(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년01월22일
(11) 등록번호 10-2628035
(24) 등록일자 2024년01월17일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6806 (2018.01) *C12N 15/10* (2017.01)
(52) CPC특허분류
C12Q 1/6806 (2018.05)
C12N 15/1093 (2013.01)
(21) 출원번호 10-2019-7038492
(22) 출원일자(국제) 2018년06월05일
심사청구일자 2021년05월31일
(85) 번역문제출일자 2019년12월26일
(65) 공개번호 10-2020-0016287
(43) 공개일자 2020년02월14일
(86) 국제출원번호 PCT/US2018/036078
(87) 국제공개번호 WO 2018/226708
국제공개일자 2018년12월13일
(30) 우선권주장
62/516,324 2017년06월07일 미국(US)
(56) 선행기술조사문헌
JP2010535513 A*
Vitak 등, NATURE METHODS, Vol. 14, No. 3, 페이지 1-10, (2017.01.30.)*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
오레곤 헬스 앤드 사이언스 유니버시티
미국 오레곤주 97239 포틀랜드 샘 잭슨 파크 로드 사우스웨스트 3181
일루미나, 인코포레이티드
미국 캘리포니아 92122 샌디에고 일루미나 웨이 5200
(72) 발명자
애디 앤드류 씨.
미국 오레곤주 97239 포틀랜드 샘 잭슨 파크 로드 사우스웨스트 3181
멀퀸 라이언
미국 오레곤주 97239 포틀랜드 샘 잭슨 파크 로드 사우스웨스트 3181
(뒷면에 계속)
(74) 대리인
특허법인아주김장리

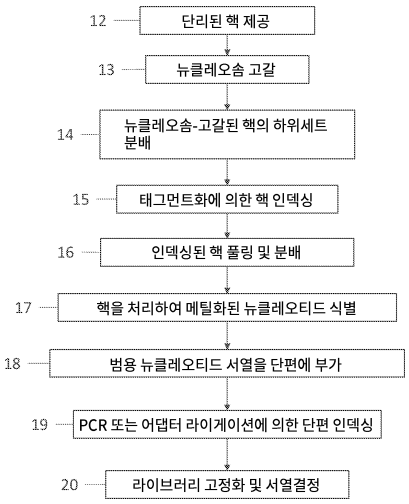
전체 청구항 수 : 총 31 항 심사관 : 김현태

(54) 발명의 명칭 **메틸화 서열결정을 위한 단일 세포 전체 게놈 라이브러리**

(57) 요 약

복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리의 제조 방법이 본 명세서에서 제공된다. 본 발명의 방법은 분할-및-풀링 조합 인덱싱 및 바이설파이트 처리 기법을 조합하여, 다수의 단일 세포의 메틸화 프로파일을 신속하게, 정확하게, 그리고 저렴하게 특성화한다.

대 표 도 - 도1



(52) CPC특허분류
C12Q 2523/101 (2013.01)
C12Q 2523/125 (2013.01)
C12Q 2525/191 (2013.01)
C12Q 2535/122 (2019.08)
C12Q 2537/143 (2013.01)
C12Q 2537/159 (2013.01)
C12Q 2563/159 (2013.01)
C12Q 2563/179 (2013.01)

(72) 발명자
스티머스 프랭크 제이.
미국 캘리포니아주 92122 샌디에이고 5200 일루미나 웨이
포코로크 드미트리 케이.
미국 캘리포니아주 92122 샌디에이고 5200 일루미나 웨이
노어버그 스티븐
미국 캘리포니아주 92122 샌디에이고 5200 일루미나 웨이

명 세 서

청구범위

청구항 1

복수의 단일 세포로부터 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리(sequencing library)의 제조 방법으로서,

(a) 복수의 세포로부터 단리된 핵을 제공하는 단계;

(b) 게놈 DNA로부터 뉴클레오솜의 결합을 해제하기 위해 상기 단리된 핵에 화학적 처리를 하여, 단리된 핵의 온전성을 유지하면서 결합된 뉴클레오솜이 더 적은, 고갈된 뉴클레오솜을 포함하는 핵을 생성하는 단계;

(c) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트를 트랜스포좀 복합체를 포함하는 제1 복수 구획 내로 분배하는 단계로서, 각 구획 내의 트랜스포좀 복합체가 다른 구획 내의 제1 인덱스 서열과 상이한 제1 인덱스 서열을 포함하는 단계;

(d) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트 내의 핵산을 복수의 핵산 단편으로 단편화시키고, 제1 인덱스 서열을 상기 핵산 단편의 적어도 하나의 가닥 내로 혼입시켜, 인덱싱된 핵을 생성하는 단계;

(e) 상기 인덱싱된 핵을 조합하여, 풀링된 인덱싱 핵을 생성하는 단계;

(f) 상기 풀링된 인덱싱 핵의 하위세트를 제2 복수 구획 내로 분배하고, 상기 인덱싱된 핵이 바이설파이트 처리를 겪게 하여, 바이설파이트로 처리된 핵산 단편을 생성하는 단계;

(g) 상기 각 구획 내의 바이설파이트로 처리된 핵산 단편을 5' 말단에 범용 뉴클레오티드 서열 및 3' 말단에 무작위 뉴클레오티드 서열을 포함하는 복수의 프라이머를 사용하여 선형 증폭에 의해 증폭시켜, 증폭된 핵산 분자를 생성하는 단계;

(h) 제2 인덱스 서열을 상기 증폭된 핵산 분자 내로 혼입하여, 두 인덱스를 포함하는 핵산 분자를 생성하는 단계로서, 각 구획 내의 제2 인덱스 서열이 다른 구획 내의 제2 인덱스 서열과 상이한 단계; 및

(i) 상기 두 인덱스를 포함하는 핵산 분자를 조합하여, 그에 의해, 복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리를 생성하는 단계를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 2

복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리의 제조 방법으로서,

(a) 복수의 세포로부터 단리된 핵을 제공하는 단계;

(b) 상기 단리된 핵이 화학적 처리를 겪게 하여, 단리된 핵의 온전성을 유지하면서 결합된 뉴클레오솜이 더 적은, 고갈된 뉴클레오솜을 포함하는 핵을 생성하는 단계;

(c) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트를 트랜스포좀 복합체를 포함하는 제1 복수 구획 내로 분배하는 단계로서, 각 구획 내의 트랜스포좀 복합체가 다른 구획 내의 제1 인덱스 서열과 상이한 제1 인덱스 서열을 포함하는 단계;

(d) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트 내의 핵산을 복수의 핵산 단편으로 단편화하고, 제1 인덱스 서열을 핵산 단편의 적어도 하나의 가닥 내로 혼입시켜, 인덱싱된 핵을 생성하는 단계;

(e) 상기 인덱싱된 핵을 조합하여, 풀링된 인덱싱 핵을 생성하는 단계;

(f) 상기 풀링된 인덱싱 핵의 하위세트를 제2 복수 구획 내로 분배하고, 상기 인덱싱된 핵이 바이설파이트 처리를 겪게 하여, 바이설파이트로 처리된 핵산 단편을 생성하는 단계;

(g) 상기 각 구획 내의 바이설파이트 처리된 핵산 단편을 범용 어댑터에 라이게이션시켜, 라이게이션된 핵산 분자를 생성하는 단계;

(h) 제2 인덱스 서열을 상기 라이게이션된 핵산 분자 내로 혼입하여, 두 인덱스를 포함하는 핵산 분자를 생성하

는 단계로서, 각 구획 내의 제2 인덱스 서열이 다른 구획 내의 제2 인덱스 서열과 상이한 것인, 단계; 및

(i) 상기 두 인덱스를 포함하는 핵산 분자를 조합하여, 그에 의해, 복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리를 생성하는 단계를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 3

제1항 또는 제2항에 있어서, 상기 화학적 처리가 핵산과 단백질 간 상호작용을 방해할 수 있는 카오트로픽제(chaotropic agent)로의 처리를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 4

제3항에 있어서, 상기 카오트로픽제가 리튬 다이아이오도살리실레이트를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 5

제1항 또는 제2항에 있어서, 상기 화학적 처리가 핵산과 단백질 간 상호작용을 방해할 수 있는 세제로의 처리를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 6

제5항에 있어서, 상기 세제가 소듐 도데실 설페이트(SDS)를 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 7

제6항에 있어서, 상기 세포가 단계 (a) 이전에 가교결합제로 처리되는, 서열결정 라이브러리의 제조 방법.

청구항 8

제7항에 있어서, 상기 가교결합제가 포름알데히드인, 서열결정 라이브러리의 제조 방법.

청구항 9

제1항 또는 제2항에 있어서, 단계 (c) 및 (f)에서 분배가, 형광 활성화 핵 분류(FANS)에 의해 수행되는, 서열결정 라이브러리의 제조 방법.

청구항 10

제1항 또는 제2항에 있어서, 상기 제1 복수 구획이 다중 웰 플레이트인, 서열결정 라이브러리의 제조 방법.

청구항 11

제1항 또는 제2항에 있어서, 상기 제2 복수 구획이 다중 웰 플레이트인, 서열결정 라이브러리의 제조 방법.

청구항 12

제1항 또는 제2항에 있어서, 상기 트랜스포좀 복합체의 각각이 트랜스포사제(transposase) 및 트랜스포존(transposon)을 포함하며, 상기 트랜스포존의 각각이 전달된 가닥을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 13

제12항에 있어서, 상기 전달된 가닥이 시토신 잔기를 포함하지 않는, 서열결정 라이브러리의 제조 방법.

청구항 14

제13항에 있어서, 상기 전달된 가닥이 제1 인덱스 서열을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 15

제14항에 있어서, 상기 전달된 가닥이 제1 범용 서열(first universal sequence) 및 제1 서열결정 프라이머 서

열을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 16

제1항에 있어서, 상기 바이설파이트로 처리된 핵산 단편의 선형 증폭이 1 내지 10회 사이클을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 17

제1항에 있어서, 상기 단계 (g)에서 프라이머의 5' 말단에서의 범용 뉴클레오티드 서열이 제2 서열결정 프라이머 서열을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 18

제1항에 있어서, 상기 단계 (g)에서 프라이머의 3' 말단에서의 무작위 뉴클레오티드 서열이 9개의 무작위 뉴클레오티드로 이루어지는, 서열결정 라이브러리의 제조 방법.

청구항 19

제1항에 있어서, 상기 단계 (h)에서 제2 인덱스 서열의 혼입이 각 구획 내의 증폭된 핵산 분자를 각각이 인덱스 서열을 포함하는 제1 범용 프라이머 및 제2 범용 프라이머와 접촉시키고, 지수적 증폭 반응을 수행하는 것을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 20

제2항에 있어서, 상기 범용 어댑터의 라이게이션 이전에, 하나 이상의 뉴클레오티드를 상기 바이설파이트로 처리된 핵산 단편의 3' 말단에 부가하여, 3' 오버행을 생성하는 단계를 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 21

제20항에 있어서, 상기 하나 이상의 뉴클레오티드의 부가가 말단 트랜스퍼라제를 사용하여 수행되는, 서열결정 라이브러리의 제조 방법.

청구항 22

제20항에 있어서, 상기 범용 어댑터가 상기 바이설파이트로 처리된 핵산 단편 내의 3' 오버행과 역 상보성인 오버행을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 23

제2항에 있어서, 상기 단계 (h)에서 제2 인덱스 서열의 혼입은 각 구획 내의 두 인덱스를 포함하는 핵산 분자를 각각이 인덱스 서열을 포함하는 제1 범용 프라이머 및 제2 범용 프라이머와 접촉시키고, 지수적 증폭 반응을 수행하는 것을 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 24

제19항 또는 제23항에 있어서, 상기 제1 범용 프라이머가 상기 두 인덱스를 포함하는 핵산 분자의 3' 말단에서의 범용 서열과 상보성인 제1 포획 서열 및 제1 앵커 서열을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 25

제19항 또는 제23항에 있어서, 상기 제2 범용 프라이머가 상기 두 인덱스를 포함하는 핵산 분자의 5' 말단에서의 범용 서열과 상보성인 제2 포획 서열 및 제2 앵커 서열을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 26

제25항에 있어서, 상기 지수적 증폭 반응이 중합효소 연쇄 반응(PCR)을 포함하는, 서열결정 라이브러리의 제조

방법.

청구항 27

제1항 또는 제2항에 있어서, 표적 핵산에 대하여 특이성을 갖는 복수의 포획 올리고뉴클레오티드를 사용한 표적 핵산의 농축을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 28

제27항에 있어서, 상기 포획 올리고뉴클레오티드가 고체 기재(solid substrate)의 표면 상에 고정화된, 서열결정 라이브러리의 제조 방법.

청구항 29

제1항 또는 제2항에 있어서, 사전결정된 크기 범위에 속하는, 두 인덱스를 포함하는 핵산 분자의 선택을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 30

제1항 또는 제2항에 있어서, 복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한, 두 인덱스를 포함하는 핵산 분자의 서열결정을 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 31

제30항에 있어서,

복수의 증폭 부위를 포함하는 표면을 제공하는 단계로서, 상기 증폭 부위가 유리 3' 말단을 갖는 적어도 2개의 부착된 단일 가닥 핵산의 집단을 포함하는 단계, 및

상기 증폭 부위를 포함하는 표면을, 각각이 개별 두 인덱스를 포함하는 핵산 분자로부터의 앰플리콘의 클론 집단을 포함하는 복수의 증폭 부위를 생성하기에 적합한 조건 하에서 서열결정 라이브러리와 접촉시키는 단계를 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

발명의 설명

기 술 분 야

[0001] 관련 출원에 대한 상호-참조

[0002] 본 출원은 2017년 6월 7일자 출원된 미국 가출원 제62/516,324호의 이익을 주장하며, 이는 본 명세서에 그의 전문이 참고로 포함된다.

[0003] 기술분야

[0004] 본 개시내용의 실시형태는 핵산을 서열결정(sequencing)하는 것에 관한 것이다. 특히, 본 명세서에서 제공되는 방법 및 조성물의 실시형태는 단일-세포 바이설파이트 서열결정 라이브러리의 생성 및 그로부터의 서열 데이터의 수득에 관한 것이다.

배 경 기 술

[0005] 고 세포 계수 단일-세포 서열결정은 전사체, 염색질-접근 가능성 및 돌연변이 차이를 통한 복합 조직 내의 집단의 분리에서 그의 효능을 나타내었다. 추가로, 단일-세포 분석은 세포 분화 궤도가 게놈-특이적 패턴, 예컨대, DNA의 메틸화에서 평가되게 한다. DNA 메틸화는 시토신으로의 공유적 부가이며; 조직 발달에서 활성 변형의 대상인 세포 유형 특이성을 갖는 표시이다. DNA 메틸화는 소듐 바이설파이트 처리의 탈아미노화 화학을 사용하여 염기 쌍 분석에서 프로빙될 수 있다.

[0006] 최근의 연구는 단일 세포 환원 표현 바이설파이트 서열결정(single cell reduced representation bisulfite sequencing: scRRBS) 또는 단일 세포 전체 게놈 바이설파이트 서열결정(single cell whole genome bisulfite sequencing: scWGBS)에서 단일-세포 투입물을 필요로하기까지 바이설파이트 서열결정을 최적화하였다. 그러나, 이들 방법은 단일 세포 반응이 분리되어 수행되는 병행 및 분리된 라이브러리 생성을 통한 단일-세포 데콘볼류션에 의존하여, 확대 가능성이 결여된다. 각각의 세포 서열결정을 위하여 완전히 새로운 세트의 시약이 필요하여, 각각의 추가의 세포에 대하여 비용이 드는 선형 스케일링을 초래한다. DNA의 바이설파이트 전환의 문제로 인하여, 액적- 또는 칩-기반의 미세유체 시스템이 단일 세포 바이설파이트 서열결정을 위해 알맞게 사용되지 않았으며, 대안적인 플랫폼을 사용하는 어떠한 이론적으로 실행 가능한 전략도 존재하지 않는다.

발명의 내용

[0007] 조성물 및 확대 가능한 고-세포 계수, 단일-세포 메틸롬(methylome) 프로파일링 검정이 본 명세서에서 제공된다. 단일-세포 전체 게놈 서열결정(scWGBS)은 본 명세서에서 제공된 단일-세포 조합 인덱싱 전략에 의해 개선되어, 세포가 대량으로 처리될 수 있고, 단일-세포 출력물이 가상 환경에서 역다중화될 수 있게 한다. 일부 실시형태에서, 본 명세서에서 제공된 방법은 트랜스포사제(transposase)-기반의 어댑터 혼입을 이용하며, 이는 기존의 방법에 비하여 증가된 효율과 훨씬 더 높은 정렬 속도를 초래한다. 2개의 서열결정 어댑터 중 하나를 첨부하기 위한 트랜스포사제의 이용은 더 적은 노이즈 판독물과 함께, 훨씬 더 효율적인 라이브러리 구축을 가능하게 하여, 이에 따라, 단일-세포-단일-웰 방법을 사용하는 10 내지 30%에 비하여 약 60%의 정렬 속도(대량 세포 전략과 유사한 속도)를 초래한다. 이는 검정의 서열결정 부분을 위하여 더욱 사용 가능한 서열 판독물 및 현저한 비용 감소를 초래한다. 단일-세포 바이설파이트 서열결정 라이브러리를 생성하기 위한 단일-세포 조합 인덱싱 전략의 이용은 최소의 충돌률과 함께 인간 및 마우스 세포의 믹스에서 입증된다. 또한, 3가지 인간 세포 유형의 믹스의 성공적인 데콘볼루션과, 공개 이용 가능한 데이터를 사용하여 세포 유형 지정을 달성하는 것이 입증된다.

[0008] 정의

[0009] 본 명세서에 사용되는 바와 같이, 용어 "유기체", "대상체"는 상호 교환 가능하게 사용되고, 동물 및 식물을 지칭한다. 동물의 예는 포유동물, 예컨대, 인간이다.

[0010] 본 명세서에 사용되는 바와 같이, 용어 "세포 유형"은 형태, 표현형, 발생 기원 또는 다른 공지되거나 인지 가능한 구별되는 세포 특징에 기초하여 세포를 식별하도록 의도된다. 다양한 상이한 세포 유형이 단일의 유기체로부터(또는 유기체의 동일한 종으로부터) 수득될 수 있다. 예시적인 세포 유형은 방광, 췌장 상피, 췌장 알파, 췌장 베타, 췌장 내피, 골수 림프모구, 골수 B 림프모구, 골수 대식구, 골수 적혈모세포, 골수 수지상, 골수 지방세포, 골수 골세포, 골수 연골세포, 전골수모세포(promyeloblast), 골수 거대핵모세포, 방광, 뇌 B 림프구, 뇌 신경교, 뉴런, 뇌 성상세포, 신경외배엽, 뇌 대식구, 뇌 미세아교세포, 뇌 상피, 피질 뉴런, 뇌 섬유모세포, 유방 상피, 결장 상피, 결장 B 림프구, 유방 상피, 유방 근상피, 유방 섬유모세포, 결장 장세포, 자궁경부

상피, 난소 상피, 난소 섬유모세포, 유관 상피, 혀 상피, 편도 수지상, 편도 B 림프구, 말초 혈액 림프모구, 말초 혈액 T 림프모구, 말초 혈액 피부 T 림프구, 말초 혈액 자연 살해자, 말초 혈액 B 림프모구, 말초 혈액 단핵구, 말초 혈액 골수모세포, 말초 혈액 단핵모세포, 말초 혈액 전골수모세포, 말초 혈액 대식구, 말초 혈액 호염기구, 간 내피, 간 비만(mast), 간 상피, 간 B 림프구, 비장 내피, 비장 상피, 비장 B 림프구, 간 간세포, 간 알렉산더, 간 섬유모세포, 폐 상피, 기관지 상피, 폐 섬유모세포, 폐 B 림프구, 폐 슈반(Schwann), 폐 편평, 폐 대식구, 폐 골모세포, 신경내분비, 폐 폐포, 위 상피, 및 위 섬유모세포를 포함하지만 이들로 제한되지 않는다.

[0011] 본 명세서에 사용되는 바와 같이, 용어 "조직"은 유기체에서 하나 이상의 특정 기능을 수행하도록 함께 작용하는 세포의 집합물 또는 응집물을 의미하도록 의도된다. 세포는 선택적으로 형태적으로 유사할 수 있다. 예시적인 조직은 눈, 근육, 피부, 힘줄, 정맥, 동맥, 혈액, 심장, 비장, 림프절, 골, 골수, 폐, 기관지, 기관, 소화관, 소장, 대장, 결장, 직장, 침샘, 혀, 담낭, 맹장, 간, 췌장, 뇌, 위, 피부, 신장, 요관, 방광, 요도, 생식선, 고환, 난소, 자궁, 나팔관, 흉선, 뇌하수체, 갑상선, 부신, 또는 부갑상선을 포함하지만 이들로 제한되지 않는다. 조직은 인간 또는 다른 유기체의 다양한 기관 중 임의의 것으로부터 유래될 수 있다. 조직은 건강한 조직 또는 건강하지 않은 조직일 수 있다. 건강하지 않은 조직은 비정상적인 메틸화를 갖는 다양한 악성종양, 예를 들어, 폐, 유방, 직장결장, 전립선, 비인두, 위, 고환, 피부, 신경계, 골, 난소, 간, 조혈 조직, 췌장, 자궁, 신장, 림프 조직 등에서의 악성종양을 포함하지만 이들로 제한되지 않는다. 악성종양은 다양한 조직학적 하위유형, 예를 들어, 암종, 선암종, 육종, 섬유선암종, 신경내분비, 또는 미분화성일 수 있다.

[0012] 본 명세서에 사용되는 바와 같이, 용어 "구획"은 어떤 것을 다른 것으로부터 분리 또는 단리하는 영역 또는 부피를 의미하도록 의도된다. 예시적인 구획은, 바이알, 관, 웰, 소적, 볼루스(boluse), 비드, 용기, 표면 특징부, 또는 물리적 힘, 예컨대, 유체 유동, 자성, 전기 전류 등에 의해서 분리된 영역 또는 부피를 포함하지만, 이들로 제한되지 않는다. 일 실시형태에서, 구획은 다중-웰 플레이트, 예컨대, 96- 또는 384-웰 플레이트의 웰이다.

[0013] 본 명세서에 사용되는 바와 같이, "트랜스포좀 복합체"는 통합 효소 및 통합 인식 부위를 포함하는 핵산을 지칭한다. "트랜스포좀 복합체"는 전위 반응을 촉매작용할 수 있는 트랜스포사제 인식 부위 및 트랜스포사제에 의해 형성된 기능적 복합체이다(예를 들어, Gunderson 등의 WO 2016/130704호 참조). 통합 효소의 예는 인테그라제 또는 트랜스포사제를 포함하지만, 이들로 제한되지 않는다. 통합 인식 부위의 예는 트랜스포사제 인식 부위를 포함하지만, 이들로 제한되지 않는다.

[0014] 본 명세서에 사용되는 바와 같이, 용어 "핵산"은 해당 분야에서의 그의 용도와 일치하는 것으로 의도되고, 자연 발생 핵산 또는 그의 기능적 유사체를 포함한다. 특히 유용한 기능적 유사체는 핵산에 서열 특이적 방식으로 혼성화할 수 있거나 또는 특정 뉴클레오티드 서열의 복제를 위한 주형으로서 사용될 수 있다. 자연 발생 핵산은 일반적으로 포스포디에스테르 결합을 함유하는 백본을 갖는다. 유사체 구조는 해당 분야에 공지된 다양한 것들 중 임의의 것을 비롯한 교대 백본 연결을 가질 수 있다. 자연 발생 핵산은 일반적으로 데옥시리보스 당(예를 들어, 데옥시리보핵산(DNA)에서 발견됨) 또는 리보스 당(예를 들어, 리보핵산(RNA)에서 발견됨)을 갖는다. 핵산은 해당 분야에 공지된 이러한 당 모이어티의 다양한 유사체 중 임의의 것을 함유할 수 있다. 핵산은 고유 또는 비-고유 염기를 포함할 수 있다. 이와 관련하여, 고유 데옥시리보핵산은 아데닌, 티민, 시토신 또는 구아닌으로 이루어진 군으로부터 선택되는 하나 이상의 염기를 가질 수 있고, 리보핵산은 우라실, 아데닌, 시토신 또는 구아닌으로 이루어진 군으로부터 선택되는 하나 이상의 염기를 가질 수 있다. 핵산에 포함될 수 있는 유용한 비-고유 염기는 해당 분야에 공지되어 있다. 비-고유 염기의 예에는 잠금 핵산(LNA) 및 가교 핵산(BNA)이 포함된다. LNA 및 BNA 염기는 DNA 올리고뉴클레오티드 내로 혼입될 수 있으며, 올리고뉴클레오티드 혼성화 강도 및 특이성을 증가시킬 수 있다. LNA 및 BNA 염기 및 이러한 염기의 용도는 당업자에게 알려져 있으며, 일상적이다.

[0015] 본 명세서에 사용되는 바와 같이, 용어 "표적"은 핵산에 대한 언급에서 사용되는 경우 본 명세서에 언급된 방법 또는 조성물과 관련하여 핵산에 대한 의미론적인 식별자로서 의도되고, 달리 명확하게 지시하는 것을 넘어서 핵산의 구조 또는 기능을 필수적으로 제한하지 않는다. 표적 핵산은 본질적으로 공지된 또는 미지의 서열의 임의의 핵산일 수 있다. 그것은 예를 들어, 게놈 DNA 또는 cDNA의 단편일 수 있다. 서열결정은 표적 분자의 전체 또는 부분의 서열의 결정을 초래할 수 있다. 표적은 1차 핵산 시료, 예컨대, 핵으로부터 유래될 수 있다. 일 실시형태에서, 표적은 각각의 표적 단편의 말단에서 범용 서열의 배치에 의해 증폭에 적합한 주형으로 가공될 수 있다. 표적은 또한 cDNA로의 역전사에 의해 1차 RNA 시료로부터 수득될 수 있다.

[0016] 본 명세서에 사용되는 바와 같이, 용어 "범용"은 뉴클레오티드 서열을 기술하기 위해 사용되는 경우, 분자가 또

한 서로 상이한 서열의 영역을 갖는 둘 이상의 핵산 분자에 공통인 서열의 영역을 지칭한다. 분자 집합물의 상이한 구성원에 존재하는 범용 서열은 범용 서열, 예를 들어, 범용 포획 서열의 일부에 상보성인 범용 포획 핵산의 집단, 예를 들어, 포획 올리고뉴클레오티드를 사용하여 다수의 상이한 핵산의 포획을 가능하게 할 수 있다. 범용 포획 서열의 비-제한적인 예는 P5 및 P7 프라이머와 동일하거나 이와 상보성인 서열을 포함한다. 유사하게, 분자의 집합물의 상이한 구성원에 존재하는 범용 서열은 범용 서열, 예를 들어, 범용 앵커 서열의 일부에 상보성인 범용 프라이머의 집단을 사용하여 다수의 상이한 핵산의 복제 또는 증폭을 가능하게 할 수 있다. 이에 따라, 포획 올리고뉴클레오티드 또는 범용 프라이머는 범용 서열에 특이적으로 혼성화할 수 있는 서열을 포함한다.

[0017] 증폭 프라이머, 예를 들어, 포획 올리고뉴클레오티드를 지칭하는 경우, 용어 "P5" 및 "P7"이 사용될 수 있다. 용어 "P5'"(P5 프라임) 및 "P7'"(P7 프라임)은 각각 P5 및 P7의 상보물을 지칭한다. 임의의 적합한 증폭 프라이머가 본 명세서에 제시된 방법에서 사용될 수 있으며, P5 및 P7의 이용이 단지 예시적인 실시형태인 것이 이해될 것이다. 유동 셀 상에서의 증폭 프라이머, 예컨대 P5 및 P7의 이용은 WO 2007/010251호, WO 2006/064199호, WO 2005/065814호, WO 2015/106941호, WO 1998/044151호 및 WO 2000/018957호의 개시내용에 의해 예시된 바와 같이 해당 분야에 공지되어 있다. 예를 들어, 임의의 적합한 정방향 증폭 프라이머는 고정화되든지 용액 중이든지, 상보성 서열로의 혼성화 및 서열의 증폭을 위하여 본 명세서에 제시된 방법에서 유용할 수 있다. 유사하게, 임의의 적합한 역방향 증폭 프라이머는 고정화되든지 용액 중이든지, 상보성 서열로의 혼성화 및 서열의 증폭을 위하여 본 명세서에 제시된 방법에서 유용할 수 있다. 당업자는 본 명세서에 제시된 바와 같은 핵산의 포획 및/또는 증폭에 적합한 프라이머 서열을 설계하고 이용하는 방법을 이해할 것이다.

[0018] 본 명세서에 사용되는 바와 같이, 용어 "프라이머" 및 그의 파생어는 일반적으로 관심 표적 서열에 혼성화할 수 있는 임의의 핵산을 지칭한다. 전형적으로, 프라이머는 뉴클레오티드가 중합효소에 의해 중합될 수 있는 기질로서 기능하지만; 일부 실시형태에서, 프라이머는 합성된 핵산 가닥 내에 혼입되어, 또 다른 프라이머가 혼성화할 수 있는 부위를 제공하여, 합성된 핵산 분자에 상보성인 새로운 가닥의 합성을 프라이밍할 수 있다. 프라이머는 뉴클레오티드 또는 그의 유사체의 임의의 조합을 포함할 수 있다. 일부 실시형태에서, 프라이머는 단일-가닥 올리고뉴클레오티드 또는 폴리뉴클레오티드이다. 용어 "폴리뉴클레오티드" 및 "올리고뉴클레오티드"는 임의의 길이의 뉴클레오티드의 중합체 형태를 지칭하도록 본 명세서에서 상호 교환 가능하게 사용되고, 리보뉴클레오티드, 데옥시리보뉴클레오티드, 그의 유사체, 또는 그의 혼합물을 포함할 수 있다. 이 용어는 등가물로서, 뉴클레오티드 유사체로부터 제조된 DNA 또는 RNA의 유사체를 포함하도록 이해되어야 하고, 단일 가닥(예컨대, 센스 또는 안티센스) 및 이중 가닥 폴리뉴클레오티드에 적용 가능하도록 이해되어야 한다. 이 용어는 본 명세서에 사용되는 바와 같이 또한 예를 들어, 역전사효소의 작용에 의해 RNA 주형으로부터 생성된 DNA에 상보성이거나 또는 이를 카피하는, cDNA를 포함한다. 이 용어는 분자의 1차 구조 만을 지칭한다. 따라서, 이 용어는 삼중-, 이중- 및 단일-가닥 데옥시리보핵산("DNA"), 뿐만 아니라 삼중-, 이중- 및 단일-가닥 리보핵산("RNA")을 포함한다.

[0019] 본 명세서에 사용되는 바와 같이, 용어 "어댑터" 및 그의 파생어, 예를 들어, 범용 어댑터는 일반적으로 본 개시내용의 핵산 분자에 라이게이션될 수 있는 임의의 선형 올리고뉴클레오티드를 지칭한다. 일부 실시형태에서, 어댑터는 시료 중에 존재하는 임의의 표적 서열의 3' 말단 또는 5' 말단과 실질적으로 비-상보성이다. 일부 실시형태에서, 적합한 어댑터 길이는 약 10 내지 100개 뉴클레오티드, 약 12 내지 60개 뉴클레오티드 및 약 15 내지 50개 뉴클레오티드의 길이의 범위이다. 일반적으로, 어댑터는 뉴클레오티드 및/또는 핵산의 임의의 조합을 포함할 수 있다. 일부 양태에서, 어댑터는 하나 이상의 위치에 하나 이상의 절단 가능한 기를 포함할 수 있다. 또 다른 양태에서, 어댑터는 프라이머, 예를 들어, 범용 프라이머의 적어도 일부와 실질적으로 동일하거나 또는 실질적으로 상보성인 서열을 포함할 수 있다. 일부 실시형태에서, 어댑터는 하류 오차 보정, 식별 또는 서열결정을 보조하도록 바코드 또는 태그를 포함할 수 있다. 용어 "어댑터(adaptor)" 및 "어댑터(adapter)"는 상호 교환 가능하게 사용된다.

[0020] 본 명세서에 사용되는 바와 같이, 용어 "각각"은 물품의 집합물에 대한 언급에 사용되는 경우, 집합물 내의 개별 물품을 식별하도록 의도되지만, 문맥에서 명백하게 다르게 언급되지 않는 한, 반드시 집합물 내의 모든 물품을 지칭하는 것은 아니다.

[0021] 본 명세서에 사용되는 바와 같이, 용어 "수송"은 유체를 통한 분자의 이동을 지칭한다. 이 용어는 수동 수송, 예컨대, 그들의 농도 구배(예를 들어, 수동 확산)에 따른 분자의 이동을 포함할 수 있다. 이 용어는 또한 능동 수송을 포함할 수 있고, 이에 의해 분자는 그들의 농도 구배를 따라 또는 그들의 농도 구배에 대하여 이동할 수 있다. 따라서, 수송은 에너지를 적용하여 목적하는 방향으로 또는 목적하는 위치, 예컨대, 증폭 부위로 하나 이

상의 분자를 이동시키는 것을 포함할 수 있다.

[0022] 본 명세서에 사용되는 바와 같이, "증폭시키다", "증폭시키는" 또는 "증폭 반응" 및 그들의 파생어는 일반적으로 핵산 분자의 적어도 일부가 적어도 하나의 추가의 핵산 분자로 복제되거나 또는 카피되는 임의의 작용 또는 과정을 지칭한다. 추가의 핵산 분자는 주형 핵산 분자의 적어도 일부 부분과 실질적으로 동일하거나 또는 실질적으로 상보성인 서열을 선택적으로 포함한다. 주형 핵산 분자는 단일-가닥 또는 이중-가닥일 수 있고, 추가의 핵산 분자는 독립적으로 단일-가닥 또는 이중-가닥일 수 있다. 증폭은 핵산 분자의 선형 또는 지수적 복제를 선택적으로 포함한다. 일부 실시형태에서, 이러한 증폭은 등온 조건을 사용하여 수행될 수 있고; 다른 실시형태에서, 이러한 증폭은 열순환을 포함할 수 있다. 일부 실시형태에서, 증폭은 단일 증폭 반응에서 복수의 표적 서열의 동시의 증폭을 포함하는 멀티플렉스 증폭이다. 일부 실시형태에서, "증폭"은 단독으로 또는 조합하여, DNA 및 RNA 기반 핵산의 적어도 일부 부분의 증폭을 포함한다. 증폭 반응은 당업자에게 공지된 증폭 과정 중 임의의 것을 포함할 수 있다. 일부 실시형태에서, 증폭 반응은 중합효소 연쇄 반응(PCR)을 포함한다.

[0023] 본 명세서에 사용되는 바와 같이, "증폭 조건" 및 그의 파생어는 일반적으로 하나 이상의 핵산 서열을 증폭시키기에 적합한 조건을 지칭한다. 이러한 증폭은 선형이거나 또는 지수적일 수 있다. 일부 실시형태에서, 증폭 조건은 등온 조건을 포함할 수 있거나, 대안적으로 열순환 조건, 또는 등온 조건과 열순환 조건의 조합을 포함할 수 있다. 일부 실시형태에서, 하나 이상의 핵산 서열을 증폭시키기에 적합한 조건은 중합효소 연쇄 반응(PCR) 조건을 포함한다. 전형적으로, 증폭 조건은 핵산, 예컨대, 하나 이상의 표적 서열을 증폭시키거나, 하나 이상의 어댑터에 라이게이션된 증폭된 표적 서열, 예를 들어, 어댑터-라이게이션된 증폭된 표적 서열을 증폭시키기에 충분한 반응 혼합물을 지칭한다. 일반적으로, 증폭 조건은 증폭 또는 핵산 합성을 위한 촉매, 예를 들어, 중합효소; 증폭될 핵산에 대하여 어느 정도의 상보성을 보유하는 프라이머; 및 뉴클레오티드, 예컨대, 핵산에 혼성화된 후 프라이머의 신장을 촉진시키기 위한 데옥시리보뉴클레오티드 트리포스페이트(dNTP)를 포함한다. 증폭 조건은 핵산으로의 프라이머의 혼성화 또는 어닐링, 프라이머의 신장 및 신장된 프라이머가 증폭 중인 핵산 서열로부터 분리되는 변성 단계가 요구될 수 있다. 필수적이지는 않지만, 전형적으로, 증폭 조건은 열순환을 포함할 수 있고; 일부 실시형태에서, 증폭 조건은 어닐링, 신장 및 분리 단계가 반복되는 복수의 사이클을 포함한다. 전형적으로, 증폭 조건은 양이온, 예컨대, Mg^{2+} 또는 Mn^{2+}를 포함하고, 다양한 이온 강도의 개질제를 또한 포함할 수 있다.

[0024] 본 명세서에 사용되는 바와 같이, "재-증폭" 및 그들의 파생어는 일반적으로 증폭된 핵산 분자의 적어도 일부가 임의의 적합한 증폭 과정(일부 실시형태에서 "2차" 증폭으로 지칭됨)을 통해 추가로 증폭되어, 재증폭된 핵산 분자를 생성하는 임의의 과정을 지칭한다. 2차 증폭은 증폭된 핵산 분자가 생성되었던 원래의 증폭 과정과 동일할 필요도 없고; 재증폭된 핵산 분자가 증폭된 핵산 분자와 완전히 동일하거나 완전히 상보성일 필요도 없으며; 필요한 것은 재증폭된 핵산 분자가 증폭된 핵산 분자 또는 그의 상보물의 적어도 일부를 포함하는 것이다. 예를 들어, 재-증폭은 1차 증폭과 상이한 표적-특이적 프라이머를 포함하는 상이한 프라이머 및/또는 상이한 증폭 조건의 사용을 포함할 수 있다.

[0025] 본 명세서에 사용되는 바와 같이, 용어 "중합효소 연쇄 반응"("PCR")은 Mullis의 미국 특허 제4,683,195호 및 제4,683,202호의 방법을 지칭하는데, 이것은 클로닝 또는 정제 없이 게놈 DNA의 혼합물 중에서 관심 폴리뉴클레오티드의 세그먼트의 농도를 증가시키는 방법을 기술한다. 관심 폴리뉴클레오티드를 증폭시키는 이러한 과정은 매우 과량의 2가지 올리고뉴클레오티드 프라이머를 요망되는 관심 폴리뉴클레오티드를 함유하는 DNA 혼합물에 도입하는 단계에 이어서, DNA 중합효소의 존재 하에서의 일련의 열순환 단계로 이루어진다. 2가지 프라이머는 관심 이중 가닥 폴리뉴클레오티드의 그들의 각각의 가닥에 상보성이다. 혼합물은 먼저 더 높은 온도에서 변성된 다음, 프라이머는 관심 분자의 폴리뉴클레오티드 내의 상보성 서열에 어닐링된다. 어닐링 후, 프라이머는 중합효소를 사용하여 신장되어 새로운 상보성 가닥의 쌍을 형성한다. 변성, 프라이머 어닐링 및 중합효소 신장 단계를 수회 반복하여(열순환으로 지칭됨) 고농도의 요망되는 관심 폴리뉴클레오티드의 증폭된 세그먼트를 수득할 수 있다. 요망되는 관심 폴리뉴클레오티드(앰플리콘)의 증폭된 세그먼트의 길이는 서로에 대한 프라이머의 상대적인 위치에 의해 결정되고, 따라서 이러한 길이는 제어 가능한 파라미터이다. 이러한 과정의 반복으로 인해서, 이 방법은 "중합효소 연쇄 반응"(이하 "PCR")으로 지칭된다. 관심 폴리뉴클레오티드의 요망되는 증폭된 세그먼트가 혼합물 중에서 우세한 핵산 서열(농도 면에서)이 되기 때문에, 그들은 "PCR 증폭된" 것으로 칭해진다. 상기에 논의된 방법에 대한 변형에서, 표적 핵산 분자를 복수의 상이한 프라이머 쌍, 일부 경우에, 관심 표적 핵산 분자당 하나 이상의 프라이머쌍을 사용하여 PCR 증폭시켜, 멀티플렉스 PCR 반응을 형성할 수 있다.

[0026] 본 명세서에 정의된 바와 같이 "멀티플렉스 증폭"은 적어도 하나의 표적-특이적 프라이머를 사용한 시료 내의 2

개 이상의 표적 서열의 선택적인 비-무작위 증폭을 지칭한다. 일부 실시형태에서, 멀티플렉스 증폭은 단일 반응 용기 내에서 표적 서열의 일부 또는 전부가 증폭되도록 수행된다. 주어진 멀티플렉스 증폭의 "플렉시" 또는 "플렉스"는 일반적으로 단일의 멀티플렉스 증폭 동안 증폭되는 상이한 표적-특이적 서열의 수를 지칭한다. 일부 실시형태에서, 플렉시는 약 12-플렉스, 24-플렉스, 48-플렉스, 96-플렉스, 192-플렉스, 384-플렉스, 768-플렉스, 1536-플렉스, 3072-플렉스, 6144-플렉스 또는 그 초과일 수 있다. 몇몇 상이한 방법(예를 들어, 젤 전기영동 이후의 농도계, 생물분석기 또는 정량적 PCR을 사용한 정량화, 표지된 프로브와의 혼성화; 비오티닐화된 프라이머의 혼입 이후의 아비딘-효소 컨쥬게이트 검출; 증폭된 표적 서열 내로의 ^{32}P-표지된 데옥시뉴클레오티드 트리포스페이트의 혼입)에 의해 증폭된 표적 서열을 검출하는 것이 또한 가능하다.

[0027] 본 명세서에 사용되는 바와 같이, "증폭된 표적 서열" 및 그의 파생어는 일반적으로 표적-특이적 프라이머 및 본 명세서에서 제공된 방법을 사용하여 표적 서열을 증폭시킴으로써 생성된 핵산 서열을 지칭한다. 증폭된 표적 서열은 표적 서열에 대하여 동일한 센스(즉, 양성 가닥) 또는 안티센스(즉, 음성 가닥) 중 어느 하나일 수 있다.

[0028] 본 명세서에 사용되는 바와 같이, 용어 "라이게이션시키는", "라이게이션" 및 그들의 파생어는 일반적으로 2개 이상의 분자를 함께 공유적으로 연결시키는 과정, 예를 들어, 2개 이상의 핵산 분자를 서로 공유적으로 연결시키는 과정을 지칭한다. 일부 실시형태에서, 라이게이션은 핵산의 인접 뉴클레오티드 간의 닉을 연결하는 것을 포함한다. 일부 실시형태에서, 라이게이션은 제1 핵산 분자의 말단과 제2 핵산 분자의 말단 사이에 공유 결합을 형성하는 것을 포함한다. 일부 실시형태에서, 라이게이션은 하나의 핵산의 5' 포스페이트기와 제2 핵산의 3' 하이드록실기 사이에 공유 결합을 형성하여, 라이게이션된 핵산 분자를 형성하는 것을 포함할 수 있다. 일반적으로 본 개시내용의 목적을 위하여, 증폭된 표적 서열을 어댑터에 라이게이션시켜, 어댑터-라이게이션된 증폭된 표적 서열을 생성할 수 있다.

[0029] 본 명세서에 사용되는 바와 같이, "리가제" 및 그의 파생어는 일반적으로 2개의 기질 분자의 라이게이션을 촉매작용시킬 수 있는 임의의 작용제를 지칭한다. 일부 실시형태에서, 리가제는 핵산의 인접 뉴클레오티드 사이의 닉의 연결을 촉매작용시킬 수 있는 효소를 포함한다. 일부 실시형태에서, 리가제는 하나의 핵산 분자의 5' 포스페이트와 또 다른 핵산 분자의 3' 하이드록실 사이에 공유 결합의 형성을 촉매작용시켜, 라이게이션된 핵산 분자를 형성할 수 있는 효소를 포함한다. 적합한 리가제는 T4 DNA 리가제, T4 RNA 리가제 및 에스케리키아 콜라이(*E. coli*) DNA 리가제를 포함할 수 있지만, 이에 제한되지 않는다.

[0030] 본 명세서에 사용되는 바와 같이, "라이게이션 조건" 및 그의 파생어는 일반적으로 2개의 분자를 서로 라이게이션시키기에 적합한 조건을 지칭한다. 일부 실시형태에서, 라이게이션 조건은 핵산 사이의 닉 또는 갭을 시일링(sealing)하기에 적합하다. 본 명세서에 사용되는 바와 같이, 용어 닉 또는 갭은 해당 분야에서의 그 용어의 사용과 일치한다. 전형적으로, 닉 또는 갭은 효소, 예컨대, 리가제의 존재 하에서 적절한 온도 및 pH에서 라이게이션될 수 있다. 일부 실시형태에서, T4 DNA 리가제는 약 70 내지 72℃의 온도에서 핵산 사이의 닉을 연결할 수 있다.

[0031] 본 명세서에 사용되는 바와 같이, 용어 "유동 셀"은 하나 이상의 유체 시약이 유동될 수 있는 고체 표면을 포함하는 챔버를 지칭한다. 본 개시내용의 방법에서 용이하게 사용될 수 있는 유동 셀 및 관련 유체 시스템 및 검출 플랫폼의 예는 예를 들어, 문헌[Bentley et al., *Nature* 456:53-59 (2008)], WO 04/018497호; US 7,057,026호; WO 91/06678호; WO 07/123744호; US 7,329,492호; US 7,211,414호; US 7,315,019호; US 7,405,281호 및 US 2008/0108082호에 기술되어 있으며, 이의 각각은 본 명세서에 참조로 포함된다.

[0032] 본 명세서에 사용되는 바와 같이, 핵산에 대한 언급에서 사용되는 경우 용어 "앰플리콘"은 핵산의 카피 생성물을 의미하고, 여기서 생성물은 핵산의 뉴클레오티드 서열의 적어도 일부와 동일하거나 그에 상보성인 뉴클레오티드 서열을 갖는다. 앰플리콘은 예를 들어, 중합효소 연장, 중합효소 연쇄 반응(PCR), 회전환 증폭(RCA), 라이게이션 신장, 또는 라이게이션 연쇄 반응을 비롯한, 주형으로서 핵산 또는 그의 앰플리콘을 사용하는 임의의 다양한 증폭 방법에 의해 생성될 수 있다. 앰플리콘은 특정 뉴클레오티드 서열의 단일의 카피(예를 들어, PCR 생성물) 또는 뉴클레오티드 서열의 다수의 카피(예를 들어, RCA의 콘카타머 생성물(concatameric product))를 갖는 핵산 분자일 수 있다. 표적 핵산의 제1 앰플리콘은 전형적으로 상보성 카피이다. 후속 앰플리콘은 제1 앰플리콘의 생성 이후에, 표적 핵산으로부터 또는 제1 앰플리콘으로부터 생성된 카피이다. 후속 앰플리콘은 표적 핵산에 실질적으로 상보성이거나 표적 핵산에 실질적으로 동일한 서열을 가질 수 있다.

[0033] 본 명세서에 사용되는 바와 같이, 용어 "증폭 부위"는 하나 이상의 앰플리콘이 생성될 수 있는 어레이 내의 또는 어레이 상의 부위를 지칭한다. 증폭 부위는 그 부위에서 생성되는 적어도 하나의 앰플리콘을 함유하거나, 보

유하거나 또는 그에 부착되도록 추가로 구성될 수 있다.

[0034] 본 명세서에 사용되는 바와 같이, 용어 "어레이"는 상대적인 위치에 따라 서로 구별될 수 있는 부위의 집단을 지칭한다. 어레이의 상이한 부위에 존재하는 상이한 분자는 그 어레이 내의 부위의 위치에 따라 서로 구별될 수 있다. 어레이의 개별 부위는 특정 유형의 하나 이상의 분자를 포함할 수 있다. 예를 들어, 부위는 특정 서열을 갖는 단일 표적 핵산 분자를 포함할 수 있거나, 부위는 동일한 서열(및/또는 그의 상보성 서열)을 갖는 몇몇 핵산 분자를 포함할 수 있다. 어레이의 부위는 동일한 기판, 즉, 기재(substrate) 상에 위치된 상이한 특징부일 수 있다. 예시적인 특징부는 기판 내의 웰, 기판 내 또는 기판 상의 비드(또는 다른 입자), 기판으로부터의 돌출부, 기판 상의 릿지(ridge) 또는 기판 내의 채널을 비제한적으로 포함한다. 어레이의 부위는 각각 상이한 분자를 보유하는 개별 기판일 수 있다. 개별 기판에 부착되는 상이한 분자는, 기판이 회합된 표면 상의 기판의 위치에 따라 또는 액체 또는 젤 중의 기판의 위치에 따라 식별될 수 있다. 개별 기판이 표면 상에 위치된 예시적인 어레이는 비제한적으로 웰 내에 비드를 갖는 것을 포함한다.

[0035] 본 명세서에 사용되는 바와 같이, 용어 "용량"은 부위 및 핵산 물질에 대한 언급에서 사용되는 경우 그 부위를 점유할 수 있는 핵산 물질의 최대양을 의미한다. 예를 들어, 이 용어는 특정 조건에서 그 부위를 점유할 수 있는 핵산 분자의 총 수를 지칭할 수 있다. 예를 들어, 핵산 물질의 총 질량 또는 특정 조건에서 그 부위를 점유할 수 있는 특정 뉴클레오티드 서열의 카피의 총수를 비롯한 다른 측정치가 마찬가지로 사용될 수 있다. 전형적으로, 표적 핵산에 대한 부위의 용량은 표적 핵산의 앰플리콘에 대한 부위의 용량과 실질적으로 동등할 것이다.

[0036] 본 명세서에 사용되는 바와 같이, 용어 "포획제"는 표적 분자(예를 들어, 표적 핵산)에 부착, 보유 또는 결합할 수 있는 물질, 화학물질, 분자 또는 그의 모이어티를 지칭한다. 예시적인 포획제는 비제한적으로 표적 핵산의 적어도 일부에 상보성인 포획 핵산(본 명세서에 포획 올리고뉴클레오티드로도 지칭됨), 표적 핵산(또는 그에 부착된 연결 모이어티)에 결합할 수 있는 수용체-리간드 결합쌍의 구성원(예를 들어, 아비딘, 스트렙트아비딘, 비오틴, 렉틴, 탄수화물, 핵산 결합 단백질, 에피토프, 항체 등), 또는 표적 핵산(또는 그에 부착된 연결 모이어티)과 공유 결합을 형성할 수 있는 화학 시약을 포함한다.

[0037] 본 명세서에 사용되는 바와 같이, 용어 "클론 집단"은 특정 뉴클레오티드 서열에 대하여 균질한 핵산의 집단을 지칭한다. 균질한 서열은 전형적으로 적어도 10개의 뉴클레오티드 길이이지만, 예를 들어, 적어도 50, 100, 250, 500 또는 1000개의 뉴클레오티드 길이를 비롯하여, 훨씬 더 길 수 있다. 클론 집단은 단일 표적 핵산 또는 주형 핵산으로부터 유래될 수 있다. 전형적으로, 클론성 집단 내의 모든 핵산이 동일한 뉴클레오티드 서열을 가질 것이다. 소수의 돌연변이(예를 들어, 증폭 인공물로 인함)가 클론성에서 벗어나지 않으면서 클론 집단에서 일어날 수 있는 것이 이해될 것이다.

[0038] 본 명세서에 사용되는 바와 같이, 조성물, 용품, 핵산 또는 핵과 관련하여 "제공하는"은 조성물, 용품, 핵산 또는 핵을 제조하거나, 조성물, 용품, 핵산 또는 핵을 구매하거나, 또는 다르게 화합물, 조성물, 용품 또는 핵을 수득하는 것을 의미한다.

[0039] 용어 "및/또는"은 열거된 요소 중 하나 또는 전부 또는 열거된 요소 중 임의의 둘 이상의 조합을 의미한다.

[0040] 단어 "바람직한" 및 "바람직하게는"은 특정 상황 하에서 특정 이익을 제공할 수 있는 본 발명의 실시형태를 지칭한다. 그러나, 다른 실시형태가 동일하거나 다른 상황 하에서 또한 바람직할 수 있다. 추가로, 하나 이상의 바람직한 실시형태의 언급은 다른 실시형태가 유용하지 않다는 것을 암시하지 않고, 본 발명의 범주로부터 다른 실시형태를 배제하도록 의도되지 않는다.

[0041] 용어 "포함한다" 및 그의 변형은 이들 용어가 설명 및 청구범위에 존재하는 경우 제한적인 의미를 갖지 않는다.

[0042] 실시형태가 용어 "포함하다", "포함한다", 또는 "포함하는" 등과 함께 본 명세서에 기재되는 경우에는 언제나, "이루어진" 및/또는 "본질적으로 이루어진"과 관련하여 기재된 다른 유사한 실시형태도 또한 제공되는 것이 이해된다.

[0043] 달리 명시되지 않는 한, 단수 표현 및 "적어도 하나"는 상호 교환 가능하게 사용되고, 하나 또는 하나 초과를 의미한다.

[0044] 또한 본 명세서에서, 종점에 의한 수치 범위의 언급은 그 범위 내에 포함되는 모든 수를 포함한다(예를 들어, 1 내지 5는 1, 1.5, 2, 2.75, 3, 3.80, 4, 5 등을 포함한다).

[0045] 별개의 단계를 포함하는 본 명세서에 개시된 임의의 방법의 경우, 단계는 임의의 실행 가능한 순서로 행해질 수

있다. 그리고, 적절한 경우, 2개 이상의 단계의 임의의 조합이 동시에 행해질 수 있다.

[0046] 본 명세서 전체에서 "일 실시형태", "실시형태", "특정 실시형태", 또는 "일부 실시형태" 등에 대한 언급은 그 실시형태와 관련하여 기재된 특정 특징부, 구성, 조성물 또는 특징이 본 개시내용의 적어도 하나의 실시형태에 포함됨을 의미한다. 따라서, 본 명세서 전체의 다양한 위치에서 이러한 어구의 존재는 본 개시내용의 동일한 실시형태를 반드시 지칭하는 것은 아니다. 추가로, 특정 특징부, 구성, 조성물 또는 특징이 하나 이상의 실시형태에서 임의의 적합한 방식으로 조합될 수 있다.

도면의 간단한 설명

[0047] 본 개시내용의 예시적인 실시형태의 하기의 상세한 설명은 하기의 도면과 함께 읽는 경우 최적으로 이해될 수 있다.

도 1은 본 개시내용에 따른 단일-세포 조합 인덱싱을 위한 일반적인 예시적 방법의 일반적인 블록 다이어그램을 보여준다.

도 2는 도 1에 일반적으로 예시된 단일-세포 조합 인덱싱을 위한 방법의 일 실시형태의 개략도를 보여준다.

도 3은 선형 증폭 후의 단편-어댑터 분자의 예시적인 실시형태의 개략도를 보여준다.

도 4는 범용 어댑터의 부가 후의 단편-어댑터 분자의 예시적인 실시형태의 개략도를 보여준다.

개략도는 반드시 일정한 비율로 도시된 것은 아니다. 도면에서 사용된 유사한 번호는 유사한 성분, 단계 등을 지칭한다. 그러나, 주어진 도면에서 성분을 지칭하기 위한 번호의 사용은 동일한 번호로 표지된 또 다른 도면에서의 성분을 제한하도록 의도되지 않는다는 것이 이해될 것이다. 또한, 성분을 지칭하기 위한 상이한 번호의 사용은 상이한 번호의 성분이 다른 번호의 성분과 동일하거나 유사할 수 없다는 것을 나타내도록 의도되지 않는다.

발명을 실시하기 위한 구체적인 내용

[0048] 본 명세서에서 제공된 방법은 복수의 세포로부터 단리된 핵을 제공하는 단계를 포함한다(도 1, 블록 12). 세포는 임의의 유기체(들)로부터, 그리고 유기체(들)의 임의의 세포 유형 또는 임의의 조직으로부터의 것일 수 있다. 방법은 세포를 해리시키는 단계(도 2, 블록 i), 및/또는 핵을 단리하는 단계(도 2, 블록 ii)를 추가로 포함할 수 있다. 세포로부터 핵을 단리하는 방법은 당업자에게 공지되어 있고, 일상적이다. 핵의 수는 적어도 2개일 수 있다. 상한은 본 명세서에 기재된 바와 같은 방법의 다른 단계에서 사용되는 장비(예를 들어, 다중-웰 플레이트)의 실시적 제한에 좌우된다. 예를 들어, 일 실시형태에서 핵의 수는 1,000,000,000개 이하, 100,000,000개 이하, 10,000,000개 이하, 1,000,000개 이하, 10,000개 이하, 또는 1,000개 이하일 수 있다. 당업자는 각각의 핵 내의 핵산 분자가 유기체의 전체 유전적 상보물을 나타내고, 인트론 서열 및 엑손 서열 둘 모두뿐만 아니라 비-코딩 조절 서열, 예컨대, 프로모터 및 인핸서 서열을 포함하는 게놈 DNA 분자라는 것을 인식할 것이다.

[0049] 일 실시형태에서, 핵은 게놈 DNA에 결합된 뉴클레오솜을 포함한다. 이러한 핵은 세포의 전체 게놈의 DNA 서열을 결정하지 않는 방법, 예컨대 sciATAC-seq에서 유용할 수 있다. 또 다른 실시형태에서, 단리된 핵은 뉴클레오솜의 핵을 고갈시키는 조건을 겪어, 뉴클레오솜-고갈된 핵을 생성한다(도 1, 블록 13, 및 도 2, 블록 ii). 이러한 핵은 세포의 전체 게놈 DNA 서열을 결정하는 것을 목적으로 하는 방법에서 유용할 수 있다. 일 실시형태에서, 뉴클레오솜-고갈을 위해 사용되는 조건은 단리된 핵의 온전성을 유지시킨다. 뉴클레오솜 고갈된 핵의 생성 방법은 당업자에게 알려져 있다(예를 들어, 문헌[Vitak et al., 2017, Nature Methods, 14(3):302-308] 참조). 일 실시형태에서, 조건은 핵산-단백질 상호작용을 파괴할 수 있는 카오트로픽제(chaotropic agent)로의 처리를 포함하는 화학적 처리이다. 유용한 카오트로픽제의 일 예에는 리튬 다이아이오도살리실레이트가 포함되지만 이로 제한되지 않는다. 또 다른 실시형태에서, 조건은 핵산-단백질 상호작용을 방해할 수 있는 세제로의 처리를 포함하는 화학 처리이다. 유용한 세제의 예에는 소듐 도데실 설페이트(SDS)가 포함되지만 이로 제한되지 않는다. 일부 실시형태에서, 세제, 예컨대 SDS가 사용되는 경우, 핵이 단리되는 세포를 단리 이전에 가교결합제로 처리한다. 가교결합제의 유용한 예에는 포름알데히드가 포함되지만, 이로 제한되지 않는다.

[0050] 본 명세서에서 제공된 방법은 핵, 예컨대 뉴클레오솜-고갈된 핵의 하위세트를 제1 복수 구획 내로 분배하는 단계(도 1, 블록 14, 도 2, 좌측 개략도)를 포함한다. 하위세트 내에, 이에 따라, 각 구획 내에 존재하는 핵의 수는 적어도 1개일 수 있다. 일 실시형태에서, 하위세트 내에 존재하는 핵의 수는 2,000개 이하이다. 핵을 하위세트 내에 분배하는 방법은 당업자에게 공지되어 있고, 일상적이다. 예에는 형광-활성화된 핵 분류(FANS)가 포함

되지만, 이로 제한되지 않는다.

[0051] 각각의 구획은 트랜스포좀 복합체를 포함한다. 트랜스포좀 복합체, 즉, 트랜스포사제 인식 부위에 결합된 트랜스포사제는 때로는 "태그먼트화(tagmentation)"라 지칭되는 과정에서 트랜스포사제 인식 부위를 핵 내의 표적 핵산 내로 삽입할 수 있다. 일부 이러한 삽입 사례에서, 트랜스포사제 인식 부위의 하나의 가닥은 표적 핵산 내로 전달될 수 있다. 이러한 가닥은 "전달된 가닥"이라 지칭된다. 일 실시형태에서, 트랜스포좀 복합체는 2개의 하위유닛을 갖는 이량체 트랜스포사제, 및 2개의 비-인접 트랜스포존 서열을 포함한다. 또 다른 실시형태에서, 트랜스포사제는 2개의 하위유닛을 갖는 이량체 트랜스포사제, 및 인접한 트랜스포존 서열을 포함한다.

[0052] 일부 실시형태는 과다활성 Tn5 트랜스포사제 및 Tn5-형 트랜스포사제 인식 부위(문헌[Goryshin and Reznikoff, *J. Biol. Chem.*, 273:7367 (1998)]) 또는 MuA 트랜스포사제 및 R1 및 R2 말단 서열을 포함하는 Mu 트랜스포사제 인식 부위(문헌[Mizuuchi, K., *Cell*, 35: 785, 1983]; 문헌[Savilahti, H, *et al., EMBO J.*, 14: 4893, 1995])의 사용을 포함할 수 있다. Tn5 모자이크 말단(Mosaic End: ME) 서열이 또한 당업자에 의해 최적화된 바와 같이 사용될 수 있다.

[0053] 본 명세서에서 제공된 조성물 및 방법의 특정 실시형태와 함께 사용될 수 있는 전위 시스템의 추가 예는 스타필로코커스 아우레우스(*Staphylococcus aureus*) Tn552(문헌[Colegio *et al., J. Bacteriol.,* 183: 2384-8, 2001]; 문헌[Kirby C *et al., Mol. Microbiol.,* 43: 173-86, 2002]), Ty1(문헌[Devine & Boeke, *Nucleic Acids Res.,* 22: 3765-72, 1994] 및 국제 공개 WO 95/23875호), 트랜스포존 Tn7(문헌[Craig, N L, *Science.* 271: 1512, 1996]; 문헌[Craig, N L, Review in: *Curr Top Microbiol Immunol.,* 204:27-48, 1996]), Tn/O 및 IS10(문헌[Kleckner N, *et al., Curr Top Microbiol Immunol.,* 204:49-82, 1996]), 마리너(Mariner) 트랜스포사제(문헌[Lampe D J, *et al., EMBO J.,* 15: 5470-9, 1996]), Tc1(문헌[Plasterk R H, *Curr. Topics Microbiol. Immunol.,* 204: 125-43, 1996]), P 요소(문헌[Gloor, G B, *Methods Mol. Biol.,* 260: 97-114, 2004]), Tn3(문헌[Ichikawa & Ohtsubo, *J Biol. Chem.* 265:18829-32, 1990]), 박테리아 삽입 서열(문헌[Ohtsubo & Sekine, *Curr. Top. Microbiol. Immunol.* 204: 1-26, 1996]), 레트로바이러스(문헌[Brown, *et al., Proc Natl Acad Sci USA*, 86:2525-9, 1989]) 및 효모의 레트로트랜스포존(문헌[Boeke & Corces, *Annu Rev Microbiol.* 43:403-34, 1989])을 포함한다. 추가 예는 IS5, Tn10, Tn903, IS911, 및 트랜스포사제 패밀리 효소의 엔지니어링된 버전을 포함한다(문헌[Zhang *et al.,* (2009) *PLoS Genet.* 5:e1000689. Epub 2009 Oct 16]; 문헌[Wilson C. et al (2007) *J. Microbiol. Methods* 71:332-5]).

[0054] 본 명세서에서 제공된 방법 및 조성물과 함께 사용될 수 있는 인테그라제의 다른 예는 레트로바이러스 인테그라제 및 이러한 레트로바이러스 인테그라제, 예컨대, HIV-1, HIV-2, SIV, PFV-1, RSV로부터의 인테그라제에 대한 인테그라제 인식 서열를 포함한다.

[0055] 본 명세서에 기재된 방법 및 조성물에서 유용한 트랜스포존 서열은 미국 특허 출원 공개 제2012/0208705호, 미국 특허 출원 공개 제2012/0208724호 및 국제 특허 출원 공개 제WO 2012/061832호에 제공되어 있다. 일부 실시형태에서, 트랜스포존 서열은 제1 트랜스포사제 인식 부위, 제2 트랜스포사제 인식 부위, 및 2개의 트랜스포사제 인식 부위 사이에 존재하는 인덱스를 포함한다.

[0056] 본 명세서에 유용한 일부 트랜스포좀 복합체는 2개의 트랜스포존 서열을 갖는 트랜스포사제를 포함한다. 일부 이러한 실시형태에서, 2개의 트랜스포존 서열은 서로에 연결되지 않고, 다시 말해서, 트랜스포존 서열은 서로 비-인접하다. 이러한 트랜스포좀의 예는 해당 분야에 공지되어 있다(예를 들어, 미국 특허 출원 공개 제2010/0120098호 참조).

[0057] 일부 실시형태에서, 트랜스포좀 복합체는 2개의 트랜스포사제 서브유닛에 결합하여 "루프형(looped) 복합체" 또는 "루프형 트랜스포좀"을 형성하는 트랜스포존 서열 핵산을 포함한다. 일례에서, 트랜스포좀은 이량체 트랜스포사제 및 트랜스포존 서열을 포함한다. 루프형 복합체는, 원래의 표적 DNA의 정보의 지시를 유지하고, 표적 DNA을 단편화하지 않으면서 트랜스포존이 표적 DNA 내로 삽입되는 것을 보장할 수 있다. 인식될 바와 같이, 루프형 구조는 표적 핵산의 물리적 연결성을 유지시키면서 목적하는 핵산 서열, 예컨대, 인덱스를, 표적 핵산 내에 삽입할 수 있다. 일부 실시형태에서, 루프형 트랜스포좀 복합체의 트랜스포존 서열은 트랜스포존 서열이 단편화되어 2개의 트랜스포존 서열을 포함하는 트랜스포좀 복합체를 생성할 수 있도록 단편화 부위를 포함할 수 있다. 이러한 트랜스포좀 복합체는, 트랜스포존이 삽입되는 이웃하는 표적 DNA 단편이 검정의 이후의 단계에서 확실하게 조립될 수 있는 코드 조합을 수용하는 것을 보장하는 데 유용하다.

[0058] 트랜스포좀 복합체는 또한, 트랜스포사제 인덱스라고도 지칭되는 적어도 하나의 인덱스 서열을 포함한다. 인덱

스 서열은 트랜스포존 서열의 부분으로서 존재한다. 일 실시형태에서, 인덱스 서열은 표적 핵산에 전달되는 트랜스포사제 인식 부위의 가닥인, 전달된 가닥 상에 존재할 수 있다. 태그 또는 바코드라고도 지칭되는 인덱스 서열은 특정 표적 핵산이 존재하였던 구획의 마커 특징으로서 유용하다. 트랜스포좀 복합체의 인덱스 서열은 각각의 구획에 대해 상이하다. 따라서, 이러한 실시형태에서, 인덱스는 특정 구획 내에 존재하는 표적 핵산 각각에 부착된 핵산 서열 태그이고, 이의 존재는 핵의 집단이 이러한 방법 단계에서 존재하였던 구획을 나타내거나 이를 식별하기 위해 사용된다.

[0059] 인덱스 서열은 최대 20개 뉴클레오티드 길이, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개일 수 있다. 4개의 뉴클레오티드 태그가 동일한 어레이 상에서 256개의 시료를 다중화하는 가능성을 제공하고, 6개 염기 태그는 동일한 어레이 상에서 4096개의 시료를 처리하는 것을 가능하게 한다.

[0060] 일 실시형태에서, 전달된 가닥은 또한 범용 서열, 제1 서열결정 프라이머 서열 또는 그의 조합을 포함할 수 있다. 범용 서열 및 서열결정 프라이머 서열이 본 명세서에 기재된다. 따라서, 전달된 가닥이 표적 핵산에 전달되는 일부 실시형태에서, 표적 핵산은 트랜스포사제 인덱스를 포함하며, 또한, 범용 서열, 제1 서열결정 프라이머 서열 또는 그의 조합을 포함한다.

[0061] 일 실시형태에서, 전달된 가닥의 시토신 뉴클레오티드는 메틸화된다. 또 다른 실시형태에서, 전달된 가닥의 뉴클레오티드는 시토신을 함유하지 않는다. 이러한 전달된 가닥 및 트랜스포사제 인덱스 서열, 범용 서열 및/또는 제1 서열결정 프라이머 서열을 포함하는 전달된 가닥 상에 존재하는 임의의 서열은 시토신-고갈된 것으로 지칭될 수 있다. 트랜스포좀 복합체 내의 시토신-고갈된 뉴클레오티드 서열의 이용은 트랜스포사제 효율에 유의미한 영향을 갖지 않는다.

[0062] 방법은 또한 인덱싱된 핵을 생성하는 단계(도 1, 블록 15 및 도 2 블록 iii)를 포함한다. 일 실시형태에서, 인덱싱된 핵을 생성하는 단계는 뉴클레오솜-고갈된 핵의 하위세트에 존재하는 핵산(예를 들어, 각각의 구획에 존재하는 핵산)을 복수의 핵산 단편으로 단편화시키는 것을 포함한다. 일 실시형태에서, 핵산의 단편화는 핵산에 존재하는 단편화 부위를 사용함으로써 달성된다. 전형적으로, 단편화 부위는 트랜스포좀 복합체를 사용함으로써 표적 핵산 내로 도입된다. 예를 들어, 루프형 트랜스포좀 복합체는 단편화 부위를 포함할 수 있다. 단편화 부위를 사용하여 표적 핵산 내로 삽입되는 인덱스 서열 간의 물리적인, 그러나 정보를 제공하지 않는 회합을 절단할 수 있다. 절단은 생화학적, 화학적 또는 기타 수단에 의해 이루어질 수 있다. 일부 실시형태에서, 단편화 부위는 다양한 수단에 의해 단편화될 수 있는 뉴클레오티드 또는 뉴클레오티드 서열을 포함할 수 있다. 단편화 부위의 예는 제한 엔도뉴클레아제 부위, RNAse를 사용하여 절단 가능한 적어도 하나의 리보뉴클레오티드, 특정 화학 작용제의 존재 하에서 절단 가능한 뉴클레오티드 유사체, 과요오드산염으로의 처리에 의해서 절단 가능한 다이올 연결, 화학 환원제를 사용하여 절단 가능한 다이설파이드기, 광화학 절단을 겪을 수 있는 절단 가능한 모이어티, 및 펩티다제 효소 또는 기타 적합한 수단에 의해 절단 가능한 펩티드를 포함하지만, 이들로 제한되지 않는다(예를 들어, 미국 특허 출원 공개 제2012/0208705호, 미국 특허 출원 공개 제2012/0208724호 및 WO 2012/061832호 참조). 단편화의 결과는 인덱싱된 핵의 집단이고, 각각의 핵은 핵산 단편을 함유하며, 핵산 단편은 적어도 하나의 가닥 상에, 특정 구획을 나타내는 인덱스 서열을 포함한다.

[0063] 다수의 구획으로부터의 인덱싱된 핵은 조합될 수 있다(도 1, 블록 16 및 도 2, 좌측 상의 개략도). 예를 들어, 2 내지 96개의 구획(96-웰 플레이트가 사용되는 경우)으로부터, 또는 2 내지 384개의 구획(384-웰 플레이트가 사용되는 경우)으로부터의 인덱싱된 핵이 조합된다. 이어서, 본 명세서에서 풀링된 인덱싱 핵이라 지칭되는 이들 조합된 인덱싱 핵의 하위세트는 제2 복수 구획 내로 분배된다. 하위세트 및 이에 따라 각 구획 내에 존재하는 핵의 수는 부분적으로 상기 방법의 이러한 단계에서 동일한 구획에 존재하게 되는 동일한 트랜스포사제 인덱스를 갖는 2개의 핵의 존재인 인덱스 충돌을 감소시키려는 목적을 기반으로 한다. 이러한 실시형태에서 하위세트에 존재하는 핵의 수는 2 내지 30개, 예컨대, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 또는 30개일 수 있다. 일 실시형태에서, 하위세트에 존재하는 핵의 수는 20 내지 24개, 예컨대, 22개이다. 하위세트 내로의 핵의 분배 방법은 당업자에게 공비되어 있으며, 일상적이다. 예에는 형광-활성화된 핵 분류(FANS)가 포함되지만, 이로 제한되지 않는다.

[0064] 분배된 인덱싱 핵을 처리하여 메틸화된 뉴클레오티드를 식별한다(도 1, 블록 17 및 도 2, 블록 iv). 부위, 예컨대 CpG 다이뉴클레오티드 서열의 메틸화를 이러한 부위의 분석을 위해 해당 분야에 사용되는 다양한 기법 중 임의의 것을 사용하여 측정할 수 있다. 하나의 유용한 방법은 메틸화된 CpG 다이뉴클레오티드 서열의 식별이다. 메틸화된 CpG 다이뉴클레오티드 서열의 식별은 시토신 전환 기반의 기술을 사용하여 결정되며, 이는 단리된 게

놈 DNA 또는 그의 단편 내의 CpG 서열의 메틸화 상태-의존성인 화학적 변형에 이어서 DNA 서열 분석에 좌우된다. 메틸화된 및 비-메틸화된 CpG 다이뉴클레오티드 서열 사이를 구별할 수 있는 화학 시약은 핵산을 절단하는 하이드라진 및 바이설파이트를 포함한다. 바이설파이트 처리에 이어서 알칼리성 가수분해는 문헌[Olek A., 1996, *Nucleic Acids Res.* 24:5064-6] 또는 문헌[Frommer et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:1827-1831]에 기재된 바와 같이 5-메틸시토신을 변형되지 않게 남겨 두어, 특이적으로 비-메틸화된 시토신을 우라실로 전환시킨다. 바이설파이트-처리된 DNA는 이후에 분자 기법, 예컨대 PCR 증폭, 서열결정 및 올리고뉴클레오티드 혼성화를 포함하는(예를 들어, 핵산 마이크로어레이를 사용하는) 검출에 의해 분석될 수 있다. 일 실시형태에서, 각 구획 내의 인덱싱된 핵은 바이설파이트 처리를 위한 조건에 노출된다. 핵산의 바이설파이트 처리는 당업자에게 공지되어 있으며, 일상적이다. 일 실시형태에서, 바이설파이트 처리는 CpG 다이뉴클레오티드의 비메틸화된 시토신 잔기를 우라실 잔기로 전환시키고, 5-메틸시토신 잔기를 변경되지 않게 남겨 둔다. 바이설파이트 처리는 바이설파이트-처리된 핵산 단편을 초래한다.

[0065] 바이설파이트-처리된 핵산 단편의 생성 후에, 단편이 하나의 또는 양 말단에 추가의 뉴클레오티드를 포함하도록 변형시킨다(도 1, 블록 18 및 도 2, 블록 v 및 vi). 일 실시형태에서, 변형은 바이설파이트-처리된 핵산 단편을 복수의 프라이머를 사용하여 선형 증폭으로 처리하는 것을 포함한다. 각각의 프라이머는 적어도 2개의 영역; 5' 말단에 범용 뉴클레오티드 서열 및 3' 말단에 무작위 뉴클레오티드 서열을 포함한다. 범용 뉴클레오티드 서열은 각각의 프라이머에서 동일하며, 일 실시형태에서, 그것은 제2 서열결정 프라이머 서열(도 2(블록 vii)에서 판독 2 프라이머로도 지칭됨)을 포함한다. 바이설파이트-처리된 핵산 단편 내의 모든 서열에 상보성인 적어도 하나의 프라이머가 존재하도록 무작위 뉴클레오티드 서열의 영역을 사용한다. 완전한 커버리지의 확률을 요망되는 수준까지 증가시키기 위해 사용될 수 있는 무작위 뉴클레오티드의 수는 일상적인 방법을 사용하여 결정될 수 있으며, 6 내지 12개의 무작위 뉴클레오티드, 예컨대 9개의 무작위 뉴클레오티드일 수 있다. 일 실시형태에서, 사이클의 수는 10회 이하의 사이클, 예컨대 9회의 사이클, 8회의 사이클, 7회의 사이클, 6회의 사이클, 5회의 사이클, 4회의 사이클, 3회의 사이클, 2회의 사이클 또는 1회의 사이클에 제한된다. 선형 증폭의 결과는 증폭된 단편-어댑터 분자이다. 단편-어댑터 분자의 예는 도 3에 나타나 있다. 단편-어댑터 분자(30)는 증폭 및/또는 서열결정을 위해 사용될 수 있는 트랜스포사제 인덱스 및 범용 서열을 포함하는 트랜스포좀 복합체(31 및 32)의 전달된 가닥으로부터 기원한 뉴클레오티드를 포함한다. 단편-어댑터 분자는 또한 핵의 게놈 DNA로부터 기원한 뉴클레오티드(33), 무작위 뉴클레오티드 서열의 영역(34) 및 범용 뉴클레오티드 서열(35)을 포함한다.

[0066] 선형 증폭은, 고정화 및 서열결정 이전에 단편-어댑터 분자의 말단을 추가로 변형시키기 위하여, 지수적 증폭 반응, 예컨대 PCR로 이어진다. 이러한 단계는 PCR에 의한 단편-어댑터 분자의 인덱싱을 초래한다(도 1, 블록 19). 단편-어댑터 분자의 말단에 존재하는 범용 서열(31, 32 및/또는 35)은 프라이머로서 제공될 수 있는 범용 앵커 서열의 결합을 위해 사용될 수 있으며, 증폭 반응에서 연장될 수 있다. 전형적으로, 2개의 상이한 프라이머가 사용된다. 한 프라이머는 단편-어댑터 분자의 한 가닥의 3' 말단에서 범용 서열과 혼성화하며, 제2 프라이머는 단편-어댑터 분자의 다른 가닥의 3' 말단에서 범용 서열과 혼성화한다. 따라서, 각각의 프라이머의 앵커 서열은 상이할 수 있다. 적합한 프라이머는 각각 추가의 범용 서열, 예컨대 범용 포획 서열 및 또 다른 인덱스 서열을 포함할 수 있다. 각각의 프라이머가 인덱스를 포함할 수 있기 때문에, 이러한 단계는 1개 또는 2개의 인덱스 서열, 예를 들어, 제2 및 선택적인 제3 인덱스의 부가를 초래한다. 제2 및 선택적인 제3 인덱스를 갖는 단편-어댑터 분자는 이중-인덱스 단편-어댑터 분자로서 지칭된다. 제2 및 제3 인덱스는 서로의 역 상보물일 수 있거나, 제2 및 제3 인덱스는 서로의 역 상보물이 아닌 서열을 가질 수 있다. 이러한 제2 인덱스 서열 및 선택적인 제3 인덱스는 분배된 인덱싱 핵이 소듐 바이설파이트로의 처리 이전에 배치되는 각 구획에 대하여 특유하다. 이러한 PCR 증폭의 결과는 도 2, 블록 vii에 나타낸 단편-어댑터 분자와 유사하거나 이와 동일한 구조를 갖는 복수의 단편-어댑터 분자 또는 단편-어댑터 분자의 라이브러리이다.

[0067] 또 다른 실시형태에서, 변형은 바이설파이트-처리된 핵산 단편을, 단편의 양 말단으로의 추가의 서열의 라이게이션을 초래하는 조건으로 처리하는 것을 포함한다. 일 실시형태에서, 블런트-말단 라이게이션이 사용될 수 있다. 또 다른 실시형태에서, 단편은 예를 들어, 단일의 데옥시뉴클레오티드, 예를 들어, 데옥시아데노신(A)을 바이설파이트-처리된 핵산 단편의 3' 말단에 부가하는 비-주형-의존적 말단 트랜스퍼라제 활성을 갖는 특정 유형의 DNA 중합효소, 예컨대 Taq 중합효소 또는 클레나우 엑소 마이너스(Klenow exo minus) 중합효소의 활성에 의해 단일의 오버행 뉴클레오티드와 함께 제조된다. 이러한 효소를 사용하여 단일의 뉴클레오티드 'A'를 단편의 각각의 가닥의 블런트 말단인 3' 말단에 부가할 수 있다. 따라서, 'A'를 Taq 또는 클레나우 엑소 마이너스 중합효소를 사용한 반응에 의해 이중-가닥 표적 단편의 각각의 가닥의 3' 말단에 부가할 수 있는 한편, 단편의 각각의 말단에 부가될 추가의 서열은 부가될 이중 가닥 핵산의 각각의 영역의 3' 말단에 존재하는 양립 가능한 'T' 오버행을 포함할 수 있다. 이러한 말단 변형은 또한 핵산의 자가-라이게이션을 방지하여, 이러한 실시형태에서

부가되는 서열에 의해 플랭킹된 바이설파이트-처리된 핵산 단편의 형성을 향한 편재가 존재하게 한다.

[0068] 본 명세서에 기재된 방법에 의한 핵산 분자의 단편화는 블런트 및 3'- 및 5'-오버행 말단의 비균질 믹스를 갖는 단편을 초래한다. 따라서, 해당 분야에 공지된 방법 또는 키트(예컨대, 루시젠(Lucigen) DNA 종결자 말단 수선 키트(terminator End Repair Kit))를 사용하여 단편 말단을 수선하여 예를 들어, 클로닝 벡터의 블런트 부위 내로의 삽입에 최적인 말단을 생성하는 것이 바람직할 수 있다. 특정 실시형태에서, 핵산의 집단의 단편 말단은 블런트 말단화된다. 더욱 특히, 단편 말단은 블런트 말단화되고, 인산화된다. 포스페이트 모이어티는 예를 들어, 폴리뉴클레오타이드 키나제를 사용한 효소적 처리를 통해 도입될 수 있다.

[0069] 일 실시형태에서, 바이설파이트-처리된 핵산 단편은 먼저 동일한 범용 어댑터(또한 '불일치된 어댑터라고 지칭'되며, 이의 일반적인 특징은 Gormley 등의 US 7,741,463호 및 Bignell 등의 US 8,053,192호에 기재되어 있음)를 바이설파이트-처리된 핵산 단편의 5' 및 3' 말단에 라이게이션시킴으로써 처리되어, 단편-어댑터 분자를 형성한다. 일 실시형태에서, 범용 어댑터는 단편-어댑터 분자를 어레이 상에 고정화시키는 것을 포함하여, 서열결정을 위해 필요한 모든 서열을 포함한다. 서열결정될 핵산이 단일의 세포로부터의 것이기 때문에, 단편-어댑터 분자의 추가의 증폭이 서열결정을 위하여 충분한 수의 단편-어댑터 분자를 달성하는데 도움이 된다.

[0070] 또 다른 실시형태에서, 범용 어댑터가 서열결정에 필요한 모든 서열을 포함하지 않는 경우, 고정화 및 서열결정 이전에 각각의 단편-어댑터 분자에 존재하는 범용 어댑터를 추가로 변형시키기 위하여 PCR 단계가 사용될 수 있다. 예를 들어, 초기 프라이머 신장 반응은 단편-어댑터 분자에 존재하는 범용 서열에 상보성인 범용 앵커 서열을 사용하여 수행되며, 여기서, 각각의 개별 단편-어댑터 분자의 양 가닥에 상보성인 신장 생성물이 형성된다. 전형적으로, PCR은 추가의 범용 서열, 예컨대 범용 포획 서열 및 또 다른 인덱스 서열을 부가한다. 각각의 프라이머가 인덱스를 포함할 수 있기 때문에, 이러한 단계는 1 또는 2개의 인덱스 서열, 예를 들어, 제2 및 선택적인 제3 인덱스의 부가, 및 어댑터 라이게이션에 의한 단편-어댑터 분자의 인덱싱을 초래한다(도 1, 블록 19). 생성된 단편-어댑터 분자는 이중-인덱스 단편-어댑터 분자로 지칭된다.

[0071] 서열결정에 필요한 모든 서열을 포함하는 범용 어댑터를 라이게이션시키는 단일 단계 방법에 의해 또는 범용 어댑터를 라이게이션시킨 다음, 범용 어댑터를 추가로 변형시키기 위하여 PCR 증폭시키는 2-단계 방법에 의해 범용 어댑터를 부가한 후에, 최종 단편-어댑터 분자는 범용 포획 서열, 제2 인덱스 서열 및 선택적인 제3 인덱스 서열을 포함할 것이다. 이들 인덱스는 선형 증폭에 의한 이중-인덱스 단편-어댑터의 생성에 기재된 제2 및 제3 인덱스와 유사하다. 제2 및 제3 인덱스는 서로의 역 상보물일 수 있거나, 제2 및 제3 인덱스는 서로의 역 상보물이 아닌 서열을 가질 수 있다. 이들 제2 및 선택적인 제3 인덱스 서열은 분배된 인덱싱 핵이 소듐 바이설파이트로의 처리 이전에 배치되는 각각의 구획에 대하여 특유하다. 각 말단으로의 범용 어댑터의 부가의 결과는 도 4에 나타낸 단편-어댑터 분자(40)와 유사하거나 동일한 구조를 갖는 복수의 단편-어댑터 분자 또는 단편-어댑터 분자의 라이브러리이다. 단편-어댑터 분자(40)는 각각 3' 유동셀 어댑터(예를 들어, P5) 및 5' 유동셀 어댑터(예를 들어, P7'), 및 인덱스(42 및 47), 예컨대 i5 및 i7로도 지칭되는 포획 서열(41 및 48)을 포함한다. 단편-어댑터 분자(40)는 또한 증폭 및/또는 서열결정을 위해 사용될 수 있는 트랜스포사제 인덱스(44) 및 범용 서열(45)을 포함하는 트랜스포좀 복합체(43)의 전달된 가닥으로부터 기원하는 뉴클레오티드를 포함한다. 단편-어댑터 분자는 또한, 핵의 게놈 DNA로부터 기원한 뉴클레오티드(46)를 포함한다.

[0072] 생성된 이중-인덱스 단편-어댑터 분자는 집합적으로 고정화된 다음, 서열결정될 수 있는 핵산의 라이브러리를 제공한다. 용어 라이브러리는 그들의 3' 및 5' 말단에 공지된 범용 서열을 함유하는 단일의 세포로부터의 단편의 집합물을 지칭한다.

[0073] 바이설파이트-처리된 핵산 단편이 추가의 뉴클레오티드를 포함하도록 변형시킨 후에, 이중-인덱스 단편-어댑터 분자를 사전결정된 크기 범위, 예컨대 150 내지 400개 뉴클레오티드 길이, 예컨대 150 내지 300개 뉴클레오티드에 대하여 선택하는 조건으로 처리할 수 있다. 생성된 이중-인덱스 단편-어댑터 분자를 풀링하고, 선택적으로, 정화 과정으로 처리하여, 비혼입된 범용 어댑터 또는 프라이머의 적어도 일부를 제거함으로써 DNA 분자에 대한 순도를 향상시킬 수 있다. 임의의 적합한 정화 과정, 예컨대 전기영동, 크기 배제 크로마토그래피 등이 사용될 수 있다. 일부 실시형태에서, 고체상 가역성 고정화 상자성 비드를 사용하여 요망되는 DNA 분자를 비부착된 범용 어댑터 또는 프라이머로부터 분리하고, 크기에 기초하여 핵산을 선택할 수 있다. 고체상 가역성 고정화 상자성 비드는 베크만 쿨터(Beckman Coulter)(Agencourt AMPure XP), 써모피셔(Thermofisher)(MagJet), 오메가 바이오텍(Omega Biotek)(Mag-Bind), 프로메가(Promega) 비드(프로메가) 및 카파 바이오시스템즈(Kapa Biosystems)(카파 퓨어 비드(Kapa Pure Beads))로부터 상업적으로 입수 가능하다.

[0074] 복수의 단편-어댑터 분자가 서열결정을 위해 제조될 수 있다. 단편-어댑터 분자를 풀링한 후에, 그들을 고정화

시키고 증폭시킨 후 서열결정한다(도 1, 블록 20). 하나 이상의 공급원으로부터의 단편-어댑터 분자를 기판에 부착시키는 방법은 해당 분야에 알려져 있다. 마찬가지로, 고정화된 단편-어댑터 분자를 증폭시키기 위한 방법은 브릿지(bridge) 증폭 및 역학적 배제를 포함하지만, 이에 제한되지 않는다. 서열결정 이전의 고정화 및 증폭 방법은 예를 들어, 문헌[Bignell et al.](US 8,053,192호), 문헌[Gunderson et al.](WO2016/130704호), 문헌[Shen et al.](US 8,895,249호) 및 문헌[Pipenburg et al.](US 9,309,502호)에 기재되어 있다.

[0075] 풀링된 시료는 서열결정을 위하여 제제 중에서 고정화될 수 있다. 서열결정은 단일 분자의 어레이로서 수행될 수 있거나, 또는 서열결정 이전에 증폭될 수 있다. 증폭은 하나 이상의 고정화된 프라이머를 사용하여 수행될 수 있다. 고정화된 프라이머(들)는 편평한 표면 상의 또는 비드의 풀 상의 론(lawn)일 수 있다. 비드의 풀은 유제의 각각의 "구획" 내에 단일 비드를 갖는 유제 내로 단리될 수 있다. "구획"당 단지 하나의 주형의 농도로, 단일 주형만이 각각의 비드 상에서 증폭된다.

[0076] 용어 "고체상 증폭"은 본 명세서에 사용되는 바와 같이 고체 지지체 상에서 수행되거나 이것과 함께 수행되어 증폭된 생성물의 전부 또는 일부가 그들이 형성될 때 고체 지지체 상에 고정화되게 하는 임의의 핵산 증폭 반응을 지칭한다. 특히, 이 용어는 고체상 중합효소 연쇄 반응(고체상 PCR) 및 고체상 등온 증폭을 포함하는데, 이것은 정방향 및 역방향 증폭 프라이머 중 하나 또는 둘 모두가 고체 지지체 상에 고정된 것을 제외하고는 표준 용액상 증폭과 유사한 반응이다. 고체상 PCR은 시스템, 예컨대, 유제(여기서 하나의 프라이머가 비드에 앵커링되고, 나머지는 자유 용액 중에 존재함) 및 고체상 겔 매트릭스 내의 콜로니 형성(여기서 하나의 프라이머가 표면에 앵커링되고, 나머지는 자유 용액 중에 존재함)을 포함한다.

[0077] 일부 실시형태에서, 고체 지지체는 패턴화된 표면을 포함한다. "패턴화된 표면"은 고체 지지체의 노출된 층 내의 또는 층 상의 상이한 영역의 배열을 지칭한다. 예를 들어, 영역 중 하나 이상은 하나 이상의 증폭 프라이머가 존재하는 특징부일 수 있다. 특징부는 증폭 프라이머가 존재하지 않는 사이(interstitial) 영역에 의해 분리될 수 있다. 일부 실시형태에서, 패턴은 행 및 열로 존재하는 특징부의 x-y 포맷일 수 있다. 일부 실시형태에서, 패턴은 특징부 및/또는 사이 영역의 반복되는 배열일 수 있다. 일부 실시형태에서, 패턴은 특징부 및/또는 사이 영역의 무작위 배열일 수 있다. 본 명세서에 언급된 방법 및 조성물에 사용될 수 있는 예시적인 패턴화된 표면은 미국 특허 제8,778,848호, 제8,778,849호 및 제9,079,148호, 및 미국 특허 공개 제2014/0243224호에 기재되어 있다.

[0078] 일부 실시형태에서, 고체 지지체는 표면 내의 웰 또는 함몰부의 어레이를 포함한다. 이것은 일반적으로 해당 분야에 공지된 바와 같이 다양한 기법, 예컨대, 비제한적으로 포토리소그래피, 스탬핑 기법, 성형 기법 및 마이크로에칭 기법을 사용하여 제작될 수 있다. 당업자에 의해 이해될 바와 같이, 사용되는 기법은 어레이 기판의 조성 및 형상에 좌우될 것이다.

[0079] 패턴화된 표면 내의 특징부는 유리, 규소, 플라스틱 또는 패턴화된 공유적으로 연결된 겔, 예컨대, 폴리(N-(5-아지도아세트아미딜펜틸)아크릴아미드-코-아크릴아미드)(PAZAM, 예를 들어, 미국 특허 공개 제2013/184796호, WO 2016/066586호, 및 WO 2015/002813호 참조)를 갖는 다른 적합한 고체 지지체 상의 웰의 어레이의 웰(예를 들어, 마이크로웰 또는 나노웰)일 수 있다. 이 과정은 다수회의 사이클을 사용한 서열결정 실행에 걸쳐 안정적일 수 있는 서열결정을 위해 사용되는 겔 패드를 생성한다. 웰로의 중합체의 공유적 연결은 다양한 용도 동안 구조화된 기판의 수명 내내 구조화된 특징부 내에 겔을 유지하는 데 도움이 된다. 그러나, 많은 실시형태에서, 겔은 웰에 공유적으로 연결될 필요는 없다. 예를 들어, 일부 조건에서 구조화된 기판의 어느 부분에도 공유적으로 부착되지 않은 실란 부재 아크릴아미드(SFA, 예를 들어, 본 명세서에 그의 전문이 참조로 포함되는 미국 특허 제8,563,477호 참조)가 겔 물질로서 사용될 수 있다.

[0080] 특정 실시형태에서, 구조화된 기판은 고체 지지체 물질을 웰(예를 들어, 마이크로웰 또는 나노웰)로 패턴화시키고, 패턴화된 지지체를 겔 물질(예를 들어, PAZAM, SFA 또는 그의 화학적으로 변형된 변이체, 예컨대, SFA의 아지도 분해된(azidolyzed) 버전(아지도-SFA))로 코팅하고, 예를 들어, 화학적 또는 기계적 폴리싱을 통해, 겔 코팅된 지지체를 폴리싱하여 웰 내에 겔을 보유하지만 웰 사이의 구조화된 기판의 표면 상의 사이 영역로부터 실질적으로 모든 겔을 제거하거나 불활성화시킴으로써 제조될 수 있다. 프라이머 핵산이 겔 물질에 부착될 수 있다. 이어서 단편-어댑터 분자의 용액을 폴리싱된 기판과 접촉시켜, 개별 단편-어댑터 분자가 겔 물질에 부착된 프라이머와의 상호작용을 통해 개별 웰에 씨딩되게 할 것이지만; 표적 핵산은 겔 물질의 부재 또는 비활성으로 인하여 사이 영역을 점유하지 않을 것이다. 단편-어댑터 분자의 증폭은 웰로 한정될 것인데, 그 이유는 사이 영역 내의 겔의 부재 또는 비활성이 성장하는 핵산 콜로니의 외측으로의 이동을 방지하기 때문이다. 이러한 과정은 편리하게 제조될 수 있고, 확대 가능하고, 종래의 마이크로- 또는 나노제작 방법을 이용한다.

[0081] 본 개시내용은 단지 하나의 증폭 프라이머가 고정화된(다른 프라이머는 통상 자유 용액 중에 존재함) "고체상" 증폭 방법을 포함하지만, 정방향 및 역방향 프라이머가 둘 모두가 고정되는 고체 지지체가 제공되는 것이 바람직하다. 실제로, 고체 지지체 상에 고정화된 '복수'의 동일한 정방향 프라이머 및/또는 '복수'의 동일한 역방향 프라이머가 존재할 것인데, 그 이유는 증폭 과정이 증폭을 유지시키기 위해 과량의 프라이머를 필요로 하기 때문이다. 본 명세서에서 정방향 및 역방향 프라이머에 대한 언급은 이에 따라 문맥에서 달리 제시하지 않는 한 '복수'의 이러한 프라이머를 포함하는 것으로 해석되어야 한다.

[0082] 숙련자에 의해 이해될 바와 같이, 임의의 주어진 증폭 반응은 증폭될 주형에 대해 특이적인 적어도 하나의 유형의 정방향 프라이머 및 적어도 하나의 유형의 역방향 프라이머가 필요하다. 그러나, 특정 실시형태에서 정방향 프라이머 및 역방향 프라이머는 동일한 서열의 주형-특이적 부분을 포함할 수 있고, 완전히 동일한 뉴클레오티드 서열 및 구조(임의의 비-뉴클레오티드 변형을 포함함)를 가질 수 있다. 다시 말해서, 단지 하나의 유형의 프라이머를 사용하여 고체-상 증폭을 수행하는 것이 가능하고, 이러한 단일-프라이머 방법은 본 발명의 범주 내에 포함된다. 다른 실시형태는 정방향 프라이머 및 역방향 프라이머를 사용할 수 있는데, 이것은 동일한 주형-특이적 서열을 함유하지만, 일부 다른 구조 특징이 상이하다. 예를 들어, 하나의 유형의 프라이머는 다른 것에 존재하지 않는 비-뉴클레오티드 변형을 함유할 수 있다.

[0083] 본 개시내용의 모든 실시형태에서, 고체-상 증폭을 위한 프라이머는 바람직하게는 단일 지점 공유적 부착에 의해 프라이머의 5' 말단에서 또는 그 근처에서 고체 지지체에 고정화되어, 그의 동족 주형에 어닐링시키기에 자유로운 프라이머의 주형-특이적 부분 및 프라이머 신장을 위해 자유로운 3' 하이드록실기를 남긴다. 해당 분야에 공지된 임의의 적합한 공유적 부착 수단이 이러한 목적을 위해 사용될 수 있다. 선택된 부착 화학은 고체 지지체의 성질, 및 그것에 적용되는 임의의 유도체화 또는 작용화에 좌우될 것이다. 프라이머 그 자체는 비-뉴클레오티드 화학적 변형일 수 있는 모이어티를 포함하여 부착을 용이하게 할 수 있다. 특정 실시형태에서, 프라이머는 5' 말단에 황-함유 친핵체, 예컨대, 포스포로티오에이트 또는 티오포스페이트를 포함할 수 있다. 고체-지지된 폴리아크릴아미드 하이드로겔의 경우에, 이러한 친핵체는 하이드로겔에 존재하는 브로모아세트아미드기에 결합할 것이다. 프라이머 및 주형을 고체 지지체에 부착시키는 더욱 특정한 수단은 WO05/065814호에 완전히 기재된 바와 같은, 중합된 아크릴아미드 및 N-(5-브로모아세트아미딜펜틸) 아크릴아미드(BRAPA)로 구성된 하이드로겔로의 5' 포스포로티오에이트 부착을 통한 것이다.

[0084] 본 개시내용의 특정 실시형태는 예를 들어, 생물분자, 예컨대, 폴리뉴클레오티드로의 공유적 부착을 가능하게 하는 반응성 기를 포함하는 중간체 물질의 층 또는 코팅을 적용함으로써 "작용화"된 비활성 기판 또는 매트릭스(예를 들어, 유리 슬라이드, 중합체 비드 등)로 구성된 고체 지지체를 이용할 수 있다. 이러한 지지체의 예는 비활성 기질, 예컨대, 유리 상에 지지된 폴리아크릴아마이드 하이드로젤을 포함하지만, 이들로 제한되지 않는다. 이러한 실시형태에서, 생물분자(예를 들어, 폴리뉴클레오티드)는 중간체 물질(예를 들어, 하이드로겔)에 직접 공유적으로 부착될 수 있지만, 중간체 물질 그 자체는 기판 또는 매트릭스(예를 들어, 유리 기판)에 비-공유적으로 부착될 수 있다. 용어 "고체 지지체로의 공유적 부착"은 이에 따라 이러한 유형의 배열을 포함하는 것으로 해석되어야 한다.

[0085] 풀링된 시료는 비드 상에서 증폭될 수 있고, 여기서 각각의 비드는 정방향 및 역방향 증폭 프라이머를 함유한다. 특정 실시형태에서, 단편-어댑터 분자의 라이브러리를 사용하여, 고체-상 증폭 및 더욱 특히 고체상 등온 증폭에 의해 미국 특허 공개 제2005/0100900호, 미국 특허 제7,115,400호, WO 00/18957호 및 WO 98/44151호에 기재된 것과 유사한, 핵산 콜로니의 클러스터링된 어레이를 제조한다. 용어 '클러스터' 및 '콜로니'는 복수의 동일한 고정화된 핵산 가닥 및 복수의 동일한 고정화된 상보성 핵산 가닥을 포함하는 고체 지지체 상의 별개의 부위를 지칭하도록 본 명세서에서 상호 교환 가능하게 사용된다. 용어 "클러스터링된 어레이"는 이러한 클러스터 또는 콜로니로부터 형성된 어레이를 지칭한다. 이와 관련하여, 용어 "어레이"는 클러스터의 정렬된 배열을 필요로 하는 것으로서 이해되지 않을 것이다.

[0086] 용어 "고체상" 또는 "표면"은 프라이머가 평평한 표면, 예를 들어, 유리, 실리카 또는 플라스틱 현미경 슬라이드에 부착된 평면 어레이 또는 유사한 유동 셀 디바이스 중 어느 하나; 1개 또는 2개의 프라이머가 비드에 부착되고, 비드가 증폭되는 비드; 또는 비드가 증폭된 후 표면 상의 비드의 어레이를 의미하도록 사용된다.

[0087] 클러스터링된 어레이는 WO 98/44151호에 기재된 바와 같은 열순환 과정, 또는 온도가 일정하게 유지되고, 신장 및 변성의 사이클이 시약의 변화를 사용하여 수행되는 과정 중 어느 하나를 사용하여 제조될 수 있다. 이러한 등온 증폭 방법은 특허 출원 제WO 02/46456호 및 미국 특허 공개 제2008/0009420호에 기재되어 있다. 등온 과정에서 유용한 더 낮은 온도로 인하여, 이것이 특히 바람직하다.

[0088] 본 명세서에 기재되거나 해당 분야에 일반적으로 공지된 증폭 방법 중 임의의 것을 범용 또는 표적-특이적 프라이머와 함께 사용하여 고정화된 DNA 단편을 증폭시킬 수 있는 것이 이해될 것이다. 증폭에 적합한 방법은 본 명세서에 그의 전문이 참조로 포함되는 미국 특허 제8,003,354호에 기재된 바와 같은, 중합효소 연쇄 반응(PCR), 가닥 치환 증폭(SDA), 전사 매개 증폭(TMA) 및 핵산 서열 기반 증폭(NASBA)을 포함하지만, 이에 제한되지 않는다. 상기 증폭 방법을 사용하여 하나 이상의 관심 핵산을 증폭시킬 수 있다. 예를 들어, 멀티플렉스 PCR을 비롯한 PCR, SDA, TMA, NASBA 등을 사용하여 고정화된 DNA 단편을 증폭시킬 수 있다. 일부 실시형태에서, 관심 폴리뉴클레오티드에 특이적으로 지향된 프라이머는 증폭 반응에 포함된다.

[0089] 폴리뉴클레오티드의 증폭을 위한 다른 적합한 방법은 올리고뉴클레오티드 신장 및 라이게이션, 회전환 증폭(RCA)(문헌[Lizardi et al., Nat. Genet. 19:225-232 (1998)]) 및 올리고뉴클레오티드 라이게이션 검정(OLA)(일반적으로 미국 특허 제7,582,420호, 제5,185,243호, 제5,679,524호 및 제5,573,907호, EP 0 320 308 B1호; EP 0 336 731 B1호; EP 0 439 182 B1호; WO 90/01069호; WO 89/12696호; 및 WO 89/09835호 참조) 기술을 포함할 수 있다. 이들 증폭 방법이 고정화된 DNA 단편을 증폭시키도록 설계될 수 있는 것이 이해될 것이다. 예를 들어, 일부 실시형태에서, 증폭 방법은 라이게이션 프로브 증폭 또는 관심 핵산에 특이적으로 지향된 프라이머를 함유하는 올리고뉴클레오티드 라이게이션 검정(OLA) 반응을 포함할 수 있다. 일부 실시형태에서, 증폭 방법은 관심 핵산에 특이적으로 지향된 프라이머를 함유하는 프라이머 신장-라이게이션 반응을 포함할 수 있다. 관심 핵산을 증폭시키도록 특이적으로 설계될 수 있는 프라이머 신장 및 라이게이션 프라이머의 비제한적인 예로서, 증폭은 미국 특허 제7,582,420호 및 제7,611,869호에 예시된 바와 같은 골든게이트(GoldenGate) 검정(미국 캘리포니아주 샌디에고 소재의 일루미나, 인코포레이티드(Illumina, Inc.))을 위해 사용된 프라이머를 포함할 수 있다.

[0090] 본 개시내용의 방법에서 사용될 수 있는 예시적인 등온 증폭 방법은 예를 들어, 문헌[Dean et al., Proc. Natl. Acad. Sci. USA 99:5261-66 (2002)]에 예시된 바와 같은 다중 치환 증폭(MDA), 또는 예를 들어, 미국 특허 제6,214,587호에 의해 예시된 바와 같은 등온 가닥 치환 핵산 증폭을 포함하지만, 이에 제한되지 않는다. 본 개시내용에 사용될 수 있는 다른 비-PCR-기반 방법은 예를 들어, 문헌[Walker et al., Molecular Methods for Virus Detection, Academic Press, Inc., 1995]; 미국 특허 제5,455,166호, 및 제5,130,238호, 및 문헌[Walker et al., Nucl. Acids Res. 20:1691-96 (1992)]에 기재된 가닥 치환 증폭(SDA) 또는 예를 들어, 문헌[Lage et al., Genome Res. 13:294-307 (2003)]에 기재된 과분지화 가닥 치환 증폭(hyper-branched strand displacement amplification)을 포함한다. 게놈 DNA의 무작위 프라이머 증폭을 위해 가닥-치환 Phi 29 중합효소 또는 Bst DNA 중합효소 큰 단편, 5'->3' 엑소-와 함께 등온 증폭 방법이 사용될 수 있다. 이들 중합효소의 사용은 그들의 높은 진행도 및 가닥 치환 활성을 이용한다. 높은 진행도는 중합효소가 10 내지 20 kb 길이인 단편을 생성시킬 수 있게 한다. 상기에 언급된 바와 같이, 낮은 진행도 및 가닥-치환 활성을 갖는 중합효소, 예컨대, 클레나우 중합효소를 사용하여 등온 조건 하에서 더 작은 단편이 생성될 수 있다. 증폭 반응, 조건 및 성분의 추가의 설명은 미국 특허 제7,670,810호의 개시내용에 상세하게 언급되어 있다.

[0091] 본 개시내용에서 유용한 또 다른 폴리뉴클레오티드 증폭 방법은 예를 들어, 문헌[Grothues et al. Nucleic Acids Res. 21(5):1321-2 (1993)]에 기재된 바와 같은, 불변 5' 영역에 이어서 무작위 3' 영역을 갖는 2-도메인 프라이머의 집단을 사용하는 태깅된 PCR(Tagged PCR)이다. 제1 라운드의 증폭을 수행하여, 무작위로 합성된 3' 영역으로부터의 개별 혼성화에 기초한 열 변성된 DNA 상에서 다수의 개시를 가능하게 한다. 3' 영역의 성질로 인해, 개시 부위는 게놈 전체에서 무작위적인 것으로 고려된다. 그 후, 비결합된 프라이머를 제거할 수 있고, 불변 5' 영역에 상보성인 프라이머를 사용하여 추가 복제가 일어날 수 있다.

[0092] 일부 실시형태에서, 등온 증폭은 배제 증폭(ExAmp)으로도 지칭되는 역학 배제 증폭(KEA)을 사용하여 수행될 수 있다. 본 개시내용의 핵산 라이브러리는 증폭 시약을 반응시켜 증폭 부위가 씨딩된 개별 표적 핵산으로부터의 실질적으로 클론 집단의 앰플리콘을 각각 포함하는 복수의 증폭 부위를 생성시키는 단계를 포함하는 방법을 사용하여 제조될 수 있다. 일부 실시형태에서, 증폭 반응은 각각의 증폭 부위의 용량을 채우기에 충분한 수의 앰플리콘이 생성될 때까지 진행된다. 이러한 방식으로 이미 씨딩된 부위를 용량까지 채우는 것은 표적 핵산이 그 부위에 랜딩하고, 증폭하여, 그 부위에서 클론 집단의 앰플리콘을 생성시키는 것을 방지한다. 일부 실시형태에서, 제2 표적 핵산이 그 부위에 도착하기 전에 증폭 부위가 용량까지 채워지지 않더라도 명백한 클론성이 달성될 수 있다. 일부 조건 하에서, 제1 표적 핵산의 증폭은, 그 부위로 수송되는 제2 표적 핵산으로부터의 카피의 생성을 효과적으로 초과하거나 압도하도록 충분한 수의 카피가 제조되는 점까지 진행될 수 있다. 예를 들어, 직경이 500 nm보다 더 작은 원형 특징부 상에서 브릿지 증폭 과정을 사용하는 실시형태에서, 제1 표적 핵산에 대한 14 사이클의 지수적 증폭 후에, 동일한 부위에서의 제2 표적 핵산으로부터의 오염은 일루미나 서열결정 플랫

폼 상에서의 합성에 의한 서열결정 분석에 악영향을 미치기에 불충한 수의 오염된 앰플리콘을 생성할 것이 결정된 바 있다.

[0093] 일부 실시형태에서, 어레이 내의 증폭 부위는 완전히 클론성일 수 있지만 그럴 필요는 없다. 오히려, 일부 응용을 위해서, 개별 증폭 부위는 제1 단편-어댑터 분자로부터의 앰플리콘이 우세하게 존재할 수 있고, 제2 표적 핵산으로부터의 낮은 수준의 오염된 앰플리콘도 또한 가질 수 있다. 어레이는, 오염 수준이 어레이의 후속 사용에 허용 가능하지 않은 영향을 갖지 않는 한 낮은 수준의 오염된 앰플리콘을 갖는 하나 이상의 증폭 부위를 가질 수 있다. 예를 들어, 어레이를 검출 응용에서 사용하려는 경우, 허용 가능한 수준의 오염은 허용 가능하지 않은 방식으로 검출 기법의 신호 대 노이즈 또는 분해능에 영향을 주지 않는 수준일 것이다. 따라서, 명백한 클론성은 일반적으로 본 명세서에 언급된 방법에 의해 제조된 어레이의 특정 사용 또는 응용에 관련될 것이다. 특정 응용을 위해 개별 증폭 부위에서 허용 가능할 수 있는 예시적인 오염 수준은 최대 0.1%, 0.5%, 1%, 5%, 10% 또는 25%의 오염된 앰플리콘을 포함하지만, 이에 제한되지 않는다. 어레이는 이들 예시적인 수준의 오염된 앰플리콘을 갖는 하나 이상의 증폭 부위를 포함할 수 있다. 예를 들어, 어레이 내의 증폭 부위 중 최대 5%, 10%, 25%, 50%, 75%, 또는 심지어는 100%가 일부 오염된 앰플리콘을 가질 수 있다. 어레이 또는 부위의 다른 집합에서, 부위 중 적어도 50%, 75%, 80%, 85%, 90%, 95% 또는 99% 또는 그 초과가 클론성이거나 또는 명백하게 클론성일 수 있다.

[0094] 일부 실시형태에서, 또 다른 사례 또는 과정이 일어나는 것을 효과적으로 배제하기에 충분히 신속한 속도로 과정이 일어나는 경우 역학적 배제가 일어날 수 있다. 예를 들어, 핵산 어레이의 제조를 고려하여, 어레이의 부위를 용액으로부터의 단편-어댑터 분자로 무작위로 씨딩하고, 단편-어댑터 분자의 카피를 증폭 과정에서 생성하여 씨딩된 부위 각각을 용량까지 채운다. 본 개시내용의 역학적 배제 방법에 따라서, 씨딩 및 증폭 과정은 증폭 속도가 씨딩 속도를 초과하는 조건 하에서 동시에 진행될 수 있다. 이와 같이, 카피가 제1 표적 핵산이 씨딩된 부위에서 제조되는 비교적 신속한 속도는 제2 핵산이 증폭을 위한 부위를 씨딩하는 것을 효과적으로 배제시킬 것이다. 역학적 배제 증폭 방법은 미국 출원 공개 제2013/0338042호의 개시내용에서 상세하게 기재된 바와 같이 수행될 수 있다.

[0095] 역학적 배제는 증폭을 개시하기 위해 상대적으로 느린 속도(예를 들어, 단편-어댑터 분자의 제1 카피의 느린 제조 속도) 대 단편-어댑터 분자의(또는 단편-어댑터 분자의 제1 카피의) 후속 카피를 제조하기 위해 상대적으로 신속한 속도를 이용할 수 있다. 이전 단락의 예에서, 역학적 배제는 단편-어댑터 분자 씨딩의 상대적으로 느린 속도(예를 들어, 상대적으로 느린 확산 또는 수송) 대 부위를 단편-어댑터 씨드의 카피로 채우기 위한 증폭이 일어나는 상대적으로 신속한 속도로 인하여 일어난다. 또 다른 예시적인 실시형태에서, 역학적 배제는 부위에 씨딩된 단편-어댑터 분자의 제1 카피의 형성의 지연(예를 들어, 지연된 또는 느린 활성화) 대 부위를 채우기 위하여 후속 카피가 제조되는 상대적으로 신속한 속도로 인해 일어날 수 있다. 이러한 예에서, 개별 부위는 몇몇의 상이한 단편-어댑터 분자로 씨딩될 수 있다(예를 들어, 몇몇 단편-어댑터 분자가 증폭 이전에 각각의 부위에서 존재할 수 있다). 그러나, 임의의 주어진 단편-어댑터 분자를 위한 제1 카피 형성은 무작위로 활성화되어, 제1 카피 형성의 평균 속도가, 후속 카피가 생성되는 속도에 비하여 상대적으로 느리게 할 수 있다. 이러한 경우, 개별 부위가 몇몇 상이한 단편-어댑터 분자로 씨딩될 수 있지만, 역학적 배제는 그들 단편-어댑터 분자 중 오직 하나가 증폭되게 할 것이다. 더욱 구체적으로, 일단 제1 단편-어댑터 분자가 증폭을 위해서 활성화되면, 그 부위는 그의 카피로 용량까지 신속하게 채워져 제2 단편-어댑터 분자의 카피가 그 부위에서 제조되는 것을 방지할 것이다.

[0096] 증폭 시약은 앰플리콘 형성을 용이하게 하고, 일부 경우에 앰플리콘 형성의 속도를 증가시키는 추가 성분을 포함할 수 있다. 예는 재조합효소이다. 재조합효소는 반복된 침습/신장을 허용함으로써 앰플리콘 형성을 용이하게 할 수 있다. 보다 구체적으로, 재조합효소는 중합효소에 의한 단편-어댑터 분자의 침습 및 앰플리콘 형성을 위한 주형으로서 단편-어댑터 분자를 사용한 중합효소에 의한 프라이머의 신장을 용이하게 할 수 있다. 이러한 과정은, 앰플리콘이 침습/신장의 각각의 라운드로부터 생성되어, 후속 라운드에서 주형으로서 제공되는 연쇄 반응으로서 반복될 수 있다. 이러한 과정은 표준 PCR보다 더 신속하게 일어날 수 있는데, 그 이유는 (예를 들어, 가열 또는 화학적 변성을 통한) 변성 사이클이 필요하지 않기 때문이다. 이와 같이, 재조합효소-촉진된 증폭은 등온적으로 수행될 수 있다. 증폭을 용이하게 하기 위해 재조합효소-촉진된 증폭 시약 중에 ATP 또는 다른 뉴클레오티드(또는 일부 경우에 그의 비-가수분해 가능한 유사체)를 포함시키는 것이 일반적으로 바람직하다. 재조합효소와 단일 가닥 결합(SSB) 단백질의 혼합물이 특히 유용한데, 그 이유는 SSB가 증폭을 추가로 용이하게 할 수 있기 때문이다. 재조합효소-촉진된 증폭을 위한 예시적인 제형은 트위스트디엑스(TwistDx)(영국 캠브릿지 소재)에 의해 트위스트앰프(TwistAmp) 키트로서 상업적으로 판매되는 것을 포함한다. 재조합효소-촉진된 증폭 시

약의 유용한 성분 및 반응 조건은 US 5,223,414호 및 US 7,399,590호에 언급되어 있다.

[0097] 앰플리콘 형성을 용이하게 하고, 일부 경우에 앰플리콘 형성 속도를 증가시키기 위해 증폭 시약 중에 포함될 수 있는 성분의 또 다른 예는 헬리카제이다. 헬리카제는 앰플리콘 형성의 연쇄 반응을 가능하게 함으로써 앰플리콘 형성을 용이하게 할 수 있다. 이러한 과정은 표준 PCR보다 더 신속하게 일어날 수 있는데, 그 이유는 (예를 들어, 가열 또는 화학적 변성을 통한) 변성 사이클이 필요하지 않기 때문이다. 이와 같이, 헬리카제-촉진된 증폭은 등온적으로 수행될 수 있다. 헬리카제와 단일 가닥 결합(SSB) 단백질의 혼합물이 특히 유용한데, 그 이유는 SSB가 증폭을 추가로 용이하게 할 수 있기 때문이다. 헬리카제-촉진된 증폭을 위한 예시적인 제형은 바이오헬릭스(Biohelix)(미국 매사추세츠주 비벌리 소재)로부터 아이소앰프(IsoAmp) 키트로서 상업적으로 판매되는 것을 포함한다. 추가로, 헬리카제 단백질을 포함하는 유용한 제형의 예는 각각이 본 명세서에 참조로 포함되는 US 7,399,590호 및 US 7,829,284호에 기재되어 있다.

[0098] 앰플리콘 형성을 용이하게 하고, 일부 경우에 앰플리콘 형성 속도를 증가시키기 위해 증폭 시약 중에 포함될 수 있는 성분의 또 다른 예는 원점 결합 단백질이다.

[0099] 표면으로의 단편-어댑터 분자의 부착 후에, 고정화 및 증폭된 단편-어댑터 분자의 서열이 결정된다. 서열결정은 임의의 적합한 서열결정 기법을 사용하여 수행될 수 있으며, 가닥 재합성을 비롯하여, 고정화되고 증폭된 단편-어댑터 분자의 서열을 결정하기 위한 방법은 해당 분야에 공지되어 있고, 예를 들어, 문헌[Bignell et al.](US 8,053,192호), 문헌[Gunderson et al.](WO2016/130704호), 문헌[Shen et al.](US 8,895,249호), 및 문헌[Pipenburg et al.](US 9,309,502호)에 기재되어 있다.

[0100] 본 명세서에 기재된 방법은 다양한 핵산 서열결정 기법과 함께 사용될 수 있다. 특히 적용 가능한 기법은 핵산이 어레이 내의 고정된 위치에 부착되어 그들의 상대적인 위치가 변화되지 않고, 그 어레이가 반복적으로 영상화되는 것이다. 예를 들어, 하나의 뉴클레오티드 염기 유형을 서로 구별하는 데 사용되는 상이한 표지와 일치하는, 상이한 색상 채널에서 영상이 수득되는 실시형태가 특히 적용 가능하다. 일부 실시형태에서, 단편-어댑터 분자의 뉴클레오티드 서열을 결정하기 위한 과정은 자동화 과정일 수 있다. 바람직한 실시형태는 합성에 의한 서열결정("SBS") 기법을 포함한다.

[0101] SBS 기법은 일반적으로 주형 가닥에 대한 뉴클레오티드의 반복적으로 부가를 통한 초기 핵산 가닥의 효소적 신장을 포함한다. 전통적인 SBS 방법에서, 단일 뉴클레오티드 단량체가 각각의 전달에서 중합효소의 존재 하에 표적 뉴클레오티드에 제공될 수 있다. 그러나, 본 명세서에 기재된 방법에서, 하나 초과의 유형의 뉴클레오티드 단량체가 전달에서 중합효소의 존재 하에 표적 핵산에 제공될 수 있다.

[0102] 일 실시형태에서, 뉴클레오티드 단량체는 잠금 핵산(LNA) 또는 브릿지된 핵산(BNA)을 포함한다. 시토신-고갈된 뉴클레오티드 서열이 트랜스포좀 복합체로부터의 전달된 가닥에 존재하는 경우의 결과와 같이, 단편-어댑터 분자가 하나 이상의 시토신-고갈된 뉴클레오티드 서열을 사용하여 생성되는 경우, 시토신-고갈된 영역에 혼성화하는 뉴클레오티드 단량체의 용융 온도가 변경된다. 뉴클레오티드 단량체에서 LNA 또는 BNA의 사용은 고정화된 단편-어댑터 분자 상에 존재하는 서열결정 프라이머 서열과 뉴클레오티드 단량체 간의 혼성화 강도를 증가시킨다.

[0103] SBS는 종결자 모이어티를 갖는 뉴클레오티드 단량체 또는 임의의 종결자 모이어티가 결여된 것들을 사용할 수 있다. 종결자가 결여된 뉴클레오티드 단량체를 사용하는 방법은 하기에 추가로 상세하게 언급된 바와 같이, 예를 들어, γ-포스페이트-표지된 뉴클레오티드를 사용한 서열결정 및 파이로시퀀싱(pyrosequencing)을 포함한다. 종결자가 결여된 뉴클레오티드 단량체를 사용하는 방법에서, 각각의 사이클에 추가되는 뉴클레오티드의 수는 일반적으로 가변적이고, 주형 서열 및 뉴클레오티드 전달 방식에 좌우된다. 종결자 모이어티를 갖는 뉴클레오티드 단량체를 사용하는 SBS 기법에 있어서, 종결자는 디데옥시뉴클레오티드를 사용하는 전통적인 생거(Sanger) 서열결정에 대한 경우에서와 같이 사용되는 서열결정 조건 하에서 효과적으로 비가역적일 수 있거나, 종결자는 솔렉사(Solexa)(현재 일루미나, 인코포레이티드)에 의해 개발된 서열결정 방법에 대한 경우에서와 같이 가역적일 수 있다.

[0104] SBS 기법은 표지 모이어티를 갖는 뉴클레오티드 단량체 또는 표지 모이어티가 결여된 것들을 사용할 수 있다. 따라서, 혼입 사례는 표지의 특징, 예컨대, 표지의 형광; 뉴클레오티드 단량체의 특징, 예컨대, 분자량 또는 전하; 뉴클레오티드의 혼입의 부산물, 예컨대, 피로포스페이트의 방출 등을 기반으로 검출될 수 있다. 2개 이상의 상이한 뉴클레오티드가 서열결정 시약 중에 존재하는 실시형태에서, 상이한 뉴클레오티드는 서로 구별 가능할 수 있거나, 대안적으로 2개 이상의 상이한 표지는 사용 중인 검출 기법 하에서 구별 가능하지 않을 수 있다. 예를 들어, 서열결정 시약 중에 존재하는 상이한 뉴클레오티드는 상이한 표지를 가질 수 있고, 그들은 솔렉사(현

재 일루미나, 인코포레이티드)에 의해 개발된 서열결정 방법에 의해서 예시된 바와 같이 적절한 광학 장치를 사용하여 구별될 수 있다.

[0105] 바람직한 실시형태는 파이로시퀀싱 기법을 포함한다. 파이로시퀀싱은 특정 뉴클레오티드가 초기 가닥 내에 혼입될 때 무기 파이로포스페이트(PPi)의 방출을 검출한다(문헌[Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. (1996) "Real-time DNA sequencing using detection of pyrophosphate release." Analytical Biochemistry 242(1), 84-9]; 문헌[Ronaghi, M. (2001) "Pyrosequencing sheds light on DNA sequencing." Genome Res. 11(1), 3-11]; 문헌[Ronaghi, M., Uhlen, M. and Nyren, P. (1998) "A sequencing method based on real-time pyrophosphate." Science 281(5375), 363]; 미국 특허 제6,210,891호; 제6,258,568호 및 제6,274,320호). 파이로시퀀싱에서, 방출된 PPi는 ATP 설퍼라제에 의해 아데노신 트리포스페이트(ATP)로 즉시 전환됨으로써 검출될 수 있고, 생성된 ATP의 수준은 루시퍼라제-생성된 광자를 통해 검출된다. 서열결정될 핵산은 어레이 내의 특징부에 부착될 수 있고, 어레이를 영상화하여, 어레이의 특징부에서의 뉴클레오티드의 혼입으로 인해 생성되는 화학발광 신호를 캡처할 수 있다. 영상은 어레이를 특정 뉴클레오티드 유형(예를 들어, A, T, C 또는 G)으로 처리한 후 얻어질 수 있다. 각각의 뉴클레오티드 유형의 부가 후에 수득되는 영상은 어레이 내의 어떤 특징부가 검출되는 지에 관하여 상이할 것이다. 영상의 이들 차이는 어레이 상의 특징부의 상이한 서열 내용을 반영한다. 그러나, 각각의 특징부의 상대적인 위치는 영상에서 변하지 않고 유지될 것이다. 영상은 본 명세서에 언급된 방법을 사용하여 저장, 처리 및 분석될 수 있다. 예를 들어, 각각의 상이한 뉴클레오티드 유형으로 어레이를 처리한 후 수득되는 영상은 가역적 종결자-기반의 서열결정 방법을 위해 상이한 검출 채널로부터 수득되는 영상에 대하여 본 명세서에 예시된 것과 동일한 방식으로 취급될 수 있다.

[0106] SBS의 또 다른 예시적인 유형에서, 순환 서열결정은 예를 들어, 개시내용이 본 명세서에 참조로 포함되는 WO 04/018497호 및 미국 특허 제7,057,026호에 기재된 바와 같이 예를 들어, 절단 가능한 또는 광표백 가능한 염료 표지를 함유하는 가역적 종결자 뉴클레오티드의 단계식 첨가에 의해서 달성된다. 이러한 접근법은 솔렉사(현재 일루미나 인코포레이티드)에 의해 상업화되고 있고, 또한 WO 91/06678호 및 WO 07/123,744호에 기재되어 있다. 종결이 역전될 수 있고, 형광 표지가 절단될 수 있는 형광-표지된 종결자의 이용 가능성은 효율적인 순환식 가역적 종결(cyclic reversible termination: CRT) 서열결정을 용이하게 한다. 중합효소는 또한 이들 변형된 뉴클레오티드로부터 효율적으로 혼입 및 신장하도록 동시-조작될 수 있다.

[0107] 바람직하게는 가역적 종결자-기반의 서열결정 실시형태에서, 표지는 SBS 반응 조건 하에서 신장을 실질적으로 억제하지 않는다. 그러나, 검출 표지는 예를 들어, 절단 또는 분해에 의해 제거 가능할 수 있다. 영상은 어레이 된 핵산 특징부 내로의 표지의 혼입 이후에 캡처될 수 있다. 특정 실시형태에서, 각각의 사이클은 어레이에 4개의 상이한 뉴클레오티드 유형을 동시에 전달하는 것을 포함하고, 각각의 뉴클레오티드 유형은 스펙트럼적으로 별개의 표지를 갖는다. 이어서 각각 4개의 상이한 표지 중 하나에 대하여 선택적인 검출 채널을 사용하여 4개의 영상이 수득될 수 있다. 대안적으로, 상이한 뉴클레오티드 유형은 순차적으로 첨가될 수 있고, 어레이의 영상은 각각의 첨가 단계 사이에 수득될 수 있다. 이러한 실시형태에서, 각각의 영상은 특정 유형의 뉴클레오티드가 혼입된 핵산 특징부를 보여줄 것이다. 상이한 특징부는 각각의 특징부의 상이한 서열 내용으로 인하여 상이한 영상에 존재하거나 존재하지 않을 것이다. 그러나, 특징부의 상대적인 위치는 영상에서 변하지 않고 유지될 것이다. 이러한 가역적 종결자-SBS 방법으로부터 수득된 영상은 본 명세서에 언급된 바와 같이 저장, 처리 및 분석될 수 있다. 영상 캡처 단계 이후에, 표지를 제거할 수 있고, 가역적 종결자 모이어티를 뉴클레오티드 첨가 및 검출의 후속 사이클을 위해 제거할 수 있다. 표지가 특정 사이클에서 검출된 이후에 그리고 후속 사이클 이전에 표지를 제거하는 것은 백그라운드 신호 및 사이클 사이의 크로스토크(crosstalk)를 감소시키는 이점을 제공할 수 있다. 유용한 표지 및 제거 방법의 예는 하기에 언급되어 있다.

[0108] 특정 실시형태에서, 뉴클레오티드 단량체의 일부 또는 전부는 가역적 종결자를 포함할 수 있다. 이러한 실시형태에서, 가역적 종결자/절단 가능한 형광단은 3' 에스테르 연결을 통해 리보스 모이어티에 연결된 형광단을 포함할 수 있다(문헌[Metzker, Genome Res. 15:1767-1776 (2005)]). 다른 접근법은 종결자 화학물질을 형광 표지의 절단로부터 분리한 바 있다(문헌[Ruparel et al., Proc Natl Acad Sci USA 102: 5932-7 (2005)]). 루파렐(Ruparel) 등은 작은 3' 알릴기를 사용하여 신장을 차단하지만, 팔라듐 촉매로의 짧은 처리에 의해 용이하게 해제될 수 있는 가역적 종결자의 개발을 기재하였다. 형광단을 장파장 UV광으로의 30초 노출에 의해 용이하게 절단될 수 있는 광 절단 가능한 링커를 통해 염기에 부착시켰다. 따라서, 이황화물 환원 또는 광 절단 중 어느 하나를 절단 가능한 링커로서 사용할 수 있다. 가역적 종결에 대한 또 다른 접근법은 dNTP 상에 체적이 큰 염료를 배치한 후 일어나는 자연 종결의 사용이다. dNTP 상의 하전된 체적이 큰 염료의 존재는 입체 및/또는 정전기 장애를 통해 효과적인 종결자로서 작용할 수 있다. 하나의 혼입 사례의 존재는 염료가 제거되지 않는 한 추가 혼

입을 방지한다. 염료의 절단은 형광단을 제거하고, 종결을 효과적으로 역전시킨다. 변형된 뉴클레오티드의 예는 또한 개시내용의 전문이 본 명세서에 참조로 포함되는 미국 특허 제7,427,673호 및 제7,057,026호에 기재되어 있다.

[0109] 본 명세서에 기재된 방법 및 시스템과 함께 사용될 수 있는 추가의 예시적인 SBS 시스템 및 방법은 미국 특허 공개 제2007/0166705호, 제2006/0188901호, 제2006/0240439호, 제2006/0281109호, 제2012/0270305호, 및 제2013/0260372호, 미국 특허 제7,057,026호, PCT 공개 제WO 05/065814호, 미국 특허 출원 공개 제2005/0100900호, 및 PCT 공개 제WO 06/064199호, 및 제WO 07/010,251호에 기재되어 있다.

[0110] 일부 실시형태는 4개 미만의 상이한 표지를 사용하여 4개의 상이한 뉴클레오티드의 검출을 사용할 수 있다. 예를 들어, SBS는 미국 특허 공개 제2013/0079232호에 포함된 문헌에 기재된 방법 및 시스템을 사용하여 수행될 수 있다. 첫 번째 예로서, 한 쌍의 뉴클레오티드 유형이 동일한 파장에서 검출될 수 있지만, 다른 것과 비교하여 그 쌍 중 하나의 구성원에 대한 강도의 차이를 기반으로 또는 그 쌍의 다른 구성원에 대하여 검출된 신호에 비하여 명백한 신호가 나타나게 하거나 사라지게 하는 그 쌍의 하나의 구성원의 변화(예를 들어, 화학적 변형, 광화학적 변형 또는 물리적 변형을 통함)를 기반으로 구별될 수 있다. 두 번째 예로서, 4개의 상이한 뉴클레오티드 유형 중 3개가 특정 조건 하에서 검출될 수 있는 반면, 제4 뉴클레오티드 유형은 그러한 조건 하에서 검출 가능한 표지가 결여되어 있거나 또는 그러한 조건 하에서 최소한으로 검출된다(예를 들어, 백그라운드 형광 등으로 인한 최소 검출). 핵산 내로의 처음 3개의 뉴클레오티드 유형의 혼입은 그들의 각각의 신호의 존재에 기초하여 결정될 수 있고, 핵산 내로의 제4 뉴클레오티드 유형의 혼입은 임의의 신호의 부재 또는 최소 검출에 기초하여 결정될 수 있다. 세 번째 예로서, 하나의 뉴클레오티드 유형은 2개의 상이한 채널에서 검출되는 표지(들)를 포함할 수 있는 반면, 다른 뉴클레오티드 유형은 채널 중 하나 이하에서 검출된다. 상기에 언급된 3개의 예시적인 구성은 상호 배타적인 것으로 간주되지 않고, 다양한 조합으로 사용될 수 있다. 모든 3개의 예를 조합한 예시적인 실시형태는 제1 채널에서 검출되는 제1 뉴클레오티드 유형(예를 들어, 제1 여기 파장에 의해 여기되는 경우 제1 채널에서 검출되는 표지를 갖는 dATP), 제2 채널에서 검출되는 제2 뉴클레오티드 유형(예를 들어, 제2 여기 파장에 의해 여기되는 경우 제2 채널에서 검출되는 표지를 갖는 dCTP), 제1 채널 및 제2 채널 둘 모두에서 검출되는 제3 뉴클레오티드 유형(예를 들어, 제1 여기 파장 및/또는 제2 여기 파장에 의해 여기되는 경우 두 채널 모두에서 검출되는 적어도 하나의 표지를 갖는 dTTP) 및 어느 하나의 채널에서도 검출되지 않거나 최소한으로 검출되는 표지가 결여된 제4 뉴클레오티드 유형(예를 들어, 표지를 갖지 않는 dGTP)을 사용하는 형광-기반의 SBS 방법이다.

[0111] 추가로, 미국 특허 공개 제2013/0079232호에 포함된 문헌에 기재된 바와 같이, 서열결정 데이터는 단일 채널을 사용하여 수득될 수 있다. 이러한 소위 1-염료 서열결정 접근법에서, 제1 뉴클레오티드 유형은 표지되지만, 표지는 제1 영상이 생성된 후에 제거되고, 제2 뉴클레오티드 유형은 제1 영상이 생성된 후에만 표지된다. 제3 뉴클레오티드 유형은 제1 영상 및 제2 영상 둘 모두에서 그의 표지를 보유하고, 제4 뉴클레오티드 유형은 두 영상 모두에서 표지되지 않은 채로 유지된다.

[0112] 일부 실시형태는 라이게이션 기법에 의한 서열결정을 사용할 수 있다. 이러한 기법은 DNA 리가제를 사용하여 올리고뉴클레오티드를 혼입시키고, 이러한 올리고뉴클레오티드의 혼입을 식별한다. 올리고뉴클레오티드는 전형적으로 올리고뉴클레오티드가 혼성화하는 서열 내의 특정 뉴클레오티드의 동일성과 상관관계가 있는 상이한 표지를 갖는다. 다른 SBS 방법에서와 같이, 영상은 핵산 특징부의 어레이를 표지된 서열결정 시약으로 처리한 후 수득될 수 있다. 각각의 영상은 특정 유형의 표지가 혼입된 핵산 특징부를 나타낼 것이다. 상이한 특징부는 각각의 특징부의 상이한 서열 내용으로 인하여 상이한 영상에서 존재하거나 존재하지 않을 것이지만, 특징부의 상대적인 위치는 영상에서 변하지 않고 유지될 것이다. 라이게이션-기반의 서열결정 방법으로부터 수득되는 영상은 본 명세서에 언급된 바와 같이 저장, 처리 및 분석될 수 있다. 본 명세서에 기재된 방법 및 시스템과 함께 사용될 수 있는 예시적인 SBS 시스템 및 방법은 미국 특허 제6,969,488호, 제6,172,218호 및 제6,306,597호에 기재되어 있다.

[0113] 일부 실시형태는 나노포어 서열결정을 사용할 수 있다(문헌[Deamer, D. W. & Akeson, M. "Nanopores and nucleic acids: prospects for ultrarapid sequencing." Trends Biotechnol. 18, 147-151 (2000)]; 문헌[Deamer, D. and D. Branton, "Characterization of nucleic acids by nanopore analysis", Acc. Chem. Res. 35:817-825 (2002)]; 문헌[Li, J., M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko, "DNA molecules and configurations in a solid-state nanopore microscope" Nat. Mater. 2:611-615 (2003)]). 이러한 실시형태에서, 단편-어댑터 분자가 나노포어를 통과한다. 나노포어는 합성 포어 또는 생물학적 막 단백질, 예컨대, α-용혈소일 수 있다. 단편-어댑터 분자가 나노포어를 통과하기 때문에, 각각의 염기-쌍은 포어의 전기

전도도의 변동을 측정함으로써 식별될 수 있다(개시내용의 전문이 본 명세서에 참조로 포함되는 미국 특허 제7,001,792호; 문헌[Soni, G. V. & Meller, "A. Progress toward ultrafast DNA sequencing using solid-state nanopores." Clin. Chem. 53, 1996-2001 (2007)]; 문헌[Healy, K. "Nanopore-based single-molecule DNA analysis." Nanomed. 2, 459-481 (2007)]; 문헌[Cockroft, S. L., Chu, J., Amorin, M. & Ghadiri, M. R. "A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution." J. Am. Chem. Soc. 130, 818-820 (2008)]). 나노포어 서열결정으로부터 수득되는 데이터는 본 명세서에 언급된 바와 같이 저장, 처리 및 분석될 수 있다. 특히, 데이터는 본 명세서에 언급된 광학 영상 및 다른 영상의 예시적인 처리에 따른 영상으로서 처리될 수 있다.

[0114] 일부 실시형태는 DNA 중합효소 활성의 실시간 모니터링을 포함하는 방법을 사용할 수 있다. 뉴클레오티드 혼입은 예를 들어, 미국 특허 제7,329,492호 및 제7,211,414호에 기재된 바와 같이 형광단-보유 중합효소와 γ-포스페이트-표지된 뉴클레오티드 간의 형광 공명 에너지 전달(FRET) 상호작용을 통해서 검출될 수 있거나, 또는 뉴클레오티드 혼입은 예를 들어, 미국 특허 제7,315,019호에 기재된 바와 같은 제로-모드 도파관으로 그리고 예를 들어, 둘 모두가 본 명세서에 참조로 포함되는 미국 특허 제7,405,281호 및 미국 특허 공개 제2008/0108082호에 기재된 바와 같은 형광 뉴클레오티드 유사체 및 조작된 중합효소를 사용하여 검출될 수 있다. 조명은 표면-테더링된 중합효소 주변의 젭토리터-스케일 체적으로 제한될 수 있어서, 형광 표지된 뉴클레오티드의 혼입은 낮은 백그라운드와 함께 관찰될 수 있다(문헌[Levene, M. J. et al. "Zero-mode waveguides for single-molecule analysis at high concentrations." Science 299, 682-686 (2003)]; 문헌[Lundquist, P. M. et al. "Parallel confocal detection of single molecules in real time." Opt. Lett. 33, 1026-1028 (2008)]; 문헌[Korlach, J. et al. "Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nano structures." Proc. Natl. Acad. Sci. USA 105, 1176-1181 (2008)]). 이러한 방법으로부터 수득되는 영상은 본 명세서에 언급된 바와 같이 저장, 처리 및 분석될 수 있다.

[0115] 일부 SBS 실시형태는 신장 생성물 내로의 뉴클레오티드의 혼입 시에 방출되는 양성자의 검출을 포함한다. 예를 들어, 방출된 양성자의 검출에 기초한 서열결정은 이온 토렌트(Ion Torrent)(미국 코네티컷주 소재의 길포드(Guilford), 라이프 테크놀로지스(Life Technologies) 자회사)로부터 상업적으로 입수 가능한 전기 검출기 및 관련 기법 또는 미국 특허 공개 제2009/0026082호; 제2009/0127589호; 제2010/0137143호; 및 제2010/0282617호에 기재된 서열결정 방법 및 시스템을 사용할 수 있다. 표적 핵산을 역학적 배제를 사용하여 증폭시키기 위한 본 명세서에 언급된 방법은 양성자를 검출하는 데 사용되는 기질에 용이하게 적용될 수 있다. 보다 구체적으로, 본 명세서에 언급된 방법을 사용하여 양성자를 검출하는 데 사용되는 앰플리콘의 클론 집단을 생성할 수 있다.

[0116] 상기 SBS 방법은 멀티플렉스 포맷으로 유리하게 수행되어 다수의 상이한 단편-어댑터 분자가 동시에 조작되게 할 수 있다. 특정 실시형태에서, 상이한 단편-어댑터 분자는 일반적인 반응 용기 내에서 또는 특정 기판의 표면 상에서 처리될 수 있다. 이것은 멀티플렉스 방식으로 서열결정 시약의 편리한 전달, 미반응 시약의 제거 및 혼입 사례의 검출을 가능하게 한다. 표면-결합된 표적 핵산을 사용한 실시형태에서, 단편-어댑터 분자는 어레이 포맷에 존재할 수 있다. 어레이 포맷에서, 단편-어댑터 분자는 전형적으로 공간적으로 구별 가능한 방식으로 표면에 결합될 수 있다. 단편-어댑터 분자 직접적인 공유적 부착, 비드 또는 다른 입자로의 부착 또는 표면에 부착된 중합효소 또는 다른 분자로의 결합에 의해 결합될 수 있다. 어레이는 각각의 부위(특징부로도 지칭됨)에서 단편-어댑터 분자의 단일 카피를 포함할 수 있거나 또는 동일한 서열을 갖는 다수의 카피가 각각의 부위 또는 특징부에 존재할 수 있다. 다수의 카피는 하기에 추가로 상세하게 기재된 바와 같은 증폭 방법, 예컨대, 브릿지 증폭 또는 에멀젼 PCR에 의해 생성될 수 있다.

[0117] 본 명세서에 언급된 방법은, 예를 들어, 적어도 약 10개의 특징부/㎠, 100개의 특징부/㎠, 500개의 특징부/㎠, 1,000개의 특징부/㎠, 5,000개의 특징부/㎠, 10,000개의 특징부/㎠, 50,000개의 특징부/㎠, 100,000개의 특징부/㎠, 1,000,000개의 특징부/㎠, 5,000,000개의 특징부/㎠, 또는 그 초과를 비롯한 다양한 밀도 중 임의의 것으로 특징부를 갖는 어레이를 사용할 수 있다.

[0118] 본 명세서에 언급된 방법의 이점은, 그들이 복수의 ㎠의 신속하고 효율적인 검출을 동시에 제공한다는 점이다. 따라서, 본 개시내용은 해당 분야에 공지된 기법, 예컨대, 상기에 예시된 것들을 사용하여 핵산을 제조 및 검출할 수 있는 통합된 시스템을 제공한다. 따라서, 본 개시내용의 통합된 시스템은 증폭 시약 및/또는 서열결정 시약을 하나 이상의 고정화된 DNA 단편에 전달할 수 있는 유체 성분을 포함할 수 있고, 시스템은 성분, 예컨대, 펌프, 밸브, 저장소, 유체 라인 등을 포함한다. 유동 셀은 표적 핵산의 검출을 위하여 통합된 시스템에서 구성되고/구성되거나 사용될 수 있다. 예시적인 유동 셀은 예를 들어, 미국 특허 공개 제2010/0111768호 및 미국 가특허 제13/273,666호에 기술되어 있다. 유동 셀에 대해서 예시된 바와 같이, 통합된 시스템의 유체 성분 중 하

나 이상이 증폭 방법 및 검출 방법을 위해 사용될 수 있다. 예로서 핵산 서열결정 실시형태를 고려하여, 통합된 시스템의 유체 성분 중 하나 이상이 본 명세서에 언급된 증폭 방법을 위해 그리고 서열결정 방법, 예컨대 상기에 예시된 것들에서 서열결정 시약의 전달을 위해 사용될 수 있다. 대안적으로, 통합된 시스템은 증폭 방법을 수행하고, 검출 방법을 수행하기 위한 개별 유체 시스템을 포함할 수 있다. 증폭된 핵산을 생성할 수 있고, 또한 핵산의 서열을 결정할 수 있는 통합된 서열결정 시스템의 예는 비제한적으로 MiSeqTM 플랫폼(미국 캘리포니아주 샌디에고 소재의 일루미나, 인코포레이티드) 및 본 명세서에 참조로 포함되는 미국 가특허 제13/273,666호에 기재된 디바이스를 포함한다.

[0119] 본 명세서에 기재된 방법의 실시 동안, 다양한 조성물이 초래될 수 있다. 예를 들어, 도 2 블록 vii 또는 도 4에 나타낸 구조를 갖는 이중-인덱스 단편-어댑터 분자를 포함하는 이중-인덱스 단편-어댑터 분자, 및 이중-인덱스 단편-어댑터 분자를 포함하는 조성물이 초래될 수 있다. 도 2 블록 vii 또는 도 4에 나타낸 구조를 갖는 이중-인덱스 단편-어댑터 분자를 포함하는 이중-인덱스 단편-어댑터 분자의 서열결정 라이브러리, 및 서열결정 라이브러리를 포함하는 조성물이 초래될 수 있다. 이러한 서열결정 라이브러리는 어레이에 결합될 수 있다.

[0120] 본 발명은 하기 실시예에 의해 예시된다. 특정 실시예, 물질, 양 및 절차는 본 명세서에 언급된 바와 같은 본 발명의 범주 및 목적에 따라 넓게 해석되어야 한다는 것이 이해되어야 한다.

[0121] 실시예

[0122] 실시예에 사용되는 시약

[0123] · 포스페이트 완충 염수(PBS, 써모 피셔, 카탈로그 10010023)

[0124] · 0.25% 트립신(써모 피셔, 카탈로그 15050057)

[0125] · 트리스(피셔, 카탈로그 T1503)

[0126] · HCl(피셔, 카탈로그 A144)

[0127] · NaCl(피셔, 카탈로그 M-11624)

[0128] · MgCl2(시그마(Sigma), 카탈로그 M8226)

[0129] · Igepal® CA-630(시그마, I8896)

[0130] · 프로테아제 억제제(로슈(Roche), 카탈로그 11873580001)

[0131] · PCR-클린(Clean) ddH20

[0132] · 리튬 3,5-다이아이오도살리실산(시그마, 카탈로그 D3635) - LAND 방법만

[0133] · 포름알데히드(시그마, 카탈로그 F8775) - xSDS 방법만

[0134] · 글리신(시그마, 카탈로그 G8898) - xSDS 방법만

[0135] · NEBuffer 2.1(NEB, 카탈로그 B7202) - xSDS 방법만

[0136] · SDS(시그마, 카탈로그 L3771) - xSDS 방법만

[0137] · 트리톤(Triton)™ X-100(시그마, 카탈로그 9002-93-1) - xSDS 방법만

[0138] · DAPI(써모 피셔, 카탈로그 D1306)

[0139] · 넥스테라(Nextera)® 키트로부터의 TD 완충제(일루미나, 카탈로그 FC-121-1031)

[0140] · 96 인덱싱 시토신-고갈된 트랜스포좀(공개된 방법을 사용하여 어셈블됨, **표 1**에 나타낸 서열)

[0141] · 9-뉴클레오티드 무작위 프라이머(**표 2**)

[0142] · 10 mM dNTP 믹스(NEB, 카탈로그 N0447)

[0143] · 클레나우(3'->5' 엑소-) 중합효소(엔자이머틱스(Enzymatics), 카탈로그 P7010-LC-L)

[0144] · 200 프루프(Proof) 에탄올

[0145] · 인덱싱된 i5 및 i7 PCR 프라이머(**표 3**)

[0146] · 카파(Kapa) HiFi™ 핫스타트(HotStart) 레디믹스(ReadyMix)

[0147] · SYBR® 그린(Green)(FMC 바이오프로덕츠(BioProducts), 카탈로그 50513)

[0148] · 퀴아퀵(QIAquick)® PCR 정제 키트(퀴아젠(Qiagen), 카탈로그 28104)

[0149] · dsDNA 고 감도 큐비트(Qubit)®(써모 피셔, 카탈로그 Q32851)

[0150] · 고 감도 생물분석기 키트(아질런트(Agilent), 카탈로그 5067-4626)

[0151] · NextSeq 서열결정 키트(고 또는 중등 150-사이클)

[0152] · 비메틸화된 람다 DNA(프로메가(Promega), 카탈로그 D1521)

[0153] · HiSeq® 2500 서열결정 키트(일루미나)

[0154] · HiSeq® X 서열결정 키트(일루미나)

[0155] · EZ-96 DNA 메틸화 MagPrep 키트(자이모 리서치(Zymo Research), 카탈로그 D5040)

[0156] · 커스텀(Custom) LNA 서열결정 프라이머(**표 4**)

[0157] · 폴리에틸렌 글리콜(PEG)

[0158] · SPRI 비드

[0159] 실시예에 사용되는 장비

[0160] · 35㎛M 세포 스트레이너(Cell Strainer)(비디 바이오사이언스즈(BD Biosciences), 카탈로그 352235)

[0161] · 96-웰 플레이트 호환 가능한 자성 랙(rack)

[0162] · 소니(Sony) SH800 세포 분리기(소니 바이오테크놀로지(Sony Biotechnology), 카탈로그 SH800) 또는 DAPI 기반의 단일 핵 분리가 가능한 기타 FACS 기기

[0163] · CFX 커넥트(Connect) RT 열 순환기(바이오-라드(Bio-Rad), 카탈로그 1855200) 또는 기타 실시간 열순환기

[0164] · 열혼합기(Thermomixer)

[0165] · 큐비트® 2.0 형광계(Fluorometer)(써모 피셔, 카탈로그 Q32866)

[0166] · 2100 생물분석기(아질런트, 카탈로그 G2939A)

[0167] · NextSeq® 500(일루미나, 카탈로그 SY-415-1001-1)

[0168] · HiSeq® 2500(일루미나)

[0169] · HiSeq® X(일루미나)

실시예에 사용되는 올리고뉴클레오티드

표 1: sciMET 트랜스포사좀-로딩된 올리고(5'-3')

역 상보물 : (5phos)CTGTCTCTTATACACATCT

명칭	i5_bsPCR	인덱스	i5_Tn5
sciMET_Tn5_1	GGTGTAGTGGGTTTGG	GTTAAGAGGAA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_2	GGTGTAGTGGGTTTGG	AGTAGGAAGAT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_3	GGTGTAGTGGGTTTGG	GAATTAGGTGT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_4	GGTGTAGTGGGTTTGG	GGAGATTAATG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_5	GGTGTAGTGGGTTTGG	TATTGTGGAAT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_6	GGTGTAGTGGGTTTGG	ATATAGATGAT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_7	GGTGTAGTGGGTTTGG	GTAAGAGGAAT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_8	GGTGTAGTGGGTTTGG	GAGAGTTATTG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_9	GGTGTAGTGGGTTTGG	AGTTAGTGTGA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_10	GGTGTAGTGGGTTTGG	GATATAGAATT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_11	GGTGTAGTGGGTTTGG	AAGGAAGTGAA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_12	GGTGTAGTGGGTTTGG	AATAAGGAAGG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_13	GGTGTAGTGGGTTTGG	GTATGGATATA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_14	GGTGTAGTGGGTTTGG	TTAGATAATGA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_15	GGTGTAGTGGGTTTGG	GGTGTTGTAAT	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_16	GGTGTAGTGGGTTTGG	GAAGTGGAGAG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_17	GGTGTAGTGGGTTTGG	TTGAGTGGTAG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_18	GGTGTAGTGGGTTTGG	GATAATGGTGA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_19	GGTGTAGTGGGTTTGG	GTGTTAATGGA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_20	GGTGTAGTGGGTTTGG	TAGGAATGGTG	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG
sciMET_Tn5_21	GGTGTAGTGGGTTTGG	ATGTATGGATA	TGGTAGAGAGGGTG AGATGTGTATAAGAGACAG

[0170]

[0171]

sciMET_Tn5_22	GGTGTAGTGGGTTTGG	TGATTGTTGGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_23	GGTGTAGTGGGTTTGG	AAGAGAATTAT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_24	GGTGTAGTGGGTTTGG	AATGGTTGGTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_25	GGTGTAGTGGGTTTGG	GGTTAATTGAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_26	GGTGTAGTGGGTTTGG	GTATAATAGTT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_27	GGTGTAGTGGGTTTGG	TTAGTTGAATT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_28	GGTGTAGTGGGTTTGG	TTGGTGAAGGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_29	GGTGTAGTGGGTTTGG	TTAATATTGAA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_30	GGTGTAGTGGGTTTGG	GTTAGAATTGG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_31	GGTGTAGTGGGTTTGG	GTTATTAATTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_32	GGTGTAGTGGGTTTGG	GATTGGTAAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_33	GGTGTAGTGGGTTTGG	TGAAGTATTGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_34	GGTGTAGTGGGTTTGG	GATGGATTATG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_35	GGTGTAGTGGGTTTGG	ATTAGTATATT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_36	GGTGTAGTGGGTTTGG	GTAGGTGTGGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_37	GGTGTAGTGGGTTTGG	AGTTGAATGTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_38	GGTGTAGTGGGTTTGG	ATTGTGAGATA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_39	GGTGTAGTGGGTTTGG	TTGTGGTGAGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_40	GGTGTAGTGGGTTTGG	TTAAGTTGGTT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_41	GGTGTAGTGGGTTTGG	TATAATAATAT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_42	GGTGTAGTGGGTTTGG	AAGGTATGAGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_43	GGTGTAGTGGGTTTGG	AGGATTATAAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_44	GGTGTAGTGGGTTTGG	AGAGTTAGGTT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_45	GGTGTAGTGGGTTTGG	ATGGATAGTAT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_46	GGTGTAGTGGGTTTGG	ATATTATGTTG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_47	GGTGTAGTGGGTTTGG	GGTGGAGATAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_48	GGTGTAGTGGGTTTGG	TGGTGGTAGTG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_49	GGTGTAGTGGGTTTGG	AGGTGAGAAGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG

[0172]

sciMET_Tn5_50	GGTGTAGTGGGTTTGG	TAGGAGGTTGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_51	GGTGTAGTGGGTTTGG	TGTATAGGTAT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_52	GGTGTAGTGGGTTTGG	TGTTATGTAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_53	GGTGTAGTGGGTTTGG	TGGAAGGTATG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_54	GGTGTAGTGGGTTTGG	AATGTAAGGAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_55	GGTGTAGTGGGTTTGG	GTTATGTTAAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_56	GGTGTAGTGGGTTTGG	TGTTATAGGTG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_57	GGTGTAGTGGGTTTGG	AAGGAGAATTG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_58	GGTGTAGTGGGTTTGG	AGAGGTGGAAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_59	GGTGTAGTGGGTTTGG	GATTAGGTGTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_60	GGTGTAGTGGGTTTGG	ATTATATAAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_61	GGTGTAGTGGGTTTGG	GAGAATATGGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_62	GGTGTAGTGGGTTTGG	GGATTGAGAGG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_63	GGTGTAGTGGGTTTGG	ATTATGGTGGT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_64	GGTGTAGTGGGTTTGG	GAAGGAAGTTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_65	GGTGTAGTGGGTTTGG	GAATATGTAAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_66	GGTGTAGTGGGTTTGG	TAGTTAATATT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_67	GGTGTAGTGGGTTTGG	TGAATGAATAG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_68	GGTGTAGTGGGTTTGG	AGGATGGATTA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_69	GGTGTAGTGGGTTTGG	AAGTGTATAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_70	GGTGTAGTGGGTTTGG	GAGGTTGAAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_71	GGTGTAGTGGGTTTGG	TGTGTAATAGG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_72	GGTGTAGTGGGTTTGG	TTGATTAGAGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_73	GGTGTAGTGGGTTTGG	TATGTGTGTGG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_74	GGTGTAGTGGGTTTGG	GAGATGAGAAT	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_75	GGTGTAGTGGGTTTGG	TGGTGAAGTGA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_76	GGTGTAGTGGGTTTGG	GTGGTAGGATG	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_77	GGTGTAGTGGGTTTGG	TGTAGGTGATA	TGGTAGAGAGGGTG	AGATGTGTATAAGAGACAG

[0173]

sciMET_Tn5_78	GGTGTAGTGGGTTTGG	GTAAGGTGTGA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_79	GGTGTAGTGGGTTTGG	AGAAGAGAGTG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_80	GGTGTAGTGGGTTTGG	GGATGTTGTAT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_81	GGTGTAGTGGGTTTGG	AAGTTATATAA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_82	GGTGTAGTGGGTTTGG	TGGAATTAAGT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_83	GGTGTAGTGGGTTTGG	TAATGAGAGGA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_84	GGTGTAGTGGGTTTGG	ATAATTGATGG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_85	GGTGTAGTGGGTTTGG	TGTGAAGAGTA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_86	GGTGTAGTGGGTTTGG	GATGAATATGT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_87	GGTGTAGTGGGTTTGG	TGAGGATAGAT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_88	GGTGTAGTGGGTTTGG	ATTAATTAGAG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_89	GGTGTAGTGGGTTTGG	GGAGAGATGGA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_90	GGTGTAGTGGGTTTGG	TAATTGAGGAA	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_91	GGTGTAGTGGGTTTGG	TTGGAATTAAT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_92	GGTGTAGTGGGTTTGG	AATGTTATTGT	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_93	GGTGTAGTGGGTTTGG	GTAGTTATTAG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_94	GGTGTAGTGGGTTTGG	TATATTGTGAG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_95	GGTGTAGTGGGTTTGG	GTGTAGGATAG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG
sciMET_Tn5_96	GGTGTAGTGGGTTTGG	AGAGAAGTTGG	TGGTAGAGAGAGGGTG	AGATGTGTATAAGAGACAG

[0174]

표 2: sciMET 9-뉴클레오티드 무작위 프라이머(5'-3')

명칭	서열
sciMET_N9_IPE2	GGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNN

표 3: sciMET PCR 프라이머(5'-3')

명칭	서열
sciMET_i7_1	CAAGCAGAAGACGGCATACGAGATcaagatgccgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_2	CAAGCAGAAGACGGCATACGAGATaacgtctagtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_3	CAAGCAGAAGACGGCATACGAGATaggtatactcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_4	CAAGCAGAAGACGGCATACGAGATttcataggacGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_5	CAAGCAGAAGACGGCATACGAGATggaggcctccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_6	CAAGCAGAAGACGGCATACGAGATttcaatataaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_7	CAAGCAGAAGACGGCATACGAGATacgtcatataGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_8	CAAGCAGAAGACGGCATACGAGATttgaccaggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_9	CAAGCAGAAGACGGCATACGAGATcggttgcgcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_10	CAAGCAGAAGACGGCATACGAGATcaaggaggtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_11	CAAGCAGAAGACGGCATACGAGATttacgatgaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_12	CAAGCAGAAGACGGCATACGAGATttgctggcatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_13	CAAGCAGAAGACGGCATACGAGATaatactcttcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_14	CAAGCAGAAGACGGCATACGAGATccaactaaccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_15	CAAGCAGAAGACGGCATACGAGATtatcctcaatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_16	CAAGCAGAAGACGGCATACGAGATgccgtcgcgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_17	CAAGCAGAAGACGGCATACGAGATccgcttgcttcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_18	CAAGCAGAAGACGGCATACGAGATtgaccgaatcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_19	CAAGCAGAAGACGGCATACGAGATgtctccagagGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

[0175]

sciMET_i7_20	CAAGCAGAAGACGGCATACGAGATaatgctagtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_21	CAAGCAGAAGACGGCATACGAGATgacgacctgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_22	CAAGCAGAAGACGGCATACGAGATagagccagccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_23	CAAGCAGAAGACGGCATACGAGATccaggccgcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_24	CAAGCAGAAGACGGCATACGAGATcaggtatggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_1	AATGATACGGCGACCACCGAGATCTACACgtatcatcgaGGTGTAGTGGGTTTGG
sciMET_i7_2	AATGATACGGCGACCACCGAGATCTACACccgcgattatGGTGTAGTGGGTTTGG
sciMET_i7_3	AATGATACGGCGACCACCGAGATCTACACattccaggtacGGTGTAGTGGGTTTGG
sciMET_i7_4	AATGATACGGCGACCACCGAGATCTACACatggaattggGGTGTAGTGGGTTTGG
sciMET_i7_5	AATGATACGGCGACCACCGAGATCTACACgacgaagcgtGGTGTAGTGGGTTTGG
sciMET_i7_6	AATGATACGGCGACCACCGAGATCTACACcttgcagtagGGTGTAGTGGGTTTGG
sciMET_i7_7	AATGATACGGCGACCACCGAGATCTACACcttggtaatgGGTGTAGTGGGTTTGG
sciMET_i7_8	AATGATACGGCGACCACCGAGATCTACACcaagtcgaccGGTGTAGTGGGTTTGG

표 4: sciMET 서열결정 프라이머(LNA, 5'-3')

[0176]

명칭	서열
sciMET_Read1	TGGTAGAGAGGGTG AGATGTGTATAAGAGATAG
sciMET_Iindex1	CTATCTCTTATACACATCT CACCCTCTCTACCA

[0177] 실시예 1

[0178] 비메틸화된 대조군 람다 DNA의 제조

[0179] 100 나노그램의 비메틸화된 람다 DNA, 5㎕의 2X TD 완충제, 5㎕의 NIB 완충제(10mM Tris-HCl pH7.4, 10MM NaCl, 3mM $MgCl_2$, 0.1% Igepal®, 1x 프로테아제 억제제) 및 4㎕의 500 nM 특유하게 인덱싱된 시토신-고갈된 트랜스포좀을 조합하였다. 혼합물을 55℃에서 20분 동안 인큐베이션시킨 다음, QIAquick® PCR 정제 컬럼을 사용하여 정제하고, 30㎕의 EB 중에 용리시켰다.

[0180] DNA의 농도를 2㎕의 혼합물을 사용하여 dsDNA 고 감도 큐비트 2.0 형광계로 정량화하였다. 농도를 17.95 pg/㎕로 희석하였으며, 이는 대략 5개의 인간 세포의 게놈 질량을 모의한다.

[0181] 실시예 2

[0182] 18% PEG SPRI 비드 혼합물의 제조

[0183] Sera-Mag 비드(1 ㎖)를 로우-바인드(low-bind) 1.5 ㎖ 튜브로 분취한 다음, 상청액이 투명해질 때까지 자성 스탠드(magnetic stand) 상에 배치하였다. 비드를 500㎕의 10 mM Tris-HCl, pH 8.0의 용액을 사용하여 세척하고, 상청액이 투명해진 후에 용액을 제거하고, 총 4회의 세척을 위하여 이러한 세척 단계를 반복하였다. 비드를 하기의 혼합물, 18% PEG 8000(질량 기준), 1M NaCl, 10mM 트리스-HCl, pH 8.0, 1mM EDTA, 0.05% 트윈(Tween)-20 중에 재현탁화시키고; 온건하게 진탕시키면서 적어도 1시간 동안 실온에서 인큐베이션시킨 다음, 18% PEG SPRI 비드를 4℃에 보관하였다. 비드를 사용 전에 실온에 도달하게 하였다.

[0184] 실시예 3

[0185] 리튬 3,5-다이아이오도살리실산(LAND) 또는 SDS(xSDS)를 사용한 핵의 제조

[0186] *A. LAND 핵 제조 및 뉴클레오솜 고갈 방법*

[0187] 세포가 현탁 세포 배양물에 존재하면, 배양물을 온건하게 분쇄하여 세포 클럼프를 파괴하고, 세포를 4℃에서 5분 동안 500xg로 회전시킴으로써 펠렛화시키고, 500㎕의 빙냉 PBS로 세척하였다.

[0188] 세포가 부착 세포 배양물에 존재하면, 배지를 흡인시키고, 세포를 37℃에서 10 ㎖의 PBS로 세척한 다음, 37℃에서 단층을 덮도록 충분한 0.25% 트립신을 첨가하였다. 37℃에서 5분 동안 또는 90%의 세포가 더 이상 표면에 부착되지 않을 때까지 인큐베이션시킨 후에, 37℃ 배지를 1:1 비로 첨가하여, 트립신을 켄칭시켰다. 세포를 4℃에서 5분 동안 500xg로 회전시킴으로써 펠렛화시킨 다음, 500㎕의 빙냉 PBS로 세척하였다.

[0189] 5분 동안 500xg로 회전시킴으로써 현탁 세포 배양물 또는 부착 세포 배양물로부터의 세포를 펠렛화시킨 다음, NIB 완충제 중 200㎕의 12.5 mM LIS(2.5㎕의 1M LIS + 197.5㎕의 NIB 완충제) 중에 재현탁화시켰다. 얼음 상에서 5분 동안 인큐베이션시킨 후에, 800㎕의 NIB 완충제를 첨가하였다. 세포를 온건하게 35㎛ 세포 스트레이너를 통과시키고, 5㎕의 DAPI(5 mg/㎖)를 첨가하였다.

[0190] *B. xSDS 핵 제조 및 뉴클레오솜 고갈 방법*

[0191] 세포가 현탁 세포 배양물에 존재하면, 배지를 온건하게 분쇄하여 세포 클럼프를 파괴하였다. 10 ㎖의 배지 중 세포에, 406㎕의 37% 포름알데히드를 첨가하고, 온건하게 진탕시키면서 실온에서 10분 동안 인큐베이션시켰다. 800 마이크로리터의 2.5M 글리신을 세포에 첨가하고, 얼음 상에서 5분 동안 인큐베이션시킨 다음, 4℃에서 8분 동안 550xg로 원심분리하였다. 10 ㎖의 빙냉 PBS로 세척한 후에, 세포를 5 ㎖의 빙냉 NIB(10mM 트리스HCl pH7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% Igepal®, 1x 프로테아제 억제제) 중에 재현탁화시키고, 온건하게 혼합하면서 얼음 상에서 20분 동안 인큐베이션시켰다.

[0192] 세포가 부착 세포 배양물에 존재하면, 배지를 흡인시키고, 세포를 37℃에서 10 ㎖의 PBS로 세척한 다음, 37℃에서 단층을 덮도록 충분한 0.25% 트립신을 첨가하였다. 37℃에서 5분 동안 또는 90%의 세포가 더 이상 표면에 부착되지 않을 때까지 인큐베이션시킨 후에, 37℃ 배지를 1:1 비로 첨가하여, 트립신을 켄칭시키고, 배지로 부피가 10 ㎖이 되게 하였다. 세포를 10 ㎖의 배지 중에 재현탁화시키고, 406㎕의 37% 포름알데히드를 첨가하고, 온건하게 진탕시키면서 실온에서 10분 동안 인큐베이션시켰다. 800 마이크로리터의 2.5 M 글리신을 세포에 첨가하고, 얼음 상에서 5분 동안 인큐베이션시켰다. 세포를 4℃에서 8분 동안 550xg로 원심분리하고, 10 ㎖의 빙냉 PBS로 세척하였다. 5 ㎖의 빙냉 NIB 중에 세포를 재현탁화시킨 후에, 그들을 온건하게 혼합하면서 얼음 상에서 20분 동안 인큐베이션시켰다.

[0193] 현탁 세포 배양물 또는 부착 세포 배양물로부터의 세포 또는 핵을 5분 동안 500xg로 회전시킴으로써 펠렛화시키고, 900㎕의 1x NEBuffer 2.1로 세척하였다. 5분 동안 500xg로 회전시킨 후에, 펠렛을 12㎕의 20% SDS와 함께 800㎕의 1x NEBuffer 2.1 중에 재현탁화시키고, 30분 동안 격렬하게 진탕시키면서 42℃에서 인큐베이션시킨 다음, 200㎕의 10% 트리톤™ X-100을 첨가하고, 30분 동안 격렬하게 진탕시키면서 42℃에서 인큐베이션시켰다. 세포를 온건하게 35㎛ 세포 스트레이너를 통과시키고, 5㎕의 DAPI(5 mg/㎖)를 첨가하였다.

[0194] 실시예 4

[0195] 핵 분류 및 태그먼트화

[0196] 태그먼트화 플레이트를 10㎕의 1x TD 완충제(1개의 플레이트에 있어서: 500㎕의 NIB 완충제 + 500㎕의 TD 완충제)를 사용하여 준비하고, 2500개의 단일의 핵을 태그먼트화 플레이트의 각 웰 내로 분류하였다. 이러한 단계에

서, 웰당 핵의 수는 웰당 핵의 수가 전체 플레이트에 대하여 일관적인 한, 약간 달라질 수 있다. 또한, 트랜스포사제 인덱스가 보존될 경우 상이한 시료를 플레이트의 상이한 웰로 다중화시킬 수 있다. 도 2에 따라 세포를 게이팅하였다. 플레이트를 5분 동안 500 x g로 회전 침강시킨 후에, 4㎕의 500 nM 특유하게 인덱싱된 시토신-고갈된 트랜스포좀을 각 웰에 첨가하였다. 밀봉 후에, 플레이트를 온건하게 진탕시키면서 55℃에서 15분 동안 인큐베이션시켰다. 그 다음, 플레이트를 얼음 상에 배치하였다. 모든 웰을 풀링한 다음, 35μM 세포 스트레이너를 통과시켰다. 5 마이크로리터의 DAPI(5 mg/㎖)를 첨가하였다.

[0197] 실시예 5

[0198] 핵의 두번째 분류

[0199] 각 웰에 대하여 5㎕의 자이모 분해 시약(2.5㎕ M-분해 완충제, 2.25㎕ H20 및 0.25㎕ 프로테이나제 K)을 사용하여 마스터 믹스를 제조하였다. 10 또는 22개의 단일 핵을 더욱 엄격한 분류 환경을 사용하여 각 웰로 분류하였다. 10개의 단일의 핵을 비메틸화된 대조군 스파이크-인을 위해 사용될 웰로 분류하였으며, 22개의 세포를 다른 웰로 분류하였다. 그 다음, 플레이트를 4℃에서 5분 동안 600 x g로 회전 침강시켰다.

[0200] 실시예 6

[0201] 분해 및 바이설파이트 전환

[0202] C-고갈된 트랜스포좀으로 사전-처리된 대략 ~35 pg(2㎕)의 비메틸화된 대조군 람다 DNA를 사용하여, 웰을 10개의 단일의 핵으로 스파이킹하였다. 플레이트를 50℃에서 20분 동안 인큐베이션시켜, 핵을 분해하고, 32.5㎕의 신선하게 제조된 자이모 CT 전환 시약을 제조처의 프로토콜에 따라 첨가하였다. 웰을 분쇄에 의해 혼합하고, 플레이트를 4℃에서 2분 동안 600 x g로 회전 침강시켰다. 플레이트를 계속하기 전에 하기의 단계: 8분 동안 98℃, 3.5시간 동안 64℃를 위하여 열순환기 상에 배치한 다음, 20시간 미만 동안 4℃에서 유지하였다. 자이모 매그바인딩(MagBinding) 비드(5㎕)를 각 웰에 첨가하고, 150㎕의 M-결합 완충제를 각 웰에 첨가하였다. 분쇄에 의해 웰을 혼합한 후에, 플레이트를 실온에서 5분 동안 인큐베이션시켰다. 상청액이 투명해질 때까지 플레이트를 96-웰 호환성 자성 랙 상에 배치하였다.

[0203] i) 자성 랙으로부터 플레이트를 제거하고, ii) 100㎕의 80% 에탄올을 각 웰에 첨가하여, 비드 펠렛 상에 흐르게 하고, iii) 플레이트를 다시 자성 랙 상에 배치한 다음, 일단 투명해지면 상청액을 제거함으로써 상청액을 제거하고, 웰을 신선한 80% 에탄올(부피 기준)로 세척하였다.

[0204] 50㎕의 M-탈설폰화 완충제를 각 웰에 첨가하고, 비드를 분쇄에 의해 완전히 재현탁화시키고, 실온에서 15분 동안 인큐베이션시키고, 플레이트를 자성 랙 상에 배치한 다음, 일단 투명해지면 상청액을 제거함으로써 탈설폰화를 달성하였다.

[0205] i) 자성 랙으로부터 플레이트를 제거하고, ii) 100㎕의 80% 에탄올을 각 웰에 첨가하여, 비드 펠렛 상에 흐르게 하고, iii) 플레이트를 다시 자성 랙 상에 배치한 다음, 일단 투명해지면 상청액을 제거함으로써 상청액을 제거하고, 웰을 신선한 80% 에탄올(부피 기준)로 세척하였다.

[0206] 비드 펠렛은 균열이 눈에 보이기 시작할 때까지 비드 펠렛이 약 10분 동안 건조되게 하였다.

[0207] 25㎕의 자이모 M-용리 완충제를 각 웰에 첨가하고, 분쇄하여 펠렛을 완전히 해리시키고, 플레이트를 55℃에서 4분 동안 가열함으로써 용리를 달성하였다.

[0208] 실시예 7

[0209] 선형 증폭

[0210] 전체 용리액을 웰당 하기의 반응 믹스를 사용하여 제조된 플레이트로 옮겼다: 16㎕의 PCR-클린 H20, 5㎕의 10X NEBuffer 2.1, 2㎕의 10 mM dNTP 믹스 및 2㎕의 10 μM 9-뉴클레오티드 무작위 프라이머.

[0211] 선형 증폭을 하기와 같이 수행하였다: i) 95℃에서 45초 동안 인큐베이션시킨 다음, 얼음 상에서 순간적으로 냉각시키고, 얼음 상에 유지함으로써 DNA가 단일-가닥이 되게 하고, ii. 일단 완전히 냉각되면, 10U 클레나우(3'->5' 엑소-) 중합효소를 각 웰에 첨가하고, iii) 플레이트를 4℃에서 5분 동안 인큐베이션시킨 다음, 온도를 +1℃/15초의 속도로 37℃까지 증가시킨 다음, 37℃에서 90분 동안 유지하였다.

[0212] 총 4 라운드의 선형 증폭을 위하여 단계 i 내지 iii을 3회 이상 반복하였다. 각 증폭을 위하여, 하기의 혼합물을 각 웰 내의 반응물에 첨가하였다: 1㎕의 10 μM 9-뉴클레오티드 무작위 프라이머, 1㎕의 10 mM dNTP 믹스,

및 1.25㎕의 4X NEBuffer 2.1. 4 라운드의 선형 증폭은 전형적으로 더 적은 라운드에 비하여 판독물 정렬 속도와 라이브러리 복잡성을 유의미하게 증가시켰다.

[0213] 웰을 하기와 같이 1.1X(웰 반응 부피에 비한 부피 기준 농도)에서 제조된 18% PEG SPRI 비드 혼합물을 사용하여 세정하였다. 플레이트를 실온에서 5분 동안 인큐베이션시키고, 자성 랙 상에 배치하고, 일단 투명해지면, 상청액을 제거하였다. 비드 펠렛을 50㎕의 80% 에탄올로 세척하였다. 임의의 잔류 액체를 제거하고, 비드 펠렛이 균열되기 시작할 때까지 건조되게 하였다. DNA를 21㎕의 10 mM 트리스-Cl(pH 8.5) 중에 용리하였다.

[0214] 실시예 8

[0215] 인덱싱 PCR 반응

[0216] 전체 용리액을 웰당 하기의 반응 믹스를 사용하여 제조된 플레이트로 옮겼다: 2㎕의 10 μM i7 인덱스 PCR 프라이머, 2㎕의 10 μM i5 인덱스 PCR 프라이머, 25㎕의 2X 카파 HiFi™ 핫스타트 레디믹스 및 0.5㎕의 100X SYBR® 그린 I. 하기의 사이클: 2분 동안 95℃, (80초 동안 94℃, 30초 동안 65℃, 30초 동안 72℃)을 사용하여 실시간 열 순환기에서 PCR 증폭을 수행하고, 일단 대다수의 웰이 측정되는 SYBR® 녹색 형광의 변곡을 보이면, 반응을 중단시켰다. 라이브러리 제조를 위한 16 내지 21회의 PCR 사이클에 변곡 안정기가 관찰되었다.

[0217] 실시예 9

[0218] 라이브러리 정화 및 정량화

[0219] 라이브러리를 웰마다 하기와 같이 0.8X(웰 반응 부피에 비한 부피 기준 농도)에서 18% PEG SPRI 비드 혼합물을 사용하여 세정하였다. 플레이트를 실온에서 5분 동안 인큐베이션시키고, 자성 랙 상에 배치하고, 일단 투명해지면, 상청액을 제거하였다. 비드 펠렛을 50㎕의 80% 에탄올로 세척하였다. 임의의 잔류 액체를 제거하고, 비드 펠렛이 균열되기 시작할 때까지 건조되게 하였다. DNA를 25㎕의 10 mM 트리스-Cl(pH 8.5) 중에 용리하였다.

[0220] 5㎕의 각 웰을 사용하여 라이브러리를 풀링하고, 2㎕을 사용하여, 제조처의 프로토콜에 따라 dsDNA 고 감도 큐비트® 2.0 형광계를 사용하여 DNA의 농도를 정량화하였다. 큐비트® 판독물을 사용하여, 라이브러리를 약 4 ng/㎕로 희석하고, 1㎕를 제조처의 프로토콜에 따라 고 감도 생물분석기 2100에서 실행하였다. 그 다음, 라이브러리를 200bp 내지 1 kbp 범위에 대하여 정량화하여, 일루미나 서열결정을 위하여 풀을 1 nM로 희석하였다.

[0221] 실시예 10

[0222] 서열결정

[0223] NextSeq® 500을 하기의 변경을 제외하고, 1 nM 시료에 대하여 제조처의 지침에 따라 시행을 위해 준비하였다. 라이브러리 풀을 0.9 pM의 농도 및 1.5 ㎖의 총 부피로 로딩하고, 카트리지 위치 10에 두고; 커스텀 플라이머를 9㎕의 100 μM 원액 서열결정 프라이머 1을 총 1.5 ㎖의 HT1 완충제로 희석함으로써 카트리지 위치 7 내로 준비하고, 100 μM 원액 농도에서 18㎕의 각각의 커스텀 인덱스 서열결정 프라이머를 총 3 ㎖의 HT1 완충제로 희석하여, 카트리지 위치 9 내로 준비하고; NextSeq® 500을 독립형 방식으로 작동시키고; SCIseq 커스텀 화학 레시피(문헌[Amini et al., 2014, *Nat. Genet*. 46, 1343-1349])를 선택하고; 이중 인덱스를 선택하고; 적절한 수의 판독 사이클을 입력하고(150이 권고됨); 인덱스 1에 대하여 10회 사이클 및 인덱스 2에 대하여 20 사이클; 모든 판독 및 인덱스를 위하여 커스텀 체크박스를 선택하였다.

[0224] 본 명세서에 인용된 모든 특허, 특허 출원 및 간행물 및 전자적으로 이용 가능한 자료(예를 들어, GenBank 및 RefSeq에서의 뉴클레오티드 서열 제출물, 및 예를 들어, SwissProt, PIR, PRF, PDB에서의 아미노산 서열 제출물, GenBank 및 RefSeq에서의 주석이 달린 코딩 영역으로부터의 번역물 포함)의 완전한 개시내용은 그들의 전문이 참조로 포함된다. 간행물에 언급된 보충 자료(예컨대, 보충 표, 보충 도면, 보충 자료 및 방법, 및/또는 보충 실험 데이터)는 마찬가지로 그들의 전문이 참조로 포함된다. 본 출원의 개시내용과 본 명세서에 참조로 포함된 임의의 문서의 개시내용(들) 간에 불일치가 존재하는 경우, 본 출원의 개시내용이 우선해야 한다. 상기 상세한 설명 및 실시예는 단지 이해의 명확성을 위해 제공된다. 불필요한 제한은 그로부터 이해되지 않아야 한다. 본 발명은 나타내고 기재된 정확한 상세사항에 제한되지 않고, 당업자에게 자명한 변형이 청구범위에 의해서 정의된 본 발명에 포함될 것이다.

[0225] 달리 지시되지 않는 한, 명세서 및 청구범위에 사용된 성분의 양, 분자량 등을 표현하는 모든 수는 모든 경우에 용어 "약"에 의해서 수식되는 것으로 이해되어야 한다. 따라서, 반대로 지시되지 않는 한, 명세서 및 청구범위에 언급된 수치 파라미터는 본 발명에 의해 수득하고자 하는 목적하는 특성에 따라서 달라질 수 있는 근사치이

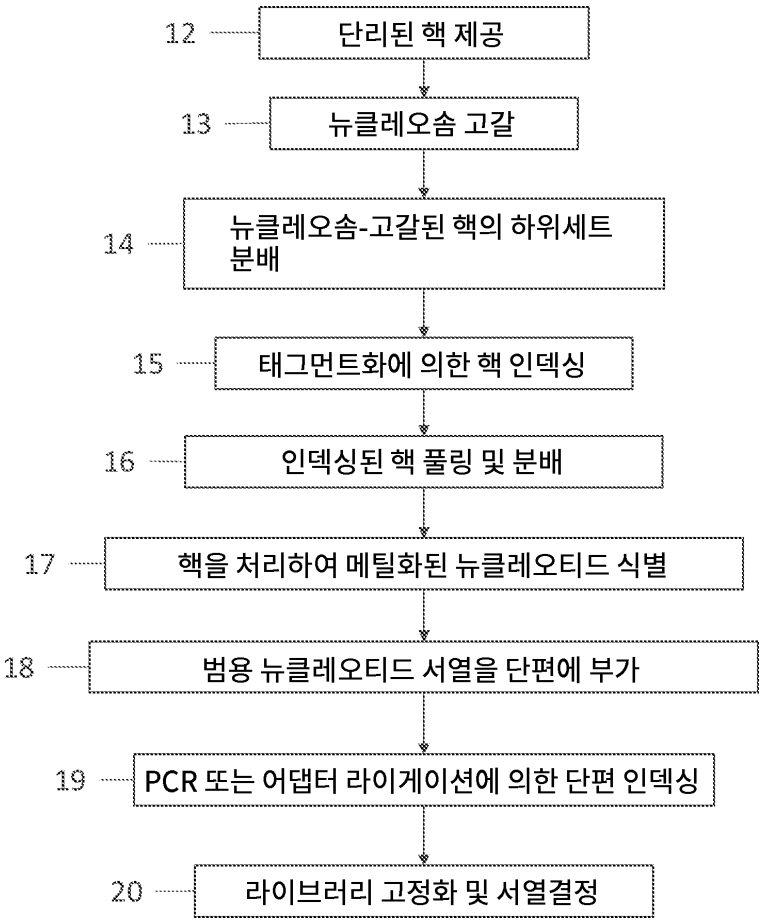
다. 적어도 그리고 청구범위의 범주와 동등한 학설을 제한하려는 시도가 아니지만, 각각의 수치 파라미터는 적어도 보고된 유효 자릿수에 비추어 그리고 통상적인 반올림 기법을 적용함으로써 해석되어야 한다.

[0226] 본 발명의 넓은 범주를 설명하는 수치 범위 및 파라미터가 근사치임에도 불구하고, 구체적인 실시예에 언급된 수치는 가능한 한 정확하게 기록된다. 그러나 모든 수치는 본질적으로 그들의 각각의 시험 측정치에서 발견된 표준 편차에서 필연적으로 발생하는 범위를 함유한다.

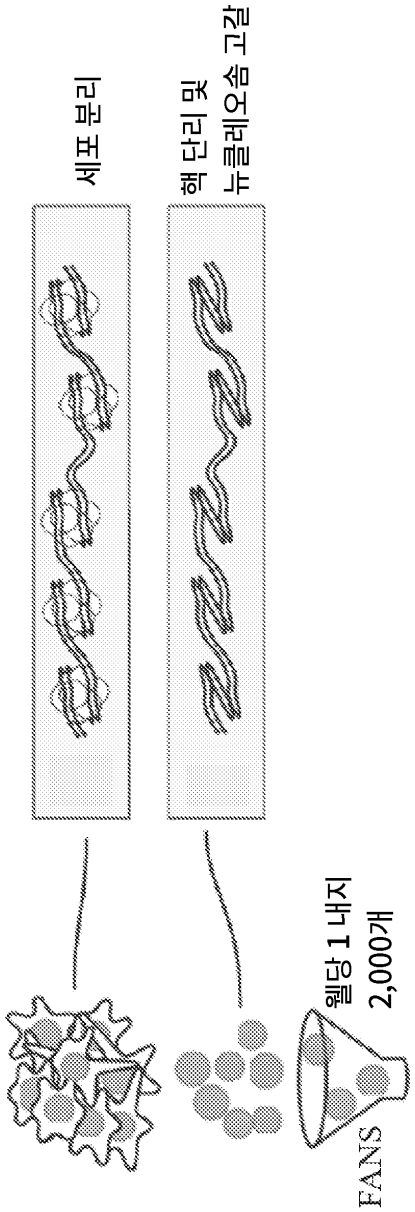
[0227] 모든 제목은 읽는 이의 편의를 위한 것이며, 달리 그렇게 명시되지 않는 한 제목 뒤에 오는 본문의 의미를 제한하기 위하여 사용되지 않아야 한다.

도면

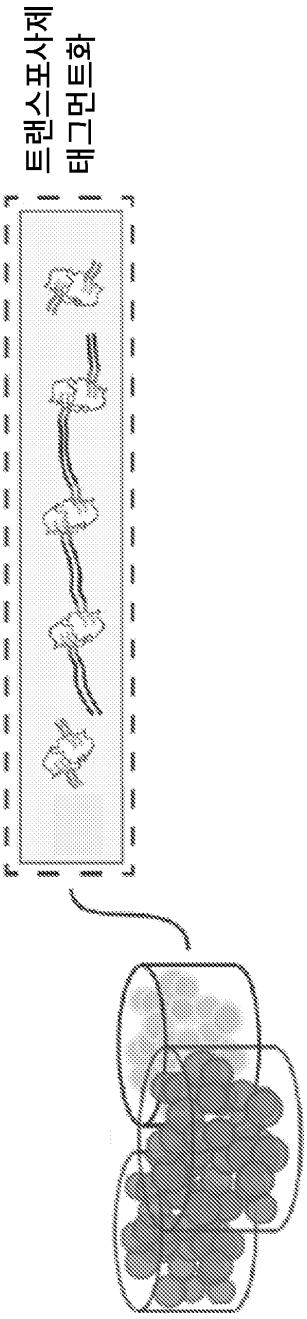
도면1



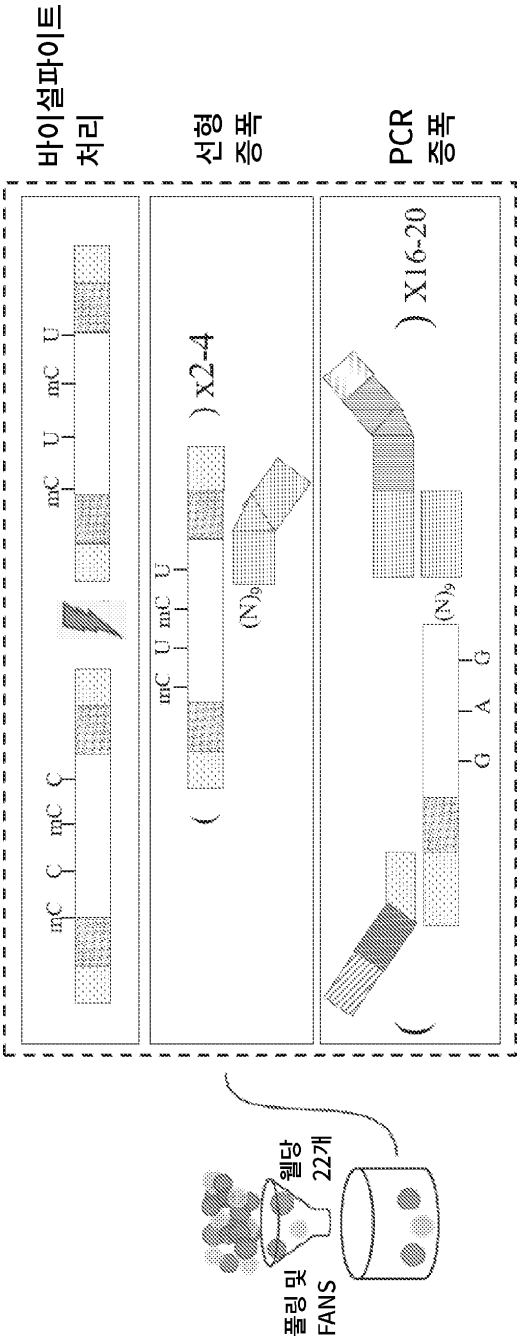
도면2a



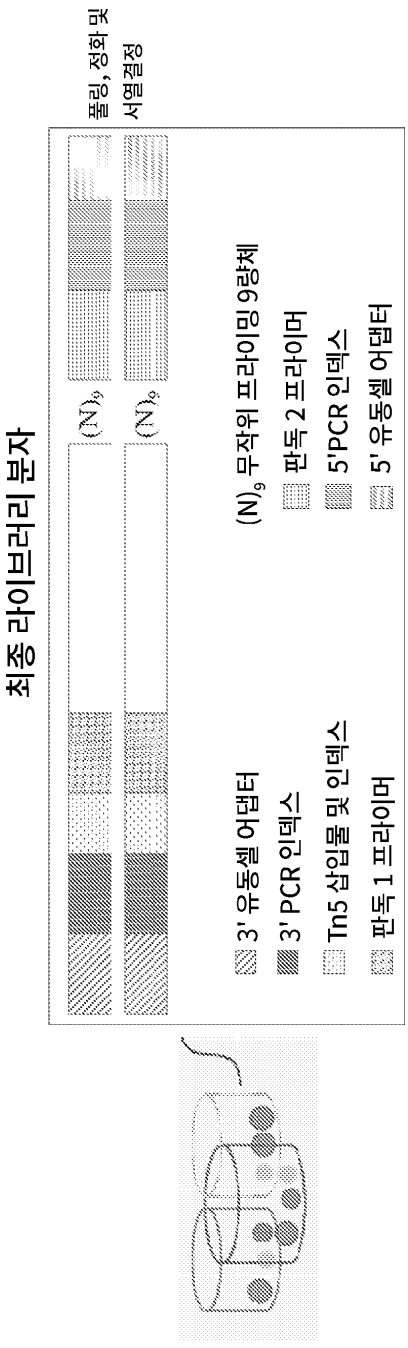
도면2b



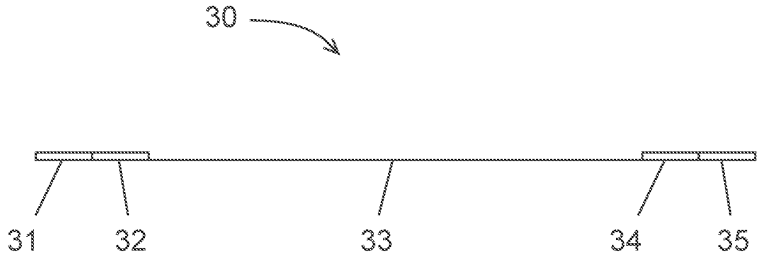
도면2c



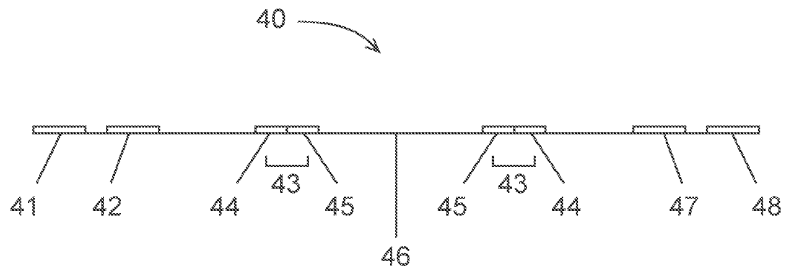
도면2d


최종 라이브러리 분자
풀링, 정화 및 서열결정
$(N)_9$
$(N)_9$
$(N)_9$ 무작위 프라이밍 9량체
3' 유동셀 어댑터
3' PCR 인덱스
Tn5 삽입물 및 인덱스
판독 1 프라이머
판독 2 프라이머
5'PCR 인덱스
5' 유동셀 어댑터

도면3


30
31
32
33
34
35

도면4



서 열 목 록

SEQUENCE LISTING

<110> OREGON HEALTH & SCIENCE UNIVERSITY

ILLUMINA, INC.

<120> SINGLE CELL WHOLE GENOME LIBRARIES FOR METHYLATION SEQUENCING

<130> IP-1584-PCT

<140> PCT/US2018/036078

<141> 2018-06-05

<150> 62/516,324

<151> 2017-06-07

<160> 132

<170> PatentIn version 3.5

<210> 1

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 1

ctgtctctta tacacatct 19

<210> 2

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 2

ggtgtagtgg gtttgggtta agaggaatgg tagagagggt gagatgtgta taagagacag 60

<210> 3

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 3

ggtgtagtgg gtttggagta ggaagattgg tagagagggt gagatgtgta taagagacag 60

<210>

4

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 4

ggtgtagtgg gtttgggaat taggtgttgg tagagagggt gagatgtgta taagagacag 60

<210> 5

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 5

ggtgtagtgg gtttggggag attaatgtgg tagagagggt gagatgtgta taagagacag 60

<210> 6

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 6

ggtgtagtgg gtttggtatt gtggaattgg tagagagggt gagatgtgta taagagacag 60

<210> 7

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 7

ggtgtagtgg gtttggatat agatgattgg tagagagggt gagatgtgta taagagacag 60

<210> 8

<211>

60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 8

ggtgtagtgg gtttgggtaa gaggaattgg tagagagggt gagatgtgta taagagacag 60

<210> 9

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 9

ggtgtagtgg gtttgggaga gttattgtgg tagagagggt gagatgtgta taagagacag 60

<210> 10

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 10

ggtgtagtgg gtttggagtt agtgtgatgg tagagagggt gagatgtgta taagagacag 60

<210> 11

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 11

ggtgtagtgg gtttgggata tagaatttgg tagagagggt gagatgtgta taagagacag 60

<210> 12

<211> 60

<

212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 12

ggtgtagtgg gtttggaagg aagtgaatgg tagagagggt gagatgtgta taagagacag 60

<210> 13

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 13

ggtgtagtgg gtttggaata aggaaggtgg tagagagggt gagatgtgta taagagacag 60

<210> 14

<211> 60

<212

> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 14

ggtgtagtgg gtttgggtat ggatatatgg tagagagggt gagatgtgta taagagacag 60

<210> 15

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 15

ggtgtagtgg gtttggttag ataatgatgg tagagagggt gagatgtgta taagagacag 60

<210> 16

<211> 60

<212>

DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 16

ggtgtagtgg gtttggggtg ttgtaattgg tagagagggt gagatgtgta taagagacag 60

<210> 17

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 17

ggtgtagtgg gtttgggaag tggagagtgg tagagagggt gagatgtgta taagagacag 60

<210> 18

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 18

ggtgtagtgg gtttggttga gtggtagtgg tagagagggt gagatgtgta taagagacag 60

<210> 19

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 19

ggtgtagtgg gtttgggata atggtgatgg tagagagggt gagatgtgta taagagacag 60

<210> 20

<211> 60

<212> DNA

<

213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 20

ggtgtagtgg gtttgggtgt taatggatgg tagagagggt gagatgtgta taagagacag 60

<210> 21

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 21

ggtgtagtgg gtttggtagg aatggtgtgg tagagagggt gagatgtgta taagagacag 60

<210> 22

<211> 60

<212> DNA

<213

> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 22

ggtgtagtgg gtttggatgt atggatatgg tagagagggt gagatgtgta taagagacag 60

<210> 23

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 23

ggtgtagtgg gtttggtgat tgttggttgg tagagagggt gagatgtgta taagagacag 60

<210> 24

<211> 60

<212> DNA

<213>

Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 24

ggtgtagtgg gtttggaaga gaattattgg tagagagggt gagatgtgta taagagacag 60

<210> 25

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 25

ggtgtagtgg gtttggaatg gttggtatgg tagagagggt gagatgtgta taagagacag 60

<210> 26

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 26

ggtgtagtgg gtttggggtt aattgagtgg tagagagggt gagatgtgta taagagacag 60

<210> 27

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 27

ggtgtagtgg gtttgggtat aatagtttgg tagagagggt gagatgtgta taagagacag 60

<210> 28

<211> 60

<212> DNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 28

ggtgtagtgg gtttggttag ttgaatttgg tagagagggt gagatgtgta taagagacag 60

<210> 29

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 29

ggtgtagtgg gtttggttgg tgaaggttgg tagagagggt gagatgtgta taagagacag 60

<210> 30

<211> 60

<212> DNA

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 30

ggtgtagtgg gtttggttaa tattgaatgg tagagagggt gagatgtgta taagagacag 60

<210> 31

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 31

ggtgtagtgg gtttgggtta gaattggtgg tagagagggt gagatgtgta taagagacag 60

<210> 32

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 32

ggtgtagtgg gtttgggtta ttaattatgg tagagagggt gagatgtgta taagagacag 60

<210> 33

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 33

ggtgtagtgg gtttgggatt ggtaagatgg tagagagggt gagatgtgta taagagacag 60

<210> 34

<211> 60

<212> DNA

<213> Artificial Sequence

<220><

221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 34

ggtgtagtgg gtttggtgaa gtattgttgg tagagagggt gagatgtgta taagagacag 60

<210> 35

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 35

ggtgtagtgg gtttgggatg gattatgtgg tagagagggt gagatgtgta taagagacag 60

<210> 36

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221>

> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 36

ggtgtagtgg gtttggatta gtatatttgg tagagagggt gagatgtgta taagagacag 60

<210> 37

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 37

ggtgtagtgg gtttgggtag gtgtggttgg tagagagggt gagatgtgta taagagacag 60

<210> 38

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221>

source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 38

ggtgtagtgg gtttggagtt gaatgtatgg tagagagggt gagatgtgta taagagacag 60

<210> 39

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 39

ggtgtagtgg gtttggattg tgagatatgg tagagagggt gagatgtgta taagagacag 60

<210> 40

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 40

ggtgtagtgg gtttggttgt ggtgagttgg tagagagggt gagatgtgta taagagacag 60

<210> 41

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 41

ggtgtagtgg gtttggttaa gttggtttgg tagagagggt gagatgtgta taagagacag 60

<210> 42

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<

223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 42

ggtgtagtgg gtttggtata ataatattgg tagagagggt gagatgtgta taagagacag 60

<210> 43

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 43

ggtgtagtgg gtttggaagg tatgagttgg tagagagggt gagatgtgta taagagacag 60

<210> 44

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223
> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 44

ggtgtagtgg gtttggagga ttataagtgg tagagagggt gagatgtgta taagagacag 60

<210> 45

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 45

ggtgtagtgg gtttggagag ttaggtttgg tagagagggt gagatgtgta taagagacag 60

<210> 46

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223>
/note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 46

ggtgtagtgg gtttggatgg atagtattgg tagagagggt gagatgtgta taagagacag 60

<210> 47

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 47

ggtgtagtgg gtttggatat tatgttgtgg tagagagggt gagatgtgta taagagacag 60

<210> 48

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 48

ggtgtagtgg gtttggggtg gagatagtgg tagagagggt gagatgtgta taagagacag 60

<210> 49

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 49

ggtgtagtgg gtttggtggt ggtagtgtgg tagagagggt gagatgtgta taagagacag 60

<210> 50

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 50
ggtgtagtgg gtttggaggt gagaagttgg tagagagggt gagatgtgta taagagacag 60
<210> 51
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 51
ggtgtagtgg gtttggtagg aggttgttgg tagagagggt gagatgtgta taagagacag 60
<210> 52
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<
400> 52
ggtgtagtgg gtttggtgta taggtattgg tagagagggt gagatgtgta taagagacag 60
<210> 53
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 53
ggtgtagtgg gtttggtgtt atgtagatgg tagagagggt gagatgtgta taagagacag 60
<210> 54
<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400

> 54

ggtgtagtgg gtttggtgga aggtatgtgg tagagagggt gagatgtgta taagagacag 60

<210> 55

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 55

ggtgtagtgg gtttggaatg taaggagtgg tagagagggt gagatgtgta taagagacag 60

<210> 56

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400>

56

ggtgtagtgg gtttgggtta tgttaagtgg tagagagggt gagatgtgta taagagacag 60

<210> 57

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 57

ggtgtagtgg gtttggtgtt ataggtgtgg tagagagggt gagatgtgta taagagacag 60

<210> 58

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 58

ggtgtagtgg gtttggaagg agaattgtgg tagagagggt gagatgtgta taagagacag 60

<210> 59

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 59

ggtgtagtgg gtttggagag gtggaagtgg tagagagggt gagatgtgta taagagacag 60

<210> 60

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 60

ggtgtagtgg gtttgggatt aggtgtatgg tagagagggt gagatgtgta taagagacag 60

<210> 61

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 61
ggtgtagtgg gtttggatta tataagatgg tagagagggt gagatgtgta taagagacag 60
<210> 62
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 62
ggtgtagtgg gtttgggaga atatggttgg tagagagggt gagatgtgta taagagacag 60
<
210> 63
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 63
ggtgtagtgg gtttggggat tgagaggtgg tagagagggt gagatgtgta taagagacag 60
<210> 64
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 64
ggtgtagtgg gtttggatta tggtggttgg tagagagggt gagatgtgta taagagacag 60
<210
> 65
<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 65

ggtgtagtgg gtttgggaag gaagttatgg tagagagggt gagatgtgta taagagacag 60

<210> 66

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 66

ggtgtagtgg gtttgggaat atgtaagtgg tagagagggt gagatgtgta taagagacag 60

<210>

67

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 67

ggtgtagtgg gtttggtagt taatatttgg tagagagggt gagatgtgta taagagacag 60

<210> 68

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 68

ggtgtagtgg gtttggtgaa tgaatagtgg tagagagggt gagatgtgta taagagacag 60

<210> 69

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 69

ggtgtagtgg gtttggagga tggattatgg tagagagggt gagatgtgta taagagacag 60

<210> 70

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 70

ggtgtagtgg gtttggaagt gtatagatgg tagagagggt gagatgtgta taagagacag 60

<210> 71

<

211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 71

ggtgtagtgg gtttgggagg ttgaagatgg tagagagggt gagatgtgta taagagacag 60

<210> 72

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 72

ggtgtagtgg gtttggtgtg taataggtgg tagagagggt gagatgtgta taagagacag 60

<210> 73

<211

> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 73

ggtgtagtgg gtttggttga ttagagatgg tagagagggt gagatgtgta taagagacag 60

<210> 74

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 74

ggtgtagtgg gtttggtatg tgtgtggtgg tagagagggt gagatgtgta taagagacag 60

<210> 75

<211>

60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 75

ggtgtagtgg gtttgggaga tgagaattgg tagagagggt gagatgtgta taagagacag 60

<210> 76

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 76

ggtgtagtgg gtttggtggt gaagtgatgg tagagagggt gagatgtgta taagagacag 60

<210> 77

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 77

ggtgtagtgg gtttgggtgg taggatgtgg tagagagggt gagatgtgta taagagacag 60

<210> 78

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 78

ggtgtagtgg gtttggtgta ggtgatatgg tagagagggt gagatgtgta taagagacag 60

<210> 79

<211> 60

<

212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 79

ggtgtagtgg gtttgggtaa ggtgtgatgg tagagagggt gagatgtgta taagagacag 60

<210> 80

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 80

ggtgtagtgg gtttggagaa gagagtgtgg tagagagggt gagatgtgta taagagacag 60

<210> 81

<211> 60

<212

> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 81

ggtgtagtgg gtttggggat gttgtattgg tagagagggt gagatgtgta taagagacag 60

<210> 82

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 82

ggtgtagtgg gtttggaagt tatataatgg tagagagggt gagatgtgta taagagacag 60

<210> 83

<211> 60

<212>

DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 83

ggtgtagtgg gtttggtgga attaagttgg tagagagggt gagatgtgta taagagacag 60

<210> 84

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 84

ggtgtagtgg gtttggtaat gagaggatgg tagagagggt gagatgtgta taagagacag 60

<210> 85

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 85

ggtgtagtgg gtttggataa ttgatggtgg tagagagggt gagatgtgta taagagacag 60

<210> 86

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 86

ggtgtagtgg gtttggtgtg aagagtatgg tagagagggt gagatgtgta taagagacag 60

<210> 87

<211> 60

<212> DNA

<

213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 87

ggtgtagtgg gtttgggatg aatatgttgg tagagagggt gagatgtgta taagagacag 60

<210> 88

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 88

ggtgtagtgg gtttggtgag gatagattgg tagagagggt gagatgtgta taagagacag 60

<210> 89

<211> 60

<212> DNA

<213
> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 89

ggtgtagtgg gtttggatta attagagtgg tagagagggt gagatgtgta taagagacag 60

<210> 90

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 90

ggtgtagtgg gtttggggag agatggatgg tagagagggt gagatgtgta taagagacag 60

<210> 91

<211> 60

<212> DNA

<213>

Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 91

ggtgtagtgg gtttggtaat tgaggaatgg tagagagggt gagatgtgta taagagacag 60

<210> 92

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 92

ggtgtagtgg gtttggttgg aattaattgg tagagagggt gagatgtgta taagagacag 60

<210> 93

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 93

ggtgtagtgg gtttggaatg ttattgttgg tagagagggt gagatgtgta taagagacag 60

<210> 94

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 94

ggtgtagtgg gtttgggtag ttattagtgg tagagagggt gagatgtgta taagagacag 60

<210> 95

<211> 60

<212> DNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 95

ggtgtagtgg gtttggtata ttgtgagtgg tagagagggt gagatgtgta taagagacag 60

<210> 96

<211> 60

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 96

ggtgtagtgg gtttgggtgt aggatagtgg tagagagggt gagatgtgta taagagacag 60

<210> 97

<211> 60

<212> DNA

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 97

ggtgtagtgg gtttggagag aagttggtgg tagagagggt gagatgtgta taagagacag 60

<210> 98

<211> 37

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<220><221> modified_base

<222> (29)..(37)

<223> a, c, t, g, unknown or other

<400> 98

ggagttcaga cgtgtgctct tccgatctnn nnnnnnn 37

<210> 99

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 99

caagcagaag acggcatacg agatcaagat gccggtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 100

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 100

caagcagaag acggcatacg agataacgtc tagtgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 101

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 101

caagcagaag acggcatacg agataggtat actcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 102

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 102

caagcagaag acggcatacg agatttcata ggacgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 103

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 103

caagcagaag acggcatacg agatggaggc ctccgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210>

104

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 104

caagcagaag acggcatacg agatttcaat ataagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 105

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 105

caagcagaag acggcatacg agatacgtca tatagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 106

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 106

caagcagaag acggcatacg agatttgacc aggagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 107

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 107

caagcagaag acggcatacg agatcggttg cgcggtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 108

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 108

caagcagaag acggcatacg agatcaagga ggtcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210>

109

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 109

caagcagaag acggcatacg agatttacga tgaagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 110

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 110

caagcagaag acggcatacg agatttgctg gcatgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 111

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 111

caagcagaag acggcatacg agataatact cttcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 112

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 112

caagcagaag acggcatacg agatccaact aaccgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 113

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 113

caagcagaag acggcatacg agattatcct caatgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210>

114

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 114

caagcagaag acggcatacg agatgccgtc gcgtgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 115

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 115

caagcagaag acggcatacg agatccgctg cttcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 116

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 116

caagcagaag acggcatacg agattgaccg aatcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 117

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 117

caagcagaag acggcatacg agatgtctcc agaggtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 118

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 118

caagcagaag acggcatacg agataatgct agtcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210>

119

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 119

caagcagaag acggcatacg agatgacgac ctgcgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 120

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 120

caagcagaag acggcatacg agatagagcc agccgtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 121

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 121

caagcagaag acggcatacg agatccaggc cgcagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 122

<211> 68

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 122

caagcagaag acggcatacg agatcaggta tggagtgact ggagttcaga cgtgtgctct 60

tccgatct 68

<210> 123

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 123

aatgatacgg cgaccaccga gatctacacg tatcatcgag gtgtagtggg tttgg 55

<210> 124

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 124

aatgatacgg cgaccaccga gatctacacc cgcgattatg gtgtagtggg tttgg 55

<210> 125

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 125

aatgatacgg cgaccaccga gatctacaca ttcaggtacg gtgtagtggg tttgg 55

<210> 126

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 126

aatgatacgg cgaccaccga gatctacaca tggaattggg gtgtagtggg tttgg 55

<210> 127

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 127

aatgatacgg cgaccaccga gatctacacg acgaagcgtg gtgtagtggg tttgg 55

<210> 128

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 128

aatgatacgg cgaccaccga gatctacacc ttgcagtagg gtgtagtggg tttgg 55

<210> 129

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 129

aatgatacgg cgaccaccga gatctacacc ttggtaatgg gtgtagtggg tttgg 55

<210> 130

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 130

aatgatacgg cgaccaccga gatctacacc aagtcgaccg gtgtagtggg tttgg 55

<210> 131

<211> 33

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 131

tggtagagag ggtgagatgt gtataagaga tag 33

<210> 132

<211> 33

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 132

ctatctctta tacacatctc accctctcta cca 33

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 2

【변경전】

복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리의 제조 방법으로서,

(a) 복수의 세포로부터 단리된 핵을 제공하는 단계;

(b) 상기 단리된 핵이 화학적 처리를 겪게 하여, 단리된 핵의 온전성을 유지하면서 결합된 뉴클레오솜이 더 적은, 고갈된 뉴클레오솜을 포함하는 핵을 생성하는 단계;

(c) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트를 트랜스포좀 복합체를 포함하는 제1 복수 구획 내로 분배하는 단계로서, 각 구획 내의 트랜스포좀 복합체가 다른 구획 내의 제1 인덱스 서열과 상이한 제1 인덱스 서열을 포함하는 단계;

(d) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트 내의 핵산을 복수의 핵산 단편으로 단편화하고, 제1 인덱스 서열을 핵산 단편의 적어도 하나의 가닥 내로 혼입시켜, 인덱싱된 핵을 생성하는 단계;

(e) 상기 인덱싱된 핵을 조합하여, 풀링된 인덱싱 핵을 생성하는 단계;

(f) 상기 풀링된 인덱싱 핵의 하위세트를 제2 복수 구획 내로 분배하고, 상기 인덱싱된 핵이 바이설파이트 처리를 겪게 하여, 바이설파이트-처리된 핵산 단편을 생성하는 단계;

(g) 상기 각 구획 내의 바이설파이트 처리된 핵산 단편을 범용 어댑터에 라이게이션시켜, 라이게이션된 핵산 분자를 생성하는 단계;

(h) 제2 인덱스 서열을 상기 라이게이션된 핵산 분자 내로 혼입하여, 두 인덱스를 포함하는 핵산 분자를 생성하는 단계로서, 각 구획 내의 제2 인덱스 서열이 다른 구획 내의 제2 인덱스 서열과 상이한 것인, 단계; 및

(i) 상기 두 인덱스를 포함하는 핵산 분자를 조합하여, 그에 의해, 복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리를 생성하는 단계를 포함하는, 서열결정 라이브러리의 제조 방법.

【변경후】

복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리의 제조 방법으로서,

(a) 복수의 세포로부터 단리된 핵을 제공하는 단계;

(b) 상기 단리된 핵이 화학적 처리를 겪게 하여, 단리된 핵의 온전성을 유지하면서 결합된 뉴클레오솜이 더 적은, 고갈된 뉴클레오솜을 포함하는 핵을 생성하는 단계;

(c) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트를 트랜스포좀 복합체를 포함하는 제1 복수 구획 내로 분배하는 단계로서, 각 구획 내의 트랜스포좀 복합체가 다른 구획 내의 제1 인덱스 서열과 상이한 제1 인덱스 서열을 포함하는 단계;

(d) 상기 고갈된 뉴클레오솜을 포함하는 핵의 하위세트 내의 핵산을 복수의 핵산 단편으로 단편화하고, 제1 인덱스 서열을 핵산 단편의 적어도 하나의 가닥 내로 혼입시켜, 인덱싱된 핵을 생성하는 단계;

(e) 상기 인덱싱된 핵을 조합하여, 풀링된 인덱싱 핵을 생성하는 단계;

(f) 상기 풀링된 인덱싱 핵의 하위세트를 제2 복수 구획 내로 분배하고, 상기 인덱싱된 핵이 바이설파이트 처리를 겪게 하여, 바이설파이트로 처리된 핵산 단편을 생성하는 단계;

(g) 상기 각 구획 내의 바이설파이트 처리된 핵산 단편을 범용 어댑터에 라이게이션시켜, 라이게이션된 핵산 분자를 생성하는 단계;

(h) 제2 인덱스 서열을 상기 라이게이션된 핵산 분자 내로 혼입하여, 두 인덱스를 포함하는 핵산 분자를 생성하는 단계로서, 각 구획 내의 제2 인덱스 서열이 다른 구획 내의 제2 인덱스 서열과 상이한 것인, 단계; 및

(i) 상기 두 인덱스를 포함하는 핵산 분자를 조합하여, 그에 의해, 복수의 단일 세포로부터의 핵산의 메틸화 상태를 결정하기 위한 서열결정 라이브러리를 생성하는 단계를 포함하는, 서열결정 라이브러리의 제조 방법.

【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 16

【변경전】

제1항에 있어서, 상기 바이설파이트-처리된 핵산 단편의 선형 증폭이 1 내지 10회 사이클을 포함하는, 서열결정 라이브러리의 제조 방법.

【변경후】

제1항에 있어서, 상기 바이설파이트로 처리된 핵산 단편의 선형 증폭이 1 내지 10회 사이클을 포함하는, 서열결정 라이브러리의 제조 방법.

【직권보정 3】

【보정항목】 청구범위

【보정세부항목】 청구항 20

【변경전】

제2항에 있어서, 상기 범용 어댑터의 라이게이션 이전에, 하나 이상의 뉴클레오티드를 상기 바이설파이트-처리된 핵산 단편의 3' 말단에 부가하여, 3' 오버행을 생성하는 단계를 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

【변경후】

제2항에 있어서, 상기 범용 어댑터의 라이게이션 이전에, 하나 이상의 뉴클레오티드를 상기 바이설파이트로 처리된 핵산 단편의 3' 말단에 부가하여, 3' 오버행을 생성하는 단계를 추가로 포함하는, 서열결정 라이브러리의 제조 방법.

【직권보정 4】

【보정항목】 청구범위

【보정세부항목】 청구항 22

【변경전】

제20항에 있어서, 상기 범용 어댑터가 상기 바이설파이트-처리된 핵산 단편 내의 3' 오버행과 역 상보성인 오버행을 포함하는, 서열결정 라이브러리의 제조 방법.

【변경후】

제20항에 있어서, 상기 범용 어댑터가 상기 바이설파이트로 처리된 핵산 단편 내의 3' 오버행과 역 상보성인 오버행을 포함하는, 서열결정 라이브러리의 제조 방법.