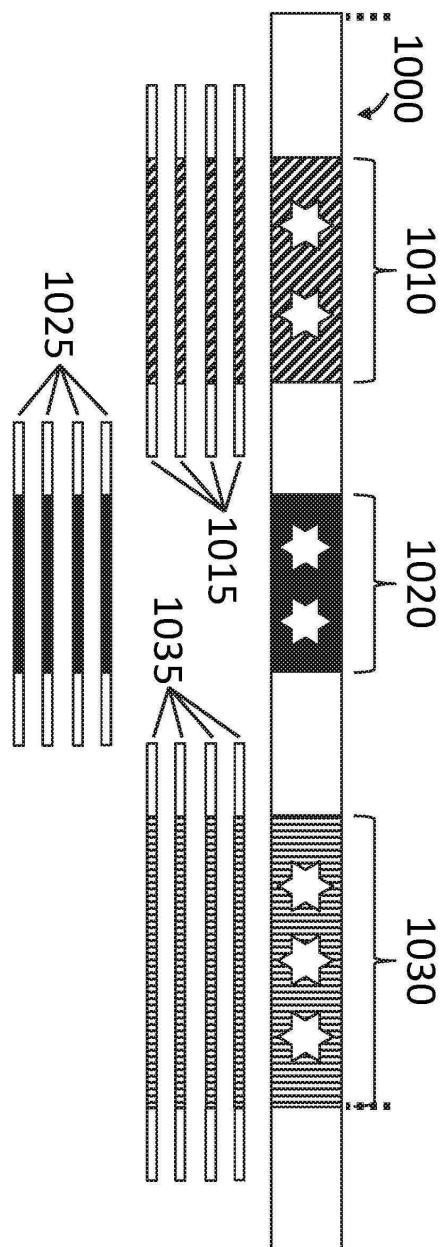
## 도면8



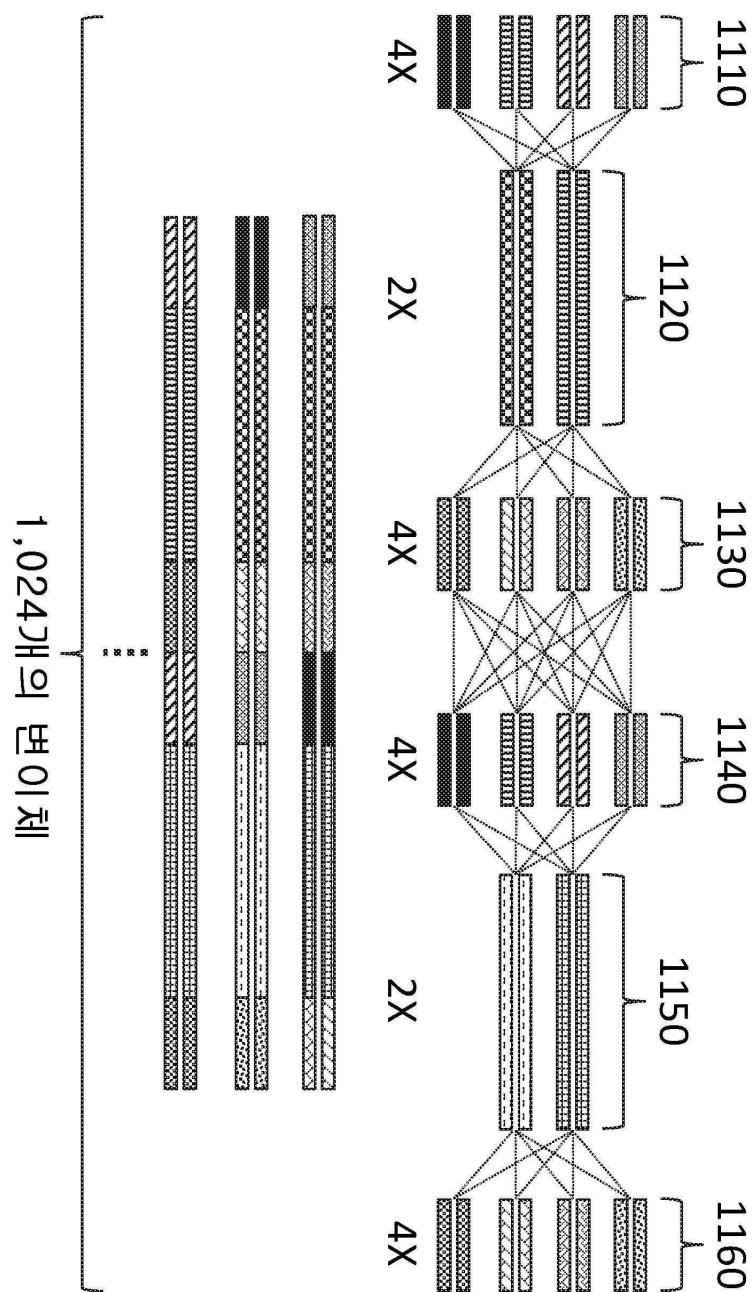
도면9



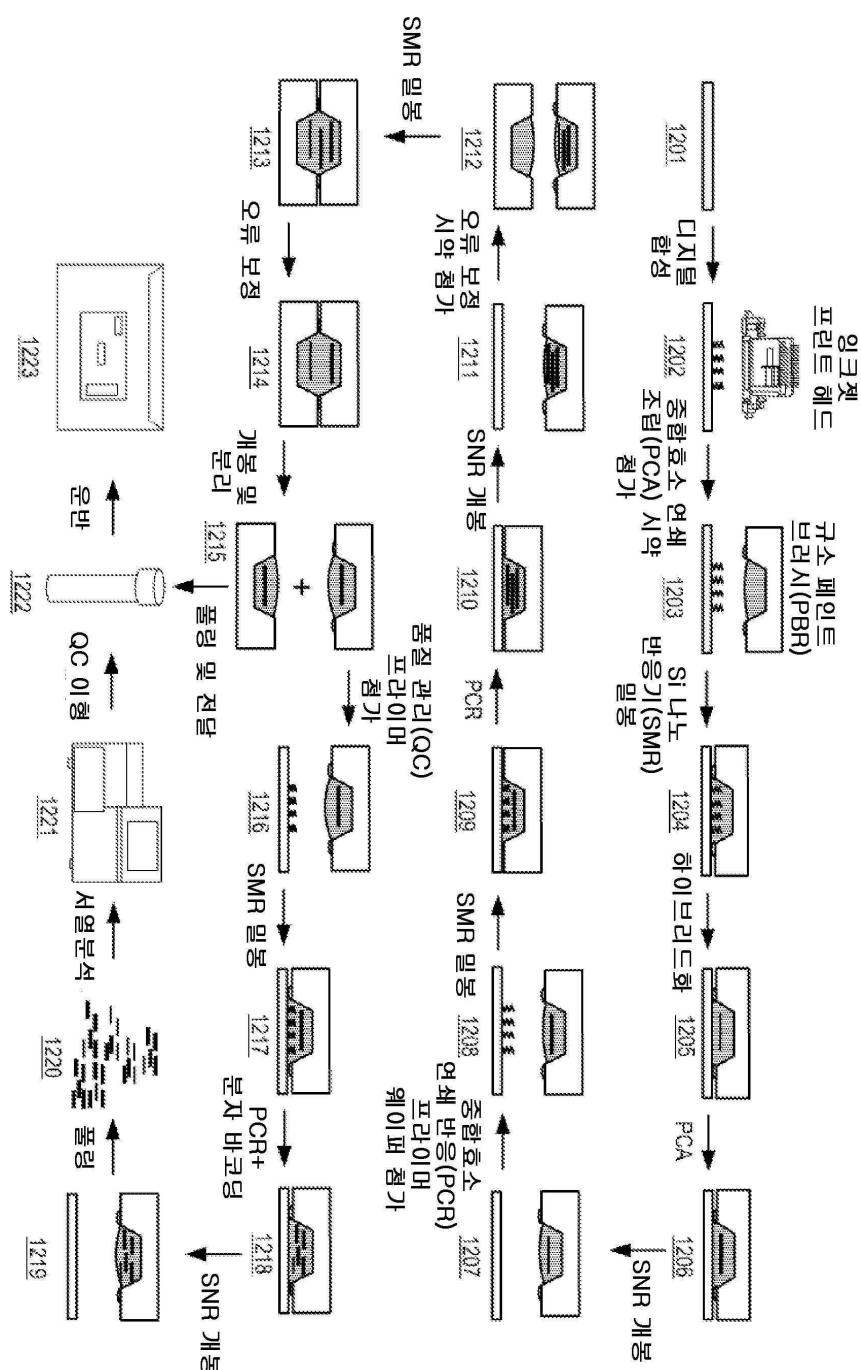
도면10



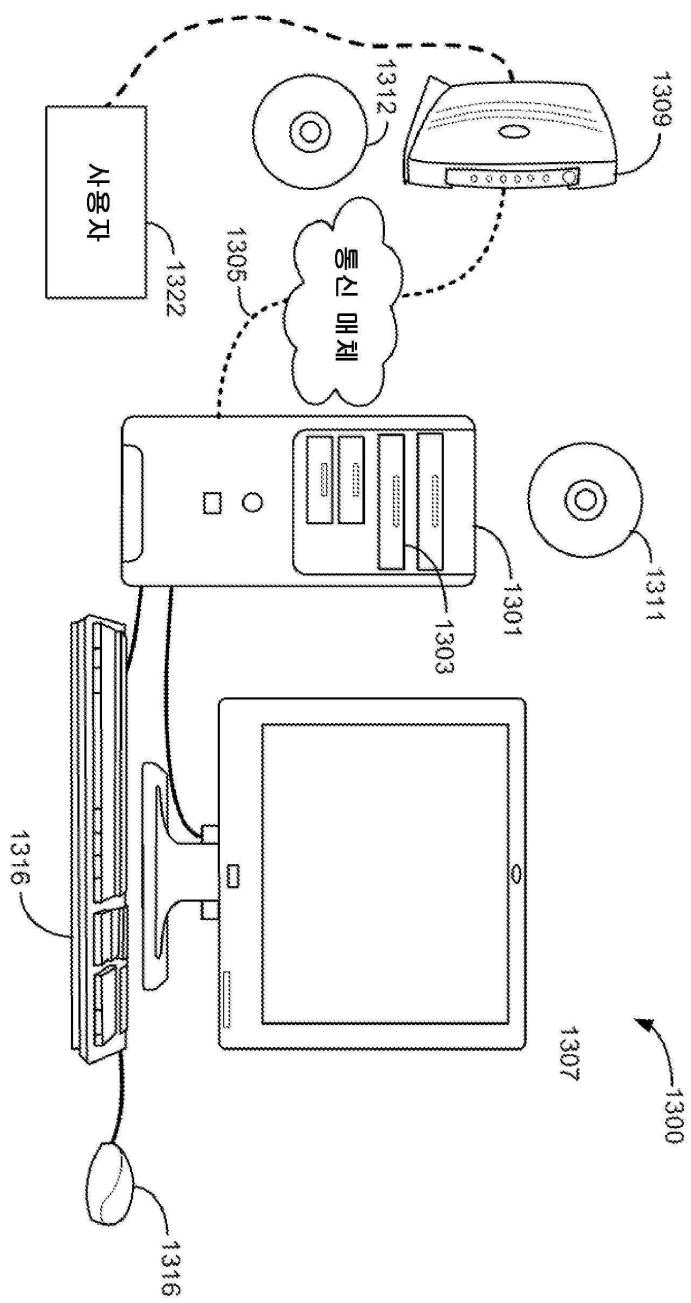
도면11



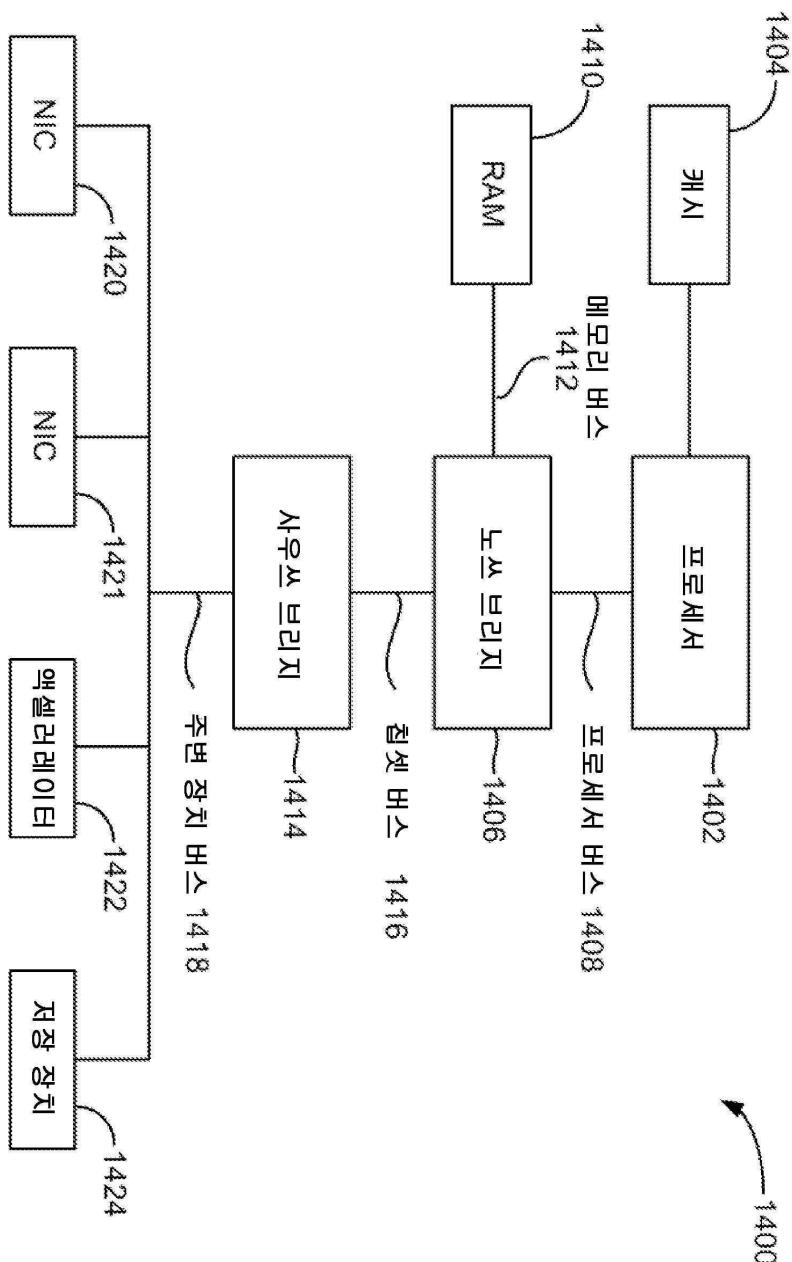
도면12



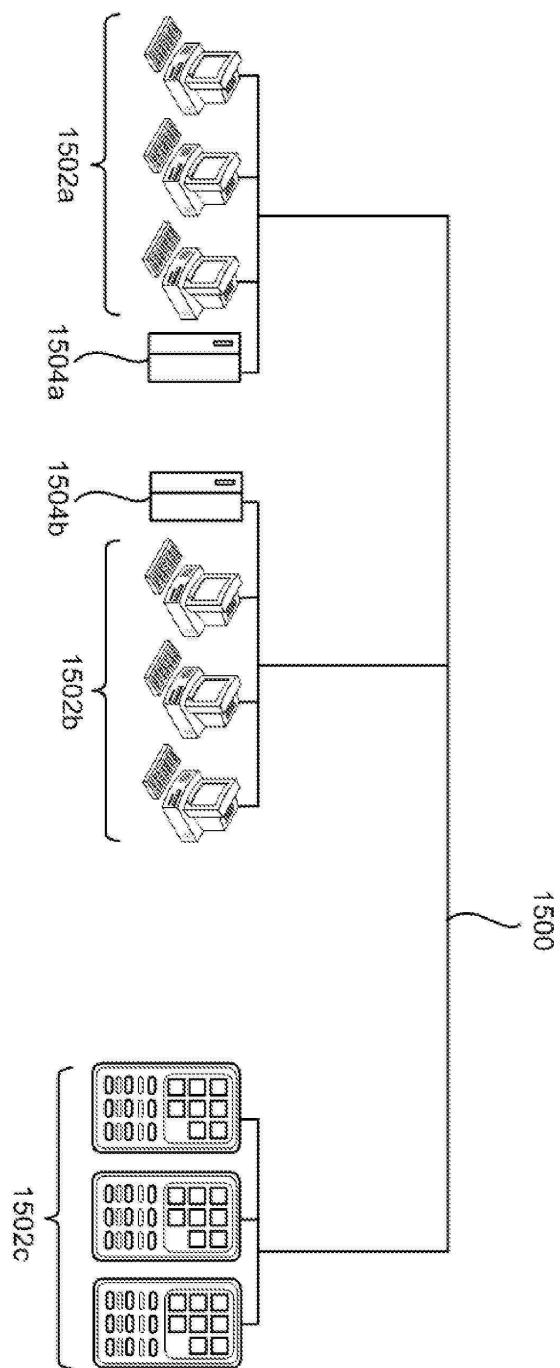
### 도면13



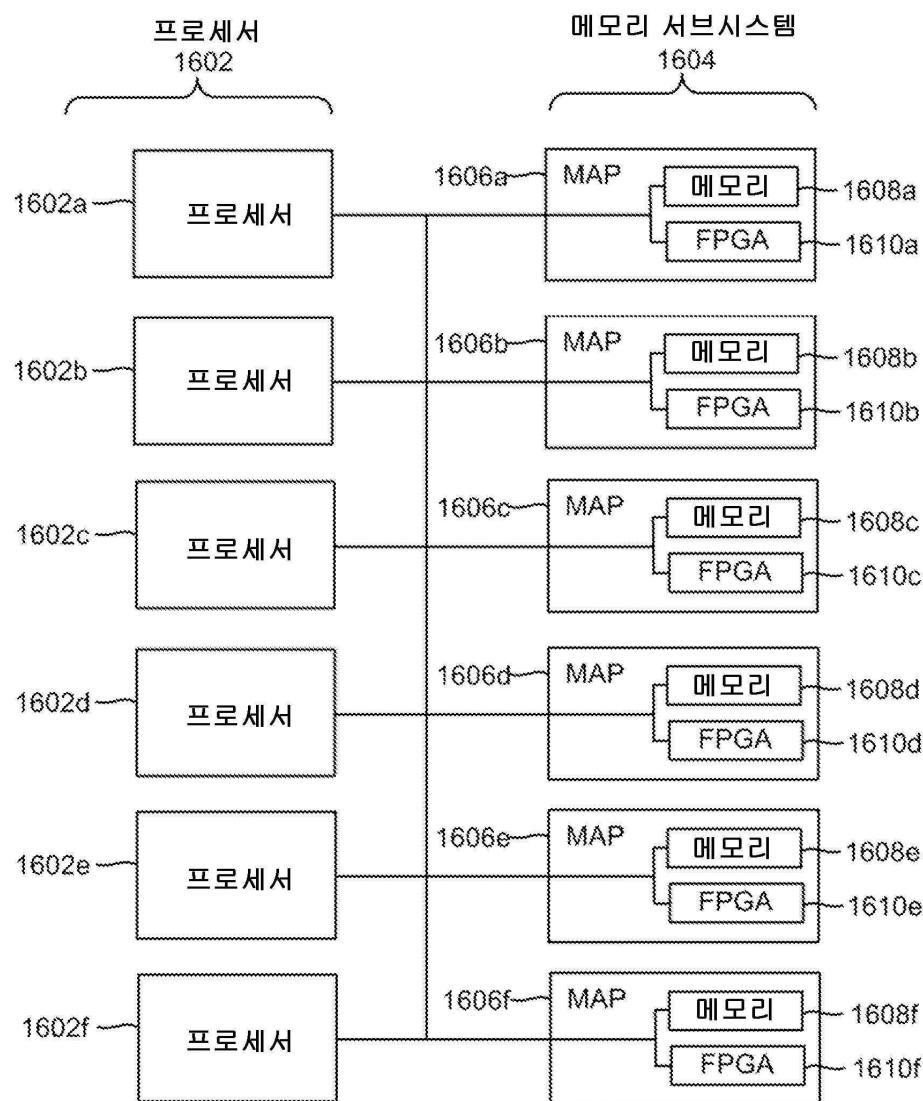
## 도면14



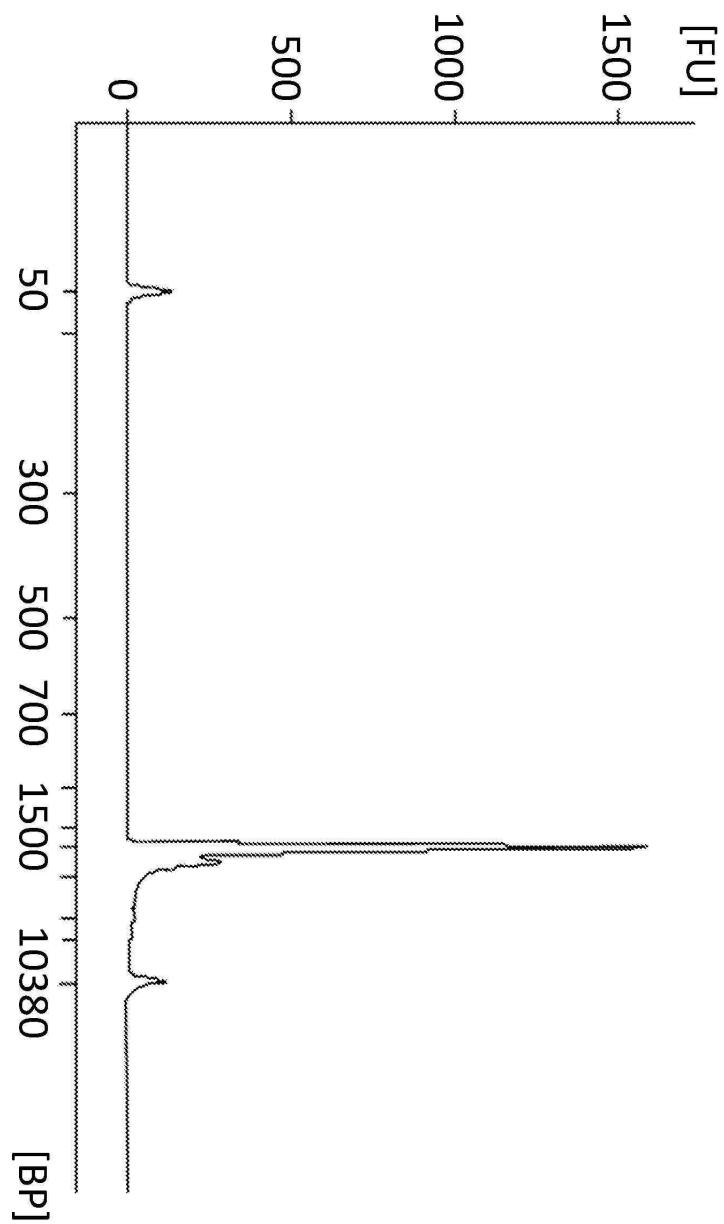
도면15



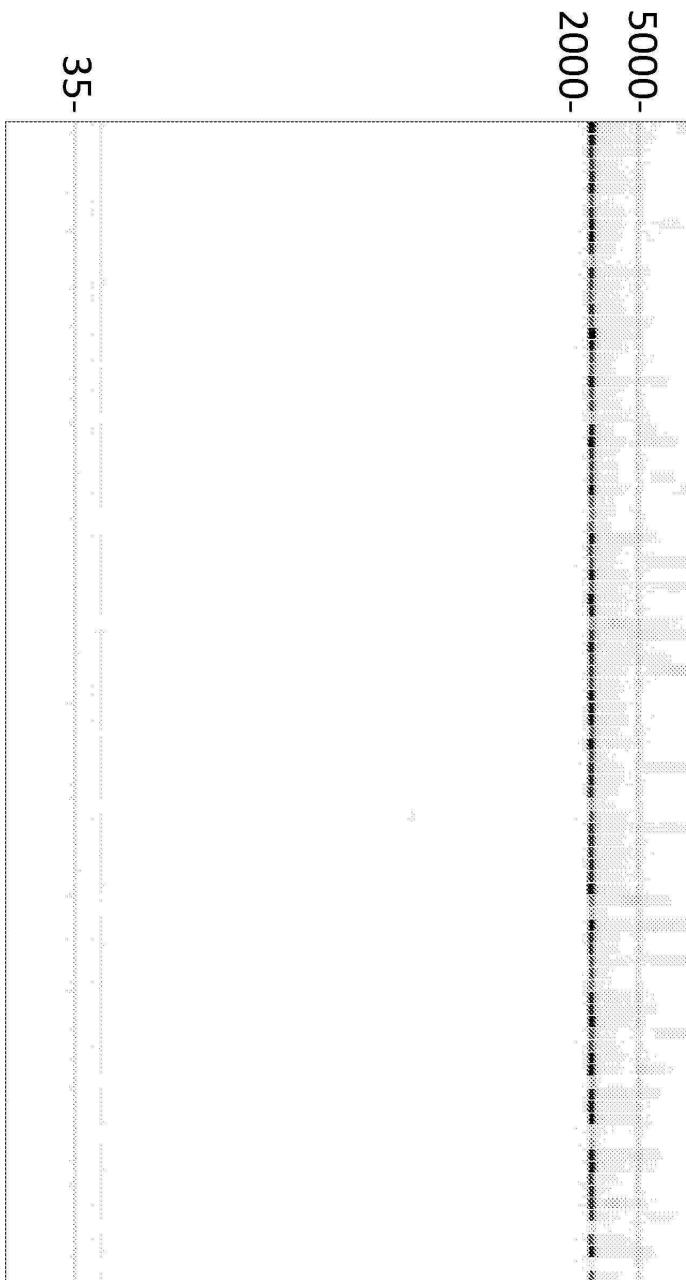
## 도면16



도면17



도면18



서 열 목 록

SEQUENCE LISTING

<110> TWIST BIOSCIENCE CORPORATION

<120> OLIGONUCLEIC ACID VARIANT LIBRARIES AND SYNTHESIS THEREOF

<130> 44854-718.601

<140><141><150> 62/354,034

<151> 2016-06-23

<150> 62/263,548

<151> 2015-12-04

<150> 62/220,879

<151> 2015-09-18

<160> 31

<170> Patent In version 3.5

<210> 1

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 1

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccat

44

<210> 2

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 2

atgttagca agggcgagga gctgttcacc ggggtggtgc ccat

44

<210> 3

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 3

atgttaagca agggcgagga gctgttcacc ggggtggtgc ccat

44

<210> 4

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223>

> Description of Artificial Sequence: Synthetic  
primer

<400> 4

atgatttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

<210> 5

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 5

atgtcttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

<210> 6

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400

> 6

atgccttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

<210> 7

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 7

atgacttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

<210> 8

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

&lt;400&gt; 8

atggctagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 9

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Description of Artificial Sequence: Synthetic

primer

&lt;400&gt; 9

atgtatagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 10

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Description of Artificial Sequence: Synthetic

primer

&lt;400&gt; 10

atgcatagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 11

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic

primer

&lt;400&gt; 11

atgcaaagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 12

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Description of Artificial Sequence: Synthetic

primer

&lt;400&gt; 12

atgaatagca agggcgagga gctgttcacc ggggtggtgc ccat	44
<210> 13	
<211> 44	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 13	
atgaaaagca agggcgagga gctgttcacc ggggtggtgc ccat	44
<210> 14	
<211> 44	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 14	
atggatagca agggcgagga gctgttcacc ggggtggtgc ccat	44
<210> 15	
<211> 44	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 15	
atggaaagca agggcgagga gctgttcacc ggggtggtgc ccat	44
<210> 16	
<211> 44	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 16	
atgtgttagca agggcgagga gctgttcacc ggggtggtgc ccat	44

&lt;210&gt; 17

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

&lt;400&gt; 17

atgtggagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 18

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

><223> Description of Artificial Sequence: Synthetic  
primer

&lt;400&gt; 18

atgcgttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 19

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

&lt;400&gt; 19

atgggttagca agggcgagga gctgttcacc ggggtggtgc ccat 44

&lt;210&gt; 20

&lt;211&gt; 62

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

&lt;400&gt; 20

agacaatcaa ccattgggg tggacagcct tgacctctag acttcggcat tttttttt 60

tt 62

<210> 21

<211> 112

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 21

cgggatcctt atcgtcatcg tcgtacagat cccgaccat ttgctgtcca ccagtcatgc	60
tagccatacc atgatgatga tcatgtatgaa aaccccgat tttttttt tt	112

<210> 22

<211> 19

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 22

atgcggggtt ctcatcatc	19
----------------------	----

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer

<400> 23

cgggatcctt atcgtcatcg	20
-----------------------	----

<210> 24

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 24

Ala Trp Ile Lys Arg Glu Gln

1 5

<210> 25

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

<220><221> MOD\_RES

<222> (1)..(1)

<223> Any amino acid

<400> 25

Xaa Trp Ile Lys Arg Glu Gln

1 5

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

<220><221> MOD\_RES

<222> (2)..(2)

<223> Any amino acid

<400> 26

Ala Xaa Ile Lys Arg Glu Gln

1 5

<210> 27

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

<220><221> MOD\_RES

<222> (3)..(3)

<223> Any amino acid

&lt;400&gt; 27

Ala Trp Xaa Lys Arg Glu Gln

1 5

&lt;210&gt; 28

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (4)..(4)

&lt;223&gt; Any amino acid

&lt;400&gt; 28

Ala Trp Ile Xaa Arg Glu Gln

1 5

&lt;210&gt; 29

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

&lt;220&gt;&lt;221&gt; MOD\_RES

&lt;222&gt; (5)..(5)

&lt;223&gt; Any amino acid

&lt;400&gt; 29

Ala Trp Ile Lys Xaa Glu Gln

1 5

&lt;210&gt; 30

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Description of Artificial Sequence: Synthetic

peptide

<220><221> MOD\_RES

<222> (6)..(6)

<223> Any amino acid

<400> 30

Ala Trp Ile Lys Arg Xaa Gln

1 5

<210> 31

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
peptide

<220><221> MOD\_RES

<222> (7)..(7)

<223> Any amino acid

<400> 31

Ala Trp Ile Lys Arg Glu Xaa

1 5