



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2007-0112785

(43) 공개일자 2007년11월27일

(51) Int. Cl.

C12Q 1/68 (2006.01)

(21) 출원번호 10-2007-7019819

(22) 출원일자 2007년08월30일

심사청구일자 없음

번역문제출일자 2007년08월30일

(86) 국제출원번호 PCT/US2006/003845

국제출원일자 2006년02월01일

(87) 국제공개번호 WO 2006/084132

국제공개일자 2006년08월10일

(30) 우선권주장

60/649,294 2005년02월01일 미국(US)

(뒷면에 계속)

(71) 출원인

에이젠코트 바이오사이언스 코오폰레이션

미국, 메사츄세츠 01915, 비벌리, 스위트 2450,
커밍스 센터 500

(72) 발명자

맥커난, 케빈

미국, 메사츄세츠 01945, 마블헤드, 페쿠오트 로
드 10

블란차드, 알란

미국, 메사츄세츠 01949, 미들타운, 캠벨 로드 4

(뒷면에 계속)

(74) 대리인

신영무

전체 청구항 수 : 총 354 항

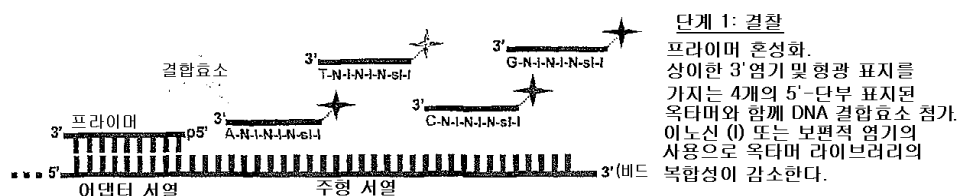
(54) 비드-기초 서열화를 위한 시약, 방법, 및 라이브러리

(57) 요약

본 발명은 단일 가닥의 주형을 따라 연속적인 사이클의 이중나선 연장을 수행함으로써 핵산 서열을 결정하는 방법을 제공한다. 그 사이클은 연장, 결찰, 및 바람직하게는 절단의 단계들을 포함한다. 어떤 실시예에서 방법은 포스포로티올레이트 연쇄를 함유하는 연장 프로브를 사용하고 그러한 연쇄를 절단하기에 적절한 제제를 사용한다. 어떤 실시예에서 방법은 비염기성 잔기 또는 손상된 염기를 함유하는 연장 프로브를 사용하고, 뉴클레오시드와 비염기성 잔기 사이의 연쇄를 절단하기에 적절한 제제 및/또는 핵산으로부터 손상된 염기를 제거하기에 적절한 제제를 사용한다. 본 발명은 최소한 2개의 구별가능하게 표지된 프로브 군을 사용하는 서열에 대한 정보를 결정하는 방법을 제공한다. 어떤 실시예에서 그 방법은 각 사이클에서 주형의 다수의 뉴클레오티드 각각으로부터 2보다 적은 비트의 정보를 얻는다. 어떤 실시예에서는 서열화 반응이 반-고체 지지체에 또는 그 위에 고정된 비트에 부착된 주형에 대해 수행된다. 본 발명은 또한 본 발명의 방법에 사용하기에 적당한 포스포로티올레이트 연쇄 또는 트리거 잔기를 함유하는 표지된 연장 프로브의 세트를 제공한다. 또한, 본 발명은 개시 올리고뉴클레오티드와 연장된 가닥을 제거하고, 상이한 개시 올리고뉴클레오티드를 사용하여 후속적인 반응을 수행함으로써 단일 주형에 대하여 다수의 서열화 반응을 수행하는 것을 포함한다. 본 발명은 주형을 제조하기 위한 효과적인 방법, 특히 다수의 상이한 주형을 동시에 서열화를 수행하기에 효과적인 방법을 더 제공한다. 본 발명은 또한 결찰 및 절단을 수행하는 방법을 제공한다. 본 발명은 또한 쌍을 이룬 태그를 함유하는 핵산 단편의 신규한 라이브러리, 다수의 상이한 주형 (예컨대 쌍을 이룬 태그를 함유함)이 부착된 마이크로입자의 제조 방법, 및 주형을 개별적으로 서열화하는 방법을 제공한다. 본 발명은 또한 자동 서열화 시스템, 유동 셀, 영상 처리 방법, 및 컴퓨터-실행가능한 지시사항 (예컨대 영상-처리 방법을 수행하기 위한) 및/또는 서열 정보를 저장하는 컴퓨터-관독가능한 매질을 제공한다. 어떤 실시예에서 서열 정보는 데이터베이스에 저장된다.

대표도 - 도6A

주기적인 올리고뉴클레오타이드 결찰 및 절단에 의한 서열화



(72) 발명자

코틀러, 레브

미국, 메사추세츠 02134, 알스톤 #3, 래드클리프
로드 6

코스타, 지나

미국, 메사추세츠 01929, 에섹스, 시아 코트 4

(30) 우선권주장

60/656,599 2005년02월25일 미국(US)

60/673,749 2005년04월21일 미국(US)

60/699,541 2005년07월15일 미국(US)

60/722,526 2005년09월30일 미국(US)

특허청구의 범위

청구항 1

주형 폴리뉴클레오타이드의 뉴클레오타이드 서열을 확인하는 방법에 있어서,

- (a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오타이드를 따라 올리고뉴클레오타이드 프로브를 거기에 결합함으로써 개시 올리고뉴클레오타이드를 연장하되, 상기 올리고뉴클레오타이드 프로브는 포스포로티올레이트 연쇄를 포함하는 단계;
- (b) 폴리뉴클레오타이드의 하나 또는 그 이상의 뉴클레오타이드를 확인하는 단계; 및
- (c) 단계 (a)와 (b)를 뉴클레오타이드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 확인 방법.

청구항 2

제 1 항에 있어서, 상기 확인 단계는 가장 최근에 결합된 올리고뉴클레오타이드 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 확인 방법.

청구항 3

제 1 항에 있어서, 상기 방법은 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단함으로써 연장가능한 프로브를 생성하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 확인 방법.

청구항 4

제 3 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 확인 방법.

청구항 5

제 1 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 확인 방법.

청구항 6

주형 폴리뉴클레오타이드의 뉴클레오타이드 서열을 결정하는 방법에 있어서,

- (a) 주형 폴리뉴클레오타이드에 혼성된, 연장가능한 말단을 가지는 프로브를 포함하는 프로브-주형 이중나선을 제공하는 단계;
- (b) 연장된 올리고뉴클레오타이드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단에 연장 올리고뉴클레오타이드 프로브를 결합시키되, 연장 프로브는 포스포로티올레이트 연쇄를 포함하는 단계;
- (c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오타이드 프로브의 바로 아래에 있는 주형 폴리뉴클레오타이드의 뉴클레오타이드 잔기인 주형 폴리뉴클레오타이드의 최소한 하나의 뉴클레오타이드를 확인하는 단계;
- (d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오타이드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및
- (e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오타이드에 있는 뉴클레오타이드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 7

제 6 항에 있어서, 상기 연장 프로브는 한 말단에서 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 8

제 6 항에 있어서, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 9

제 6 항에 있어서, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오티드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 10

제 6 항에 있어서, 상기 방법은 연장 프로브가 결찰 단계에서 연장가능한 말단에 결찰되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 11

제 6 항에 있어서, 상기 생성 단계는 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 12

제 11 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 13

제 6 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 14

제 6 항에 있어서, 상기 단계 (a)는 별도의 분취액으로, 다수의 구별되는 프로브-주형 이중나선을 제공하는 것을 포함하고, 각각의 구별되는 이중나선은 주형 폴리뉴클레오티드에 혼성화된 개시 올리고뉴클레오티드 프로브를 포함하되, 각 이중나선의 주형 폴리뉴클레오티드는 동일하지만, 각 이중나선의 개시 올리고뉴클레오티드 프로브는 주형 폴리뉴클레오티드의 상이한 서열에 결합되며, 상기 단계 (b) 내지 (e)가 각 분취액과는 독립적으로 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 15

제 14 항에 있어서, 상기 각 분취액에 대하여, 상기 연장 올리고뉴클레오티드 프로브가 한 말단에서 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 16

제 15 항에 있어서, 상기 각 분취액에 대하여, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 17

제 15 항에 있어서, 상기 각 분취액에 대하여, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오티드 트리포스페이트의 존재하에 핵산 중합효소를 사용하여 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 18

제 15 항에 있어서, 상기 방법은 연장 프로브가 결찰 단계에서 연장가능한 말단에 결찰되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 19

제 15 항에 있어서, 상기 생성 단계는 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제를 사용하여 포스포로티올레이트 연쇄를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 20

제 19 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 21

제 15 항에 있어서, 상기 절찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 22

제 6 항에 있어서, 상기 방법은 (f) 절찰된 프로브와 개시 올리고뉴클레오티드를 주형으로부터 제거하는 단계; (g) 주형 폴리뉴클레오티드의 상이한 서열에 결합된 두 번째 올리고뉴클레오티드를 사용하여 단계 (a)를 반복하는 단계; 및 (h) 단계 (b) 내지 (e)를 반복하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 23

제 22 항에 있어서, 상기 방법은 주형 폴리뉴클레오티드의 상이한 서열에 결합된 개시 올리고뉴클레오티드를 사용하여 여러 번 반복되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 24

제 23 항에 있어서, 상기 연장 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 25

제 23 항에 있어서, 상기 각 반복에 대하여, 상기 확인 단계는 가장 최근에 절찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 26

제 23 항에 있어서, 상기 각 반복에 대하여, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오티드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 27

제 23 항에 있어서, 상기 방법은 연장 프로브가 절찰 단계에서 연장가능한 말단에 절찰되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 28

제 23 항에 있어서, 상기 생성 단계는 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 29

제 28 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 30

제 23 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 31

제 22 항에 있어서, 상기 제거 단계는 결찰된 프로브, 개시 올리고뉴클레오티드, 및 주형을 물에 약 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15 mM의 중황산 나트륨 (NaHSO₄)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 32

제 22 항에 있어서, 상기 제거 단계는 결찰된 프로브, 개시 올리고뉴클레오티드, 및 주형을 약 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO₄), 예컨대 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO₄)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 33

부착 지점에서 지지체에 부착된 주형 폴리뉴클레오티드의 뉴클레오티드 서열을 확인하는 방법에 있어서,

(a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 올리고뉴클레오티드 프로브를 거기에 결합함으로써 개시 올리고뉴클레오티드를 연장하되, 이때 연장은 주형을 따라 지지체에 대한 부착 지점을 향하여 진행되는 단계;

(b) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및

(c) 단계 (a)와 (b)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 부착 지점에서 지지체에 부착된 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 34

제 33 항에 있어서, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 35

제 33 항에 있어서, 상기 올리고뉴클레오티드 프로브가 포스포로티올레이트 연쇄를 함유하고, 상기 생성 단계는 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 36

제 35 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 37

제 35 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 38

부착 지점에서 지지체에 부착된 주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

(a) 주형 폴리뉴클레오티드에 혼성된, 연장가능한 말단을 가지는 프로브를 포함하는 프로브-주형 이중나선을 제공하는 단계;

(b) 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단

에 연장 올리고뉴클레오타이드 프로브를 결합시키는 단계;

(c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오타이드 프로브의 바로 아래에 있는 주형 폴리뉴클레오타이드의 뉴클레오타이드 잔기인 주형 폴리뉴클레오타이드의 최소한 하나의 뉴클레오타이드를 확인하는 단계;

(d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오타이드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및

(e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오타이드에 있는 뉴클레오타이드의 서열이 결정될 때까지 반복하는 단계를 포함하는 부착 지점에서 지지체에 부착된 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 39

제 38 항에 있어서, 상기 연장 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 40

제 38 항에 있어서, 상기 확인 단계는 가장 최근에 결합된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 41

제 38 항에 있어서, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오타이드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오타이드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 42

제 38 항에 있어서, 상기 방법은 연장 프로브가 결합 단계에서 연장가능한 말단에 결합되지 않을 때면 언제나 연장된 올리고뉴클레오타이드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 43

제 38 항에 있어서, 상기 올리고뉴클레오타이드 프로브가 포스포로티올레이트 연쇄를 함유하고, 상기 생성 단계는 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 44

제 43 항에 있어서, 상기 절단제는 AgNQ인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 45

제 38 항에 있어서, 상기 방법은 (f) 결합된 프로브와 개시 올리고뉴클레오타이드를 주형으로부터 제거하는 단계; (g) 주형 폴리뉴클레오타이드의 상이한 서열에 결합된 두 번째 올리고뉴클레오타이드를 사용하여 단계 (a)를 반복하는 단계; 및 (h) 단계 (b) 내지 (e)를 반복하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 46

제 45 항에 있어서, 상기 방법은 주형 폴리뉴클레오타이드의 상이한 서열에 결합된 개시 올리고뉴클레오타이드를 사용하여 여러 번 반복되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 47

제 38 항에 있어서, 상기 결합 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 48

제 38 항에 있어서, 상기 주형이 실질적으로 평면이고 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 49

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 확인하는 방법에 있어서,

- (a) 반-고체 지지체에 또는 그 위에 고정된 마이크로입자에 부착된 주형 폴리뉴클레오티드를 제공하는 단계;
- (b) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 올리고뉴클레오티드 프로브를 거기에 결합함으로써 개시 올리고뉴클레오티드를 연장하되, 올리고뉴클레오티드 프로브는 잘라지기 쉬운 연쇄를 포함하는 단계;
- (c) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및
- (d) 단계 (b)와 (c)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 50

제 49 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 51

제 49 항에 있어서, 상기 주형이 실질적으로 평면이고 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 52

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

- (a) 주형 폴리뉴클레오티드에 혼성된, 연장가능한 말단을 가지는 프로브를 포함하는 프로브-주형 이중나선을 제공하되, 프로브-주형 이중나선은 반-고체 지지체에 또는 그 위에 삽입된 마이크로입자에 부착되는 단계;
- (b) 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단에 연장 올리고뉴클레오티드 프로브를 결합시키되, 연장 프로브는 포스포로티올레이트 연쇄를 포함하는 단계;
- (c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오티드 프로브의 바로 아래에 있는 주형 폴리뉴클레오티드의 뉴클레오티드 잔기인 주형 폴리뉴클레오티드의 최소한 하나의 뉴클레오티드를 확인하는 단계;
- (d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오티드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및
- (e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오티드에 있는 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 53

제 52 항에 있어서, 상기 결합 및 생성 단계는 반-고체 지지체에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 54

제 52 항에 있어서, 상기 주형이 실질적으로 평면이고 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 55

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

- (a) 마이크로입자의 존재 하에 에멀션의 구획에서 주형 폴리뉴클레오티드 분자를 증폭시킴으로써 주형 폴리뉴클레오티드의 클론 집단이 부착된 마이크로입자가 생성되는 단계;
- (b) 에멀션으로부터 마이크로입자를 회수하는 단계;
- (c) 반-고체 지지체에 또는 그 위에 마이크로입자를 삽입하는 단계;
- (d) 연장된 이중나선을 형성하기 위하여, 주형 폴리뉴클레오티드를 따라 거기에 올리고뉴클레오티드 프로브를 결합시킴으로써 개시 올리고뉴클레오티드를 연장시키되, 올리고뉴클레오티드 프로브는 잘라지기 쉬운 연쇄를 포함하는 단계;
- (e) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및
- (f) 단계 (d)와 (e)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 56

제 55 항에 있어서, 상기 방법은 (i) 상이한 서열을 포함하는 다수의 주형 폴리뉴클레오티드 분자가 에멀션의 개별적인 구획에서 증폭되고; (ii) 각각에 상이한 서열을 가지는 주형 폴리뉴클레오티드의 클론 집단이 부착된 다수의 마이크로입자가 에멀션으로부터 회수되고 지지체에 또는 그 위에 삽입되며; (iii) 단계 (d), (e) 및 (f)가 삽입된 마이크로입자에 부착된 클론 집단에 대해 동시에 수행됨으로써 다수의 서열이 동시에 결정되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 57

최소한 2개의 구별가능하게 표지된 올리고뉴클레오티드 프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보를 결정하는 방법에 있어서,

- (a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 올리고뉴클레오티드 프로브를 거기에 결합시킴으로써 개시 올리고뉴클레오티드를 연장하되, 올리고뉴클레오티드 프로브는 구별가능하게 표지된 올리고뉴클레오티드 프로브 군의 한 구성원인 단계;
- (b) 올리고뉴클레오티드 프로브와 결합된 표지를 검출하는 단계; 및
- (c) 프로브 군 이름의 순서 리스트가 얻어질 때까지 단계 (a)와 (b)를 반복하는 단계; 및
- (d) 프로브 군 이름의 순서 리스트를 사용하여 뉴클레오티드의 서열에 대한 하나 또는 그 이상의 가능성을 제거하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 58

제 57 항에 있어서, 상기 단계 (d)가 프로브 군 이름의 순서 리스트의 코드를 해석하여 서열을 결정하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 59

제 57 항에 있어서, 상기 방법은 주형 폴리뉴클레오티드에 혼성된, 연장가능한 말단을 가지는 개시 올리고뉴클레오티드 프로브를 포함하는 프로브-주형 이중나선을 제공하는 것을 포함하고, 연장 단계는 상기 연장가능한 말단에 올리고뉴클레오티드 프로브를 결합시킴으로써 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하되, 연장 단계에서 연장가능한 말단에 더 이상 올리고뉴클레오티드 프로브가 결합되지 않을 때면 언제나 어떠한 남아있는 연장가능한 말단을 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 60

제 57 항에 있어서, 각 프로브 군의 올리고뉴클레오티드 프로브가 한 말단에 비-연장성 부분을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 61

제 57 항에 있어서, 상기 방법은 각각의 검출 단계 후에, (f) 만약 연장가능한 말단이 이미 존재하는 것이 아니라면, 생성된 말단이 가장 최근에 결찰된 올리고뉴클레오타이드 프로브가 결찰된 말단과는 상이하도록 가장 최근에 결찰된 올리고뉴클레오타이드 프로브에 대해 연장가능한 말단을 생성하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 62

제 61 항에 있어서, 상기 올리고뉴클레오타이드 프로브는 포스포로티올레이트 연쇄를 포함하고, 연장가능한 프로브 말단은 Ag, Hg, Cu, Mn, Zn 및 Cd로 이루어지는 군으로부터 선택된 원자를 포함하는 절단제로 포스포로티올레이트 연쇄를 절단함으로써 생성되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 63

제 62 항에 있어서, 상기 절단제는 AgNO_3 인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 64

제 57 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 65

제 57 항에 있어서, 상기 주형이 실질적으로 평면인 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 66

제 57 항에 있어서, 상기 집합은 2개의 구별가능하게 표시된 프로브 군을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 67

제 57 항에 있어서, 상기 집합은 3개의 구별가능하게 표시된 프로브 군을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 68

제 57 항에 있어서, 상기 집합은 4개의 구별가능하게 표시된 프로브 군을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 69

제 57 항에 있어서, 상기 집합은 4개 이상의 구별가능하게 표시된 프로브 군을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 70

제 57 항에 있어서, 상기 올리고뉴클레오타이드 프로브가 뉴클레오타이드들이 독립적으로 선택되지 않는 구속된 부분을 포함하고, 상기 서열이 상이한 구속된 부분을 가지는 올리고뉴클레오타이드 프로브가 코드화에 따라 프로브 군에 지정되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 71

제 57 항에 있어서, 상기 올리고뉴클레오타이드 프로브가 표 1에 제시된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째 및 네 번째 프로브 군으로 지정되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보

결정 방법.

청구항 72

제 58 항에 있어서, 주형의 최소한 하나의 뉴클레오타이드가 공지된 정체를 가지되, 코드해석 단계는:

- (i) 공지된 뉴클레오타이드의 정체 및 그것의 근접한 뉴클레오타이드가 공지된 정체의 뉴클레오타이드에 인접한 뉴클레오타이드의 반대쪽에서 결찰된 프로브의 구축된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 공지된 정체의 뉴클레오타이드에 인접한 주형의 뉴클레오타이드에 정체를 지정하는 단계;
- (ii) 그것의 근접한 뉴클레오타이드가 연속되는 뉴클레오타이드의 반대쪽에서 결찰된 프로브의 구축된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계; 및
- (iii) 서열이 결정될 때까지 단계 (ii)를 반복하는 단계들을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 73

제 58 항에 있어서, 상기 방법은,

- (a) 뉴클레오타이드가 공지된 정체를 가지도록 주형의 뉴클레오타이드의 정체를 결정하는 단계를 더 포함하되, 코드 해석 단계는:
 - (i) 공지된 뉴클레오타이드의 정체 및 그것의 근접한 뉴클레오타이드가 공지된 정체의 뉴클레오타이드에 인접한 뉴클레오타이드의 반대쪽에서 결찰된 프로브의 구축된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 공지된 정체의 뉴클레오타이드에 인접한 주형의 뉴클레오타이드에 정체를 지정하는 단계;
 - (ii) 그것의 근접한 뉴클레오타이드가 연속되는 뉴클레오타이드의 반대쪽에서 결찰된 프로브의 구축된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계; 및
 - (iii) 서열이 결정될 때까지 단계 (ii)를 반복하는 단계들을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 74

제 73 항에 있어서, 상기 결정 단계는 만약 이중나선에 인접한 위치에서 주형에 상보한다면 표지된 뉴클레오타이드의 통합이 허용되는 조건하에서 중합효소의 존재 하에 표지된 뉴클레오타이드와 주형-프로브 이중나선을 접촉시키는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 75

제 58 항에 있어서, 상기 코드해석 단계는 프로브 군 이름의 순서 리스트로부터 최소한 하나의 후보 서열을 생성하는 단계와, 주형의 뉴클레오타이드 서열로서 후보 서열을 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 76

제 75 항에 있어서, 상기 생성 단계는 최소한 4개의 후보 서열을 생성하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 77

제 75 항에 있어서, 상기 생성 단계는,

- (i) 뉴클레오타이드 서열의 첫 번째 뉴클레오타이드에 대하여 정체를 가정하는 단계;
- (ii) 첫 번째 뉴클레오타이드에 상응하는 프로브 군의 이름을 토대로 하여 인접한 뉴클레오타이드에 대하여 가능한 정체를 결정함으로써 첫 번째 뉴클레오타이드에 인접한 뉴클레오타이드에 대하여 정체를 지정하는 단계;
- (iii) 그것의 정체가 가장 최근에 지정된 뉴클레오타이드에 상응하는 프로브 군의 이름을 토대로 하여 연속되는 뉴클레오타이드에 대하여 가능한 정체를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계;

(iv) 후보 서열이 생성될 때까지 단계 (iii)을 반복하는 단계; 및

(v) 원하는 수의 후보 서열이 생성될 때까지, 각 반복시 상이한 정체가 첫 번째 뉴클레오타이드에 대해 가정되는 단계 (i)-(iv)를 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 78

제 75 항에 있어서, 상기 선택 단계는 최소한 하나의 후보 서열을 하나 또는 그 이상의 공지된 서열과 비교하고, 예정된 정도의 동일성이 하나 또는 그 이상의 공지된 서열과 가장 근접하게 동일한 것을 나타내는 후보 서열을 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 79

제 78 항에 있어서, 상기 주형이 관심 있는 유기체로부터 유도되고, 상기 비교 단계는 최소한 하나의 후보 서열을 그 유기체로부터 얻어진 서열을 함유하는 데이터베이스의 서열과 비교하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 80

제 78 항에 있어서, 상기 비교 단계는 최소한 하나의 후보 서열을 다수의 비교 서열을 함유하는 데이터베이스의 서열과 비교하는 것을 포함하고, 다수의 비교 서열의 각각은 결정된 폴리뉴클레오타이드의 서열에 대해 가능한 대체 서열을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 81

제 75 항에 있어서, 상기 선택 단계는:

(i) 구별가능하게 표지된 코드화된 프로브 군의 두 번째 집합을 사용하여 주형으로부터 프로브 군 이름의 두 번째 순서 리스트를 얻되, 프로브 군의 두 번째 집합의 프로브 군은 프로브 군의 첫 번째 집합의 프로브 군에 상이하게 코드화되는 단계;

(ii) 프로브 군 이름의 두 번째 순서 리스트로부터 최소한 하나의 비교 서열을 생성하는 단계;

(iii) 최소한 하나의 후보 서열의 부분을 최소한 하나의 비교 서열의 부분과 비교하는 단계; 및

(iv) 단계 (c)에서 비교된 부분에 걸쳐 비교서열과 예정된 수준의 동일성을 나타내거나 가장 가깝게 동일한 후보 서열을 주형의 뉴클레오타이드 서열로서 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 82

제 81 항에 있어서, 상기 비교된 부분이 단일 디뉴클레오타이드인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 83

제 81 항에 있어서, 상기 두 번째 프로브 군 이름의 순서 리스트가 단지 하나의 요소를 함유하는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 84

제 57 항에 있어서, 상기 각 프로브 군의 올리고뉴클레오타이드 프로브가 구조 $5'-(XY)(N)_k N_B^*-3'$ 또는 $3'-(XY)(N)_k N_B^*-5'$ 의 구조를 가지며, N은 어떠한 뉴클레오시드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, XY는 프로브의 구속된 부분이고, X와 Y는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않는 뉴클레오시드를 나타내며, X와 Y는 최소한 2-배 축퇴성이고, 최소한 하나의 뉴클레오시드간 연쇄는 잘라지기 쉬운 연쇄이며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 Y 상에 또는 $(N)_k$ 의 어떠한 뉴클레오시드 상에 존재할 수 있어야 하는 것인 주형

폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 85

제 84 항에 있어서, 상기 잘라지기 쉬운 연쇄가 포스포로티올레이트 연쇄인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 86

제 84 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 87

제 86 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 88

제 84 항에 있어서, 상기 4개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군이 사용되고, 프로브의 구속된 부분에 대한 상이한 서열을 가지는 올리고뉴클레오타이드 프로브는 표 1에 설명된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째, 및 네 번째 프로브 군으로 지정되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 89

제 57 항에 있어서, 상기 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 90

제 57 항에 있어서, 상기 검출 단계는 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 2보다 적은 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 91

최소한 2개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보를 결정하는 방법에 있어서,

- (a) 연장가능한 말단을 가지는 이중나선 부분과 서열화될 단일-가닥의 부분을 포함하는 프로브-주형 복합체를 최소한 두 개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군과 접촉시킴으로써, 이중나선 부분에 바로 인접한 주형의 부분에 상보하는 부분을 포함하는 올리고뉴클레오타이드 프로브 사이에 혼성화가 일어나도록 하는 단계;
- (b) 혼성된 올리고뉴클레오타이드 프로브를 연장가능한 말단에 결합시킴으로써 연장된 이중나선을 포함하는 프로브-주형 복합체를 생성하는 단계;
- (c) 결합된 프로브와 결합된 표지를 검출하는 단계;
- (d) 만약 이미 존재하는 것이 아니라면 연장된 이중나선 상에 연장가능한 프로브 말단을 생성하는 단계; 및
- (e) 프로브 군 이름의 순서 리스트가 얻어질 때까지 단계 (a)에서 (d)를 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 92

제 91 항에 있어서, 상기 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보 결정 방법.

청구항 93

제 91 항에 있어서, 상기 검출 단계는 주형의 최소한 2개의 뉴클레오티드의 각각으로부터 동시에 2보다 적은 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 94

프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보를 결정하는 방법에 있어서,

(a) 연속적인 사이클의 연장, 결찰, 검출, 및 절단을 수행하되, 검출 단계는 어떠한 개별적인 뉴클레오티드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2개의 뉴클레오티드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 단계; 및

(b) 단계 (a)에서 얻어진 정보를 최소한 한 비트의 추가의 정보와 조합하여 서열을 결정하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 95

제 94 항에 있어서, 상기 최소한 한 비트의 추가의 정보는 주형의 뉴클레오티드의 정체, 최소한 하나의 공지 서열과 후보 서열을 비교함으로써 얻어진 정보; 및 올리고뉴클레오티드 프로브 군의 두 번째 집합을 사용하여 방법을 반복함으로써 얻어진 정보로 이루어지는 군으로부터 선택된 항목을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열에 대한 정보 결정 방법.

청구항 96

서열화 에러로부터 단일 뉴클레오티드 다형성을 구별하는 방법에 있어서,

(a) 제 58 항의 방법을 사용하여 다수의 주형을 서열화하되, 주형은 단일 핵산 서열의 중복하는 단편들을 나타내는 단계;

(b) 단계 (a)에서 얻어진 서열들을 정렬하는 단계; 및

(c) 만약 서열이 실질적으로 첫 번째 부분을 가로질러 동일하다면 서열화 에러를 나타내는 서열과 실질적으로 두 번째 부분을 가로질러 상이한 서열 사이를 결정하되, 각 부분은 최소한 3 뉴클레오티드의 길이를 가지는 단계를 포함하는 단일 뉴클레오티드 다형성 구별 방법.

청구항 97

서열화 에러로부터 단일 뉴클레오티드 다형성을 구별하는 방법에 있어서,

(a) 단일 핵산 서열의 중복하는 단편들을 나타내는 다수의 주형을 사용하여 제 58 항의 단계 (a)부터 (c)를 수행함으로써 프로브 군의 다수의 순서 리스트를 얻는 단계;

(b) 단계 (a)에서 얻어진 프로브 군의 순서 리스트들을 정렬하여 최소한 90% 동일한 리스트 내에서 정렬된 영역을 얻는 단계; 및

(c) 만약 리스트가 정렬된 영역 내에서 오직 한 위치에서만 다르다면 서열화 에러를 나타내는 프로브 군의 순서 리스트들 사이의 차이를 결정하는 단계; 또는

(d) 만약 리스트가 정렬된 영역 내에서 둘 또는 그 이상의 인접한 위치에서 다르다면 단일 뉴클레오티드 다형성을 나타내는 프로브 군의 순서 리스트들 사이의 차이를 결정하는 단계를 포함하는 단일 뉴클레오티드 다형성 구별 방법.

청구항 98

최소한 2개의 구별가능하게 표지된 올리고뉴클레오티드 프로브 군의 집합으로서, 상기 각 프로브 군의 프로브들은 구속된 부분과 구속되지 않은 부분을 포함하고, 구속된 부분의 각 위치는 최소한 2-배 축퇴성이며, 각 군의 프로브들은 잘라지기 쉬운 뉴클레오티드 간 연쇄를 포함하는 집합.

청구항 99

제 98 항에 있어서, 상기 각 프로브가 결합 효소에 의해 연장될 수 없는 말단을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 100

제 98 항에 있어서, 상기 각 프로브가 결합 효소에 의해 연장될 수 없는 말단을 포함하고, 각 프로브가 잘라지기 쉬운 연쇄와 결합 효소에 의해 연장될 수 없는 말단 사이의 위치에서 검출가능한 부분을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 101

제 98 항에 있어서, 상기 잘라지기 쉬운 연쇄가 포스포로티올레이트 연쇄인 것인 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 102

제 98 항에 있어서, 상기 집합은 2개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 103

제 98 항에 있어서, 상기 집합은 3개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 104

제 98 항에 있어서, 상기 집합은 4개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 105

제 98 항에 있어서, 상기 집합은 4개 이상의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 106

제 98 항에 있어서, 상기 프로브가 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 검출가능한 부분을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 107

최소한 2개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합으로서, 상기 각 프로브 군의 올리고뉴클레오타이드 프로브가 구조 $5'-(X)_j(N)_kN_B-3'$ 또는 $3'-(X)_j(N)_kN_B-5'$ 을 가지며, N은 어떠한 뉴클레오시드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, $(X)_j$ 는 각 X가 뉴클레오시드를 나타내는 프로브의 구속된 부분이고, $(X)_j$ 의 뉴클레오시드들은 동일하거나 상이하지만 독립적으로 선택되지 않으며, 각 X는 최소한 2-배 축퇴성이고, j는 2와 5 사이에 있으며, k는 1과 100 사이에 있고, 포괄적이며, 각 프로브는 프로브 말단에 있는 $(X)_j$ 의 뉴클레오시드 이외의 위치에서 검출가능한 부분을 포함하고, 상기 각 프로브 군의 프로브들은 동일한 표지를 포함하고, 상이한 프로브 군의 프로브들은 상이한 구별가능한 표지를 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 108

제 107 항에 있어서, 상기 최소한 하나의 뉴클레오시드간 연쇄가 잘라지기 쉬운 연쇄인 것인 구별가능하게 표시된 코드화된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 109

제 107 항에 있어서, 상기 잘라지기 쉬운 연쇄가 포스포로티올레이트 연쇄인 것인 구별가능하게 표시된 올리고

뉴클레오티드 프로브 군의 집합.

청구항 110

제 107 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 111

제 110 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 112

제 107 항에 있어서, 상기 세트가 4개의 프로브 군으로 구성되고, 상기 각 프로브 군의 올리고뉴클레오티드 프로브가 구조 $5'-(XY)(N)_kN_B^*-3'$ 또는 $3'-(XY)(N)_kN_B^*-5'$ 의 구조를 가지며, N은 어떠한 뉴클레오시드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, XY는 프로브의 구속된 부분이고, X와 Y는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않는 뉴클레오시드를 나타내며, X와 Y는 최소한 2-배 축퇴성이고, 최소한 하나의 뉴클레오시드간 연쇄는 잘라지기 쉬운 연쇄이며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오시드 상에 또는 Y 상에 존재할 수 있어야 하는 것인 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 113

제 112 항에 있어서, 상기 잘라지기 쉬운 연쇄가 포스포로티올레이트 연쇄인 것인 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 114

제 112 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 115

제 114 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 116

제 112 항에 있어서, 상기 프로브의 구속된 부분에 대한 상이한 서열을 가지는 올리고뉴클레오티드 프로브가 표 1에 설명된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째, 및 네 번째 프로브 군으로 지정되는 것인 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 117

최소한 2개의 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합으로서, 상기 각 프로브 군의 올리고뉴클레오티드 프로브가 구조 $5'-(X)_j(N)_kN_B-3'$ 또는 $3'-(X)_j(N)_kN_B-5'$ 을 가지며, N은 어떠한 뉴클레오시드 또는 비염기성 잔기를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, $(X)_j$ 는 각 X가 뉴클레오시드 또는 비염기성 잔기를 나타내는 프로브의 구속된 부분이고, 단 X_j 는 뉴클레오티드를 나타내고 $(X)_j$ 의 뉴클레오시드들은 동일하거나 상이하지만 독립적으로 선택되지 않으며, 각 X는 최소한 2-배 축퇴성이고, j는 2와 5 사이에 있으며, k는 1과 100 사이에 있고, 포괄적이며, 각 프로브는 프로브 말단에 있는 $(X)_j$ 의 뉴클레오시드 이외의 위치에서 검출가능한 부분을 포함하고, 상기 각 프로브 군의 프로브들은 동일한 표지를 포함하고, 상이한 프로브 군의 프로브들은 상이한 구별가능한 표지를 포함하는 구별가능하게 표시된 올리고뉴클레오티드 프로브 군의 집합.

청구항 118

제 117 항에 있어서, 상기 최소한 하나의 뉴클레오타이드간 연쇄가 잘라지기 쉬운 연쇄인 것인 구별가능하게 표지된 코드화된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 119

제 117 항에 있어서, 상기 잘라지기 쉬운 연쇄가 뉴클레오타이드와 비염기성 잔기 사이에 있는 것인 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 120

제 117 항에 있어서, 상기 올리고뉴클레오타이드 프로브가 트리거 잔기를 포함하는 구별가능하게 표지된 코드화된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 121

제 117 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 122

제 121 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 123

제 117 항에 있어서, 상기 세트가 4개의 프로브 군으로 구성되고, 상기 각 프로브 군의 올리고뉴클레오타이드 프로브가 구조 $5'-(XY)(N)_kN_B^*-3'$ 또는 $3'-(XY)(N)_kN_B^*-5'$ 의 구조를 가지며, N은 어떠한 뉴클레오타이드 또는 비염기성 잔기를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, XY는 프로브의 구속된 부분이고, X와 Y는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않는 뉴클레오타이드를 나타내며, X와 Y는 최소한 2-배 축퇴성이고, 최소한 하나의 뉴클레오타이드간 연쇄는 잘라지기 쉬운 연쇄이며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오타이드 상에 또는 Y 상에 존재할 수 있어야 하는 것인 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 124

제 123 항에 있어서, 상기 잘라지기 쉬운 연쇄가 뉴클레오타이드와 비염기성 잔기 사이에 있는 것인 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 125

제 123 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 126

제 125 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 127

제 123 항에 있어서, 상기 올리고뉴클레오타이드 프로브가 트리거 잔기를 포함하는 구별가능하게 표지된 코드화된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 128

제 123 항에 있어서, 상기 프로브의 구속된 부분에 대한 상이한 서열을 가지는 올리고뉴클레오타이드 프로브가 표 1에 설명된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째, 및 네 번째 프로브 군으로 지정되는 것

인 구별가능하게 표지된 프로브 군의 집합.

청구항 129

최소한 2개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합을 포함하는 키트.

청구항 130

제 129 항에 있어서, 상기 프로브가 잘라지기 쉬운 뉴클레오시드간 연쇄를 포함하는 키트.

청구항 131

제 130 항에 있어서, 상기 잘라지기 쉬운 뉴클레오시드간 연쇄가 포스포로티올레이트 연쇄인 것인 키트.

청구항 132

제 129 항에 있어서, 상기 키트가 결합 효소, 포스포로티에이트 연쇄를 절단할 수 있는 제제, 포스파타제, 중합 효소, 지지체, 완충제, 열안정성 중합효소, 뉴클레오타이드, 에멀션을 제조하기 위한 시약, 및 겔을 제조하기 위한 시약으로 이루어지는 군으로부터 선택된 최소한 하나의 항목을 더 포함하는 키트.

청구항 133

다수의 주형 폴리뉴클레오타이드를 제조하는 방법에 있어서,

(a) 다수의 마이크로입자를 가역성 반-고체 지지체에 또는 그 위에 삽입하여 첫 번째 배열의 마이크로입자를 형성하는 단계; 및

(b) 반-고체 지지체 내에서 출발 주형 폴리뉴클레오타이드를 증폭시킴으로써 증폭 후에 각 마이크로입자가 분자들이 부착된 주형의 클론 집단을 가지는 단계를 포함하는 다수의 주형 폴리뉴클레오타이드 제조 방법.

청구항 134

제 133 항에 있어서, 상기 방법은,

(a) 반-고체 지지체를 용해시키는 단계; 및 (b) 마이크로입자를 수집하는 단계를 더 포함하는 다수의 주형 폴리뉴클레오타이드 제조 방법.

청구항 135

제 134 항에 있어서, 상기 방법은 다른 반-고체 지지체에서 또는 그 위에서 마이크로입자의 두 번째 배열을 형성하는 단계를 더 포함하되, 두 번째 배열의 마이크로입자는 그것들이 첫 번째 배열에 존재하는 것보다 더 큰 밀도로 존재하는 다수의 주형 폴리뉴클레오타이드 제조 방법.

청구항 136

포스포로티올레이트 연쇄를 포함하는 올리고뉴클레오타이드 프로브를 포함하는 키트로서, 상기 프로브가 검출가능한 부분으로 표지되는 것인 키트.

청구항 137

제 136 항에 있어서, 상기 검출가능한 부분이 형광 염료인 것인 키트.

청구항 138

제 136 항에 있어서, 상기 키트가 포스포로티올레이트 연쇄를 절단할 수 있는 제제를 더 포함하는 키트.

청구항 139

제 136 항에 있어서, 상기 키트가 결합 효소를 더 포함하는 키트.

청구항 140

제 136 항에 있어서, 상기 키트가 결합 효소와 포스포로티올레이트 연쇄를 절단할 수 있는 제제를 더 포함하는

키트.

청구항 141

제 136 항에 있어서, 상기 키트가 결합 효소, 포스포로티에이트 연쇄를 절단할 수 있는 제제, 포스파타제, 중합 효소, 지지체, 완충제, 열안정성 중합효소, 뉴클레오티드, 에멀션을 제조하기 위한 시약, 및 겔을 제조하기 위한 시약으로 이루어지는 군으로부터 선택된 최소한 하나의 항목을 더 포함하는 키트.

청구항 142

제 136 항에 있어서, 상기 키트가 포스포로티올레이트 연쇄를 포함하는 형광 표지된 올리고뉴클레오티드 프로브를 다수 함유하여서 프로브의 상이한 말단 뉴클레오티드에 상응하는 프로브가 직접 스펙트럼으로 분해가능한 형광 염료를 운반하는 것인 키트.

청구항 143

형태 $5'-O-P-O-X-O-P-S-(N)_k N_B^*-3'$ 의 올리고뉴클레오티드로서, 상기 N은 어떠한 뉴클레오티드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오티드를 나타내며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오티드 상에 존재할 수 있어야 하는 것인 올리고뉴클레오티드 프로브.

청구항 144

제 143 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오티드를 포함하는 올리고뉴클레오티드 프로브.

청구항 145

제 143 항에서 설명된 것과 같은 올리고뉴클레오티드 프로브의 세트로서, 상기 세트가 다수의 형광 표지된 올리고뉴클레오티드 프로브를 함유함으로써 프로브의 상이한 뉴클레오티드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오티드 프로브 세트.

청구항 146

형태 $5'-N_B^*(N)_k-S-P-O-X-3'$ 의 올리고뉴클레오티드 프로브로서, 상기 N은 어떠한 뉴클레오티드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오티드를 나타내며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오티드 상에 존재할 수 있어야 하는 것인 올리고뉴클레오티드 프로브.

청구항 147

제 146 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오티드를 포함하는 올리고뉴클레오티드 프로브.

청구항 148

제 146 항에서 설명된 것과 같은 올리고뉴클레오티드 프로브의 세트로서, 상기 세트가 다수의 형광 표지된 올리고뉴클레오티드 프로브를 함유함으로써 프로브의 상이한 뉴클레오티드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오티드 프로브 세트.

청구항 149

형태 $5'-O-P-O-X-O-(N)_k-O-P-S-(N)_i N_B^*-3'$ 의 올리고뉴클레오티드 프로브로서, 상기 N은 어떠한 뉴클레오티드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오티드를 나타내며, $(k+i)$ 는 1과 100 사이에 있고, k는 1과 100 사이에 있으며, i는 0과 99 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_i$ 의 어떠한 뉴클레오티드 상에 존재할 수 있어야 하는 것인 올

리고뉴클레오타이드 프로브.

청구항 150

제 149 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오타이드를 포함하는 올리고뉴클레오타이드 프로브.

청구항 151

제 149 항에 있어서, 상기 $i=0$ 인 것인 올리고뉴클레오타이드 프로브.

청구항 152

제 149 항에서 설명된 것과 같은 올리고뉴클레오타이드 프로브의 세트로서, 상기 세트가 다수의 형광 표지된 올리고뉴클레오타이드 프로브를 함유함으로써 프로브의 상이한 뉴클레오타이드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오타이드 프로브 세트.

청구항 153

형태 $5'-N_B^*(N)_i-S-P-O-(N)_k-O-P-O-X-3'$ 의 올리고뉴클레오타이드 프로브로서, 상기 N은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오타이드를 나타내며, $(k+i)$ 는 1과 100 사이에 있고, k는 1과 100 사이에 있으며, i는 0과 99 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_i$ 의 어떠한 뉴클레오타이드 상에 존재할 수 있어야 하는 것인 올리고뉴클레오타이드 프로브.

청구항 154

제 153 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오타이드를 포함하는 올리고뉴클레오타이드 프로브.

청구항 155

제 153 항에 있어서, 상기 $i=0$ 인 것인 올리고뉴클레오타이드 프로브.

청구항 156

제 153 항에서 설명된 것과 같은 올리고뉴클레오타이드 프로브의 세트로서, 상기 세트가 다수의 형광 표지된 올리고뉴클레오타이드 프로브를 함유함으로써 프로브의 상이한 뉴클레오타이드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오타이드 프로브 세트.

청구항 157

$3'-XNNNNsINI-5'$, $3'-XNNNNsIII-5'$, $3'-XNNNNsNII-5'$, 및 $3'-XNNNNIsII-5'$ 으로 이루어지는 군으로부터 선택되는 형태의 올리고뉴클레오타이드 프로브로서, 상기 X와 N이 어떠한 뉴클레오타이드를 나타내고, "s"는 잘라지기 쉬운 연쇄를 나타내며, 잘라지기 쉬운 연쇄와 올리고뉴클레오타이드의 5' 말단 사이에 있는 최소한 하나의 잔기는 X의 정체에 상응하는 표지를 포함하는 올리고뉴클레오타이드 프로브.

청구항 158

제 157 항에 있어서, 상기 s가 포스포로티올레이트 연쇄를 나타내는 것인 올리고뉴클레오타이드 프로브.

청구항 159

두 번째 폴리뉴클레오타이드에 첫 번째 폴리뉴클레오타이드를 결합하는 방법에 있어서,

- (a) 반-고체 지지체에 또는 그 위에 고정된 첫 번째 폴리뉴클레오타이드를 제공하는 단계;
- (b) 첫 번째 폴리뉴클레오타이드를 두 번째 폴리뉴클레오타이드 및 결합 효소와 접촉시키는 단계; 및
- (c) 결합에 적당한 조건하에서 결합 효소의 존재 하에 첫 번째와 두 번째 폴리뉴클레오타이드를 유지하는 단계를

포함하는 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 160

제 159 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 공유 또는 비공유 연쇄를 통하여 반-고체 지지체에 직접 부착되는 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 161

제 159 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 반-고체 지지체에 또는 그 위에 고정된 지지체에 부착되는 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 162

제 161 항에 있어서, 상기 반-고체 지지체가 겔이고 지지체가 마이크로입자인 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 163

제 161 항에 있어서, 상기 반-고체 지지체가 겔이고 지지체가 자성 마이크로입자인 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 164

폴리뉴클레오티드를 절단하는 방법에 있어서,

- (a) 반-고체 지지체에 또는 그 위에 고정된 폴리뉴클레오티드를 제공하되, 폴리뉴클레오티드는 잘라지기 쉬운 연쇄를 포함하는 단계;
- (b) 폴리뉴클레오티드를 절단제와 접촉시키는 단계; 및
- (c) 절단에 적당한 조건하에서 절단제의 존재 하에 폴리뉴클레오티드를 유지하는 단계를 포함하는 폴리뉴클레오티드 절단 방법.

청구항 165

제 164 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 공유 또는 비공유 연쇄를 통하여 반-고체 지지체에 직접 부착되는 것인 폴리뉴클레오티드 절단 방법.

청구항 166

제 164 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 반-고체 지지체에 또는 그 위에 고정된 지지체에 부착되는 것인 폴리뉴클레오티드 절단 방법.

청구항 167

제 166 항에 있어서, 상기 반-고체 지지체가 겔이고 상기 지지체가 마이크로입자인 것인 폴리뉴클레오티드 절단 방법.

청구항 168

제 166 항에 있어서, 상기 반-고체 지지체가 겔이고 상기 지지체가 자성 마이크로입자인 것인 폴리뉴클레오티드 절단 방법.

청구항 169

중량에 의한 기포 치환을 제공하기 위해 배향된 유동 셀을 포함하는 자동 서열화 기기.

청구항 170

초당 40,000개의 뉴클레오티드를 확인할 수 있는 자동 서열화 시스템.

청구항 171

하루에 8.6Gb의 서열 정보를 생성하는 자동 서열화 시스템.

청구항 172

하루에 48Gb의 서열 정보를 생성하는 자동 서열화 시스템.

청구항 173

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 확인하는 방법에 있어서,

(a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 거기에 올리고뉴클레오티드 프로브를 결합시킴으로써 개시 올리고뉴클레오티드를 연장하되, 올리고뉴클레오티드 프로브는 트리거 잔기를 포함하는 단계;

(b) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계;

(c) 올리고뉴클레오티드 프로브를 절단제를 사용하여 절단함으로써 연장가능한 프로브 말단을 생성하는 단계; 및

(d) 뉴클레오티드 서열이 결정될 때까지 단계 (a), (b) 및 (c)를 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 174

제 173 항에 있어서, 상기 확인 단계는 가장 최근에 결합된 올리고뉴클레오티드 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 175

제 173 항에 있어서, 상기 올리고뉴클레오티드 프로브가 비염기성 잔기, 손상된 염기, 또는 테옥시이노신을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 176

제 173 항에 있어서, 상기 올리고뉴클레오티드 프로브가 손상된 염기를 포함하고, 상기 방법은 손상된 염기를 제거하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 177

제 176 항에 있어서, 상기 제거 단계는 연장된 이중나선을 DNA 글리코실라제와 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 178

제 173 항에 있어서, 상기 방법은 올리고뉴클레오티드 프로브를 절단제로 절단함으로써 연장가능한 프로브 말단을 생성하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 179

제 178 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제, Endo V, 및 과요오드산염으로 이루어지는 군으로부터 선택되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 180

제 178 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 181

제 178 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열

확인 방법.

청구항 182

제 173 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 183

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

- (a) 주형 폴리뉴클레오티드에 혼성된, 연장가능한 말단을 가지는 프로브를 포함하는 프로브-주형 이중나선을 제공하는 단계;
- (b) 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단에 연장 올리고뉴클레오티드 프로브를 결합시키되, 연장 프로브는 트리거 잔기를 포함하는 단계;
- (c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오티드 프로브의 바로 아래에 있는 주형 폴리뉴클레오티드의 뉴클레오티드 잔기인 주형 폴리뉴클레오티드의 최소한 하나의 뉴클레오티드를 확인하는 단계;
- (d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오티드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및
- (e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오티드에 있는 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 184

제 183 항에 있어서, 상기 연장 프로브가 비염기성 잔기, 손상된 염기, 또는 테옥시이노신을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 185

제 183 항에 있어서, 상기 각 연장 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 186

제 183 항에 있어서, 상기 확인 단계는 가장 최근에 결합된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 187

제 183 항에 있어서, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오티드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 188

제 183 항에 있어서, 상기 방법은 연장 프로브가 결합 단계에서 연장가능한 말단에 결합되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 189

제 183 항에 있어서, 상기 생성 단계는 AP 엔도뉴클레아제 및 과요오드산염으로 이루어지는 군으로부터 선택되는 절단제를 사용하여 올리고뉴클레오티드 프로브를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 190

제 189 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 191

제 189 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 192

제 183 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 193

제 183 항에 있어서, 상기 (a) 단계는 별도의 분취액으로, 다수의 구별되는 프로브-주형 이중나선을 제공하는 것을 포함하고, 각각의 구별되는 이중나선은 주형 폴리뉴클레오티드에 혼성화된 개시 올리고뉴클레오티드 프로브를 포함하되, 각 이중나선의 주형 폴리뉴클레오티드는 동일하지만, 각 이중나선의 개시 올리고뉴클레오티드 프로브는 주형 폴리뉴클레오티드의 상이한 서열에 결합되며, 상기 단계 (b) 내지 (e)가 각 분취액과는 독립적으로 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 194

제 193 항에 있어서, 상기 각 분취액에 대해, 상기 연장 올리고뉴클레오티드 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 195

제 194 항에 있어서, 상기 각 분취액에 대하여, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 196

제 194 항에 있어서, 상기 각 분취액에 대하여, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오티드 트리포스페이트의 존재하에 핵산 중합효소를 사용하여 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 197

제 194 항에 있어서, 상기 방법은 연장 프로브가 결찰 단계에서 더 이상 연장가능한 말단에 결찰되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 198

제 194 항에 있어서, 상기 생성 단계는 AP 엔도뉴클레아제 및 과요오드산염으로 이루어지는 군으로부터 선택되는 절단제를 사용하여 올리고뉴클레오티드 프로브를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 199

제 198 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 200

제 198 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 201

제 194 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 202

제 183 항에 있어서, 상기 방법은 (f) 결찰된 프로브와 개시 올리고뉴클레오티드를 주형으로부터 제거하는 단계; (g) 주형 폴리뉴클레오티드의 상이한 서열에 결합된 두 번째 올리고뉴클레오티드를 사용하여 단계 (a)를 반복하는 단계; 및 (h) 단계 (b)에서 (e)를 반복하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 203

제 202 항에 있어서, 상기 방법은 주형 폴리뉴클레오티드의 상이한 서열에 결합된 개시 올리고뉴클레오티드를 사용하여 여러 번 반복되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 204

제 202 항에 있어서, 상기 제거 단계는 결찰된 프로브, 개시 올리고뉴클레오티드, 및 주형을 물에 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15 mM의 중황산 나트륨 (NaHSO_4)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 205

제 202 항에 있어서, 상기 제거 단계는 결찰된 프로브, 개시 올리고뉴클레오티드, 및 주형을 약 2%의 SDS, 약 200mM의 NaCl, 및 약 10mM의 중황산 나트륨 (NaHSO_4), 예컨대 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO_4)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 206

제 203 항에 있어서, 상기 연장 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 207

제 203 항에 있어서, 상기 각 반복에 대하여, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 208

제 203 항에 있어서, 상기 각 반복에 대하여, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오시드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오티드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 209

제 203 항에 있어서, 상기 방법은 연장 프로브가 결찰 단계에서 연장가능한 말단에 결찰되지 않을 때면 언제나 연장된 올리고뉴클레오티드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 210

제 203 항에 있어서, 상기 생성 단계는 AP 엔도뉴클레아제 및 과요오드산염으로 이루어지는 군으로부터 선택되는 절단제를 사용하여 올리고뉴클레오티드 프로브를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 211

제 210 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 212

제 210 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 213

제 203 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 214

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 확인하는 방법에 있어서,

(a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 올리고뉴클레오티드 프로브를 거기에 결찰함으로써 개시 올리고뉴클레오티드를 연장하되, 올리고뉴클레오티드 프로브는 트리거 잔기를 함유하고, 연장된 주형을 따라 지지체에 대한 부착 지점을 향하여 진행되는 단계;

(b) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및

(c) 단계 (a)와 (b)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 215

제 214 항에 있어서, 상기 확인 단계는 가장 최근에 결찰된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 216

제 214 항에 있어서, 상기 생성 단계는 AP 엔도뉴클레아제, Endo V, 및 과요오드산염으로 이루어지는 군으로부터 선택되는 절단제를 사용하여 올리고뉴클레오티드 프로브를 절단하는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 217

제 216 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 218

제 216 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 219

제 216 항에 있어서, 상기 결찰 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 220

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

(a) 주형 폴리뉴클레오티드에 혼성된 프로브를 포함하는 프로브-주형 이중나선을 제공하되, 프라이머는 연장가능한 말단을 가지는 단계;

(b) 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단

에 연장 올리고뉴클레오타이드 프로브를 결합시키되, 연장은 주형을 따라 지지체에 대한 부착 지점을 향하여 진행되고, 올리고뉴클레오타이드 프로브는 트리거 잔기를 함유하는 단계;

(c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오타이드 프로브의 바로 아래에 있는 주형 폴리뉴클레오타이드의 뉴클레오타이드 잔기인 주형 폴리뉴클레오타이드의 최소한 하나의 뉴클레오타이드를 확인하는 단계;

(d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오타이드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및

(e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오타이드에 있는 뉴클레오타이드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 221

제 220 항에 있어서, 상기 각 연장 프로브가 한 말단에 비-연장성 부분을 가지는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 222

제 220 항에 있어서, 상기 확인 단계는 가장 최근에 결합된 연장 프로브에 부착된 표지를 검출하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 223

제 220 항에 있어서, 상기 확인 단계는 상기 비-연장성 부분을 제거하고, 하나 또는 그 이상의 표지된 사슬-종결 뉴클레오타이드 트리포스페이트의 존재하에 핵산 중합효소로 상기 연장된 올리고뉴클레오타이드 프로브를 연장하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 224

제 220 항에 있어서, 상기 방법은 연장 프로브가 결합 단계에서 연장가능한 말단에 결합되지 않을 때면 언제나 연장된 올리고뉴클레오타이드 프로브를 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 225

제 220 항에 있어서, 상기 생성 단계는 절단제를 사용하여 올리고뉴클레오타이드 프로브를 절단하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 226

제 225 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제, Endo V, 및 과요오드산염으로 이루어지는 군으로부터 선택되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 227

제 225 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 228

제 225 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 결정 방법.

청구항 229

제 220 항에 있어서, 상기 방법은 (f) 결합된 프로브와 개시 올리고뉴클레오타이드를 주형으로부터 제거하는 단계; (g) 주형 폴리뉴클레오타이드의 상이한 서열에 결합된 두 번째 올리고뉴클레오타이드를 사용하여 단계 (a)를 반복하는 단계; 및 (h) 단계 (b)에서 (e)를 반복하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드

드 서열 결정 방법.

청구항 230

제 229 항에 있어서, 상기 방법은 주형 폴리뉴클레오티드의 상이한 서열에 결합된 개시 올리고뉴클레오티드를 사용하여 여러 번 반복되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 231

제 220 항에 있어서, 상기 결합 및 생성 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 232

제 220 항에 있어서, 상기 주형이 실질적으로 평면인 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 233

제 229 항에 있어서, 상기 제거 단계는 결합된 프로브, 개시 올리고뉴클레오티드, 및 주형을 물에 약 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15 mM의 중황산 나트륨 (NaHSO_4)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 234

제 229 항에 있어서, 상기 제거 단계는 결합된 프로브, 개시 올리고뉴클레오티드, 및 주형을 약 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO_4), 예컨대 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO_4)을 함유하고 있는 용액과 접촉시키는 것을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 235

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 확인하는 방법에 있어서,

- (a) 반-고체 지지체에 또는 그 위에 고정된 마이크로입자에 부착된 주형 폴리뉴클레오티드를 제공하는 단계;
- (b) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오티드를 따라 올리고뉴클레오티드 프로브를 거기에 결합함으로써 개시 올리고뉴클레오티드를 연장하되, 올리고뉴클레오티드 프로브는 트리거 잔기를 포함하는 단계;
- (c) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및
- (d) 단계 (b)와 (c)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 236

제 235 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 237

제 235 항에 있어서, 상기 주형이 실질적으로 평면인 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 238

제 235 항에 있어서, 상기 올리고뉴클레오티드가 잘라지기 쉬운 연쇄를 포함하거나 포함하도록 쉽게 변형될 수 있고, 잘라지기 쉬운 연쇄는 뉴클레오시드와 비염기성 잔기 사이에 있는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 확인 방법.

청구항 239

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

- (a) 주형 폴리뉴클레오티드에 혼성된, 연장가능한 말단을 가지는 프로브를 포함하는 프로브-주형 이중나선을 제공하되, 프로브-주형 이중나선은 반-고체 지지체에 또는 그 위에 삽입된 마이크로입자에 부착되는 단계;
- (b) 연장된 올리고뉴클레오티드 프로브를 함유하는 연장된 이중나선을 형성하기 위하여, 상기 연장가능한 말단에 연장 올리고뉴클레오티드 프로브를 결합시키되, 연장 프로브는 트리거 잔기를 포함하는 단계;
- (c) 연장된 이중나선에서, (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오티드 프로브의 바로 아래에 있는 주형 폴리뉴클레오티드의 뉴클레오티드 잔기인 주형 폴리뉴클레오티드의 최소한 하나의 뉴클레오티드를 확인하는 단계;
- (d) 연장가능한 말단이 이미 존재하는 것이 아니라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 올리고뉴클레오티드 프로브 상에서 연장가능한 말단을 생성하는 단계; 및
- (e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오티드에 있는 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 240

제 239 항에 있어서, 상기 결합 및 생성 단계는 반-고체 지지체에서 수행되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 241

제 239 항에 있어서, 상기 주형이 실질적으로 평면인 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 242

주형 폴리뉴클레오티드의 뉴클레오티드 서열을 결정하는 방법에 있어서,

- (a) 마이크로입자의 존재 하에 에멀션의 구획에서 주형 폴리뉴클레오티드 분자를 증폭시킴으로써 주형 폴리뉴클레오티드의 클론 집단이 부착된 마이크로입자가 생성되는 단계;
- (b) 에멀션으로부터 마이크로입자를 회수하는 단계;
- (c) 반-고체 지지체에 또는 그 위에 마이크로입자를 삽입하는 단계;
- (d) 연장된 이중나선을 형성하기 위하여, 주형 폴리뉴클레오티드를 따라 거기에 올리고뉴클레오티드 프로브를 결합시킴으로써 개시 올리고뉴클레오티드를 연장시키되, 올리고뉴클레오티드 프로브는 트리거 잔기를 포함하는 단계;
- (e) 폴리뉴클레오티드의 하나 또는 그 이상의 뉴클레오티드를 확인하는 단계; 및
- (f) 단계 (d)와 (e)를 뉴클레오티드의 서열이 결정될 때까지 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 243

제 242 항에 있어서, 상기 방법은 (i) 상이한 서열을 포함하는 다수의 주형 폴리뉴클레오티드 분자가 에멀션의 개별적인 구획에서 증폭되고; (ii) 각각에 상이한 서열을 가지는 주형 폴리뉴클레오티드의 클론 집단이 부착된 다수의 마이크로입자가 에멀션으로부터 회수되고 지지체에 또는 그 위에 삽입되며; (iii) 단계 (d), (e) 및 (f)가 삽입된 마이크로입자에 부착된 클론 집단에 대해 동시에 수행됨으로써 다수의 서열이 동시에 결정되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 결정 방법.

청구항 244

최소한 2개의 구별가능하게 표지된 올리고뉴클레오티드 프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레

오티드의 뉴클레오타이드 서열에 대한 정보를 결정하는 방법에 있어서,

(a) 연장된 이중나선을 형성하기 위하여 주형 폴리뉴클레오타이드를 따라 올리고뉴클레오타이드 프로브를 거기에 결합시킴으로써 개시 올리고뉴클레오타이드를 연장하되, 올리고뉴클레오타이드 프로브는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 한 구성원이고 트리거 잔기를 함유하는 단계;

(b) 올리고뉴클레오타이드 프로브와 결합된 표지를 검출하는 단계; 및

(c) 프로브 군 이름의 순서 리스트가 얻어질 때까지 단계 (a)와 (b)를 반복하는 단계; 및

(d) 프로브 군 이름의 순서 리스트를 사용하여 뉴클레오타이드의 서열에 대한 하나 또는 그 이상의 가능성을 제거하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 245

제 244 항에 있어서, 상기 단계 (d)는 프로브 군 이름의 순서 리스트의 코드를 해석하여 서열을 결정하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 246

제 244 항에 있어서, 상기 방법은 주형 폴리뉴클레오타이드에 혼성된, 연장가능한 말단을 가지는 개시 올리고뉴클레오타이드 프로브를 포함하는 프로브-주형 이중나선을 제공하는 것을 포함하고, 연장 단계는 상기 연장가능한 말단에 올리고뉴클레오타이드 프로브를 결합시킴으로써 연장된 올리고뉴클레오타이드 프로브를 함유하는 연장된 이중나선을 형성하되, 연장 단계에서 연장가능한 말단에 더 이상 올리고뉴클레오타이드 프로브가 결합되지 않을 때면 언제나 어떠한 남아있는 연장가능한 말단을 캡핑하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 247

제 244 항에 있어서, 상기 각 프로브 군의 올리고뉴클레오타이드 프로브는 한 말단에 비-연장성 부분을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 248

제 244 항에 있어서, 상기 방법은 각각의 검출 단계 후에, (f) 만약 연장가능한 말단이 이미 존재하는 것이 아니라면, 생성된 말단이 가장 최근에 결합된 올리고뉴클레오타이드 프로브가 결합된 말단과는 상이하도록 가장 최근에 결합된 올리고뉴클레오타이드 프로브에 대해 연장가능한 말단을 생성하는 단계를 더 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 249

제 248 항에 있어서, 상기 연장가능한 프로브 말단은 AP 엔도뉴클레아제, Endo V, 및 과요오드산염으로 이루어지는 군으로부터 선택되는 절단제를 사용하여 올리고뉴클레오타이드를 절단함으로써 생성되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 250

제 249 항에 있어서, 상기 절단제는 AP 엔도뉴클레아제인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 251

제 249 항에 있어서, 상기 절단제는 엔도뉴클레아제 VIII인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 252

제 244 항에 있어서, 상기 연장 단계는 반-고체 지지체에서 또는 그 위에서 수행되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 253

제 244 항에 있어서, 상기 주형이 실질적으로 평면인 견고한 기판에 부착된 마이크로입자에 부착되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 254

제 244 항에 있어서, 상기 집합은 2개의 구별가능하게 표지된 프로브 군을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 255

제 244 항에 있어서, 상기 집합은 3개의 구별가능하게 표지된 프로브 군을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 256

제 244 항에 있어서, 상기 집합은 4개의 구별가능하게 표지된 프로브 군을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 257

제 244 항에 있어서, 상기 집합은 4개 이상의 구별가능하게 표지된 프로브 군을 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 258

제 244 항에 있어서, 상기 올리고뉴클레오티드 프로브가 뉴클레오티드가 독립적으로 선택되지 않는 구속된 부분을 포함하고, 상기 서열이 다른 구속된 부분을 가지는 올리고뉴클레오티드 프로브가 코드화에 따라 프로브 군으로 지정되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 259

제 244 항에 있어서, 상기 올리고뉴클레오티드 프로브가 표 1에서 설명된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째, 및 네 번째 프로브 군으로 지정되는 것인 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 260

제 245 항에 있어서, 상기 주형의 최소한 하나의 뉴클레오티드가 공지된 정체를 가지되, 상기 코드해석 단계는:

- (i) 공지된 뉴클레오티드의 정체 및 그것의 근접한 뉴클레오티드가 공지된 정체의 뉴클레오티드에 인접한 뉴클레오티드의 반대쪽에서 결합된 프로브의 구속된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 공지된 정체의 뉴클레오티드에 인접한 주형의 뉴클레오티드에 정체를 지정하는 단계;
- (ii) 그것의 근접한 뉴클레오티드가 연속되는 뉴클레오티드의 반대쪽에서 결합된 프로브의 구속된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 연속되는 뉴클레오티드에 정체를 지정하는 단계; 및
- (iii) 서열이 결정될 때까지 단계 (ii)를 반복하는 단계를 포함하는 주형 폴리뉴클레오티드의 뉴클레오티드 서열 정보 결정 방법.

청구항 261

제 245 항에 있어서,

- (a) 주형의 뉴클레오티드의 정체를 결정함으로써 뉴클레오티드가 공지된 정체를 가지는 단계를 더 포함하되, 코드 해석 단계는:

- (i) 공지된 뉴클레오티드의 정체 및 그것의 근접한 뉴클레오티드가 공지된 정체의 뉴클레오티드에 인접한 뉴클레오티드의 반대쪽에서 결합된 프로브의 구속된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 공지된 정체의 뉴클레오티드에 인접한 주형의 뉴클레오티드에 정체를 지정하는 단계;

(ii) 그것의 근접한 뉴클레오타이드가 연속되는 뉴클레오타이드의 반대쪽에서 결합된 프로브의 구속된 부분의 가능한 서열과 정체가 일치하는 지를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계; 및

(iii) 서열이 결정될 때까지 단계 (ii)를 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 262

제 261 항에 있어서, 상기 결정 단계가 만약 이중나선에 인접한 위치에서 주형에 상보한다면 표시된 뉴클레오타이드의 통합이 허용되는 조건하에서 중합효소의 존재 하에 표시된 뉴클레오타이드와 주형-프로브 이중나선을 접촉시키는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 263

제 245 항에 있어서, 상기 코드해석 단계는 프로브 군 이름의 순서 리스트로부터 최소한 하나의 후보 서열을 생성하는 단계와, 주형의 뉴클레오타이드 서열로서 후보 서열을 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 264

제 263 항에 있어서, 상기 생성 단계는 최소한 4개의 후보 서열을 생성하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 265

제 263 항에 있어서, 상기 생성 단계는,

(i) 뉴클레오타이드 서열의 첫 번째 뉴클레오타이드에 대하여 정체를 가정하는 단계;

(ii) 첫 번째 뉴클레오타이드에 상응하는 프로브 군의 이름을 토대로 하여 인접한 뉴클레오타이드에 대하여 가능한 정체를 결정함으로써 첫 번째 뉴클레오타이드에 인접한 뉴클레오타이드에 대하여 정체를 지정하는 단계;

(iii) 그것의 정체가 가장 최근에 지정된 뉴클레오타이드에 상응하는 프로브 군의 이름을 토대로 하여 연속되는 뉴클레오타이드에 대하여 가능한 정체를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계;

(iv) 후보 서열이 생성될 때까지 단계 (iii)을 반복하는 단계; 및

(v) 원하는 수의 후보 서열이 생성될 때까지, 각 반복시 상이한 정체가 첫 번째 뉴클레오타이드에 대해 가정되는 단계 (i)-(iv)를 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 266

제 263 항에 있어서, 상기 선택 단계는 최소한 하나의 후보 서열을 하나 또는 그 이상의 공지된 서열과 비교하고, 하나 또는 그 이상의 공지된 서열과 예정된 정도의 동일성을 나타내거나 또는 가장 근접하게 동일한 후보 서열을 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 267

제 266 항에 있어서, 상기 주형이 관심 있는 유기체로부터 유도되고, 상기 비교 단계는 최소한 하나의 후보 서열을 그 유기체로부터 얻어진 서열을 함유하는 데이터베이스의 서열과 비교하는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 268

제 266 항에 있어서, 상기 비교 단계는 최소한 하나의 후보 서열을 다수의 비교 서열을 함유하는 데이터베이스의 서열과 비교하는 것을 포함하고, 다수의 비교 서열의 각각은 결정된 폴리뉴클레오타이드의 서열에 대해 가능한 대체 서열을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 269

제 263 항에 있어서, 상기 선택 단계는,

(i) 구별가능하게 표시된 코드화된 프로브 군의 두 번째 집합을 사용하여 주형으로부터 프로브 군 이름의 두 번째 순서 리스트를 얻되, 프로브 군의 두 번째 집합의 프로브 군은 프로브 군의 첫 번째 집합의 프로브 군에 상이하게 코드화되는 단계;

(ii) 프로브 군 이름의 두 번째 순서 리스트로부터 최소한 하나의 비교 서열을 생성하는 단계;

(iii) 최소한 하나의 후보 서열의 부분을 최소한 하나의 비교 서열의 부분과 비교하는 단계; 및

(iv) 단계 (c)에서 비교된 부분에 걸쳐 비교서열과 예정된 수준의 동일성을 나타내거나 가장 가깝게 동일한 후보 서열을 주형의 뉴클레오타이드 서열로서 선택하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 270

제 269 항에 있어서, 상기 비교된 부분이 단일 디뉴클레오타이드인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 271

제 269 항에 있어서, 상기 프로브 군 이름의 두 번째 순서 리스트가 단지 하나의 요소를 함유하는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 272

제 244 항에 있어서, 상기 각 프로브 군의 올리고뉴클레오타이드 프로브가 구조 $5'-(XY)(N)_kN_B^*-3'$ 또는 $3'-(XY)(N)_kN_B^*-5'$ 을 가지며, N은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, XY는 프로브의 구속된 부분이고, X와 Y는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않는 뉴클레오타이드를 나타내며, X와 Y는 최소한 2-배 축퇴성이고, 최소한 하나의 뉴클레오타이드간 연쇄는 뉴클레오타이드와 비염기성 잔기 사이의 잘라지기 쉬운 연쇄이거나 또는 뉴클레오타이드와 손상된 염기를 포함하는 잔기 사이의 연쇄이며, k는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 Y 상에 또는 $(N)_k$ 의 어떠한 뉴클레오타이드 상에 존재할 수 있어야 하는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 273

제 272 항에 있어서, 상기 최소한 하나의 뉴클레오타이드간 연쇄가 뉴클레오타이드와 비염기성 잔기 사이의 잘라지기 쉬운 연쇄인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 274

제 272 항에 있어서, 상기 검출가능한 부분이 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 275

제 274 항에 있어서, 상기 절단가능한 링커가 이황화 결합을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 276

제 272 항에 있어서, 상기 4개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군이 사용되고, 프로브의 구속된 부분에 대한 상이한 서열을 가지는 올리고뉴클레오타이드 프로브는 표 1에 설명된 24개의 코드화 중 하나에 따라 첫 번째, 두 번째, 세 번째, 및 네 번째 프로브 군으로 지정되는 것인 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 277

제 244 항에 있어서, 상기 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주

형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 278

제 244 항에 있어서, 상기 검출 단계는 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 2보다 적은 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 279

최소한 2개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보를 결정하는 방법에 있어서,

- (a) 연장가능한 말단을 가지는 이중나선 부분과 서열화될 단일-가닥의 부분을 포함하는 프로브-주형 복합체를 최소한 두 개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군과 접촉시킴으로써, 이중나선 부분에 바로 인접한 주형의 부분에 상보하는 부분을 포함하는 올리고뉴클레오타이드 프로브 사이에 혼성화가 일어나도록 하되, 프로브 군의 프로브는 트리거 잔기를 포함하는 단계;
- (b) 혼성된 올리고뉴클레오타이드 프로브를 연장가능한 말단에 결합시킴으로써 연장된 이중나선을 포함하는 프로브-주형 복합체를 생성하는 단계;
- (c) 결합된 프로브와 결합된 표지를 검출하는 단계;
- (d) 만약 이미 존재하는 것이 아니라면 연장된 이중나선 상에 연장가능한 프로브 말단을 생성하는 단계; 및
- (e) 프로브 군 이름의 순서 리스트가 얻어질 때까지 단계 (a)에서 (d)를 반복하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 280

제 279 항에 있어서, 상기 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 281

제 279 항에 있어서, 상기 검출 단계는 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 2보다 적은 비트의 정보를 얻는 것을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 282

올리고뉴클레오타이드 프로브 군의 첫 번째 집합을 사용하여 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열에 대한 정보를 결정하는 방법에 있어서,

- (a) 연속적인 사이클의 연장, 결합, 검출, 및 절단을 수행하되, 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2개의 뉴클레오타이드의 각각으로부터 동시에 평균 2 비트의 정보를 얻는 것을 포함하는 단계; 및
- (b) 단계 (a)에서 얻어진 정보를 최소한 한 비트의 추가의 정보와 조합하여 서열을 결정하는 단계를 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 283

제 282 항에 있어서, 상기 최소한 한 비트의 추가의 정보는 주형의 뉴클레오타이드의 정체, 최소한 하나의 공지 서열과 후보 서열을 비교함으로써 얻어진 정보; 및 올리고뉴클레오타이드 프로브 군의 두 번째 집합을 사용하여 방법을 반복함으로써 얻어진 정보로 이루어지는 군으로부터 선택된 항목을 포함하는 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열 정보 결정 방법.

청구항 284

최소한 2개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합으로서, 상기 각 프로브 군의 프로브들

은 구속된 부분과 구속되지 않은 부분을 포함하고, 구속된 부분의 각 위치는 최소한 2-배 축퇴성이며, 각 군의 프로브들은 트리거 잔기를 포함하는 집합.

청구항 285

제 284 항에 있어서, 상기 각 프로브가 결합 효소에 의해 연장될 수 없는 말단을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 286

제 284 항에 있어서, 상기 각 프로브가 결합 효소에 의해 연장될 수 없는 말단을 포함하고, 각 프로브가 잘라지기 쉬운 연쇄와 결합 효소에 의해 연장될 수 없는 말단 사이의 위치에서 검출가능한 부분을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 287

제 284 항에 있어서, 상기 집합은 2개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 288

제 284 항에 있어서, 상기 집합은 3개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 289

제 284 항에 있어서, 상기 집합은 4개의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 290

제 284 항에 있어서, 상기 집합은 4개 이상의 프로브 군을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 291

제 284 항에 있어서, 상기 프로브가 절단가능한 링커에 의해 부착되거나, 광표백이 가능하거나, 또는 둘 다인 검출가능한 부분을 포함하는 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합.

청구항 292

최소한 2개의 구별가능하게 표시된 올리고뉴클레오타이드 프로브 군의 집합을 포함하는 키트로서, 상기 올리고뉴클레오타이드 프로브가 트리거 잔기를 포함하는 키트.

청구항 293

제 292 항에 있어서, 상기 트리거 잔기가 비염기성 잔기, 테옥시이노신, 또는 손상된 염기를 포함하는 잔기인 것인 키트.

청구항 294

제 292 항에 있어서, 상기 프로브가 뉴클레오시드와 비염기성 잔기 사이에 있는 잘라지기 쉬운 연쇄를 포함하는 키트.

청구항 295

제 292 항에 있어서, 상기 키트가 결합 효소, 포스포로티에이트 연쇄를 절단할 수 있는 제제, 포스파타제, 중합 효소, 지지체, 완충제, 열안정성 중합효소, 뉴클레오타이드, 에멀션을 제조하기 위한 시약, 및 겔을 제조하기 위한 시약으로 이루어지는 군으로부터 선택된 최소한 하나의 항목을 더 포함하는 키트.

청구항 296

트리거 잔기를 포함하는 올리고뉴클레오타이드 프로브를 포함하는 키트로서, 상기 프로브가 잘라지기 쉬운 연쇄를 포함하거나 포함하도록 쉽게 변형될 수 있고, 상기 프로브는 검출가능한 부분으로 표시되는 것인 키트.

청구항 297

제 296 항에 있어서, 상기 검출가능한 부분이 형광 염료인 것인 키트.

청구항 298

제 296 항에 있어서, 상기 키트가 잘라지기 쉬운 연쇄를 절단할 수 있는 제제를 더 포함하는 키트.

청구항 299

제 296 항에 있어서, 상기 키트가 결합 효소를 더 포함하는 키트.

청구항 300

제 296 항에 있어서, 상기 키트가 결합 효소와 잘라지기 쉬운 연쇄를 절단할 수 있는 제제를 더 포함하는 키트.

청구항 301

제 296 항에 있어서, 상기 키트가 결합 효소, 잘라지기 쉬운 연쇄를 절단할 수 있는 제제, 포스파타제, 중합효소, 지지체, 완충제, 열안정성 중합효소, 뉴클레오타이드, 에멀션을 제조하기 위한 시약, 및 겔을 제조하기 위한 시약으로 이루어지는 군으로부터 선택된 최소한 하나의 항목을 더 포함하는 키트.

청구항 302

제 301 항에 있어서, 상기 키트가 다수의 형광 표시된 올리고뉴클레오타이드 프로브를 함유하고, 상기 프로브는 뉴클레오시드와 트리거 잔기 사이에 잘라지기 쉬운 연쇄를 포함함으로써 프로브의 상이한 말단 뉴클레오타이드에 상응하는 프로브들이 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반하는 것인 키트.

청구항 303

형태 $5'-O-P-O-X-O-P-S-(N)_kN_B^*-3'$ 의 올리고뉴클레오타이드로서, 상기 N은 어떠한 뉴클레오타이드 또는 비염기성 잔기를 나타내고, 단 최소한 하나의 N은 트리거 잔기여야 하며, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내고, *는 검출가능한 부분을 나타내며, X는 뉴클레오타이드를 나타내고, k는 1과 100 사이에 있으며, 포괄적이고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오타이드 상에 존재할 수 있어야 하는 것인 올리고뉴클레오타이드 프로브.

청구항 304

제 303 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오타이드를 포함하는 올리고뉴클레오타이드 프로브.

청구항 305

제 303 항에서 설명된 것과 같은 올리고뉴클레오타이드 프로브의 세트로서, 상기 세트가 다수의 형광 표시된 올리고뉴클레오타이드 프로브를 함유함으로써 프로브의 상이한 뉴클레오타이드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오타이드 프로브 세트.

청구항 306

형태 $5'-N_B^*(N)_k-S-P-O-X-3'$ 의 올리고뉴클레오타이드 프로브로서, 상기 N은 어떠한 뉴클레오타이드 또는 비염기성 잔기를 나타내고, 단 최소한 하나의 N은 트리거 잔기여야 하고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오타이드를 나타내며, k는 1과 100 사이에 있고,

포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 $(N)_k$ 의 어떠한 뉴클레오티드 상에 존재할 수 있어야 하는 것인 올리고뉴클레오티드 프로브.

청구항 307

제 306 항에 있어서, 상기 프로브가 최소한 하나의 축퇴성-감소 뉴클레오티드를 포함하는 올리고뉴클레오티드 프로브.

청구항 308

제 306 항에서 설명된 것과 같은 올리고뉴클레오티드 프로브의 세트로서, 상기 세트가 다수의 형광 표지된 올리고뉴클레오티드 프로브를 함유함으로써 프로브의 상이한 뉴클레오티드 X에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반할 수 있는 것인 올리고뉴클레오티드 프로브 세트.

청구항 309

3'-XNNNNRINI-5', 3'-XNNNNRIII-5', 3'-XNNNNRNII-5', 3'-XNNNNIRII-5', 3'-XNNNNRNI-5', 3'-XNNNNRII-5', 3'-XNNNNRII-5' 및 3'-XNNNNIRI-5'으로 이루어지는 군으로부터 선택되는 형태의 올리고뉴클레오티드 프로브로서, 상기 X와 N은 어떠한 뉴클레오티드를 나타내고, "R"은 트리거 잔기를 나타내며, 트리거 잔기와 올리고뉴클레오티드의 5' 말단 사이에 있는 최소한 하나의 잔기는 X의 정체에 상응하는 표지를 포함하는 올리고뉴클레오티드 프로브.

청구항 310

제 309 항에 있어서, 상기 R이 테옥시리보스 잔기를 나타내는 것인 올리고뉴클레오티드 프로브.

청구항 311

두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법에 있어서,

(a) 반-고체 지지체에 또는 그 위에 고정된 첫 번째 폴리뉴클레오티드를 제공하는 단계;

(b) 첫 번째 폴리뉴클레오티드를 두 번째 폴리뉴클레오티드 및 결합 효소와 접촉시키는 단계; 및

(c) 결합에 적당한 조건하에서 결합 효소의 존재 하에 첫 번째와 두 번째 폴리뉴클레오티드를 유지하되, 최소한 하나의 폴리뉴클레오티드는 트리거 잔기를 포함하는 단계를 포함하는 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 312

제 311 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 공유 또는 비공유 연쇄를 통하여 반-고체 지지체에 직접 부착되는 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 313

제 311 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 반-고체 지지체에 또는 그 위에 고정된 지지체에 부착되는 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 314

제 313 항에 있어서, 상기 반-고체 지지체가 겔이고 지지체가 마이크로입자인 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 315

제 313 항에 있어서, 상기 반-고체 지지체가 겔이고 지지체가 자성 마이크로입자인 것인 두 번째 폴리뉴클레오티드에 첫 번째 폴리뉴클레오티드를 결합하는 방법.

청구항 316

폴리뉴클레오티드를 절단하는 방법에 있어서,

- (a) 반-고체 지지체에 또는 그 위에 고정된 폴리뉴클레오티드를 제공하되, 폴리뉴클레오티드는 트리거 잔기를 포함하는 단계;
- (b) 폴리뉴클레오티드를 절단제와 접촉시키는 단계; 및
- (c) 절단에 적당한 조건하에서 절단제의 존재 하에 폴리뉴클레오티드를 유지하는 단계를 포함하는 폴리뉴클레오티드 절단 방법.

청구항 317

제 316 항에 있어서, 상기 트리거 잔기가 비염기성 잔기, 데옥시이노신, 또는 손상된 염기를 포함하는 잔기인 것인 폴리뉴클레오티드 절단 방법.

청구항 318

제 316 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 공유 또는 비공유 연쇄를 통하여 반-고체 지지체에 직접 부착되는 것인 폴리뉴클레오티드 절단 방법.

청구항 319

제 316 항에 있어서, 상기 첫 번째 폴리뉴클레오티드가 반-고체 지지체에 또는 그 위에 고정된 지지체에 부착되는 것인 폴리뉴클레오티드 절단 방법.

청구항 320

제 319 항에 있어서, 상기 반-고체 지지체가 겔이고 상기 지지체가 마이크로입자인 것인 폴리뉴클레오티드 절단 방법.

청구항 321

제 320 항에 있어서, 상기 반-고체 지지체가 겔이고 상기 지지체가 자성 마이크로입자인 것인 폴리뉴클레오티드 절단 방법.

청구항 322

마이크로입자의 집단을 제조하기 위한 성분들의 집합으로서, 상기 집합은

- (a) 개별적인 마이크로입자들이 최소한 첫 번째와 두 번째의 그것에 부착된 프라이머의 집단을 가지고, 첫 번째 집단의 프라이머가 두 번째 집단의 프라이머의 서열과 상이한 서열을 가지는 것인 마이크로입자의 집단과;
- (b) 핵산 단편들의 라이브러리를 포함하며, 각 핵산 단편은 관심 있는 첫 번째 및 두 번째 핵산 절편을 함유하고, 첫 번째 및 두 번째 프라이머는 관심 있는 첫 번째 및 두 번째 핵산 절편 외부에 위치한 보편적 서열에 상응하는 것인 성분들의 집합.

청구항 323

제 322 항에 있어서, 상기 첫 번째 및 두 번째 핵산 절편이 쌍을 이룬 태그의 5' 및 3' 태그인 것인 성분들의 집합.

청구항 324

제 322 항에 있어서, 상기 핵산 단편이 증폭 프라이머에 대한 하나 또는 그 이상의 프라이머 결합 부위를 포함하는 내부 어댑터를 함유함으로써 각 핵산 절편이 PCR을 사용하여 증폭될 수 있는 것인 성분들의 집합.

청구항 325

제 324 항에 있어서, 상기 집합은 내부 어댑터의 프라이머 결합 부위에 상보하는 프라이머를 더 포함하는 성분들의 집합.

청구항 326

실질적으로 동일한 핵산 서열의 첫 번째 집단이 부착되어 있고 실질적으로 동일한 핵산 서열의 두 번째 집단이

부착된 마이크로입자로서, 상기 핵산의 첫 번째 집단은 관심 있는 첫 번째 핵산 절편을 포함하고 핵산의 두 번째 집단은 관심 있는 두 번째 핵산 절편을 포함하는 마이크로입자.

청구항 327

제 326 항에 있어서, 상기 관심 있는 첫 번째 및 두 번째 핵산 절편이 쌍을 이룬 태그의 5' 태그 및 3' 태그인 것인 마이크로입자.

청구항 328

제 326 항에 있어서, 상기 관심 있는 첫 번째 및 두 번째 핵산 절편이 자연 발생의 연속적인 핵산에서 서로에게서 예정된 거리 내에서 발견되는 5' 태그와 3' 태그인 것인 마이크로입자.

청구항 329

제 327 항 또는 제 328 항에 있어서, 상기 관심 있는 첫 번째 및 두 번째 핵산 절편이 단일한 더 큰 핵산 절편으로부터 증폭된 것인 마이크로입자.

청구항 330

제 329 항에 있어서, 상기 증폭이 PCR 에멀션의 개별적인 구획에서 수행되는 것인 마이크로입자.

청구항 331

제 326 항에서 설명된 것과 같은 마이크로입자의 집단으로서, 상기 개별적인 마이크로입자는 실질적으로 동일한 핵산 서열의 첫 번째 및 두 번째 집단이 부착되어 있고, 그것들은 다른 개별적인 마이크로입자에 부착된 실질적으로 동일한 핵산 서열의 첫 번째 및 두 번째 집단과는 최소한 부분적으로 상이한 것인 마이크로입자의 집단.

청구항 332

제 331 항에 있어서, 상기 개별적인 마이크로입자에 부착된 핵산 서열의 첫 번째 집단이 관심 있는 첫 번째 핵산 절편을 포함하고, 개별적인 마이크로입자에 부착된 핵산 서열의 두 번째 집단은 관심 있는 두 번째 핵산 절편을 포함하는 마이크로입자의 집단.

청구항 333

제 332 항에 있어서, 상기 개별적인 마이크로입자에 부착된 핵산 서열의 첫 번째 및 두 번째 집단에서 관심 있는 첫 번째 및 두 번째 핵산 절편이 쌍을 이룬 태그의 5' 태그 및 3' 태그이거나 그것들을 포함하는 마이크로입자의 집단.

청구항 334

제 332 항에 있어서, 상기 개별적인 마이크로입자에 부착된 핵산 서열의 첫 번째 및 두 번째 집단에서 관심 있는 첫 번째 및 두 번째 핵산 절편이 단일한 더 큰 핵산 단편으로부터 증폭되는 것인 마이크로입자의 집단.

청구항 335

제 334 항에 있어서, 상기 증폭이 PCR 에멀션의 개별적인 구획에서 수행되는 것인 마이크로입자의 집단.

청구항 336

제 331 항에 있어서, 상기 첫 번째 및 두 번째의 핵산 서열이 PCR 에멀션의 개별적인 구획에서 증폭을 수행함으로써 개별적인 마이크로입자에 부착되고, 상기 개별적인 구획의 최소한 일부가 단일한 마이크로입자 및 첫 번째 및 두 번째 핵산 절편을 함유하는 단일 핵산 단편을 함유하는 것인 마이크로입자의 집단.

청구항 337

제 331 항 내지 제 336 항 중 어느 한 항의 마이크로입자의 집단을 포함하는 배열.

청구항 338

제 337 항에 있어서, 상기 마이크로입자가 반-고체 지지체에 또는 그 위에 고정되는 것인 배열.

청구항 339

실질적으로 동일한 핵산 서열의 첫 번째 집단이 부착되어 있고 실질적으로 동일한 핵산 서열의 두 번째 집단이 부착된 마이크로입자의 제조 방법에 있어서,

- (a) 첫 번째 프라이머 집단이 부착되어 있고 두 번째 프라이머 집단이 부착된 마이크로입자를 제공하는 단계;
- (b) 첫 번째 및 두 번째 핵산 절편 양 옆에 마이크로입자에 부착된 첫 번째 및 두 번째 프라이머의 서열에 상응하는 첫 번째 및 두 번째 프라이머 결합 영역이 있고, 두 영역은 최소한 하나의 추가의 프라이머 결합 영역에 의해 분리되는 첫 번째 및 두 번째 핵산 절편을 포함하는 핵산 단편을 제공하는 단계; 및
- (c) 마이크로입자와 핵산 단편을 적당한 증폭 시약 및 프라이머의 존재하에 증폭이 일어날 수 있는 조건하에서 배양함으로써, 첫 번째 및 두 번째 핵산 서열이 증폭되고 그 증폭의 결과로서 마이크로입자에 부착되는, 첫 번째 및 두 번째 핵산 서열을 증폭하는 단계를 포함하는 마이크로입자 제조 방법.

청구항 340

제 339 항에 있어서, 상기 첫 번째 및 두 번째 프라이머 결합 영역이 각각 증폭 프라이머의 결합을 위한 영역과 서열화 프라이머의 결합을 위한 영역을 포함하는 마이크로입자 제조 방법.

청구항 341

제 339 항에 있어서, 상기 최소한 하나의 추가의 프라이머 결합 영역이 2개의 증폭 프라이머의 결합을 위한 영역을 포함함으로써, 각 핵산 서열의 양 옆에 한 쌍의 증폭 프라이머에 대한 프라이머 결합 부위가 있게 되고, 그로써 2개의 핵산 서열이 각각 PCR에 의해 증폭되는 것인 마이크로입자 제조 방법.

청구항 342

제 339 항에 있어서, 상기 증폭이 PCR 에멀션의 구획에서 수행되는 것인 마이크로입자 제조 방법.

청구항 343

제 339 항에 있어서, 상기 첫 번째 및 두 번째 핵산 서열이 각각 태그를 포함하고, 그 태그는 쌍을 이룬 태그의 5' 태그와 3' 태그인 것인 마이크로입자 제조 방법.

청구항 344

실질적으로 동일한 핵산 서열의 상이한 집단이 부착된 마이크로입자의 집단의 제조 방법에 있어서, 상기 방법은 제 339 항의 방법을 수행하는 것을 포함하고, 상기 증폭이 PCR 에멀션의 다수의 구획에서 수행되며, 상기 구획의 최소한 일부는 단일 마이크로입자와 첫 번째 및 두 번째 핵산 절편을 포함하는 단일 핵산 단편을 함유하고, 상기 개별적인 핵산 단편의 첫 번째 및 두 번째 핵산 서열은 첫 번째 및 두 번째와 상이한 것인 마이크로입자 제조 방법.

청구항 345

제 344 항의 마이크로입자의 집단을 반-고체 지지체에 또는 그 위에 고정하는 것을 포함하는 마이크로입자의 배열의 제조 방법.

청구항 346

핵산 서열화를 수행하는 방법에 있어서,

- (a) 마이크로입자에 부착된 핵산 분자의 첫 번째 집단으로부터 서열 정보를 얻는 단계와;
- (b) 동일한 마이크로입자에 부착된 핵산 분자의 두 번째 집단으로부터 서열 정보를 얻되, 첫 번째 및 두 번째 핵산 분자의 서열은 최소한 부분적으로 상이한 것인 단계를 포함하는 핵산 서열화 수행 방법.

청구항 347

제 346 항에 있어서, 단계 (a)와 (b)의 서열 정보가 차례로 얻어지는 것인 핵산 서열화 수행 방법.

청구항 348

제 346 항에 있어서, 상기 첫 번째 핵산 분자 집단이 쌍을 이룬 태그의 5' 태그를 포함하고, 두 번째 핵산 분자 집단이 3' 태그를 포함하는 핵산 서열화 수행 방법.

청구항 349

마이크로입자의 집단에 대해 제 346 항의 방법을 수행하는 것을 포함하는 서열화 방법에 있어서, 상기 개별적인 마이크로입자는 그것에 부착된 첫 번째 및 두 번째 핵산 분자를 가지고, 그것들은 다른 개별적인 마이크로입자에 부착된 첫 번째 및 두 번째 핵산 분자의 서열과 최소한 부분적으로 서열이 상이한 것인 서열화 방법.

청구항 350

제 349 항에 있어서, 상기 개별적인 마이크로입자 상에서 첫 번째 핵산 분자 집단의 서열화가 동시에 수행되며, 상기 개별적인 마이크로입자 상에서 두 번째 핵산 분자 집단의 서열화가 동시에 수행되는 것인 서열화 방법.

청구항 351

물에 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15mM의 중황산 나트륨 (NaHSO₄)을 포함하는 조성물.

청구항 352

제 351 항에 있어서, 상기 조성물의 pH가 2.0 과 3.0 사이에 있는 것인 조성물.

청구항 353

제 351 항에 있어서, 상기 조성물이 물에 약 2%의 SDS, 약 200mM의 NaCl, 및 약 10 mM의 중황산 나트륨 (NaHSO₄)을 포함하는 조성물.

청구항 354

제 353 항에 있어서, 상기 조성물의 pH가 2.0 과 3.0 사이에 있는 조성물.

명세서

기술 분야

<1> 본 발명은 비드-기반 서열화를 위한 시약, 방법, 및 라이브러리에 관한 것이다.

배경 기술

<2> 핵산 서열화 기술은 기초 연구 분야로부터 임상 진단에 이르기까지 광범위한 분야에서 매우 중요하다. 그러한 기술로부터 활용할 수 있는 결과들은 그 특이성 정도가 다양한 정보를 포함한다. 예를 들어 유용한 정보는 특정한 폴리뉴클레오티드가 참조 폴리뉴클레오티드와 서열이 다른 지를 결정하고, 특정한 폴리뉴클레오티드 서열이 샘플에 존재하는 지를 확인하고, 폴리뉴클레오티드 내의 하나 또는 그 이상의 뉴클레오티드의 정체와 같은 부분적인 서열 정보를 결정하고, 폴리뉴클레오티드 내의 뉴클레오티드들의 정체와 순서를 결정하는 것 등으로 구성될 수 있다.

<3> DNA 가닥은 전형적으로 4가지 유형의 하위단위, 즉 아데닌(A), 시토신(C), 구아닌(G) 및 티미딘(T) 염기를 함유하는 데옥시리보뉴클레오티드로 구성되는 중합체이다. 이들 하위단위는 서로 간에 한 데옥시리보스기의 5' 탄소를 다음 기의 3' 탄소에 연결하는 공유 인산이에스테르 결합에 의해 부착된다. 대부분의 자연 발생 DNA는 두 개의 이러한 가닥으로 구성되는데, 두 가닥은 반대방향으로 평행하게 배열되며 상보하는 염기, 즉 A와 T 및 G와 C 사이에 형성된 수소 결합에 의해 결합되어 있다.

<4> DNA 서열화는 먼저 사슬 종결 또는 디데옥시뉴클레오티드 방법의 개발 (Sanger, et al., *Proc. Natl. Acad. Sci.* 74:5463- 5467, 1977) 및 화학적 분해 방법 (Maxam & Gilbert, *Proc. Natl. Acad. Sci.* 74:560-564,

1977)의 개발로 대규모로 수행 가능하게 되었는데, 전자의 방법이 가장 광범위하게 사용되어 이를 개선하여 자동화되었다. 특히 형광 표지된 사슬 터미네이터의 사용은 자동 DNA 서열화기(sequencer)의 개발에 중요한 핵심이었다. 상기 언급된 두 가지 접근법에 공통적인 것은 상이한 크기의 표지된 DNA 단편들의 하나 또는 그 이상의 집합의 생성과, 그런 다음 단편의 3' 말단에 있는 뉴클레오타이드의 정체(사슬 종결 방법에서) 또는 단편으로부터 가장 최근에 제거된 뉴클레오타이드의 정체(화학적 분해 방법의 경우에)를 결정하기 위해 길이에 따라 분리되어야 하는 것이었다.

- <5> 비록 현재 활용되는 서열화 기술들이 많은 완전한 게놈의 서열화와 같은 주요한 획기적인 사건을 이루는 것을 가능하게 하였지만, 이들 기술은 많은 단점을 가지며, 많은 영역에서 개선되어야 할 필요성이 여전히 남아있다. 표지된 DNA 단편들의 분리는 전형적으로 폴리아크릴아미드 겔 전기영동을 사용하여 이루어졌다. 그러나 이 단계는 많은 맥락에서 서열화의 속도 및 정확성을 제한하는 주요한 병목인 것으로 증명되었다. 모세관 전기영동(CAE)이 인간 게놈 프로젝트를 완성하게 한 돌파구인 것으로 증명되었지만 (Venter, et al., *Science*, 291:1304-1351, 2001; Lander, et al., *Nature*, 409:860-921, 2001), 상당한 단점들이 남아있다. 예를 들어 CAE는 여전히 시간 소모적인 분리 단계를 필요로 하고 또한, 정확하지 않을 수 있는, 크기에 따른 분류단계를 포함한다.
- <6> 사슬 종결 방법에 대한 다양한 대체방법들이 제안되어 왔다. "합성에 의한 서열화"로서 언급되기도 하는 한 접근법에서, 올리고뉴클레오타이드 프라이머는 먼저 표적 주형에 혼성화된다. 그런 다음 프라이머는 연속적인 사이클의 상이하게 표지된 뉴클레오타이드의 중합효소-촉매적(catalyzed) 첨가에 의해 연장되며, 그 뉴클레오타이드가 성장하는 가닥 안으로 통합되는 것이 검출된다. 표지의 정체는 주형의 상보하는 뉴클레오타이드를 확인하는 데 사용된다. 이와 달리, 각각의 뉴클레오타이드를 사용하여 동시에 다수의 반응이 수행될 수 있고, 특정한 뉴클레오타이드를 사용하는 반응에 표지된 뉴클레오타이드가 통합되는 것은 주형의 상보하는 뉴클레오타이드를 확인할 수 있게 해준다 (Melamede, 미국 특허 제 4,863,849호; Cheeseman, 미국 특허 제 5,302,509호, Tsien et al., 국제 출원 제 WO91/06678호; Rosenthal et al., 국제 출원 제 WO93/21340호; Canard et al., *Gene*, 148: 1-6 (1994); Metzker et al., *Nucleic Acids Research*, 22: 4259-4267 (1994)).
- <7> 어떠한 상당한 길이의 폴리뉴클레오타이드를 효과적으로 서열화하기 위해서는 중합효소는 각 사이클에서 정확하게 한 뉴클레오타이드에 결합하는 것이 바람직하다. 그러므로 일반적으로 사슬 터미네이터로 작용하는, 즉 그것들의 결합이 중합효소에 의한 추가의 연장을 방지하는 뉴클레오타이드를 사용하는 것이 필요하다. 그러면 결합한 뉴클레오타이드는 효소적으로 또는 화학적으로 반드시 변형되어, 중합효소가 다른 뉴클레오타이드와 결합하도록 한다. 사슬 터미네이터로서 작용할 수 있지만 그것들이 결합한 후에는 후속되는 단계에서 연장될 수 있는 다양한 뉴클레오타이드 유사체들이 제시되었다. 그러한 "가역성 터미네이터"는, 예를 들면, 미국 특허 제 5,302,509호; 제 6,255,475호; 제 6,309,836호; 제 6,613,513호에서 설명되었다. 그러나 고효율로 중합효소에 의해 결합될 수 있는 가역성 터미네이터를 확인하는 것은 어려운 것으로 증명되었는데, 이는 아마도 주어진 뉴클레오타이드의 작은 크기, 뉴클레오타이드가 터미네이터로서 작용할 수 있는 능력에 영향을 주는 변형 및 성장하는 폴리뉴클레오타이드 가닥으로의 통합에도 영향을 준다는 사실 때문이다.
- <8> 다른 접근법은 DNA 중합반응 동안 방출되는 피로포스페이트 (PPi)의 검출을 토대로 하는 피로서열화를 포함한다 (미국 특허 제 6,210,891호 및 제 6,258,568호 참조). 전기영동에 의한 분리의 요구는 피할 수 있었지만 피로서열화 역시 여전히 그것의 광범위한 적용성을 제한하는 수많은 결점을 가진다 (Franca, et al., *Quarterly Reviews of Biophysics*, 35(2): 169-200, 2002). 혼성화에 의한 서열화 또한 대체방법으로 제시되었지만 (미국 특허 제 5,202,231호; 제 WO 99/60170; 제 WO 00/56937; Drmanac, et al., *Advances in Biochemical Engineering/Biotechnology*, 11:16-101, 2002), 매우 유사한 서열들을 구별하는 데 있어서 예러의 잠재성을 포함하는 많은 결점을 지니고 있다. 한 가닥의 모든 염기를 표지화한 후 샘플 가닥 중에서 차례로 절단된 3' 말단 뉴클레오타이드를 검출하는 것을 포함하는, 엑소뉴클레아제에 의한 단일-분자 서열화는 이론적으로는 긴 DNA 분자의 서열을 신속하게 결정하기 위한 매우 강력한 방법이다 (Stephan, et al., *J. Biotechnol.*, 86:255-267, 2001). 그러나 이 가능성을 현실화하기 전에 극복해야 할 다양한 기술적 장애가 남아있다 (Stephan, et al., 2001).
- <9> 특정 서열 변화를 기초로 한 진단 목적의 시험은 이미 다양한 상이한 질병에 대해 사용되고 있다. 인간 게놈의 서열화는, 예방 치료를 포함하여 치료법이 환자의 특정한 유전적 구성에 맞춰지거나 특정한 대립 형질 또는 돌연변이의 확인을 기초로 선택될 개인화된 의료 시대의 도래를 예고하는 것으로 널리 여겨진다. 이런 경우 HIV와 같은 병원체의 서열 변종의 신속하고 정확한 결정에 대한 요구가 증가하고 있다. 그러므로 정확하고 신속한 서열 결정에 대한 수요는 가까운 미래에 더욱 확대될 것이 분명하다. 그러므로 모든 유형의 서열 결정법에 대한

개선된 방법이 필요하다.

<10> 발명의 개요

<11> 본 발명은 단편의 분리를 수행할 필요성을 피하면서 또한 어떤 실시예에서는 효소 중합효소를 사용할 필요를 피하는 새롭고 개선된 서열화 방법을 제공한다. 배경기술에서 논의된 방법의 대체방법은 미국 특허 5,740,341호와 6,306,597호 (Macevicz)에서 설명된다. 그 방법들은 단일-가닥의 주형을 따라 이중나선이 연장되는 반복되는 사이클을 토대로 한다. 이들 방법의 바람직한 실시예에서, 뉴클레오티드는 각 사이클마다 확인된다. 본 발명은 이들 방법에 개선된 점들을 제공한다. 개선점으로 인해 이들 방법을 효과적으로 실행할 수 있고, 이는 많은 양을 처리해야할 서열화에 특히 적합하다. 또한 본 발명은 단일-가닥의 주형을 따라 이중나선이 연장되는 반복된 사이클을 포함하지만 각 사이클 동안 어떠한 개별적인 뉴클레오티드의 확인을 포함하지 않는 서열 결정에 대한 방법을 제공한다.

<12> 한 측면으로, 본 발명은 단일-가닥의 주형을 따라 이중나선이 연장되고, 표지된 연장 프로브를 결찰(ligation)하고, 표지를 검출하는 연속적인 사이클을 토대로 하는 개선된 서열화 방법을 제공한다. 일반적으로 연장은 개시 올리고뉴클레오티드와 주형에 의해 형성된 이중나선으로부터 시작된다. 개시 올리고뉴클레오티드는 연장된 이중나선을 형성하기 위하여 그것의 말단에 올리고뉴클레오티드 프로브를 결찰해 연장되고, 그것은 다음에 연속적인 결찰 사이클에 의해 반복적으로 연장된다. 각 사이클 동안에 주형의 하나 또는 그 이상의 뉴클레오티드의 정체는 연속적으로 결찰된 올리고뉴클레오티드 프로브 상의 또는 그것과 결찰된 표지를 확인함으로써 결정된다. 새롭게 첨가된 프로브의 표지는 또한 결찰 전 및/또는 결찰 후에 검출될 수 있다. 일반적으로 결찰 후에 표지가 검출되는 것이 바람직하다.

<13> 바람직한 실시예에서 프로브는 말단 위치에 (이중나선의 성장하는 핵산 가닥에 결찰된 뉴클레오티드로부터 프로브의 반대쪽 말단에) 비-연장성 부분을 가지고 있어서 단일 사이클에서 연장된 이중나선에 단지 한 번의 연장만이 일어난다. "비-연장성"이란 변형 없이는 리가아제(ligase)에 대한 기질로서 작용하지 않는 부분을 의미한다. 예를 들어 그 부분은 5' 인산기 또는 3' 수산기가 없는 뉴클레오티드 잔기일 수 있다. 그 부분은 결찰(ligation)을 방해하는 차단기가 부착된 뉴클레오티드일 수 있다. 본 발명의 바람직한 실시예에서 비-연장성 부분은 결찰 후에 제거되어 연장할 수 있는 말단이 생성되고, 이로써 이중나선은 후속 사이클에서 더욱 연장될 수 있다.

<14> 비-연장성 부분의 제거를 가능하게 하기 위해, 본 발명의 어떤 실시예에서는 프로브는 실질적으로 인산이에스테르 결함을 절단하지 않을 조건 하에서 절단될 수 있는 뉴클레오시드간 연쇄(linkage)를 최소한 하나 함유한다. 그런 연쇄는 본원에서 "잘라지기 쉬운 뉴클레오시드간 연쇄" 또는 "잘라지기 쉬운 연쇄"로 언급된다. 잘라지기 쉬운 뉴클레오시드간 연쇄의 절단은 비-연장성 부분을 제거하고, 연장할 수 있는 프로브 말단을 생성하거나 연장할 수 있는 프로브 말단을 형성하기 위해 변형될 수 있는 말단 잔기를 남긴다. 잘라지기 쉬운 뉴클레오시드간 연쇄는 프로브 내의 어떠한 두 개의 뉴클레오시드가 사에 위치할 수 있다. 바람직한 것은 잘라지기 쉬운 연쇄가 새롭게 형성된 결함으로부터 최소한 7개의 뉴클레오티드만큼 떨어져서 (즉 그만큼 멀리) 위치하는 것이다. 연장할 수 있는 말단에 결찰된 말단 뉴클레오티드와 잘라지기 쉬운 연쇄 사이의 연장 프로브에 있는 뉴클레오티드는 주형에 완벽하게 혼성화할 필요는 없다. 이들 뉴클레오티드는 "스페이스(spacer)"로서 작용할 수 있고, 간격 내에 있는 각 뉴클레오티드에 대한 사이클을 수행하지 않고 주형을 따라 간격을 두고 위치한 뉴클레오티드를 확인할 수 있게 한다.

<15> 잘라지기 쉬운 뉴클레오시드간 연쇄와 표지는 바람직하게는 잘라지기 쉬운 뉴클레오시드간 연쇄의 절단으로 연장 프로브가 표지된 부분과 성장하는 핵산 가닥의 일부를 보유하는 부분으로 분리되어, 표지된 부분이 확산될 수 있도록 (예컨대 온도를 상승시킬 때) 위치한다. 예를 들어 표지는 결찰되는 뉴클레오티드로부터 반대쪽 말단에 있는 연장 프로브의 말단 뉴클레오티드에 부착될 수 있다. 이와 달리, 표지는 많은 접근법 중 어떠한 것을 사용하여 제거될 수 있다.

<16> 본 발명자들은 인산이에스테르 결함의 가교 산소 원자들 중 하나가 황 원자에 의해 대체되는, 포스포로티올레이트(phosphorothiolate) 연쇄가 특히 유용한 잘라지기 쉬운 뉴클레오시드간 연쇄라는 것을 발견하였다. 포스포로티올레이트 연쇄에서 황 원자는 한 뉴클레오시드의 3' 탄소 또는 인접한 뉴클레오시드의 5' 탄소에 부착될 수 있다.

<17> 상기에서 설명된 방법들의 어떤 실시예에서 다수의 서열화 반응이 수행된다. 반응은 첫 번째 결찰이 일어나는 말단이 주형에 관련하여 상이한 위치에 위치하게 되도록 주형의 상이한 서열에 혼성화하는 개시 올리고뉴클레오

티드를 사용한다. 예를 들어 첫 번째 결찰이 일어나는 위치는 한 뉴클레오티드가 증가할 때마다 상호 간에 이동하거나 "상(phase) 밖에" 있을 수 있다. 그러므로 동일한 길이의 올리고뉴클레오티드 프로브를 사용한 각 연장 사이클 후에 동일한 상대적 상은 상이한 주형 상의 개시 올리고뉴클레오티드의 말단을 사이에 존재한다. 상기 반응은 동시에, 각각이 동일한 주형의 복제물을 함유하는 별도의 구획에서, 또는 연속적으로, 즉 첫 번째 개시 올리고뉴클레오티드를 사용하여 서열 정보를 얻은 후 주형으로부터 연장된 이중나선을 제거한 후 주형의 상이한 서열에 혼성화하는 개시 올리고뉴클레오티드를 사용하여 추가 반응(들)을 수행함으로써 수행될 수 있다.

<18> 다른 측면으로 본 발명은 다양한 핵산 조작을 사용하는 용액을 제공한다. 한 실시예에서 본 발명은 본질적으로 물에 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15mM의 중황산 나트륨 (NaHSO₄)을 함유하는 또는 그것들로 구성되는 용액을 제공한다. 용액은 물 중의 약 2%의 SDS, 약 200mM의 NaCl, 및 약 10mM의 중황산 나트륨 (NaHSO₄)을 함유하거나 본질적으로 그것들로 구성된다. 예를 들어 한 실시예에서 용액은 물 중의 2% SDS, 200mM의 NaCl 및 10mM의 중황산 나트륨 (NaHSO₄)을 함유한다. 다른 실시예에서 용액은 본질적으로 물 중의 2% SDS, 200mM의 NaCl 및 10mM의 중황산 나트륨 (NaHSO₄)으로 구성된다. 어떤 실시예에서 용액은 2.0과 3.0 사이, 예컨대 2.5의 pH를 갖는다. 용액은 이중-가닥 핵산, 예컨대 이중-가닥 DNA를 개별적인 가닥으로 분리하는데, 즉 이중-가닥 핵산을 변성 (용융)시키는 데 유용하다. 특정 실시예에서 두 가닥은 모두 DNA이다. 다른 실시예에서 두 가닥은 모두 RNA이다. 다른 실시예에서 한 가닥은 DNA이고 다른 가닥은 RNA이다. 다른 실시예에서 하나 또는 두 가닥 모두 RNA와 DNA를 모두 함유한다. 다른 실시예에서 하나 또는 두 가닥은 A, G, C, 또는 T 이외의 최소한 하나의 뉴클레오티드를 함유한다. 어떤 실시예에서 하나 또는 두 가닥은 비-자연 발생 뉴클레오티드를 함유한다. 또 다른 실시예에서 하나 또는 그 이상의 잔기는 트리거(trigger) 잔기, 예컨대, 비염기성 잔기 또는 손상된 염기이다. 어떤 실시예에서 하나 또는 그 이상의 잔기는 보편적인 염기를 함유한다. 어떤 실시예에서 하나 또는 두 가닥은 잘라지기 쉬운 연쇄를 함유한다.

<19> 이중-가닥 핵산은 전체적으로 또는 부분적으로 이중-가닥일 수 있다. 그것들은 용액 중에서 유리되어 있거나 하나 또는 두 가닥 모두 고체 또는 반고체 지지체 또는 기관과 물리적으로 결합될 수 있다 (예컨대 공유적으로 또는 비 공유적으로 부착될 수 있다). 특히 주의할 것은 이들 용액 중에서 배양(incubatio)된 이중-가닥 핵산은 열 또는 겔 적층 탈피를 유발할 수 있는 (예컨대 핵산이 폴리아크릴아미드 겔과 같은 반고체 지지체에 위치하거나 부착될 때) 또는 스트렙타비딘 (streptavidin;SA)-비오틴 (biotin) 결합과 같은 비공유 결합을 파괴할 수 있는 (예컨대 핵산이 SA-비오틴 결합을 통하여 지지체 또는 기관에 부착될 때) 엄격한 변성제(denaturant)의 부재하에 단일 가닥으로 효과적으로 분리된다는 것이다. 한 실시예에서 용액은 핵산 중 하나가 SA-비오틴 결합을 통하여 비드에 부착된 이중-가닥 핵산을 분리하기 위해 사용된다.

<20> 본 발명은 또한 이중 가닥 핵산을 상기에서 언급된 어떠한 용액, 예컨대 약 1.0 내지 3.0%의 SDS, 약 100 내지 300mM의 NaCl, 및 5 내지 15mM의 중황산 나트륨 (NaHSO₄)을 함유하는 수용액, 예컨대 1.0 내지 3.0%의 SDS, 100 내지 300mM의 NaCl, 및 5 내지 15mM의 중황산 나트륨 (NaHSO₄)을 함유하는 용액과 접촉시키는 단계를 포함하는, 이중-가닥 핵산의 가닥을 분리하는 방법을 제공한다. 한 실시예에서 용액은 약 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO₄), 예컨대 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO₄)을 함유한다. 한 실시예에서 용액은 본질적으로 물에 2%의 SDS, 200mM의 NaCl, 및 10mM의 중황산 나트륨 (NaHSO₄)을 함유한다. 어떤 실시예에서 용액은 2.0과 3.0 사이, 예컨대 2.5의 pH를 갖는다. 어떤 실시예에서 이중-가닥 핵산은 용액 중에서 배양된다. 다른 실시예에서 이중-가닥 핵산 (바람직하게는 지지체 또는 기관에 부착되어 있음)은 용액으로 세척된다. 어떤 실시예에서 이중-가닥 핵산은 최소한 10%의 이중-가닥 핵산 분자가 단일 가닥으로 분리되기에 충분한 시간 동안 용액과 접촉된다. 어떤 실시예에서 이중-가닥 핵산은 용액과 최소한 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99% 또는 그 이상의 이중-가닥 핵산이 단일 가닥으로 분리되기에 충분한 시간 동안 접촉된다. 예시적인 실시예에서 이중-가닥 핵산은 15초 내지 3시간 동안 용액과 접촉된다. 다른 실시예에서 이중-가닥 핵산은 용액과 1분 내지 1시간 동안 접촉된다. 어떤 실시예에서 이중-가닥 핵산은 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 또는 60분 동안 용액과 접촉된다. 방법은 배양 주기 후에 용액을 제거하거나 또는 핵산의 전부 또는 일부를 용액으로부터 제거하는 추가 단계를 포함할 수 있다.

<21> 용액은 본원에서 설명된 많은 서열화 방법의 하나 또는 그 이상의 단계에서 사용될 수 있고 이들 방법 중 어떠한 것에서도 사용될 수 있다. 예를 들어 용액은 주형으로부터 연장된 이중나선을 분리하기 위해 사용될 수 있다. 용액은 더 이상 연장된 이중나선에 부착되지 않는 연장 프로브의 부분을 제거하기 위하여 잘라지기 쉬운 연쇄의 절단 후에 사용될 수 있다. 용액은 또한 삼중-가닥 핵산의 가닥을 분리하거나 상호 간에 혼성화된 자체-

상보 부분을 함유하는 단일 핵산 가닥의 이중-가닥 영역을 분리하는 데 사용될 수 있다.

- <22> 다른 측면으로, 본 발명은 최소한 두 개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군(family)을 사용하여 서열에 대한 정보를 얻는 방법을 제공한다. 프로브 군 중의 프로브는 구속되지 않는 부분과 구속된 부분을 함유한다. 상기에서 설명된 방법에서와 같이, 연장은 개시 올리고뉴클레오타이드와 주형에 의해 형성된 이중나선으로부터 시작된다. 개시 올리고뉴클레오타이드는 연장된 이중나선을 형성하기 위하여 그것의 말단에 올리고뉴클레오타이드 프로브를 결합하여 연장되고, 그것은 다음에 연속적인 결합 사이클에 의해 반복적으로 연장된다. 프로브는 말단 위치에 (이중나선의 성장하는 핵산 가닥에 결합된 뉴클레오타이드로부터 프로브의 반대쪽 말단에) 비-연장성 부분을 가지므로 단일 사이클에서 연장된 이중나선에 단지 한 번의 연장만이 일어난다. 각 사이클 동안에, 연속적으로 결합된 프로브 상의 또는 프로브와 결합된 표지가 검출되고, 비-연장성 부분은 제거되거나 연장 가능한 말단이 생성되도록 변형된다. 표지는 프로브가 속하는 프로브 군에 상응한다.
- <23> 연장, 결합, 및 검출의 연속적인 사이클로 연속해서 연속적으로 결합된 프로브가 속하는 프로브 군의 순서 리스트가 생성된다. 프로브 군의 순서 리스트는 서열에 대한 정보를 얻기 위해 사용된다. 그러나 새롭게 결합된 프로브가 속하는 프로브 군에 대해 아는 것 자체로는 주형의 뉴클레오타이드의 정체에 결정하기에 충분하지 않다. 그 대신, 새롭게 결합된 프로브가 속하는 프로브 군에 대해 아는 것은 프로브의 구속된 부분의 서열에 대한 가능성으로서 특정 서열을 제거하지만 각 위치에 있는 뉴클레오타이드의 정체에 대한 최소한 두 가지 가능성을 남겨둔다. 그러므로 새롭게 결합된 프로브의 구속된 부분에 있는 뉴클레오타이드에 대하여 반대 위치에 위치한 주형의 뉴클레오타이드 (그 프로브의 구속된 부분에 있는 뉴클레오타이드에 상보하는 뉴클레오타이드)의 정체에 대해서는 최소한 두 가지 가능성이 남아있다.
- <24> 특정 실시예에서, 원하는 수의 사이클을 수행한 후에, 프로브 군 정체의 순서 연속물(series)을 사용하여 후보 서열의 세트가 생성된다. 후보 서열 세트는 목표물을 이루기 위하여 충분한 정보를 제공할 수 있다. 본 발명의 바람직한 실시예에서, 후보 서열 중에서 정확한 서열을 선택하기 위하여 하나 또는 그 이상의 추가의 단계가 수행된다. 예를 들어 서열은 공지된 서열의 데이터베이스와 비교될 수 있고, 데이터베이스의 서열 중 하나와 가장 밀접한 후보 서열이 정확한 서열로서 선택된다. 다른 실시예에서, 주형에 대해 상이하게 코드화된 세트의 프로브 군을 사용하여 연장, 결합, 검출, 및 절단의 연속적인 사이클에 의해 다른 회기(round)의 서열화가 수행되고, 두 번째 회기에서 얻어진 정보는 정확한 서열을 선택하기 위해 사용된다. 다른 실시예에서 최소한 하나의 정보 항목은 서열을 결정하기 위해 프로브 군 정체의 순서 리스트로부터 얻어진 정보와 조합된다.
- <25> 본 발명은 또한 주형이 프로브 군을 사용하여 서열화될 때 에러 조사를 수행하는 방법을 제공한다. 방법 중 어떤 것은 단일 뉴클레오타이드 다형성 (SNP)과 서열화 에러 사이를 구분한다.
- <26> 본 발명은 또한 최소한 2개의 관심 있는 절편 (예컨대 최소한 두 개의 태그)와 최소한 세 개의 프라이머 결합 영역 (PBR)을 함유하는 핵산 단편 (예컨대 DNA 단편)을 제공하며, 그로써 각각 관심 있는 절편에 상응하는 최소한 두 개의 구별되는 주형이 각 단편으로부터 증폭될 수 있다. "프라이머 결합 영역"은 올리고뉴클레오타이드가 증폭 프라이머, 서열화 프라이머, 개시 올리고뉴클레오타이드 등으로서 작용할 수 있도록 혼성화할 수 있는 핵산의 부분이다. 그러므로 프라이머 결합 영역은 적당한 상보 올리고뉴클레오타이드의 선택을 가능하게 하기 위해 공지된 서열을 가져야만 한다. 본원과 도면에서 사용되는 바와 같이, 본 발명의 방법에 사용된 핵산 가닥의 부분은, 방법의 실시에서, 프라이머가 실제로 영역에 결합하거나 핵산 가닥의 상보하는 가닥의 상응하는 부분에 결합하는 것에 관계없이 프라이머 결합 영역으로서 언급될 수 있다. 그러므로 핵산 부분은 본 발명의 방법에 사용될 때, 프라이머가 실제로 그 영역에 결합하거나 (프라이머의 서열이 영역의 서열과 상보하거나 실질적으로 상보하는 경우) 또는 영역의 상보물에 결합하거나 (프라이머의 서열이 프라이머 결합 영역의 서열에 동일하거나 실질적으로 동일한 경우)에 관계없이 프라이머 결합 영역으로서 언급될 수 있다. 관심 있는 절편은 그것에 대한 서열 정보가 요구되는 핵산의 어떠한 절편이다. 예를 들어 관심 있는 서열은 태그일 수 있고, 본 명세서의 목적의 경우 관심 있는 절편이 태그라고 가정될 것이다 (본원 및 다른 곳에서 "말단 태그"로서 언급됨). 그러나 본 발명은 태그인 관심 있는 절편에 제한되지 않는다는 것이 인지되어야 한다. 어떤 실시예에서 최소한 두 개의 태그는 쌍을 이루는 태그이다. 핵산 단편은 하나 또는 그 이상의 태그, 예컨대 하나 또는 그 이상의 쌍을 이루는 태그, 예컨대 2, 3, 4, 5 또는 그 이상의 쌍의 쌍을 이루는 태그를 함유할 수 있다. 본 발명은 또한 그러한 핵산 단편을 함유하는 라이브러리, 및 주형과 라이브러리를 제조하기 위한 방법을 제공한다.
- <27> 본 발명은 추가로 최소한 두 개의 구별되는 집단의 핵산이 부착된 마이크로 입자, 예컨대 비드를 제공하는데, 이때 최소한 두 개의 집단은 각각 다수의 실질적으로 동일한 핵산으로 구성되며, 집단은 단일 핵산 단편으로부터 증폭 (예컨대 PCR 증폭)에 의해 생성되었다. 어떤 실시예에서 단일 핵산 단편은 5' 태그와 3' 태그를 함유하

며, 이때 5' 및 3' 태그는 쌍을 이룬 태그이다. 단일 핵산 단편이 한 쌍의 5' 태그와 3' 태그를 함유하는 어떤 실시예에서, 마이크로입자에 부착된 핵산의 집단 중 하나는 최소한 한 부분의 5' 태그를 포함하고, 마이크로입자에 부착된 핵산의 집단 중 하나는 최소한 한 부분의 3' 태그를 포함한다. 바람직한 실시예에서 집단 중 하나는 완전한 5' 태그를 포함하고 한 집단은 완전한 3' 태그를 포함한다.

<28> 핵산 단편은 다중 PBR을 함유하는데, 그것들 중 최소한 하나는 태그 사이에 위치하고 그것들 중 최소한 두 개는 그 태그를 함유하는 핵산 단편의 부분 양 옆에 있어서, 최소한 일부의 5' 태그를 포함하는 영역이 증폭될 수 있으며, 최소한 일부의 3' 태그를 포함하는 영역이 증폭될 수 있어서 핵산의 두 개의 구별되는 집단이 생성될 수 있다. 바람직한 실시예에서, 전체 5' 태그 및 전체 3' 태그가 증폭될 수 있다. 예를 들어, 핵산 단편은 5' 태그 양 옆에 있는 첫 번째 및 두 번째 프라이머 결합 부위를 함유할 수 있으며, 또한 3' 태그 양 옆의 세 번째 및 네 번째 프라이머 결합 부위를 함유할 수 있다. 첫 번째와 두 번째 프라이머 결합 부위에 결합하는 프라이머를 사용한 PCR 증폭으로 5' 태그가 증폭된다. 세 번째와 네 번째 프라이머 결합 부위에 결합하는 프라이머를 사용한 PCR 증폭으로는 3' 태그가 증폭된다. 프라이머는 각 프라이머로부터의 연장이 증폭될 태그를 함유하는 DNA 단편의 영역을 향하여 진행되도록 선택되어야 한다는 것이 인지될 것이다. 이와 달리, 첫 번째 프라이머 결합 부위는 한 태그의 상류에 위치할 수 있고, 두 번째 프라이머 결합 부위는 다른 태그의 하류에 위치할 수 있으며, 세 번째 프라이머 결합 부위는 두 태그 사이에 위치할 수 있다. 세 번째 프라이머 결합 부위는 태그들 중 하나를 증폭하는 PCR 증폭에 대한 전방 프라이머에 대한 프라이머 결합 부위로서 작용하고, 다른 태그를 증폭하는 PCR 증폭에 대한 역 프라이머에 대한 결합 부위로서도 작용한다. 그러므로 한 실시예에서 본 발명은 최소한 두 개의 구별되는 집단의 핵산이 부착되는 마이크로입자, 예컨대 비드를 제공하는데, 이때 최소한 두 집단은 각각 실질적으로 동일한 다수의 핵산으로 구성되고, 첫 번째 구별되는 집단은 5' 태그를 포함하며 두 번째 구별되는 집단은 3' 태그를 포함한다.

<29> 본 발명은 또한 마이크로입자, 예컨대 비드의 집단을 제공하는데, 이때 개별적인 마이크로입자는 거기에 부착된 최소한 두 개의 구별되는 핵산 집단을 가지며, 이 때 최소한 두 개의 집단은 각각 실질적으로 동일한 다수의 핵산으로 구성되며, 집단은 단일 DNA 단편으로부터 증폭 (예컨대 PCR 증폭)에 의해 생성되었다. 실질적으로 동일한 집단은 예컨대 5' 태그와 3' 태그일 수 있다. 본 발명은 추가로 그러한 마이크로입자의 배열과, 실질적으로 동일한 핵산의 집단의 서열화를 포함하는 서열화 방법을 제공한다. 예를 들어 한 실시예에서, 개별적인 마이크로입자에 부착된 실질적으로 동일한 핵산의 두 집단은 각각 상이한 프라이머 결합 영역 (PBR)을 포함하고, 따라서 상이한 서열화 프라이머를 사용함으로써, 집단 중 하나는 다른 집단으로부터 간섭받지 않고 서열화될 수 있다. 만약 실질적으로 동일한 핵산의 둘 이상의 실질적으로 동일한 집단이 단일 마이크로입자에 부착된다면, 집단은 각각 독특한 PBR을 가질 수 있어서, 주어진 PBR에 결합하는 프라이머는 마이크로입자에 부착된 핵산의 실질적으로 동일한 다른 집단에 존재하는 PBR에 결합하지 않는다. 그러므로 본 발명의 방법은 실질적으로 동일한 핵산의 최소한 두 개의 상이한 집단 (예컨대 5' 태그를 함유하는 주형의 다수의 복사물과 3' 태그를 함유하는 주형의 다수의 복사물)이 부착된 마이크로입자의 생성을 가능하게 하는데, 이때 태그는 쌍을 이룬 태그이다. 본 발명의 방법에 따르면, 주형은 서열화 프라이머에 대한 결합 부위를 제공하는 상이한 PBR을 함유한다. 그러므로 5' 태그를 함유하는 주형의 PBR에 상보하는 서열화 프라이머를 선택함으로써, 3' 태그를 함유하는 주형이 또한 동일한 마이크로입자에 존재할지라도 3' 태그를 함유하는 주형으로부터 간섭받는 일 없이 5' 태그로부터 서열 정보를 얻을 수 있다. 3' 태그를 함유하는 주형의 PBR에 상보하는 서열화 프라이머를 선택함으로써, 5' 태그를 함유하는 주형이 또한 동일한 마이크로입자에 존재할지라도 5' 태그를 함유하는 주형으로부터 간섭받는 일 없이 3' 태그로부터 서열 정보를 얻을 수 있다. 쌍을 이룬 태그 둘 다 동일한 마이크로입자에 존재한다는 사실은 그것들이 선행 기술에서와 같이 단일 주형 내에 존재한다면 그런 것처럼, 5' 및 3'의 쌍을 이룬 태그의 서열이 상호 간에 결합할 수 있다는 것을 의미한다.

<30> 본 발명은 또한 예컨대 실질적으로 평면인 지지체에 또는 그 위에 배열된 주형을 서열화하기 위해 사용될 수 있는 자동 서열화 시스템을 제공한다. 본 발명은 나아가 하드 디스크, CD, 집(zip) 디스크, 플래시 메모리 등과 같은 컴퓨터-판독가능 매질 상에 저장될 수 있는 영상 처리 방법을 제공한다. 바람직한 어떤 실시예에서 시스템은 일 초당 40,000 또는 그 이상의 뉴클레오티드를 확인한다. 바람직한 어떤 실시예에서 시스템은 하루 (24시간)에 8.6 기가바이트 (Gb) 또는 그 이상의 서열 데이터를 생성한다. 바람직한 어떤 실시예에서 시스템은 하루에 48Gb 또는 그 이상의 서열 정보 (뉴클레오티드 확인)를 생성한다.

<31> 또한 본 발명은 본 발명의 서열화 방법을 적용함으로써 생성된 정보를 저장하는 컴퓨터-판독가능한 매질을 제공한다. 정보는 데이터베이스로 저장될 수 있다.

<32> 본 출원은 다양한 특허, 특허 출원, 잡지의 기사, 및 다른 공보를 언급하는데, 그것들은 모두 본원에 참조로 삽

입된다. 또한, 다음의 표준 참조 작업이 본원에 참조로 삽입된다: Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., edition as of July 2002; Sambrook, Russell, and Sambrook, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001. 본 명세서와 참조로 삽입된 어떠한 문헌 사이에 불일치가 있는 경우에 본 명세서가 기준이며, 어떤 충돌 또는 불일치가 존재하는 지를 결정하는 것은 본 발명자들이 결정할 사항이며 어떠한 때에라도 이루어질 수 있다는 것이 인지되어야 한다.

발명의 상세한 설명

<78> 정의

<79> 본 발명의 설명을 이해하기 용이하도록 하기 위하여 다음의 정의를 제공한다. 일반적으로 다르게 규정되지 않는 용어들은 당해 기술분야에서 일반적으로 허용되는 것과 같은 의미 또는 의미들로 제공되는 것으로 인지되어야 한다.

<80> 본원에서 사용되는 "비염기성 잔기(abasic residue)"는 질소 함유 염기의 제거 또는 질소 함유 염기의 충분한 부분의 제거 후에 남아있어서 그 결과의 분자가 더 이상 뉴클레오타이드 또는 뉴클레오타이드의 특징적인 수소 결합에 참여할 수 없는 뉴클레오타이드 또는 뉴클레오타이드 부분의 구조를 가지는 잔기이다. 비염기성 잔기는 뉴클레오타이드 또는 뉴클레오타이드로부터 질소 함유 염기를 제거하여 생성될 수 있다. 그러나 용어 "비염기성"은 잔기의 구조적 특징을 언급하기 위해 사용된 것이고, 잔기가 생성되는 방식과는 무관하다. 용어 "비염기성 잔기"와 "비염기성 부위"는 본원에서 퓨린 또는 피리미딘 염기가 없는 핵산 내의 잔기를 언급하기 위해 사용된다.

<81> 본원에서 사용되는 "비퓨린/비피리미딘(apurinic/apyrimidinic;AP) 엔도뉴클레아제(endonuclease)"는 폴리뉴클레오타이드의 비염기성 잔기의 5'측, 3'측, 또는 5'과 3'측 둘 다에서 결합을 절단하는 효소를 언급한다. 본 발명의 어떤 실시예에서 AP 엔도뉴클레아제는 AP 리아제(lyase)이다. AP 엔도뉴클레아제의 실례로는, 그것들에 한정되는 것은 아니지만, 대장균 엔도뉴클레아제 VIII 및 그것의 동족체(homologue)와 대장균 엔도뉴클레아제 III 및 그것의 동족체가 있다. 특이한 효소, 예컨대 대장균 Endo VIII, Endo V 등과 같은 엔도뉴클레아제에 대한 언급은 당해 기술분야에서 동족체인 것으로 인지되고 손상된 염기의 제거 및/또는 비염기성 잔기 또는 다른 트리거 잔기를 함유하는 DNA의 절단과 관련하여 유사한 생화학적 활성을 가지는 것으로서 인지된 다른 종으로부터의 동족체를 포함하는 것으로 의도되었음이 인지되어야 한다.

<82> 본원에서 사용되는 용어 "배열(array)"은 지지체 매트릭스에 또는 그 위에 분포된 실체들의 집합을 말한다; 바람직하게는, 개별적인 실체는 다양한 기법들 중 어떠한 것에 의해서든 배열의 뚜렷한 특징이 확인되는 것을 가능하게 하기에 충분한 상호간 거리에 놓이게 된다. 예를 들어 실체는 핵산 분자, 핵산 분자들의 클론 집단, 마이크로입자 (임의로 핵산 분자들이 부착된 클론 집단을 가짐) 등일 수 있다. 동사로서 사용될 때 용어 "배열" 및 그것의 변용은 배열을 형성하는 어떠한 과정, 예컨대 실체를 지지체 매트릭스에 또는 그 위에 분포시키는 과정을 말한다.

<83> "손상된 염기(damaged base)"는 A, G, C, 또는 T와는 상이한 퓨린 또는 피리미딘 염기로서, DNA 글리코실라제에 의해 DNA로부터 제거하기 위한 기질이 되기 쉽다. 우라실은 본 발명의 목적을 위해 손상된 염기로 간주된다. 발명의 어떤 실시예에서 손상된 염기는 히포크산틴(hypoxanthine)이다.

<84> 폴리뉴클레오타이드의 집단 중 하나인 폴리뉴클레오타이드의 위치와 관련하여 "축퇴성(degenerate)"은 뉴클레오타이드의 부분을 형성하는 염기의 정체가 집단의 상이한 구성원들 중에서도 달라지는 위치를 차지하는 것을 의미한다. 그러므로 집단은 그것의 서열이 축퇴성 위치에서는 상이한 개별적인 구성원들을 함유한다. 용어 "위치(position)"는 일반적으로 5' 또는 3' 말단과 관련하여 폴리뉴클레오타이드의 각 뉴클레오타이드에 지정되는 숫자 값을 말한다. 예를 들어, 연장 프로브의 3' 말단에 있는 뉴클레오타이드는 위치 1로 지정될 수 있다. 그러므로 구조 3'-XXXXXXX-5'의 연장 프로브 풀(pool)에서 N은 위치 4에 있다. 위치 4는 만약 풀의 상이한 구성원에서 N의 정체가 다를 수 있다면 축퇴성인 것으로 간주된다. 연장 프로브의 풀은 또한 위치 N에서 축퇴성이라고 말한다. 위치는 만약 그것이 어떠한 k개의 상이한 정체를 가지는 뉴클레오타이드에 의해 차지될 수 있다면 k-배 축퇴성이라고 말한다. 예를 들어 2개의 상이한 염기 중 어느 하나를 포함하는 뉴클레오타이드가 차지하고 있는 위치는 2-배 축퇴성이다.

<85> "서열에 대한 정보를 결정하는 것"은 "서열 결정(sequence determination)"을 포함하고, 또한 서열에 대하여 하나 또는 그 이상의 가능성을 제거하는 것과 같은 다른 수준의 정보를 포함한다. 폴리뉴클레오타이드에 대해 서열

결정을 수행하는 것은 전형적으로 완벽하게 상보하는 (100% 상보하는) 폴리뉴클레오타이드의 서열에 관한 동등한 정보를 낳는 것이고, 따라서 직접 완벽하게 상보하는 폴리뉴클레오타이드에 대해 수행된 서열 결정과 동등하다.

<86> 다수의 구성요소, 예컨대 한 올리고뉴클레오타이드 프로브 분자 또는 그것의 부분의 뉴클레오타이드와 관련하여 "독립적인" 것은 각 구성요소의 정체는 제한되지 않으며 어떠한 다른 구성요소의 정체에 의해 제한되지 않는다는 것, 예컨대 각 구성요소의 정체는 어떠한 다른 구성요소(들)의 정체와 관계없이 선택된다는 것을 의미한다. 그러므로 하나 또는 그 이상의 구성요소의 정체를 아는 것이 어떠한 다른 구성요소의 정체성에 관한 모든 정보를 다 제공하는 것은 아니다. 예를 들어 서열 NNNN의 뉴클레오타이드는 만약 각 N의 정체가 A, G, C, 또는 T라면 어떠한 다른 N의 정체와 관계없이 독립적이다.

<87> "결찰(ligation)"은 주형-구동된 반응에서 둘 또는 그 이상의 핵산, 예컨대 올리고뉴클레오타이드 및/또는 폴리뉴클레오타이드의 말단 사이의 공유 결합 또는 연쇄를 형성하는 것을 의미한다. 결합 또는 연쇄의 특징은 광범위하게 달라질 수 있으며 결찰은 효소적으로 또는 화학적으로 수행될 수 있다.

<88> 본원에서 사용되는 용어 "마이크로입자"는 50 미크론 또는 그 이하, 바람직하게는 10 미크론 또는 그 이하의 가장 작은 횡단면 크기를 가지는 입자를 말한다. 어떤 실시예에서 가장 작은 횡단면 크기는 대략 3 미크론 이하, 대략 1 미크론 이하, 대략 0.5 미크론 이하, 예컨대 대략 0.1, 0.2, 0.3, 또는 0.4 미크론이다. 마이크로입자는 그것들에 한정되지는 않지만 유리 (예컨대 제어진 기공 유리), 실리카, 지르코니아, 교차 결합된 폴리스티렌, 폴리아크릴레이트, 폴리메틸메타크릴레이트, 이산화 티탄, 라텍스, 폴리스티렌 등을 포함하는 다양한 무기 또는 유기 물질로 구성될 수 있다. 예컨대 미국 특허 6,406,848호에는 다양한 적당한 물질 및 다른 고려사항들이 기재되어 있다. Dynal사 (Oslo, Norway)로부터 구할 수 있는 다이나 비드 (Dyna beads)는 본 발명에 사용되는 시판중인 마이크로입자의 실례이다. 자기적으로 반응하는 마이크로입자가 사용될 수 있다. 바람직한 특정 마이크로입자의 자성 반응성은 증폭 후에 마이크로입자-부착된 주형의 손쉬운 수집과 농축을 가능하게 하며, 추가 단계들 (예컨대 세척, 시약 제거 등)을 촉진한다. 본 발명의 어떤 실시예에서 상이한 모양을 가지는 (예컨대 일부는 구형이고 나머지는 비구형임) 마이크로입자 집단이 사용된다.

<89> 본원에서 사용되는 용어 "마이크로스피어(microsphere)" 또는 "비드"는 50 미크론 이하, 바람직하게는 10 미크론 이하의 직경을 가지는 실질적으로 구형의 마이크로입자를 말한다. 어떤 실시예에서 직경은 대략 3 미크론 이하, 대략 1 미크론 이하, 대략 0.5 미크론 이하, 예컨대 대략 0.1, 0.2, 0.3, 또는 0.4 미크론이다. 본 발명의 어떤 실시예에서 단일분산형 마이크로스피어의 집단이 사용된다. 즉 마이크로스피어는 실질적으로 균일한 크기이다. 예를 들어 마이크로입자의 직경은 5% 이하, 예컨대 2% 이하, 또는 1% 이하의 변동 계수를 가질 수 있다. 그러나 다른 실시예에서 마이크로입자의 집단의 변동 계수는 5% 이상, 예컨대 5%, 5%와 10% 사이 (포괄적), 10%와 25% 사이 (포괄적) 등이다. 어떤 실시예에서 마이크로입자의 혼합된 집단이 사용된다. 예를 들어 각각이 5% 미만의 변동 계수를 가지는 두 집단의 혼합물이 사용될 수 있고, 그 결과 단일분산형이 아닌 혼합된 집단이 생성된다. 실례로서, 직경이 1 미크론인 마이크로스피어와 3 미크론인 마이크로스피어의 혼합물이 사용될 수 있다. 본 발명의 어떤 실시예에서 추가의 정보는 서열화가 단일분산형이 아닌 집단의 마이크로스피어에 부착된 주형을 사용하여 수행될 때 마이크로스피어의 크기에 의해 제공된다. 예를 들어 주형의 상이한 라이브러리가 상이한 크기의 마이크로스피어에 부착될 수 있다. 또한, 더 적은 수의 주형 분자가 더 작은 입자에 부착될 수 있기 때문에, 신호의 세기는 달라질 수 있으며, 그것은 다중나선 서열화를 촉진할 수 있다.

<90> 본원에서 사용되는 용어 "핵산 서열"은 핵산 물질 자체를 언급할 수 있으며, 특이한 핵산, 예컨대 DNA 또는 RNA 분자를 생화학적으로 특징짓는 서열 정보 (즉 5가지 염기 문자 A, G, C, T, 또는 U 중에서 선택된 문자의 연속)에 제한되지 않는다. 본원에 제시된 핵산은 다른 언급이 없는 한 5'→3' 방향으로 표시된다.

<91> "뉴클레오타이드(nucleoside)"는 당 분자에 결합된 질소 함유 염기를 포함한다. 본원에서 사용되는 이 용어는 문헌에 설명되어 있는 것과 같이, 그것들의 2'-데옥시 및 2'-히드록실 형태의 천연 뉴클레오타이드 및 뉴클레오타이드 유사체를 포함한다 (Kornberg and Baker, DNA Replication, 2nd Ed.(Freeman, San Francisco, 1992)). 예를 들어 천연 뉴클레오타이드는 아데노신, 티미딘, 구아노신, 시티딘, 우리딘, 데옥시아데노신, 데옥시티미딘, 데옥시구아노신, 및 데옥시시티딘을 포함한다. 뉴클레오타이드 "유사체"는 변형된 염기 부분 및/또는 변형된 당 부분을 가지는 합성 뉴클레오타이드, 예컨대 문헌에 기재되어 있는 것들을 말한다 (Scheit, Nucleotide Analogs, John Wiley, New York, 1980). 그러한 유사체는 결합 특성을 증강시키고, 축퇴성을 감소시키며, 특이성을 증가시키는 등의 목적으로 디자인된 합성 뉴클레오타이드를 포함한다. 뉴클레오타이드 유사체로는 2-아미노아데노신, 2-티오티미딘, 피롤로-피리미딘, 3-메틸 아데노신, C5-프로피닐시티딘, C5-프로피닐우리딘, C5-브로모우리딘, C5-플루오로우리딘, C5-요오도우리딘, C5-메틸시티딘, 7-테아자아데노신, 7-테아자구아노신, 8-옥소아데노신, 8-옥소구아

노신, 0(6)-메틸구아닌, 2-티오시티딘 등이 있다. 뉴클레오시드 유사체는 본원에서 언급되는 보편적인 염기 중 어느 것이라도 포함할 수 있다.

- <92> 본원에서 사용된 용어 "유기체(organism)"는 복제될 수 있고 서열 결정을 위해 관심 있는 대상인 핵산을 포함한다. 그것은 플라스미드; 바이러스; 원핵, 고세균 및 진핵 세포, 셀 라인, 진균, 원생동물, 식물, 동물 등을 포함한다.
- <93> 프로브와 주형 폴리뉴클레오타이드의 돌출하는 가닥과 관련하여 "완벽하게 매치된 이중나선"은 한 가닥으로부터 돌출하는 가닥이 다른 것과 이중 가닥의 구조를 형성하여 이중 가닥의 구조에 있는 각각이 뉴클레오시드가 반대 가닥에 있는 뉴클레오시드와 함께 왓슨-크릭 염기쌍을 이루는 것을 의미한다. 이 용어는 또한 테옥시이노신, 2-아미노퓨린 염기를 가지는 뉴클레오시드 등과 같은 뉴클레오시드 유사체의 쌍을 이루는 것도 포함하며, 그것은 쌍을 이루는 것이 수소 결합의 형성을 포함하거나 포함하지 않거나 프로브의 축퇴성을 감소시키기 위해 사용될 수 있다.
- <94> 용어 "다수성(plurality)"은 하나 이상을 의미한다.
- <95> 용어 "다형성(polymorphism)"은 당해 기술분야에서 사용되는 원래의 의미로 제공되며, 동일한 종의 개체 중에서 게놈 서열에 차이가 있는 것을 말한다. "단일 뉴클레오타이드 다형성 (SNP)"은 단일 위치에서의 다형성을 나타낸다.
- <96> "폴리뉴클레오타이드", "핵산", 또는 "올리고뉴클레오타이드"는 뉴클레오시드간 연쇄에 의해 연결된 뉴클레오시드의 선형 중합체 (테옥시리보뉴클레오시드, 리보뉴클레오시드, 또는 그것들의 유사체를 포함함)를 나타낸다. 전형적으로 폴리뉴클레오타이드는 최소한 3개의 뉴클레오시드를 포함한다. 본 발명의 어떤 실시예에서 연장 프로브의 하나 또는 그 이상의 뉴클레오시드는 보편적인 염기를 포함한다. 통상 올리고뉴클레오타이드의 크기는 몇 개의 단량체 유닛, 예컨대 3 내지 4개로부터 수백개의 단량체 유닛까지 걸쳐 있다. 올리고뉴클레오타이드와 같은 폴리뉴클레오타이드가 연속되는 문자들에 의해, 예컨대 "ATGCTG"와 같이 표시될 때는 언제나 뉴클레오타이드는 왼쪽으로부터 오른쪽으로 5'→3' 순서이며, 다른 언급이 없는 한 "A"는 테옥시아데노신을 나타내고, "C"는 테옥시시티딘을 나타내며, "G"는 테옥시구아노신을 나타내고, "T"는 티미딘을 나타낸다는 것이 인지될 것이다. 문자 A, C, G, 및 T는 당해 기술분야에서 공인된 것과 같이 염기 자체, 뉴클레오시드, 또는 염기를 포함하는 뉴클레오타이드를 나타내기 위해 사용될 수 있다.
- <97> 자연 발생 폴리뉴클레오타이드에서 뉴클레오시드간 연쇄는 전형적으로 인산이에스테르 결합이며, 하위단위는 "뉴클레오타이드"로서 언급된다. 그러나 다른 뉴클레오시드간 연쇄, 예컨대 포스포로티올레이트 연쇄를 포함하는 올리고뉴클레오타이드 프로브가 본 발명의 특정 실시예에 사용된다. 비-인산이에스테르 연쇄로 그러한 올리고뉴클레오타이드 프로브를 구성하는 하나 또는 그 이상의 하위단위는 포스페이트기를 포함하지 않을 것임이 인지될 것이다. 그러한 뉴클레오타이드의 유사체는 본원에서 사용되는 용어 "뉴클레오타이드"의 범주 안에 속하는 것으로 간주되며, 인산이에스테르 연쇄가 아닌 하나 또는 그 이상의 뉴클레오시드간 연쇄를 포함하는 핵산은 여전히 "폴리뉴클레오타이드", "올리고뉴클레오타이드" 등으로 언급된다. 다른 실시예에서 올리고뉴클레오타이드 프로브와 같은 폴리뉴클레오타이드는 AP 엔도뉴클레아제 민감 부위를 함유하는 연쇄를 포함한다. 예를 들어 올리고뉴클레오타이드 프로브는 비염기성 잔기, DNA 글리코실라제에 의해 제거되기 위한 기질인 손상된 염기를 함유하는 잔기, 또는 AP 엔도뉴클레아제에 의한 절단에 대한 기질인 다른 잔기 또는 연쇄를 함유할 수 있다. 다른 실시예에서 올리고뉴클레오타이드 프로브는 이당 뉴클레오시드를 함유한다.
- <98> 용어 "프라이머"는 전형적으로 길이가 약 10 내지 100 뉴클레오타이드의 폴리뉴클레오타이드로서, 표적과 혼성화하여 표적 폴리뉴클레오타이드 또는 "주형"과 결합하는 짧은 폴리뉴클레오타이드를 나타낸다. 프라이머는 바람직하게는 표적에 상보하는 폴리뉴클레오타이드의 주형-지시된 합성에 대한 개시 지점을 제공하는데, 그런 합성은 적절한 효소(들), 보조인자, 기질, 예컨대 뉴클레오타이드, 올리고뉴클레오타이드 등의 존재하에 일어날 수 있다. 프라이머는 전형적으로 그곳으로부터 연장될 수 있는 말단을 제공한다. DNA 중합효소와 같은 중합효소에 의해 촉매된 합성에 대한 프라이머의 경우에 (예컨대 "합성에 의한 서열화", 중합효소 사슬 반응 (PCR) 증폭 등에서), 프라이머는 전형적으로 유리 3' OH 기를 가지거나, 가지도록 변형될 수 있다. 전형적으로, PCR 반응은 증폭될 영역의 한계를 정하는 "상류" (또는 "전방") 프라이머와 "하류" (또는 "역") 프라이머를 포함하는 한 쌍의 프라이머 (첫 번째와 두 번째 증폭 프라이머)를 사용한다. 연속적인 사이클의 연장, 결찰 (및 임의로 절단)에 의한 합성에 대한 프라이머의 경우, 프라이머는 전형적으로 DNA 결합 효소에 대한 기질로서 작용하는 유리 5' 포스페이트기 또는 3' OH기를 가지거나, 가지도록 변형될 수 있다.

- <99> 본원에서 사용되는 "프로브 군(family)"은 각각이 동일한 표지를 포함하는 프로브들의 그룹을 나타낸다.
- <100> 본원에서 폴리뉴클레오티드와 관련하여 사용되는 "서열 결정", "뉴클레오티드 서열의 결정", "서열화" 등의 용어는 폴리뉴클레오티드의 부분적인 결정뿐 아니라 전체적인 서열 정보의 결정을 포함한다. 즉 이 용어는 서열 비교, 핑거프린팅, 및 표적 폴리뉴클레오티드에 관한 유사한 수준의 정보뿐만 아니라 관심 있는 영역 안에 있는 표적 폴리뉴클레오티드의 각 뉴클레오티드의 발현을 확인하고 순서를 정하는 것을 포함한다. 본 발명의 어떤 실시예에서 "서열 결정"은 단일 뉴클레오티드의 확인을 포함하는 한편, 다른 실시예에서는 하나 이상의 뉴클레오티드가 확인된다. 본 발명의 어떤 실시예에서는 그 자체로는 단일 사이클에서 어떤 뉴클레오티드를 확인하는 것이 불충분한 서열 정보를 모은다. 뉴클레오티드, 뉴클레오티드, 및/또는 염기의 확인은 본원에서 동등한 것으로 간주한다. 폴리뉴클레오티드에 대하여 서열 결정을 수행하는 것은 전형적으로 완벽하게 상보하는 (100% 상보하는) 폴리뉴클레오티드의 서열에 관한 동등한 정보를 산출하며, 따라서 완벽하게 상보하는 폴리뉴클레오티드에 대하여 직접 수행된 서열 결정과 동등하다는 것이 주지된다.
- <101> 본원에서 사용되는 "서열화 반응"은 연장, 절찰, 및 검출의 사이클의 한 세트를 말한다. 연장된 이중나선이 주형으로부터 제거되고 두 번째 세트의 사이클이 주형에 대해 수행될 때, 그 결과로 생성되는 서열 정보가 단일 서열을 생성하기 위해 조합될 수 있지만, 각 세트의 사이클은 별도의 서열화 반응으로 간주한다.
- <102> 본원에서 사용되는 "반-고체(semi-solid)"는 고체와 액체 성분을 둘 다 가지고 있는 압축가능한 매트릭스를 나타내는 것으로, 그 안의 액체는 고체 매트릭스 요소들 사이에 있는 기공, 공간 또는 다른 틈새를 차지한다. 예시적인 반-고체 매트릭스로는 폴리아크릴아미드, 셀룰로스, 폴리아미드 (나일론), 및 교차 결합된 아가로스, 텍스트란 및 폴리에틸렌 글리콜로 만들어진 매트릭스들을 들 수 있다. 반-고체 지지체는 기관으로도 언급되는 두 번째 지지체, 예컨대 실질적으로 평면이고 견고한 지지체 위에 제공될 수 있으며, 그 지지체는 반-고체 지지체를 지지한다.
- <103> 본원에서 사용되는 "지지체(support)"는 그 위에 또는 그 안에서 핵산 분자, 마이크로입자 등이 고정될 수 있는, 즉 그것에 핵산 분자나 마이크로입자가 공유적으로 또는 비 공유적으로 부착되거나, 또는 그 안에 또는 그 위에 핵산 분자나 마이크로입자가 부분적으로 또는 전체적으로 삽입되어서 그것들이 상호 간에 자유롭게 확산되거나 이동하는 것이 거의 또는 완전히 방지되는 매트릭스를 말한다.
- <104> "트리거 잔기"는 핵산에 존재할 때 핵산이 절단제(cleavage agent) (예컨대 효소, 질산 은 등) 또는 제제의 조합에 의해 절단되는 것을 그렇지 않으면 트리거 잔기를 포함하지 않은 동일할 핵산보다 더 쉽게 만들거나, 및/또는 핵산이 그러한 절단에 대해 보다 더 민감하게 만드는 잔기를 생성하기 위해 변형되기가 쉬운 잔기이다. 그러므로 트리거 잔기가 핵산에 존재하는 것은 핵산의 잘라지기 쉬운 연쇄의 존재를 초래할 수 있다. 예를 들어 비염기성 잔기는 트리거 잔기인데, 왜냐하면 핵산 내의 비염기성 잔기의 존재로 인해 핵산이 AP 엔도뉴클레아제와 같은 효소에 의해 절단되기 쉽게 되기 때문이다. 손상된 염기를 함유하는 뉴클레오티드도 트리거 잔기인데, 왜냐하면 손상된 염기를 핵산에 포함하는 뉴클레오티드의 존재가 예컨대 DNA 글리코실라제에 의한 손상된 염기의 제거 후에, 핵산이 AP 엔도뉴클레아제와 같은 효소에 의해 절단되기 쉽게 만들기 때문이다. 절단 부위는 트리거 잔기와 인접한 잔기 사이의 결합이거나 트리거 잔기로부터 제거된 하나 또는 그 이상의 잔기인 결합일 수 있다. 예를 들어 테옥시이노신은 트리거 잔기인데, 왜냐하면 핵산 내의 테옥시이노신의 존재로 인해 핵산은 대장균 엔도뉴클레아제 V 및 그것의 동족체에 의한 절단이 더 쉽기 때문이다. 그러한 효소는 두 번째 인산이 에스테르 결합 3'을 테옥시이노신으로 절단한다. 본원에 개시된 프로브들 중 어느 것이든지 하나 또는 그 이상의 트리거 잔기를 포함할 수 있다. 트리거 잔기는 그럴 필요는 없지만 리보스 또는 데옥시리보스 부분을 포함할 수 있다. 바람직하게도 절단제는 트리거 잔기의 부재시에 실질적으로 핵산을 절단하지 않지만 동일한 조건하에서 트리거 잔기를 함유하는 핵산에 대해 상당한 절단 활성을 나타내는 제제이며, 상기 조건은 핵산이 절단제에 더 민감하게 되도록 변형하는 제제의 존재를 포함할 수 있다. 예를 들어 바람직하게도 만약 절단제가 길이와 조성은 동일하되 그것들 중 하나가 트리거 잔기를 함유하고 다른 것은 트리거 잔기를 함유하지 않는 핵산을 함유하는 조성물에 존재할 때, 트리거 잔기를 함유하는 핵산이 절단될 가능성은 최소한 10; 25; 50; 100; 250; 1000; 2500; 5000; 10,000; 25,000; 50,000; 100,000; 250,000; 500,000; 1,000,000 또는 그 이상이며, 이것은 트리거 잔기를 함유하는 핵산이 절단될 가능성만큼 크며, 예컨대 트리거 잔기를 함유하지 않지만 그 외에는 동일한 핵산이 절단될 가능성에 대한 트리거 잔기를 함유하는 핵산이 절단될 가능성의 비율은 1 내지 10^6 , 또는 이 범위에 속한 어떠한 정수이다. 이 비율은 특정 핵산 및 트리거 잔기의 위치와 뉴클레오티드 함량에 따라 달라질 수 있음이 인지될 것이다.
- <105> 바람직하게도 만약 트리거 잔기를 함유하는 핵산이 절단제에 의한 절단에 대해 핵산이 민감하게 되도록 변형되

어야 할 필요가 있다면, 그러한 변형은 적당한 변형제(들)의 존재하에 쉽게 일어날 수 있는데, 예를 들어 변형은 타당한 수율로 및 타당한 시간 주기에 일어난다. 예를 들어, 본 발명의 어떤 실시예에서 트리거 잔기를 함유하는 핵산의 최소한 50%, 최소한 60%, 최소한 70%, 바람직하게는 최소한 80%, 최소한 90% 또는 더욱 바람직하게는 최소한 95%가 예컨대 24시간, 바람직하게는 12시간 내에, 보다 바람직하게는 1분 내지 4시간 이내에 변형된다.

<106> 다양한 적당한 트리거 잔기와 상응하는 절단제가 본원에서 예시된다. 본원에서 설명된 것들과 유사한 활성을 가지는 모든 트리거 잔기와 절단제가 사용될 수 있다. 당업자라면 특정한 트리거 잔기와 절단제 조합이 본 발명에 사용하기에 적당한지의 여부, 예컨대 트리거 잔기를 함유하는 핵산에 대한 절단 효율과 속도, 절단제의 선택성 등이 본 발명의 방법에 사용하기에 적당한지의 여부를 결정할 수 있을 것이다. "트리거 잔기"는 상기에서 주지된 바와 같이, 트리거 잔기가 절단에 대한 증가된 민감성을 부여하는 능력이 일반적으로, 트리거 잔기가 발견되는 특정 서열 맥락에 상당히 의존하지 않는다는 점에서, 그 맥락이 변형 및/또는 절단에 대한 민감성에 어느 정도 영향을 미칠 수 있을지라도, 단순히 제한 효소 부위를 형성하는 뉴클레오티드와는 구별된다는 것이 주지된다. 물론 주변의 뉴클레오티드에 따라 트리거 잔기는 제한 부위의 일부를 형성할 수 있다. 그러므로 대부분 경우 제한 효소이며 비-서열 특이적 절단 능력을 가진 효소의 사용이 배제되지 않는다 할지라도 절단제는 제한 효소가 아니다.

<107> 본원에서 사용되는 "보편적인 염기(universal base)"는 일반적으로 자연 발생 핵산에서 발견되는 하나 이상의 염기와 "쌍"을 이룰 수 있고, 따라서 이중나선의 그러한 자연 발생 염기를 대체할 수 있는 염기이다. 염기는 각각의 자연 발생 염기와 쌍을 이룰 필요는 없다. 예를 들어 어떤 염기는 오직 또는 선택적으로 퓨린과, 또는 오직 또는 선택적으로 피리미딘과 쌍을 이룬다. 어떤 바람직한 보편적 염기 (전체적으로 보편적 염기)는 자연 발생 핵산에서 일반적으로 발견되는 염기 중 어느 것과도 쌍을 이룰 수 있으며, 따라서 이중나선의 이들 염기 중 어느 것이라도 대체할 수 있다. 염기는 각각의 자연 발생 염기와 동등하게 쌍을 이룰 필요는 없다. 만약 프로브 혼합물이 모든 자연 발생 뉴클레오티드와 쌍을 이루지 않는 보편적 염기를 포함하는 (하나 또는 그 이상의 위치에서) 프로브를 함유한다면, 특정 프로브의 그 위치에서 둘 또는 그 이상의 보편적 염기를 활용함으로써 최소한 하나의 보편적 염기는 A와 쌍을 이루고, 최소한 하나의 보편적 염기는 G와 쌍을 이루며, 최소한 하나의 보편적 염기는 C와 쌍을 이루고, 최소한 하나의 보편적 염기는 T와 쌍을 이루는 것이 바람직할 것이다.

<108> 당해 기술분야에는 많은 보편적 염기들이 공지되어 있는데, 예를 들면 그것들에 한정되는 것은 아니지만 히포크산틴, 3-니트로피롤, 4-니트로인돌, 5-니트로인돌, 4-니트로벤즈이미다졸, 5-니트로인다졸, 8-아자-7-데아자아데닌, 6H,8H-3,4-디히드로피리미도[4,5-c][1,2]옥사진-7-온 (P. Kong Thoo Lin. and D.M. Brown, Nucleic Acids Res., 1989, 17, 10373-10383), 2-아미노-6-메톡시아미노퓨린 (D.M. Brown and P. Kong Thoo Lin, Carbohydrate Research, 1991, 216, 129-139) 등이 있다. 히포크산틴은 하나의 바람직한 완전히 보편적인 염기이다. 히포크산틴을 포함하는 뉴클레오시드로는 그것들에 한정되는 것은 아니지만 이노신, 이소이노신, 2'-데옥시이노신, 및 7-데아자-2'-데옥시이노신, 2-아자-2'-데옥시이노신이 있다.

<109> 추가의 보편적 염기는 예를 들면 관련 문헌에 기재되어 있는 바와 같이 당해 기술분야에 공지되어 있다 (Loakes, D. and Brown, D. M., Nucl. Acids Res. 22:4039-4043, 1994; Ohtsuka, E. et al., J. Biol. Chem. 260(5):2605-2608, 1985; Lin, P.K.T. and Brown, D. M., Nucleic Acids Res. 20(19):5149-5152, 1992; Nichols, R. et al., Nature 369(6480):492-493, 1994; Rahmon, M. S. and Humayun, N. Z., Mutation Research 377 (2):263-8, 1997; Berger, M., et al., Nucleic Acids Research, 28(15):2911-2914, 2000; Amosova, O., et al., Nucleic Acids Res. 25 (10):1930-1934, 1997; and Loakes, D., Nucleic Acids Res. 29(12):2437-47, 2001). 보편적 염기는 반드시 그럴 필요는 없지만, 반대로 위치한 염기와 수소 결합을 형성할 수 있다. 보편적 염기는 왓슨-크릭 또는 비-왓슨-크릭 상호작용 (예컨대 후그스틴(Hoogsteen) 상호작용)을 통해 수소 결합을 형성할 수 있다.

<110> 본 발명의 어떤 실시예에서는 보편적 염기를 포함하는 올리고뉴클레오티드 프로브를 사용하기보다는 비염기성 잔기를 포함하는 올리고뉴클레오티드 프로브가 사용된다. 비염기성 잔기는 4가지 자연 발생 뉴클레오티드 중 어떠한 것의 반대 위치를 차지할 수 있고, 따라서 보편적 염기를 포함하는 뉴클레오티드와 동일한 기능을 수행할 수 있다. 본 발명의 어떤 실시예에서 비염기성 잔기에 인접한 연쇄는 AP 엔도뉴클레아제에 의해 절단되지만, 비염기성 잔기는 또한 본원에서 설명된 바와 같이 다른 잘라지기 쉬운 연쇄 (예컨대 포스포로티올레이트)가 존재하고 다른 절단 시약이 사용되는 실시예에서도 사용된다 (즉 보편적 염기의 기능을 수행한다).

<111> A. 연장, 결합, 및 절단의 연속적인 사이클에 의한 서열화

- <112> 발명의 한 측면의 전체적인 도식은 도 1A에 도표로 도시되며, 일반적으로 미국 특허 5,740,341호 및 6,306,597호 (둘 다 Macevicz)에서 교시되는 방법을 닮았다. 편리를 목적으로, 이들 특허는 본원에서 "Macevicz"로서 포괄적으로 언급될 것이다. 특허 Macevicz는 폴리뉴클레오타이드의 뉴클레오타이드의 서열을 확인하는 방법을 교시하는데, 그 방법은 (a) 연장된 이중나선을 형성하기 위하여 폴리뉴클레오타이드를 따라 거기에 올리고뉴클레오타이드를 결합하여 개시 올리고뉴클레오타이드를 연장하는 단계; (b) 폴리뉴클레오타이드의 하나 또는 그 이상의 뉴클레오타이드를 확인하는 단계; 및 (c) 단계 (a)와 (b)를 뉴클레오타이드의 서열이 결정될 때까지 반복하는 단계로 이루어진다.
- <113> Macevicz는 추가로 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열을 결정하는 방법을 교시하는데, 그 방법은 다음의 단계들로 이루어진다: (a) 주형 폴리뉴클레오타이드에 혼성화된 개시 올리고뉴클레오타이드를 포함하는 프로브-주형 이중나선을 제공하는 단계, 상기 프로브는 연장가능한 프로브 말단을 가진다; (b) 상기 연장가능한 프로브 말단에 연장 올리고뉴클레오타이드 프로브를 결합하여 연장된 올리고뉴클레오타이드 프로브를 함유하는 연장된 이중나선을 형성하는 단계; (c) 연장된 이중나선에서 (1) 방금 결합된 연장 프로브에 상보하거나 또는 (2) 연장된 올리고뉴클레오타이드 프로브의 바로 아래쪽에 있는 주형 폴리뉴클레오타이드의 뉴클레오타이드 잔기인 주형 폴리뉴클레오타이드의 최소한 하나의 뉴클레오타이드를 확인하는 단계; (d) 만약 연장가능한 프로브 말단이 전에 존재하지 않는 것이라면 생성된 말단이 마지막 연장 프로브가 결합된 말단과는 상이하도록 연장된 프로브 상에 연장가능한 프로브 말단을 생성하는 단계; 및 (e) 단계 (b), (c) 및 (d)를 주형 폴리뉴클레오타이드의 뉴클레오타이드 서열이 결정될 때까지 반복하는 단계. 이들 방법의 어떤 실시예에서 각각의 연장 프로브는 개시 올리고뉴클레오타이드 프로브에서 멀리 떨어져 있는 말단에 사슬-종결 부분을 갖는다. 어떤 실시예에서 재생 단계는 연장된 올리고뉴클레오타이드 프로브의 화학적으로 잘라지기 쉬운 뉴클레오시드간 연쇄를 절단하는 것을 포함한다.
- <114> 도 1A를 참조로 하면, 미지의 서열 및 결합 영역(40)을 가지고 있는 폴리뉴클레오타이드 영역(50)을 포함하는 폴리뉴클레오타이드 주형(20)이 지지체(10)에 부착된다. 결합 영역(40)의 말단 말단에 있는 뉴클레오타이드(41)과, 폴리뉴클레오타이드 영역(50)의 근접한 말단에 있는 뉴클레오타이드(51)은 서로 인접한다. 개시 올리고뉴클레오타이드(30)은 결합 영역(40)의 위치에서 이중나선을 형성하기 위하여 결합 영역(40)과 혼성화하도록 제공된다. 개시 올리고뉴클레오타이드(30)은 또한 본원에서 "프라이머"로서 언급되며, 결합 영역(40)은 "프라이머 결합 영역"으로서 언급될 수 있다. 이중나선은 반드시 그럴 필요는 없지만 바람직하게는 매치된 이중나선일 수 있다. 개시 올리고뉴클레오타이드는 연장가능한 말단(31)을 갖는다. 도 1A에서 개시 올리고뉴클레오타이드는 연장가능한 말단(31)이 반대 뉴클레오타이드(41)에 위치하도록 결합 영역에 결합한다. 그러나 개시 올리고뉴클레오타이드는 아래에서 추가로 논의되는 바와 같이, 결합 영역의 그 밖의 곳에서도 결합할 수 있다. 길이가 N인 연장 올리고뉴클레오타이드 프로브(60)는 개시 올리고뉴클레오타이드에 인접한 주형에 혼성화한다. 연장 올리고뉴클레오타이드 프로브의 말단 뉴클레오타이드(61)은 연장가능한 말단(31)에 결합된다.
- <115> 말단 뉴클레오타이드(61)은 폴리뉴클레오타이드 영역(50)의 첫 번째 미지의 뉴클레오타이드에 상보한다. 그러므로 말단 뉴클레오타이드(61)의 정체는 뉴클레오타이드(51)의 정체를 설정한다. 바람직하게도 뉴클레오타이드(51)은 말단 뉴클레오타이드(61)로서 A, G, C, 또는 T를 가지는 것으로 알려져 있는 연장 프로브와 결합된 표지(도시되지 않음)를 검출함으로써 확인된다. 표지는 검출 후에 제거된다. 도 2는 상이한 3' 말단 뉴클레오타이드를 가지는 연장 프로브에 대해 상이한 표지, 예컨대 상이한 색의 형광발색단을 지정하기 위한 도식을 도시한다.
- <116> 결합 및 검출에 이어서, 연장가능한 프로브 말단은 만약 프로브(60)가 이미 그러한 말단을 갖지 않았다면 연장 프로브(60) 상에서 생성된다. 바람직하게는 길이가 N인 두 번째 연장 프로브(70)는 연장 프로브(60)에 인접한 주형에 어닐(anneal)되고 프로브(60)의 연장가능한 말단에 결합된다. 연장 프로브(70)의 말단 뉴클레오타이드(71)의 정체는 폴리뉴클레오타이드(50)의 반대로 위치한 뉴클레오타이드(52)의 정체를 설정한다. 그러므로 말단 뉴클레오타이드(71)은 연장 프로브의 "서열 결정 부분"을 구성하고, 그것은 그것의 혼성화 특이성이 그것으로부터 주형의 하나 또는 그 이상의 뉴클레오타이드의 정체를 결정할 수 있는 기초로서 사용된다는 것을 의미한다. 전형적으로 연장 프로브의 추가의 뉴클레오타이드는 주형과 혼성화하겠지만, 그것의 정체가 특정한 표지와 관련되어 있는 프로브의 뉴클레오타이드들만이 주형의 뉴클레오타이드를 확인하는 데 사용된다.
- <117> 본 발명의 바람직한 실시예에서 연장가능한 말단의 생성은 아래에서 더 상세하게 설명되는 바와 같이 뉴클레오시드간 연쇄의 절단을 포함한다. 바람직하게도 절단은 또한 표지를 제거한다. 절단은 연장 프로브로부터 많은 뉴클레오타이드 M을 제거한다 (도시되지 않음). 그러므로 이중나선은 각 사이클에서 N-M 뉴클레오타이드에 의해 연장되며, 주형의 N-M의 간격에 위치한 뉴클레오타이드가 확인된다. 주어진 주형의 다수의 복사물은 전형적으로 단일 지지체에 부착될 것이며, 서열화 반응은 이들 주형에 대해 동시에 수행될 것임이 인지되어야 한다.

- <118> Macevicz는 올리고뉴클레오타이드 프로브가 일반적으로 다음 연장 사이클의 연장된 이중나선을 생성하기 위해 개시 올리고뉴클레오타이드 또는 연장된 이중나선에 결합될 수 있어야 하고; 결합은 프로브가 결합 전에 주형과 이중나선을 형성해야 한다는 점에서 주형-구동되어야 하며; 프로브는 단일 연장 사이클에서 동일한 주형에 대한 다중 프로브 결합을 방지하기 위해 차단 부분을 포함하여야 하고; 프로브는 결합 후에 연장가능한 말단을 재생하기 위해 처리되거나 변형될 수 있어야 하며; 프로브는 연속적인 결합 후에 주형에 관련한 서열 정보의 획득을 허용하는 신호화 부분 (즉 검출가능한 부분)을 포함하여야 한다는 것을 교시한다.
- <119> Macevicz는 특정한 적당한 개시 올리고뉴클레오타이드, 연장 올리고뉴클레오타이드 프로브, 주형, 결합 부위, 및 그러한 성분들을 합성하고, 디자인하고, 제조하고, 또는 얻기 위한 다양한 방법의 특징들을 교시한다. Macevicz는 또한 특정한 적당한 결합 효소, 결합 조건, 및 다양한 적당한 표지를 교시한다. Macevicz는 또한 새롭게 결합된 연장 프로브에 표시된 사슬-종결 뉴클레오타이드를 부가하기 위하여 중합효소 연장을 사용하는 또 다른 확인 방법을 교시한다. 부가된 뉴클레오타이드의 정체는 주형의 반대쪽에 위치한 뉴클레오타이드를 확인한다.
- <120> 당업자에 의해 인지되는 것과 같이, 주형, 개시 올리고뉴클레오타이드, 연장 프로브, 프라이머 등에 대한 언급은 일반적으로 단일 분자보다는 관련된 영역 내에서 실질적으로 동일한 핵산 분자의 집단 또는 풀을 의미한다. 그러므로 예를 들어 "주형"은 일반적으로 다수의 실질적으로 동일한 주형 분자를 의미하고; "프로브"는 일반적으로 실질적으로 동일한 다수의 프로브 분자를 의미한다. 하나 또는 그 이상의 위치에서 축퇴성인 프로브의 경우에, 특별한 프로브를 포함하는 프로브 분자의 서열은 축퇴성 위치에서 상이할 것, 즉 특정 프로브를 구성하는 프로브 분자의 서열은 실질적으로 비축퇴성 위치(들)에서만 동일할 것임이 인지될 것이다. 설명을 목적으로 하자면 단일 형태는 단일 분자와 실질적으로 동일한 분자들의 집단을 포함하는 것으로 이해되어야 한다. 단일 핵산 분자 (즉 한 분자)를 나타내는 것으로 의도되는 경우에는 용어 "주형 분자", "프로브 분자", "프라이머 분자" 등이 사용될 것이다. 어떤 경우에 실질적으로 동일한 핵산 분자의 집단의 여러 특성이 분명하게 나타날 것이다.
- <121> 실질적으로 동일한 핵산 분자의 집단은 화학적 합성, 세포에서의 생물학적 합성, 하나 또는 그 이상의 출발 핵산 분자로부터의 시험관 내 효소적 증폭 등을 포함한 다양한 공지된 방법 중 어느 것을 사용하여 얻거나 제조될 수 있다. 예를 들어 당해 기술분야에 잘 알려져 있는 방법들을 사용하여 관심 있는 핵산은 그것을 적당한 발현 벡터, 예컨대 DNA 또는 RNA 플라스미드에 삽입하고, 그런 다음 그 안에서 복제될 세포, 예컨대 박테리아 세포 안에 도입됨으로써 클로닝(cloning)될 수 있다. 그런 다음 관심 있는 핵산의 복사물을 함유하는 플라스미드 DNA 또는 RNA가 세포로부터 분리된다. 게놈 DNA는 바이러스, 세포 등으로부터 분리되거나 mRNA의 역 전사에 의해 제조된 cDNA는 또한 일반적으로 클로닝 또는 시험관 내 증폭의 중간 단계를 수행하는 것이 바람직하긴 하지만 그러한 중간 단계 없이 실질적으로 동일한 핵산 분자 (예컨대 그것의 서열을 결정하고자 하는 주형 폴리뉴클레오타이드)의 집단의 공급원일 수 있다.
- <122> 집단의 구성원들은 100% 동일할 필요는 없다는 것, 예컨대 특정 횟수의 "에러"가 합성 과정 동안 일어날 수 있다는 것이 인지될 것이다. 바람직하게도 최소한 50%의 집단의 구성원은 참조 핵산 분자 (즉 서열 비교를 위한 기초로서 사용된 규정된 서열의 분자)와 최소한 90%, 또는 더욱 바람직하게는 최소한 95% 동일하다. 보다 바람직하게는, 최소한 60%, 최소한 70%, 최소한 80%, 최소한 90%, 최소한 95%, 최소한 99%, 또는 그 이상의 집단의 구성원이 참조 핵산 분자와 최소한 90%, 또는 더욱 바람직하게는 최소한 95% 동일하거나, 또는 더 더욱 바람직하게는 99% 동일하다. 바람직하게도 최소한 95% 또는 더욱 바람직하게는 최소한 99%의 집단의 구성원의 참조 핵산 분자에 대한 % 동일성은 최소한 98%, 99%, 99.9% 또는 그 이상이다. % 동일성은 두 개의 임의로 배열된 서열을 비교하고, 두 서열 모두에서 동일한 핵산 염기 (A, T, C, G, U, 또는 I)가 발생하는 위치의 수를 결정하여 매치된 위치들의 수를 산출하고, 매치된 위치의 수를 위치의 총 수로 나누고, 그 결과를 100으로 곱하여 서열 동일성의 백분율을 구함으로써 계산될 수 있다. 어떤 경우에 핵산 분자, 예컨대 주형, 프로브, 프라이머 등은 또한 주형, 프로브, 또는 프라이머 기능을 수행하지 않는 부분을 함유하는 더 큰 핵산 분자의 일부분일 수 있다는 것이 인지될 것이다. 그 경우에 집단의 개별적 구성원은 그 부분과 관련하여 실질적으로 동일할 필요는 없다.
- <123> Macevicz는 주형이 비드와 같은 지지체에 부착되어 있고 연장이 도 1A에서 도시된 것처럼 지지체의 말단에 위치한 주형의 말단을 향하여 진행되는 방법을 교시한다. 그러므로 결합 영역은 미지의 서열보다 지지체에 더 가깝게 위치하고, 연장된 이중나선은 지지체로부터 멀어지는 방향으로 성장한다. 그러나 본 발명자들은 이 방법이 유익하게도 결합 영역이 지지체 말단인 주형의 말단에 위치하고, 연장이 지지체를 향하여 안쪽으로 진행되는 대체 접근법을 사용하여 실시될 수 있다는 것을 뜻밖에도 발견하였다. 이 실시예는 도 1B에 도시되는데, 도면에서 다양한 구성요소들이 도 1A에서와 같은 번호로 표시된다. 본 발명자들은 주형의 말단으로부터 지지체를 향하여

"안쪽으로" 서열화하는 것이 월등한 결과를 산출하는 것으로 판단하였다. 특히 주형의 말단으로부터 비드와 같은 지지체를 향하여 서열화하는 것은 지지체로부터 바깥쪽으로 서열화하는 것보다 더 높은 결찰 효율을 초래한다.

<124> Macevicz에 의해 추가로 교시되는 것으로서, 바람직하게도 올리고뉴클레오타이드 프로브는 예정된 길이의 모든 가능한 서열의 올리고뉴클레오타이드를 포함하는 혼합물로서 주형에 적용된다. 예를 들어 구조가 $NNNNN((N)_k)$ 로 표시할 수도 있다, 여기서 k 는 6)이고 길이가 6 뉴클레오타이드인 가능한 서열을 모두 함유하는 프로브의 혼합물은 4^6 (4096)개의 프로브 중을 함유할 것이다. 일반적으로 프로브의 구조는 $X(N)_kN^*$ 이다 (N 은 어떠한 뉴클레오타이드를 나타내고, k 는 1과 100 사이의 수이며, $*$ 는 표지를 나타내고, X 는 그것의 정체가 표지에 상응하는 뉴클레오타이드를 나타낸다). 어떤 실시예에서 k 는 1과 100 사이, 1과 50 사이, 1과 30 사이, 1과 20 사이, 예컨대 4와 10 사이일 수 있다. 하나 또는 그 이상의 뉴클레오타이드는 보편적 염기를 포함할 수 있다. 일반적으로 프로브는 N 에 의해 표시되는 위치에서 4-배 축퇴성이거나 N 에 의해 표시되는 하나 또는 그 이상의 위치에서 축퇴성-감소 뉴클레오타이드를 포함한다. 필요하다면 혼합물은 상보하는 서열과 완벽하게 매치된 이중나선이 유사한 안정성 또는 유리 결합 에너지를 가지는 프로브의 하위세트 (긴장 부류)로 나누어질 수 있다. 하위세트는 Macevicz에 의해 교시된 바와 같이 별도의 혼성화 반응에 사용될 수 있다.

<125> 프로브 혼합물의 복잡성 (즉 상이한 서열의 수)은 소위 축퇴성-감소 뉴클레오타이드 또는 뉴클레오타이드 유사체를 사용하는 것을 포함한 많은 방법에 의해 감소될 수 있다. 예를 들어 8가지 뉴클레오타이드의 모든 가능한 서열을 함유하는 프로브의 라이브러리는 4^8 개의 프로브를 함유할 것이다. 프로브의 수는 4^6 으로 감소될 수 있는 한편, 2개의 위치에서 보편적 염기를 사용함으로써 옥타머(8량체) 라이브러리의 다양한 바람직한 특징, 예컨대 길이를 보유한다. 본 발명은 상기에서 언급된 또는 상기에서 언급된 참고문헌에서 설명된 보편적 염기들 중 어떠한 것든지 사용되는 것을 포함한다.

<126> 실시예에 따르면, 연장된 이중나선 또는 개시 올리고뉴클레오타이드는 아래에서 한층 더 설명되는 바와 같이, 올리고뉴클레오타이드 프로브에 의해 5'→3' 방향으로 또는 3'→5' 방향으로 연장될 수 있다. 일반적으로 올리고뉴클레오타이드 프로브는 주형과 완벽하게 매치된 이중나선을 형성할 필요는 없지만 그러한 결합이 바람직할 수 있다. 주형의 단일 뉴클레오타이드가 각 연장 사이클에서 확인되는 실시예에서 완벽한 염기 쌍을 이루는 것만이 특정 뉴클레오타이드를 확인하는 데 필요하다. 예를 들어 올리고뉴클레오타이드 프로브가 효소적으로 연장된 이중나선에 결찰되는 실시예에서 완벽한 염기쌍을 이루는 것, 즉 적절한 왓슨-크릭 염기쌍이 결찰된 프로브의 말단 뉴클레오타이드와 주형에 있는 그것의 상보물 사이에 필요하다. 일반적으로 그러한 실시예에서 프로브의 나머지 뉴클레오타이드는 주형을 따라 예정된 부위에서, 또는 염기 번호에서 다음 결찰이 일어날 것을 보장하는 "스페이서"로서 작용한다. 즉 그것들의 쌍 형성, 또는 그것의 결핍은 추가의 서열 정보를 제공하지는 않는다. 마찬가지로 염기 확인을 위해 중합효소 연장에 좌우되는 실시예에서 프로브는 주로 스페이서로서 작용하고, 따라서 주형에 대한 특이적 혼성화는 중요하지 않다.

<127> 상기에서 설명된 방법은 서열의 부분적인 결정, 즉 주형에서 상호 간에 공간적으로 떨어져 있는 개별적인 뉴클레오타이드의 확인을 가능하게 한다. 본 발명의 바람직한 실시예에서, 보다 완전한 정보를 모으기 위해서는 각 반응이 상이한 개시 올리고뉴클레오타이드 i 를 사용하는 다수의 반응이 수행된다. 개시 올리고뉴클레오타이드 i 는 결합 영역의 상이한 부분에 결합한다. 바람직하게도 개시 올리고뉴클레오타이드는 상이한 개시 올리고뉴클레오타이드의 연장가능한 말단이 결합 영역에 혼성화될 때 상호 간에 1 뉴클레오타이드만큼 상쇄되는 위치에서 결합한다. 예를 들어 도 3에서 알 수 있는 바와 같이, 서열화 반응 $1...N$ 이 수행된다. 개시 올리고뉴클레오타이드 $i_1...i_n$ 은 동일한 길이를 가지며, 그것들의 말단 뉴클레오타이드 (31, 32, 33 등)이 결합 영역(40)의 연속적인 인접한 위치 (41, 42, 43 등)에 혼성화하도록 결합한다. 그러써 연장 프로브 $e_1...e_n$ 은 주형의 연속적인 인접한 영역에서 결합하고, 개시 올리고뉴클레오타이드의 연장가능한 말단에 결찰된다. i_n 에 결찰된 프로브 e_n 의 말단 뉴클레오타이드 (61)는 폴리뉴클레오타이드 영역(50)의 뉴클레오타이드(55), 즉 주형의 첫 번째 미지 폴리뉴클레오타이드에 상보한다. 연장, 결찰, 및 검출의 두 번째 사이클에서 프로브 e_{12} 의 말단 뉴클레오타이드(71)는 폴리뉴클레오타이드 영역(50)의 뉴클레오타이드(56), 즉 미지 서열의 두 번째 뉴클레오타이드에 상보한다. 마찬가지로, 개시 올리고뉴클레오타이드 i_2, i_3, i_4 등으로 개시된 이중나선에 결찰된 연장 프로브의 말단 뉴클레오타이드들은 미지 서열(50)의 세 번째, 네 번째, 및 다섯 번째 뉴클레오타이드와 상보할 것이다. 개시 올리고뉴클레오타이드는 점진적으로 폴리뉴클레오타이드 영역(50)에 가깝게 가기보다는 그것으로부터 점점 더 멀리 떨어진 영역에 결합할 수 있다는 것이 인지될 것이다.

- <128> 연장 프로브의 비-말단 뉴클레오타이드의 스페이서 기능은 개시 올리고뉴클레오타이드가 상응하는 대다수 사이클이 어떠한 주어진 주형에 대해 수행되어야 할 필요없이 결합하는 위치로부터 상당히 제거되는 주형의 위치에서 서열 정보를 획득하는 것을 가능하게 한다. 예를 들어 길이가 N인 프로브의 연속적인 결합 사이클과, 계속해서, 연장 프로브로부터 단일한 말단 뉴클레오타이드를 제거하기 위한 절단에 의해, N-1 뉴클레오타이드만큼 간격이 있는 뉴클레오타이드가 연속적인 회기(round)에서 확인된다. 예를 들어 주형에서 위치 1, N, 2N-1, 3N-2, 4N-3 및 5N-4에 있는 뉴클레오타이드는 6 사이클에서 확인될 수 있는데, 그때 주형의 위치 1에 있는 뉴클레오타이드는 주형에 대해 개시 올리고뉴클레오타이드가 결합함으로써 형성된 이중나선에 있는 연장가능한 프로브 말단에 결합되는 뉴클레오타이드 반대쪽에 있는 뉴클레오타이드이다. 마찬가지로 만약 절단으로 인해 2개의 뉴클레오타이드가 길이가 N인 연장 프로브로부터 제거된다면, 상호간에 N-2 뉴클레오타이드만큼 떨어져 있는 위치에 있는 뉴클레오타이드가 연속적인 회기에서 확인된다. 예를 들어 주형에서 위치 1, N-1, 2N-3, 3N-5, 4N-7에 있는 뉴클레오타이드는 6 사이클로 확인될 수 있다. 그러므로 만약 프로브가 길이가 8 뉴클레오타이드이고, 각 사이클에서 2개의 뉴클레오타이드가 제거된다면, 위치 1, 7, 13, 19, 및 25에 있는 뉴클레오타이드가 확인된다. 그러므로 주형의 첫 번째 뉴클레오타이드로부터 거리 X만큼 떨어진 뉴클레오타이드를 확인하기 위해 필요한 사이클의 수는 X와 비슷하기보다는 X/M와 비슷하며, 이때 M은 절단 후에도 남아있는 연장 프로브의 길이이다.
- <129> 예를 들어 도 3B에 도시된 개략도는 주형의 매 6번째 염기를 판독하기 위해 디자인된 연장 프로브를 사용하여 연장, 결합, 및 절단 방법을 사용한 순(net) 결과를 도시한다. 결합 영역 내에서 상쇄되는 위치에 결합하는 6개의 개시 올리고뉴클레오타이드를 사용하여 연속적으로 스트리핑(striping)하고 서열화한 후 그 결과들을 조합함으로써, 모든 주형의 염기들은 규정된 길이에 걸쳐 명료하게 된다. 예를 들어 만약 6 반응의 각각에 대하여 10회의 연속적인 결합이 수행된다면, 그 결과의 판독 길이는 60개의 연속적인 염기쌍일 것이며, 한편 15회의 연속적인 결합이 각 반응에 대해 수행된다면, 그 결과의 판독 길이는 90개의 연속적인 염기쌍일 것이다.
- <130> 어떠한 이론에 구속되는 것을 바라지는 않지만, 본 발명자들은 이 접근법과는 대조적으로, 합성 방법에 의한 대부분의 연속적인 서열화가 궁극적으로는 긴 판독 길이의 가능성을 제한하는 여러 축적과 대립한다고 제시한다. 본원에 설명되는 어떤 방법들의 유익한 특징은 그것들이 매 n번째 염기의 확인을 가능하게 하고 (프로브의 절단 가능한 부분의 위치에 따라), 그로써 주어진 수의 사이클(y) 후에 당업자는 $n^*y-(n-1)^{th}$ 염기 (예컨대 절단 부위의 3'쪽에 있는 6 염기의 프로브를 사용하여 15 사이클 후에 전술한 실시예의 71번째 염기, 또는 20 사이클 후에 115번째 염기)에 도달할 수 있다. n-1, n-2 등의 위치에서 개시 올리고뉴클레오타이드를 "재설정"하는 능력은 주어진 판독 길이에 대한 연속적인 여러 축적을 크게 최소화하는데 (단계의 제거(dephasing) 또는 축소를 통하여), 그 이유는 연장된 프로브를 주형으로부터 스트리핑하고 새로운 개시 올리고뉴클레오타이드를 혼성화하는 과정이 바탕값 신호(background signal)를 0으로 효과적으로 재설정하기 때문이다. 예를 들어 합성에 의한 중합효소 기반 서열화와 본원에서 설명된 결합 기반 접근법을 비교하면, 각 연장 사이클에서 노이즈에 대한 신호 비율이 99:1이라면, 중합효소 기반 접근법에 대한 100 사이클 후 비율은 37:63이 될 것이고, 결합 효소 기반 방법에 대해서는 85:15가 될 것이다. 결합 효소 기반 방법에 대한 순 결과는 판독 길이에 있어서 중합효소 기반 접근법을 증가하여 크게 증가한다.
- <131> 주형의 각각의 선행하는 뉴클레오타이드에 대하여 사이클을 수행하는 데 필요하다면 요구되는 것보다 적은 수의 사이클을 사용하여 뉴클레오타이드를 확인하는 능력은 많은 이유로 중요하다. 특히 방법의 각 단계가 100% 효율로 발생할 것 같지는 않다. 예를 들어 어떤 주형은 연장 프로브에 연속적으로 결합될 수 없고; 어떤 연장 프로브는 절단될 수 없는 식이다. 그러므로 각 사이클에서 주형의 상이한 복사물에 대해 일어나는 반응은 점진적으로 단계가 제거되며, 그것으로부터 유용하고 정확한 정보가 얻어질 수 있는 주형의 수는 감소한다. 그러므로 개시 올리고뉴클레오타이드의 연장가능한 말단으로부터 멀리 떨어져 있는 적은 수 이상의 위치에 위치한 뉴클레오타이드를 판독하는 데 필요한 사이클의 수를 최소화하는 것이 특히 바람직하다. 그러나 연장 프로브의 길이를 증가시키는 것은 프로브 혼합물의 더 큰 복잡성을 초래하며, 그것은 각각의 개별적인 프로브 서열의 효과적인 농도를 감소시킨다. 본원에서 설명되는 바와 같이, 축퇴성-감소 뉴클레오타이드는 복잡성을 감소시키기 위해 사용될 수 있지만 혼성화 길이의 감소 및/또는 결합 효율의 저하를 초래할 수 있다. 본 발명자들은 결과를 최적화하기 위하여 이들 경합하는 인자들의 균형을 맞추는 필요가 있음을 깨달았다. 그러므로 본 발명의 바람직한 실시예에서 길이가 8 뉴클레오타이드인 연장 프로브를 사용하고, 선택된 위치에서 축퇴성-감소 뉴클레오타이드를 사용하였다. 또한 본 발명자들은 적절한 잘라지기 쉬운 연쇄와 절단 조건 및 절단 단계의 효율 (즉 각 절단 단계에서 연속적으로 절단되는 연쇄의 백분율)을 최적화하기 위한 시간, 및 적절한 연쇄에 대한 그것의 특이성을 선택하는 것의 중요성을 인지하였다.

<132> B. 올리고뉴클레오타이드 연장 프로브 디자인

<133> Macevicz는 축퇴성-감소 뉴클레오시드 유사체가 올리고뉴클레오타이드 연장 프로브에 사용될 수 있다고 언급한 한편으로, 그는 그 위치에서 연장 프로브에 그러한 잔기를 포함하는 잔기를 포함하기 위해 특히 바람직한 구체적인 위치를 교시하지 않았고, 축퇴성-감소 뉴클레오시드를 통합하는 특별한 프로브 구조 (즉 서열)을 교시하지 않았다. 본 발명자들은 올리고뉴클레오타이드 연장 프로브의 특정 위치에서 및 특정 수로 축퇴성-감소 뉴클레오시드 (예컨대 보편적 염기를 포함하는 뉴클레오시드)를 활용하는 것이 특히 유익할 것임을 인지하였다. 예를 들어 본 발명의 어떤 실시예에서 위치 6 또는 그 이상의 숫자 (X로부터 계수하여)의 위치에 있는 모든 또는 대부분의 뉴클레오타이드는 보편적 염기를 포함한다. 예를 들어 위치 6 또는 그 이상의 숫자에 있는 뉴클레오타이드의 최소한 50%, 최소한 60%, 최소한 70%, 최소한 80%, 최소한 90%, 또는 최소한 100%가 보편적 염기를 포함할 수 있다. 뉴클레오타이드는 동일한 보편적 염기를 포함할 필요는 없다. 본 발명의 어떤 실시예에서 히포크산틴 및/또는 니트로-인돌(nitro-indole)이 보편적 염기로서 사용된다. 예를 들어 이노신과 같은 뉴클레오시드가 사용될 수 있다.

<134> 본 발명자들은 길이가 6 뉴클레오타이드 이상인 연장 프로브를 사용할 때 월등한 결과를 얻을 수 있으며, 그때에 연장가능한 프로브 말단에 결합될 뉴클레오타이드로부터 계수하여 위치 6 또는 그 이상의 숫자의 위치에 있는 하나 또는 그 이상의 뉴클레오타이드가 축퇴성-감소 뉴클레오타이드임을, 예컨대 보편적 염기를 포함하는 것을 (즉 만약 가장 근접한 뉴클레오타이드가 위치 1이라 간주한다면, 위치 6 또는 그 이상의 위치에 있는 하나 또는 그 이상의 뉴클레오타이드는 보편적 염기를 포함한다), 예를 들어 옥타머 프로브의 경우 위치 6 또는 그 이상의 위치에 있는 1, 2, 또는 3개의 뉴클레오타이드가 보편적 염기를 포함하는 것을 인지하였다. 예를 들어 3'→5' 방향으로의 서열화에 대해 구조가 3'-XNNNNsINI-5' (X와 N은 어떠한 뉴클레오타이드를 나타내고, "s"는 잘라지기 쉬운 연쇄를 나타낸다)인 프로브가 사용될 수 있고, 따라서 절단은 3' 말단로부터 계수하여 다섯 번째와 여섯 번째 잔기 사이에서 일어나며, 잘라지기 쉬운 연쇄와 5' 말단 사이에 있는 최소한 하나의 잔기는 바람직하게는 X의 정체에 상응하는 표지를 가진다. 다른 디자인은 3'-XNNNNsNII-5'이다. 또 다른 프로브 디자인은 3'-XNNNNsIII-5'이다. 이 디자인은 가장 적절한 복잡성을 가지는 1024개의 상이한 종의 프로브 혼합물을 생성하고, 상당한 아데닐화 생성물의 형성을 방지하기에 충분히 길며 (실시예 1 참조), 절단 후에 남아있는 결과의 연장 생성물이 변형되지 않은 DNA로 구성될 수 있다는 장점을 갖는다. 한 가지 결점은 이 프로브가 프라이머를 한 번에 단지 5 염기씩만 연장할 수 있다는 것이다. 판독 길이가 연장 길이 곱하기 사이클의 수의 함수이기 때문에, 연장 길이에 대한 각 추가의 염기는 1×사이클 수만큼 판독 길이를 증가시킬 가능성을 갖는다 (예컨대 20 사이클이 사용되면 20 염기). 다른 프로브 디자인은 하나 또는 그 이상의 이노신 (또는 다른 보편적 염기)를 절단 후의 연장 프로브의 말단에 남김으로써 6 염기, 또는 더 긴 연장된 이중나선이 생성된다. 예를 들어 프로브 3'-XNNNNIsII-5'를 사용하면, 이중나선은 한 번에 6 염기씩 연장되며 접합부에 5' 이노신을 남긴다. 이들 각각의 디자인에서 잘라지기 쉬운 연쇄와 5' 말단 사이에 있는 최소한 하나의 잔기는 바람직하게는 X의 정체에 상응하는 표지를 갖는다. 본 발명의 어떤 구체예에서 연장가능한 프로브 말단에 결합될 뉴클레오타이드 반대쪽 말단으로부터 계수하여 프로브의 말단으로부터 세 번째 뉴클레오타이드는 보편적 염기를 포함한다 (즉 만약 말단이 위치 K라고 생각된다면 위치 K-2에 있는 뉴클레오타이드는 보편적 염기를 포함한다).

<135> 본 발명의 어떤 실시예에서 고정된 핵산 (LNA) 염기는 개시 올리고뉴클레오타이드 프로브, 연장 프로브, 또는 둘 다에서 하나 또는 그 이상의 위치에서 사용된다. 고정된 핵산은 예를 들면 미국 특허 6,268,490호와 문헌에 설명되어 있다_(Koshkin, AA, et al., Tetrahedron, 54:3607-3630, 1998; Singh, SK, et al., Chem. Comm., 4:455-456, 1998). LNA는 표준 포스포라미다이트 화학을 사용하여 자동 DNA 합성기에 의해 합성될 수 있고, 또한 자연 발생 뉴클레오타이드 및/또는 뉴클레오타이드 유사체를 함유하는 올리고뉴클레오타이드 안에 통합될 수 있다. 그것들은 또한 아래에서 설명되는 것과 같은 표지와 함께 합성될 수 있다.

<136> C. 주형 및 지지체 제조 방법

<137> Macevicz는 다수의 실질적으로 동일한 주형 분자를 포함하는 주형이 종래의 중합효소 사슬 반응 (PCR) 방법에서와 같이 먼저 예컨대 튜브 또는 다른 용기 안에서의 증폭에 의해 합성되는 방법을 교시한다. Macevicz는 증폭된 주형 분자가 합성 후에 바람직하게는 자성 마이크로입자 (예컨대 비드)와 같은 지지체에 부착되는 것을 교시한다.

<138> 본 발명자들은 서열화될 주형이 바람직하게도 지지체 상에서 또는 지지체 자체 안에서, 예컨대 지지체, 예컨대 마이크로입자 또는 다양한 반-고체 지지체 물질, 예를 들면 한 쌍의 증폭 프라이머 중 하나가 PCR 반응을 수행하기 전에 부착되는 겔 매트릭스를 사용함으로써 합성될 수 있다는 것을 발견하였다. 이 접근법은 주형 분자를 합성 후에 지지체에 부착시키는 별도의 단계에 대한 필요성을 피할 수 있다. 그러므로 상이한 서열의 다수의 주

형 종은 편리하게도 동시에 증폭될 수 있다. 예를 들어 아래에서 설명되는 방법에 따르면, 마이크로입자 상에서의 합성 결과 개별적인 마이크로입자의 집단이 생성되고, 개별적인 마이크로입자는 각각 거기에 부착된 특정 주형 분자 (또는 그것의 상보물)의 다수의 복사물이 부착되어 있고, 각 마이크로입자에 부착된 주형 분자들은 다른 마이크로입자에 부착된 주형 분자들과는 서열이 다르다. 그러므로 각각의 지지체는 그것에 주형의 클론 집단을 가지는데, 예를 들면 지지체 A는 그것에 부착된 주형 X의 다수의 복사물을 가질 것이고; 지지체 B는 그것에 부착된 주형 Y의 다수의 복사물을 가질 것이며; 지지체 C는 그것에 부착된 주형 Z의 다수의 복사물을 가질 것이다. "주형의 클론 집단", "핵산의 클론 집단" 등은 바람직하게는 관심 있는 단일 주형 분자(출발 주형)로부터 출발하는 증폭의 연속적인 회기에 의해 생성된, 실질적으로 동일한 주형 분자들의 집단을 의미한다. 실질적으로 동일한 주형 분자는 출발 주형 또는 그것의 상보물에 실질적으로 동일할 수 있다.

<139> 증폭은 전형적으로 PCR을 사용하여 수행되지만, 다른 증폭 방법도 사용될 수 있다 (하기 참조). 클론 집단의 구성원들은 반드시 100% 동일하지 않아도 된다는 것, 예를 들어 특정 수의 "에러"가 합성 과정 중에, 예컨대 증폭되는 동안 발생할 수 있다는 것이 인지되어야 할 것이다. 클론 집단의 구성원의 최소한 50%가 출발 주형 분자 (또는 그것의 상보물)에 대해 최소한 90% 동일할 수 있으며, 보다 바람직하게는 최소한 95% 동일하다. 집단의 구성원의 최소한 60%, 최소한 70%, 최소한 80%, 최소한 90%, 최소한 95%, 최소한 99%, 또는 그 이상이 출발 주형 분자 (또는 그것의 상보물)에 대해 최소한 90% 동일할 수 있으며, 보다 바람직하게는 최소한 95%, 더욱 바람직하게는 최소한 99% 동일한 것이다. 바람직하게는 출발 주형 분자 (또는 그것의 상보물)에 대한 집단의 구성원의 최소한 95%, 보다 바람직하게는 최소한 99%의 % 동일성은 최소한 98%, 99%, 99.9% 또는 그 이상이다.

<140> 증폭 프라이머는 다양한 기법들 중 어느 것을 사용하든지 지지체에 부착될 수 있다. 예를 들어 프라이머의 한 말단 (5' 말단)은 결합 쌍의 한 구성원(예컨대 비오틴)으로 기능화될 수 있고, 지지체는 결합 쌍의 다른 구성원 (예컨대 스트렙타아비딘)으로 기능화될 수 있다. 어떠한 유사한 결합 쌍이라도 사용될 수 있다. 예를 들어 규정된 서열의 핵산 태그는 지지체에 부착될 수 있고, 상보하는 핵산 태그를 가지는 프라이머는 지지체에 부착된 핵산 태그에 혼성화될 수 있다. 다양한 링커와 교차결합제가 또한 사용될 수 있다.

<141> PCR을 수행하기 위한 방법은 당해 기술분야에 잘 알려져 있고, 예를 들면 미국 특허 4,683,195호, 4,683,202호 및 4,965,188호와 문헌에 설명되어 있다 (Dieffenbach, C. and Dveksler, GS, *PCR Primer: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2003). 마이크로입자 상에서 핵산을 증폭하기 위한 방법은 당해 기술분야에 잘 알려져 있으며, 예를 들면 표준 PCR은 마이크로적정 디쉬 또는 튜브의 웰에서 프라이머가 부착된 비드 상에서 수행될 수 있다 (예컨대 비드는 실시예 12에서와 같이 제조된다). PCR이 편리한 증폭 방법이지만 당해 기술분야에 공지되어 있는 다른 많은 방법들 중 어느 것이든지 또한 사용될 수 있다. 예를 들어 다중 가닥 대체 증폭법, 헬리카제 대체 증폭법 (HDA), 닉 번역법, Q 베타 복제 효소 증폭법, 회전 서클 증폭법, 및 기타 등온성 증폭법 등이 사용될 수 있다.

<142> 주형 분자는 다양한 공급원 중 어느 것으로부터든지 얻을 수 있다. 예를 들어 DNA는 피험체로부터 얻어지거나 유도될 수 있는 샘플로부터 분리될 수 있다. 단어 "샘플"은 그것에 대해 서열 결정이 수행될 주형의 모든 공급원을 나타내는 광범위한 의미로 사용된다. 구 "로부터 유도된"은 피험체로부터 직접 얻어진 샘플 및/또는 샘플 중의 핵산이 추가로 주형 분자를 얻기 위해 처리될 수 있음을 나타내기 위해 사용된다. 샘플의 공급원은 어떠한 바이러스, 원핵, 고세균, 또는 진핵 종일 수 있다. 본 발명의 어떤 실시예에서 공급원은 사람이다. 샘플은 예컨대 세포를 함유하고 있는 혈액 또는 다른 체액; 정자; 생검 등일 수 있다. 관심 있는 어떠한 유기체로부터의 게놈 또는 미토콘드리아 DNA가 서열화될 수 있다. cDNA가 서열화될 수도 있다. RNA는 또한 예컨대 RT-PCR과 같이 당해 기술분야에서 공지되어 있는 방법을 사용하여 cDNA를 생성하는 첫 번째 역 전사에 의해 서열화될 수 있다. 상이한 샘플 및/또는 피험체로부터 얻어지는 DNA의 혼합물이 조합될 수 있다. 샘플은 다양한 방법 중 어떠한 것에 의해서 처리될 수 있다. 핵산은 공지된 방법을 사용하여 샘플로부터 분리, 정제, 및/또는 증폭될 수 있다. 물론 유기체로부터 유도되지 않는 전체적으로 인공적인, 합성 핵산, 재조합 핵산도 또한 서열화될 수 있다.

<143> 주형은 이중 또는 단일-가닥 형태로 제공될 수 있다. 전형적으로 주형이 초기에 이중 가닥 형태로 제공될 때 두 가닥은 그 결과로서 분리될 것이며 (예컨대 DNA는 변성될 것이다), 두 가닥 중 하나만이 증폭되어 주형 분자의 국소화된, 예컨대 마이크로입자에 부착되고, 반-고체 지지체에 또는 그 위에 고정되는 등의 주형 분자의 국소화된 클론 집단이 생성된다.

<144> 주형은 다양한 추가의 방법으로 선택되거나 처리될 수 있다. 예를 들어 메틸-민감성 제한 효소 (예컨대 MspI)로 처리되어 있는 DNA로부터 얻어진 주형이 사용될 수 있다. DNA 단편을 초래할 그러한 처리는 증폭 전에 수행될 수 있다. 메틸화된 염기를 함유하는 단편은 증폭되지 않는다. 하이포메틸화된 주형으로부터 얻어진 서열 정보는

하이포메틸화에 대한 선택이 수행되지 않은 동일한 공급원으로부터 유도된 주형으로부터 얻어진 서열 정보와 비교될 수 있다.

<145> 주형은 라이브러리에 삽입되거나, 그 안에 제공되거나, 그것으로부터 유도될 수 있다. 예를 들어 하이포메틸화된 라이브러리는 당해 기술분야에 공지되어 있다. 주형을 라이브러리 안에 삽입하는 것은 주형의 말단에 추가의 뉴클레오티드 서열, 예컨대 태그, 프라이머 또는 개시 올리고뉴클레오티드를 위한 결합 부위 등이 쉽게 연결되는 것을 가능하게 한다. 예를 들어 어떤 스트래티지는 다수의 결합 부위, 예컨대 증폭 프라이머를 위한 결합 부위, 개시 올리고뉴클레오티드를 위한 결합 부위, 포획체를 위한 결합 부위 등을 가지는 태그의 첨가를 가능하게 한다.

<146> 다양한 적당한 라이브러리가 당해 기술분야에 공지되어 있다. 예를 들어 특별한 관심 있는 라이브러리 및 그것을 구성하기 위한 방법들이 USSN 10/978,224, PCT 공보 제 WO 2005/042781호 및 제 WO 2005/082098호의 문헌에 설명되어 있다 (Shendure, J., et al., Science, 309(5741):1728-32, 2005, Scienceexpress, 4 Aug. 2005 (www.scienceexpress.org)). 물론 그러한 라이브러리를 제조하는 다른 방법들도 사용될 수 있다는 것이 인지될 것이다. 특히 관심 있는 어떤 라이브러리는 그것들의 각각이 서열화 단계에서 사용된 증폭 및/또는 서열화 프라이머에 상보하는, 즉 이들 서열이 프라이머 결합 영역 (PBR)으로서 작용하는 서열에 의해 분리된 관심 있는 2개의 핵산 절편을 함유하는 다수의 핵산 단편 (전형적으로 DNA)을 함유한다. 특히 관심 있는 실시예에서 핵산 절편은 자연 발생 DNA의 연속적인 조각의 부분이다. 예를 들어 절편은 상기에서 언급된 참조문헌에서 설명된 것과 같이 게놈 DNA의 연속적인 조각의 5'과 3' 말단으로부터 유래될 수 있다. 그러한 핵산 절편은 본원에서 상기 언급된 참조문헌과 일치하는 방식으로, "태그" 또는 "말단 태그"로서 언급된다. 단일한 연속 핵산으로부터 유도된 2개의 태그, 예컨대 5'과 3' 말단은 그러므로 "한 쌍의 태그", "쌍을 이룬 태그", 또는 "2태그"로서 언급된다. "한 쌍의 태그"는 단수로 사용될지라도 2개의 태그를 포함한다는 것이 인지될 것이다. 그것으로부터 한 쌍의 태그가 유도되는 예정된 크기 한계 범위 내의 DNA의 연속적인 조각을 선택함으로써, 2개의 태그를 분리하는 거리가 필연적으로 발생하게 된다.

<147> 서열화 및/또는 증폭 프라이머에 상보하는 서열에 의해 분리되는 것 외에, 라이브러리의 핵산 단편들은 또한 전형적으로 태그 양 옆의 서열화 및/또는 증폭 프라이머에 상보하는 서열을 함유한다. 즉 첫 번째의 그런 서열은 단편의 5' 말단에 더 가까운 태그에 대해 5' 쪽에 위치할 수 있고, 두 번째 그런 서열은 단편의 3' 말단에 더 가깝게 위치한 태그에 대해 3' 쪽에 위치할 수 있다. 그것으로부터 태그가 유도될 수 있는 연속적인 핵산에 존재하는 것과 같은 2개의 태그의 위치는 반드시 그럴 필요는 없지만, 다양한 실시예에서 라이브러리의 DNA 단편에 있는 태그의 위치에 상보할 수 있다는 것이 주지된다.

<148> 핵산 단편과 태그는 상이한 크기 범위를 가질 수 있다. 전형적으로 핵산 단편은 예를 들면 길이가 80 내지 300 뉴클레오티드, 예컨대 100 내지 200, 100 내지 150, 대략 150 뉴클레오티드, 대략 200 뉴클레오티드일 수 있다. 태그는 예컨대 15 내지 25 뉴클레오티드 길이, 예컨대 대략 17 내지 18 뉴클레오티드 길이일 수 있다. 이들 길이는 예시적인 것이고, 위의 범위로 제한하려는 의도는 아니라는 것이 주지되어야 한다. 더 짧거나 더 긴 단편 및/또는 태그가 사용될 수 있다.

<149> 단일한 연속적인 핵산으로부터 쌍을 이룬 태그를 얻는 것이 라이브러리를 구성하기 위한 편리한 방법을 제공하는 한편으로, 쌍을 이룬 태그의 중요한 측면은 그것들이 원래 유도되었던 핵산에서 일정 거리 ("분리 거리")만큼 서로로부터 분리된다는 사실이며, 이때 분리 거리는 예정된 범위의 거리 내에 있다. 태그가 예정된 범위 내에 속하는 분리 거리만큼 분리되어 있다는 사실은 참조 서열 (예컨대 참조 게놈 서열)에 대하여 태그 서열이 일렬 배열되는 것을 가능하게 한다. 어떠한 이론에 구속되는 것을 바라지는 않지만, 이것은 참조 게놈과 관련하여 여전히 정확한 서열의 대체를 가능하게 하는 한편으로 더 짧은 판독 길이의 사용을 가능하게 하는 게놈 재서열화와 같은 특정 용도에 유익할 수 있다. 한 쌍의 태그의 5'과 3' 태그는 핵산의 더 큰 조각의 절편, 예컨대 게놈 DNA를 나타내고 (즉 그 절편의 서열을 가지고), 그 절편은 DNA의 자연 발생 조각에서 서로간 예정된 거리 안에, 예컨대 게놈 DNA의 조각 안에 위치한다. 예를 들어 본 발명의 어떤 실시예에서 한 쌍의 태그의 5' 및 3' 태그는 DNA의 자연 발생 조각에서 서로 최대 500 뉴클레오티드 내에, 서로 최대 1kB 내에, 서로 최대 2kB 내에, 서로 최대 5kB 내에, 서로 최대 10kB 내에, 서로 최대 20kB 내에 위치한 DNA의 절편을 나타낸다. 어떤 실시예에서 한 쌍의 태그의 5' 및 3' 태그는 DNA의 자연 발생 조각에서 500 뉴클레오티드와 2kB 떨어진 사이에, 예컨대 700 뉴클레오티드와 1.2kB 떨어진 사이에, 대략 1kB 떨어져 위치한다. 한 쌍의 태그에서 두 태그가 분리되는 정확한 거리는 중요한 것은 아니며 일반적으로는 알려져 있지 않다는 것이 주지된다. 또한, 태그는 원래 핵산의 더 큰 조각으로부터 얻어지지만, 단어 "태그"는 그것의 원래의 서열 맥락에 존재하거나 라이브러리 단편, 라이브러리 단편으로부터의 증폭 생성물, 서열화될 주형 등에 존재하거나 간에 태그의 서열을 가지는 모든 핵산 절

편에 적용된다.

<150> 핵산 단편 (예컨대 라이브러리 분자)은 다음의 구조를 가질 수 있다:

<151> 링커 1 - 태그 1 - 링커 3 - 태그 1 - 링커 2

<152> 태그 1과 태그 2는 한 쌍의 태그의 5' 및 3' 태그일 수 있다. 태그 중 어느 것이든지 하나는 5' 태그이거나 3' 태그일 수 있다. 링커 1 및 링커 2는 하나 또는 그 이상의 프라이머에 대한 프라이머 결합 영역을 함유한다. 어떤 실시예에서 링커 1과 링커 2는 각각 증폭 프라이머에 대한 PBR과 서열화 프라이머에 대한 PBR을 함유한다. 각 링커의 프라이머들은 모여있을 수 있으며(nested), 서열화 프라이머 PBR은 증폭 프라이머 PBR에 대해 안쪽에 위치한다. 링커 3은 태그 1과 태그 2의 서열화를 가능하게 하기 위한 하나 또는 그 이상의 서열화 프라이머에 대한 PBR을 함유할 수 있다. 용어 "링커"는 라이브러리의 다중 핵산 단편에 존재하는, 예컨대 실질적으로 라이브러리의 모든 단편에 존재하는 핵산 서열을 말한다. 링커는 라이브러리가 구성되는 동안 실제로 연결 기능을 수행할 수도 있고 그렇지 않을 수도 있으며, 단순히 주어진 라이브러리의 대부분의 또는 모든 구성원에 공통적인 규정된 서열로 간주할 수 있다. 그러한 서열은 또한 "보편적 서열"로서 언급된다. 그러므로 링커 또는 그것의 부분에 상보하는 핵산은 라이브러리의 다중 구성원에 혼성화할 것이고, 라이브러리의 대부분의 또는 모든 분자에 대한 증폭 프라이머 또는 서열화 프라이머로 사용될 수 있을 것이다.

<153> 본 발명의 어떤 실시예에서, 핵산 단편은 다음의 구조를 가진다:

<154> 링커 1 - 태그 1 - 내부 어댑터 - 태그 2 - 링커 2

<155> 태그 1과 태그 2 및 링커 1과 링커 2는 상기에서 설명된 바와 같이 PBR을 함유한다. 내부 어댑터는 2개의 프라이머 결합 영역을 함유하는데, 그것은 아래에서 추가로 논의되는 바와 같이 IA 및 IB로 언급될 것이다. 이들 PBR은 2개의 뚜렷한 실질적으로 동일한 핵산의 집단이 부착된 마이크로입자를 생성하는 데 사용되고, 이때 핵산은 태그 1을 포함하는 집단의 것이고 태그 2를 포함하는 다른 집단의 핵산이다. 2개의 뚜렷한 핵산 집단은 최소한 부분적으로 상이한 서열을 가진다. 예컨대 그것들은 태그 영역의 서열이 상이하다. 내부 어댑터는 2개의 프라이머 결합 영역 사이에 스페이서 영역을 함유한다. 스페이서 영역은 비염기성 잔기를 함유할 수 있으며, 그것은 스페이서를 통해 중합효소가 연장되는 것을 방지할 것이다. 물론 스페이서를 통해 중합효소 연장을 방해할 어떠한 다른 차단기를 함유하는 스페이서 영역이 사용될 수 있다.

<156> 다른 실시예에서, 핵산 단편은 하나 또는 그 이상의 추가 태그 (예컨대 2, 4, 6개, 등) 및 하나 또는 그 이상의 추가의 내부 어댑터를 포함한다. 예를 들어 핵산 단편은 다음의 구조를 가진다:

<157> 링커 1 - 태그 1 - 내부 어댑터 1 - 태그 2 - 링커 2 - 태그 3 - 내부 어댑터 2 - 태그 4 - 링커 3

<158> 본 발명의 그러한 단편의 핵산 단편 및 라이브러리, 둘 또는 그 이상의 실질적으로 동일한 핵산 집단을 함유하는 마이크로입자, 및 그러한 마이크로입자의 배열은 본원에서 설명되는 결찰-기반 서열화 방법 이외의 다양한 서열화 방법에 사용될 수 있다. 예를 들어 FISSEQ, 피로서열화 등과 같은 서열화 방법들이 사용될 수 있다 (WO 2005/082098 참조). 물론 결찰-기반 방법도 또한 유익하게 사용될 수 있다. 본원에서 설명된 결찰-기반 방법의 맥락에서 용어 "서열화 프라이머"는 "개시 올리고뉴클레오타이드"를 의미하는 것으로 인식될 수 있음이 인지될 것이다.

<159> 본 발명의 어떤 실시예에서 서열화될 주형은 에멀션의 개별적인 수성 구획 ("반응기"로도 불림)에서 PCR에 의해 합성된다. 바람직하게는 구획은 각각 적당한 첫 번째 증폭 프라이머가 부착된 비드와 같은 특별한 지지체, 주형의 첫 번째 복사물, 두 번째 증폭 프라이머, PCR 반응에 필요한 성분들 (예컨대 뉴클레오타이드, 중합효소, 보조인자, 등)을 함유한다. 에멀션을 제조하는 방법은 예를 들면 미국 특허 제 6,489,103호 (Griffiths); 제 5,830,663호 (Embleton); 및 미국 공보 제20040253731호 (Ghadessy)에 설명되어 있다. 마이크로입자에 부착된 주형의 클론 집단을 생성하기 위해 에멀션의 개별적인 구획 내에서 PCR을 수행하는 방법은 예를 들면 문헌 (Dressman, D., et al, *Proc. Natl. Acad. Sci.*, 100(15):8817-8822, 2003)과 PCT 공보 제 WO 2005/010145호에 설명되어 있다.

<160> 상기 언급된 참조문헌에 설명된 방법, 또는 그것들의 변형방법은 서열화를 위해 마이크로입자에 부착된 주형의 클론 집단을 생성하는 데 사용될 수 있다. 바람직하며 비-제한적인 실시예에서 PCR에 적당한 짧은 (<500 뉴클레오타이드) 주형은 보편적 어댑터 서열을 상이한 표적 서열(주형)의 각 말단에 부착함으로써 (예컨대 결찰 (ligation)에 의하여) 생성될 수 있다. (이 문맥에서 보편적이라 함은 동일한 어댑터 서열이 각각의 주형에 부착되어 PCR 증폭 프라이머의 단일 쌍을 사용하여 증폭될 수 있는 "적절한" 주형이 생성된다는 것을 의미한다.)

벌크 PCR 반응은 적응된 주형, 하나의 유리 증폭 프라이머, 두 번째 증폭 프라이머가 부착된 마이크로입자, 및 다른 PCR 시약 (예컨대 중합효소, 보조인자, 뉴클레오타이드 등)을 사용하여 제조된다. 수성 PCR 반응은 오일 상 (가벼운 미네랄 오일 및 계면활성제를 함유함)과 1:2 비율로 혼합된다. 이 혼합물을 유-중-수 에멀션을 생성하기 위해 와동(vortex)한다. 1 ml의 혼합물이 에멀션 내에서 4×10^9 이상의 수성 구획을 생성하기에 충분하며, 각각은 잠재적인 PCR 반응기이다. 에멀션 샘플의 분취액은 마이크로적정 플레이트 웰 (예컨대 96 웰 플레이트, 384 웰 플레이트 등)에 분배되고 마이크로입자 상에서 고-상 PCR 증폭을 이루기 위하여 열적으로 순환된다. 클론을 확보하기 위하여, 마이크로입자와 주형 농도는 조심스럽게 조절되고, 이로 인해 반응기는 거의 하나 이상의 비드 또는 주형 분자를 함유하지 않는다. 예를 들어 본 발명의 어떤 실시예에서 반응기의 최소한 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 또는 그 이상은 단일 비드와 단일 주형을 함유한다. 따라서, 주형의 각 클론 집단의 구성원들은, 마이크로입자에 이들이 부착된 결과, 서로 가깝게 공간적으로 위치한다. 일반적으로 주형의 부착 지점은 실질적으로 입자의 표면에 균일하게 분포될 수 있다.

<161> 개별적인 마이크로입자가 한 쌍의 태그의 5' 태그와 3' 태그가 부착된 증폭된 핵산 단편의 뚜렷한 집단을 가진 마이크로입자의 집단을 제조하기 위하여 PCR 에멀션 방법을 사용하는 것이 특히 주목된다. 달리 표현하자면, 개별적인 입자가, 상기에서 설명된 것과 같이 증폭되어 이들에 부착된 것들과 같은, 라이브러리과 상이한 핵산 단편을 가지는 마이크로입자의 집단을 제조하는 것에 특히 관심이 있다.

<162> 에멀션에서 DNA를 증폭하기 위한 것으로 당해 기술분야에 공지되어 있는 방법들 (예컨대 상기에서 언급된 참조 문헌에서 설명된 것들)은 큰 핵산 분자의 증폭을 이루고 이들 분자를 마이크로입자에 부착하는 능력의 관점에서 제한된다. 예를 들어 PCR 효율은 더 긴 증폭산물(amplicon)에 의해 지수적으로 감소하는 것이 증명되었다. PCR 효율의 이런 감소는 PCR 에멀션에서 증폭될 수 있고 그러한 증폭을 통하여 마이크로입자에 부착될 수 있는, 상기에서 설명된 것들과 같은, 쌍을 이룬 태그와 프라이머 결합 부위를 함유하는 핵산 단편을 사용했을 때의 효율을 감소시킨다. 그러므로 한 쌍의 태그의 첫 번째 및 두 번째 태그를 함유하는 실질적으로 동일한 핵산 단편의 단일 집단이 PCR 에멀션으로 증폭되고 그러한 증폭을 통하여 비드에 부착되는 방법들은 많은 한계점에 시달린다.

<163> 본 발명은 더 작은 증폭산물을 사용할 수 있으면서도 한 쌍의 5' 및 3' 태그를 함유하는 단일 핵산 단편이 마이크로입자에 증폭을 통해 부착될 때 발생하는 쌍을 이룬 태그 정보를 여전히 보존하는 접근법을 제공한다. 본 발명은 최소한 두 집단이 각각 다수의 실질적으로 동일한 핵산으로 구성되고, 실질적으로 동일한 첫 번째 핵산 집단이 관심 있는 첫 번째 핵산 절편, 예컨대 5' 태그를 포함하고, 두 번째 핵산 집단이 관심 있는 두 번째 핵산 절편, 예컨대 3' 태그를 포함하는, 최소한 두 개의 구별되는 핵산 집단이 부착된 마이크로입자, 예컨대 비드를 제공한다. 첫 번째 및 두 번째 핵산 집단은 두 개의 태그를 함유하고 또한 태그 양 옆에 있으면서 태그들을 분리시키는 적절하게 위치한 프라이머 결합 부위를 함유하는 더 큰 단일 핵산 단편으로부터 증폭됨으로써 두 개의 증폭 반응이 차례로, 또는 바람직하게는 동시에 PCR 에멀션의 단일 반응기에서 마이크로입자 및 증폭 시약의 존재하에 수행될 수 있다. 마이크로입자는 그것에 부착된 두 개의 상이한 프라이머 집단을 가지는데, 하나는 핵산 단편의 태그 중 하나의 외부에 있는 프라이머 결합 영역의 서열에 상응하고, 다른 하나는 핵산 단편의 다른 태그의 외부에 있는 프라이머 결합 영역, 즉 두 개의 태그 양 옆에 있는 프라이머 결합 영역의 서열에 상응한다.

<164> 또한 본 발명에는 두 개의 태그 사이에 위치한 프라이머 결합 영역에 결합하는 프라이머가 제공되고, 그로써 각각이 태그 중 하나를 함유하는 핵산 단편의 일부분을 증폭시키는 두 개의 별도의 PCR 반응이 수행된다. 증폭된 핵산 절편은 서로 다른 추가의 프라이머 결합 영역들을 함유한다. 이들 추가의 프라이머 결합 영역은 핵산 절편에 존재하고 증폭 프라이머에 대한 PBR의 안쪽에 위치한다, 즉 그것들은 모여 있다. 이들 추가의 PBR은 두 개의 상이한 서열화 프라이머에 대한 결합 영역으로서 작용한다. 그러므로 두 개의 상이한 서열화 프라이머 중 하나 또는 다른 하나를 실질적으로 동일한 핵산 절편의 두 집단이 부착된 마이크로입자에 적용함으로써, 두 개의 핵산 절편 중 하나 또는 다른 하나는 다른 핵산 절편의 존재로 인한 간섭 없이 서열화될 수 있다. 각각의 핵산 절편은 그것이 증폭되는 핵산 단편보다 상당히 더 짧으며, 그로써 에멀션-기반 PCR이 쌍을 이룬 태그를 함유하는 단편의 라이브러리를 사용하여 수행될 수 있는 효율을 개선시키는 한편, 한 쌍의 태그의 태그들 사이의 결합을 여전히 보존한다.

<165> 상기에서 설명된 방법은 도 34와 도 35의 다양한 패널을 참조로 하여 더 잘 이해될 수 있는데, 도면에서 동일한 서열을 가지는 핵산 부분들은 동일한 색으로 표시된다. 상기의 설명은 도 34 및 도 35와 일치하게 설명될 것이다. 도 34A와 도 35A는 동일한 단계를 도시하는데, 도 35A는 한층 더 상세한 설명을 제시한다. 도 34A 및 도 35A에서 알 수 있는 바와 같이, 두 개의 태그 (태그 1 및 태그 2)를 함유하는 쌍을 이룬 말단 라이브러리 단편

이 내부 어댑터 카세트 (IA-IB)와 양 옆에 있는 독특한 링커 서열 (P1 및 P2)를 사용하여 구성된다. 내부 어댑터 카세트와 양 옆에 있는 링커 서열은 둘 다 PCR 증폭과 DNA 서열화 둘 다를 제공하는 뉴클레오티드 서열을 함유한다. PCR 프라이머 영역은 모여 있는 DNA 서열화 프라이머의 사용이 가능해지도록 디자인된다. DNA 캡처 마이크로입자 (비드)는 독특한 양 옆의 링커 서열에 동일한 두 개의 올리고뉴클레오티드 서열을 부착함으로써 생성된다. PCR 증폭을 위해서는, P1 및 P2 서열을 가지는 올리고뉴클레오티드가 결합된 DNA 캡처 마이크로입자가 단일한 2-태그 라이브러리 단편 (즉 한 쌍의 태그의 5' 태그 및 3' 태그)을 함유하는 라이브러리 단편과 용액-기반 PCR 프라이머를 함유하는 반응 안으로 삽입된다.

<166> 용액-기반한 양 옆의 링커 프라이머 (P1과 P2)는 비교를 위해 제한적 양으로 내부 어댑터 프라이머 (IA 및 IB)에 첨가되고, PCR-생성된 태그 생성물의 효과적인 드라이브-투-비드 증폭을 촉진하기 위해 작용할 것이다 (즉 $P1 \ll IB$, $P2 \ll IA$). 필요하다면 프라이머의 양을 적절하게 조절하는 것이 또한 핵산 집단이 실질적으로 동일한 수의 핵산, 예컨대 첫 번째 집단에 속하는 개별적인 마이크로입자에 대하여 대략 절반의 핵산과 두 번째 집단에 속하는 개별적인 마이크로입자에 대하여 대략 절반을 함유하는 것을 보장할 수 있다. 그러므로 필요하다면 비대칭 PCR의 형태가 상이한 집단의 비율을 조절하기 위해 사용될 수 있다.

<167> 도 34B 및 도 35B에서 알 수 있는 바와 같이 (도 35B는 도 34B와 관련하여 추가로 상세한 점들을 제공한다), 증폭되는 동안 단일 쌍의 말단 라이브러리 단편은 4개의 올리고뉴클레오티드 프라이머 (P1, P2, IA 및 IB)의 존재 하에 두 개의 독특한 PCR 생성물을 생성할 것이다. 한 집단은 P1과 IA가 양 옆에 있는 태그 1을 함유하고, 두 번째 집단은 P2와 IB가 양 옆에 있는 태그 2를 함유한다.

<168> 증폭에 이어 마이크로입자가 초기 라이브러리 단편으로부터 생성된 태그 1과 태그 2에 상응하는 두 개의 독특한 PCR 집단으로 로딩(loading)될 것이다. 따라서 각각의 태그는 독특한 세트의 프라이밍 영역을 함유하여 도 34C, 도 35C 및 도 35D에서 알 수 있는 바와 같은 각 태그의 연속적인 서열화가 가능해진다. 도 35C 및 도 35D는 상이한 서열화 프라이머를 사용하여 태그 1과 태그 2의 순차적인 서열화를 도시한다. 다양한 서열화 방법 중 어느 것이든지 사용될 수 있다.

<169> 상기 방법은 핵산 서열의 두 개 이상의 구별되는 집단, 예컨대 4, 6, 8, 12, 16, 20 집단 (이들 집단은 2, 3, 4, 6, 8, 10개의 쌍을 이룬 태그를 포함한다)이 부착된 마이크로입자를 생성하기 위해 사용될 수 있다. 각 집단은 두 태그의 경우에 상기에서 설명된 바와 같이, 개별적으로 각 서열에 독특한 프라이머 결합 영역을 제공함으로써 서열화될 수 있다.

<170> 본 발명은 도 34 및 도 35에서 도시되고 상기에서 설명된 구조를 가지는 핵산 단편, 그러한 단편의 라이브러리, 그러한 단편으로부터의 핵산 절편이 부착된 마이크로입자, 개별적인 마이크로입자에 다른 마이크로입자에 부착된 것과 다른 서열의 핵산 집단이 부착된 그러한 마이크로입자의 집단, 마이크로입자의 배열, 핵산 단편으로부터 핵산 절편 (태그)을 증폭하기 위한 증폭 프라이머, 마이크로입자에 부착된 핵산 절편을 서열화하기 위한 서열화 프라이머, 단편, 라이브러리 및 마이크로입자의 제조 방법, 및 마이크로입자에 부착된 핵산의 서열화 방법을 포함한다. 본 발명은 전술한 성분들의 어떠한 조합과, 임의로 하나 또는 그 이상의 효소, 완충제, 또는 증폭, 서열화 등에 유용한 다른 시약을 함유하는 키트(kit)를 포함한다.

<171> 필요에 따라 주형이 부착되는 마이크로입자를 풍부하게 하기 위해 다양한 방법이 사용될 수 있다. 예를 들어 마이크로입자에 부착된 증폭 생성물의 일부에 상보하는 올리고뉴클레오티드 (포획제)가 다른 (바람직하게는 더 큰) 마이크로입자, 마이크로적정 웰, 또는 다른 표면과 같은 포획체에 부착되는 혼성화-기반 방법이 사용될 수 있다. 증폭 생성물의 일부는 표적 영역으로서 언급될 수 있다. 표적 영역은 증폭되는 동안 주형에, 예컨대 미지의 서열을 가진 주형 부분의 한쪽 말단에 통합될 수 있다. 예를 들어 표적 영역은 마이크로입자에 부착되지 않은 증폭 프라이머에 존재할 수 있어서, 상보하는 부분이 증폭된 주형에 존재한다. 그러므로 다수의 상이한 주형은 동일한 표적 영역을 포함할 수 있고, 따라서 단일 포획제(capture agent)가 다수의 상이한 주형에 혼성화되어 포획제로서 유일한 단일 올리고뉴클레오티드 서열을 사용하는 다수 마이크로입자의 포획이 가능하다. 증폭된 마이크로입자는 혼성화가 일어날 수 있는 조건하에서 포획제에 노출된다. 그 결과로서, 증폭된 주형이 부착된 마이크로입자들이 포획제를 경유하여 포획제(capture entity)에 부착된다. 그런 다음 부착되지 않은 마이크로입자들이 제거되고, 보유된 마이크로입자들이 방출된다 (예컨대 온도를 상승시킴으로써). 특별한 포획체가 사용되는 어떤 실시예에서, 혼성화 후에 마이크로입자들이 부착된 포획체로 구성되는 응집체가 마이크로입자들이 부착되어 있지 않은 특정 포획체로부터 및 포획체에 부착되지 않은 마이크로입자로부터, 예컨대 글리세롤과 같은 점성 용액 중에서의 원심분리에 의해 분리된다. 크기, 밀도 등을 기초로 한 다른 분리 방법들도 또한 사용될 수 있다. 그러나 풍부하게 하기 위해 사용될 수 있는 많은 방법들 중 한 가지가 혼성화이다. 예를 들어 (예컨대 합

성 중에) 주형에 통합될 수 있는 많은 상이한 리간드 중 하나에 대한 친화력을 가진 포획제가 사용될 수 있다. 많은 회기의 풍부화가 사용될 수 있다.

<172> 도 14A는 첫 번째 증폭 프라이머가 부착된 비드 상(phase)에서 형광 표시된 두 번째 증폭 프라이머와 과잉의 주형을 사용하여 PCR 반응이 수행되는 유-중-수 에멀션의 구획 영상을 도시한다. 수성 반응기는 확산된 유리 프라이머로부터 약하게 형광을 내는 한편, 비드는 고-상 증폭의 결과로서 비드 상에 축적되는 프라이머들로부터 강하게 형광을 나타낸다 (즉 형광 프라이머가 첫 번째 증폭 프라이머를 경유하여 비드에 부착된 증폭된 주형 안에 통합된다). 비드 신호는 상이한 크기의 반응기들에서 균일하다.

<173> 증폭에 이어서, 마이크로입자가 수집되고 (예컨대 자성 입자의 경우 자석을 사용함으로써), 본원에서 설명되는 바와 같은 연장, 결찰, 및 절단의 반복된 사이클에 의해 서열화되기 위해 사용된다. 본 발명의 어떤 실시예에서 마이크로입자는 아래에서 설명되는 바와 같이 서열화 전에 반-고체 지지체에 또는 그 위에 배열된다. 실시예 12, 13, 14 및 15는 (i) 마이크로입자상에서의 주형의 합성을 위해 증폭 프라이머가 부착된 마이크로입자를 제조하기 위해 (실시예 12); (ii) PCR을 수행하기 위한 다수의 반응기를 포함하는 에멀션의 제조를 위해 (실시예 13); (iii) 에멀션의 구획에서 PCR 증폭을 위해 (실시예 13); (iv) 에멀션을 깨고 마이크로입자를 회수하기 위해 (실시예 13); (v) 클론 주형 집단에 부착된 마이크로입자를 풍부하게 하기 위해 (실시예 14); (vi) 반-고체 폴리아크릴아미드 지지체에 대한 기관으로서 작용하는 유리 슬라이드의 제조를 위해 (실시예 15); 및 (vii) 마이크로입자를 중합하지 않은 아크릴아미드와 혼합하여 주형이 부착되어 있고 기관 위의 아크릴아미드에 삽입되어 있는 마이크로입자의 배열을 형성하기 위하여 (실시예 15) 사용될 수 있는 대표적이고 비 제한적인 방법들을 추가로 상세히 설명한다. 실시예 15는 또한 중합효소 트래핑에 대한 프로토콜을 설명하는데, 그것은 반-고체 지지체에서 PCR을 수행할 때 특정 방법에서 사용된다. 당업자는 이들 방법에 대한 많은 수정된 방법이 사용될 수 있음을 인지할 것이다.

<174> 본 발명의 다른 실시예에서 주형은 반-고체 지지체, 예컨대 그 안에 적당한 증폭 프라이머가 고정되어 있는 겔에서 PCR에 의해 증폭된다. PCR 반응에 필요한 주형, 추가의 증폭 프라이머, 및 시약들은 반-고체 지지체 안에 존재한다. 한 쌍의 증폭 프라이머 중 하나 또는 둘 다는 반-고체 지지체에 적당한 연결 부분, 예컨대 아크리다이트 기를 통하여 부착된다. 부착은 중합되는 동안 일어날 수 있다. 추가의 시약 (예컨대 주형, 두 번째 증폭 프라이머, 중합효소, 뉴클레오타이드, 보조인자 등)은 반-고체 지지체의 형성 전에 (예컨대 겔 형성 전 액체 중에) 존재하거나, 또는 하나 또는 그 이상의 시약은 그것이 형성된 후에 반-고체 지지체 안으로 확산될 수 있다. 반-고체 지지체의 기공 크기는 확산이 허용되도록 선택된다. 당해 기술분야에서 잘 알려져 있는 바와 같이, 폴리아크릴아미드 겔의 경우 기공 크기는 주로 아크릴아미드 단량체의 농도에 의해 결정되며 어느 정도는 교차결합제에 의해 결정된다. 다른 반-고체 지지체 물질의 경우에도 유사한 고려사항들이 적용된다. 원하는 기공 크기를 이루기에 적절한 교차결합제 및 농도가 선택될 수 있다. 발명의 어떤 실시예에서 첨가제, 예컨대 양이온성 지질, 폴리아미드, 다가 양이온, 등이 중합반응 전에 용액에 포함되며, 그것은 마이크로입자를 둘러싸는 겔 내 미셀 또는 응집체를 형성한다. 미국 특허 5,705,628호, 5,898,071호 및 6,534,262호에 개시된 방법들이 또한 사용될 수 있다. 예를 들어 다양한 "균집 시약"이 클론 PCR에 대한 비드 가까이에서 DNA를 균집하기 위해 사용될 수 있다. SPRI® 자성 비드 기법 및/또는 조건이 또한 사용될 수 있다. 예를 들어 미국 특허 5,665,572호에는 10%의 폴리에틸렌 글리콜 (PEG)의 존재하에 효과적인 PCR 증폭이 증명되어 있다. 발명의 방법의 어떤 실시예에서 증폭 (예컨대 PCR), 결찰, 또는 이 두 가지 모두는 베타인, 폴리에틸렌 글리콜, PVP-40, 등과 같은 시약의 존재하에 수행된다. 이들 시약은 에멀션에 존재하는 용액에 첨가될 수 있고, 및/또는 반-고체 지지체에 확산될 수 있다.

<175> 반-고체 지지체는 실질적으로 평면인 견고한 기관 위에 위치하거나 조립될 수 있다. 어떤 바람직한 실시예에서 기관은 전형적인 표지 (예컨대 형광 표지, 양자 돛트, 플라스몬 공명 입자, 나노클러스터)의 여기와 검출에 사용된 여기 및 방출 파장의 방사선, 예컨대 대략 400 내지 900nm의 방사선에 대해 투명하다. 유리, 플라스틱, 석영 등의 물질이 적당하다. 반-고체 지지체는 기관에 점착할 수 있고 임의로 다양한 방법 중 어느 것이든지 사용하여 기관에 고정될 수 있다. 기관은 점착력 또는 결합을 증강하는 물질, 예컨대 실란, 폴리라이신 등으로 코팅될 수도 있고 되지 않을 수도 있다. 미국 특허 제 6,511,803호에는 반-고체 지지체에서 PCR을 사용하여 주형의 클론 집단을 합성하는 방법, 실질적으로 평면의 기관 위에서 반-고체 지지체를 제조하는 방법 등이 설명되어 있다. 유사한 방법들이 본 발명에 사용될 수 있다. 기관은 반-고체 기관의 형성 전에 액체를 함유하도록 웰 또는 함몰된 부분을 가질 수 있다. 이와 달리, 도드라진 장벽 또는 마스크가 이 목적에 대해 사용될 수 있다.

<176> 상기 접근법은 클론 주형의 공간적으로 위치한 집단을 생성하기 위해 에멀션에서 반응기를 사용하는 것에 대한 대체방법을 제공한다. 클론 집단은 반-고체 지지체에서 불연속 위치에 존재함으로써, 새롭게 결찰된 연장 프로

브를 검출할 목적으로 서열화되는 동안 각 집단으로부터, 예컨대 영상에 의해 신호가 얻어질 수 있다. 본 발명의 어떤 실시예에서는 둘 또는 그 이상의 구별되는 클론 집단이 단일 핵산 단편으로부터 증폭되고 반-고체 지지체의 구별되는 위치에서 혼합물로서 존재한다. 혼합물의 각 클론 집단은 태그를 포함할 수 있어서, 예컨대 불연속적인 위치는 5' 태그를 함유하는 단편과 3' 태그를 함유하는 단편을 함유한다. 5' 태그와 3' 태그를 포함하는 클론 주형은 상이한 서열화 프라이머를 함유하므로, 그것들은 상호 독립적으로 서열화될 수 있다. 이 접근법은 마이크로입자상에서 실질적으로 동일한 핵산의 다수의 집단을 생성하고 단일 마이크로입자로부터 쌍을 이룬 태그의 두 구성원들에 대한 서열 정보를 얻기 위해 상기에서 설명된 접근법과 동일하다.

<177> 일반적으로 본 발명의 어떠한 방법에 사용하기 위한 반-고체 지지체는 약 100 미크론 또는 그 이하의 두께, 예컨대 약 50 미크론 이하의 두께, 예컨대 약 20 내지 40 미크론의 두께 (포괄적)를 가지는 층을 형성한다. 실질적으로 평면의 표면을 가지는 커버 슬립 또는 유사한 다른 물체가 반-고체 지지체 물질 위에, 바람직하게는 중합반응 전에 놓이게 되어 균일한 겔 층이 생성되는 것을 보조하여, 예컨대 실질적으로 평면인 및/또는 실질적으로 두께가 균일한 겔 층이 형성될 수 있다.

<178> 본 발명의 또 다른 실시예에서는, 상기 방법의 변형 방법이 사용될 수 있는데, 그 방법에서 주형은 적당한 증폭 프라이머가 부착된 마이크로입자 상에서의 PCR에 의해 합성되며, 이때 마이크로입자는 주형 합성 전에 반-고체 지지체에 또는 그 위에 고정된다. 즉 마이크로입자는 전체적으로 또는 부분적으로 반-고체 지지체에 삽입된다. 일반적으로 마이크로입자는 그것들이 밑에 놓여있는 기판 상에서는 정지상태일 수 있지만, 반-고체 지지체 물질에 의해 완전히 둘러싸인다. 그러므로 마이크로입자는 반-고체 지지체가 파괴되지 않는 한 서로에 대하여 실질적으로 고정된 위치에서 유지된다. 이런 접근법은 클론 주형의 공간적으로 위치한 집단을 생성하기 위하여 에멀션을 사용하는 것에 대한 또 다른 대체법을 제공한다. 마이크로입자는 반-고체 지지체의 형성 전에 액체와 혼합될 수 있다. 이와 달리, 마이크로입자는 실질적으로 평면인 기판 위에 배열될 수 있고, 중합반응, 교차결합 등이 일어나기 전에 마이크로입자 배열에 액체가 첨가될 수 있다. 마이크로입자에는 첫 번째 증폭 프라이머가 부착된다. 두 번째 증폭 프라이머는 반드시 그럴 필요는 없지만 반-고체 지지체에 부착될 수 있다. 추가의 시약들 (예컨대 주형, 두 번째 증폭 프라이머, 중합효소, 뉴클레오타이드, 보조인자 등)은 반-고체 지지체가 형성되기 전에 존재할 수 있고 (예컨대 겔 형성 전의 액체에), 또는 하나 또는 그 이상의 이들 시약은 겔이 형성된 후에 반-고체 지지체에 확산될 수 있다. 반-고체 기판은 일반적으로 상기에서 설명된 바와 같이, 예컨대 유리 슬라이드 위에서 형성된다.

<179> 본 발명의 어떤 실시예에서 겔은 가용화되어 (예컨대 소화 또는 탈중합 또는 용해) 클론 주형 집단이 부착된 마이크로입자는 주형 합성 후에 편리하게 회수될 수 있다 (예컨대 자성 입자의 경우 자석을 사용함으로써). 가용화, 소화, 탈중합, 용해 등이 일어날 수 있는 겔은 본원에서 "가역성"으로 언급된다. 종래의 폴리아크릴아미드 중합반응은 중합반응을 개시하기 위하여 적절한 촉매 (예컨대 N,N,N',N'-테트라메틸에틸렌디아민(TEMED))와 함께 교차결합제로서 N-N' 메틸렌비스아크릴아미드(BIS)의 사용을 포함한다. 가역성 겔을 생성하기 위하여 N-N' 디알릴타르타르디아미드(DATD)와 같은 대체 교차결합제도 사용될 수 있다. 이 화합물은 구조적으로는 BIS와 유사하지만 과요오드산에 의해, 예컨대 과요오드산 나트륨을 함유하는 용액 중에서 절단될 수 있는 시스-디올기를 가지고 있다 (Anker, H.S.: F.E.B.S. Lett., 7:293, 1970). 그러므로 DATD 겔은 쉽게 가용화될 수 있다. 교차결합제로서 DATD를 사용하여 만들어진 겔은 매우 투명하고 유리에 잘 결합한다. 가역성 겔을 형성하는 DATD-유사 성질을 가지는 다른 교차결합제는 에틸렌 디아크릴레이트이다 (Choules, G. L. and Zimm, B. S.: Anal. Biochem., 13:336-339, 1965). N,N'-비스아크릴릴시스타민 (BAC)은 가역성 폴리아크릴아미드 겔을 형성하기 위해 사용될 수 있는 다른 교차결합제이다. 과요오드산염에서 녹을 수 있는 겔을 형성하기 위해 사용될 수 있는 다른 교차결합제는 N,N'-(1,2-디히드록시에틸렌)비스아크릴아미드(DHEBA)이다. 다양한 다른 물질 중 어느 것이라도 가역성 반-고체 지지체를 형성할 수 있으면 또한 사용될 수 있다. 예를 들어 열-가역성 중합체, 예컨대 플루로닉 (Pluronic)(BASF사로부터 시판됨)이 사용될 수 있다. 플루로닉은 폴리(에틸렌 옥사이드)-폴리(프로필렌 옥사이드)-폴리(에틸렌 옥사이드) (PEO-PPG-PEO) 삼중블록 공중합체의 군이다 (Nace, V. M., et al., *Nonionic Surfactants*, Marcel-Dekker, NY, 1996). 이들 물질은 상승된 온도 (예컨대 실온보다 높은 온도)에서는 반-고체 (겔)이 되고 냉각하면 액화된다. 플루로닉을 화학적으로 유도하기 위하여, 예컨대 그것에 프라이머의 부착을 용이하게 하기 위하여 다양한 방법들이 사용될 수 있다 (Neff, J. A. et al., J. Biomed. Mater. Res., 40:511, 1998; Prud'homme, RK, et al., Langmuir, 12:4651, 1996).

<180> 가용화 후에 마이크로입자는 수집되어 연장, 결찰, 및 절단의 반복된 사이클을 사용하여 서열화될 수 있다. 서열화 전에 마이크로입자는 두 번째 반-고체 지지체에 또는 그 위에 그것들이 첫 번째 반-고체 지지체에 또는 그 위에 존재할 때보다 더 높은 밀도로 배열될 수 있다. 반-고체 지지체는 전형적으로 그 자체가 실질적으로 평면

이고 견고한 기관, 예컨대 유리 슬라이드에 의해 지지된다.

<181> 그러므로 두 가지 일반적인 접근법이 반-고체 지지체에 또는 그 위에 삽입된 클론 주형 집단을 포함하고 있는 마이크로입자의 배열을 가지는 반-고체 지지체를 생성하기 위하여 사용될 수 있다. 첫 번째 접근법은 반-고체 지지체에는 존재하지 않는 마이크로입자 상에서 증폭을 수행하는 것 (예컨대 에멀션-기반 PCR에 의해)과 그런 다음 그 마이크로입자를 반-고체 지지체에 또는 그 위에 고정하는 것을 포함한다. 두 번째 일반적 접근법은 반-고체 지지체에 또는 그 위에 마이크로입자를 고정하는 것과 그런 다음 증폭을 수행하는 것을 포함한다. 어느 경우든지, 마이크로입자의 응집(clumping)을 감소시키기 위한 밋/또는 마이크로입자를 실질적으로 단일 촛점의 평면에 배열하기 위한 과정을 사용하는 것이 바람직할 것이다. 예를 들어 입자를 폴리아크릴아미드 겔에 고정시킬 때, 단량체와 교차결합제의 농도는 입자가 완전한 중합반응 전에 용액의 바닥에 가라앉음으로써 그것들이 그 밑에 놓여있는 기관에 가라앉고 그로써 단일 평면에 배열되도록 선택된다. 본 발명의 어떤 실시예에서 실질적으로 평면의 표면을 가지는 물체, 예컨대 커버 슬립은 마이크로입자를 함유하는 액체 아크릴아미드 (또는 반-고체 지지체를 형성할 수 있는 다른 물질)의 상부에 놓이게 됨으로써, 아크릴아미드가 "샌드위치" 구조의 두 층 사이에 갇히게 된다. 그런 다음 샌드위치는 뒤집어지고, 그로써 중력의 작용에 의해 마이크로입자가 가라앉게 되고 커버 슬립 (또는 실질적으로 평면 표면을 가지는 다른 물체) 위에 정지하게 된다. 중합반응 후에 커버 슬립은 제거된다. 이렇게 마이크로입자는 반-고체 지지체의 표면에 밀접한 (예컨대 표면에 정접한), 실질적으로 단일 평면에 삽입된다.

<182> 마이크로입자와 같은 지지체를 상기에서 설명한 바와 같이 반-고체 매트릭스에 고정하는 것보다는, 마이크로입자는 본 발명의 어떤 실시예에서는 공유적으로 또는 비 공유적으로 그것을 고정하기 위하여 반-고체 지지체를 사용하지 않는 실질적으로 평면의 견고한 기관에 부착된다. 마이크로입자를 기관, 예컨대 유리, 플라스틱, 석영, 실리콘 등에 부착하기 위한 다양한 방법이 당해 기술분야에 공지되어 있다. 기관은 코팅되거나 (예컨대 스펀-코팅) 되지 않을 수 있고 또는 부착을 촉진하는 물질 (다양한 중합체 중 어느 것) 또는 제제로 기능화될 수 있다. 코팅은 얇은 막, 자체-조립되는 단층 등일 수 있다. 마이크로입자, 마이크로입자에 부착된 부분, 또는 마이크로입자에 부착된 올리고뉴클레오티드 (예컨대 주형) 중 어느 것이든지 부착될 수 있다.

<183> 일반적으로 서로에 대한 친화력을 나타냄으로써 결합 쌍을 형성하는 어떠한 쌍의 분자든지 마이크로입자 또는 주형을 기관에 부착하기 위해 사용될 수 있다. 결합 쌍의 첫 번째 구성원은 공유적으로 또는 비 공유적으로 기관에 부착되며, 결합 쌍의 두 번째 구성원도 공유적으로 또는 비 공유적으로 마이크로입자에 부착된다. 첫 번째 결합 파트너는 링커를 통하여 기관에 부착될 수 있다. 두 번째 결합 파트너는 링커를 통하여 마이크로입자 또는 주형에 부착될 수 있다. 예를 들어 한 접근법에 따르면, 슬라이드 또는 다른 적당한 기관이 아민-반응성 기로 변형된다 (예컨대 아민-반응성 기를 함유하는 PEG 링커를 사용하여). 아민-반응성 기는 수성 조건 하에서 (예컨대 pH 8.0에서) 아민, 예컨대 어떠한 단백질, 예를 들면 스트렙타아비딘 중의 라이신과 반응한다. 그러므로 아민을 내포하는 부분으로 기능화된 마이크로입자는 기관 위에 고정된다. 아민을 내포하는 부분은 단백질이거나 적당하게 기능화된 핵산, 예컨대 DNA 주형일 수 있다. 다수의 부분이 비드에 부착될 수 있다. 예를 들어 비드는 기관에 비드를 부착시키기 위하여 NHS 에스테르와 반응하는 단백질이 부착되어 있을 수 있고, 또한 비드가 기관에 부착된 후에 서열화될 수 있는 DNA 주형이 부착되어 있을 수 있다. 한 말단에 아민-반응성 NHS 부분을 가진 중합체 테터(tether)를 내포하는 적당하게 코팅된 슬라이드는 예컨대 Schott Nexterion사 (Schott North America, Inc., Elmsford, NY 10523)로부터 시판중이다. 이와 달리, 코팅된 슬라이드 (예컨대 비오틴-코팅된 슬라이드)는 Accelr8 Technology Corporation (Denver, CO)으로부터 구할 수 있다. 그들의 OptiChemTM 기법은 마이크로입자를 기관에 부착하는 단지 한 가지 방법을 나타낸다 (미국 특허 제 6,844,028호 참조). 이와 달리, 마이크로입자는 비오틴으로 비드 상의 폴리뉴클레오티드를, 예컨대 비오틴-디데옥시ATP 및/또는 비오틴-데옥시ATP와 함께 말단 트랜스페라제를 사용하여 기능화된 후, 비오틴-스트렙타아비딘 연쇄를 촉진하는 조건 하에서 스트렙타아비딘-코팅된 슬라이드 (예컨대 Accelr8 Technology Corporation, Denver, CO.로부터 구할 수 있다) 와 접촉하여 기관에 부착될 수 있다.

<184> 일반적으로 당해 기술분야에 공지되어 있는 광범위한 방법 중 어느 것이든지 올리고뉴클레오티드 프라이머, 프로브, 주형 등과 같은 핵산을 변형하여, 그러한 핵산이 마이크로입자 또는 다른 지지체 또는 기관에 부착되는 것을 촉진하기 위해 사용될 수 있다. 또한 당해 기술분야에 공지되어 있는 광범위한 방법 중 어느 것이든지 핵산이 부착하기 쉽도록 마이크로입자 또는 다른 지지체를 변형하고, 마이크로입자가 지지체 또는 기관에 부착하는 것을 촉진하기 위해 사용될 수 있다. 마이크로입자는 원하는 기능성의 부착을 촉진하는 표면 화학을 가지는 것들이 활용될 수 있다. 이들 표면 화학의 어떤 실례로는 그것들에 한정되는 것은 아니지만, 지방족 및 방향족 아민을 포함한 아미노기, 카르복실산, 알데하이드, 아미드, 클로로메틸기, 히드라지드, 히드록실기, 술포네이트

및 술페이트가 있다. 이들 기는 핵산에 존재하는 기와 반응할 수 있고, 또는 핵산은 반응성 기의 부착에 의해 변형될 수 있다. 또한 대다수의 안정한 이중 기능성 기가 당해 기술분야에 잘 알려져 있는데, 이를테면 동형의 이작용성(homobifunctional) 및 이형의 이작용성(heterobifunctional) 링커이다. (Pierce Chemical Technical Library, available at the Web site having URL www.piercenet.com (원래는 1994-95년에 Pierce Catalog에 공개됨) and G. T. Hermanson, Bioconjugate Techniques, Academic Press, Inc., 1996; 또한 미국 특허 제 6,632,655호 참조).

<185> 본원에서 설명된 방법에 따라 형성된 마이크로입자의 배열은 일반적으로 무작위적이다. 본원에서 사용되는 바와 같이, 용어 "무작위로 패턴화된" 또는 "무작위"는 지지체 위의 실체(특징)의 무질서한, 비-데카르트식 분포 (달리 표현하면, 예정된 지점에서 또는 격자의 x-와 y-축을 따른 위치들에서 또는 규정된 '클락(clock) 위치'에서 배열되지 않은, 방사상 패턴의 중심으로부터 각도 또는 반경), 즉 의도한 디자인 (또는 그러한 디자인이 이용될 수 있는 프로그램)을 통해 또는 개별적인 실체의 대체에 의해 이루어지지 않는 것을 말한다. 그러한 실체의 "무작위로 패턴화된" 또는 "무작위" 배열은 실체의 풀을 포함하고 있는 용액, 에멀션, 에어로졸, 증기 또는 건조체제를 지지체 위에 또는 그 안에 드롭(drop), 분무, 도포, 도말 또는 분배하고, 그것들이 지지체의 또는 그 위의 특정 부위에 지시되는 어떠한 방식으로든지 간접받는 일 없이 그것들이 지지체 위에 또는 그 안에 정착되도록 함으로써 이루어질 수 있다. 예를 들어 실체는 반-고체 지지체에 대한 선구체(예컨대 아크릴아미드 단량체)를 함유하는 용액 중에 현탁될 수 있다. 그런 다음 그 용액은 두 번째 지지체 위에 분포되고 반-고체 지지체가 두 번째 지지체 위에 형성된다. 실체는 반-고체 지지체에 또는 그 위에 삽입된다. 물론 비-무작위 배열도 사용될 수 있다. 일반적으로 본원에서 사용되는 배열 형성 방법은, 예를 들면 폴리뉴클레오티드의 합성이 기관 위의 예정된 위치에서 개별적인 뉴클레오티드 하위단위의 순차적인 적용에 의해 일어나는 방법과는 구별된다.

<186> 도 14B (상부)는 그 위에 폴리아크릴아미드 겔이 있는 슬라이드의 형광 영상(1인치에서 3인치)을 도시한다. 비드에 부착된 주형에 혼성화된 형광 표지된 올리고뉴클레오티드를 가지는 비드(1 미크론 직경)는 겔에 고정되어 있다. 영상은 슬라이드당 대략 280×10^6 개의 비드를 영상화하기에 충분한 비드 표면 밀도 (즉 비드가 위치한 영역 내에서의 기관의 단위 면적당 비드의 수)를 도시한다. 표면 밀도 및 영상화 가능한 면적은 단일 슬라이드 위에서 최소한 500×10^6 개의 비드를 영상화하기에 충분하다. 예를 들어 도 14B(바닥)은 비드가 폴리아크릴아미드 겔과 같은 반-고체 지지체 층에 삽입될 투명한 면적을 둘러싸고 있는 Teflon® 마스크를 포함한 슬라이드의 개략적인 도면이다. 이 마스크의 면적은 864mm^2 이다. 500×10^6 개의 비드가 있을 때, 표면 밀도는 mm^2 당 578,000 비드이다. 1 미크론 비드의 빈틈없이 채워진 6각형 배열은 mm^2 당 1,155,000 비드를 제공하므로, 이 실시예는 이론상 최대 밀도의 52%를 가지는 배열을 유발한다. 더 작고 더 큰 수의 비드, 그리고 더 크거나 적은 비드 표면 밀도가 이 특정 실시예 외에 사용될 수 있음이 인지될 것이다.

<187> 마이크로입자는 실질적으로 평면인 반-고체 지지체 위에, 또는 다른 지지체 또는 기관 위에 다양한 밀도로 배열될 수 있으며, 밀도는 다양한 방법으로 규정될 수 있다. 예를 들어 밀도는 실질적으로 평면인 배열이 단위 면적당 마이크로입자 (예컨대 구형 마이크로입자)의 수의 관점에서 표시될 수 있다. 발명의 어떤 실시예에서 실질적으로 평면인 배열의 단위 면적당 마이크로입자의 수는 육각형 배열의 마이크로입자의 수의 최소한 80%이다 ("육각형 배열"이란 배열의 모든 마이크로입자가 미국 특허 6,406,848호에서 설명된 바와 같이 동일 면적의 다른 인접한 마이크로입자를 최소한 6회 접촉하게 되는 마이크로입자의 실질적으로 평면인 배열을 의미한다). 그러나 발명의 다른 실시예에서 밀도는 더 낮다. 예를 들어 실질적으로 평면인 배열의 단위 면적당 마이크로입자의 수는 육각형 배열의 마이크로입자의 수의 80% 이하, 70% 이하, 60% 이하, 또는 50% 이하이다. 어떠한 이론에 구속되는 것을 원하지는 않지만, 예를 들어 효소, 프라이머, 보조인자 등과 같은 시약의 적절한 확산을 가능하게 하고, 특정 시약이 마이크로입자에 대한 차등 친화력을 가지거나 그 안에 캡처되는 경우 발생할 수 있는 시약 분할 효과를 피하기 위해서 더 낮은 밀도를 활용하는 것이 바람직할 수 있다. 그러한 효과는 배열상의 다른 위치에서 상이한 반응 조건을 조래할 수 있고, 심지어는 이들 시약에 의해 배열상의 특정 위치로의 접근을 방해할 수도 있다. 이런 문제점은 반응이 유동 셀에서 수행될 때 악화될 수 있는데, 왜냐하면 시약이 유동 셀을 통해 일정 방향을 향해 가는 방식으로 이동하기 때문이다. 발명의 어떤 실시예에서 혼합 장치, 예컨대 기계식 또는 음향식 수단에 의해 유체 혼합이 이루어지는 장치가 유동 셀의 챔버에 포함된다. 많은 적당한 혼합 장치가 당해 기술분야에 공지되어 있다.

<188> 본 발명의 서열화 방법은 모든 유형의 배열 포맷, 이를테면 마이크로입자의 배열 또는 주형의 배열 그 자체일 수 있는 무작위와 비-무작위 배열 두 가지로 배열된 주형을 사용하여 실행될 수 있다. 예를 들어 그 위에 주형이 배열되어 있는 지지체는 미국 특허 5,641,658호와 PCT 공보 W00018957에 설명되어 있다. 배열은 필터, 멤브

레인 (예컨대 나일론), 금속 표면 등과 같은 광범위한 기관 위에 위치할 수 있다. 그 위에서 연장, 결찰, 및 절단의 반복된 사이클에 의해 서열화가 수행될 수 있는 배열 포맷 추가의 실례는 섬유 광학 번들의 개별적인 광학 섬유의 말단 또는 기부 말단에서 웰에 위치한 비드의 배열이다. 예컨대 미국 공보 및 특허 제 6,023,540호; 제 6,429,027호, 제 20040185483호, 제 2002187515호, PCT 출원 제US98/05025 호, 및 PCT 제 US98/09163호 및 PCT 공보 제 W00039587호에서 설명된 비드 배열 및 "배열의 배열" 참조. 주형이 부착된 비드는 본원에서 설명된 바와 같이 배열될 수 있다. 증폭은 바람직하게는 배열이 형성되기 전에 수행된다. 그러한 기관 위에 형성된 배열은 반드시 실질적으로 평면일 필요는 없다.

<189> 다른 실시예에서 PCR은 기관 또는 지지체에 부착된 올리고뉴클레오타이드를 포함하는 배열상에서 수행된다 (예컨대 미국 특허 제 5,744,305호; 제 5,800,992호; 제 6,646,243호 및 관련 특허 (Affymetrix); PCT 공보 제 W02004029586호; 제W003065038호; 제 W003040410호 (Nimblegen) 참조). 일반적으로 그러한 올리고뉴클레오타이드는 3' 또는 5' 말단을 가진다. 필요에 따라, 상기 말단은 예컨대 이미 존재하고 있는 것이 아니라면, 포스페이트기 또는 OH 기를 3' 말단에 첨가하여 변형될 수 있다. 지지체 또는 기관에 부착된 올리고뉴클레오타이드에 상보하는 영역을 포함하는 주형 분자는 올리고뉴클레오타이드에 혼성화하고, PCR은 배열 위에서 제자리 수행되어 배열 위의 각 위치에서 클론 주형 집단이 생성된다. 배열에 부착된 올리고뉴클레오타이드는 증폭 프라이머 중 하나로 기능화될 수 있다. 그런 다음 주형이 본원에서 설명된 결찰-기반 방법을 사용하여 서열화된다. 서열화는 또한 미국 공보 제 20030068629호에서 설명된 것과 같은 배열에서 주형에 대해 수행될 수 있다.

<190> 표면에 DNA 배열을 제조하기 위한 또 다른 방법이 사용될 수 있다. 예를 들어 말단 알데히드 기로 변형된 알칸티올(alkanethiol)이 금 표면에 자기조립 분자막막(SAM)을 제조하기 위해 사용될 수 있다. 단층의 알데히드기는 아민-변형된 올리고뉴클레오타이드 또는 다른 아민-내포 생체분자와 반응하여 슈프(Schiff) 염기를 형성하고, 그것은 계속해서 나트륨 시아노보로하이드라이드(sodium cyanoborohydride)로 처리함으로써 안정한 이차 아민으로 환원될 수 있다 (Peelen & Smith, Langmuir, 21(1):266-71, 2005). 그런 다음 주형의 PCR 증폭이 수행될 수 있다. 이와 달리, 주형의 클론 집단이 부착된 마이크로입자는 마이크로입자 또는 주형 위의 아민기 또는 입자에 부착된 올리고뉴클레오타이드와 그러한 표면과의 반응에 의해 표면에 부착될 수 있다.

<191> 클론 주형 집단이 부착된 마이크로입자를 얻는 또 다른 방법은 미국 특허 제 5,604,097호에 설명되어 있는 "고상 클로닝" 접근법으로서, 그 방법은 마이크로입자 상에 폴리뉴클레오타이드를 분류하기 위한 올리고뉴클레오타이드 태그를 사용하여 동일한 서열의 폴리뉴클레오타이드들만이 어떠한 특정 마이크로입자에 부착되도록 한다.

<192> 발명의 어떤 실시예에서, 연장, 결찰 및 절단의 반복된 사이클에 의한 서열화는 서열화 시약 (예컨대 연장 프로브, 결합 효소, 포스포타제 등)을 지지체에 또는 지지체 위에 주형의 클론 집단이 고정되어 있는 겔과 같은 반-고체 지지체 안에 확산시키고, 그로써 각각의 클론 집단이 지지체의 공간적으로 구별되는 영역에 위치하게 하여 수행된다. 어떤 실시예에서 주형은 상기에서 설명된 것처럼 반-고체 지지체에 직접 부착된다. 그러나 바람직한 실시예에서 주형은 마이크로입자와 같은 두 번째 지지체에 고정되고, 그 입자가 계속해서 또한 상기에서 설명된 것처럼 반-고체 지지체에 또는 그 위에 고정된다.

<193> 하기 실시예 1에서 설명되는 것과 같이, 본 발명자들은 인내를 요하는 결찰과 절단이 폴리아크릴아미드 겔에서 고정된 비드에 부착된 주형에 대하여 수행될 수 있음을 발견하였다. 그러므로 본 발명은 다음의 단계들로 이루어지는 두 번째 폴리뉴클레오타이드에 첫 번째 폴리뉴클레오타이드를 결찰하는 방법을 제공한다: (a) 반-고체 지지체에 또는 그 위에 고정된 첫 번째 폴리뉴클레오타이드를 제공하는 단계; (b) 첫 번째 폴리뉴클레오타이드를 두 번째 폴리뉴클레오타이드 및 결합 효소와 접촉시키는 단계; 및 (c) 첫 번째 및 두 번째 폴리뉴클레오타이드를 결합 효소의 존재 및 결찰을 위한 적당한 조건하에 유지하는 단계. 적당한 조건이란 사용될 특정 결합 효소에 대한 적절한 완충제, 보조인자, 온도, 시간 등을 제공하는 것을 포함한다. 바람직한 실시예에서 반-고체 지지체는 아크릴아미드 겔과 같은 겔이다. 한층 더 바람직한 실시예에서, 첫 번째 폴리뉴클레오타이드는 비드와 같은 지지체에 대한 부착의 결과로서 반-고체 지지체에 또는 그 위에 고정되며, 비드와 같은 지지체는 그 자체로서 지지체 매트릭스에 부분적으로 또는 완전하게 삽입되어 반-고체 지지체에 또는 그 위에 고정된다. 이와 달리, 첫 번째 폴리뉴클레오타이드는 아크리다이트(acrydite) 부분과 같은 결합을 통하여 반-고체 지지체에 직접 부착될 수 있다. 그 결합은 공유적이거나 비공유적(예컨대 비오틴-아비딘 상호작용을 통하여)이다. 미국 특허 제 6,511,803호에는 발명의 바람직한 지지체, 즉 폴리아크릴아미드 겔에 핵산 분자를 부착하기 위해 사용될 수 있는 다양한 방법이 설명되어 있다.

<194> 본 발명은 나아가 다음 단계들로 이루어지는 폴리뉴클레오타이드의 절단 방법을 제공한다: (a) 반-고체 지지체에 또는 그 위에 고정된, 잘라지기 쉬운 연쇄를 포함하고 있는 폴리뉴클레오타이드를 제공하는 단계; (b) 절단제와

폴리뉴클레오티드를 접촉시키는 단계; 및 (c) 절단에 적당한 조건하에서 절단제의 존재하에 폴리뉴클레오티드를 유지하는 단계. 적당한 조건은 특정 절단제에 대해 적절한 완충제, 온도, 시간 등을 제공하는 것을 포함한다. 바람직한 실시예에서 반-고체 지지체는 아크릴아미드 겔과 같은 겔이다. 보다 더 바람직한 실시예에서 폴리뉴클레오티드는 비드와 같은 지지체에 대한 부착의 결과로서 반-고체 지지체에 고정되며, 비드와 같은 지지체는 그 자체로서 반-고체 지지체에 고정된다. 이와 달리, 폴리뉴클레오티드는 아크리다이트 부분과 같은 결합을 통하여 반-고체 지지체에 직접 부착될 수 있다. 그 결합은 공유적이거나 비공유적(예컨대 비오틴-아비딘 상호작용을 통하여)이다.

<195> Macevicz는 특정 서열을 가지는 단일 주형 종을 서열화하는 것을 개시하였다. 그러나 그는 상이한 서열을 가지는 다수의 주형을 동시에 서열화하기 위하여 자신의 방법을 동시에 수행하는 가능성에 대해서는 논의하지 않았다. 본 발명자들은 고처리량 방식으로 서열화를 효과적으로 수행하기 위하여 상기에서 설명된 것과 같이 다수의 지지체 (예컨대 비드)를 제조하고, 그로써 각각의 지지체가 그것에 부착된 특정 서열의 주형을 가지게 되고, 각 지지체에 부착된 주형에 대해 동시에 본원에 설명된 방법을 수행하는 것이 바람직하다는 것을 발견하였다. 이 접근법의 특정 실시예에서 다수의 그런 지지체는 슬라이드와 같은 평면 기판에 또는 그 위에 배열된다. 어떤 실시예에서 지지체는 겔에 또는 그 위에 배열된다. 지지체는 무작위 방식으로 배열될 수 있다. 즉 기판 위의 각 지지체의 위치는 예정되어 있지 않다. 지지체는 규칙적으로 공간 간격을 두고 위치하거나 가로와 세로 등으로 순서가 정해진 배열로 위치할 필요는 없다. 바람직하게는 지지체는 많은 또는 대부분의 지지체로부터 개별적인 신호를 검출할 수 있는 밀도로 배열된다. 어떤 바람직한 실시예에서 지지체는 주로 단일 초점의 평면에 분포된다. 동일한 서열의 주형이 부착된 다수의 지지체가 품질 조절을 목적으로 포함될 수 있다. 서열화 반응은 각각의 지지체에 부착된 주형에 대해 동시에 수행된다.

<196> 신호는 다양한 영상화 양식을 포함한, 다양한 수단들 중 어떠한 것을 사용하여 수집될 수 있다. 바람직하게도, 검출 전에 서열화가 기판 (예컨대 기판 위에 위치한 반-고체 지지체에 삽입된 비드) 위에 배열된 마이크로입자에 대해 수행되는 실시예의 경우 영상화 장치는 1 μ m 이하의 해상도를 가진다. 예를 들어 CCD 카메라가 고정되어 있는 스캐닝 현미경, 또는 충분한 해상도를 가지는 마이크로배열 스캐너가 사용될 수 있다. 이와 달리, 비드는 형광 검출을 위해 장착된 현미경에 부착된 유동 셀 또는 유체공학 워크스테이션을 통과할 수 있다. 신호를 수집하기 위한 다른 방법은 섬유 광학 번들을 포함한다. 적절한 영상 획득 및 처리 소프트웨어가 사용될 수 있다.

<197> 발명의 어떤 실시예에서 서열화는 마이크로 유체공학 장치에서 수행된다. 예를 들어 주형이 부착된 비드를 장치 안에 로딩하고 시약은 그곳을 통과하여 흘릴 수 있다. 예컨대 PCR을 사용하는 주형 합성이 또한 장치에서 수행될 수 있다. 미국 특허 제 6,632,655호는 적당한 마이크로 유체공학 장치의 실례를 소개한다.

<198> D. 상이한 개시 올리고뉴클레오티드를 사용하는 제-개시를 이용한 서열화

<199> 본 발명의 바람직한 실시예에서 첫 번째 개시 올리고뉴클레오티드를 연장함으로써 생성된 연장된 가닥은 충분한 수의 사이클 후에 주형으로부터 제거되고, 두 번째 개시 올리고뉴클레오티드가 결합 영역에 어닐(anneal)되어 연장, 결합 및 검출의 사이클이 이어진다. 그 과정은 어떠한 수의 상이한 개시 올리고뉴클레오티드를 사용하여 반복된다. 연장 프로브가 절단되는 실시예에서, 바람직하게는 사용된 상이한 개시 올리고뉴클레오티드의 수(따라서, 반응의 수)는 프로브의 말단 부분의 방출 후에 주형에 혼성화된 채로 남아있는 연장 프로브의 부분의 길이와 같다. 그러므로 이 실시예에 따르면, 서열 정보 (예컨대 각 뉴클레오티드의 순서와 정제)는 단일 지지체에 부착된 주형으로부터 얻어질 수 있는 한편, 여전히 연속적인 뉴클레오티드가 각 사이클에서 확인되었다면 필요로 하는 것보다 실질적으로 더 적은 수의 사이클을 사용하여 서열로 실제 판독될 수 있다.

<200> 개시 올리고뉴클레오티드가 순차적으로 동일한 주형에 결합하는 실시예는 Macevicz에 의해 교시된 방법과 같은, 주형을 다수의 분취액으로 나누는 것을 필요로 하는 접근법 이상의 장점들을 갖는다. 예를 들어 개시 올리고뉴클레오티드를 동일한 주형에 적용하면 다수의 분취액으로부터 얻어진 데이터를 계속해서 조합하고, 나중에도 조합해야 할 필요가 없어진다. 지지체가 무작위 형식으로 배열되어서 개별적인 지지체의 위치가 예정되지 않은 실시예에서, 각각이 동일한 서열의 주형이 부착된 다수의 지지체로부터 얻어지는 부분적인 서열 정보를 신뢰할 만한 수준으로 조합하는 것은 어렵거나 불가능할 것이다.

<201> E. 단일 주형에 대한 각 사이클에서 다중 뉴클레오티드의 확인

<202> Macevicz는 연장, 결합 및 검출의 각 사이클에서 주형의 단일한 뉴클레오티드를 확인하는 것을 교시하였다. 그러나 본 발명자들은 그 방법이 각 사이클의 주형에 있는 다수의 뉴클레오티드의 확인이 가능하도록 변형될 수 있음을 발견하였다. 이 경우에 연장 프로브는 연장된 이중나선에 접해 있는 둘 또는 그 이상, 바람직하게는 연

속적인 뉴클레오타이드의 정체가 표지로부터 결정될 수 있도록 표지된다. 달리 표현하자면, 연장 프로브의 서열 결정 부분은 단일 뉴클레오타이드 이상이고 전형적으로는 근접한 뉴클레오타이드, 바로 옆에 있는 뉴클레오타이드, 및 가능한 하나 또는 그 이상의 추가의, 바람직하게는 연속적인 뉴클레오타이드를 포함하고, 이것들은 모두 주형에 특이적으로 혼성화한다. 예를 들어 염기 A, G, C, 및 T를 확인하기 위하여 4개의 표지를 사용하기보다는 16개의 구별가능하게 표지된 프로브 또는 프로브 조합이 16개의 가능한 디뉴클레오타이드 AA, AG, AC, AT, GA, GG, GC, GT, CA, CG, CC, CT, TA, TG, TC 및 TT를 확인하기 위해 사용된다. 구별가능하게 표지된 각각의 연장 프로브의 서열 결정 부분은 이들 디뉴클레오타이드의 하나에 상보한다. 보다 많은 표지를 활용하는 유사한 방법은 각 사이클에서 더 긴 뉴클레오타이드 서열의 확인을 가능하게 한다.

<203> F. 표지

<204> 본원에서 용어 "표지"는 그것에 의하여 상이한 종 (예컨대 상이한 말단 뉴클레오타이드를 가지고 있는 프로브들) 이 서로 간에 구별될 수 있는 프로브가 부착된 또는 그것들과 결합된 어떠한 검출가능한 부분 또는 다수의 검출 가능한 부분을 나타내는 광범위한 의미로 사용된다. 그러므로 표지와 특이한 검출가능한 부분 사이에 일대일 대응성이 있지 않아도 된다. 예를 들어 다수의 검출가능한 부분은 단일 프로브에 부착하여 프로브가 상이한 검출 가능한 부분을 가지거나 그것에 검출가능한 부분의 세트가 부착된 프로브로부터 구별되는 것을 허용하는 조합된 신호가 생성될 수 있다. 예를 들어 검출가능한 부분의 조합은 "조합 다색 코딩(combinatorial multicolor coding)"으로서 언급되는 표지화 도식에 따라 사용될 수 있다 (미국 특허 제 6,632,609호 및 Speicher, et al., *Nature Genetics*, 12:368-375, 1996).

<205> 본 발명의 프로브는 다양한 방법, 이를테면 형광 또는 화학발광 부분, 비색 부분, 기질과 접촉할 때 검출가능한 신호를 생성하는 효소적 부분 등의 부착에 의하여 직접 또는 간접적으로 표지될 수 있다. Macevicz는 프로브가 예컨대 Menchen 등의 미국 특허 제 5,188,934호와 Begot 등의 PCT 출원 제 PCT/US90105565호에 의해 개시된 바와 같이 형광 염료로 표지될 수 있음을 교시하였다. 용어 "형광 염료" 및 "형광발색단"은 본원에서 규정된 여기 과정에서 빛 에너지를 흡수하고 상이한 과정에서 빛 에너지를 방출하는 부분을 나타내기 위해 사용된다. 바람직하게는 표지는 스펙트럼으로 분해될 수 있는 주어진 프로브의 혼합물과 함께 사용하기 위해 선택된다. 본원에서 사용된 것과 같이 "스펙트럼으로 분해될 수 있는"은 표지가 작동 조건하에서 그것들의 스펙트럼 특성, 특히 형광 방출 파장을 토대로 구별될 수 있음을 의미한다. 예를 들어 하나 또는 그 이상의 말단 뉴클레오타이드의 정체는 최대 빛 방출 강도의 뚜렷한 파장과, 또는 아마도 상이한 파장에서의 강도의 비율과 상관관계가 있을 수 있다. 표지의 검출 및 확인을 위해 사용되는 표지의 스펙트럼 특성(들)은 본원에서 "색"으로 언급된다. 표지는 자주 특이한 스펙트럼 특성, 예컨대 단일한 검출가능한 부분으로 구성되는 표지의 경우 최대 방출 강도의 빈도, 또는 다수의 검출가능한 부분으로 구성되는 표지들의 경우 방출 피크의 빈도를 토대로 확인된다는 것이 인지될 것이다.

<206> 바람직하게는 4개의 스펙트럼으로 분해가능한 각 형광 염료와 프로브의 4개의 말단 뉴클레오타이드 사이에 일대일 대응이 가능한 4개의 프로브가 제공된다. 스펙트럼으로 분해가능한 염료의 세트는 특허 제 4,855,225호와 제 5,188,934호; 국제 출원 제 PCT/US90/05565; 및 Lee 등의 *Nucleic Acids Researchs*, 20: 2471-2483 (1992)에서 개시된다. 어떤 실시예에서, FITC, HEXTM, Texas Red, 및 Cy5로 구성되는 세트가 바람직하다. 수많은 적당한 형광 염료가, 예를 들면 Molecular Probes사(Eugene OR)에 의해 시판중이다. 형광 염료의 구체적인 실례로는, 그것들에 한정되는 것은 아니지만 다음과 같은 것들이 있다: Alexa Fluor 염료 (Alexa Fluor 350, Alexa Fluor 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 633, Alexa Fluor 660 및 Alexa Fluor 680), AMCA, AMCA-S, BODIPY 염료 (BODIPY FL, BODIPY R6G, BODIPY TMR, BODIPY TR, BODIPY 530/550, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY 630/650, BODIPY 650/665), CAL 염료, 카르복시로다민 6G, 카르복시-X-로다민 (ROX), 케스케이드 블루, 케스케이드 옐로우, 시아닌 염료 (Cy3, Cy5, Cy3.5, Cy5.5), 단실, 다폭실, 디알킬아미노쿠마린, 4',5'-디클로로-2',7'-디메톡시-플루오레신, DM-NERF, 에오신, 에리트로신, 플루오레신, FAM, 히드록시쿠마린, IRDyes (IRD40, IRD 700, IRD 800), JOE, 리사민 로다민 B, 마리나 블루, 메톡시쿠마린, 나프토플루오레신, 오레곤 그린 488, 오레곤 그린 500, 오레곤 그린 514, 오이스터 염료, 태평양 블루, PyMPO, 파이렌, 로다민 6G, 로다민 그린, 로다민 레드, 로돌 그린, 2',4',5',7'-테트라-브로모술폰-플루오레신, 테트라메틸-로다민(TMR), 카르복시테트라메틸로다민 (TAMRA), 텍사스 레드, 텍사스 레드-X. 추가의 설명에 대해서는 *The Handbook of Fluorescent Probes and Research Products*, 9th ed., Molecular Probes, Inc.를 참조한다.

<207> 어떤 형광기는 그것들 자체가 직접적으로 검출되기보다는, 비방사성 형광 공명 에너지 전달 (FRET) 과정에서 에

너지를 다른 기에 전달하고, 두 번째 기가 검출된 신호를 생성한다. 퀀처(quencher)의 사용 또한 본 발명의 범주 내에 있다. "퀀처"는 가까이 위치하고 있을 때 여기된 형광 표지의 에너지를 흡수할 수 있고 가시광의 방출 없이 그 에너지를 방산(dissipating)할 수 있는 부분을 말한다. 퀀처의 실례로는, 그것들에 한정되는 것은 아니지만 DABCYL (4-(4'-디메틸아미노페닐아조)벤조산)숙신이미딜 에스테르, 디아릴로다민 카르복실산, 숙신이미딜 에스테르(QSY-7), 및 4',5'-디니트로플루오레신 카르복실산, 숙신이미딜 에스테르 (QSY-33)(모두 Molecular Probes사 제품), 퀀처1 (Q1; Epoch사 제품), 또는 "블랙홀 퀀처" BHQ-1, BHQ-2, 및 BHQ-3 (BioSearch사 제품)이 있다.

<208> 상기에서 언급한 다양한 검출가능한 부분 외에 본 발명은 또한 스펙트럼으로 분해가능한 양자 도트, 금속 나노입자 또는 나노클러스터, 등의 사용을 포함하는데, 그것들은 직접 올리고뉴클레오타이드 프로브에 부착될 수 있거나 중합체 매트릭스에 삽입되거나 그것과 결합한 후 프로브에 부착될 수 있다. 상기에서 언급한 바와 같이, 검출가능한 부분은 그 자체로서 직접 검출되지 않아도 좋다. 예를 들어 그것들은 검출되는 기관 위에서 작용할 수 있고 또는 검출가능하게 되도록 변형이 필요할 수도 있다.

<209> 상기에서 설명된 것처럼 본 발명의 어떤 실시예에서 표지는 다수의 검출가능한 부분으로 구성된다. 이들 검출가능한 부분으로부터 조합된 신호는 프로브를 확인하기 위해 사용되는 색을 발산한다. 예를 들어 특정 서열의 "자주색" 프로브는 "푸른색"과 "적색"의 검출가능한 부분이 프로브에 부착함으로써 구성될 수 있다. 이와 달리, 뚜렷한 색을 동일한 서열을 가지지만 상이한 검출가능한 프로브로 표지되어 혼합된 프로브를 생성하는 2종의 프로브를 조합함으로써 생성될 수 있다. 그러므로 특정 서열의 "자주색" 프로브는 그 서열을 가지는 2종의 프로브를 구성함으로써 제조될 수 있다. "적색" 검출가능한 부분이 첫 번째 종에 부착되고, "푸른색" 검출가능한 부분이 두 번째 종에 부착된다. 이들 2종의 분취액이 혼합된다. 분취액을 상이한 비율로 혼합함으로써 다양한 색의 농도가 약간씩 다른 다양한 자주색이 생성될 수 있다. 이런 접근법은 많은 장점을 부여한다. 먼저, 이 접근법으로 더 적은 수의 검출가능한 부분을 사용하여 다수의 구별가능한 프로브를 생성할 수 있다. 둘째, 프로브를 혼합함으로써 특정한 검출가능한 부분과 특정한 뉴클레오타이드 사이의 상호작용으로부터 유발될 수 있는 편향성(bias)을 감소시키는 것을 보조할 수 있는 중복도(degree of redundancy)를 제공할 수 있다.

<210> 본 발명의 어떤 실시예에서, 검출가능한 부분은 절찰 및 검출 후에 검출가능한 부분의 제거를 가능하게 하는 절단가능한 연쇄에 의해 올리고뉴클레오타이드 연장 프로브에 있는 뉴클레오타이드에 부착된다. 다양한 상이한 절단가능한 연쇄 중 어느 것이라도 사용될 수 있다. 본원에서 사용되는 용어 "절단가능한 연쇄"는 검출가능한 부분을 뉴클레오타이드에 결합시키는 화학적 부분을 말하며, 그것은 필요에 따라, 본질적으로는 뉴클레오타이드를 변경시키거나 그것이 부착하는 핵산 분자 없이 뉴클레오타이드로부터 검출가능한 부분을 제거하기 위해 절단될 수 있다. 절단은 예를 들면 연쇄의 성질에 따라 산 또는 염기 처리에 의해, 또는 연쇄의 산화 또는 환원에 의해, 또는 빛 처리 (광학적 절단)에 의해 이루어질 수 있다. 절단가능한 연쇄와 절단제의 실례는 문헌에 설명되어 있다 (Shirnkus et al., 1985, Proc. Natl. Acad. Sci. USA 82:2593-2597; Soukup et al., 1995, Bioconjug. Chem. 6:135-138; Shimikus et al., 1986, DNA 5:247-255; and Herman and Fenn, 1990, Meth. Enzymol. 184:584-588).

<211> 예를 들어 미국 특허 제 6,511,803호에서 설명된 바와 같이, 이황화 연쇄는 환원될 수 있고 그로써 디티오프레이톨(DTT)과 같은 티올 화합물 환원제를 사용하여 절단될 수 있다. 형광발색단은 반응성 아릴 아미노기를 가진 있는 뉴클레오타이드 (예컨대 dCTP)와 같이, 포함(conjugation)에 사용할 수 있는 술폰히드릴기(SH)를 가지는 형광발색단 (예컨대 SH기를 가지고 있는 시아닌 5 또는 시아닌 3 형광발색단; New England Nuclear--DuPont)이 활용가능하다. 반응성 피리딜디티올은 술폰히드릴기와 반응하여 술폰히드릴(sulfhydryl) 결합을 형성하고, 그것은 디티오프레이톨(dithiothreitol)과 같은 환원제로 절단될 수 있다. NHS-에스테르 유형의 이작용성 교차결합제(Pierce)가 반응성 아릴 아미노기를 포함하는 데옥시뉴클레오타이드를 피리딜디티올기에 결합시키기 위해 사용될 수 있고, 그것은 계속해서 형광발색단 상에 있는 SH와 반응하여 본 발명의 방법에 유용한 절단가능한 뉴클레오타이드-형광발색단 복합체인 이황화 결합이 형성된다. 이와 달리, 뉴클레오타이드와 형광발색단 사이의 시스-글리콜 연쇄가 과요오드산염에 의해 절단될 수 있다. 다양한 절단가능한 연쇄가 미국 특허 제 6,664,079호와 제 6,632,655호, 미국 공개 출원 제 20030104437호, 제 WO 04/18497호 및 제 WO 03/48387호에서 설명된다.

<212> 본 발명의 다른 실시예에서는, 빛과 같은 전자기 에너지에 노출시켜 (광학 표백) 검출할 수 없게 만들 수 있는 검출가능한 부분이 사용된다.

<213> 절단가능한 연쇄에 의해 프로브에 부착되는 표지를 가지거나 또는 광표백될 수 있는 표지를 가지는 연장 프로브를 사용하는 본 발명의 그런 실시예에서, 서열화 방법은 전형적으로 절찰 및 표지 검출이 수행된 후에 하나 또

는 그 이상의 사이클에서 절단 또는 광표백의 단계를 포함할 것이다. 상기에서 언급된 바와 같이 올리고뉴클레오타이드 연장 프로브에 존재하는 잘라지기 쉬운 연쇄의 절단은 완료될 필요는 없다(즉, 100% 이하의 새롭게 결합된 프로브는 그들이 결합되었던 사이클 내에서 절단될 수 있다). 그러한 프로브는 일반적으로 비-연장성 말단을 포함하거나 캡핑되어 있기 때문에, 그것들은 연속적인 사이클에 사용될 수 없다. 그러나 프로브를 절단하지 못한다는 것은 표지가 그 프로브가 결합된 주형 분자와 결합한 채로 남아있고, 그것은 후속적인 사이클에서 노이즈를 증가시킬 수 있는 바탕값 수준(즉 바탕값 형광)에 기여한다. 표지를 제거하기 위해 절단 또는 광표백하거나 표지가 검출될 수 없도록 통합하는 것은 이 바탕값을 감소시키고 신호 대 노이즈 비율을 개선한다. 절단 또는 광표백은 매 사이클마다 자주, 또는 그보다 덜한 빈도로, 예컨대 격일로, 매 3일마다, 또는 매 5일마다 또는 더 많은 사이클 주기로 수행될 수 있다. 본 발명의 어떤 실시예에서, 실제로 절단가능한 링커를 절단하기 위해 어떠한 단계를 추가할 필요는 없다. 예를 들어 DTT와 같은 절단제는 이미 미결합 연장 프로브를 제거하기 위해 사용될 수 있는 세척 완충액에 존재할 수 있다.

<214> G. 바람직한 잘라지기 쉬운 연쇄

<215> 본 발명자들은 최소한 하나의 포스포로티올레이트 연쇄를 가지는 연장 프로브가 특히 연장, 결합, 검출 및 절단의 연속적인 사이클에 의해 서열화하는 방법의 실시예에 유용하다는 것을 발견하였다. 그러한 연쇄에서 인산이에스테르 결합의 가교 산소 원자들 중 하나는 황 원자에 의해 대체된다. 포스포로티올레이트 연쇄는 도 4A에서 도시된 바와 같은 5'-S-포스포로티올레이트 연쇄(3'-O-P-S-5')이거나 도 4B에서 도시된 바와 같은 3'-S-포스포로티올레이트 연쇄(3'-S-P-O-5')이다. 3'-O-P-S-5' 또는 3'-S-P-O-5'로서 표시되는 연쇄에 있는 인 원자는 도 4A와 도 4B에서 알 수 있는 바와 같이 2개의 비-가교 산소 원자에 부착될 수 있다(전형적인 인산이에스테르 결합에서와 같이)는 것이 인지되어야 한다. 이와 달리, 인 원자는 다양한 다른 원자 또는 기, 예컨대 어떠한 S, CH₃, BH₃ 등에 부착될 수 있다. 그러므로 본 발명의 한 측면은 포스포로티올레이트 연쇄를 포함하는 표지된 올리고뉴클레오타이드 프로브이다. 프로브가 본원에서 설명된 서열화 방법에 특별히 사용될 수 있는 한편, 프로브는 또한 다양한 다른 목적을 위해서도 사용될 수 있다. 특히 본 발명은 (i) 형태가 5'-O-P-O-X-O-P-S-(N)_kN_B^{*}-3'인 올리고뉴클레오타이드; 및 (ii) 형태가 5'-N_B^{*}(N)_k-S-P-O-X-3'인 올리고뉴클레오타이드를 제공한다. 이들 각각의 프로브에서 N은 어떠한 뉴클레오타이드를 나타내고, N_B는 절단에 의해 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, X는 뉴클레오타이드를 나타내며, k는 1과 100 사이의 수이다. 어떤 실시예에서 k는 1과 50 사이, 1과 30 사이, 1과 20 사이, 예컨대 4와 10 사이이며, 단, 검출가능한 부분은 N_B 대신에, 또는 그것 외에 (N)_k의 어떠한 뉴클레오타이드상에 존재할 수 있어야 한다. 이들 프로브 중 어떠한 것에 있는 말단 뉴클레오타이드는 포스페이트기 또는 히드록실기를 포함할 수도 있고 포함하지 않을 수도 있다. 나아가 인 원자는 일반적으로 바람직한 실시예에서 2개의 추가적(비-가교) 산소 원자에 부착될 것이다.

<216> 5'-S-포스포로티올레이트 또는 3'-S-포스포로티올레이트 연쇄를 함유하는 올리고뉴클레오타이드를 합성하기 위한 방법은 당해 기술분야에 알려져 있으며, 이들 방법중 어떤 것은 자동 고상 올리고뉴클레오타이드 합성법을 따를 수 있다. 합성 과정은 문헌에서 찾아볼 수 있다(Cook, AF, J. Am. Chem. Soc., 92:190-195, 1970; Chladek, S. et al., J. Am. Chem. Soc., 94:2079-2084, 1972; Rybakov, VN, et al., Nucleic Acids Res., 9:189-201, 1981; Cosstick, R. and VyIe, JS, J. Chem. Soc. Chem. Commun., 992-992, 1988; Mag, M., et al., Nucleic Acids Res., 19(7):1437-1441, 1991; Xu, Y, and Kool, ET, Nucleic Acids Res., 26(13):3159-3164, 1998; Cosstick, R. and Vyle, JS, Tetrahedron Lett., 30:4693-4696, 1989; Cosstick, R. and Vyle, JS, Nucleic Acids Res., 18:829-835, 1990; Sun, SG and Piccirilli, JA, Nucl. Nucl., 16:1543-1545, 1997; Sun SG, et al., RNA, 3:1352-1363, 1997; Vyle, JS, et al., Tetrahedron Lett., 33:3017-3020, 1992; Li, X., et al., J. Chem. Soc. Perkin Trans., 1:2123-22129, 1994; Liu, XH and Reese, CB, Tetrahedron Lett., 37:925-928, 1996; Weinstein, LB, et al., J. Am. Chem. Soc., 118:10341-10350, 1996; 및 Sabbagh, G., et al., Nucleic Acids Res., 32(2):495-501, 2004). 또한 본 발명자들은 새로운 합성 방법을 개발하였다. 예를 들어, 도 7에 dA의 3'-포스포로아미다이트에 대한 합성 도식이 도시되어 있다. 유사한 도식이 dG의 3'-포스포로아미다이트를 합성하기 위해 사용될 수 있다. 이들 포스포로아미다이트는 퓨린 뉴클레오시드와 결합된 3'-S-포스포로티올레이트 연쇄를 함유하고 있는 올리고뉴클레오타이드를, 예컨대 자동 DNA 합성기를 사용하여 합성하기 위해 사용될 수 있다.

<217> 포스포로티올레이트 연쇄는 다양한 금속-함유제를 사용하여 절단될 수 있다. 금속은 예를 들면 Ag, Hg, Cu, Mn,

Zn, 또는 Cd이다. 바람직하게는 제제는 Ag^+ , Hg^{++} , Cu^{++} , Mn^{++} , Zn^+ 또는 Cd^+ 양이온을 제공하는 수용성 염이다 (다른 산화 상태의 이온을 제공하는 염도 사용될 수 있다). I_2 도 또한 사용될 수 있다. 은-함유 염, 예컨대 질산 은 (AgNO_3), 또는 Ag^+ 이온을 제공하는 다른 염이 특히 바람직하다. 적당한 조건은 예를 들면 50mM AgNO_3 를 약 22 내지 37°C에서 약 10분 이상, 예컨대 30분을 포함한다. 바람직하게는 pH는 4.0과 10.0 사이이며, 더욱 바람직하게는 5.0과 9.0 사이, 예컨대 약 6.0과 8.0 사이, 예컨대 약 7.0이 좋다 (Mag, M., et al., Nucleic Acids Res., 19(7):1437-1441, 1991). 예시적인 프로토콜을 하기 실시예 1에 제공한다.

<218> 5'→3' 방향으로의 서열화는 3'-O-P-S-5' 연쇄를 함유하는 연장 프로브를 사용하여 수행될 수 있다. 도 5A는 형태 5'-O-P-O-X-O-P-S-NNNN N_B^* -3' (여기서, N은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며 (예를 들어 N_B 는 3' 히드록실기가 없거나 부착된 차단 부분을 가지는 뉴클레오타이드이다), *는 검출가능한 부분을 나타내고, X는 검출가능한 부분에 상응하는 정체를 가지는 뉴클레오타이드를 나타낸다)인 연장 프로브를 사용하는 혼성화, 결찰 및 절단의 단일 사이클을 도시한다. 이와 달리, 대다수의 차단 부분 중 어느 하나가 3' 말단 뉴클레오타이드에 부착되어 다중 결합을 방해할 수 있다. 예를 들어 벌키(bulky) 기가 뉴클레오타이드의 당 부분, 예컨대 2' 또는 3'에 부착되면, 결합이 방해될 것이다. 형광 표지는 적절한 벌키 기로서 작용할 수 있다.

<219> 결합 영역(40)과 미지 서열의 폴리뉴클레오타이드 영역(50)을 함유하는 주형은 지지체, 예컨대 비드에 부착된다. 바람직한 실시예에서, 도 5A에서 알 수 있는 것과 같이 결합 영역은 지지체에 부착하는 지점으로부터 주형의 반대쪽 말단에 위치한다. 연장가능한 말단 (이 경우 유리 3' OH기)을 가지는 개시 올리고뉴클레오타이드(30)이 결합 영역(40)에 어닐린다. 연장 프로브(60)는 폴리뉴클레오타이드 영역(50)에서 주형에 혼성화한다. 뉴클레오타이드 X는 주형의 미지의 뉴클레오타이드 Y와 상보하는 염기쌍을 형성한다. 연장 프로브(60)는 개시 올리고뉴클레오타이드에 결찰된다 (예컨대 T4 결합 효소를 사용하여). 결찰 후에 연장 프로브(60)에 부착된 표지가 검출된다 (표시되지 않음). 표지는 뉴클레오타이드 X의 정체에 상응한다. 그러므로 뉴클레오타이드 Y는 뉴클레오타이드 X에 상보하는 뉴클레오타이드로서 확인된다. 그런 다음 연장 프로브(60)가 포스포로티올레이트 연쇄에서 절단되어 (예컨대 AgNO_3 또는 Ag^+ 이온을 제공하는 다른 염을 사용하여), 그 결과 연장된 이중나선이 형성된다. 절단으로 연장된 이중나선의 3' 말단에 포스페이트기가 남는다. 포스파타제 처리를 사용하여 연장된 이중나선 상에 연장가능한 프로브 말단을 생성한다. 이런 과정이 원하는 수의 사이클 동안 반복된다.

<220> 바람직한 실시예에서 서열화는 3'-S-P-O-5' 연쇄를 함유하는 연장 프로브를 사용하여 3'→5' 방향으로 수행된다. 도 5B는 형태 5'- N_B^* -NNNN-S-P-O-X-3' (여기서 N은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며 (예를 들어 N_B 는 3' 히드록실기가 없거나 부착된 차단 부분을 가지는 뉴클레오타이드이다), *는 검출가능한 부분을 나타내고, X는 검출가능한 부분에 상응하는 정체를 가지는 뉴클레오타이드를 나타낸다)의 연장 프로브를 사용하는 혼성화, 결찰 및 절단의 단일 사이클을 도시한다.

<221> 결합 영역(40)과 미지 서열의 폴리뉴클레오타이드 영역(50)을 함유하는 주형은 지지체, 예컨대 비드에 부착된다. 바람직한 실시예에서, 도 5B에서 알 수 있는 것과 같이 결합 영역은 지지체에 부착하는 지점으로부터 주형의 반대쪽 말단에 위치한다. 연장가능한 말단 (이 경우 유리 5' 포스페이트기)을 가지는 개시 올리고뉴클레오타이드(30)이 결합 영역(40)에 어닐린다. 연장 프로브(60)는 폴리뉴클레오타이드 영역(50)에서 주형에 혼성화한다. 뉴클레오타이드 X는 주형의 미지의 뉴클레오타이드 Y와 상보하는 염기쌍을 형성한다. 연장 프로브(60)는 개시 올리고뉴클레오타이드에 결찰된다 (예컨대 T4 결합 효소를 사용하여). 결찰 후에 연장 프로브(60)에 부착된 표지가 검출된다 (표시되지 않음). 표지는 뉴클레오타이드 X의 정체에 상응한다. 그러므로 뉴클레오타이드 Y는 뉴클레오타이드 X에 상보하는 뉴클레오타이드로서 확인된다. 그런 다음 연장 프로브(60)가 포스포로티올레이트 연쇄에서 절단되어 (예컨대 AgNO_3 또는 Ag^+ 이온을 제공하는 다른 염을 사용하여), 그 결과 연장된 이중나선이 형성된다. 절단으로 연장된 이중나선의 5' 말단에 모노포스페이트기가 남고 따라서 연장가능한 말단을 생성하기 위하여 추가의 단계를 수행할 필요는 없다. 이런 과정이 원하는 수의 사이클 동안 반복된다.

<222> 이 도식의 많은 변형이 사용될 수 있음이 인지될 것이다. 예를 들어 프로브는 6 뉴클레오타이드보다 짧거나 길 수 있고; 표지는 3' 말단 뉴클레오타이드 상에 있을 필요가 없으며; P-S 연쇄는 어떠한 2개의 인접한 뉴클레오타이드 사이에 위치할 수 있는 등이다. 상기에서 설명된 실시예에서 연속적인 사이클의 연장, 결찰, 검출, 및 절단은

인접하게 위치한 뉴클레오타이드의 확인을 초래한다. 그러나 연장 프로브의 말단부 (즉 결찰이 일어나는 곳과 반대쪽 말단)에 더 가깝게 P-S 연쇄를 놓음으로써, 연속적으로 확인된 뉴클레오타이드는 상기에서 설명되고 도 1과 6에서 알 수 있는 바와 같이, 주형을 따라 일정 간격으로 공간배치될 것이다.

<223> 도 6A-6F는 단일 주형에 대하여 연속적으로 수행된 여러 서열화 반응을 보다 상세하게 도표로 예시한 도면이다. 서열화는 3'-S-P-O-5'를 함유하는 연장 프로브를 사용하여 3'→5' 방향으로 수행된다. 각각의 서열화 반응은 다중 사이클의 연장, 결찰, 검출, 및 절단을 포함한다. 반응은 주형의 상이한 부분에 결합하는 개시 올리고뉴클레오타이드를 활용한다. 연장 프로브의 길이는 8 뉴클레오타이드이며 프로브의 3' 말단으로부터 계수하여 여섯 번째와 일곱 번째 뉴클레오타이드 사이에 위치한 포스포로티올레이트 연쇄를 함유한다. 뉴클레오타이드 2-6은 스페이서로 작용하여 각 반응으로 주형을 따라 일정간격으로 배치된 다수의 뉴클레오타이드의 확인이 가능해진다. 다수의 반응을 시리즈로 수행하고 각 반응으로부터 얻어진 부분적인 서열 정보를 적절하게 조합함으로써 주형의 부분의 완전한 서열이 결정된다.

<224> 도 6A는 주형의 어댑터 서열 (상기에서는 결합 영역으로서 언급됨)에 혼성화하여 연장가능한 이중나선을 제공하는 첫 번째 개시 올리고뉴클레오타이드 (도 6A-6F에서는 프라이머로서 언급됨)를 사용하는 개시를 도시한다. 도 6B-6D는 주형의 매 여섯 번째 염기가 판독되는 여러 사이클의 뉴클레오타이드 확인을 도시한다. 도 6B에서, 주형 서열의 첫 번째 미지 뉴클레오타이드에 상보하는 3' 말단 뉴클레오타이드를 가지는 첫 번째 연장 프로브가 주형에 결합하고, 프라이머의 연장가능한 말단에 결찰된다. 연장 프로브에 부착된 표지가 3' 말단 뉴클레오타이드로서 A를 가지는 것으로 프로브를 확인하게 되고, 그로써 주형 서열에 있는 첫 번째 미지의 뉴클레오타이드를 A로서 확인한다. 도 6C는 AgNO₃를 사용하여 포스포로티올레이트 연쇄에서 연장 올리고뉴클레오타이드를 절단하는 것과 표지가 부착되는 연장 프로브의 부분을 방출하는 것을 도시한다. 도 6D는 연장, 결찰, 및 절단의 추가의 사이클을 도시한다. 프로브가 길이가 5 뉴클레오타이드인 스페이서를 함유하기 때문에 서열화 반응은 주형의 매 여섯 번째 뉴클레오타이드를 확인하게 된다.

<225> 원하는 수의 사이클을 수행한 후에 첫 번째 개시 올리고뉴클레오타이드를 포함하는 연장된 가닥이 제거되고, 첫 번째 개시 올리고뉴클레오타이드가 결합한 곳과는 상이한 부분의 결합 영역에 결합하는 두 번째 개시 올리고뉴클레오타이드가 주형에 혼성화한다. 도 6E는 두 번째 개시 올리고뉴클레오타이드로 개시 반응이 수행된 후, 여러 사이클의 뉴클레오타이드 확인이 이루어지는 두 번째 서열화 반응을 도시한다. 도 6F는 여러 사이클의 뉴클레오타이드 확인이 수반되는 세 번째 개시 올리고뉴클레오타이드를 사용하는 개시 반응을 도시한다. 두 번째 개시 올리고뉴클레오타이드로부터의 연장으로 첫 번째 서열화 반응에서 확인된 뉴클레오타이드로부터 상이한 "프레임"에 있는 매 여섯 번째의 염기의 확인이 가능해진다.

<226> 포스포로티올레이트 연쇄를 포함하는 연장 프로브가 본 발명의 어떤 실시예에서는 바람직하지만, 다양한 다른 잘라지기 쉬운 연쇄가 유익하게 사용될 수 있다. 예를 들어 자연 발생 핵산에서 발견되는 O-P-O 연쇄의 다수의 변형이 알려져 있다 (Micklefield, J. Curr. Med. Chem., 8:1157-1179, 2001). P-O 결합을 함유하는 본원에서 설명된 어떠한 구조든지 잘라지기 쉬운 P-O 결합을 함유하도록 변형될 수 있다. 예를 들어 NH-P-O 결합은 NH-P-S 결합으로 바뀔 수 있다.

<227> 본 발명의 어떤 실시예에서 연장 프로브는 핵산이 절단제 또는 절단제의 조합에 의한 절단에 민감하도록 만드는 트리거 잔기를 포함하며, 트리거 잔기는 임의로 그런 후에 변형체에 의해 변형된다. 특히 본 발명자들은 DNA 수복에 포함된 효소들이 연장, 결찰, 검출, 및 절단의 연속적인 사이클에 의해 서열화하기 위한 방법의 실시에 사용하기에 유익한 절단제라는 것을 발견하였다. 일반적으로 트리거 잔기, 예컨대 손상된 염기 또는 비염기성 잔기의 연장 프로브 내의 존재는 프로브가 하나 또는 그 이상의 DNA 수복 효소에 의한 절단에 민감하게 만들 수 있고, 임의로 DNA 글리코실라제에 의한 변형이 수반된다. 그러므로 AP 엔도뉴클레아제와 같은 DNA 수복에 포함된 효소에 의한 절단에 대한 기질인 연쇄를 포함하는 연장 프로브는 본 발명에 사용된다. DNA 수복에 포함된 효소, 예컨대 DNA 글리코실라제에 의한 변형의 기질인 잔기들을 함유하는 연장 프로브는 또한 본 발명에서 특별히 사용되며, 그 변형은 프로브를 AP 엔도뉴클레아제에 의한 절단에 민감하게 만든다. 어떤 실시예에서 연장 프로브는 비염기성 잔기를 포함한다. 즉 그것은 퓨린 또는 피리미딘 염기가 부족하다. 비염기성 잔기와 인접한 뉴클레오타이드 사이의 연쇄는 AP 엔도뉴클레아제에 의한 절단에 민감하고 따라서 잘라지기 쉬운 연쇄이다. 발명의 어떤 실시예에서 비염기성 잔기는 2' 데옥시리보스를 포함한다. 어떤 실시예에서 연장 프로브는 손상된 염기를 포함한다. 손상된 염기는 손상된 염기를 제거하는 효소, 예컨대 DNA 글리코실라제에 대한 기질이다. 손상된 염기가 제거된 후, 그 결과의 비염기성 잔기와 인접한 뉴클레오타이드 사이의 연쇄는 AP 엔도뉴클레아제에 의한 절단에 민감하고, 따라서 본 발명에 따르는 잘라지기 쉬운 연쇄로 간주한다.

- <228> 상이한 많은 AP 엔도뉴클레아제가 본 발명에서 절단제로서 사용된다. AP 엔도뉴클레아제의 2개의 주요 부류가 비염기성 잔기에 인접한 연쇄를 절단하는 그것들의 메커니즘을 토대로 구별되었다. 부류 I AP 엔도뉴클레아제, 예컨대 대장균의 엔도뉴클레아제 III (Endo III) 및 엔도뉴클레아제 VIII (Endo VIII) 및 사람 동족체 hNTH1, NEIL1, NEIL2, 및 NEIL3는 AP 잔기의 3' 쪽에 있는 DNA를 절단하여 3' 말단 포스페이트와 5' 말단 포스페이트를 포함하는 3' 부분을 가지는 5' 부분을 생성하는 AP 리아제이다. 부류 II AP 엔도뉴클레아제, 예컨대 대장균의 엔도뉴클레아제 IV (Endo IV) 및 엑소뉴클레아제 III (Exo III)는 AP 부위의 DNA 5'을 절단하여 그 결과 생성되는 단편의 말단에 3' OH와 5' 데옥시리보스 포스페이트를 생성한다 (Doublié, S., et al., *Proc. Natl. Acad. Sci* 101(28), 10284-10289, 2004; Haltiwanger, B.M., et al., *Biochem J.*, 345, 85-89, 2000; Levin, J. and Demple, B., *Nucl. Acids. Res*, 18(17), 1990); 그 외에 다양한 부류 I 및 부류 II AP 엔도뉴클레아제 및 그것들이 DNA로부터 손상된 염기를 제거하고 및/또는 비염기성 잔기를 함유하는 DNA를 절단하는 조건에 대한 추가의 논의에 대해서는 전술한 참고문헌들을 참조한다). 당업자는 다른 유기체 (예컨대 효모)에 존재하고 본 발명에 사용되는 다양한 동족체가 있음을 인지할 것이다.
- <229> 어떤 효소들은 그것들이 손상된 염기를 제거하여 AP 잔기를 생성하고 또한 글리코실라제 활성에 의해 생성된 AP 부위에 대해 3'의 인산이에스테르 공격을 절단하는 리아제 활성을 나타내는 두 가지 글리코실라제 활성을 가진다는 점에서 이작용성(bifunctional)이다. 그러므로 이들 이중 활성 효소는 AP 엔도뉴클레아제이면서 DNA 글리코실라제이기도 하다. 예를 들어 Endo VIII은 N-글리코실라제와 AP-리아제 둘 다로서 작용한다. N-글리코실라제 활성은 이중-가닥의 DNA로부터 손상된 피리미딘을 방출하여 비퓨린 (AP 부위)을 생성한다. AP-리아제 활성은 AP 부위에 대해 3'과 5'을 절단하여 5' 포스페이트와 3' 포스페이트를 생성한다. 엔도뉴클레아제 VIII에 의해 인지되고 제거되는 손상된 염기는 우레아, 5,6-디히드록시티민, 티민 글리콜, 5-히드록시-5-메틸히단톤, 우라실 글리콜, 6-히드록시-5,6-디히드로티민 및 메틸타르트로닐우레아를 포함한다 (Dizdaroglu, M., et al., *Biochemistry*, 32,12105-12111, 1993 and Hatahet, Z. et al., *J. Biol. Chem.*, 269,18814-18820, 1994; Jiang, D., et al., *J. Biol. Chem.*, 272(51), 32220-32229, 1997; Jiang, D., et al., *J. Bact.*, 179(11), 3773-3782, 1997).
- <230> Fpg (포름아미도피리미딘 [fapy]-DNA 글리코실라제)(8-옥소구아닌 DNA 글리코실라제로도 알려져 있다)가 또한 N-글리코실라제와 AP-리아제로서 작용한다. N-글리코실라제 활성은 이중 가닥의 DNA로부터 손상된 퓨린을 방출시켜서 비퓨린 (AP 부위)을 생성한다. AP-리아제 활성은 AP 부위에 대해 3'과 5' 양쪽을 절단함으로써 AP 부위를 제거하고 하나의 염기 갭(gap)을 생성한다. Fpg에 의해 인지되고 제거되는 손상된 염기로는 7,8-디히드로-8-옥소구아닌 (8-옥소구아닌), 8-옥소아데닌, fapy-구아닌, 메티-fapy-구아닌, fapy-아데닌, 아플라톡신 B1-fapy-구아닌, 5-히드록시-시토신 및 5-히드록시-우라실이 있다 (Tchou, J. et al. *J. Biol Chem.*, 269, 15318-15324, 1994; Hatahet, Z. et al. *J. Biol. Chem.*, 269, 18814-18820, 1994; Boiteux, S., et al, *EMBO J.*, 5, 3177-3183, 1987; Jiang, D., et al., *J. Biol. Chem.*, 272(51), 32220-32229, 1997; Jiang, D., et al., *J. Bact.*, 179(11), 3773-3782, 1997).
- <231> 많은 DNA 글리코실라제와 AP 엔도뉴클레아제가 상업적으로, 예컨대 New England Biolabs사(Ipswich, MA)에서 시판중이다.
- <232> 본 발명의 어떤 실시예에서는, AP 엔도뉴클레아제에 의한 절단에 대한 기질인 부위를 포함하는 연장 프로브가 포스포로티올레이트 연쇄를 함유하는 연장 프로브를 위한 상기에서 설명된 것과 같은 서열화 방법에서 또는 서열화 방법 AB (하기 설명 참조)에서 사용된다. 이들 방법 중 어떠한 것에서든, 연장 프로브가 성장하는 핵산 가닥에 결합된 후, 연장 프로브는 AP 엔도뉴클레아제를 사용하여 절단되어 표지를 포함하는 프로브의 부분이 제거된다.
- <233> 특정 AP 엔도뉴클레아제에 따라, 또 서열화가 3'→5' 또는 5'→3' 방향으로 수행되는 지에 따라, 연장된 이중나선 상에 연장가능한 프로브 말단을 생성하기 위하여 절단 후에 연장된 이중나선을 폴리뉴클레오티드 키나제 또는 포스파타제로 처리하는 것이 필요하거나 바람직할 수 있다 (연장가능한 말단의 설명에 대해서는 도 5A 및 5B 참조). 그러므로 본 발명의 어떤 방법에서 연장가능한 말단은 폴리뉴클레오티드 키나제 또는 포스파타제를 사용한 처리에 의해 생성된다. 당업자들은 다양한 효소에 대해 적절한 완충제가 사용될 것이고, 추가적인 세척 단계가 효소를 제거하고 상기 방법의 후속 단계에 적절한 조건을 제공하기 위해 포함될 수 있음을 인지할 것이다.
- <234> 다른 실시예에서 연장 프로브는 DNA 글리코실라제에 의한 제거를 위한 기질인 손상된 염기를 포함한다. 광범위한 세포독성 및 돌연변이성 DNA 염기가 상이한 DNA 글리코실라제에 의해 제거되며, 그것은 DNA에 대한 손상 후에 염기 절출 수복 경로를 개시한다 (Krokan, H.E., et al, *Biochem J.*, 325 (Pt 1):1-16, 1997). DNA 글리코

실라제는 손상된 염기와 데옥시리보스 사이의 N-글리코시드 결합을 절단함으로써 유리 염기를 방출시키고 비퓨린/비피리미딘 (AP) 염기를 남긴다. 어떤 구체예에서 연장 프로브는 우라실 잔기를 포함하는데, 그것은 우라실-DNA 글리코실라제 (UDG)에 의해 제거된다. UDG는 오늘날까지 연구된 살아있는 모든 유기체에서 발견되며, 이들 효소의 대다수는 당해 기술분야에 알려져 있고, 본 발명에도 사용된다 (Frederica, et al, Biochemistry, 29, 2353-2537, 1990; Krokan, 상기 동일). 예를 들어 포유류 세포는 최소한 4가지 유형의 UDG: 미토콘드리아 UNG1 및 핵 UNG2, SMUG1, TDG, 및 MBD4 (Krokan, et al., Oncogene, 21, 8935-8948, 2002)를 포함한다. UNG1 및 UNG2는 대장균 Ung의 전형이 된 고도로 보존된 군에 속한다.

<235> 발현 프로브가 손상된 염기를 포함하는 실시예에서, 연장 프로브가 연장가능한 프로브 말단에 결합된 후에 연장된 이중나선은 손상된 염기를 제거하는 글리코실라제와 접촉하게 됨으로써, 비염기성 잔기가 생성된다. 글리코실라제에 의해 제거되기 쉬운 손상된 염기를 포함하는 연장 프로브를 "잘라지기 쉬운 연쇄를 포함하도록 쉽게 변형될 수 있는" 것으로 간주한다. 그런 다음 연장된 이중나선은 AP 엔도뉴클레아제와 접촉하게 되고, 그 효소는 상기에서 설명한 것과 같이, 비염기성 잔기와 인접한 뉴클레오타이드 사이의 연쇄를 절단한다. 본 발명의 어떤 실시예에서 DNA 글리코실라제이기도 하고 AP 엔도뉴클레아제이기도 한 이중 활성 효소는 이들 두 가지 반응을 모두 수행하기 위해 사용된다. 어떤 실시예에서 손상된 염기를 함유하는 연장된 이중나선은 DNA 글리코실라제 및 AP 엔도뉴클레아제와 접촉한다. 효소는 본 발명의 다양한 실시예에서 조합으로 또는 차례로 (즉 글리코실라제 먼저, 나중에 엔도뉴클레아제) 사용될 수 있다.

<236> 본 발명의 어떤 실시예에서 연장 프로브는 데옥시이노신인 트리거 잔기를 포함한다. 상기에서 주지된 바와 같이, 데옥시이노신 3' 엔도뉴클레아제로도 불리는 대장균 엔도뉴클레아제 V (Endo V), 및 그것들의 동족체는 데옥시이노신 잔기에 대해 두 번째 인산이에스테르 결합 3'에 있는 데옥시이노신을 함유하는 핵산을 절단하여 3' OH와 5' 포스페이트 말단을 남긴다. 그러므로 이 결합은 연장 프로브에서 잘라지기 쉬운 연쇄로서 작용한다. Endo V와 그것의 절단 특성은 당해 기술분야에 공지되어 있다 (Yao, M. and Kow Y. W., *J. Biol. Chem.*, 271, 30672-30673 (1996); Yao, M. and Kow Y. W., *J Biol. Chem.*, 270, 28609-28616 (1995); He, B, et al., *Mutat. Res.*, 459, 109-114 (2000)). 데옥시이노신 외에 Endo V는 또한 데옥시우리딘, 데옥시크산토신, 및 데옥시크사노신을 인지한다 (Hitchcock, T. et al., *Nuc. Acids Res.*, 32(13), 32(13) (2004)). 포유류 동족체, 예컨대 mEndo V도 또한 절단 활성을 나타낸다 (Moe, A., et al., *Nuc. Acids Res.*, 31(14), 3893-3900 (2004)). Endo V가 데옥시이노신을 포함하는 프로브에 대한 바람직한 절단제인 한편, 다른 절단 시약도 데옥시이노신을 포함하는 프로브를 절단하기 위해 사용될 수 있다. 예를 들어 손상된 염기로서, 히포크산틴은 적절한 DNA 글리코실라제에 의해 제거되기 쉽고, 그 결과의 비염기성 잔기를 함유하는 연장 프로브는 다시 엔도뉴클레아제에 의해 절단되기 쉽다.

<237> 만약 데옥시이노신이 트리거 잔기로서 사용된다면, 프로브의 다른 곳, 특히 연장가능한 프로브 말단에 결합될 말단과 트리거 분자 사이의 위치에 있는 데옥시이노신을 사용하는 것을 피하는 것이 바람직할 수 있다는 것이 인지될 것이다. 그러므로 만약 프로브가 하나 또는 그 이상의 보편적 염기를 포함한다면, 데옥시이노신 이외의 뉴클레오타이드가 사용될 수 있다. 또한 트리거 잔기를 함유하는 핵산을 특정 절단제에 의한 절단에 민감하게 하는 트리거 잔기가 연장 프로브에 사용되는 경우, 동일한 절단제에 의하여 트리거 절단될 프로브에 (또는 그 연장 프로브와 함께 서열화 반응에 사용될 다른 프로브에) 다른 잔기를 포함하는 것을 피하는 것이 바람직할 것이다.

<238> 본 발명은 트리거 잔기를 포함하는 핵산을 절단하는 모든 효소의 사용을 포함한다. 추가 효소는 New England Biolabs®, Inc와 같은 효소 공급업체의 카탈로그를 찾아봄으로써 확인될 수 있다. New England Biolabs 카탈로그, 2005년판 (New England Biolabs, Ipswich, MA 01938-2723)은 본원에 참조로 삽입된 것이며, 본 발명은 트리거 잔기를 함유하는 핵산을 절단하는, 거기에 개시된 모든 효소, 또는 그러한 효소의 동족체의 사용까지 고려한다. 사용될 수 있는 다른 효소로는 예컨대 hOGG1 및 그것의 동족체가 있다 (Radicella, JP, et al., *Proc Natl Acad Set USA.*, 94(15):8010-5, 1997).

<239> 손상된 염기, 비염기성 잔기 등과 같은 트리거 잔기를 함유하는 올리고뉴클레오타이드를 합성하기 위한 방법은 당해 기술분야에 공지되어 있다. AP 엔도뉴클레아제에 대한 기질인 부위를 함유하는 올리고뉴클레오타이드, 예컨대 비염기성 잔기를 함유하는 올리고뉴클레오타이드를 합성하기 위한 방법은 당해 기술분야에 공지되어 있고, 일반적으로 자동 고상 올리고뉴클레오타이드 합성을 따른다. 어떤 실시예에서 비염기성 잔기의 원하는 위치에 우리딘을 함유하는 올리고뉴클레오타이드가 합성된다. 그런 다음 그 올리고뉴클레오타이드는 우라실을 제거하는 UDG와 같은 효소로 처리됨으로써, 우리딘이 올리고뉴클레오타이드에 존재하는 경우에는 언제나 비염기성 잔기가 생성된다.

- <240> 본 발명의 어떤 실시예에서 올리고뉴클레오타이드 프로브는 문헌에 설명되는 것과 같은 이당 뉴클레오시드를 포함한다 (Nauwelaerts, K., et al, *Nuc. Acids. Res.*, 31(23), 2003). 결찰 반응 후에 연장된 이중나선은 과요오드산염 (NaIO₄)를 사용하여 절단되고, 염기 (예컨대 NaOH)로 처리되어 표지가 제거되어, 유리 3' OH와 P5-OP₃H₂기가 생성된다. 서열화가 3'→5' 또는 5'→3'으로 수행되는 지에 따라 폴리뉴클레오타이드 키나제 또는 포스파타제로 연장된 이중나선을 처리함으로써 연장가능한 말단을 생성하는 것이 필요하거나 바람직할 수 있다. 그러므로 본 발명의 어떤 실시예에서 연장가능한 말단은 폴리뉴클레오타이드 키나제 또는 포스파타제를 사용한 처리에 의해 생성된다.
- <241> 이당 뉴클레오시드를 포함하는 폴리뉴클레오타이드는 비염기성 잔기를 포함하는 것으로 간주한다. 예를 들어 한 뉴클레오타이드의 3' OH기와 다음 뉴클레오타이드의 5' 포스페이트기 사이에 삽입된 리보스 잔기를 함유하는 폴리뉴클레오타이드는 비염기성 잔기를 포함하는 것으로 간주한다.
- <242> **캡핑**
- <243> 어떤 경우에, 연장가능한 말단을 가지는 모든 프로브보다 적은 수의 프로브가 연장, 결찰, 및 절단의 각 사이클에서 연속적인 결찰 반응에 참여한다. 만약 그러한 프로브가 연속되는 사이클에 참여한다면, 각 뉴클레오타이드 확인 단계의 정확성은 점차로 감소할 것임이 인지될 것이다. 비록 발명자들이 포스포로티올레이트 연쇄를 함유하는 연장 프로브의 사용이 고효율로 결합되는 것을 허용한다는 것을 발견하였지만, 발명의 어떤 실시예에서는 캡핑 단계가 포함되어 미래의 사이클에 참여하는 것으로부터 결찰을 진행하지 않는 그런 연장가능한 말단이 방지된다. 3'-O-P-S-5' 포스포로티올레이트 연쇄를 함유하는 연장 프로브를 사용하여 5'→3' 방향으로 서열화될 때, 캡핑은 DNA 중합효소와 비-연장성 부분을 포함하는 결찰되지 않은 연장가능한 말단을 연장함으로써 수행될 수 있다. 3'-S-P-O-5' 포스포로티올레이트 연쇄를 함유하는 연장 프로브를 사용하여 3'→5' 방향으로 서열화될 때, 캡핑은 주형을 포스파타제로, 예컨대 결찰 또는 검출 후에 처리함으로써 수행될 수 있다. 다른 캡핑 방법도 사용될 수 있다.
- <244> H. 올리고뉴클레오타이드 프로브 군을 사용하는 서열화
- <245> 상기에서 설명된 "방법 A"로서 포괄적으로 언급된 서열화 방법에서, 어떠한 특정 연장 프로브에 부착된 표지와 프로브의 말단 (즉 연장된 이중나선의 연장가능한 프로브 말단에 결찰된 말단)에 있는 하나 또는 그 이상의 뉴클레오타이드의 정체 사이의 직접적이고 공지된 대응성이 있다. 그러므로 새롭게 결찰된 연장 프로브의 표지를 확인하는 것은 주형의 하나 또는 그 이상의 뉴클레오타이드의 정체를 확인하기에 충분하다. 본 발명은 "방법 AB"로서 포괄적으로 언급되며, 뉴클레오타이드 확인에 대한 상이한 접근법을 채택하는 연장, 결찰, 및 바람직하게는 절단의 연속적인 사이클을 포함하는 추가적인 서열화 방법을 제공한다.
- <246> 본 발명은 최소한 2개의 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합을 사용하는 서열화 방법 AB를 제공한다. 각 프로브 군은 표지를 토대로 이름이 정해지는데, 예를 들면, "레드", "블루", "옐로우", "그린"과 같다. 상기에서 설명된 방법에서와 같이, 연장은 개시 올리고뉴클레오타이드와 주형에 의해 형성된 이중나선으로부터 출발한다. 개시 올리고뉴클레오타이드는 올리고뉴클레오타이드 프로브를 그것의 말단에 결찰해 연장된 이중나선을 형성하고, 그런 다음 반복적으로 결찰의 연속적인 사이클에 의해 연장됨으로써 연장된다. 프로브는 말단 위치 (이중나선의 성장하는 핵산 가닥에 결찰된 뉴클레오타이드로부터 프로브의 반대쪽 말단에 있는)에 비-연장성 부분을 가짐으로써 연장된 이중나선의 단지 단일한 연장만이 단일 사이클에서 일어난다. 각 사이클 동안에 연속적으로 결찰된 프로브 위의 또는 그것과 결합된 표지가 검출되고, 비-연장성 부분이 제거되거나 변형되어 연장성 말단이 생성된다. 표지의 검출은 프로브가 속한 프로브 군의 명칭을 확인한다.
- <247> 연장, 결찰, 및 검출의 연속적인 사이클은 질서정연한 일련의 표지 이름을 초래한다. 표지는 연속적인 위치에서 주형에 혼성화하는 연속적으로 결찰된 프로브가 속하는 프로브 군에 상응한다. 프로브는 결찰 후에 주형의 상이한 뉴클레오타이드의 반대쪽에 위치한 근접한 말단을 가진다. 그러므로 프로브 군 이름의 순서와 주형의 뉴클레오타이드 순서 사이에는 대응성이 있다.
- <248> 잘라지기 쉬운 연쇄가 연장 프로브의 근접한 뉴클레오시드와 인접한 뉴클레오시드 사이에 위치한 본 발명의 실시예에서, 프로브 군 이름의 순서 리스트는 연장된 올리고뉴클레오타이드 프로브가 각 사이클에서 한 뉴클레오타이드씩 연장되기 때문에 단일한 개시 올리고뉴클레오타이드로부터 시작하는 연장, 결찰, 검출, 및 절단의 연속적인 사이클에 의해 얻어질 수 있다. 만약 잘라지기 쉬운 연쇄가 두 개의 다른 뉴클레오시드 사이에 위치한다면, 프로브 군 이름의 순서 리스트는 서열화 방법 A에 대해 설명된 것과 같이, 결합 반응 내에서 상이한 위치에 혼성화하는 개시 올리고뉴클레오타이드가 사용되는 다수의 서열화 반응으로부터 얻어진 결과들로부터 조립된다.

- <249> 새롭게 결찰된 프로브가 어떤 프로브 군에 속하게 되는지를 아는 것 그 자체로서는 주형의 뉴클레오타이드의 정체를 결정하기에는 충분하지 않다. 실제로, 프로브 군 이름을 결정하는 것은 프로브의 최소한 일부분의 서열에 대한 가능성으로서 뉴클레오타이드의 특정한 조합을 제거하지만, 각 뉴클레오타이드의 정체에 대한 최소한 두 가지의 가능성을 남긴다. 그러므로 추가의 정보 없이 프로브 군 이름에 대한 지식은 때로 새롭게 결찰된 프로브의 뉴클레오타이드들에 대하여 반대쪽 위치에 위치한 주형의 뉴클레오타이드의 정체에 대하여 최소한 2가지 가능성을 남긴다. 그러므로 어떠한 단일 사이클의 연장, 결찰, 검출 (및 임의로 절단)도 그 자체로는 주형의 어떠한 뉴클레오타이드든지 확인하지 못한다. 그러나 그것은 주형에 대한 하나 또는 그 이상의 가능한 서열의 제거를 허용하고, 그로써 서열에 대한 정보를 제공한다. 본 발명의 특정 실시예에서, 하기에서 설명되는 것과 같은 프로브 및 프로브 군의 적절한 디자인을 사용하여 주형의 서열이 여전히 결정될 수 있다. 본 발명의 어떤 실시예에서 서열화 방법 AB는 그러므로 다음의 두 단계: 프로브 군 이름의 순서 리스트가 얻어지는 첫 번째 단계와, 순서 리스트가 주형의 서열을 결정하기 위해 암호 해독되는 두 번째 단계를 포함한다.
- <250> 다른 언급이 없는 한, 서열화 방법 A와 AB는 일반적으로 프로브를 합성하고, 주형을 제조하고, 연장, 결찰, 절단, 및 검출의 단계들을 수행하기 위하여 유사한 방법들을 사용한다.
- <251> **서열화 방법 AB에 대한 올리고뉴클레오타이드 연장 프로브와 프로브 군의 특징**
- <252> 서열화 방법 AB에 사용하기 위한 프로브 군은 각 프로브 군이 상이한 서열의 다수의 표지된 올리고뉴클레오타이드 프로브를 포함하고, 서열의 각 위치에서 프로브 군은 그 위치에서 상이한 염기를 가지는 최소한 2개의 프로브를 포함하는 것을 특징으로 한다. 바람직하게도 프로브는 잘라지기 쉬운 뉴클레오시드간 연쇄를 포함한다. 잘라지기 쉬운 연쇄는 프로브의 어느 곳에서든지 위치할 수 있다. 바람직하게는 프로브는 한 말단에서 결합 효소에 의해 연장될 수 없는 부분을 가진다. 바람직하게도 프로브는 잘라지기 쉬운 연쇄와 결합 효소에 의해 연장될 수 없는 부분 사이의 위치에서 표지되어, 프로브가 연장가능한 프로브 말단에 결찰된 후 잘라지기 쉬운 연쇄의 절단으로 연장가능한 프로브 말단에 결찰된 미표지된 부분과 미표지된 부분에 더 이상 부착되지 않는 표지된 부분이 생성된다.
- <253> 각 프로브 군의 프로브들은 바람직하게는 최소한 j 뉴클레오시드 X 를 포함하는데, 여기서 j 는 최소한 2이고, 각 X 는 프로브 군의 프로브 중에서도 최소한 2-배 축퇴성이다. 각 프로브 군의 프로브들은 나아가 최소한 k 뉴클레오시드 N 을 포함하는데, 여기서 k 는 최소한 2이고, N 은 어떠한 뉴클레오시드를 나타낸다. 일반적으로 $j+k$ 는 100 또는 그 이하이고, 전형적으로는 30 이하이다. 뉴클레오시드 X 는 프로브 중 어디에든지 위치할 수 있다. 뉴클레오시드 X 는 연속적인 위치에 위치할 필요는 없다. 유사하게 뉴클레오시드 N 은 연속적인 위치에 위치할 필요는 없다. 달리 표현하자면, 뉴클레오시드 X 와 N 은 산제할 수 있다. 그럼에도 불구하고, 뉴클레오시드 X 는 5'→3' 서열을 가지는 것으로 간주되며, 뉴클레오시드는 연속적일 필요는 없는 것으로 이해된다. 예를 들어, 구조 $X_A N X_G N N X_C N$ 의 프로브에서 뉴클레오시드 X 는 서열 AGC를 가지는 것으로 간주할 것이다. 마찬가지로, 뉴클레오시드 N 도 서열을 가지는 것으로 간주할 수 있다.
- <254> 뉴클레오시드 X 는 동일하거나 상이할 수 있지만, 독립적으로 선택되지는 않는다. 즉 각 X 의 정체는 프로브 내의 하나 또는 그 이상의 다른 뉴클레오시드 X 의 정체에 의해 구속된다. 그러므로 일반적으로 단지 뉴클레오시드 X 의 특정 조합만이 어떠한 특정 프로브에 및 어떠한 특정 프로브 군의 프로브 내에 존재한다. 달리 표현하면, 각 프로브에서, 뉴클레오시드 X 의 서열은 길이 j 의 모든 가능한 서열의 하위세트를 나타낼 수 있다. 그러므로 X 의 하나 또는 그 이상의 뉴클레오타이드의 정체는 하나 또는 그 이상의 다른 뉴클레오시드에 대하여 가능한 정체를 제한한다.
- <255> 뉴클레오시드 N 은 바람직하게는 독립적으로 선택되며, A, G, C, 또는 T (또는 임의로, 축퇴성-감소 뉴클레오시드)일 수 있다. 바람직하게도 뉴클레오시드 N 의 서열은 길이 k 의 가능한 모든 서열을 나타내며, 단 하나 또는 그 이상의 N 은 축퇴성-감소 뉴클레오시드일 수 있다. 그러므로 프로브는 두 부분을 함유하는데, 뉴클레오시드 N 으로 이루어지는 한 부분은 구속되지 않은 부분으로서 언급되고, 뉴클레오시드 X 로 이루어지는 다른 부분은 구속된 부분으로서 언급된다. 상기에서 설명된 것과 같이, 부분들은 연속적인 뉴클레오시드일 필요는 없다. 구속된 부분과 구속되지 않은 부분을 함유하는 프로브는 본원에서 부분적으로 구속된 프로브로서 언급된다. 바람직하게도 구속된 부분의 하나 또는 그 이상의 뉴클레오시드는 프로브의 근접한 말단에, 즉 본 발명의 상이한 실시예에서 올리고뉴클레오타이드 프로브의 5' 또는 3' 말단의 어느 한 쪽일 수 있는 연장가능한 프로브 말단에 결찰된 뉴클레오시드를 함유하는 말단에 있다.
- <256> 어떠한 올리고뉴클레오타이드 프로브의 구속된 부분이 특정 서열을 유일하게 가질 수 있기 때문에, 프로브의 구속

된 부분에 있는 하나 또는 그 이상의 뉴클레오타이드의 정체는 아는 것은 하나 또는 그 이상의 다른 뉴클레오타이드에 대한 정보를 제공한다. 정보는 하나 또는 그 이상의 다른 뉴클레오타이드를 정확하게 결정하기에 충분할 수도 있고 그렇지 않을 수도 있지만, 구속된 부분에 있는 하나 또는 그 이상의 다른 뉴클레오타이드의 정체에 대한 하나 또는 그 이상의 가능성을 제거하기에는 충분할 것이다. 서열화 방법 AB의 특정한 바람직한 실시예에서 프로브의 구속된 부분에 있는 한 뉴클레오타이드의 정체는 아는 것은 구속된 부분에 있는 각각의 다른 뉴클레오타이드를 정확하게 확인하기에, 즉 구속된 부분을 포함하는 뉴클레오타이드의 정체와 순서를 결정하기에 충분하다.

<257> 상기에서 설명된 서열화 방법에서와 같이, 주형에 상보하는 연장 프로브의 가장 근접한 뉴클레오타이드는 개시 올리고뉴클레오타이드의 연장가능한 말단에 (연장, 결합, 및 검출의 첫 번째 사이클에서) 및 후속적인 연장, 결합, 및 검출의 사이클에서 연장된 올리고뉴클레오타이드 프로브의 연장가능한 말단에 결합된다. 검출은 새롭게 결합된 프로브가 속하는 프로브 군의 이름을 결정한다. 프로브의 구속된 부분에 있는 각 위치가 최소한 2-배 축퇴성이기 때문에 프로브 군의 이름은 그 자체로서는 구속된 부분의 어떠한 뉴클레오타이드를 확인하지 못한다. 그러나 구속된 부분의 서열이 길이 j 의 가능한 모든 서열의 하위세트 중 하나이기 때문에, 프로브 군을 확인하는 것은 구속된 부분의 서열에 대한 특정한 가능성을 제거한다. 프로브의 구속된 부분은 그것의 서열 결정 부분을 구성한다. 그러므로 프로브의 구속된 부분에 있는 하나 또는 그 이상의 다른 뉴클레오타이드의 정체에 대한 하나 또는 그 이상의 가능성을 그것이 속하는 프로브 군을 확인함으로써 제거하는 것은 연장 프로브가 혼성화된 주형에 있는 뉴클레오타이드의 정체에 대한 하나 또는 그 이상의 가능성을 제거한다. 본 발명의 바람직한 실시예에서 부분적으로 구속된 프로브는 어떠한 두 개의 뉴클레오타이드 사이의 잘라지기 쉬운 연쇄를 포함한다.

<258> 어떤 실시예에서 부분적으로 구속된 프로브는 일반 구조식 $(X)_j(N)_k$ 를 가지는데, 식에서 X 는 뉴클레오타이드를 나타내고, $(X)_j$ 는 각 위치에 있는 최소한 2-배의 축퇴성이어서 X 가 상이한 염기쌍 특이성을 가지는 최소한 2개의 뉴클레오타이드 중 어느 것일 수 있으며, N 은 어떠한 뉴클레오타이드를 나타내고, j 는 최소한 2이며, k 는 1과 100 사이에 있고, 프로브 말단에 있는 X 이외의 최소한 하나의 N 또는 X 는 검출가능한 부분을 포함한다. 바람직하게는 $(N)_k$ 는 각 위치에서 독립적으로 4-배 축퇴성이어서, 각 프로브에서 $(N)_k$ 는 길이 k 의 가능한 모든 서열을 나타내고, 단 $(N)_k$ 에 있는 하나 또는 그 이상의 위치는 축퇴성-감소 뉴클레오타이드가 차지할 수 있다. $(X)_j$ 에 있는 뉴클레오타이드는 동일하거나 상이할 수 있지만 독립적으로 선택되지는 않는다. 달리 표현하면, 각 프로브에서 $(X)_j$ 는 길이 j 의 가능한 모든 서열의 하위세트를 나타낼 수 있다. 그러므로 $(X)_j$ 의 하나 또는 그 이상의 뉴클레오타이드의 정체는 하나 또는 그 이상의 다른 뉴클레오타이드에 대하여 가능한 정체를 제한한다. 그러므로 프로브는 구속되지 않은 부분인 $(N)_k$ 와 구속된 부분인 $(X)_j$ 의 두 부분을 함유한다.

<259> 본 발명의 바람직한 어떤 실시예에서 부분적으로 구속된 프로브는 $5'-(X)_j(N)_kN_B^*-3'$ 또는 $3'-(X)_j(N)_kN_B^*-5'$ 의 구조를 가지는데, 이때 N 은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, $*$ 는 검출가능한 부분을 나타내고, $(X)_j$ 는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구속된 부분이고, $(X)_j$ 에 있는 뉴클레오타이드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않으며, 최소한 하나의 뉴클레오타이드간 연쇄는 잘라지기 쉬운 연쇄이고, j 는 최소한 2이며, k 는 1과 100 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 프로브 말단에 있는 X 이외의 어떠한 뉴클레오타이드 N 또는 X 상에 존재할 수 있어야 한다. 잘라지기 쉬운 연쇄는 $(X)_j$ 의 두 개의 뉴클레오타이드 사이에, $(X)_j$ 의 가장 멀리 있는 뉴클레오타이드와 $(N)_k$ 의 가장 근접한 뉴클레오타이드 사이에, $(N)_k$ 내에 있는 뉴클레오타이드들 사이에, 또는 $(N)_k$ 의 말단 뉴클레오타이드와 N_B 사이에 있을 수 있다. 바람직하게는 잘라지기 쉬운 연쇄는 포스포로티올레이트 연쇄이다.

<260> 본 발명의 보다 바람직한 실시예에서 프로브는 $5'-(XY)(N)_kN_B^*-3'$ 또는 $3'-(XY)(N)_kN_B^*-5'$ 의 구조를 가지는데, 이때 N 은 어떠한 뉴클레오타이드를 나타내고, N_B 는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, $*$ 는 검출가능한 부분을 나타내고, XY 는 프로브의 구속된 부분이고, 이때 X 와 Y 는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않는 뉴클레오타이드를 나타내며, X 와 Y 는 최소한 2-배 축퇴성이고, 최소한 하나의 뉴클레오타이드간 연쇄는 잘라지기 쉬운 연쇄이며, k 는 1과 100 사이에 있고, 포괄적이며, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 프로브 말단에 있는 X 이외의 어떠한 뉴클레오타이드 N 또는 X 상에 존재할 수 있어야 한다. 바람직하게는 잘라지기 쉬운 연쇄는 포스포로티올레이트 연쇄이다. 구조 $5'-(XY)(N)_kN_B^*-3'$ 를 가지는 프로브는 $5' \rightarrow 3'$ 방향의

서열화에 사용된다. 구조 3'-(XY)(N)_kN_B^{*}-5'를 가지는 프로브는 3'→5' 방향의 서열화에 사용된다.

<261> 특정한 바람직한 프로브의 구조는 다음과 같이 상세한 설명에서 표시된다. 5'→3' 방향으로의 서열화에 대해 부분적으로 구축된 프로브는 구조 5'-O-P-O-(X)_j(N)_k-O-P-S-(N)_iN_B^{*}-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, (X)_j는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, (X)_j에 있는 뉴클레오시드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않고, j는 최소한 2이며, (k+i)는 1과 100 사이에 있고, k는 1과 100 사이에 있으며, i는 0과 99 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 발명의 어떤 실시예에서 (X)_j는 (XY)인데, 이때 X와 Y는 최소한 2-배 축퇴성이고, 동일하거나 상이하지만 독립적으로 선택되지 않는 뉴클레오티드를 나타낸다. 발명의 어떤 실시예에서 i는 0이다.

<262> 5'→3' 방향으로의 서열화에 대한 다른 바람직한 프로브는 구조 5'-O-P-O-(X)_j(N)_k-O-P-S-(N)_iN_B^{*}-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, (X)_j는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, (X)_j에 있는 뉴클레오티드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않고, j는 최소한 2이며, i는 1과 100 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 발명의 어떤 실시예에서 (X)_j는 (XY)인데, 이때 위치 X와 Y는 최소한 2-배 축퇴성이고, X와 Y는 동일하거나 상이하지만 독립적으로 선택되지 않는 뉴클레오시드를 나타낸다. 5'→3' 방향으로의 서열화에 대한 또 다른 바람직한 프로브는 구조 5'-O-P-O-(X)_j-O-P-S-(X)_k(N)_iN_B^{*}-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, (X)_j-O-P-S-(X)_k는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, (X)_j-O-P-S-(X)_k에 있는 위치는 최소한 2-배 축퇴성이고 동일하거나 상이할 수 있지만 독립적으로 선택되지 않으며, j와 k는 둘 다 최소한 1이고, (j+k)는 최소한 2 (예컨대 2, 3, 4, 또는 5)이며, i는 1과 100 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 발명의 어떤 실시예에서 j와 k는 둘 다 1이다.

<263> 3'→5' 방향으로의 서열화에 대해 부분적으로 구축된 프로브는 구조 5'-N_B^{*}(N)_i-S-P-O-(N)_k-O-P-O-(X)_j-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, (X)_j는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, (X)_j에 있는 뉴클레오시드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않고, j는 최소한 2이며, (k+i)는 1과 100 사이에 있고, k는 1과 100 사이에 있으며, i는 0과 99 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 발명의 어떤 실시예에서 (X)_j는 (XY)인데, 이때 X와 Y는 최소한 2-배 축퇴성이고, 동일하거나 상이하지만 독립적으로 선택되지 않는 뉴클레오시드를 나타낸다. 발명의 어떤 실시예에서 i는 0이다.

<264> 3'→5' 방향으로의 서열화에 대한 다른 바람직한 프로브는 구조 5'-N_B^{*}(N)_i-S-P-O-(X)_j-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, (X)_j는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, (X)_j에 있는 뉴클레오시드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않고, j는 최소한 2이며, i는 1과 100 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 발명의 어떤 실시예에서 (X)_j는 (XY)인데, 이때 X와 Y는 최소한 2-배 축퇴성이고, 동일하거나 상이하지만 독립적으로 선택되지 않는 뉴클레오시드를 나타낸다. 발명의 어떤 실시예에서 j는 부분적으로 구축된 프로브의 어떤 것에서 2와 5 사이, 예컨대 2, 3, 4, 또는 5이다.

<265> 3'→5' 방향으로의 서열화에 대한 또 다른 바람직한 프로브는 구조 5'-N_B^{*}(N)_i-S-P-O-(X)_k-O-P-O-(X)_j-3'을 가지는데, 이때 N은 어떠한 뉴클레오시드를 나타내고, N_B는 결합 효소에 의해 연장될 수 없는 부분을 나타내며, *는 검출가능한 부분을 나타내고, -(X)_k-O-P-O-(X)_j는 각 위치에서 최소한 2-배 축퇴성인 프로브의 구축된 부분이며, -(X)_k-O-P-O-(X)_j에 있는 뉴클레오시드는 동일하거나 상이할 수 있지만 독립적으로 선택되지 않으며, j와 k는 둘 다 최소한 1이고, (j+k)는 최소한 2 (예컨대 2, 3, 4, 또는 5)이며, i는 1과 100 사이에 있고, 단 검출가능한 부분은 N_B 대신에 또는 그것 외에 (N)_i의 어떠한 뉴클레오시드 상에 존재할 수 있어야 한다. 어떤 실시예에서 j와 k는 둘 다 1이다.

<266> 잘라지기 쉬운 연쇄가 (X)_j의 가장 근접한 뉴클레오시드와 (X)_j의 다음으로 가장 근접한 뉴클레오시드 사이에 위치한 발명의 실시예에서, 프로브 군 이름의 순서 리스트는 연장된 올리고뉴클레오티드 프로브가 각 사이클에서 한 뉴클레오티드씩 연장되기 때문에 단일한 개시 올리고뉴클레오티드로부터 시작하는 연장, 결합, 검출, 및 절단의 연속적인 사이클에 의해 얻어질 수 있다. 잘라지기 쉬운 연쇄가 2개의 다른 뉴클레오시드 사이에 위치한 발명의 실시예에서, 프로브 군 이름의 순서 리스트는 서열화 방법 A에 대해 설명된 것과 같이, 결합 반응 내에서 상이한 위치에 혼성된 개시 올리고뉴클레오티드가 사용되는 다수의 서열화 반응으로부터 얻어지는 결과로부터 더 조립된다.

<267> 상기에서 설명된 것들 이외의 다수의 구조 중 어느 것을 가지는 프로브들이 서열화 방법 AB에 사용될 수 있음이 인지될 것이다. 예를 들어 프로브는 XNY(N)_k (구축된 뉴클레오시드 X와 Y는 인접하지 않는다) 또는 XIY(N)_k (I는 보편적 염기이다)와 같은 구조를 가질 수 있다. 추가의 가능성을 나타내는 것들로는 (N)_kX(N)_l, (N)_iX(N)_jY(N)_kZ(N)_l, (N)_iX(N)_jYIZ(N)_l, 및 (N)_iX(N)_jY(N)_kZ(I)_l이 있다. 상기에서 설명된 프로브에서와 같이, 이들 프로브는 잘라지기 쉬운 연쇄, 검출가능한 부분, 및 결합 효소에 의해 연장될 수 없는 한 말단에 있는 부분을 포함한다. 바람직하게도 프로브는 결합 효소에 의해 연장될 수 없는 부분으로부터 프로브의 반대쪽 말단에 있는 뉴클레오티드에 부착된 검출가능한 부분을 포함하지 않는다. 이들 구조 중 어느 것, 또는 다른 것을 가지는 프로브들을 포함하는 프로브 군은 각 프로브 군이 상이한 서열의 다수의 표지된 프로브를 포함하고, 서열의 각 위치에서 프로브 군은 그 위치에서 상이한 염기를 가지는 프로브를 최소한 2개 포함해야 하는 기준을 충족한다. 바람직하게도 각 프로브에 있는 뉴클레오시드의 총 수는 100 이하, 예컨대 30 이하이다.

<268> **올리고뉴클레오티드 연장 프로브 군의 코드화.** 본 발명의 서열화 방법은 코드화된 프로브 군을 사용한다. "코드화"란 하나의 규정된 세트의 서열을 가지는 부분을 포함하는 프로브와 특정 표지를 결합시켜서 규정된 서열 세트의 구성원인 서열을 가지는 부분을 포함하는 프로브들이 표지로 표지된 도식을 말한다. 일반적으로 코드화는 하나 또는 그 이상의 프로브를 가지는 다수의 구별가능한 표지 각각과 결합함으로써, 각각의 구별가능한 표지는 상이한 그룹의 프로브와 결합하고, 각 프로브는 단지 하나의 표지 (검출가능한 부분의 조합을 포함할 수 있다)로 표지된다. 바람직하게는 프로브의 각 그룹의 프로브들은 각각 동일한 규정된 세트의 서열의 구성원인 서열을 가지는 부분을 포함한다. 그 부분은 단일 뉴클레오시드일 수도 있고 또는 길이가 다수의, 예컨대 2, 3, 4, 5, 또는 더 많은 뉴클레오시드일 수 있다. 부분의 길이는 프로브의 전체 길이의 단지 작은 분획만을 구성할 수도 있고 또는 전체 프로브를 구성할 수도 있다. 규정된 세트의 서열은 부분의 길이에 따라 단지 단일 서열을 함유하거나 어떠한 수의 상이한 서열을 함유할 수 있다. 예를 들어 만약 부분이 단일 뉴클레오시드라면, 규정된 세트의 서열은 최대 4 요소 (A, G, C, T)를 가질 수 있다. 만약 부분의 길이가 2개의 뉴클레오시드라면, 규정된 세트의 서열은 16가지의 요소를 가질 수 있을 것이다 (AA, AG, AC, AT, GA, GG, GC, GT, CA, CG, CC, CT, TA, TG, TC, TT). 일반적으로 규정된 세트의 서열은 가능한 서열의 총 수보다 적은 요소를 함유할 것이고, 코드화는 규정된 세트의 서열을 하나 이상 사용할 것이다.

<269> 본원에서 설명된 서열화 방법 A는 일반적으로 프로브의 근접한 뉴클레오시드 (즉 연장가능한 프로브 말단에 결합된 뉴클레오시드)와 표지의 정체성 사이의 직접적인 대응성이 있는 간단한 코드화를 가지는 한 세트의 프로브를 사용한다. 근접한 뉴클레오시드는 주형에서 그것과 혼성된 뉴클레오티드와 상보하며, 그래서 새롭게 결합된 프로브의 근접한 뉴클레오시드의 정체는 연장된 이중나선의 반대 위치에 위치한 주형의 뉴클레오티드의 정체를 결정한다. 일반적 의미에서 본원에서 설명되는 다른 서열화 방법에 사용되는 프로브는 구조 X(N)_k를 가지는데, 이때 X는 근접한 뉴클레오시드이고, 각각의 뉴클레오시드 N은 4-배 축퇴성이어서, 길이 k의 모든 가능한 서열이 프로브를 구성하는 올리고뉴클레오티드 프로브 분자의 폴로 표시된다. 그러므로 예를 들어 어떤 올리고뉴클레오티드 프로브 분자들은 위치 k=1에서 A를 함유할 것이고, 다른 것들은 위치 k=1에서 G를 함유할 것이며, 다른 것

들은 위치 $k=1$ 에서 C를 함유할 것이고, 다른 것들은 위치 $k=1$ 에서 T를 함유할 것이며, 다른 위치 k 에 대해서도 유사하며, 이때에 $(N)_k$ 에 인접한 뉴클레오타이드는 위치 $k=1$ 을 차지하는 것으로 여겨진다; $(N)_k$ 의 다음 뉴클레오타이드는 위치 $k=2$ 를 차지하는 것으로 여겨진다. 그러나 어떠한 주어진 올리고뉴클레오타이드 프로브에서, X는 단지 단일 염기쌍 특이성을 나타내며, 그것은 전형적으로 특정 뉴클레오타이드 정체, 예컨대 A, G, C, 또는 T에 상응한다. 그러므로 X는 전형적으로 특정 프로브를 구성하는 프로브 분자의 풀에서 균일하게 A, G, C, 또는 T이다. 도 2는 구조 $X(N)_k$ 를 가지는 프로브에 대한 적당한 코드화를 도시한다. 이 코드화에 따르면 $X=C$ 를 가지는 프로브는 표지 "레드"로 지정되고; $X=A$ 를 가지는 프로브는 표지 "옐로우"로 지정되며; $X=G$ 를 가지는 표지는 표지 "그린"으로 지정되고; $X=T$ 를 가지는 표지는 표지 "블루"로 지정된다. 그러므로 표지의 서열 결정 부분과 그것의 표지 사이에는 일대일 대응성이 있다.

<270> 새롭게 결찰된 연장 프로브의 정체는 연장 프로브의 가장 근접한 뉴클레오타이드의 정체에 상응하는 상기 접근법은 표지의 정체성이 연장 프로브의 가장 근접한 뉴클레오타이드의 정체에만 상응하지 않고 오히려 연장 프로브의 가장 근접한 2 또는 그 이상의 뉴클레오타이드의 정체에 상응하는 코드화를 포함하는 것으로 확대될 수 있어서, 주형의 다중 뉴클레오타이드의 정체는 단일 사이클의 연장, 결찰, 및 검출 (보통은 절단이 이어짐)로 결정될 수 있다는 것이 인지될 것이다. 그러나, 그러한 코드화는 올리고뉴클레오타이드 연장 프로브의 단일 서열과 표지를 여전히 결합시킬 것이므로, 주형의 반대쪽에 위치한 상보하는 뉴클레오타이드의 정체는 확인될 수 있다. 예를 들어 상기에서 설명된 것과 같이 단일 사이클에서 2개의 뉴클레오타이드를 확인하기 위해서는, 각각이 상응하는 표지를 가진 (즉 16개의 구별가능한 표지) 16개의 상이한 올리고뉴클레오타이드 프로브가 필요할 것이다.

<271> 서열화 방법 AB는 프로브를 표지에 결합시키는 대체 접근법을 사용한다. 표지의 정체성과 프로브의 서열 결정 부분의 서열 사이의 일대일 대응성보다는 동일한 표지는 상이한 서열 결정 부분들을 가지는 다수의 프로브에 지정된다. 프로브는 부분적으로 구속되고, 프로브의 구속된 부분은 그것의 서열 결정 부분이다. 그러므로 동일한 표지는 각각이 상이한 서열을 가지는 구속된 부분을 가지는 다수의 상이한 프로브에 지정되고, 이때 서열은 규정된 세트의 서열 중 하나이다. 상기에서 언급된 바와 같이, 동일한 표지를 포함하는 프로브는 "프로브 군"을 구성한다. 방법은 각각이 상이한 서열을 가지는 구속된 부분을 가지는 다수의 프로브를 포함하는 다수의 그러한 프로브 군을 사용하며, 이때 서열은 규정된 세트의 서열 중 하나이다.

<272> 다수의 프로브 군은 프로브 군의 "집합"으로서 언급된다. 프로브 군의 집합의 각 프로브 군의 프로브들은 집합의 다른 프로브 군을 표지하기 위해 사용된 표지와는 구별되는 표지로 표지된다. 각 프로브 군은 바람직하게는 그 자체만의 규정된 세트의 서열을 가진다. 바람직하게도 각 프로브 군의 프로브의 구속된 부분은 동일한 길이이고, 바람직하게는 프로브 군의 집합의 프로브 군의 구속된 부분은 동일한 길이이다. 바람직하게도 프로브 군의 집합의 프로브 군에 대한 규정된 서열의 세트의 조합은 구속된 부분의 모든 가능한 서열을 포함한다. 바람직하게도 프로브 군의 집합은 4개의 구별가능하게 표지된 프로브 군을 포함하거나, 그것을 구성한다. 바람직하게도 프로브의 구속된 부분은 길이가 2 뉴클레오타이드이다.

<273> 구별가능하게 표지된 프로브 군의 광범위하게 상이하게 코드화된 집합은 상기 기준을 만족할 것이며, 본 발명의 방법을 실시하는 데 사용될 수 있다. 그러나, 프로브 군의 특정한 집합이 바람직하다. 부분적으로 구속된 프로브를 포함하는 4개의 구별가능하게 표지된 프로브 군의 바람직한 집합에 대한 예시적인 코드화는 도 25A에 도시되어 있다. 도 25A에 도시된 것처럼, 구속된 부분은 프로브의 2개의 대부분의 3' 뉴클레오타이드를 구성한다. 프로브 군은 표지된 "레드", "옐로우", "그린", 및 "블루"이다. 각 프로브 군의 프로브들은 그것의 서열이 규정된 세트의 서열 중 하나인 구속된 부분을 포함하고, 규정된 세트는 각 프로브 군에 대해 달라진다. 예를 들어 프로브의 근접한 말단일 것으로 여겨지는 각 서열의 3' 말단에서 시작하여, "레드" 프로브 군에 대한 서열의 규정된 세트는 {CT, AG, GA, TC}이고; "옐로우" 프로브 군에 대한 서열의 규정된 세트는 {CC, AT, GG, TA}이며; "그린" 프로브 군에 대한 서열의 규정된 세트는 {CA, AC, GT, TG}이고; "블루" 프로브 군에 대한 서열의 규정된 세트는 {CG, AA, GC, TT}이다. 각각의 규정된 세트는 다른 세트의 하나에 존재하는 어떠한 구성원을 함유하지 않으며, 그것이 바람직한 특징이다. 또한 프로브 군의 집합의 프로브 군에 대한 규정된 서열의 세트의 조합은 길이 2의 가능한 모든 서열, 즉 가능한 모든 디뉴클레오타이드를 포함한다. 이런 프로브 군의 집합의 바람직하지만 반드시 그럴 필요는 없는 다른 특징은 프로브의 구속된 부분에 있는 각 위치가 4-배 축퇴성이라는 것, 즉 A, G, C, 또는 T 중 어느 하나에 의해 차지될 수 있다는 것이다. 이 프로브 군의 집합의 바람직하지만 반드시 그럴 필요는 없는 또 다른 특징은 각 세트의 규정된 서열 내에서 단지 하나의 서열만이 어떠한 위치에서, 예컨대 가장 근접한 위치에서 또는 어떠한 다른 위치에서 어떠한 특이한 뉴클레오타이드를 가진다는 것이다. 필수적인 것은 아니지만 특히 바람직한 것은 규정된 서열의 각 세트 내에서 단지 하나의 서열만이 가장 근접한 뉴클레오타이드가 위치 1에 있을 것이라고 생각할 때, 구속된 부분 내에서 위치 2 또는 그보다 높은 수의 위치에 어떠한 특이한

뉴클레오시드를 가지는 것이다. 예를 들어 레드 프로브 군에 대한 규정된 세트의 서열에서 단지 하나의 서열이 위치 2에 있는 T를 가지고; 단지 하나의 서열이 위치 2에 G를 가지며; 단지 하나의 서열이 위치 2에 A를 가지고; 단지 하나의 서열이 위치 2에 C를 가진다.

<274>

도 25A에서 도시된 것과 같은 어떠한 특정 코드화가 제공될 때, 프로브 군 중 하나의 프로브의 구속된 부분에 있는 하나 또는 그 이상의 뉴클레오시드의 정체는 아는 것은 그 프로브의 구속된 부분에 있는 다른 뉴클레오타이드에 대한 정보를 제공하는 것이다. 가장 일반적인 의미로, 프로브 군의 프로브의 구속된 부분에 있는 하나 또는 그 이상의 뉴클레오시드의 정체는 아는 것은, 그 프로브 군에 대한 규정된 세트의 서열이 그 위치에서 그 정체를 가지는 뉴클레오시드를 가지는 서열을 함유하지 않을 것이기 때문에 다른 위치들 중 한 위치에 있는 뉴클레오시드에 대한 하나 또는 그 이상의 가능한 정체를 제거하기에 충분한 정보를 제공한다. 전형적으로 프로브 군의 프로브의 구속된 부분에 있는 하나 또는 그 이상의 뉴클레오시드의 정체는 아는 것은 다수의 뉴클레오시드, 예컨대 각각의 다른 뉴클레오시드에 대한 하나 또는 그 이상의 가능한 정체를 제거하기에 충분한 정보를 제공한다. 바람직한 코드화에 대하여, 프로브 군의 프로브의 구속된 부분에 있는 하나 또는 그 이상의 뉴클레오시드의 정체는 아는 것은 프로브의 각각의 다른 뉴클레오시드에 대한 모든, 그러나 한 가지 가능성을 제거한다. 예를 들어 도 25A에 도시된 코드화된 프로브 군의 경우, 만약 프로브가 레드 군의 한 구성원인 것으로 알려져 있고, 또한 가장 근접한 뉴클레오시드가 C라고 알려져 있다면, 인접한 뉴클레오시드는 반드시 T여야 한다. 유사하게, 만약 프로브가 그린 군의 한 구성원인 것으로 알려져 있고, 또한 가장 근접한 뉴클레오시드가 G라고 알려져 있다면, 인접한 뉴클레오시드는 반드시 T여야 한다. 그러므로 구속된 부분에 있는 한 뉴클레오시드의 정체는 아는 것은 하나를 제외한 다른 뉴클레오시드에 대한 모든 가능성을 제거하는 것이고, 따라서 다른 뉴클레오시드의 정체는 완전히 명시된다. 그러나 아직은 프로브의 구속된 부분에 있는 최소한 하나의 뉴클레오시드의 정체를 알지 않고서는 구속된 부분의 각 위치에 있는 뉴클레오시드가 A, G, C, 또는 T일 수 있기 때문에, 그것이 속하는 프로브 군의 이름을 아는 것에만 기초하여 프로브의 어떠한 구체적인 뉴클레오시드의 정체에 대해 어떠한 정보든지 얻는 것은 불가능하다. 도 25B는 서열화 방법 AB를 사용하여 프로브 군의 바람직한 집합 (상부 패널)과 결찰, 검출 및 절단의 사이클 (하부 패널)을 도시한다.

<275>

본 발명자들은 길이가 2 뉴클레오시드이고 도 25A에 도시된 것과 같은 프로브 군의 집합의 유익한 특징을 가진 구속된 부분을 함유하는 프로브 군의 24개의 집합을 디자인하였다. 이들 프로브 군은 프로브가 속하는 프로브 군의 이름을 아는 것과, 프로브의 한 뉴클레오시드의 정체를 아는 것이 구속된 부분에 있는 다른 뉴클레오시드를 정확하게 확인하기에 충분하다는 점에서 최대한 정보를 제공한다. 이것은 모든 프로브에 대한, 그리고 각 구속된 부분의 모든 뉴클레오시드에 대한 경우이다. 프로브 군의 24개의 바람직한 집합의 각각에 대한 코드화 도식은 하기 표 1에 제시한다. 표 1은 프로브 군의 각 집합에 대하여 1부터 24까지의 코드화 ID 등급을 지정한다. 각 코드화는 서열화 방법 AB에 사용하기 위하여 일반 구조 $(XY)_N$ 의 바람직한 프로브 군의 집합의 구속된 부분을 규정하고, 그로써 집합 자체를 규정한다. 표 1에서 코드화 ID 아래 줄에 있는 1의 값은 그 코드화에 따르면, 각각 첫 번째 줄과 두 번째 줄에 표시된 것과 같은 뉴클레오시드 X와 Y를 포함하는 프로브가 첫 번째 프로브 군으로 지정되는 것을 나타내고; (ii) 코드화 ID 아래 줄에 있는 2의 값은 그 코드화에 따르면, 각각 첫 번째 줄과 두 번째 줄에 표시된 것과 같은 뉴클레오시드 X와 Y를 포함하는 프로브가 두 번째 프로브 군으로 지정되는 것을 나타내며; (iii) 코드화 ID 아래 줄에 있는 3의 값은 그 코드화에 따르면, 각각 첫 번째 줄과 두 번째 줄에 표시된 것과 같은 뉴클레오시드 X와 Y를 포함하는 프로브가 세 번째 프로브 군으로 지정되는 것을 나타내고; (iv) 코드화 ID 아래 줄에 있는 4의 값은 그 코드화에 따르면, 각각 첫 번째 줄과 두 번째 줄에 표시된 것과 같은 뉴클레오시드 X와 Y를 포함하는 프로브가 첫 번째 프로브 군으로 지정되는 것을 나타낸다. 값 1, 2, 3 및 4는 각각 표지를 나타낸다. 예를 들어 코드화 9는 도 25A에 도시된 프로브 군의 집합을 규정하고, 이때 1은 블루를 나타내며, 2는 그린을 나타내고, 3은 레드를 나타내며, 4는 옐로우를 나타낸다. 표지에 대한 값의 지정은 임의적인 것, 예컨대 1은 그린, 레드, 또는 옐로우를 똑같이 잘 나타낼 수 있을 것임이 인지될 것이다. 값 1, 2, 3, 및 4와 표지 사이의 결함을 바꾸는 것은 각 프로브 군의 프로브 세트를 변화시키지는 않겠지만, 단순히 상이한 표지와 각 프로브 군을 결합시킬 것이다.

<276>

표 1: 올리고뉴클레오타이드 프로브 군 코드화

<277>

		코드화 ID																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
X	Y																									
A	A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

C	A	2	4	3	2	2	4	3	2	2	3	4	3	2	3	3	4	2	3	4	4	2	4	3	4
G	A	4	3	2	3	3	2	4	4	3	2	2	4	4	2	4	3	4	2	3	2	3	2	4	3
T	A	3	2	4	4	4	3	2	3	4	4	3	2	3	4	2	2	3	4	2	3	4	3	2	2
A	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
C	C	1	1	1	1	1	1	1	1	4	4	3	4	4	4	4	3	3	4	3	3	3	3	4	3
G	C	3	4	4	4	4	3	3	3	1	1	1	1	3	3	3	4	1	1	1	1	4	4	3	4
T	C	4	3	3	3	3	4	4	4	3	3	4	3	1	1	1	1	4	3	4	4	1	1	1	1
A	G	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
C	G	4	2	4	4	4	2	4	4	1	1	1	1	1	1	1	1	4	2	2	2	4	2	2	2
G	G	1	1	1	1	2	4	2	2	4	4	4	2	2	4	2	2	2	4	4	4	1	1	1	1
T	G	2	4	2	2	1	1	1	1	2	2	2	4	4	2	4	4	1	1	1	1	2	4	4	4
A	T	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
C	T	3	3	2	3	3	3	2	3	3	2	2	2	3	2	2	2	1	1	1	1	1	1	1	1
G	T	2	2	3	2	1	1	1	1	2	3	3	3	1	1	3	1	3	3	2	3	2	3	2	2
T	T	1	1	1	1	2	2	3	2	1	1	1	1	2	3	1	3	2	2	3	2	3	2	3	3

<278> 바람직한 프로브 군의 집합을 규정하기 위해 표 1의 사용을 한층 더 예시하기 위하여 코드화 17을 생각해보기로 한다. 이 코드화에 따르면, 구축된 부분 AA, GC, TG, 및 CT를 가지는 프로브는 표지 1 (예컨대 레드)로 지정되고; 구축된 부분 CA, AC, GG, TT를 가지는 프로브는 표지 2 (예컨대 옐로우)로 지정되며; 구축된 부분 TA, CC, AG, 및 GT를 가지는 프로브는 표지 3 (예컨대 그린)으로 지정되고; 구축된 부분 GA, TC, CG, 및 AT를 가지는 프로브는 표지 4 (예컨대 블루)로 지정된다. 그 결과의 프로브 군 집합은 도 26에 도시된다.

<279> 도 27A 내지 도 27C는 프로브 군의 바람직한 24개의 집합을 도식적으로 규정하기 위한 대체 방법을 나타낸다. 방법은 도 27A에 도시된 것과 같은 다이어그램을 사용한다. 그러한 다이어그램의 첫 번째 줄은 첫 번째 염기를 나타낸다. 각 표지는 4개의 상이한 염기 서열에 부착되며, 그 염기 서열의 각각은 첫 번째 줄의 염기를 선택된 표지의 줄에 있는 염기와 병렬시킴으로써 제공된다. 예를 들어 만약 제목이 "첫 번째 염기"로 표시되어 있는 줄에 A가 있다면, 서열 AA를 가지는 구축된 부분을 포함하는 프로브가 프로브 군 1 (표지 1)로 지정되고; 서열 AC를 가지는 구축된 부분을 포함하는 프로브가 프로브 군 2 (표지 2)로 지정되며; 서열 AG를 가지는 구축된 부분을 포함하는 프로브가 프로브 군 3 (표지 3)으로 지정되고; 서열 AT를 가지는 구축된 부분을 포함하는 프로브가 프로브 군 4 (표지 4)로 지정된다. 프로브 군에 대한 지정은 C, G, 또는 T로 시작하는 구축된 부분을 가지는 프로브에 대해서도 유사한 방식으로 이루어진다. 그러므로 도 27A에 도시된 것과 같은 염기들로 채워진 다이어그램은 도 27B에 도시된 것과 같은 코드화로 번역는데, 세트 {AA, CC, GG, TT}에 구축된 부분을 가지는 프로브는 프로브 군 1로 지정되고; 세트 {AC, CA, GT, TG}에 구축된 부분을 가지는 프로브는 프로브 군 2로 지정되며; 세트 {AG, CT, GC, TA}에 구축된 부분을 가지는 프로브는 프로브 군 3으로 지정되고; 세트 {AT, CG, GA, TC}에 구축된 부분을 가지는 프로브는 프로브 군 4로 지정된다. 도 27C는 프로브 군의 24개의 바람직한 집합의 각각을 생성하기 위해 도 27A의 다이어그램의 그림자진 부분 대신에 삽입될 수 있는 다이어그램을 도시한다. 서열화 방법 AB에서 프로브 군의 바람직한 집합을 사용하는 방법은 하기에서 더 자세하게 설명된다.

<280> 표 1에 의해 규정된 코드화된 프로브 군의 24개의 집합은 서열화 방법 AB에서 사용하기 위한 프로브 군의 집합의 바람직한 실시예만을 나타낸다. 프로브 군 이름을 아는 것이 구축된 부분에 있는 하나 또는 그 이상의 뉴클레오타이드의 정체성에 대한 지식과 함께 하나 또는 그 이상의 뉴클레오타이드에 대한 정보를 제공하는 동일한 기본적인 원리를 사용하는 광범위한 다른 코드화 도식, 프로브 군, 및 프로브 구조들도 사용될 수 있다. 프로브 군의 바람직한 집합과 비교할 때 덜 바람직한 집합의 프로브 군은 일반적으로 (i) 최소한 어떤 프로브에 대해서는, 프로브 군 이름과 뉴클레오타이드 정체성을 얹으로써 제공된 정보의 양이 더 적고; 또는 (ii) 최소한 어떤 프로브에 대해서는, 프로브 군의 이름을 얹으로써 제공된 정보의 양이 더 많기 때문에 덜 바람직하다.

<281> 일반적으로 덜 바람직한 프로브 군 집합은 바람직한 프로브 군 집합이 사용되는 방법과 유사한 방식으로 서열화 방법 AB를 수행하기 위해 사용될 수 있다. 그러나 코드 해석에 필요한 단계들은 다를 수 있다. 예를 들어 어떤 상황에서 후보 서열들을 상호 비교하는 것은 최소한 서열의 일부를 결정하는 데 충분할 수도 있다.

<282> 프로브가 길이가 2 뉴클레오타이드인 구축된 부분을 포함하는 덜 바람직한 프로브 군 집합의 실례는 도 28에 도시되어 있다. 이 코드화에 따르면, 세트 {AA, AC, GA, GC}에 구축된 부분을 가지는 프로브는 프로브 군 1로 지정

되고; 세트 {CA, CC, TA, TC}에 소속된 부분을 가지는 프로브는 프로브 군 2로 지정되며; 세트 {AG, AT, GG, GT}에 소속된 부분을 가지는 프로브는 프로브 군 3으로 지정되고; 세트 {CG, CT, TG, TT}에 소속된 부분을 가지는 프로브는 프로브 군 4로 지정된다. 이 집합의 프로브 군에서, 프로브 군의 이름을 아는 것은 그것의 표지가 프로브 군의 이름을 결정하기 위해 검출되었던 새롭게 결찰된 연장 프로브의 근접한 뉴클레오타이드 반대쪽에 위치한 주형에 있는 뉴클레오타이드의 정체에 대한 특정 가능성을 제거한다. 예를 들어 만약 프로브 군 이름이 1이라면, 새롭게 결찰된 연장 프로브의 근접한 뉴클레오타이드는 A 또는 G여서, 주형의 상보하는 뉴클레오타이드는 반드시 T 또는 C여야 한다. 소속된 부분에는 각 위치에 최소한 두 가지의 가능성이 있기 때문에 뉴클레오타이드는 정확하게 확인될 수 없지만, 몇 가지 가능성을 무시하기에 충분한 정보는 바람직한 집합의 프로브 군이 사용될 때 상황과는 대조적으로 단일 사이클로부터 얻어진다.

<283> 본 발명의 어떤 실시예에서, 소속된 부분이 길이가 3 뉴클레오타이드인 부분적으로 소속된 프로브가 사용된다. 그것의 소속된 부분이 길이 3의 모든 가능한 서열을 포함하는 프로브를 함유하기 위해, 프로브 군 집합은 $4^3=64$ 개의 상이한 프로브를 포함하여야 하고, 그것이 바람직하다. 도 29A는 길이가 3 뉴클레오타이드인 (트리뉴클레오타이드) 소속된 부분을 포함하는 프로브를 포함하는 프로브 군 집합에 대하여 소속된 부분을 생성하기 위해 사용될 수 있는 다이어그램을 도시한다. 이 도면은 4세트의 A, G, C, 및 T로 표시된 열과 1, 2, 3 및 4로 표시된 프로브 군 이름의 줄을 나타낸다. 각 세트의 4열은 내부의 뉴클레오타이드 정체와는 반대쪽 칸에 있다. 트리뉴클레오타이드에 대하여 프로브 군을 결정하기 위해, 트리뉴클레오타이드 중 첫 번째 뉴클레오타이드를 확인하는 문자로 표시된 열이 선택된다. 그 열 내에서, 트리뉴클레오타이드 중 두 번째 뉴클레오타이드를 가지는 줄이 선택된다. 트리뉴클레오타이드는 줄의 상부에 표시된 프로브 군에 지정된다. 예를 들어 트리뉴클레오타이드 "TCG"를 프로브 군에 지정하기 위해 다음의 과정이 수반된다: 마지막 뉴클레오타이드가 "G"이기 때문에 고려할 것은 "G"를 함유하는 칸 반대쪽에 위치한 4열의 세트, 즉 세 번째 세트의 열에 제한된다. 첫 번째 뉴클레오타이드가 "T"이기 때문에 고려할 사항은 추가로 4의 세트의 마지막 열로 제한된다. 프로브 군 지정은 중간 뉴클레오타이드를 함유하는 줄의 제목에 의해 결정된다. 중간 뉴클레오타이드가 "C"이기 때문에 트리뉴클레오타이드는 프로브 군 1로 지정된다. 유사한 과정으로 다음의 프로브 군 지정이 산출된다: AAA = 1; ATA = 2; AGA = 3; GTA = 4; GAG = 1; TGG = 2 등. 과정은 가능한 모든 트리뉴클레오타이드가 프로브 군에 지정될 때까지 계속된다.

<284> 도 29B는 3 뉴클레오타이드 길이의 소속된 부분을 가지는 프로브를 포함하는 프로브 군의 집합에 대해 추가의 소속된 부분을 제조하기 위한 과정을 도시한다. 그 과정은 상기에서 설명된 프로브 군의 24개의 바람직한 집단의 각각으로부터 그러한 집합을 제조하기 위해 사용되는데, 그 집합에서 소속된 부분은 길이가 2 뉴클레오타이드이며 집합은 4개의 프로브 군을 함유한다. 바람직한 프로브 군 집합을 나타내는 예시적인 다이어그램은 도면의 상부에 도시된다. 이 다이어그램의 줄은 직접적으로 상부 다이어그램의 각 줄에 지정된 색에 따라 도면의 하부에 있는 줄로 도표화된다. 그러므로 상부 다이어그램의 줄은 좌측으로부터 우측으로 블루, 그린, 옐로우, 그리고 레드이다. 하부 다이어그램의 줄 1에 있는 염기들은 위에서 아래로 블루, 그린, 옐로우, 그리고 레드이며, 이때 4 뉴클레오타이드의 각 세트는 상부 다이어그램의 줄에 상응한다. 하부 다이어그램에서 줄 2, 3, 및 4는 줄 1의 4 뉴클레오타이드의 각 세트를 아래쪽으로 점진적으로 움직임으로써 생성된다.

<285> "프로브 군"는 각각이 동일한 표지를 가진 다수의 상이한 프로브를 포함하는 단일한 "슈퍼-프로브"인 것으로 여겨질 수 있다는 것이 인지될 것이다. 이 경우 프로브를 구성하는 프로브 분자들은 일반적으로 프로브의 어떠한 부분을 가로지르는 실질적으로 동일한 분자의 집단이 아닐 것이다. 용어 "프로브 군"의 사용은 어떠한 제한하는 효과를 가지는 것으로 의도된 것은 아니지만 편의를 위해 그러한 "슈퍼-프로브"를 구성할 프로브의 특성을 설명하기 위해 사용된다.

<286> 코드 해석

<287> 상기에서 설명된 것처럼, 최소한 2개의 구별가능하게 표시된 프로브를 포함하는 프로브 군의 집합을 사용하는 연장, 결찰, 검출 및 절단의 연속적인 사이클은 단일한 서열화 반응으로부터 또는 주형의 상이한 부위로부터 순서 리스트 안으로 개시하는 다중 서열화 반응에서 결정된 프로브 군 이름을 조립하는 것으로부터 프로브 군 이름의 순서 리스트를 산출한다. 수행된 사이클의 수는 원하는 서열의 길이와 대략 같아야 한다. 순서 리스트는 실질적인 양의 정보를 함유하지만 관심 있는 서열을 즉시 산출할 형태는 아니다. 추가의 단계(들), 그것들 중 최소한 하나가 서열에 대한 추가의 정보 중 최소한 하나의 항목을 모으는 것을 포함하는 단계(들)이 관심 있는 서열을 가장 그럴듯하게 나타내는 서열을 얻기 위해 수행되어야 한다. 관심 있는 서열을 가장 그럴듯하게 나타내는 서열은 본원에서는 "정확한" 서열로서 언급되고, 프로브 군의 순서 리스트로부터 정확한 서열을 끌어내는 과정은 "코드해석"으로서 언급된다. 상기에서 설명된 것과 같은 "순서 리스트"의 구성요소들은 리스트를 생성하

는 동안에 또는 그 후에 재배열될 수 있으며, 단 그때에 리스트의 구성요소와 주형의 뉴클레오타이드 사이의 대응성을 포함한 정보 내용이 유지되어야 하고, 재배열, 단편화, 및/또는 치환이 적절하게 코드해석 과정 중에 고려되어야 한다 (아래에서 논의됨). 그러므로 용어 "순서 리스트"는 상기에서 설명된 것과 같이 생성된 순서 리스트의 재배열된, 단편화된, 및/또는 치환된 버전을 포함하는 것으로 의도되며, 단 그러한 재배열된, 단편화된, 및/또는 치환된 버전은 실질적으로 동일한 정보 내용을 포함해야 한다.

<288> 순서 리스트는 다양한 접근법을 사용하여 코드가 해석될 수 있다. 이들 접근법 중 일부는 프로브 군 이름의 순서 리스트로부터 최소한 하나의 후보 서열의 세트를 생성하는 것을 포함한다. 후보 서열의 세트는 목적을 이루기 위해 충분한 정보를 제공한다. 바람직한 실시예에서 하나 또는 그 이상의 추가의 단계는, 후보 서열들 중으로부터 또는 후보 서열이 비교되는 서열 세트로부터 관심 있는 서열을 가장 그럴듯하게 나타내는 서열을 선택하기 위해 수행된다. 예를 들어 한 접근법에서 최소한 하나의 후보 서열의 최소한 일부분이 최소한 하나의 다른 서열과 비교된다. 정확한 서열은 비교를 토대로 선택된다. 본 발명의 어떤 실시예에서 코드해석은 방법을 반복하고 프로브 군의 원래의 집합으로부터 상이하게 코드화된 프로브 군의 집합을 사용하여 프로브 군 이름의 두 번째 순서 리스트를 얻는 것을 포함한다. 프로브 군의 두 번째 순서 리스트로부터의 정보는 정확한 서열을 결정하기 위해 사용된다. 어떤 실시예에서 프로브 군의 교대로 코드화된 집합을 사용하여 연장, 결찰 및 검출의 하나 정도의 적은 사이클로부터 얻어진 정보는 정확한 서열을 선택할 수 있을 정도로 충분하다. 달리 표현하자면, 교대로 코드화된 프로브 군을 사용하여 확인된 첫 번째 프로브 군은 후보 서열이 정확한지를 결정하기에 충분한 정보를 제공한다.

<289> 다른 코드해석 접근법은 어떠한 활용가능한 서열화 방법에 의해, 예컨대 서열화 방법 A의 단일 사이클에 의해 주형의 최소한 하나의 뉴클레오타이드를 특이적으로 확인하는 것을 포함한다. 하나 또는 그 이상의 뉴클레오타이드(들)에 대한 정보는 프로브 군 이름의 순서 리스트를 코드 해석하기 위한 "열쇠"로서 사용된다. 이와 달리, 서열화된 주형의 부분은 그것의 서열이 미지인 영역 외에 공지된 서열의 영역을 포함할 수 있다. 만약 서열화 방법 AB가 미지 서열과 최소한 하나의 공지 서열의 한 뉴클레오타이드를 포함하는 주형의 부분에 적용된다면, 공지 서열은 프로브 군 이름의 순서 리스트의 코드를 해석하기 위한 "열쇠"로서 사용될 수 있다. 다음 단원에서는 후보 서열을 생성하는 방법이 설명된다. 그 다음 단원에서는 후보 서열을 사용하여 공지 서열과 비교함으로써, 두 번째 세트의 후보 서열과 비교함으로써, 그리고 공지 뉴클레오타이드 정체를 활용함으로써 정확한 서열을 선택하는 것에 대해 설명된다.

<290> **후보 서열의 생성**

<291> 서열화된 주형의 영역은 연장, 결찰, 및 절단의 연속적인 사이클에 의해 제조되는 연장된 이중나선에 상보한다는 것이 인지될 것이다. 그러므로 연장된 이중나선에 대한 후보 서열을 생성하는 것은 서열화된 주형의 영역에 대한 후보 서열을 생성하는 것에 동등하다. 실제로 당업자는 서열화된 주형의 영역에 대한 후보 서열을 생성할 수 있거나 또는 연장된 이중나선에 대한 후보 서열을 생성하고 서열화된 주형의 영역에 대한 후보 서열을 결정하기 위하여 그것들의 상보물을 취할 수 있다. 후자의 접근법이 본원에서 설명된다. 프로브 군 이름의 리스트로부터 후보 서열을 생성하기 위하여, 프로브 군 리스트의 첫 번째 구성원이 고려된다. 그 프로브 군과 결합된 구속된 부분의 세트는 구속된 부분의 길이에 동등한 길이까지 서열의 초기 뉴클레오타이드에 대한 가능성을 제한한다. 예를 들어 만약 구속된 부분이 디뉴클레오타이드라면, 연장된 이중나선의 첫 번째 디뉴클레오타이드에 대한 가능한 서열은 그 프로브 군에 속하는 프로브에서 일어나는 그런 구속된 부분에 한정된다 (그로써 서열화된 주형의 영역에 있는 첫 번째 디뉴클레오타이드에 대한 가능한 서열들은 그 프로브 군에 속하는 프로브에서 일어나는 구속된 부분에 상보적이다). 첫 번째 디뉴클레오타이드에 대한 가능성은 전형적으로 컴퓨터에 의해 기록된다. 유사하게, 연장된 이중나선의 두 번째 디뉴클레오타이드 (즉 첫 번째 디뉴클레오타이드로부터 한 뉴클레오타이드 별첨된 디뉴클레오타이드)에 대해 가능한 서열은 그 두 번째 프로브 군에 속하는 프로브에서 일어나는 구속된 부분들에 제한된다 (그리고 따라서 주형의 두 번째 디뉴클레오타이드, 즉 첫 번째 디뉴클레오타이드로부터 한 뉴클레오타이드 별첨된 디뉴클레오타이드에 대해 가능한 서열은 그 두 번째 프로브 군에 속하는 프로브에서 일어나는 구속된 부분에 상보하는 그런 조합에 제한된다). 두 번째 디뉴클레오타이드에 대해 가능한 서열도 또한 기록된다. 연속되는 디뉴클레오타이드에 대한 가능성도 마찬가지로, 결정해야 할 서열의 원하는 길이에 상응하는 디뉴클레오타이드에 대해 가능성이 다 기록되었거나 또는 리스트에 더 이상의 프로브 군이 없을 때까지 기록된다.

<292> 가능성을 기록하는 과정의 대표적인 실례가 도 30에 도시되는데, 도면에서 프로브 군 이름의 리스트는 도 25A에 도시된 프로브 군 집합을 사용하여 생성되는 것으로 여겨진다. 도 30에서 가장 왼쪽에 있는 줄은 위에서 아래의 순서: 옐로우, 그린, 레드, 블루로 프로브 군의 리스트를 도시한다. 리스트의 각 프로브 군에 상응하는 디뉴클레오타이드에 대한 서열 가능성은 도면의 오른쪽에 도시된다. 뉴클레오타이드 위치는 서열 가능성의 위에 표시된다.

서열은 위치 1부터 시작하므로, 첫 번째 디뉴클레오타이드는 위치 1과 2를 차지하고; 두 번째 디뉴클레오타이드는 위치 2와 3을 차지한다. 엘로우 프로브 군에 대해 가능성은 도 30에 도시된 것과 같이 CC, AT, GG, 및 TA이다. 그런 프로브 군에 대한 가능성은 CA, AC, GT, 및 TG이다. 각 디뉴클레오타이드의 가능한 서열을 기록하는 과정은 원하는 서열 길이에 도달할 때까지 계속된다.

<293> 가능성 세트가 생성된 후, 첫 번째 가정은 도 30의 위치 1로서 표시된 서열의 5' 위치에 있을 것으로 가정되는 후보 서열의 첫 번째 뉴클레오타이드의 정체에 대해 이루어진다. 첫 번째 가정은 그 뉴클레오타이드가 A이거나, 그 뉴클레오타이드가 G이거나, 그 뉴클레오타이드가 C이거나, 또는 그 뉴클레오타이드가 T일 수 있다는 것이다.

<294> 각 후보에 대해 가능한 서열은 인접한 디뉴클레오타이드의 가능한 서열에 의해 제한되는 것이 관찰될 것인데, 왜냐하면 인접한 디뉴클레오타이드는 중복하기 때문에, 즉 첫 번째 디뉴클레오타이드의 두 번째 뉴클레오타이드는 또한 두 번째 디뉴클레오타이드의 첫 번째 뉴클레오타이드이기 때문이다. 예를 들어 만약 첫 번째 뉴클레오타이드가 C일 것으로 가정되면, 첫 번째 디뉴클레오타이드는 CC여야 한다. 만약 첫 번째 디뉴클레오타이드가 CC이면, 두 번째 디뉴클레오타이드는 그것의 첫 번째 위치에 C를 가져야 한다. 첫 번째 위치에 C를 가지는 두 번째 디뉴클레오타이드에 대해 유일하게 가능한 서열은 CA이기 때문에, 두 번째 디뉴클레오타이드가 CA임이 명백하다. 그러므로 첫 번째 3 뉴클레오타이드의 서열은 CCA이어야 한다. 마찬가지로 세 번째 디뉴클레오타이드에 대한 가능한 서열은 두 번째 디뉴클레오타이드의 가능한 서열에 의해 제한된다. 만약 두 번째 디뉴클레오타이드가 CA라면, 세 번째 디뉴클레오타이드는 그것의 첫 번째 위치에 A를 가지는 유일한 가능성 때문에 AG임이 틀림없다. 그러므로 처음 4개의 뉴클레오타이드의 서열은 CCAG이다. 이 과정을 계속하면 처음 5 뉴클레오타이드에 대하여 5'-CCAGC-3'의 서열이 나온다. 따라서 CCAGC는 첫 번째 후보 서열이다.

<295> 두 번째 후보 서열은 첫 번째 뉴클레오타이드가 A라는 가정에서 비롯된다. 이 가정으로 첫 번째 디뉴클레오타이드에 대하여 AT가 생성된다. 첫 번째 디뉴클레오타이드에 대하여 AT의 서열과 일치하는 두 번째 디뉴클레오타이드에 대해 유일하게 가능한 서열은 TG이다. 두 번째 디뉴클레오타이드에 대하여 TG의 서열과 일치하는 세 번째 디뉴클레오타이드에 대하여 유일하게 가능한 서열은 GA이다. 세 번째 디뉴클레오타이드에 대하여 GA의 서열과 일치하는 네 번째 디뉴클레오타이드에 대해 유일하게 가능한 서열은 AA이다. 이들 디뉴클레오타이드를 전체 길이의 후보 서열로 조립하면 ATGAA가 나온다. 마찬가지로 첫 번째 뉴클레오타이드가 G라는 가정은 후보 서열 GGTCG를 낳고, 첫 번째 뉴클레오타이드가 T라는 가정은 후보 서열 TACTT를 낳는다. 그러므로 4개의 후보 서열이 생성되고, 각각은 서열의 첫 번째 뉴클레오타이드일 것으로 가정되는 상이한 뉴클레오타이드로부터 시작된다.

<296> 가정이 다른 뉴클레오타이드 중 하나보다는 첫 번째 뉴클레오타이드에 대해 이루어져야 할 필요는 없다. 예를 들어 가정은 네 번째 뉴클레오타이드의 정체에 대해 이루어진 것과 동일할 수 있는데, 그 경우 후보 서열은 주형을 따라 뒤쪽으로 (즉 3'→5' 방향으로) 이동함으로써 생성될 수 있을 것이다. 예를 들어 네 번째 뉴클레오타이드가 T라고 가정하는 것은 네 번째 디뉴클레오타이드가 반드시 TT여야 하고; 세 번째 디뉴클레오타이드가 반드시 CT여야 하며; 두 번째 디뉴클레오타이드가 AC여야 하고; 첫 번째 디뉴클레오타이드가 반드시 CC여야 하는 것을 의미한다. (뉴클레오타이드는 그것의 정체에 서열에서 3'→5'으로 이동함에 의해 생성될지라도 5'→3' 방향으로 쓰여진다.) 이와 달리, 가정은 서열의 중간에 있는 어떠한 뉴클레오타이드에 대해 이루어질 수 있고, 디뉴클레오타이드 정체는 5'→3'과 3'→5' 두 방향으로 이동함으로써 생성된다. 한 뉴클레오타이드에 대한 가정이 없을 때, 각각의 뉴클레오타이드의 정체는 각 위치를 A, G, C, 또는 T가 차지할 수 있기 때문에 완전히 결정되지 못한 채로 남아있음이 인지될 것이다.

<297> 프로브 군의 바람직한 집합을 사용할 때, 어떠한 단일 뉴클레오타이드 (예컨대 첫 번째 뉴클레오타이드)의 정체를 가정함으로써 하나 및 유일하게 하나의 후보 서열이 생성된다. 그러나 덜 바람직한 프로브 군 집합이 사용될 때, 하나 이상의 뉴클레오타이드에 대한 정체를 가정하는 것이 필요할 수도 있다, 즉 첫 번째 뉴클레오타이드에 대한 정체를 가정하는 것은 서열의 나머지를 완전하게 구체화하지 못한다. 예를 들어 덜 바람직한 프로브 군의 집합은 그것의 규정된 서열이 AA 및 AC인 구성원들을 가진 군을 포함할 것이다. 그러한 경우에 첫 번째 뉴클레오타이드가 A라고 가정하는 것은 두 번째 뉴클레오타이드에 대한 두 가지 가능성을 남긴다. 덜 바람직한 프로브 군의 집합을 사용하여 서열화하는 것은 아래에서 추가로 논의된다. 만약 구축된 부분이 비연속적인 뉴클레오타이드를 구성한다면, 상기에서 설명된 접근법은 여전히 약간 변형되어 사용될 수 있다는 것이 인지될 것이다.

<298> **후보 서열을 공지 서열과 비교하는 것에 의한 서열 확인**

<299> 일반적으로 연장된 이중나선의 후보 서열이 상기에서 설명된 바와 같이 결정되었다면, 서열화될 주형의 영역에 대한 상응하는 후보 서열은 그것들의 상보물을 취함으로써 얻어진다. 어떤 경우에, 후보 서열은 그 자체로서 목적을 이루기에 충분한 정보를 제공할 것이다. 예를 들어 만약 서열화의 목적이 단순히 어떤 서열 가능성을 배제

하는 것이라면, 후보 서열을 그러한 가능성과 비교함으로써 충분할 것이다. 도 30에 도시된 후보 서열은 서열화되고 있는 영역이 예를 들어 폴리A 꼬리의 일부가 아니었다는 결정을 가능하게 할 것이다. 더 긴 서열은 서열화되고 있는 영역이 벡터의 일부가 아니라는 것을 확증할 수 있을 것이다.

<300>

많은 경우에 정확한 서열을 명백하게 결정하는 것이 바람직할 것이다. 본 발명의 바람직한 실시예에 따르면, 정확한 서열은 서열화될 주형의 영역에 대한 후보 서열을 공지 서열의 세트와 비교함으로써 확인된다. 공지된 서열의 세트는 예를 들면 관심 있는 특정 유기체에 대한 서열의 세트일 수 있다. 예를 들어 만약 사람 DNA가 서열화된다면, 후보 서열은 사람 드래프트 게놈 서열과 비교될 수 있다. 대중이 활용할 수 있는 인간 게놈 서열 공급원에 대한 안내를 위해서는 URL www.ncbi.nih.gov/genome/guide/human의 웹 사이트를 참조한다. 다른 실례로서, 만약 감염성 병원체 (예컨대 피험체로부터 분리된 박테리아 또는 바이러스)로부터 유도된 핵산이 서열화된다면, 그 박테리아 또는 바이러스의 변이 군주의 서열을 함유하고 있는 데이터베이스가 연구될 수 있다. 완전하거나 부분적인 서열을 함유하고 있는 그러한 많은 유기체-특이적 데이터베이스가 당해 기술분야에 알려져 있고, 서열화 노력을 가속화하고 있기 때문에 더욱 널리 활용될 것이다. 대표적인 실례는 마우스에 대한 데이터베이스 (예컨대 URL www.ncbi.nlm.nih.gov/genome/seq/MmHome.html을 가지는 웹사이트 참조), 사람 면역결핍성 바이러스에 대한 데이터베이스 (예컨대 URL hiv-web.lanl.gov/content/hiv-db/mainpage.html을 가지는 웹사이트 참조), 말라리아 중 플라스모디움 팔시파룸(*Plasmodium falciparum*)에 대한 데이터베이스 (예컨대 URL www.tigr.org/tdb/edb2/pfal/htmls/index.shtml을 가지는 웹사이트 참조) 등을 포함한다. 물론 유기체-특이적 세트의 서열을 사용할 필요는 없다. 광범위한 유기체 및 바이러스로부터의 서열을 함유하는 GenBank(URL www.ncbi.nlm.nih.gov/Genbank/을 가지는 웹사이트)와 같은 데이터베이스가 연구될 수도 있다. 데이터베이스는 주형이 유도되는 유기체 또는 바이러스로부터의 어떠한 서열까지 함유할 필요는 없다. 일반적으로 서열은 게놈 서열, cDNA 서열, EST 등일 수 있다. 다중 서열이 연구될 수도 있다.

<301>

단순히 연구를 수행하는 것이 목적을 이루기 위해 충분할 수 있다. 예를 들어 만약 바이러스 핵산이 환자로부터 분리된다면, 후보 서열을 그 바이러스의 공지 서열 세트와 비교함으로써 매치되는 서열을 조사하지 않더라도 바이러스 핵산이 그 바이러스로부터의 서열을 함유하는지 또는 함유하지 않는 지를 결정할 수 있다. 매치의 존재는 환자가 바이러스로 감염되었는지를 확증하는 한편, 매치의 부재는 환자가 바이러스로 감염되지 않았음을 나타내는 것일 것이다.

<302>

어떤 실시예에서 공지 서열의 세트는 더 협의적인 범위의 서열을 함유하며, 그것은 그것에 대해 서열화가 수행되는 목적에 특이하게 맞춰질 수 있다. 그러므로 서열화되는 핵산에 대한 정보는 공지 서열의 세트를 선택하기 위해 사용될 수 있다. 예를 들어 만약 주형이 특정한 유전자의 서열을 나타내는 것으로 알려져 있다면, 그 공지 서열은 관심 있는 주어진 유전자좌(locus)에서 유전자, 돌연변종 및 야생형 서열의 상이한 대립형질을 나타낼 수 있다. 단지 후보 서열을 어떤 후보 서열이 정확한지를 결정하기 위하여 단일한 공지 서열과 비교하는 것만이 필요할 것이다. 예를 들어 본 발명의 어떤 실시예에서 주형은 관심 있는 영역을 함유하는 DNA를 증폭하여 얻어진다 (예컨대 관심 있는 영역 양 옆에 있는 프라이머를 사용하여). 관심 있는 영역은 돌연변이 또는 다형성, 예컨대 특정 질병과 관련된 돌연변이 또는 다형성이 존재할 수 있는 부위를 포함할 수 있다. 만약 주형이 관심 있는 특정 유전자로부터의 서열을 나타낸다면, 후보 서열은 단지 그 영역에 대한 단일한 참조 서열, 예컨대 서열의 야생형이나 돌연변이 형태와 비교될 필요가 있다. 달리 표현하면, 만약 주형의 서열의 부분 또는 전부가 알려져 있다면, 다수의 공지 서열과의 비교를 수행할 필요가 없을 수도 있다. 대신에 공지 서열의 전부 또는 부분을 포함하는 후보 서열이 정확한 것으로 선택된다. 예를 들어 *BRCA1* 및 *BRCA2* 유전자의 돌연변이는 유방암의 증가하는 위험과 관련된 것으로 알려져 있고, 어떤 피험체가 그러한 돌연변이를 운반하는 지를 결정하는 데 상당한 관심이 있다. 만약 주형이 *BRCA1* 유전자로부터의 서열을 포함하는 것으로 알려져 있다면, 예컨대 유전자의 일부를 포함하는 관심 있는 영역의 양 옆에 있는 프라이머가 주형의 클론 집단을 생성하는 데 사용되었다면, 후보 서열은 정확한 서열을 결정하기 위하여 단지 야생형 또는 돌연변이 *BRCA1* 서열에 대해 비교될 필요가 있다.

<303>

보다 일반적인 경우에 후보 서열을 공지 서열의 세트와 비교하는 것은 후보 서열들 중 어느 하나에 유사한 어떠한 공지 서열을 확인할 것이다. 만약 충분한 길이의 후보 서열이 제공되지만 한다면, 데이터베이스가 하나 이상의 후보 서열과 동일하거나 밀접하게 닮은 서열을 함유할 가능성은 매우 작다. 달리 표현하면, 만약 후보 서열이 충분히 길다면, 하나 이상의 후보 서열이 공지 서열의 세트로 표시될 것 같지는 않다. 후보 서열은 "매치"인 것으로 여겨지는 어떠한 서열과 비교된다. 전형적으로 매치가 존재한다는 것을 수립하기 위해 필요한 동일성의 정도에 대한 경계를 설정하는 것이 바람직할 것이다. 예를 들어 공지 서열은 후보서열과 공지 서열이 최소한 50%, 최소한 60%, 최소한 70%, 최소한 80%, 최소한 90%, 최소한 95%, 최소한 99%, 또는 심지어 100% 동일하다면 매치인 것으로 여겨질 것이다. 전형적으로 % 동일성은 길이가 최소한 10 뉴클레오타이드, 예컨대 10 내지 15

뉴클레오타이드, 15 내지 20 뉴클레오타이드, 20 내지 25 뉴클레오타이드, 25 내지 30 뉴클레오타이드 등의 윈도우에 대해 평가될 것이다. 윈도우의 길이는 다양한 상이한 기준, 예를 들면 그것들에 한정되는 것은 아니지만 다수의 공지 서열에 있는 서열의 수, 다수의 공지 서열의 정체 또는 공급원 등에 따라 선택될 수 있다. 예를 들어 만약 후보 서열이 GenBank와 같은 큰 데이터베이스의 서열에 대해 비교된다면, 더 적은 수의 서열을 함유하고 있는 데이터베이스가 사용되는 것보다 더 긴 길이를 사용하는 것이 바람직할 것이다. 발명의 어떤 실시예에서 서열은 서로 인접할 필요는 없는 다수의 상이한 윈도우를 가로질러 비교된다. 바람직하게는 윈도우의 조합된 길이는 최소한 10 뉴클레오타이드 길이, 예컨대 10 내지 15 뉴클레오타이드, 15 내지 20 뉴클레오타이드, 20 내지 25 뉴클레오타이드, 25 내지 30 뉴클레오타이드 등이다. 어떤 경우에 공지 서열의 세트에 있는 다수의 서열은 매치될 수 있다. 서열은 예를 들면 주형이 유도되는 것과 같은 유기체에서 발견되는 상동성 유전자, 상이한 유기체로부터의 상동성 유전자, 위유전자(pseudogene), cDNA 및 게놈 서열 등을 나타낼 수 있다.

<304> 일반적으로 공지 서열의 세트에 있는 서열을 가장 밀접하게 닮은 후보 서열이 정확한 서열로서 선택된다. 이와 달리, 예컨대 만약 서열화 방법이 고비용의 에러로 수행될 수 있다는 믿음에 타당한 이유가 있다면, 데이터베이스로부터의 상응하는 서열을 정확한 것으로서 선택하는 것이 바람직할 것이다. 예를 들어 만약 에러 비율이 예정된 경계값 이상인 것으로 알려진다면 데이터베이스로부터의 서열을 정확한 서열로서 선택하는 것이 바람직할 것이다.

<305> 다중 후보 서열에 대해 매치가 발견될 가능성을 보장하기 위해서 필요한 길이는 예컨대 그것들에 한정되는 것은 아니지만 공지 서열의 특정 세트, 매치를 수용하기 위한 경계 등을 포함한 다양한 고려사항에 따라 좌우될 것이다. 일반적으로 길이가 ~25 내지 26 뉴클레오타이드인 서열이 전형적인 유기체의 게놈에서 일단 유일하게 나타날 것이다. 그러므로 대략 이 길이의 후보 서열을 생성하는 것이 정확한 서열을 확인하는 데에 충분하다. 일반적으로 후보서열은 최소한 10 뉴클레오타이드의 길이여야 하고, 바람직하게는 최소한 15, 최소한 20 뉴클레오타이드 길이, 예컨대 20 내지 25, 25 내지 30, 30 내지 35, 35 내지 40, 45 내지 50, 또는 심지어는 그 이상의 길이여도 좋다.

<306> **첫 번째 후보 서열의 세트를 두 번째 후보 서열의 세트와 비교하는 것에 의한 서열 확인**

<307> 발명의 어떤 실시예에서 코드해석은 첫 번째 코드화에 따라 코드화된 프로브 군의 첫 번째 집합을 사용하여 프로브 군의 첫 번째 순서 리스트를 생성하고, 그것으로부터 후보 서열의 첫 번째 세트를 생성한 후, 두 번째 코드화에 따라 코드화된 두 번째 프로브 군 집합을 사용하여 동일한 주형으로부터 두 번째 프로브 군의 순서 리스트를 생성하고 그것으로부터 두 번째 후보 서열의 세트를 생성함으로써 수행된다. 새롭게 합성된 DNA 가닥은 두 개의 서열화 반응 사이에 주형으로부터 제거되거나, 또는 동일한 서열의 주형이 두 번째 프로브 군 집합을 사용하여 서열화된다. 후보 서열의 세트가 비교된다. 어떤 집합의 프로브 군이 사용되는 지에 관계없이, 후보 서열 중 하나는 정확한 서열일 것인 한편으로 다른 것은 정확한 것이 아니라는 것 (또는 기껏해야 부분적으로 정확하다는 것)이 인지될 것이다. 그러므로 모든 세트의 후보 서열이 정확한 서열을 함유할 것이지만, 대부분의 경우 어떠한 주어진 세트의 후보 서열의 다른 후보 서열은 다른 세트의 후보 서열에서 발견되는 것들과는 다를 것이다. 그러므로, 단순히 2 세트의 후보 서열을 비교함으로써 정확한 서열이 결정될 수 있다. 2개의 상이하게 코드화된 집합의 프로브 군을 사용하여 동일한 길이의 후보 서열을 생성할 필요는 없다. 발명의 바람직한 실시예에서 두 번째 프로브 군 집합을 사용하여 생성된 후보 서열은 2 뉴클레오타이드처럼 짧거나, 또는 동등하게 두 번째 프로브 군 집합을 사용하여 생성된 프로브 군의 순서 리스트는 1 요소 (즉 결찰과 검출의 단일 사이클)처럼 짧을 수도 있다.

<308> 도 31A 내지 도 31C는 두 개의 구별가능하게 표지된 바람직한 프로브 군을 사용하여 후보 서열을 생성하는 것과 코드 해석하는 것의 실례를 도시한다. 도 31A는 첫 번째 코드화에 따라 코드화된 프로브 군의 바람직한 집합을 도시한다. 도 31B는 프로브 군 옐로우, 그린, 레드, 블루의 순서 리스트로부터 4개의 후보 서열의 생성 ("2314"로 표시될 수 있는데, 레드=1, 옐로우=2, 그린=3, 블루=4이다)과, 정확한 서열이 CAGGC (굵게 표시됨)로 가정되는 것을 도시한다. 도 31C는 두 번째 코드화에 따라 코드화된 프로브 군의 바람직한 집합을 도시한다. 주형에 있는 첫 번째 디뉴클레오타이드가 CA이기 때문에, 옐로우 프로브 군에 있는 가장 위에 있는 프로브는 첫 번째 연장 사이클에서 연장가능한 말단에 결찰될 것이다. 이것은 첫 번째 디뉴클레오타이드에 대한 다음 세트의 후보서열: CA, TC, GG, AT를 유발한다. 첫 번째 프로브 군 집합을 사용하여 생성된 후보 서열 중에서 오직 CAGGC 서열만이 이들 디뉴클레오타이드 중 어떠한 것을 사용하여 시작한다. 그러므로 그것이 정확한 서열이어야 한다. 일반적으로 첫 번째와 두 번째 프로브 군 집합이 다음 기준을 만족하는 것이 바람직하다: 첫 번째와 두 번째 프로브 군 집합이 비교될 때 (i) 첫 번째 집합에 있는 프로브 군의 각각의 4개 프로브 중 3개가 두 번째 집합의 새로운 프로브 군에 지정되어야 하고; (ii) 3개의 제외된 프로브 각각이 두 번째 집합의 상이한 프로

브 군에 다시 지정되어야 한다.

<309> **공지 뉴클레오타이드 정체를 사용하는 프로브 군의 순서 리스트의 코드 해석**

<310> 상기에서 설명된 바와 같이, 후보 서열은 연장된 이중나선 또는 주형의 단일 뉴클레오타이드에 대한 정체를 가정함으로써 생성될 수 있다. 사용된 특이한 프로브 군 집합에 따라, 일반적으로 그것은 최소한 4개의 후보 서열을 생성하는 데 필요할 것이다. 그러나 다수의 후보 서열의 생성은 만약 주형에서의 최소한 하나의 뉴클레오타이드의 정체 (및 따라서 연장된 이중나선에서의 정체)가 알려진다면 피할 수 있다. 그 경우 단일 후보 서열을 생성할 필요가 있을 뿐이다. 후보 서열을 생성하는 방법은 상기에서 설명된 것과 동일하다. 주형의 최소한 하나의 뉴클레오타이드의 정체는 어떠한 서열화 방법, 이를테면 그것들에 한정되는 것은 아니지만 서열화 방법 A, 구별가능하게 표지된 뉴클레오타이드와 증합효소를 사용하여 개시 올리고뉴클레오타이드로부터 프라이머 연장법 등을 사용하여 결정될 수 있다. 주형의 하나 또는 그 이상의 뉴클레오타이드는 서열화 방법 AB 이외의 서열화 방법을 사용하여 먼저 서열화될 수 있고, 그런 다음 개시 올리고뉴클레오타이드와 어떠한 연장 생성물이 제거될 수 있으며, 동일한 주형에 대해 서열화 방법 AB를 사용한 서열화가 수행된다 (또는 그 역).

<311> 다른 접근법은 단순히 그것의 서열이 결정될 부분 외에 공지된 정체의 하나 또는 그 이상의 공지 뉴클레오타이드를 함유하는 주형을 서열화하는 것이다. 예를 들어 개시 올리고뉴클레오타이드가 결합하는 영역과 미지의 서열이 서열화되기 시작하는 영역 사이의 주형의 부분은 하나 또는 그 이상의 정체가 공지된 뉴클레오타이드를 포함할 수 있다. 주형의 이 부분에 서열화 방법 AB를 적용함으로써, 서열의 하나 또는 그 이상의 뉴클레오타이드의 정체는 예정될 것이고, 따라서 단일 후보서열을 생성하는 데 사용될 수 있으며, 그것은 정확한 서열이 될 것이다.

<312> 그러므로 상기에서 설명된 방법은 (i) 어떤 정체가 공지된 정체의 뉴클레오타이드 및 그것의 근접한 뉴클레오타이드가 공지된 정체의 뉴클레오타이드에 인접한 뉴클레오타이드의 반대쪽에서 결합된 프로브의 구축된 부분의 가능한 서열과 일치하는 지를 결정함으로써 공지된 정체의 뉴클레오타이드에 인접한 주형의 뉴클레오타이드에 정체를 지정하는 단계; (ii) 그것의 근접한 뉴클레오타이드가 연속적인 뉴클레오타이드의 반대쪽에서 결합된 프로브의 구축된 부분의 가능한 서열과 그 정체가 일치하는 지를 결정함으로써 연속되는 뉴클레오타이드에 정체를 지정하는 단계; 및 (iii) 서열이 결정될 때까지 단계 (ii)를 반복하는 단계를 포함한다. 이들 단계는 연장된 이중나선에 대해 동일한 단계들을 수행하는 것과 동등한데, 왜냐하면 연장된 이중나선과 서열화될 주형의 영역 사이에는 정확한 대응성이 있기 때문이다.

<313> **덜 바람직한 프로브 군을 사용하는 서열화**

<314> 덜 바람직한 프로브 군 집합은 바람직한 프로브 군 집합이 사용되는 방법과 유사한 방식으로 서열화 방법 AB를 수행하기 위해 사용될 수 있다. 그러나 그 결과는 많은 관점에서 다를 수 있다. 예를 들어 서열의 어떤 부분들은 후보 서열로부터 추가의 정보를 필요로 하지 않고서도 전체적으로 확인될 수 있다. 도 32는 도 28에 도시된 것과 같이 코드화된 프로브 군의 덜 바람직한 집합을 사용하여 서열을 결정한 실패를 도시한다. 서열 결정은 일반적으로 바람직한 프로브 군 집합에 대해 설명된 것과 같이 진행된다. 관심 있는 주형은 서열 "GCATGA"를 가지며, 그 결과 프로브 군의 순서 리스트로서 "12341"이 생성된다. 위치 1에 있는 뉴클레오타이드가 A라고 가정하면 후보 서열로서 "ACATGA"가 생성된다. 그러나 바람직한 프로브 군 집합을 사용하는 경우와는 달리, 표지 "1"이 첫 번째 뉴클레오타이드로서 A를 가지는 2개의 상이한 디뉴클레오타이드, 즉 "AA" 및 "AG"와 결합되므로, 두 번째 뉴클레오타이드에 대해서는 두 가지 가능성이 있다. 그러므로 위치 1에 있는 뉴클레오타이드가 A라고 가정하면 두 번째 후보 서열로서 "ACATGC"가 생성된다. 위치 1에 있는 뉴클레오타이드가 G라고 가정하면 후보 서열로서 "GCATGA"가 생성되고 또한 후보 서열로서 "GCATGC"가 생성된다. 표지 "1"이 위치 1에서 C 또는 T를 가지는 어떠한 디뉴클레오타이드와도 결합되지 않기 때문에 "C"나 "T"로 시작되는 후보 서열은 생성되지 않는다. 도 32는 상호간에 배열된 4가지 후보 서열을 도시한다. 모든 후보 서열의 중간 4개의 뉴클레오타이드는 CATG인 것이 관찰될 것이다. 그러므로 정확한 서열은 위치 2-5에 CATG를 포함하여야 한다. 만약 이들 뉴클레오타이드만이 관심 있는 것이라면, 추가의 코드해석 단계를 수행할 필요는 없다.

<315> 상기에서 언급한 것과 같이, 프로브 군의 집합은 4개의 상이한 프로브 군을 구성할 필요는 없지만, 2보다 큰, 4^N 개까지의 어떠한 수를 구성할 수 있으며, 이때 N은 구축된 부분의 길이이다. 그러나 만약 4보다 적은 군이 사용된다면 4 이상의 후보 서열을 생성할 필요가 있으며, 한편 만약 4 이상의 프로브 군이 사용된다면, 추가의 표지가 필요할 것이다. 이들 및 다른 이유들로 4개의 프로브 군을 구성하는 집합이 바람직하다.

<316> **상호간에 후보 서열을 비교함에 의한 서열 확인**

<317> 발명의 어떤 실시예에서 관심 있는 서열의 부분 또는 전체는 상호간에 후보 서열을 비교함으로써 결정될 수 있

다. 일반적으로 그러한 비교는 어떤 후보 서열이 그것의 전체 길이를 통해 정확한지를 결정하는 데에는 충분하지 않을 것이다. 그러나 만약 둘 또는 그 이상의 후보 서열이 동일하거나 서열의 일부에 걸쳐 충분히 유사하다면, 이 정보는 상기에서 설명된 바와 같이 그 부분 내에서 주형의 뉴클레오타이드의 서열을 명백하게 확인하는 데 충분할 것이다.

<318> 필요에 따라 주형은 다르게 코드화된 프로브 군을 사용하여 한 번 또는 그 이상 추가로 확인된 서열을 가지는 추가의 부분들을 생성하기 위해 서열화될 수 있다. 이들 부분은 원하는 길이의 서열을 조립하기 위해 조합될 수 있다.

<319> **프로브 군을 사용한 에러 보정**

<320> 동일한 DNA 서열의 전부 또는 부분을 나타내는 다수의 주형을 서열화하고 그 서열들을 일렬배열하는 것이 때로 바람직하다. 만약 주형이 관심 있는 영역을 부분적으로만 함유한다면, 더 긴 서열이 중복하는 단편들을 조립함으로써 얻어진다. 예를 들어 유기체의 게놈을 서열화할 때 전형적으로 DNA는 단편화되고, 충분한 단편들이 서열화되어서 각 신장된 DNA가 여러 (예컨대 4-12) 상이한 단편들에서 나타난다. 중복하는 서열들을 더 긴 서열로 조립하기 위한 컴퓨터 소프트웨어가 당업자에게 공지되어 있다.

<321> 종래의 서열화 방법이 사용될 때, 다수의 단편이 단편들 중 하나 (변칙적인 단편으로서 언급됨)가 그 영역 내에 있는 단일 위치에서 다른 것들과 다른 것을 제외하면 영역에 걸쳐 완벽하게 일렬배열되는 경우는 흔하다. 분리된 차이가 서열화 에러를 나타내는지 또는 본래의 차이 (예컨대 단일 뉴클레오타이드 다형성)가 그 위치에 존재하는지를 결정하는 것이 문제가 될 수도 있다.

<322> 본 발명은 서열화 방법 AB를 사용하여 에러 조사를 수행하는 신규한 방법을 제공한다. 그 방법에 따르면, 동일하게 신장된 DNA를 나타내는 단편을 포함하는 주형이 상기에서 설명된 것과 같이 구별가능하게 표지된 프로브의 집합을 사용하여 서열화되고, 그 결과 각 주형에 대한 프로브 군의 순서 리스트가 생성된다. 프로브 군의 순서 리스트는 일렬로 정렬된다. 만약 단일 위치에서 다른 단편과 다른 한 리스트만을 제외하고 여러 리스트가 예정된 길이, 예컨대 리스트의 10, 15, 20, 또는 25 또는 그 이상의 요소에 걸쳐 완벽하게 정렬된다면, 그 차이는 서열화 에러에 돌릴 수 있다. 만약 실제 다형성이 존재한다면, 변칙적인 단편으로부터 생성된 순서 프로브 리스트는 둘 또는 그 이상의 인접한 위치에서 다른 단편으로부터 생성된 순서 프로브 리스트와 다를 것이다.

<323> 예를 들어 표 1의 코드화 4를 사용하는 바람직한 프로브 군의 집합을 사용하는 서열화 방법 AB를 서열 5'-CAGACGACAAGTATAATG-3'을 포함하는 주형에 적용하면 다음: "23324322132444142"의 프로브 군 순서 리스트가 아래와 같이 생성된다:

<324> 23324322132444142

<325> CAGACGACAAGTATAATG

<326> 만약 실제 SNP (예컨대 CAGACGAGAAGTATAATG, 밑줄친 뉴클레오타이드는 다형성 부위를 나타낸다)가 있다면, 그것은 리스트: 23324333132444142 (밑줄친 SNP의 결과로서 일어난 변화를 나타낸다)의 두 개의 연속적인 요소의 변화를 초래한다. 프로브 군의 순서 리스트와 SNP를 함유하는 서열 사이의 대응성은 아래와 같다:

<327> 23324333132444142

<328> CAGACGAGAAGTATAATG

<329> 그러나 결찰된 연장 프로브와 결합된 표지를 확인할 때의 에러는 프로브 군의 순서 리스트에 단일 에러를 초래하고 전방의 그 지점으로부터 그 결과의 후보 서열에 변화를 초래한다. 예를 들어 일곱 번째 결찰된 연장 프로브 23324332132444142 (밑줄친 숫자는 잘못 확인된 표지를 나타낸다)와 결합된 표지를 결정할 때의 에러는 그 결과의 후보 서열을 CAGACGAGTTCATATTAC로 바꾸고, 이 서열에서 밑줄친 부분은 서열화 에러의 결과로서 일어난 변화를 나타낸다. 프로브 군의 순서 리스트와 서열 사이의 대응성은 아래와 같다:

<330> 23324332132444142

<331> CAGACGAGTTCATATTAC

<332> 3개의 염기, 4개의 표지 도식을 사용할 때, SNP를 함유하는 단편은 변칙적인 단편에 대한 프로브 군의 순서 리스트의 3개의 연속적인 차이를 초래하는 한편, 서열화 에러는 단지 하나의 차이만을 초래한다. 예를 들어 도 29에 도시된 것과 같이 코드화된 프로브 군의 집합이 사용될 때 서열 CAGACGACAAGTATAATG에 대한 프로브 군의 순

서 리스트는 다음과 같다:

- <333> 2322224132412244
- <334> CAGACGACAAGTATAATG
- <335> SNP를 함유하는 변칙적인 단편, 예컨대 CAGACGAGAAGTATAATG는 SNP를 함유하지 않는 단편으로부터 생성된 순서 리스트에 비하여 3개의 연속적인 위치에서 상이한 프로브 군의 순서 리스트를 초래할 것이고, 다음과 같다:
- <336> 2322213332412244
- <337> CAGACGAGAAGTATAATG
- <338> 서열화 에러는 프로브 군의 순서 리스트에 오직 하나의 차이를 초래할 것이고 에러 전방의 지점으로부터 완전히 상이하게 생성된 후보서열을 초래하게 될 것이다.
- <339> 그러므로 단편 (변칙적 단편)으로부터 생성된 프로브 군의 순서 리스트가, 동일한 스트레치의 DNA를 나타내지만 분리된 단일 위치에 있는 다른 순서 리스트와는 상이한 다른 단편으로부터 생성된 프로브 군의 순서 리스트와 정렬될 때, 차이를 함유하는 순서 리스트가 서열화 에러 (프로브 군의 잘못된 확인)를 나타내는 것 같다. 단편 (변칙적 단편)으로부터 생성된 프로브 군의 순서 리스트가, 동일한 신장된 DNA를 나타내지만 2 또는 그 이상의 연속적인 위치에 있는 다른 순서 리스트와는 상이한 다른 단편으로부터 생성된 프로브 군의 순서 리스트와 정렬될 때, 변칙적인 단편이 SNP를 함유하는 것 같다. 바람직하게는 프로브 군의 순서 리스트의 정렬된 부분은 길이가 최소한 3 또는 4 요소, 바람직하게는 최소한 6, 8, 또는 그 이상의 길이이다. 바람직하게는 정렬된 부분은 최소한 66% 동일하거나, 최소한 70% 동일하거나, 최소한 80% 동일하거나, 최소한 90% 동일하거나, 또는 그 이상, 예컨대 100% 동일하다.
- <340> 유사하게, 단편에 대한 후보 서열이, 서열의 첫 번째 부분에 걸쳐 동일한 신장된 DNA를 나타내지만 서열의 두 번째 부분에 걸쳐서는 다른 단편에 대한 후보 서열과 실질적으로 상이한 다른 단편에 대한 후보 서열과 정렬될 때, 서열화 에러가 발생하는 것 같다. 단편에 대한 후보 서열이, 서열의 두 부분에 걸쳐 동일한 스트레치의 DNA를 나타내지만 단일 위치에서는 상이한 다른 단편에 대한 후보 서열들과 정렬될 때, 변칙적인 단편은 SNP를 함유하는 것 같다. 바람직하게는 후보 서열의 정렬된 부분은 길이가 최소한 4 뉴클레오티드이다. 바람직하게는 정렬된 부분은 최소한 66% 동일하거나, 최소한 70% 동일하거나, 최소한 80% 동일하거나, 최소한 90% 동일하거나, 또는 그 이상, 예컨대 100% 동일하다.
- <341> 그러므로 본 발명은 서열화 에러로부터 단일 뉴클레오티드 다형성을 구별하는 방법을 제공하는데, 그 방법은 (a) 서열화 방법 AB를 사용하여 다수의 주형을 서열화하는 단계, 이때 주형은 단일 핵산 서열의 중복하는 단편들을 나타내고; (b) 단계 (a)에서 얻어진 서열들을 정렬하는 단계; 및 (c) 만약 서열이 첫 번째 부분을 가로질러 실질적으로 동일하고 두 번째 부분을 가로질러 실질적으로 상이하다면 서열화 에러를 나타내는 서열들 사이의 차이를 결정하는 단계로 이루어지며, 이때 각 부분의 길이는 최소한 3 뉴클레오티드이다. 본 발명은 추가로 서열화 에러로부터 단일한 뉴클레오티드 다형성을 구별하는 방법을 제공하는데, 그 방법은 (a) 단일 핵산 서열의 중복하는 단편들을 나타내는 다수의 주형을 사용하여 서열화 방법 AB를 사용함으로써 프로브 군의 다수의 순서 리스트를 얻는 단계; (b) 단계 (a)에서 얻어진 프로브 군의 순서 리스트를 최소한 90% 동일한 리스트 내에 배열된 영역을 얻기 위하여 정렬하는 단계; 및 (c) 만약 리스트가 정렬된 영역 내의 한 위치에서만 상이하다면 서열화 에러를 나타내는 프로브 군의 순서 리스트 사이의 그 차이를 결정하는 단계; 또는 (d) 만약 리스트가 정렬된 영역 내의 둘 또는 그 이상의 연속적인 위치에서 상이하다면 단일 뉴클레오티드 다형성을 나타내는 프로브 군의 순서 리스트 사이의 차이를 결정하는 단계로 이루어진다.
- <342> **위치를 벗어난 정보 수집**
- <343> 당해 기술분야에 잘 알려져 있는 것과 같이, "비트"(2진 숫자)는 기초 2로 한 자리 숫자를 나타내는, 달리 표현하자면 1 또는 0을 말하며, 디지털 데이터의 가장 작은 단위를 나타낸다. 뉴클레오티드가 4개의 상이한 정체 중 어느 하나를 가질 수 있기 때문에 뉴클레오티드의 정체를 구체화하는 것은 2 비트를 필요로 한다는 것이 인지될 것이다. 예를 들어 A, G, C, 및 T는 각각 00, 01, 10, 및 11로 표시될 수 있을 것이다. 구별가능하게 표지된 프로브 군의 바람직한 집합의 프로브 군의 이름을 구체화하는 것은 4개의 구별가능하게 표지된 프로브 군이 있기 때문에 2 비트를 필요로 한다.
- <344> 가장 관련적인 형태의 서열화에서 및 서열화 방법 A에서, 각각의 뉴클레오티드는 불연속적인 단위로서 확인되고, 동시에 한 뉴클레오티드에 상응하는 정보가 모아진다. 각 검출 단계는 단일 뉴클레오티드로부터 2 비

트의 정보를 얻는다. 대조적으로 서열화 방법 AB로는 각 검출 단계에서 다수의 뉴클레오타이드의 각 뉴클레오타이드로부터 2보다 적은 비트의 정보를 얻는 한편, 바람직한 프로브 군의 집합이 사용될 때 여전히 검출 단계당 2 비트의 정보를 얻는다. 프로브 군의 순서 리스트의 각 프로브 군 이름은 주형의 최소한 2개의 뉴클레오타이드의 정체를 나타내며, 이때 정확한 수는 프로브의 서열 결정 부분의 길이에 의해 결정된다. 예를 들어 표 1의 코드화 4에 따라 코드화된 프로브 군의 집합을 사용하여 서열 5'-CAGACGACAAGTATAATG-3'으로부터 얻어진 프로브 군의 순서 리스트를 생각해보자:

<345> 23324322132444142

<346> CAGACGACAAGTATAATG

<347> 프로브 군 2는 디뉴클레오타이드 CA가 프로브 군 2의 프로브에 존재하는 명시된 부분 중 하나이기 때문에 리스트의 첫 번째 프로브 군이다. 프로브 군 3은 디뉴클레오타이드 AG가 프로브 군 3의 프로브에 존재하는 명시된 부분 중 하나이기 때문에 리스트의 두 번째 프로브 군이다. 상기에서 언급된 바와 같이, 4개의 프로브 군이 있기 때문에, 각 프로브 군 정체는 2 비트의 정보를 나타낸다. 그러므로 각 검출 단계는 2개의 뉴클레오타이드에 대하여 2 비트의 정보를 모으고, 그 결과 각 뉴클레오타이드로부터 평균 한 비트의 정보가 발생한다.

<348> 그러므로 본 발명은 서열을 결정하는 방법을 제공하는데, 그 방법은 연장, 결찰, 및 검출의 다중 사이클을 포함하며, 검출 단계는 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 주형의 최소한 2 뉴클레오타이드의 각각으로부터 평균 2 비트의 정보를 동시에 얻는 것을 포함한다. 본 발명은 나아가 첫 번째 올리고뉴클레오타이드 프로브 군 집합을 사용하여 주형 폴리뉴클레오타이드에 있는 뉴클레오타이드들의 서열을 결정하는 방법을 제공하는데, 그 방법은 (a) 순차적인 사이클의 연장, 결찰, 검출, 및 절단을 수행하는 단계, 이때에 평균 2 비트의 정보가 어떠한 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻는 일 없이 각 사이클 동안 주형의 최소한 2 뉴클레오타이드의 각 뉴클레오타이드로부터 동시에 얻어지고; 및 (b) 단계 (a)에서 얻어진 정보를 최소한 한 비트의 추가의 정보와 조합하여 서열을 결정하는 단계로 이루어진다. 발명의 다양한 실시예에서 최소한 한 비트의 추가의 정보는 주형의 뉴클레오타이드의 정체, 후보 서열을 최소한 하나의 공지 서열과 비교함으로써 얻어진 정보; 및 두 번째 올리고뉴클레오타이드 프로브 군 집합을 사용하여 방법을 반복함으로써 얻어진 정보로 이루어지는 군으로부터 선택된 항목을 포함한다.

<349> 그러므로 방법이 개별적인 뉴클레오타이드로부터 2 비트의 정보를 얻지 않는 한편, 평균 2 비트의 정보는 각 사이클에서 주형으로부터 모아지며, 단 프로브 군의 바람직한 집합이 사용될 때는 위치를 벗어난 방식으로 모아진다. 2 또는 3 프로브 군의 집합이 사용될 때 2보다 적은 비트의 정보는 각 사이클 동안에 모아진다.

<350> 위치를 벗어난 정보 수집은 상기에서 설명된 것과 같은 여러 조사 방법의 적용을 허용하는 것을 포함하여 많은 장점을 갖는다. 또한 주형의 각 뉴클레오타이드가 바람직한 실시예에서 한 번 이상 조사되기 때문에, 위치를 벗어난 정보 수집은 특정 뉴클레오타이드와 결합된 형광발색단을 검출함에 있어서 체계적인 편향성을 피하는 것을 보조할 수 있다.

<351> 본원에서 설명된 프로브 군 및 프로브 군의 집합은 프로브의 연장, 결찰, 및 절단의 연속적인 사이클을 포함하는 방법에 더불어 다양한 서열화 방법에 사용될 수 있다. 본 발명은 또한 상기에서 설명된 것과 같은 서열 및 구조를 가지는 프로브 군 및 프로브 군 집합을 제공하는데, 이때 프로브는 임의로 잘라지기 쉬운 연쇄를 함유하지 않는다. 예를 들어 프로브는 단지 인산이에스테르 골격 연쇄를 함유할 수 있거나 및/또는 트리거 잔기를 함유하지 않을 수 있다. 발명의 어떤 실시예에서 프로브 군은 연장과 결찰의 연속적인 사이클을 사용하여 서열화를 수행하기 위해 사용될 수 있지만, 각 사이클 동안에 절단은 포함하지 않는다. 예를 들어 프로브 군은 W02005021786에 및 당해 기술분야의 그 밖의 곳에서 설명된 것과 같은 결찰-기반 방법에서 사용될 수 있다. 그러한 방법에서 프로브 군을 사용하기 위하여 프로브 상의 표지는 절단가능한 링커, 예컨대 W02005021786에 개시된 것과 같은 것들에 의해 부착되어야 하며, 그로써 핵산의 잘라지기 쉬운 연쇄를 절단함이 없이 제거될 수 있다. 그러한 방법은 예를 들어 W02005021786에 설명된 결찰 카세트보다는 프로브 군을 사용하여 동시에 또는 순차적으로 다중 반응을 수행한 후 프로브 군의 리스트를 조립함으로써, 프로브 군의 순서 리스트를 생성하는 데 사용될 수 있다. 리스트는 상기에서 설명된 것처럼 코드해석된다.

<352> I. 키트

<353> 본 발명의 상이한 실시예들을 수행하기 위하여 다양한 키트가 제공될 수 있다. 어떤 키트는 포스포로티올레이트 연쇄를 포함하는 연장 올리고뉴클레오타이드 프로브를 포함한다. 키트는 추가로 하나 또는 그 이상의 개시 올리고뉴클레오타이드를 포함할 수 있다. 키트는 포스포로티올레이트 연쇄를 절단하기에 적당한 절단제, 예컨대 AgNO_3 와

그 안에서 절단이 수행되는 적절한 완충제를 함유할 수 있다. 어떤 키트는 손상된 염기 또는 비염기성 잔기를 함유하는 뉴클레오시드와 같은 트리거 잔기를 포함하는 연장 올리고뉴클레오타이드 프로브를 포함한다. 키트는 추가로 하나 또는 그 이상의 개시 올리고뉴클레오타이드를 포함할 수 있다. 키트는 뉴클레오시드와 인접한 비염기성 잔기 사이의 연쇄를 절단하기에 적당한 절단 시약 및/또는 폴리뉴클레오타이드로부터 손상된 염기를 제거하기에 적당한 시약, 예컨대 DNA 글리코실라제를 함유할 수 있다. 어떤 키트는 2당 뉴클레오타이드를 포함하는 올리고뉴클레오타이드 프로브를 함유하고, 절단 시약으로서 과요오드산염을 함유한다. 어떤 실시예에서 키트는 구별가능하게 표지된 올리고뉴클레오타이드 프로브 군의 집합을 함유한다.

<354> 키트는 추가로 결합 시약 (예컨대 결합 효소, 완충제 등)과 발명의 특정 실시예를 실시하기 위한 지시사항을 포함할 수 있다. 사용될 수 있는 다른 효소들, 예컨대 포스파타제, 중합효소를 위한 적절한 완충제가 포함될 수 있다. 어떤 경우에 이들 완충제는 동일할 수 있다. 키트는 또한 주형을 고착시키기 위한 지지체, 예컨대 자성 비드를 포함할 수 있다. 비드는 PCR 증폭을 수행하기 위한 프라이머로 기능화될 수 있다. 다른 임의의 성분으로는 세척 용액; PCR 증폭을 위한 주형을 삽입하기 위한 벡터; 증폭 프라이머, 열안정성 중합효소, 뉴클레오타이드와 같은 PCR 시약; 에멀션을 제조하기 위한 시약; 겔을 제조하기 위한 시약 등이 있다.

<355> 어떤 바람직한 키트에서, 포스포로티올레이트 연쇄를 포함하는 형광 표지된 올리고뉴클레오타이드가 프로브의 상이한 말단 뉴클레오타이드에 상응하는 프로브가 뚜렷한 스펙트럼으로 분해가능한 형광 염료를 운반하도록 제공된다. 보다 바람직하게는, 4개의 그러한 프로브가 제공되어 각각의 4개의 스펙트럼으로 분해가능한 형광 염료와 프로브의 4개의 가능한 말단 뉴클레오타이드 사이의 일대일 대응성이 가능하다.

<356> 식별자, 예컨대 바 코드, 방사성 주파수 ID 태그, 등이 키트에 또는 키트 위에 존재할 수 있다. 식별자는 예컨대 품질 조절, 재고 조절, 트랙킹, 워크 스테이션 사이의 이동 등의 목적에 대해 키트를 독특하게 확인하기 위해 사용될 수 있다.

<357> 키트는 일반적으로 하나 또는 그 이상의 용기를 포함함으로써 특정한 개별적인 시약이 별도로 수납될 수 있다. 키트는 또한 시판용의 상대적으로 밀폐된 공간에 개별적인 용기를 넣기 위한 수단, 예컨대 그 안에 지시사항, 스티로폼과 같은 충전 물질 등을 넣을 수 있는 플라스틱 박스를 포함할 수 있다.

<358> J. 자동 서열화 시스템

<359> 본 발명은 다수의 주형으로부터 동시에, 즉 실질적으로 동시에 서열 정보를 모으기 위하여 사용될 수 있는 다양한 자동 서열화 시스템을 제공한다. 바람직하게는 주형은 실질적으로 평면인 기관 위에 배열된다. 도 21은 본 발명의 시스템 중 하나의 사진을 도시한다. 사진의 상부에서 볼 수 있는 것과 같이, 본 발명의 시스템은 CCD 카메라, 형광 현미경, 이동가능한 스테이지, 펠트에 유동 셀, 온도 조절기, 유체 조작 장치 및 전용 컴퓨터를 포함한다. 이들 부품에 대해 다양한 치환이 이루어질 수 있다는 것이 인지될 것이다. 예를 들어 또 다른 영상 캡처 장치가 사용될 수 있다. 이 시스템에 대한 추가의 상세한 내용은 실시예 9에서 제공된다.

<360> 본 발명의 자동 서열화 시스템과 관련된 영상 처리 방법 및 소프트웨어는 본원에 설명된 결찰-기반 방법과 다른 방법들, 이를테면 그것들에 한정되는 것은 아니지만 합성에 의한 형광 제자리 서열화와 같은 합성 방법에 의한 서열화 (FISSEQ) (Mitra RD, et al., Anal Biochem., 320(1):55-65, 2003)를 포함한 다양한 서열화 방법을 실시하기 위해 사용될 수 있다는 것이 인지될 것이다. 본원에서 설명된 결찰-기반 서열화 방법에 대한 경우와 같이, FISSEQ는 반-고체 지지체에 또는 그 위에 직접 고정된 주형에 대해, 반-고체 지지체에 또는 그 위에 있는 마이크로입자에 고정된 주형에 대해, 기관에 직접 부착된 주형에 대해 실시될 수 있다.

<361> 본 발명의 시스템의 한 가지 중요한 측면은 유동 셀이다. 일반적으로 유동 셀은 그것을 통해 유체가 흐를 수 있는 입력 및 출력 구멍을 가지는 챔버를 포함한다 (미국 특허 제 6,406,848호 및 제 6,654,505호 및 PCT 공개 번호 제 W098053300호 참조). 유체의 흐름은 다양한 시약이 첨가되고 유동 셀 안에 위치한 실체들 (예컨대 주형, 마이크로입자, 분석물 등)로부터 제거되는 것을 가능하게 한다.

<362> 바람직하게는 발명의 서열화 시스템에 사용하기에 적당한 유동 셀은 기관, 예컨대 실질적으로 평면인 기관, 예를 들면 슬라이드가 장착되어서 유체가 그 기관의 표면 위로 흐를 수 있고, 원도우가 조명, 여기, 신호 획득 등을 하는 것을 가능하게 하는 위치를 포함한다. 발명의 방법에 따라서 마이크로입자와 같은 실체는 전형적으로 그것이 유동 셀 안에 놓이기 전에 기관 위에 배열된다.

<363> 발명의 어떤 실시예에서 유동 셀은 수직으로 배향되어 있어서 유동 셀의 상부로부터 기포가 빠져나가는 것을 가능하게 한다. 유동 셀은 유체가 유동 셀의 바닥으로부터 상부로 지나갈 수 있도록, 예컨대 유입구가 셀의 바닥에 있고 출구가 셀의 상부에 있도록 배열된다. 도입될 수 있는 어떠한 기포든지 부력이 있기 때문에 기포는 조

명 윈도우를 가리는 일 없이 출구로 빠르게 부양한다. 기포가 액체의 밀도에 비해 더 낮은 밀도에 의해 액체의 표면으로 상승하는 것이 가능한 이 접근법은 본원에서 "중량에 의한 기포 치환"으로서 언급된다. 그러므로 본 발명은 중량에 의한 기포 치환을 가능하게 하기 위하여 배향된 유동 셀을 포함하는 서열화 시스템을 제공한다. 바람직하게도 직접 또는 간접적으로 부착된 (예컨대 공유적으로 또는 비 공유적으로 기관에 연결되어 있음) 마이크로입자를 가지는 기관 또는 기관에 점착성인 또는 고정된 반-고체 지지체에 또는 그 위에 고정된 마이크로입자를 가지는 기관은 유동 셀 안에 수직으로 장착된다. 즉 기관의 가장 큰 평면 표면이 바닥면에 대해 수직이다. 바람직한 실시예에서 마이크로입자는 지지체 또는 기관에 또는 그 위에 고정되기 때문에, 마이크로입자는 실질적으로 서로에 대해 고정된 위치를 유지하며, 그것은 영상의 연속적인 획득 및 영상 등록을 용이하게 한다.

<364> 도 24A 내지 도 24J는 다양한 배향의, 본 발명의 유동 셀 또는 그것의 일부의 개략적인 다이어그램을 도시한다. 본 발명의 유동 셀은 다양한 목적, 이를테면 그것에 한정되는 것은 아니지만 분석 방법 (예컨대 서열화와 같은 핵산 분석 방법, 혼성화 분석 등), 단백질 분석 방법, 결합 분석, 스크리닝 분석 등의 목적 중 어느 하나에 대해 사용될 수 있다. 유동 셀은 또한 합성을 수행하기 위해, 예컨대 조합 라이브러리를 제조하기 위해 사용될 수 있다.

<365> 도 22는 본 발명의 다른 자동 서열화 시스템의 개략적인 다이어그램을 도시한다. 유동 셀은 온도-조절된, 자동화된 스테이지 (실시예 9에서 설명되는 것과 유사함)에 장착되고, 유체 조작 시스템, 예컨대 다중-구멍 밸브가 부착된 시린지(syringe) 펌프에 부착된다. 스테이지는 하나의 유동 셀이 영상화되는 동안 다른 단계, 예컨대 연장, 결찰, 및 절단이 다른 유동 셀에서 계속해서 수행될 수 있도록 하기 위해 다수의 유동 셀을 수용한다. 이 접근법은 값비싼 광학 시스템의 활용을 최대화하는 한편, 처리량을 증가시킨다.

<366> 유체 라인에는 기포를 검출하고 시약 용도를 모니터하기 위하여 광학 및/또는 컨덕턴스 센서가 장착되어 있다. 유체공학 시스템의 온도 조절 및 센서는 시약이 긴 시간 동안 적절한 온도에서 안정하게 유지되지만 어닐링, 결찰 및 절단 단계 동안 온도 변동을 피하기 위해 시약이 유동 셀에 유입됨에 따라 작업 온도로 상승하는 것을 보장한다. 시약은 바람직하게는 로딩의 에러를 방지하기 위하여 키트에 미리 팩킹된다.

<367> 광학은 4개의 카메라를 포함하는데, 각각 4개의 필터 세트 중 하나를 통하여 한 영상을 취한다. 광표백의 효과를 감소시키기 위하여, 조명 광학이 단지 영상화될 면적만을 조명하도록, 필드의 가장자리의 다중 조명을 피하도록 공학처리될 수 있다. 영상화 광학은 표준 고감도-고정된 현미경 대물렌즈와 표준 빔-스플리터 및 필터로부터 이루어질 수 있다. 표준 2,000×2,000 픽셀 CCD 카메라가 영상을 얻기 위해 사용될 수 있다. 시스템은 적절한 기계식 지지체를 광학을 위해 삽입한다. 조명 강도는 바람직하게는 분석 소프트웨어에 의하여 추후 사용을 위하여 모니터링되고 기록된다.

<368> 다수의 영상 (예컨대 각각의 실시예에서 대략 1800 또는 그 이상의 비-중복 영상 필드)을 신속하게 얻기 위하여, 시스템은 바람직하게는 빠른 자동초점 시스템을 사용한다. 영상 자체의 분석을 토대로 하는 자동초점 시스템은 당해 기술분야에 잘 알려져 있다. 이것들은 일반적으로 포커싱(focusing) 사건당 최소한 5 프레임을 필요로 한다. 이것은 포커싱 영상을 얻기 위하여 필요한 엑스트라 조명의 관점에서 보면 느리기도 하고 비싸기도 한다 (광표백을 증가시킴). 발명의 어떤 실시예에서 또 다른 자동초점 시스템, 예컨대 기계식 시스템이 반응할 수 있을 정도로 빠르게 초점을 맞출 수 있는 독립된 광학을 토대로 하는 시스템이 사용된다. 그러한 시스템은 기술분야에 알려져 있으며, 예를 들면 소비자 CD 플레이어에 사용되는, CD가 회전함에 따라 실시간으로 서브-미크론의 포커싱을 유지하는 초점 모으기 시스템이 있다.

<369> 발명의 어떤 실시예에서 시스템은 원격 조작으로 작동된다. 특이한 프로토콜을 실행하기 위한 스크립트가 중심 데이터베이스에 저장될 수 있고 각 서열화 작동에 대해 다운로드될 수 있다. 샘플은 샘플 트래킹과 최종 데이터를 가지는 샘플과의 결합의 통합성을 유지하기 위해 바코드화될 수 있다. 중심에서의 실시간 모니터링으로 과정 에러의 빠른 해결이 가능할 것이다. 어떤 실시예에서 기기에 의해 모아진 영상들은 즉시 중심의 다중 테라바이트 저장 시스템 및 하나 또는 그 이상의 프로세서(들)의 बैं크로 업로드된다. 중심의 데이터베이스로부터의 트래킹 데이터를 사용하여 프로세서(들)은 영상을 분석하고 서열 데이터를 생성하며, 임의로 매트릭스, 예컨대 바탕 형광 수준과 비드 밀도를, 예를 들면 기기 성능에 대한 정보를 얻기 위하여 처리한다.

<370> 펌프, 스테이지, 카메라, 필터, 온도 조절을 적절하게 차를 정하고 영상 데이터에 주석을 달고 저장하기 위하여 조절 소프트웨어가 사용된다. 유저 인터페이스가 예를 들면 오퍼레이터가 기기를 세팅하고 유지하는 것을 보조하기 위하여 제공되고, 그것은 바람직하게도 슬라이드를 로딩/언로딩하고 유체 라인을 프라임하기 위한 스테이지의 위치를 정하는 기능을 포함한다. 오퍼레이터에게 다양한 작동 파라미터들, 예컨대 온도, 스테이지 위치, 전류 광학 필터 형태, 작동 프로토콜의 상태 등을 나타내기 위해 디스플레이 기능이 포함될 수 있다. 바람

직하게는 시약의 양 및 샘플 ID와 같은 트랙킹 데이터를 기록하기 위해 데이터베이스를 향한 인터페이스가 포함된다.

<371> K. 영상 및 데이터 처리 방법

<372> 본 발명은 컴퓨터 코드 (예컨대 소프트웨어)가 컴퓨터 판독가능한 매질에 기록됨에 따라 최소한 부분적으로 실행될 수 있는 다양한 영상 및 데이터 처리 방법을 제공한다. 하기 실시예 9와 10에서 한층 더 상세한 설명이 제시된다. 또한 일반적으로 서열화 방법 A와 B는 대체로 포함된 처리 단계들, 예컨대 다중 서열화 반응에서 모은 데이터 트랙을 유지하고, 그러한 데이터를 조립하고, 후보 서열을 생성하고, 서열 비교를 수행하는 등의 단계들을 수행하기 위하여 적절한 컴퓨터 소프트웨어를 사용한다.

<373> L. 서열 정보를 저장하는 컴퓨터-판독가능한 매질

<374> 또한 본 발명은 본 발명의 서열화 방법을 적용함으로써 생성된 정보를 저장하는 컴퓨터-판독가능한 매질을 제공한다. 정보는 가공하지 않은 데이터 (즉 추가로 처리되거나 분석되지 않은 데이터), 처리되거나 분석된 데이터 등을 포함한다. 데이터는 영상, 숫자 등을 포함한다. 정보는 데이터베이스로, 즉 전형적으로 검색을 용이하게 하기 위해 배열된, 예컨대 컴퓨터 메모리에 저장된 정보 (예컨대 데이터)의 집합으로 저장될 수 있다. 정보는 예를 들면 서열 및 서열과 관련된 어떠한 정보, 예컨대 서열의 부분, 참조 서열과 서열의 비교, 서열 분석의 결과, 계층 정보, 예컨대 다형성 정보 (예컨대 특정 주형이 다형성을 함유하는지의 여부) 또는 돌연변이 정보 등, 연쇄 정보 (즉 예컨대 염색체에 있는 다른 핵산 서열에 대하여 핵산 서열의 물리적 위치에 속하는 정보), 질병 관련 정보 (즉 피험체의 신체적 특징에 대한 질병에 대한 민감성의 존재와 상관관계가 있는 정보, 예컨대 피험체의 대립형질) 등을 포함한다. 정보는 샘플 ID, 피험체 ID 등과도 관련이 있을 수 있다. 샘플, 피험체 등과 관련된 추가의 정보로는 이를테면 그것들에 한정되는 것은 아니지만, 샘플의 공급원, 샘플에 대해 수행된 처리 단계, 정보의 해석, 샘플 또는 피험체의 특징 등이 있다. 본 발명은 또한 상기 언급된, 예컨대 컴퓨터-판독가능한 매질에 저장된 정보들 중 어느 하나를 컴퓨터-판독가능한 형식으로 수용하는 것을 포함하는 방법을 포함한다. 그 방법은 추가로 바람직하게는 컴퓨터-판독가능한 매질에 저장된 정보를 토대로 한 진단, 예후 또는 전조가 되는 정보를 제공하는 단계 또는 그 정보를 단순히 세 번째 부분에 제공하는 단계를 포함할 수 있다.

<375> 다음 실시예는 본 발명을 예시할 목적으로 제공되며 본 발명을 제한하려는 의도는 아니다.

실시예

<376> 실시예 1: 포스포로티올레이트가 첨가된 올리고뉴클레오타이드의 효과적인 절단 및 결찰

<377> 이 실시예는 3'-S 포스포로티올레이트 연쇄를 함유하는 연장 올리고뉴클레오타이드의 효과적인 결찰 및 절단을 증명하는 실험을 설명한다.

<378> 재료 및 방법

<379> 결찰 서열화 프로토콜

<380> **주형 제조:** 올리고뉴클레오타이드 결찰 및 절단 사이클에 의한 서열화의 잠재성을 평가하고, 이 방법의 어떤 측면들에 있어서 변화를 주었을 때의 효과를 조사하기 위하여 두 세트의 비드-기반 주형 집단을 제조하였다. 바람직한 실시를 위하여, 실시예들에 설명된 대로, 올리고뉴클레오타이드 결찰 및 절단 사이클에 의해 가닥들을 3'→5' 방향으로 연장한다. 따라서, 결찰 효율을 평가하기 위하여, 주형 모델을 5' 말단에서 비드에 결합시키고, 3' 말단에 동일한 결합 영역을 갖도록 디자인하였다. 한 세트는 이중 비오틴 부분을 통해 스트랩트아비딘-코팅된 자성 비드 (1 미크론)에 결합된 짧은 (70bp) 올리고뉴클레오타이드들로 구성되었다. 이들 짧은 주형 집단에 속한 각각은 동일한 프라이머 결합 영역 (40bp)과 3' 말단에 독특한 서열 영역 (30bp)을 갖도록 디자인되었다. 짧은 올리고뉴클레오타이드 주형 집단은 결찰 서열화 주형 1-7 (LST1-7)로 명명하였다.

<381> 비드-기반 주형 집단의 두 번째 세트는 각 주형 집단에 183bp의 스페이서 서열(사람 p53 엑손으로부터)을 삽입함으로써 유도된 PCR-생성된 긴 DNA 단편 (232pb)으로부터 디자인하였다. 주형은 이중 비오틴-함유 전방 프라이머와, 짧은 주형 집단과 동일한 30개 염기의 독특한 3' 말단 서열을 함유하는 역 프라이머를 이용하여 증폭시켰다. 주형을 수산화나트륨-함유 완충액을 이용하여 가닥 중 하나를 용융함으로써 단일-가닥으로 만들었다. 이들 긴 주형 집단은 공동 계류종인 특허출원에 설명된 짧은-단편 쌍-말단 라이브러리로부터 생성된 종들을 모방하도록 디자인하였으며, 긴-LST1-7로 명명하였다.

<382> **프라이머 혼성화:** 100 μM FAM-표지된 프라이머 2.5 μL를 100 μL의 1X 클레노우 (Klenow) 완충액과 예비 혼합하

였다. 이 용액을 완충액을 제거한 후 주형을 부착시킨 30 μ L의 자성 비드($10^6/\mu$ L) 분취액에 첨가하고, 그 결과의 용액을 잘 혼합하였다. 주형/프라이머 혼성화가 일어나도록 놓아둔 후 (혼성화 반응은 65°C에서 2분, 40°C에서 2분 그리고 얼음에서 2분 동안 수행되었다), 프라이머/완충액을 제거하고, 1E 완충액으로 비드를 3회 세척한 다음, TENT 완충액 (10mM 트리스, 2mM EDTA, 30mM NaOAc, 및 0.01% 트리톤 X-100 함유) 중에 300 μ L($10^6/\mu$ L)로 재현탁하였다.

<383> **결찰 1:** 다음으로, LigSeq-FAM과 혼성화된 2.5×10^6 LST7 비드를 1 μ L의 100 μ M LST7-1 노나머, 4 μ L의 5X T4 결합 효소 완충액(Invitrogen), 14 μ L의 H₂O, 및 1 μ L의 T4 결합 효소 (1u/ μ L, Invitrogen)를 함유하는 혼합물 중에서 37°C에서 30분 동안 배양하였다.

<384> **절단 1:** 그런 다음, 비드를 100 μ L씩의 LSWash1 (1X TE, 30mM 아세트산 나트륨, 0.01% 트리톤 X-100 함유)로 3회 세척하였다; 이 용액의 100 μ L 분취액을 취하여 분석용으로 따로 보관하였다. 그런 다음, 비드(1X)를 100 μ L의 30mM 아세트산 나트륨으로 세척하였다. 이 용액에 50 μ L의 50mM AgNO₃를 첨가하고, 그 결과의 혼합물을 37°C에서 20분 동안 배양하였다. AgNO₃를 제거하고, 100 μ L의 30mM 아세트산 나트륨으로 비드를 1회 세척하였다. 그런 다음, 비드를 100 μ L씩의 LSWash1로 3회 세척하고, 90 μ L의 Wash(TENT 완충액) 중에 재현탁한 다음, 이 용액의 10 μ L 분취액을 취하여 분석용으로 따로 보관하였다.

<385> **결찰 2:** TENT 완충액을 제거한 후에, 비드를 14 μ L의 H₂O 중에 재현탁하고, 1 μ L의 100 μ M LST7-5 노나머, 4 μ L의 5X T4 결합 효소 완충액(Invitrogen), 및 1 μ L의 T4 결합 효소(1u/ μ L, Invitrogen)를 함유하는 혼합물과 함께 37°C에서 30분 동안 배양하였다.

<386> **절단 2:** 비드를 100 μ L씩의 LSWash1 (1X TE, 30mM 아세트산 나트륨, 0.01% 트리톤 X-100)로 3회 세척하고, 45 μ L의 Wash 1E에 재현탁하였다. 이 혼합물의 15 μ L 분취액을 취하여 분석용으로 따로 보관하였다. 그런 다음, 비드를 100 μ L의 30mM 아세트산 나트륨으로 한번 더 세척하고, 5 μ L의 20mM 아세트산 나트륨 중에 재현탁하였다. 비드에 50 μ L의 50mM AgNO₃를 첨가하고, 이 혼합물을 37°C에서 20분 동안 배양하였다. AgNO₃를 제거한 후, 100 μ L의 30mM 아세트산 나트륨으로 비드를 한번 더 세척하였다. 그런 다음, 비드를 100 μ L씩의 LSWash1로 3회 세척하고, 30 μ L의 Wash 1E 중에 재현탁하였다. 이 혼합물의 20 μ L 분취액을 취하여 분석용으로 따로 보관하였다.

<387> **결과**

<388> 이 실험은 도 8을 참조하여 더욱 잘 이해될 것이다. 도 8의 상부는 실험 과정의 전체적인 윤곽을 나타낸다. 개시 올리고뉴클레오타이드 (프라이머)를 주형 (LST7로 표시)에 혼성화한 다음, 비오틴 결합을 통해 비드에 부착하였다. 개시 올리고뉴클레오타이드는 5' 포스페이트를 함유하였고, 3' 말단에서 FAM으로 형광 표지하였다. 2개의 9-량체 (노나머) 올리고뉴클레오타이드 프로브(첫 번째 절단가능한 올리고 및 두 번째 절단가능한 올리고)를 포스포로티올레이트 티미딘 염기(sT)(밀줄)를 함유하도록 합성하였다. 첫 번째 절단가능한 프로브를 T4 DNA 결합 효소를 사용하여 프라이머의 연장가능한 말단에 결합한 다음, 질산은을 사용하여 절단하였다. 절단에 의해 연장 프로브의 말단 5개 뉴클레오타이드가 제거되었고, 프라이머에 결합된 채로 남은 프로브 부분에 연장가능한 말단이 생성되었다. 그런 다음, 이 연장가능한 말단에 두 번째 절단가능한 프로브를 결합한 다음 유사하게 절단하였다.

<389> 형광 모세관 전기영동 겔 이동 분석을 이용하여 결찰 및 절단 단계들을 모니터하였다. 이 분석에서는 5' 포스페이트가 올리고뉴클레오타이드 프로브가 도입되는 결찰 기질로서 작용할 수 있도록 주형 가닥에 프라이머를 혼성한다 (형광발색단이 이동성-기반 모세관 겔 전기영동의 리포터로서 작용한다). 각 단계 후에 비드의 분취액을 분석용으로 취하였다. 올리고뉴클레오타이드 프로브의 결찰에 이어서, 자석을 사용하여 자성 비드를 수집하고, 열변성에 의해 주형 비드로부터 프라이머와 그것에 결합된 프로브(들)로 구성된 결합된 종들을 방출시킨 다음, 표지된 크기 표준물질 (리사민 래더; 크기 범위 15-120 뉴클레오타이드; 크로마토그램에서 한 세트의 오렌지색 피크로서 나타남, 도 8 참조)을 구비한 자동 DNA 서열화 기기 (ABI 3730)를 이용하여 형광 모세관 전기영동을 수행하였다. 전형적인 겔 이동에서 가능한 피크들은, i) 프라이머 피크 (연장이 없거나 프라이머 연장의 결핍으로 인한), ii) 아데닐화 피크 (DNA 결합 효소의 작용에 의한 비생산적 결찰 접합부의 5' 말단에 아데노신 잔기의 부착으로 인한 - 도 8F의 메커니즘 참조, 또한 Lehman, I.R., *Science*, 186:790-797, 1974 참조) 및 iii) 완성 피크 (올리고 프로브의 부착으로 인한)를 포함한다. 결찰 효율의 평가에 겔 이동 분석법을 이용하는 것의 한 이점은 피크 밀면적이 각 종들의 농도와 직접 상관된다는 점이다.

- <390> 도 8A는 T4 DNA 결합 효소 및 인산이에스테르 결합만을 함유하는 정확한 매치 프로브를 이용하여 수행된 대조표준 결합을 나타낸다 (도 8A의 좌측에 도시됨). 오렌지색 피크가 크기 마커이다. 좌측에 있는 청색 피크는 결합이 없을 때의 프라이머의 위치를 나타낸다. 정확한 매치 프로브 결합의 결과 좌측으로 이동한다 (화살표). 도 8B는 동일한 조건에서 티올레이트 T 염기를 함유하는 프로브를 이용하여 수행된 결합을 나타낸다 (도 8B의 좌측에 도시됨). 대조표준 프로브에서 관찰된 것과 동일한 이동이 보였다(화살표). 그런 다음, 결합된 포스포로티올레이트 프로브를 함유하는 비드-연결된 주형 집단을 질산은과 함께 배양하여 프로브를 절단하였다. 겔 이동 분석에 의하여 좌측 이동된 4bp 절단 생성물이 증명됨으로써 절단은 효과적임이 확인되었다 (도 8C). 예상된 절단 생성물을 도 8C의 좌측에 나타낸다. 그런 다음, 절단된 비드-기반 주형 집단을 두 번째 결합에 노출시켰으며, 이것은 우측 이동된 13bp 연장 생성물의 출현에 의해 생산적 결합임이 증명되었다 (도 8D). 예상된 절단 생성물을 도 8D의 좌측에 나타낸다. 두 번째 절단으로 확인된 대로 효과적인 다수의 절단 단계가 달성될 수 있으며, 이것은 예상된 좌측 이동된 8bp 절단 생성물에 의해 증명된다 (도 8E).
- <391> 이들 결과는 포스포로티올레이트 결합을 함유하는 프로브의 연속적인 결합 및 절단을 증명한다.
- <392> 이들 실험에서 결합이 100% 완성도까지 진행되지 않았다는 것은 분명하지만, T4 DNA 결합 효소를 이용한 다른 실험들에서보다는 더 높은 완성도가 관찰되었다 (아래 참조). 결합은 완벽히 진행되는 것이 물론 바람직하지만, 그것이 반드시 필요하지는 않다. 예를 들어, 상기 설명된 결합 단계 후에 5' 포스포타세로 처리함으로써 어떠한 결합되지 않은 5' 말단을 "캡핑(cap)"하는 것이 가능하다. 그러나, 이 경우, 결합 가능한 분자의 축소로 인해 연속적 결합의 수가 제한될 것이다. 주어진 수의 연속 결합에 있어서, 판독 길이는 각 결합/절단 사이클 후에 남은 프로브의 길이, 및 서열화 반응의 수에 좌우될 것이며, 각 서열화 반응 후에는 프라이머의 제거 및 프라이머 결합 부위의 다른 부분에 결합된 프라이머의 혼성화가 뒤따르는데, 이것은 주어진 주형에서 수행될 수 있다 (이것은 "재설정"의 수로도 언급된다). 이것으로 인해 프로브의 5' 말단을 향해 절단가능한 결합이 위치된 더 긴 프로브의 사용이 논의된다. 본 발명자들의 실험에서는 옥타머 및 더 긴 프로브보다 헥사머 프로브가 더 많은 양의 결합 불가능한 아데닐화 생성물을 초래한다. 따라서, 옥타머 및 더 긴 프로브는 실질적으로 완벽하게 결합될 것이다 (아래 참조). 게다가, 헥사머 프로브의 5' 말단에 형광 부분을 첨가하는 것은 결합의 효율을 감소시키는 것 같지만, 옥타머 프로브에 형광 부분을 첨가하는 것은 거의 또는 전혀 영향이 없었다. 이런 이유들 때문에 옥타머 또는 더 긴 프로브의 사용이 바람직할 것으로 생각된다.
- <393> 추가 실험들 (하기 설명된)에서 포스포로티올레이트 결합 및 축퇴성-감소 뉴클레오티드를 함유하는 프로브의 결합 및 절단; 결합된 연장 프로브의 3' 말단 특이성 및 선택성; 겔 내부 결합 및 절단; 신호 손실이 최소화된 프라이머 혼성화와 제거의 연속 사이클; 3'→5' 연장에 대한 T4 또는 Taq 결합 효소의 100% 충실도; 및 결합된 연장 프로브의 4-색 스펙트럼 분해능력이 증명되었다. 이 방법들을 수행하기 위한 자동화 시스템을 구성하였다.
- <394> **실시예 2: 축퇴성-감소 뉴클레오티드를 함유하는 포스포로티올레이트 첨가된 올리고뉴클레오티드의 효과적인 절단 및 결합**
- <395> 프로브 길이에 대한 고려사항과 경합하는 것은 연장된 올리고뉴클레오티드의 충실도 및 뒤이은 결합 효율에 대한 그것의 효과이다. T4 DNA 결합 효소의 충실도는 접합부 다음의 5번째 염기 이후에서 급격히 감소하는 것으로 나타났다 (Luo et al., Nucleic Acid Res., 24:3071-3078 및 3079-3085, 1996). 만약 새로운 결합 접합부의 5' 쪽에 미스매치가 도입된다면, 결합 효율은 축소에 의하여 감소될 수 있지만, 마땅 신호의 단계 제거나 증가는 생성되지 않을 것이다 (합성 방법에 의한 중합효소-기반 서열화에서 부딪히는 주요 장애).
- <396> 바람직하게도, 프로브 세트는 어떤 DNA 서열과도 혼성화할 수 있어야 하는데, 이로써 특성화되지 않은 DNA의 새로운 서열화가 허용된다. 그러나, 표지된 프로브 세트의 복잡성은 4-배 축퇴성 염기들의 길이 및 수에 따라서 지수함수적으로 성장한다. 게다가, 복합체 프로브 세트는 합성에 있어서 더욱 도전이 되는 동시에, 모든 프로브 종들은 대략 동일한 표시형태를 유지하므로 정제하기가 어려워진다. 또한 각 종들에 대해 일정한 농도를 유지하기 위해서는 더 높은 농도의 프로브 혼합물이 필요하다. 이런 복잡성을 관리하는 한 방식은 4-배 축퇴성 염기들 대신에 어떤 위치에 디옥시이노신과 같은 보편적 염기를 결합시킨 뉴클레오티드를 사용하는 것이다.
- <397> 4-배 축퇴성 염기 (N; 등몰량의 A, C, G, T)와 옥타머 내의 여러 위치에서 보편적 염기인 이노신 (I)을 사용하여 12개의 올리고뉴클레오티드 프로브를 디자인하였다 (이노신은 B-DNA의 4개 표준 염기 중 하나와 두 자리 수소결합을 할 수 있다: 이노신 염기쌍의 안정성 순서는 I:C > I:A > I:T ~ I:G). 이들 프로브 디자인을 평가하는 한 목적은 이노신 염기의 존재하에 여전히 효과적인 결합을 지지하는 한편으로 얼마나 낮은 옥타머 복잡성이 이루어질 수 있을 것인지를 결정하기 위한 것이었다.

- <398> 초기 연구에서 몇 개의 올리고뉴클레오타이드 프로브를 T4 DNA 결합 효소를 이용하여 비드-기반 주형(긴-LST1)에 결합하였다. 결합시, 형광발색단-표지된 프라이머 (3'FAM 프라이머)가 결합된 올리고뉴클레오타이드 프로브의 양에 비례하여 우측으로 이동한다. 프로브 디자인 NI8-9가 최고 수준의 완성도를 나타냈으며, 이때 프로브의 효과적인 결합으로 인해 프라이머 집단의 99%를 초과하는 양이 우측으로 이동한다 (도 9 참조). 이들 반응은 25℃에서 수행되었다; 반응 온도가 37℃까지 증가하였을 때, 결합은 다소 덜 효과적이었고, 완성률은 더욱 가변적이었다.
- <399> 이 데이터의 더욱 면밀하게 조사한 결과, 접합부 (밀줄)의 3' 쪽의 최초 5개의 뉴클레오타이드 내에 더 적은 이노신 염기가 있는 프로브가 더 높은 결합 효율을 나타냈다. 결합 효율에 대한 잠재적인 서열 내용의 효과를 더욱 조사 평가하기 위하여, 결합 접합부의 3'의 최초 5개 염기 내에 오직 1개의 이노신 잔기를 갖는 4개의 올리고뉴클레오타이드 프로브 디자인을 모든 주형에 대해서 스크리닝하였다. 도 10은 T4 DNA 결합 효소를 사용하는 다수의 주형에 대해 선택된 프로브 조성을 갖는 겔 이동 분석을 이용하여 평가된 결합 완성도를 나타낸다. 이들 초기 실험으로부터의 데이터는 결합 효율을 나타냈는데, 이노신 잔기가 결합 접합부 (밀줄)의 3' 위치의 최초 5개 내에 위치할 때 완성도는 가변적이고 서열-의존적이다. 그러나, 올리고뉴클레오타이드 프로브 디자인 NI8-9에서는 옥타머의 효과적인 결합이 일관성 있게 관찰되었으며, 이때 완성도는 시험된 모든 주형에 대해 99%를 초과하는 것으로 증명되었다,
- <400> 어떤 이론에 의해 결부되는 것을 원하지는 않지만, 이 데이터 (아데닐화 중간체의 존재를 포함하여)는 T4 DNA 결합 효소의 코어 DNA 결합 부위 내의 바람직하지 못한 이노신 염기쌍이 효소 결합 및 뒤이은 결합을 감소시킬 만큼 충분히 DNA 단백질 복합체를 탈안정화한다는 결론을 지지한다. 그러나, 흥미로운 문제는 이러한 탈안정화 이노신 염기쌍이 결합된 올리고뉴클레오타이드 프로브의 충실도에 영향을 미칠 것인가의 여부였다.
- <401> **실시예 3: 프로브 결합의 충실도**
- <402> Taq DNA 결합 효소와 같은 박테리아 NAD-의존성 결합 효소는 결합 접합부 전체적으로 높은 서열 충실도를 갖는 것으로 보고되었으며, 이때 3' 쪽의 미스매치는 본질적으로 Nick-폐쇄 활성을 갖지 않지만, 5' 쪽의 미스매치는 어떤 정도까지는 관용적이다 (Luo et al., Nucleic Acid Res., 24:3071-3078 및 3079-3085, 1996). 한편, T4 DNA 결합 효소는 긴축도가 다소 낮은 것으로 보고되었는데, 이것은 접합부의 3'과 5' 쪽 모두에 미스매치를 허용한다. 따라서, 본 발명자들의 시스템 환경에서 Taq DNA 결합 효소와 비교하여 T4 DNA 결합 효소를 사용한 프로브 결합의 충실도를 평가하는 것에 흥미가 있었다.
- <403> 본 발명자들은 표준 ABI 서열화 기술을 이용하여 결합된 올리고뉴클레오타이드의 서열 충실도를 평가하기 위한 두 가지 방법을 개발하였다. 첫 번째 방법은 결합 생성물을 클로닝하여 서열화하도록 디자인하였다. 이 방법에서는 결합 연장 생성물을 어댑터 서열에 부착하고 클로닝한 다음 박테리아로 형질전환시켰다. 각 콜로니를 채집하고 서열화하여, 결합 접합부 전체 각 위치에서 미스매치 빈도를 정량적으로 평가하였다. 두 번째 방법은 결합 생성물을 직접 서열화하도록 디자인하였다. 이 접근법에서는 단일-가닥 결합 생성물을 비드-기반 주형으로부터 변형시키고, 상보성 프라이머를 이용하여 직접 서열화하였다. 정확도가 낮은 위치가 결과의 서열 추적에서 다중 중복 피크를 나타내는데, 이것은 정량적 평가를 제공하며, 이 위치에서의 서열 충실도에 대한 지표이다.
- <404> 첫 번째 방법을 이용하여 T4 및 Taq DNA 결합 효소에 의한 프로브 결합의 상대적 충실도를 평가하였다. 단일 비드-기반 주형 집단 (LST1)을 보편적 서열화 프라이머에 혼성한 다음, 이것을 개시 올리고뉴클레오타이드로 사용하였다. 그런 다음, 축퇴성 올리고뉴클레오타이드 프로브 (N7A, 3'ANNNNNN5', 2000pmole)의 존재하에 T4 DNA 결합 효소 (1×10^6 비드 당 15U) 또는 Taq DNA 결합 효소 (1×10^6 비드 당 60U)와 함께 37℃에서 30분 동안 용액-기반 결합 반응을 수행하였다 (도 11, 패널 A). 결합 생성물을 클로닝하고 서열화한 다음, 결합 접합부의 3' 쪽에서 각 DNA 결합 효소의 위치 충실도를 평가하였다 (위치 1-8) (도 11, 패널 B 및 C). 이 결과는 최초 5개 위치에서는 T4 DNA 결합 효소가 Taq DNA 결합 효소와 본질적으로 동일한 수준의 충실도를 가지지만, 위치 6-8에서는 충실도가 더 낮음을 나타냈다. 이들 결과는 3개의 축퇴성, 이노신-함유 프로브 디자인 (3'-NNNNNIII-5', 3'-NNNNNINI-5', 및 3'-NNNINNI-5')에 대한 7개 주형 (LST1-7) 전부의 결합 접합부 전체 DNA 서열을 평가한 후속 클로닝 실험에 의해 더 입증되었다. 이 연구에 의해 T4 DNA 결합 효소가 위치 6-8에서는 결합 접합부 전체에서 낮은 서열 충실도를 가지지만, 시험한 모든 주형에서 최초 5개 위치에서는 높은 충실도를 나타냈음이 확인되었다 (데이터는 도시하지 않음).
- <405> 직접 서열화 방법을 이용하여 축퇴성, 이노신-함유 프로브에 대한 T4 DNA 결합 효소의 충실도를 평가하였다. 올리고뉴클레오타이드 프로브를 T4 DNA 결합 효소 및 비드-기반 주형을 함유한 결합 반응에서 25℃ 및 37℃에서 평가하였다. 올리고뉴클레오타이드 프로브 결합 효율을 겔 이동 분석을 이용하여 평가하였다 (도 12, 패널 A).

ABI3730x1 DNA 분석기를 이용, 결합 반응의 직접 서열화를 수행하여 올리고뉴클레오타이드 프로브 결합에서 T4 DNA 결합 효소의 충실도를 평가하였다 (도 12, 패널 B). 정확한 매치 올리고 프로브와 2개의 대표적 축퇴성 이노신-함유 올리고 프로브 (NI8-9 및 NI8-11)의 결합은 99%를 초과하는 완성도와 매우 낮은 미스매치 빈도 (서열화 추적에서 다중 피크가 존재하지 않음)를 제공하였다, 이 데이터는 효과적으로 결합된 프로브는 또한 높은 서열 충실도를 제공한다는 것을 시사한다.

<406> 추가의 실험에서, 단일 비드-기반 주형 집단 (LST1)을 5' 포스페이트를 함유한 보편적 서열화 프라이머에 혼성한 다음, 이것을 개시 올리고뉴클레오타이드로 사용하였다. 37°C에서 30분 동안 T4 DNA 결합 효소 (250,000개 비드 당 1U)를 사용하여 축퇴성, 이노신-함유 올리고뉴클레오타이드 프로브 (3'NNNNNiii5', 3'NNNNNiNi5', 또는 3'NNNiNNNi5', 600pmole)의 존재하에 용액-기반 결합 반응을 수행하였다. 결합 생성물을 클로닝하고, 콜로니를 채집하여 서열화하였다. 결합 접합부의 전체 각 위치에서 나타난 클론의 수를 계산하여 서열 충실도를 결정하였다. 이 결과를 도 12, 패널 C-F에 표로 작성한다. 이들 연구는 T4 DNA 결합 효소를 사용한 축퇴성, 이노신-함유 프로브의 3'→5' 결합이 처음 1-5 위치에서 높은 수준의 충실도를 가짐을 증명한다.

<407> 실시예 4: 겔 내부 결합 및 절단

<408> 올리고뉴클레오타이드 결합 사이클 방법을 조사, 개발 및 최적화하기 위한 초기 실험을 상기 설명된 바와 같이 용액 중에서 비드-기반 주형을 이용하여 수행하였다. 두 번째 세트의 실험에서는 결합 및 절단을 슬라이드 위의 폴리아크릴아미드 겔에 삽입된 비드-기반 주형 상에서 수행하였다.

<409> 단일-가닥 DNA 주형의 클론 집단이 각기 부착된 수 백만개의 비드와 5% 폴리아크릴아미드를 혼합하여 슬라이드를 제조하였으며, 중합은 유리 슬라이드 위에서 일어나도록 하였다. Teflon®마스크를 이용하여 비드-함유 폴리아크릴아미드 용액을 봉하였다. 도 14(상부)는 Cy3-표지된 프라이머가 혼성된 주형이 부착된 비드가 폴리아크릴아미드 겔 내에 고정된 슬라이드 일부분에 대한 형광 영상을 나타낸다 (이 슬라이드는 다른 실험에 사용된 것이며, 여기 사용된 슬라이드의 대표적인 것이다). 도 14(하부)는 폴리아크릴아미드 용액을 봉하는 Teflon 마스크가 장착된 슬라이드의 개략적인 도식을 나타낸다.

<410> 적합한 용액에 수동으로 슬라이드를 침지하거나, 또는 자동 라미나르 유동 셀에 슬라이드를 놓아둠으로써 반응물을 슬라이드에 도입하였다. 초기 연구에 의해 이러한 슬라이드 위의 폴리아크릴아미드 매트릭스 내에 고정된 비드에 부착된 주형에서 효과적인 겔 내부 결합이 실제로 수행될 수 있음이 확인되었다. 도 15에 나타난 실험에서는 단일-가닥 DNA 주형 비드를 아크릴아미드와 DATD를 함유하는 슬라이드 위에 고정하였다. 중합 후에, 보편적인 3' 형광발색단-표지된 5' 포스포릴레이트 프라이머 (Seq 프라이머)를 겔에 확산시켜 혼성화를 수행하였다 (패널 A). 슬라이드를 세척하여 미결합 Seq 프라이머를 제거하고, T4 DNA 결합 효소 (10U)와 올리고뉴클레오타이드 프로브를 함유한 결합 각테일을 위에 바른 다음, 37°C에서 30분 동안 배양하였다. 그런 다음, 슬라이드를 과요오드산 나트륨(0.1M)을 함유하는 완충액 중에서 배양하여 아크릴아미드 중합체를 효소절단하고 비드-기반 주형 집단을 방출시켰다. 결합된 생성물을 열에 의해 주형 가닥으로부터 변성시켜 수집한 다음, 상기 설명된 겔 이동 분석을 이용하여 분석하였다. T4 DNA 결합 효소 없이 수행한 겔 내부 결합 반응에서는 결합하지 않은 서열화 프라이머의 대표적인 단일 피크가 나타났다 (패널 B). T4 DNA 결합 효소의 존재하에 옥타머 프로브를 사용하여 수행된 결합 반응은 효과적인 겔 내부 올리고뉴클레오타이드 결합을 나타냈으며, 비드-기반 주형 집단의 99%를 초과하는 양이 효과적으로 결합되었다 (패널 C).

<411> 실시예 5: 4-색 검출

<412> 검출 효율을 최대화하기 위해서는 각각의 가능한 염기 추가 생성물에 대응하는 상이한 표지를 붙인 한 세트의 올리고뉴클레오타이드 프로브를 사용하는 것이 바람직하다. 이것은 도 15에 개략적으로 도시된 것과 같이, 적합한 여기 및 방출 필터를 구비한 본 발명의 자동 서열화 기기에서 모델링하였다. 프로브 특이성 및 선택성의 문제를 다루기 위해 3세트의 옥타머 프로브를 디자인하였다. 첫 번째 세트는 4개의 독특한 주형 집단에 상보하는 4개의 옥타머를 포함하였으며, 이것들은 상이한 3' 염기 및 5' 염료 표지를 가졌다. 두 번째 세트는 독특한 3' 염기와 5' 염료를 갖는 7개의 독특한 옥타머를 포함하였다. 세 번째 세트는 4개의 축퇴성, 이노신-함유 옥타머를 갖는 프로브 디자인에 상응하였으며, 각각 상이한 5' 염료 표지에 의해 확인되는 독특한 3' 말단 염기를 가진다.

<413> 4-색 스펙트럼 정체성을 확인하기 위하여, 프로브 세트 #1을 사용, 4개의 독특한 주형 집단을 검출하였다 (도 16 참조). 비드에 부착되고 폴리아크릴아미드에 삽입된 4개의 독특한 단일-가닥 주형 집단을 함유하는 슬라이드를 제조하였다 (패널 A). 각 비드는 부착된 주형의 클론 집단을 가졌다. 5' 포스페이트를 함유하는 보편적 서열화 프라이머를 제자리 혼성화한 다음, 4개의 독특한 형광발색단 프로브 (Cy5, CAL610, CAL560, FAM; 각각

100pmole)를 함유한 올리고뉴클레오타이드 프로브 혼합물과 T4 DNA 결합 효소 (10U/슬라이드)을 사용하여 결합 반응을 수행하였다. 슬라이드를 37℃에서 30분 동안 배양하고 세척하여 미결합 프로브들을 제거하였다. 밝은 광원에서 슬라이드를 영상화하자 가벼운 염기에 대한 백색 영상 (패널 B)이 만들어졌으며, 형광 여기는 4개의 밴드패스 필터 (FITC, Cy3, TxRed, 및 Cy5)를 이용한다. 형광 영상 캡처를 결합 후 수행하였다. 각 집단을 위-착색하고 (패널 C), 영상 값들에 대한 스펙트럼 정체성을 도표화하여 최소 신호 중복을 확인한다 (패널 D).

<414> 실시예 6: 겔에서 결합 특이성 및 선택성의 증명

<415> 3' 말단 특이성을 확인하기 위하여, 프로브 세트 #2를 사용하여, 단일 주형 집단을 계통적으로 조사하였다 (도 17 참조). 단일 주형 집단 (LST1.T)이 부착되어 폴리아크릴아미드 겔에 삽입된 비드를 사용하여 슬라이드를 제조한 다음, 보편적 서열화 프라이머를 사용하여 제자리(in situ) 혼성화하였다 (패널 A). 한 개의 3' 염기만 차이가 있는 4개의 5' 말단-표지된 프로브들로 이루어진 올리고뉴클레오타이드 프로브 혼합물을 이용하여 T4 DNA 결합 효소 (10U/슬라이드)로 겔 내부 결합 반응을 수행하였다. 슬라이드를 37℃에서 30분 동안 배양하고 세척하여 미결합 프로브 집단을 제거하였다. 백색 광원에서 슬라이드를 영상화하자 염기 영상이 만들어졌으며 (패널 B), 형광 여기는 4개의 밴드패스 필터 (FITC, Cy3, TxRed, 및 Cy5)를 이용한다. 결합 후 수행된 형광 영상 캡처에 의해 T4 DNA 결합 효소를 사용한 겔 내부 결합 후에 단일 FAM-기반 프로브 집단 (청색 스팟)이 존재하며, 스펙트럼 중복은 없었음이 확인되었다 (패널 C, D). 이 데이터는 T4 DNA 결합 효소의 특이성이 엄격하며 결합 접합부의 최초 3' 말단 염기에 의해 결정된다는 것을 증명한다.

<416> 3' 말단 특이성 및 선택성을 더욱 입증하기 위하여, 프로브 세트 #2를 사용하여 하나의 염기 차이만을 함유하며 상이한 양으로 존재하는 비드-기반 주형 집단의 혼합물을 확인하였다. 비드 혼합물을 이용하여 슬라이드를 제조했는데, 각 비드에는 단일 뉴클레오타이드 다형성 (LST1; A, G, C 또는 T)을 각기 갖는 4개의 주형 집단 중 하나가 부착되며, 이것을 도 18의 패널 A에 나타낸다. 비드를 슬라이드 위의 폴리아크릴아미드 겔 안에 삽입시켰다. 비드-기반 주형 집단을 패널 D에 개략적으로 나타낸 대로 다양한 상이한 빈도로 사용하였다. 슬라이드를 보편적 서열화 프라이머와 제자리 혼성화하였다. T4 DNA 결합 효소 (10U/슬라이드) 및 한 개의 3' 염기만 차이가 있는 4개의 5' 말단-표지된 프로브를 등몰량(equimolar amount) 함유하는 올리고뉴클레오타이드 프로브 혼합물을 사용하여 겔 내부 결합 반응을 수행하였다. 슬라이드를 37℃에서 30분 동안 배양하고 세척하여 미결합 프로브 집단을 제거하였다. 백색 광원에서 슬라이드를 영상화하자 염기 영상이 만들어졌으며 (패널 B), 형광은 4개의 상이한 밴드패스 필터 (FITC, Cy3, TxRed, 및 Cy5)를 이용한다. 각 프로브 영상 위를 덮고 위-착색하였다 (패널 C). 비드-콜링 소프트웨어를 이용하여 형광 영상을 계수하였다. 이 결과를 패널 D에 제시하며, 관찰된 결합 빈도 (Obs)가 예상 빈도 (Exp)와 상관된다는 것이 확인된다. 이 데이터는 다수 주형의 존재하에 결합한 후의 높은 프로브 특이성과 프로브 선택성을 증명하고, 단일 뉴클레오타이드 다형성 (SNP), 즉 결합에 의한, 집단 내 상이한 개체들에서 게놈 DNA의 스트레치에서 단일 뉴클레오타이드 염기에서 발생하는 변화를 검출하는 능력을 증명한다.

<417> 실시예 7: 4-색 축퇴성 이노신-함유 연장 프로브를 이용한 겔에서의 결합 특이성 및 선택성의 증명

<418> 4-색 축퇴성, 이노신-함유 올리고뉴클레오타이드 프로브 풀을 사용하여 프로브 결합의 특이성 및 선택성을 평가하기 위해 프로브 세트 #3을 이용한 다른 세트의 실험을 수행하였다. 결과를 도 19에 제시한다. 비드에 4개의 독특한 단일-가닥 주형 집단이 상이한 양으로 존재하도록 하여 비드-기반 슬라이드를 상기 설명된 대로 제조한 다음, 보편적 서열화 프라이머와 제자리 혼성화하였다 (패널 A). 5개의 축퇴성 염기 (N; 복잡도 $4^5 = 1024$), 2개의 보편적 염기 (I, 이노신), 그리고 특이적 5' 형광발색단에 상응하는 3' 말단에 있는 1개의 공지 뉴클레오타이드 (G-Cy5, A-CAL 610, T-CAL560, A-FAM; 각각 600pmole)를 갖도록 디자인된 옥타머들로 구성된 프로브 풀을 사용하여 T4 DNA 결합 효소 (10U/슬라이드)의 존재하에 겔 내부 결합 반응을 수행하였다. 슬라이드를 37℃에서 30분 동안 배양하고 세척하여 미결합 프로브 집단을 제거하였다. 슬라이드를 백색 광원에서 영상화하자 염기 영상이 만들어졌으며 (패널 B), 형광은 4개의 밴드패스 필터 (FITC, Cy3, TxRed, 및 Cy5)를 이용한다. 각 프로브 영상을 덮고 위-착색하였다 (패널 C). 비드-콜링 소프트웨어를 이용하여 형광 영상을 계수하고 각 결합 생성물의 빈도를 표로 만들었다 (패널 D); 처리하지 않은 원래의 데이터와 비드 신호 값들의 상위 90%를 표시하는 필터링된 데이터의 스펙트럼 산란 플롯이 패널 E에 도시된다. 이 데이터는 각 주형의 공지 농도에 기초하여 관찰된 결합 빈도(Obs)가 예상 빈도(Exp)와 상관되었음을 증명한다. 이것에 의해, 축퇴성 및 보편적 염기-함유 프로브 풀을 T4 DNA 결합 효소와 함께 사용하여 겔 내부 결합의 특이성 및 선택성을 획득할 수 있음이 확인된다.

<419> 실시예 8: 겔에서 개시 올리고뉴클레오타이드 혼성화 및 제거의 반복 사이클의 증명

<420> 자동 유동 셀 (아래 참조)에 장착된 현미경 슬라이드 위의 겔에 고정된 주형에서 수행된 실험으로, 개시 올리고뉴클레오타이드 어닐링 및 스트리핑의 복수 사이클이 신호 손실을 최소화하면서 슬라이드 위의 겔에 삽입된 비드

에 부착된 주형에 적용될 수 있음을 확인하였다. 44 염기의 형광 표지된 개시 올리고뉴클레오타이드를 사용하였다. 도 20에 도시된 대로, 10 사이클에 걸쳐 최소 신호 손실이 발생하였다. 개시 올리고뉴클레오타이드는 도 20에서 프라이머로서 언급된다. 상기 나타난 대로, 중합효소 기초 서열화-합성 과정의 주된 단점 중 하나는 개별적인 주형 가닥에서 포지티브 및 네거티브 단계 제거가 모두 발생하려는 경향이다. 포지티브 단계 제거는 뉴클레오타이드가 성장중인 가닥에 잘못 결합되었을 때 발생하며, 이로 인해 그 특정 가닥의 염기 서열은 남은 주형으로부터 얻어진 서열을 넘어 $n+1$ 염기 콜만큼 상을 벗어나게 된다. 더욱 흔한 네거티브 단계제거는 가닥들이 완전히 연장되지 않았을 때 발생하며, 이 결과 바탕 염기 콜이 성장중인 가닥보다 부족하게 된다 ($n-1$). 연장 생성물을 효과적으로 스트림핑하고, 다른 위치에 있는 개시 올리고뉴클레오타이드와 혼성하여 주형을 "재설정"하는 능력은 거의 또는 전혀 신호 축소 없이 매우 긴 관독 길이를 허용한다.

<421> 실시예 9: 자동 서열화 시스템

<422> 이 실시예는 하나 이상의 주형으로부터 서열 정보를 모으는데 사용될 수 있는 대표적인 본 발명의 자동 서열화 시스템을 설명한다. 바람직하게, 주형은 유리 현미경 슬라이드와 같은 실질적으로 평면인 기관 위에 놓인다. 예를 들어, 주형은 기관 위에 배열된 비드에 부착될 수 있다. 이 시스템의 사진이 도 21에 제시된다. 시스템은 올림푸스 에피-형광 현미경 본체를 토대로 하며, 자동화된 자동-초점 스테이지 및 CCD 카메라가 구비된다 (측면 장착). 회전 홀더에 있는 4개의 필터 큐브가 다양한 여기 및 방출 파장에서 4-색 검출을 허용한다. 펠티에 온도 제어장치를 구비한 슬라이드와 같은 기관을 받아들이도록 개방 및 폐쇄될 수 있는 유동 셀 (겔과 같은 반-고체 지지체를 함유하는 면적의 가장자리를 밀봉하기 위한 개스켓을 구비함)이 스테이지 위에 장착된다. 유동 셀의 수직 배향이 본 발명 시스템의 중요한 측면인데, 이로 인해 유동 셀의 상부로부터 공기 기포가 빠져나갈 수 있다. 각 세척 단계 전에 셀을 완전히 공기로 채워서 모든 시약을 배출시킬 수 있다. 유동 셀은 2개의 9-포트 카브로(Cavro) 시린지 펌프가 구비된 유체 핸들러에 연결되며, 이로써 4개의 상이하게 표지된 프로브 혼합물, 절단 시약, 어떤 다른 바람직한 시약, 효소 평형화 완충액, 세척 완충액 및 공기가 단일 포트를 통해 유동 셀로 전달된다. 시스템의 작동은 완전 자동이며, 다중 I/O 포트가 구비된 전용 컴퓨터를 이용하여 제어 소프트웨어를 통해 프로그래밍된다. 코크 센시캠(Cooke Sensicam) 카메라가 1.3 메가픽셀 쿨드 CCD와 결합되지만, 더 낮은 감도나 높은 감도의 카메라가 사용될 수도 있다 (예를 들어, 4 메가픽셀, 8 메가픽셀 등이 사용될 수 있다). 유동 셀은 1 미크론 피쳐 크기를 갖는 0.25 미크론 스테이지를 이용한다.

<423> 실시예 10: 영상 획득 및 처리 방법

<424> 이 실시예는 표지된 핵산이 부착된 비드 배열로부터 영상을 획득하고 처리하는 대표적인 방법을 설명한다. 정확한 피쳐 확인 및 정렬이 각 획득된 영상의 신뢰성 있는 분석을 위해 중요하다. 피쳐는 먼저 각 비드에 대해 가장 강한 픽셀을 제외한 나머지를 전부 버림으로써 확인된다. 주어진 영상에 대한 픽셀 값을 막대그래프로 그리고 바탕값에 해당하는 픽셀을 버린 다음 나머지 픽셀 값을 정렬한다. 모든 비드의 강도가 대략 동일한 균일한 영상들에서는 이 알고리즘에 의해 픽셀 값의 하위 80-90%가 제거된다. 그런 다음, 상위 10-20%에 있는 값을 갖는 픽셀을 스캔하여 4 픽셀 반경 안에서 국소 최고점에 있는 것들을 확인한다. 그런 다음, 이 영역에서의 평균 강도와 주변부의 평균 강도를 기록한다. 이들 값으로 정규 분포도를 형성한 다음, 그 값이 정규 분포 밖에 있는 픽셀은 제거한다. 초기에 무시된 픽셀 퍼센트, 원형 영역 크기 및 정규 분포도에서 후보 비드들을 제거하는 컷 오프 값은 모두 매개 변수화하여, 필요에 따라 조정될 수 있다. 정렬은 정렬 세트에서 각 영상에 대한 피쳐 매트릭스를 만들으로써 달성된다. 그런 다음, 결과의 매트릭스에서 가장 빈번한 x, y 좌표 오프셋을 조사하여 최적 정렬을 확인한다.

<425> 연장 프로브 추가 전에 Cy5 채널 (서열화 프라이머에 해당함)에서 비드 영상들이 수집된다. 이들 영상을 사용하여 각 비드에 대해 형광 단위(RFU 값)로서 위치 좌표와 원래의 신호 강도를 표시한 피쳐 맵을 작성한다. 후속되는 각각의 이중나선 연장에서, Cy3-표지된 뉴클레오타이드를 첨가하기 전과 후에 영상 세트를 획득한다. 이들 영상과 본래 Cy5 영상을 정렬한 다음, RFU 값을 각 비드에 할당하고 기록한다. 각 염기 첨가마다 표지되지 않은 (연장 전) 영상과 표지된 (형광-첨가) 영상 간 강도의 차이를 뺌으로써 기준선 보정을 적용한다. 그런 다음, 기준선 차감 값을 각 피쳐의 Cy5 영상에서 발견된 강도에 의해 표준화하여 근거를 정하고, 이것에 의해 비드가 연장되었는지 아닌지를 판단한다 (즉, 비드에 부착된 듀플렉스가 연장되었다면 비드가 연장된 것으로 생각한다). 이들 방법을 이용하여, 슬라이드 당 약 1,300개의 영상이 있을 때 영상당 수천 개의 피쳐가 분석될 수 있고, 이로써 실험 1회당 5-100백만 개의 주형 중들이 분석될 수 있다. 이 알고리즘은 더 이상의 효율 증진을 위해 나중에 MATLAB로부터 C++로 쉽게 이식될 수 있도록 디자인되었다.

<426> 실시예 11: 비드 정렬 및 추적 및 서열 코드해석

- <427> 이 실시예는 표지된 핵산이 부착된 비드 배열로부터의 영상을 처리하고 획득된 데이터로부터 서열을 결정하는 대표적인 방법을 설명한다.
- <428> 영상 분석은 비드의 크기와 일치하는 직경의 제로-적분 원형 탐-햇 커널을 이용하여 영상을 감는 것에서 시작한다. 이에 의해서 바탕값이 0으로 자동적으로 표준화되는 동시에 국소 최고점들을 통해 각 비드의 중심이 확인된다. 최고점들이 배치되는데, 다른 국소 최고점들과 분리된 것들이 정렬 포인트로서 사용된다. 이들 정렬 포인트는 시계열 방식으로 각 영상에 대해 계산된다. 각 영상 쌍에 대해서, 정렬 포인트를 비교하고, 공통 정렬 포인트 전체의 평균 변위에 기초하여 변위 벡터를 계산한다. 이것은 픽셀-이하 분해능의 병렬식 영상 변위를 제공한다.
- <429> N개의 영상에 대해서는, $N*(N-1)/2$ 의 병렬 변위가 있지만, 이들 중 단지 N-1만이 독립적인데, 그 나머지가 독립된 세트로부터 계산될 수 있기 때문이다. 예를 들어, 1번 영상과 2번 영상 사이의 변위와 1번 영상과 3번 영상 사이의 변위의 측정은 2번과 3번 영상 사이의 변위를 내포한다. 2번과 3번 영상 사이에서 측정된 변위가 내포된 변위와 동일하지 않다면 이 측정치들은 불일치한다. 이 불일치의 크기를 정렬 알고리즘이 얼마나 잘 작동하는지를 평가하는 미터로서 사용할 수 있다. 본 발명의 초기 시험들에서는 불일치가 각 치수에서 일반적으로 0.1 픽셀 미만인 것으로 나타난다 (도 23 참조).
- <430> 일단 영상이 시계열 방식으로 정렬되면, 각 비드를 추적하는 두 가지 방식이 있다. 비드 대부분이 다른 비드와 접촉하지 않을 만큼 비드 밀도가 낮다면, 각 개별 비드의 광학 질량중심이 확인될 수 있고, 비드 주변 영역을 적분하여 비드 밀도를 계산한다. 비드 대부분이 접촉할 만큼 비드 밀도가 높다면, 비드 주변의 어두운 바탕 밴드로 인해서 개별 비드를 확인하는 것은 불가능하다. 그러나, 픽셀-이하 분해능으로 정렬된 모든 영상에 있어서, 인접 픽셀의 상관성을 계산함으로써 동일한 비드에 속한 픽셀을 확인하는 것이 가능하다. 매우 상관성 있는 픽셀 쌍들이 충실하게 동일한 비드에 할당될 수 있다. 유사한 기술이 DNA 서열화 셀에서 레인을 추적하는데 적용되고 있으며 그 결과는 양호하다 (Blanchard, A. P., Sequence-specific effects on the incorporation of dideoxynucleotides by a modified T7 polymerase, California Institute of Technology, 1993). 일단 비드가 전체 4-색 시계열 방식을 통해 추적되면, 프로브 올리고뉴클레오타이드의 가장 3' 쪽의 염기에 해당하는 색을 구별함으로써 서열의 코드를 해석한다.
- <431> **실시예 11: 처리량 계산**
- <432> 일반적으로 서열화 시스템의 처리량은 1일 당 기계가 생성할 수 있는 영상의 수와 영상당 서열 데이터의 뉴클레오타이드(염기)의 수에 의해 주로 한정된다. 바람직하게 기계는 카메라가 일정한 일을 유지하도록 디자인되므로, 100% 카메라 이용률에 기초하여 계산한다. 각 비드를 4가지 색으로 영상화하여 1개 염기의 정체성을 결정하기 위해서는 1대의 카메라에 의한 4개의 영상, 2대의 카메라에 의해 2개씩의 영상, 또는 4대의 카메라에 의한 1개씩의 영상이 사용될 수 있다. 4대 카메라의 영상 작성이 다른 나머지 선택사항보다 처리량을 극적으로 높일 수 있으며, 바람직한 시스템은 이 접근법을 이용한다.
- <433> 본 발명자들의 초기 시험은 5.4 제곱 마이크로미터로 표시한, 비드 당 50 픽셀의 픽셀 밀도가 표준 영상 분석에 적합한 밀도를 제공함을 밝혔다. 4 메가픽셀 CCD 카메라 (현재는 흔한 기종임)를 이용하면, 1대의 CCD 카메라가 약 80,000개의 비드 (우리의 현재 영상 데이터에 기초하여)를 영상화할 수 있다. 별도의 카메라로 4개의 영상을 캡처하여 유동 셀의 다음번 필드로 이동시키는 데는 1.5초 이하의 시간이 걸릴 것이다. 비드의 75%가 유용한 정보를 제공할 경우, 우리는 처리하지 않은 서열 데이터로부터 약 $80,000 \text{ 개 비드} \times 0.75 / 1.5 = 40,000 \text{ 염기/초}$ 의 데이터를 수집할 수 있을 것이다.
- <434> 100% 카메라 이용률을 유지하는데 있어서 한 중요한 문제는 한 사이클의 결찰/절단 화학을 수행하는데 걸리는 시간과 전체 유동 셀을 영상화하는데 필요한 시간을 일치시키는 것이다. 연장, 절단 및 결찰 사이클에 걸리는 시간의 합리적인 어려움은 1 1/2 시간(5,400초)이다. 이 5,400초는 1,800개의 영상 필드, 또는 적합한 유동 셀 크기인 약 15mm x 45mm의 면적을 수용할 것이다. 4대의 카메라를 이용하는 시스템의 처리량에 대한 추정 어려움은 15mm x 45mm 유동 셀에서 초당 40,000개 염기이다. 이것은 본 발명자들이 ABI3730x1 서열화 기계를 이용하여 달성했던 약 650개 염기 판독 길이 (20개 염기/초)에서 1일 당 28회의 처리량에 기초했을 때, 약 2,000대의 ABI3730x1 서열화 기계와 동일한 것이다. 영상당 200,000개까지 비드 밀도를 2.5배 증가시킴으로써 초당 100,000개 염기까지 전체적인 처리량 증가가 가능하며, 이것은 대략 5,000대의 ABI3730x1 기계와 동일한 것이다. 이 처리량 수준에서 1일 당 총 출력은 1일 당 약 8.6Gb이며, 따라서 12X 인간 게놈 서열을 완성하는데 필요한 시간은 약 4.2일일 것이다.

- <435> 본원에 설명된 본 발명의 서열화 방법은 여러 다양한 상이한 서열화 시스템, 영상 캡처 및 처리 방법 등을 이용하여 실시될 수 있음이 주지된다. 예를 들어, 미국특허 제 6,406,848호 및 제 6,654,505호와 PCT 공보 제 W098053300호를 참조한다.
- <436> **실시예 12: 그 위에서 주형 합성을 하기 위한 마이크로입자의 제조방법**
- <437> 이 실시예는 증폭 프라이머가 부착됨으로써 주형이 증폭 (예를 들어, PCR)될 수 있고, 그 결과 주형 분자의 클론 집단이 각 마이크로입자에 부착될 수 있는 마이크로입자 (이 실시예에서는 자성 비드)의 제조 프로토콜을 설명한다. 일반적으로, 증폭 비드에는 클론 PCR 반응에 필요한 1개의 프라이머가 부착된다. 이 프라이머는 공유 커플링될 수 있거나, 또는 예를 들어 비오틴 표지되어 비드 표면의 스트렙트아비딘에 결합될 수 있다. 비드를 표준 PCR 반응 (예를 들어, 마이크로적정 플레이트의 웰, 튜브 등에서), 실시예 13에서 설명된 에멀션 PCR 반응 등에서 사용하여, 주형 분자의 클론 집단이 부착된 비드를 얻을 수 있다.
- <438> 재료
- <439> 1x TE: 10mM 트리스(pH 8) 1mM EDTA
- <440> 1x PCR 완충액: (ThermoPol Buffer, NEB)
- <441> 20mM 트리스-HCl (pH 8.8)
- <442> 10mM KCl
- <443> 10mM (NH₄)₂SO₄
- <444> 2mM MgSO₄
- <445> 0.1% 트리톤 X-100
- <446> 1M 베타인(1x PCR-B 완충액에만 첨가)
- <447> 1x 결합 & 세척 완충액:
- <448> 5mM 트리스 HCl(pH 7.5)
- <449> 0.5mM EDTA
- <450> 1M NaCl
- <451> DNA 캡처 프라이머 (20-mer, 500 μM 스톡)
- <452> **이중 비오틴-(HEG)5-P1:**
- <453> 5'-이중 비오틴-(HEG)5-CTA AGG TAG CGA CTG TCC TA-3'
- <454> (HEG)5 = 헥사에틸렌 글리콜 링커, 18개 탄소 함유 스페이서, 사용될 수 있는 많은 상이한 스페이서 부분들 중 하나. 예를 들어, 비드 표면으로부터 올리고의 P1 프라이머 부분을 들어올리기 위해 스페이서를 포함하는 것이 유용하다. 본원에 설명된 프라이머 중 어느 것이라도 이러한 스페이서 부분에 결합될 수 있다.
- <455> 디날(Dynal) 스톡 자성 비드(1μm 직경) = 10mg/ml (7-12 x 10⁶ 비드/μL)
- <456> 방법
- <457> 1. 50 μL의 비드 (약 450 x 10⁶ 비드)를 제거한다.
- <458> 2. 200 μL의 1x TE 완충액을 첨가하고 잘 혼합한다. 자석을 이용하여 분리한다.
- <459> 3. 200 μL의 1x TE 완충액으로 1회 세척한다. 자석을 이용하여 분리한다.
- <460> 4. 100 μL의 결합/세척 완충액에 재현탁한다.
- <461> 5. 3 μL의 P1 올리고를 가한다 (500 μM 스톡 = 1500pmol).
- <462> 6. 실온에서 30분 이상 회전시킨다.

7. 200 μ L의 1x TE 완충액으로 3회 세척한다.

8. 50 μ L(초기 부피)의 1x TE 완충액에 재현탁한다.

9. DNA 캡처 비드를 사용하기 전까지 4℃에 또는 얼음이 있는 장소에 보관한다. 비드는 1주일 이내에 사용되어야 한다 (비드는 1주일 이상 보관시에 응집하려는 경향을 나타낼 것이다).

실시예 13: 에멀션 중의 마이크로입자상에서 PCR을 수행하는 방법

이 실시예는 에멀션 중의 마이크로입자상에서 PCR을 수행하여 클론 주형이 부착된 마이크로입자를 생성하는데 사용될 수 있는 방법을 설명한다. 마이크로입자 (아래 사용된 명명법으로는 DNA 비드)를 먼저 제1의 프라이머 (P1)로 기능화한다. 제2의 프라이머 (P2)가 액상에 존재하며, 여기서 PCR 반응이 일어난다. 바람직하다면, 예를 들어 (20배 낮은) 낮은 농도의 P1이 액상에 포함될 수도 있다. 이렇게 하여 액상에서 주형이 신속히 구성되며, 이것이 추가의 증폭을 위한 기질이다. 용액 중 P1이 고갈됨에 따라서 반응은 마이크로입자에 부착된 P1의 이용을 향해 나아간다. P1_P2 degen10이, P1 및 P2와 혼성화하여, PCR에 의한 증폭, 및 올리고뉴클레오티드 집단에 4^{10} 의 복잡도를 제공하는 대략 10개의 축퇴성 염기 (올리고뉴클레오티드 합성 도중 통합됨)의 스트레치를 제공하는 서열을 갖는 올리고뉴클레오티드 주형 (100bp)이다.

I. 에멀션 프로토콜 (1 μ m 비드)

1. 오일상의 제조:

Span 80(7%)

Tween 80(0.4%)

라이트 미네랄 오일 중에서 제조됨

새로 제조된 오일상만을 사용

전체 오일상 = 450 μ l

2. 액상의 제조: (2×10^6 방울이 제조될 것으로 추정됨, 방울당 115fL)

시약 (스톡)	반응물 당 (μ L)	최종
dH ₂ O	156.0	---
MgCl ₂ 완충액(10X)	32.0	1X
dNTP(100mM ea)	11.3	각각 3.5mM
MgCl ₂ (1M)	7.3	23mM
베타인(5M)	32.0	0.5M
P1(프라이머 1)(10 μ M)	1.6	11.25pmole
P2(프라이머 2)(200 μ M)	40.0	5625pmole
P1_P2 degen10(100pM)	6.6	$5.9 \times 10^7 / \mu$ l
DNA 비드(8M/ μ l)	25.0	150M/에멀션
백금 Taq (5U/ μ l)	9.0	0.28U/ μ l

전체 액상 부피 = 320 μ l

최종 반응물 = 255 μ l 액상 : 450 μ L 오일 상

3. 액상 튜브를 얼음으로 옮겨서 에멀션에 첨가할 때까지 둔다.

4. 450 μ l 오일 상을 2mL 세포동결 용기(cryovial)에 넣는다.

5. IKA 선회장치가 부착된 포말 어댑터에 세포동결 용기를 똑바로 놓는다. 선회장치를 2500rpm으로 설정한다.

6. 액상의 분취액 (3개 분취액, 각각 85 μ l = 255 μ l)를 진동중인 오일 상에 첨가한다. 용기의 끝을 튜브에 넣어 교반하고 있는 2mL 세포동결 용기에 단일분산된 액상을 넣고, 그 끝에서부터 진동중인 오일 상으로 액상을 서서

히 분산시킨다. 나머지 액상을 사용하여 2회 반복 첨가한다.

7. 에멀션을 2500rpm에서 24분 동안 계속 진동시킨다.

8. 에멀션의 약 100 μ l 분취액을 96-웰 플레이트 (총 = 4개 웰)로 옮긴다. 또한, 액상을 유지하고 있는 분취액 (65 μ l)을 용액-기반 PCR 대조표준 반응을 위해 분리된 웰에 옮긴다. 플레이트를 밀봉하고 다음 절에 개략적으로 설명된 사이클을 행한다.

II. 에멀션 증폭 (1 μ m 비드)

1. 1 μ m 비드 에멀션에 대한 PCR 사이클링 매개변수 (프라이머 Tms = 62°C):

프로그램: DTB-PCR

94°C, 2분 n=1

94°C, 15초

57°C, 30초 n=100

70°C, 60초

55°C, 5분 n=1

10°C, 임의의 시간 동안

2. 사이클링 시간은 약 6시간이다.

3. 사이클링 후 에멀션을 관찰한다. 연속적인 에멀션은 균일한 호박색을 나타낼 것이며, 액상의 분리는 관찰되지 않을 것이다. "파괴"된 에멀션 (용액을 벗어남)은 튜브의 바닥에 뚜렷한 액상이 있는 것이다. 이 비드 집단은 클론되지 않을 것이므로 액상을 수집하여 없앤다.

4. 명시야(bright field) 현미경을 이용하여 사이클 후 에멀션을 평가한다. 사이클을 행한 에멀션의 2 μ L 분취액을 취하여 유리 슬라이드 위에 적하한다. 에멀션 샘플을 22 x 60mm 유리 커버슬립으로 덮는다.

5. 20X 대물렌즈로 에멀션을 관찰한다. 비드는 바람직하게 단일분산되어야 하며, 방울 대부분이 단일 비드를 함유해야 한다.

주: 만일 에멀션 샘플이 많은 수의 다중-비드 방울을 함유한다면, 에멀션 반응물을 1개의 1.5mL 에펜드로프 튜브에 모아서 6000rpm에서 15초간 회전시킨다. 튜브 바닥에 모인 비드 현탁액을 제거한다. 이 집단은 유리 비드와 단일-비드 방울보다 무거운 다중-비드 방울들로 이루어질 것이며, 따라서 간단한 회전 후에 튜브 바닥에 가라앉을 것이다. 이 비드 집단은 클론되지 않으며, 따라서 후속 처리과정 전에 없애야 한다. 단계 4와 단계 5를 반복함으로써 에멀션을 재평가하여 에멀션 샘플에 있는 단일 비드-함유 방울의 완전성을 확인한다.

6. 다음 단원에 개략적으로 나타낸 프로토콜을 이용하여 에멀션을 붕괴(파괴)한다.

III. 에멀션 파괴 및 용융 (1 μ m 비드)

비드 파괴 세척 (BBW) 완충액

2% 트리톤 X-100; 2% Tween 20; 10mM EDTA

용융 용액: 100mM NaOH

1x TE: 10mM 트리스(pH 8) 1mM EDTA

1x 결합 & 세척 (B/W) 완충액

5mM 트리스-HCl(pH 7.5)

0.5mM EDTA

1M NaCl

1. 하나의 1.5mL 에펜드로프 튜브에 각 에멀션 세트 (4 분취액)를 모은다.

2. 800 μ l BBW 완충액을 가한다. 반응 튜브를 10초간 선회시켜 에멀션을 파괴한다.

- <511> 3. 800rpm에서 2분 동안 회전시킨다.
- <512> 4. 상부 80 μ l (주로 오일 상)를 제거한다. DNA 비드는 튜브 바닥에 펠릿 상태로 있을 것이다.
- <513> 5. 800 μ l BBW를 첨가하고 선회시킨 다음 8000rpm에서 2분 동안 회전시킨다. 상부 600 μ l를 제거한다.
- <514> 6. 600 μ l의 1x TE로 2회 추가 세척하고, 자석을 사용하여 각 세척액을 교환한다.
- <515> 8. 비드 펠릿에 50 μ l의 용융 용액을 첨가하고 강하게 피펫팅하여 샘플을 재현탁한다. 튜브를 간헐적으로 툭툭 치면서 실온에서 5분 동안 용융 용액 중에서 비드를 배양한다.
- <516> 9. 자석에 튜브를 놓고 용융 용액을 제거한다. 100 μ l 용융 용액으로 1회 세척하여 두 번째 가닥을 완전히 제거한다.
- <517> 10. 비드 펠릿을 1x TE로 2회 세척하고 20 μ l의 TE 완충액에 재현탁하여 4℃에서 또는 다음 단계가 부화 단계라면 20 μ l의 1x B/W 완충액에 보관한다. 비드가 응집될 것 같은 경우, 1x PCR-B 완충액으로 교환한다.
- <518> 11. 계속해서 부화 프로토콜을 행한다 (선택적).
- <519> **실시예 14: 클론 주형 집단이 부착된 마이크로입자의 부화 방법**
- <520> 이 실시예는 주형 증폭이, 예를 들어 PCR 에멀션 중에서 연속적으로 발생한 마이크로입자의 부화 방법을 설명한다. 이 방법으로 캡처 올리고뉴클레오타이드가 부착된 더 큰 마이크로입자를 사용할 수 있다. 캡처 올리고뉴클레오타이드는 주형에 존재하는 뉴클레오타이드 영역에 상보하는 뉴클레오타이드 영역을 포함한다.
- <521> I. 에멀션 부화 (1 μ m)
- <522> A. 부화 비드의 제조 (포획체)
- <523> 부화 비드:
- <524> 스페로테크(Spherotech) 스트렙트아비딘-코팅된 폴리스티렌 비드 (약 6.5 μ m)
- <525> 비드 스톡 (0.5% w/v): 33,125 비드/ μ l
- <526> 프로토콜당: (33,125 비드/ μ l)(800 μ l) = 26.5 x 10⁶ 비드
- <527> 용법:
- <528> 에멀션당 119백만 개의 비드 - 에멀션 클론의 어림값 (2%): 에멀션당 약 3M 주형-포지티브 비드. 추정된 주형-포지티브 에멀션 비드 당 2-3개의 부화 비드 첨가 = 에멀션 반응 당 10백만 개의 부화 비드.
- <529> 부화 올리고뉴클레오타이드 (포획체):
- <530> P2-부화 (35-량체, T_m = 73℃)
- <531> 5'-이중 비오틴-18-탄소 스페이서-ttaggaccgttatagttaggatgatgcattaccctg-3'
- <532> (또는)
- <533> P2-부화 (예를 들어, 35-량체 이하, T_m = 52℃)
- <534> 5'-이중 비오틴-18-탄소 스페이서-ggtgatgcattaccctg-3'
- <535> 글리세롤 용액 - 60%(v/v):
- <536> 6ml 글리세롤
- <537> 4ml 뉴클레아제-유리 H₂O
- <538> 1. 800 μ l의 비드를 제거하고 13,000rpm에서 1분 동안 원심분리하여 B/W 완충액으로 교환한다. 500 μ l의 B/W 완충액으로 1회 세척하고 100 μ l의 B/W 완충액에 재현탁한다.
- <539> 2. 20 μ l의 부화 올리고를 첨가한다 (500 μ l 스톡 = rxn 당 10,000pmole)
- <540> 3. 실온에서 1시간 동안 비드 반응물을 회전시킨다.

- <541> 4. 비드를 500 μ l의 1x TE 완충액로 3회 세척한다. 세척 사이사이에 13000rpm에서 1분 동안 원심분리하여 비드를 펠릿으로 만든다.
- <542> 5. 비드를 25 μ l의 B/W 완충액에 재현탁한다. 농도 = 1M 부화 비드/ μ l
- <543> 주: 20-30 μ l의 1x B/W 완충액에 네 개의 부화된 에멀션 집단을 모으면 약 40M의 주형-포지티브 비드가 수득된다. 그런 다음, 다중 슬라이드가 행해질 수 있다.
- <544> B. 부화 과정
- <545> 1. 20 μ l의 부화 비드를 에멀션-유도된 비드 (20 μ l)를 함유하는 튜브에 넣는다. 비드 혼합물을 부드럽게 피펫팅하여 재현탁한다 (또는 모든 추정된 주형-포지티브 에멀션 비드에 대해 2-3개 부화 비드가 발생하는 비율을 사용한다).
- <546> 2. 비오틴화된 P2-부화 프라이머로 코팅된 부화 비드를 사용할 경우, 비드 혼합물을 65℃에서 2분 동안 배양한다. 튜브를 꺼내어 10분 동안 얼음에 둔다.
- <547> 주: 초기 실험은 100-사이클 PCR (예를 들어, P2 PCR)에 사용된 프라이머 서열을 함유하는 부화 비드를 사용하는 것이, 주형이 없는 방울에서 비드에 대해 유도된 프라이머:이량체 종들을 함유하는 비드를 부화시키는 능력 때문에 부화에 있어 덜 효과적일 수 있음을 시사하였다. 만일 상기 설명된 P2-부화 프라이머가 로딩된 부화 비드를 사용할 경우, 이 짧은 프라이머의 감소된 Tm으로 인해 비드 혼합물을 50℃에서 2분 동안 배양한다.
- <548> 3. 비드 혼합물을 300 μ l의 60% 글리세롤 용액을 함유하는 1.5mL 에펜드로프 튜브에 넣는다.
- <549> 4. 13,000rpm에서 1분 동안 원심분리한다.
- <550> 5. 회전 후, 네거티브 비드는 튜브 바닥에 펠릿을 형성할 것이다. 주형 비드가 부착된 부화 비드는 글리세롤 상의 상부에 부유할 것이다. 상부 상의 비드 집단을 수집하여 깨끗한 1.5mL 에펜드로프 튜브로 옮긴다.
- <551> 주: 튜브 바닥에 펠릿을 형성한 비드 (주형을 갖지 않는 비드)는 주형-포지티브 비드에 대해 개략적으로 나타낸 것과 동일한 세척 과정에 따라 세척되고 자석을 이용하여 분석될 수 있다.
- <552> 6. 상부 상에서 취한 비드에 1ml의 뉴클레아제-유리 H₂O를 첨가하여 글리세롤 농도를 희석시킨다. 부드럽게 피펫팅하여 비드 혼합물을 재현탁한다. 13,000rpm에서 1분 동안 회전시킨다.
- <553> 7. 회전 후, 상층액을 제거하고 100 μ l의 TE로 2회 세척한다.
- <554> 8. 세척된 비드 펠릿에 용융 용액을 100 μ l를 가한다. 실온에서 5분 동안 튜브를 회전시킨다.
- <555> 9. 용융 용액을 100 μ l를 첨가하고, 자석을 이용하여 주형 비드를 분리한다.
- <556> 10. 100 μ l의 TE로 2회 세척하여 비-자성 부화 비드를 제거하고, 자석을 이용하여 부화 비드로부터 DNA 비드를 분리하여 취한다.
- <557> 11. 주형 비드를 10-20 μ l의 1x TE에 재현탁한다. 비드가 응집할 것 같은 경우, 1x PCR-B 완충액으로 희석한다.
- <558> 12. 주형-함유 비드를 다른 부화된 집단과 함께 모아서 다음 실시예에 설명된 대로 슬라이드 위에 로딩할 수 있다.
- <559> **실시예 15: 반-고체 지지체에 또는 그 위에 고정된 마이크로입자 배열의 제조방법**
- <560> 이 실시예는 주형이 부착된 마이크로입자가 슬라이드 위에 위치된 반-고체 지지체에 고정된 (예를 들어, 삽입된) 슬라이드의 제조를 설명한다. 이러한 슬라이드를 폴로니(Polony) 슬라이드라고 할 수 있다. 이 실시예에서 사용된 반-고체 지지체는 폴리아크릴아미드이다. 프로토콜들 중 하나는 주형 근처에서 중합효소 분자를 붙잡아 증폭을 증진시키는 방법을 사용한다.
- <561> **슬라이드 제조**
- <562> A. 유리 슬라이드: 결합-실란 처리
- <563> 결합-실란은 유리 슬라이드 표면에 아크릴아미드 겔의 부착을 쉽게 한다. 사용하기 전 슬라이드를 결합-실란으로 전처리하여야 한다.

<564> 주의:

<565> ** 결합-실란 용액은 화학 후드에 보관한다.

<566> ** 결합-실란은 자극성이다. 용액을 제조할 때 실험실에서 작업한다.

<567> ** 원액의 결합-실란 용액이 반출되지 않도록 한다.

<568> ** 랙에서 이동시킬 때 슬라이드의 표면을 만지지 않도록 한다.

<569> **결합-실란 용액의 제조:**

<570> 1. 1L의 플라스틱 용기에 1L의 dH₂O와 교반 막대 1개를 넣는다.

<571> 220 μ l의 농축 아세트산을 첨가한다 (pH가 3.5로 된다). 4ml의 결합-실란 시약을 첨가한다. 교반 플레이트를 이용하여 용액을 15분 이상 혼합한다.

<572> **슬라이드 처리:**

<573> 2. 거꾸로 된 플라스틱 384-웰 플레이트에 슬라이드를 놓는다 (동일한 방향으로 마주보도록).

<574> 3. dH₂O로 세정하여 슬라이드를 세척하고 잘 버린다.

<575> 4. 100% 에탄올로 세정하고 잘 버린다.

<576> 5. dH₂O로 다시 세정하고 잘 버린 다음, 환기구 및 UV 광원이 작동중인 조직 배양 후드에 넣는다. 세척된 슬라이드를 건조시킨다 (약 30분).

<577> 6. 플레이트를 플라스틱 용기에 넣고 슬라이드를 결합-실란 용액으로 덮는다.

<578> 7. 용액과 슬라이드를 1시간 동안 반응시킨다. 용기를 간헐적으로 교반하여 유리에 결합-실란이 고르게 코팅되도록 한다.

<579> 8. 배양 후 슬라이드를 dH₂O로 3회 세정한다.

<580> 9. 100% 에탄올로 1회 세정하고 잘 버린다.

<581> 10. 사용하기 전에 슬라이드를 충분히 건조시킨다.

<582> 11. 결합-실란-처리된 슬라이드는 건조기에 보관한다.

<583> **B. 아크릴아미드-기반 슬라이드 (작은 마스크)**

<584> * 비-포획 프로토콜

<585> 1. 모든 시약은 얼음 위에 둔다. 다음의 차게한 시약들을 1.5mL 에펜드로프 튜브에 넣는다:

<586>

시약	amt(μ l) 2 슬라이드	1 슬라이드
1x TE	13	6.5
비드(1-3M, 1x TE로 희석)	10	5
Rhinohide	1	0.5
40% 아크릴아미드:비스(19:1, F/S)	5	2.5
TEMED(5%, 1x TE 중에)	2	1
APS(0.5%, 신선하게 제조)	3	1.5
총	34 μ l	17 μ l

<587> 혼합물을 강하게 피펫팅하여 비드를 교란시킨다.

<588> 슬라이드 당 17 μ l를 로딩하고 유리 커버 슬립으로 덮는다.

<589> 실온에서 60분 동안 거꾸로 하여 중합한다.

- <590> 깨끗한 면도날로 커버 슬립을 제거한다.
- <591> 슬라이드를 침지하고 1E 완충액으로 15분씩 2회 세척한다 (미결합 비드를 제거할 목적).
- <592> 비드가 삽입된 슬라이드를 4℃에서 Wash 1E 중에 보관한다.
- <593> 2. 형광발색단-표지된 서열화 프라이머와 삽입된 비드 집단을 혼성화한다. 1x PCR-B 완충액을 함유하는 코플린 (Coplin) 병에 슬라이드를 간단히 침지시켜 Wash 1E에서 1x PCR-B 완충액으로 슬라이드를 평형화한다.
- <594> 3. 1.5mL 에펜드로프 튜브에서 99 μ L의 1x PCR 완충액에 1-6 μ L의 (100 μ M 원액) 프라이머를 가한다. 아크릴아미드 매트릭스 위에 100 μ L의 프라이머 용액을 한 방울씩 떨어뜨리고 유리 커버 슬립이나 밀봉 개스킷을 덮는다.
- <595> 4. <DEVIN> 프로그램을 사용하여, 슬라이드를 가열하여 프라이머와 삽입된 비드를 혼성화한다 (65℃에서 2분, 30℃에서 서서히 어닐링). Wash 1E 중에서 2분씩 2회 슬라이드를 세척한다. 슬라이드를 결찰 기초 서열화에 바로 사용한다.
- <596> * 포획 프로토콜
- <597> 1. ssDNA 주형 비드를 1M/ μ L로 제조한다. [슬라이드 당 4-5M 비드를 갖는 폴로니 슬라이드 제조].
- <598> 2. 비드 혼합물을 30 μ L의 1x PCR 완충액에 재현탁한다.
- <599> 3. 1 μ L의 서열화 프라이머 (100 μ M 원액)을 가하고 잘 혼합한다.
- <600> 4. 65℃에서 2분 동안 가열한다.
- <601> 5. 꺼내어 5분 동안 얼음에 둔다.
- <602> 6. 80 μ L의 1x TE로 3회 세척한다.
- <603> 7. 자석을 사용하여 용액을 모두 제거한다.
- <604> 8. 아래에 개략적으로 나타낸 대로 시약을 첨가한다:

시약	amt (μ L) 2 슬라이드
1x 완충액	1.5
10x 완충액	2.0
고 농도(HC) 효소	16.0
40% 아크릴아미드:비스(19:1, F/S)	14.4
Rhinohide	2.0
TEMED(5%, 1x TE 중에)	2.0
APS(0.5%, 신선하게 제조)	1.5
총	39.4 μ L

- <606> 혼합물을 피펫팅하여 비드를 교란시킨다.
- <607> 슬라이드 당 17 μ L를 로딩하고 유리 커버 슬립을 덮는다.
- <608> 9. 바람직하게는 거꾸로 하여, 예를 들어 MJ Research Tetrad PCR 기계에서 <Pol-1> 사이클링 프로파일을 이용하여 증합한다.
- <609> 10. 깨끗한 면도날로 커버 슬립을 제거한다. 슬라이드를 침지하고 10분씩 1E 완충액 중에서 2회 세척한다 (미결합 비드를 제거할 목적).
- <610> 11. 폴로니 슬라이드를 결찰-기반 서열화에 바로 사용한다.
- <611> 12. 비드가 삽입된 폴로니 슬라이드를 4℃에서 Wash 1E 중에서 개스킷 안에 보관할 수 있다.
- <612> **실시예 16: 고체 지지체에 부착된 마이크로입자 배열의 제조방법**
- <613> 이 실시예는 주형이 부착된 마이크로입자가 고체 지지체에 부착된 슬라이드의 제조를 설명한다.

- <614> 1. 유리 슬라이드를 반응성 NHS를 갖는 폴리머 테터를 이용하여 제조한 다음 -20℃에 보관한다 (슬라이드 H, 제품 No. 1070936; Schott Nexterion; Schott North America, Inc., 뉴욕 엘름스포트).
- <615> 2. 사용하기 전 건조제의 존재하에 슬라이드를 실온에서 평형화한다.
- <616> 3. 슬라이드를 50ml의 1x PBS (300mM 인산나트륨, pH 8.7) 중에서 5분 동안 세척한다. 세척을 2회 반복한다.
- <617> 4. 용액에서 슬라이드를 꺼내서 집착 개스킷을 덮는다 (샘플 로딩을 하기 위해).
- <618> 5. 별도의 튜브에서, 1x PBS, pH 8.7에 100-400백만 개의 단백질-코팅된 또는 DNA-코팅된 비드를 나눠 넣는다. DNA는, 예를 들어 서열화를 위한 DNA 주형일 수 있다. DNA는, 예를 들어 NHS와 반응하는 아민 링커를 포함할 수 있다.
- <619> 6. 비드 샘플을 완충액을 교환하여 1x PBS, pH 8.7로 3회 세척한다.
- <620> 7. 비드를 125mL의 1x PBS, pH 8.7에 재현탁한다.
- <621> 8. 비드 용액을 슬라이드 개스킷 안으로 로딩하여 슬라이드의 표면을 고르게 코팅한다.
- <622> 9. 어두운 방안에 슬라이드를 넣고 반응물을 실온에서 1-2시간 배양한다.
- <623> 10. 배양 후 미결합 비드 용액을 제거하고, 50ml의 1x TE (10mM 트리스, 1mM EDTA, pH 8)로 슬라이드를 옮긴다.
- <624> 11. 슬라이드를 50ml의 1x TE로 일정하게 교반하면서 세척 당 15분씩 5회 세척한다.
- <625> 12. 슬라이드는 4℃에서 1x TE 중에 몇 주 동안 보관할 수 있다.
- <626> 13. 필요에 따라, 백색 광원(WL)을 이용한 명시야 영상 분석에 의해, 또는 형광발색단-기반 염료에 부착된 상보성 DNA 올리고뉴클레오타이드를 이용한 형광에 의해 비드 집단을 평가할 수 있다. DNA 주형은, 예를 들어 결찰-기반 서열화를 이용하여 서열화될 수 있다.
- <627> 도 33A는 비드가 부착된 슬라이드의 개략적인 다이어그램을 나타낸다. 작은 비율의 DNA 주형 분자만이 슬라이드에 부착됨을 주지한다. 1 마이크로 비드 (Dynabeads MyOne 스트랩트아비딘 비드; Dynal Biotech, Inc., 제품 No. 650.01)를 사용하였다. 그러나 광범위한 비드들이 사용될 수 있다.
- <628> 도 33B는 슬라이드에 부착된 비드 집단을 도시한다. 하부 패널은 백색 광원(좌측)과 형광 현미경 아래에서의 슬라이드의 동일한 영역을 나타낸다. 상부 패널은 비드 밀도 범위를 나타낸다.
- <629> **등가물 및 범위**
- <630> 당업자는 본원에 설명된 본 발명의 특정 실시예들에 대한 많은 등가물을 인정하거나, 또는 통상적인 실험을 이용하여 확인할 수 있을 것이다. 본 발명의 범위는 상기 설명에 제한되지 않으며, 첨부된 청구범위에 제시된 것과 같다. 다음의 청구항들에서, "한", "하나" 및 "그"와 같은 말들은 문맥으로부터 명백한 것이 아니라면 하나 또는 하나 이상을 의미할 수 있다. 어떤 군의 하나 이상의 구성원들 사이에 "또는"을 포함하는 청구항 또는 설명은, 문맥으로부터 명백한 것이 아니라면, 주어진 생물 또는 방법에 관하여 군 구성원 중 하나, 하나 이상, 또는 전부가 존재하거나, 사용되거나, 또는 관련될 경우 만족되는 것으로 여겨진다.
- <631> 나아가서, 본 발명은 기재된 청구항들 중 하나 이상으로부터 하나 이상의 제한, 요소, 절, 설명 용어 등이 다른 청구항에 도입된 모든 변형, 조합, 및 치환을 포함한다는 것이 이해되어야 한다. 특히, 다른 청구항에 종속된 어떤 청구항은 동일한 기초 청구항에 종속된 어떤 다른 청구항에서 발견된 하나 이상의 제한을 포함하도록 변형될 수 있다.
- <632> 또한, 어떤 하나 이상의 실시예는 배제된다는 것이 본원에 명백히 제시되지 않더라도 청구항으로부터 명백히 배제될 수 있음이 이해되어야 한다. 또한, 명세서 및/또는 청구범위는 서열화에 사용되는 시약 (예를 들어, 주형, 마이크로스피어, 프로브, 프로브 군 등)을 개시하며, 당업자가 다른 식으로 이해하거나, 또는 명세서에 다른 식으로 암시되지 않는 한, 이러한 개시는 또한 본원에 개시된 특정 방법, 또는 본 분야에 공지된 다른 방법에 따라 이 시약을 사용하는 서열화 방법도 포함한다는 것이 이해되어야 한다. 또한, 명세서 및/또는 청구범위가 서열화 방법을 개시하는 경우, 당업자가 다른 식으로 이해하거나, 또는 명세서에서 그러한 방법에서 이 시약의 사용이 명백히 배제되지 않는 한, 본원에 개시된 시약들 중 어떠한 하나 이상의 시약이 이 방법에 사용될 수 있다. 나아가, 명세서 또는 청구범위에 서열화에 사용되는 특정한 성분들이 개시된 경우, 본 발명은 이 시약들

을 제조하는 방법을 또한 포함함이 이해되어야 한다. 용어 "성분"은 주형, 주형이 부착된 마이크로입자, 라이브러리 등을 포함하는 서열화에 사용된 어떠한 항목을 언급하는데 폭넓게 사용된다. 더욱이, 도면도 본 명세서의 통합된 부분이며, 본 발명은, 주형이 부착된 마이크로입자와 같은 도면에 도시된 구조와 도면에 개시된 방법을 포함한다.

- <633> 본원에서 범위가 주어졌을 때는 그 중점이 포함된다. 나아가, 문맥으로부터 분명하고 당업자에게 이해되는 바가 아니라면, 문맥상 명확히 다르게 제시되지 않는 한, 범위로서 표현된 수치들은 본 발명의 상이한 실시예들에서 표시된 범위 안의 어떤 특정한 값 또는 하위범위를 그 범위의 하한 단위의 1/10까지로 추정할 수 있다.

산업상 이용 가능성

- <634> 본 발명은 단편의 분리를 수행할 필요성을 피하면서 또한 어떤 경우에는 중합효소를 사용할 필요를 피하는 새롭고 개선된 서열화 방법을 제공한다. 본 발명의 방법에서 뉴클레오티드는 각 사이클에서 확인되고, 본 발명의 방법에 의해 폴리뉴클레오티드의 서열화를 효과적으로 실행할 수 있고 많은 양을 한번에 처리할 수 있다.

도면의 간단한 설명

- <33> 도 1A는 개시 후 이어지는 2 사이클의 연장, 결찰 및 확인을 개략적으로 도시한 도면이다.
- <34> 도 1B는 연장이 지지체를 향하여 주형의 유리 말단로부터 내부로 진행되는 실시예에서 개시 후 이어지는 2 사이클의 연장, 결찰 및 확인을 개략적으로 도시한 도면이다.
- <35> 도 2는 프로브의 3' 염기의 정체가 형광발색단(fluorophore)의 색을 확인함으로써 결정되는 올리고뉴클레오티드 프로브에 색을 지정하기 위한 도식을 도시한다.
- <36> 도 3A는 주형의 결합 영역의 상이한 위치에서의 개시 올리고뉴클레오티드의 혼성화와 이어지는 연장 프로브의 결찰에 의해 유발되는 연장된 이중나선을 개략적으로 도시한다.
- <37> 도 3B는 주형 분자의 매 6번째의 염기를 판독하도록 디자인된 연장 프로브를 이용하여 연장, 결찰, 및 절단 방법을 사용함으로써 연속적인 서열을 조립하는 것을 개략적으로 도시한다.
- <38> 도 4A는 5'-S-포스포로티올레이트 연쇄 (3'-O-P-S-5')를 도시한다.
- <39> 도 4B는 3'-S-포스포로티올레이트 연쇄 (3'-S-P-O-5')를 도시한다.
- <40> 도 5A는 3'-O-P-S-5' 포스포로티올레이트 연쇄를 가지는 연장 프로브를 사용하여 5'→3' 방향으로 서열화하기 위한 연장, 결찰, 및 절단의 단일 사이클을 개략적으로 도시한다.
- <41> 도 5B는 3'-S-P-O-5' 포스포로티올레이트 연쇄를 가지는 연장 프로브를 사용하여 3'→5' 방향으로 서열화하기 위한 연장, 결찰, 및 절단의 단일 사이클을 개략적으로 도시한다.
- <42> 도 6A 내지 도 6F는 단일 주형에 대해 수행된 여러 서열화 반응을 보다 상세한 도표로 예시하는 도면이다. 반응은 주형의 상이한 부분들에 결합하는 개시 올리고뉴클레오티드를 활용한다.
- <43> 도 7은 dA와 dG의 3'-포스포로아미다이트(phosphoroamidite)에 대한 합성 도식을 보여주는 개략도이다.
- <44> 도 8A 내지 도 8E는 포스포로티올레이트 연쇄를 함유하는 연장 프로브의 연속적인 결찰 및 절단의 2 사이클을 증명하는 겔 시프트 분석의 결과를 도시한다.
- <45> 도 8F는 DNA 결합 효소에 의한 결찰의 메커니즘을 개략적으로 도시한다.
- <46> 도 9는 축퇴성 이노신(inosine)-함유 올리고뉴클레오티드 프로브의 결찰 효율을 증명하는 겔 시프트 분석 결과를 도시한다.
- <47> 도 10은 다중 주형에 대한 축퇴성 이노신-함유 올리고뉴클레오티드 프로브의 결찰 효율을 증명하는 겔 시프트 분석 결과를 도시한다.
- <48> 도 11은 3'→5' 연장을 위한 두 개의 DNA 결합 효소 (T4 DNA 결합 효소 및 Taq DNA 결합 효소)의 각각의 충실성을 평가하기 위해 수행된 분석의 결과를 도시한다.
- <49> 도 12는 축퇴성 이노신-함유 올리고뉴클레오티드 프로브의 결찰 효율을 증명하는 겔 시프트 분석 결과 (A) 및 올리고뉴클레오티드 프로브 결찰시 T4 DNA 결합 효소의 충실성을 평가하기 위해 수행된 결찰 반응의 직접적인

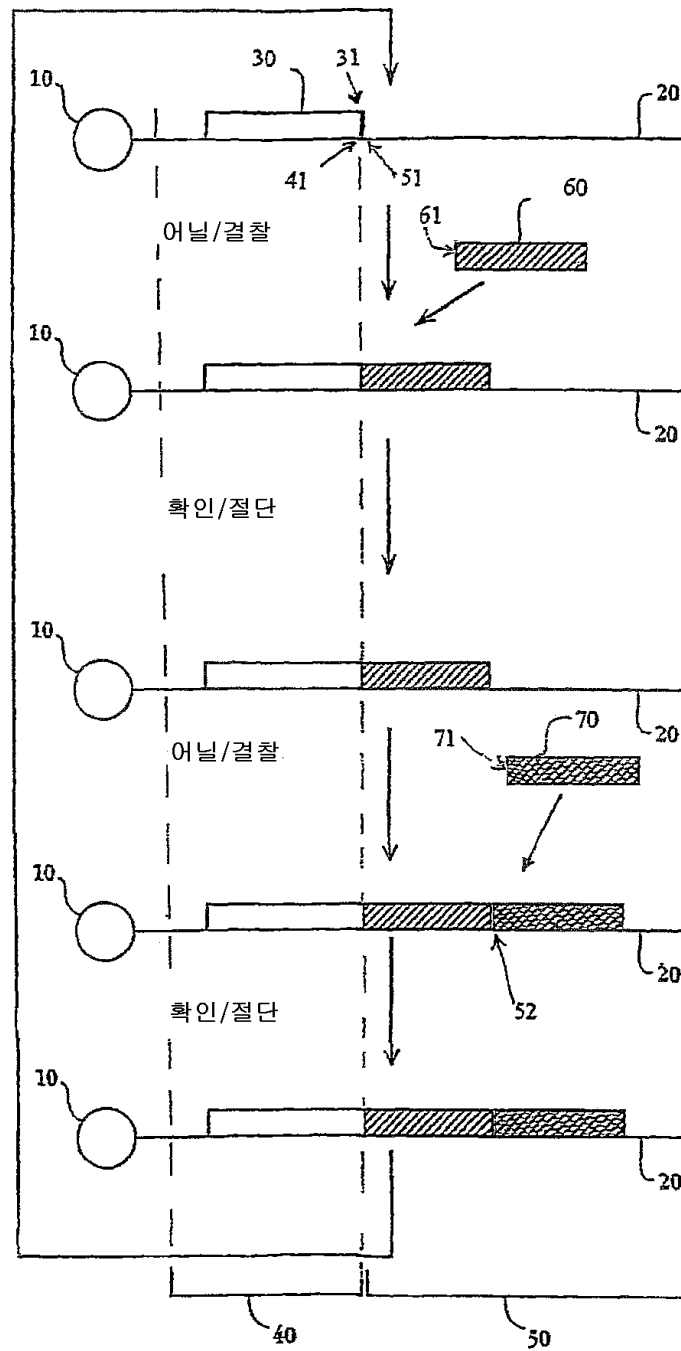
서열화 분석의 결과 (B)를 도시한다. 결과들은 패널 C-F로 표로 만들었다.

- <50> 도 13A 내지 도 13C는 비드-기반 주형이 슬라이드 상의 폴리아크릴아미드(polyacrylamide) 겔에 삽입될 때 겔 내부 (in-gel) 결합반응을 증명하는 실험의 결과를 도시한다. 도 13A는 결합 반응의 개략도이다. 겔 내부 결합 반응은 T4 DNA 결합 효소의 부재하에 (B) 및 존재하에 (C) 수행되었다.
- <51> 도 14A는 형광 표지된 두 번째 증폭 프라이머와 과량의 주형을 사용하는, 첫 번째 증폭 프라이머가 부착된 비드 상에서 수행된 에멀션 PCR 반응의 영상을 도시한다.
- <52> 도 14B (상부)는 Cy3-표지된 올리고뉴클레오타이드가 혼성화되어 있고, 폴리아크릴아미드 겔 내에서 고정되어 있는 주형이 부착된 비드가 그 위에 있는 슬라이드의 일부의 형광 영상을 도시한다. (이 슬라이드는 상이한 실험에서 사용되었지만, 본원에서 사용된 대표적인 슬라이드이다.) 도 14B (하부)는 폴리아크릴아미드 용액을 넣기 위하여 테플론 마스크가 장착된 슬라이드를 개략적으로 도시한다.
- <53> 도 15는 프로브 특이성과 선택성의 문제를 해결하기 위해 디자인된 3 세트의 표지된 올리고뉴클레오타이드 프로브를 도시하며, 또한 4개의 스펙트럼으로 분해가능한 표지의 세트에 대한 여기 및 방출 값을 도시한다.
- <54> 도 16은 올리고뉴클레오타이드 프로브의 4-색 스펙트럼 정체를 확인하는 실험의 결과를 도시한다. 4개의 독특한 단일-가닥의 주형 집단을 함유하는 슬라이드 (A)에 대해 혼성화 반응을 수행하고, 4개의 독특한 형광발색단 프로브를 함유한 올리고뉴클레오타이드 프로브 혼합물을 사용한 결합 반응을 밝은 빛 아래에서 영상화하고 (B) 결합 반응 전 및 후에 4개의 밴드패스 필터 (bandpass filter)를 사용하여 형광을 여기하였다. 개별적인 집단을 위 (pseudo)-착색하였다 (C). 최소한의 신호 중복을 나타낸 스펙트럼 상의 정체를 (D)에 도표화 하였다.
- <55> 도 17은 올리고뉴클레오타이드 프로브의 결합 특이성을 확인하는 실험을 도시한다. 도 17(A)는 결합의 개략적인 윤곽을 도시한다. 도 17(B)는 밝은 빛 영상이고, 도 17(C)는 결합 후의 폴리아크릴아미드 겔에 삽입된 비드 집단의 상응하는 형광 영상이다. 도 17(D)는 결합 전 (pre) 또는 후 (post)의 각 표지로부터 검출된 형광을 도시한다.
- <56> 도 18은 올리고뉴클레오타이드 연장 프로브의 결합 특이성과 선택성을 확인하는 다른 실험을 도시한다. 도 18(A)는 결합의 개략적인 윤곽을 도시한다. 도 18(B)는 밝은 빛 영상이고, 도 18(C)는 결합 후의 폴리아크릴아미드 겔에 삽입된 비드 집단의 상응하는 형광 영상이다. 도 18(D)는 집단의 특정 연장 프로브의 비율을 기초로 하여 예상된 빈도와 관찰된 빈도 사이의 높은 상관관계를 보여주는, 예상된 결합 빈도 대 관찰된 결합 빈도를 도시한다.
- <57> 도 19는 축퇴성 및 보편적 염기를 함유하는 올리고뉴클레오타이드 연장 프로브 풀이 특이적이고 선택적인 겔 내부 결합을 부여하기 위해 사용될 수 있는 것을 확인하는 실험을 도시한다. 도 19(A)는 결합 후의 4개의 상이하게 표지된 축퇴성 이노신-함유 프로브 풀을 예시하는, 결합 실험의 개략적인 윤곽을 도시한다. 도 19(B)는 밝은 빛 영상이고, 도 19(C)는 결합 후의 폴리아크릴아미드 겔에 삽입된 비드 집단의 상응하는 형광 영상이다. 도 19(D)는 집단의 특정 연장 프로브의 비율을 기초로 하여 예상된 빈도와 관찰된 빈도 사이의 높은 상관관계를 보여주는, 예상된 결합 빈도 대 관찰된 결합 빈도를 도시한다. 도 19(E)는 비드 신호 값의 상위 90%를 나타내는 원래의 가공하지 않은 데이터와 여과된 데이터의 산란 도표를 도시한다.
- <58> 도 20은 개시 올리고뉴클레오타이드 (프라이머)의 주형에 대한 혼성화 및 스트리핑(stripping)의 연속적인 사이클에서 검출된 신호를 도시하는 차트이다. 도면에서 알 수 있는 바와 같이, 최소한의 신호 손실은 10회 사이클에 걸쳐 발생하였다.
- <59> 도 21은 예컨대 실질적으로 평면인 지지체에 또는 그 위에 배열된 주형으로부터의 서열 정보를 모으기 위해 사용될 수 있는 자동 서열화 시스템의 사진이다. 또한 도면에는 시스템의 다양한 성분의 작동을 조절하고, 수집된 영상 데이터를 처리하고 저장하며, 사용자 인터페이스를 제공하는 등을 위한 전용 컴퓨터가 도시된다. 도면의 하부 부분은 중량에 의한 기포 치환을 이루기 위해 배열된 유동 셀의 확대도를 도시한다.
- <60> 도 22는 실질적으로 평면인 지지체에 또는 그 위에 배열된 주형을 서열화하기 위해 사용될 수 있는 고처리량 (high throughput) 자동 서열화 기기를 개략적으로 도시한다.
- <61> 도 23은 30 프레임 이상의 최소한의 불일치를 예시하는, 배열 불일치의 산란 도표를 도시한다.
- <62> 도 24A 내지 도 24I는 다양한 상이한 관점에서 바라본 본 발명의 유동 셀 또는 그것의 일부분들을 개략적으로 도시한다.

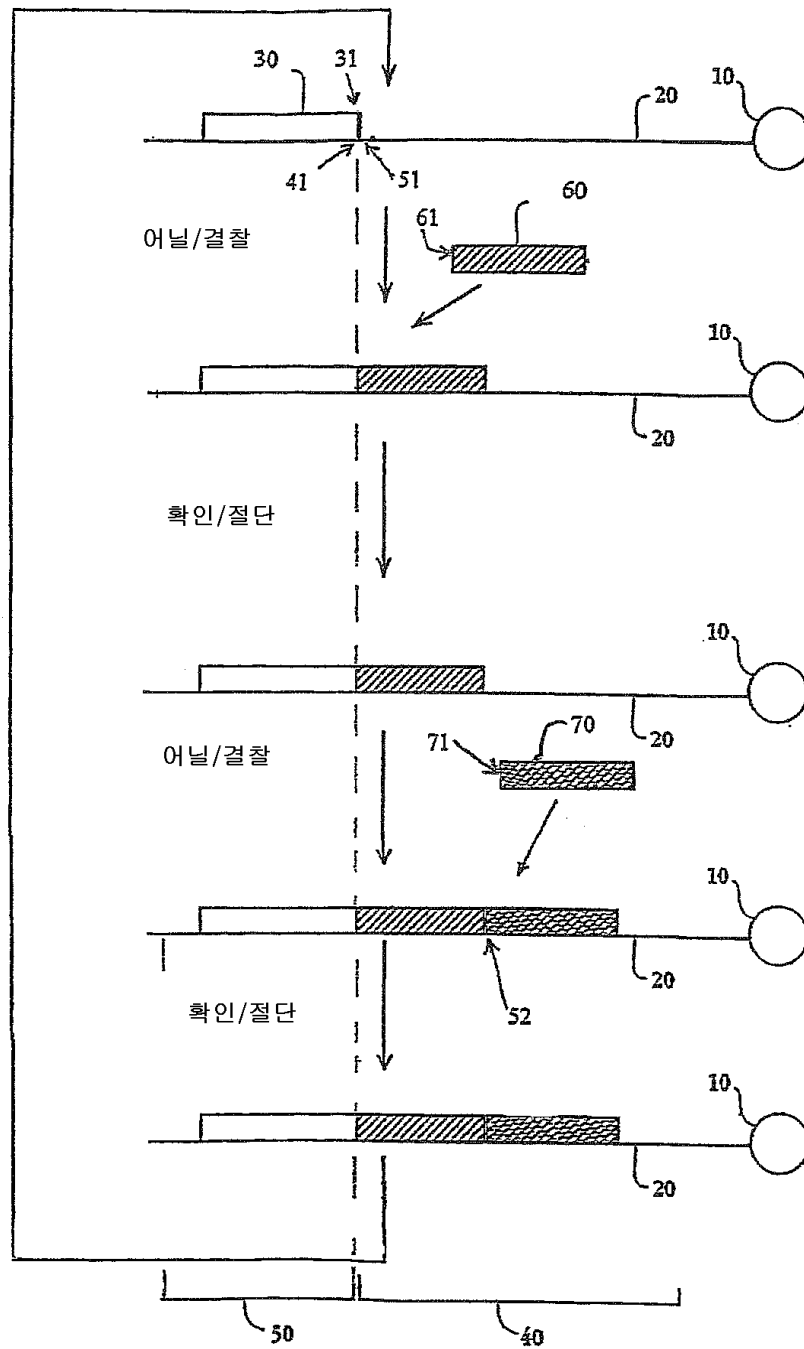
- <63> 도 25A는 길이가 2 뉴클레오티드인 구속된 부분을 포함하는 부분적으로 구속된 프로브를 포함하는 프로브 군의 바람직한 집합에 대한 예시적인 코드화를 도시한다.
- <64> 도 25B는 프로브 군의 바람직한 집합 (상부 패널)과 결합, 검출, 및 절단의 사이클 (하부 패널)을 도시한다.
- <65> 도 26은 길이가 2 뉴클레오티드인 구속된 부분을 포함하는 부분적으로 구속된 프로브를 포함하는 프로브 군의 다른 바람직한 집합에 대한 예시적인 코드화를 도시한다.
- <66> 도 27A 내지 도 27C는 표 1에서 규정된 프로브 군의 24개의 바람직한 집합을 개략적으로 규정하기 위한 대체 방법을 도시한다.
- <67> 도 28은 프로브가 길이가 2 뉴클레오티드인 구속된 부분을 포함하는 프로브 군의 덜 바람직한 집합을 도시한다.
- <68> 도 29A는 길이가 3 뉴클레오티드인 구속된 부분을 가지는 프로브를 포함하는 프로브 군의 집합에 대한 구속된 부분을 생성하기 위해 사용될 수 있는 다이어그램을 도시한다.
- <69> 도 29B는 프로브 군의 24개의 바람직한 집합으로부터 길이가 3 뉴클레오티드인 구속된 부분을 가지는 프로브를 포함하는 프로브 군의 집합에 대한 구속된 부분을 생성하기 위해 사용될 수 있는 지도화 도식의 다이어그램을 도시한다.
- <70> 도 30은 프로브 군의 집합을 사용하여 서열 결정이 수행되는 방법을 도시한다. 바람직한 세트의 프로브 군을 사용하는 실시예가 도시된다.
- <71> 도 31A 내지 도 31C는 서열 결정이 후보 서열을 생성하기 위해 첫 번째 집합의 프로브 군을 사용하고 암호를 해독하기 위해 두 번째 집합의 프로브 군을 사용하여 수행되는 방법을 도시한다.
- <72> 도 32는 서열 결정이 프로브 군의 덜 바람직한 집합을 사용하여 수행되는 방법을 도시한다.
- <73> 도 33A는 비드가 부착된 슬라이드의 개략적인 다이어그램을 도시한다. DNA 주형은 비드에 부착되어 있다.
- <74> 도 33B는 슬라이드에 부착된 비드의 집단을 도시한다. 하부 패널은 백색 광 (좌편) 및 형광 현미경 하의 슬라이드의 동일한 영역을 나타낸다. 상부 패널은 비드 밀도의 범위를 나타낸다.
- <75> 도 34A 내지 도 34C는 개별적인 핵산 집단으로서 핵산 단편 (주형)에 존재하는 한 쌍의 태그의 두 개의 태그를 증폭하고 그것들을 증폭 과정을 통해 마이크로입자에 캡처하기 위한 계획을 도시한다.
- <76> 도 35A와 도 35B는 도 35의 도식에 대한 프라이머 디자인과 증폭의 상세한 부분을 도시한다. 핵산 단편 (주형)의 두 가닥이 모두 명백하게 도시된다. 프라이머와 동일한 서열을 가지는 프라이머 결합 영역이 같은 색으로 표시된다. 예를 들어 P1은 진한 파란색으로 표시되고, 마이크로 입자 위에 및 용액 중에 존재하는 프라이머 P1은 주형의 표시된 가닥의 대응하여 착색된 부분과 동일한 서열을 가지는 것을 나타낸다. P1으로 표시된 주형의 진한 파란색 영역은, 실제로는 상응하는 프라이머 (P1)이 다른 가닥의 상보하는 부분에 결합하고 프라이머 P1과 동일한 서열을 가진다고 하여도, 프라이머 결합 영역으로서 언급될 수 있다.
- <77> 도 35C와 도 35D는 각각 도 35A 및 도 35B의 방법에 의해 생성된 마이크로입자에 부착된 첫 번째 및 두 번째 태그의 서열화를 도시한다.

도면

도면1A



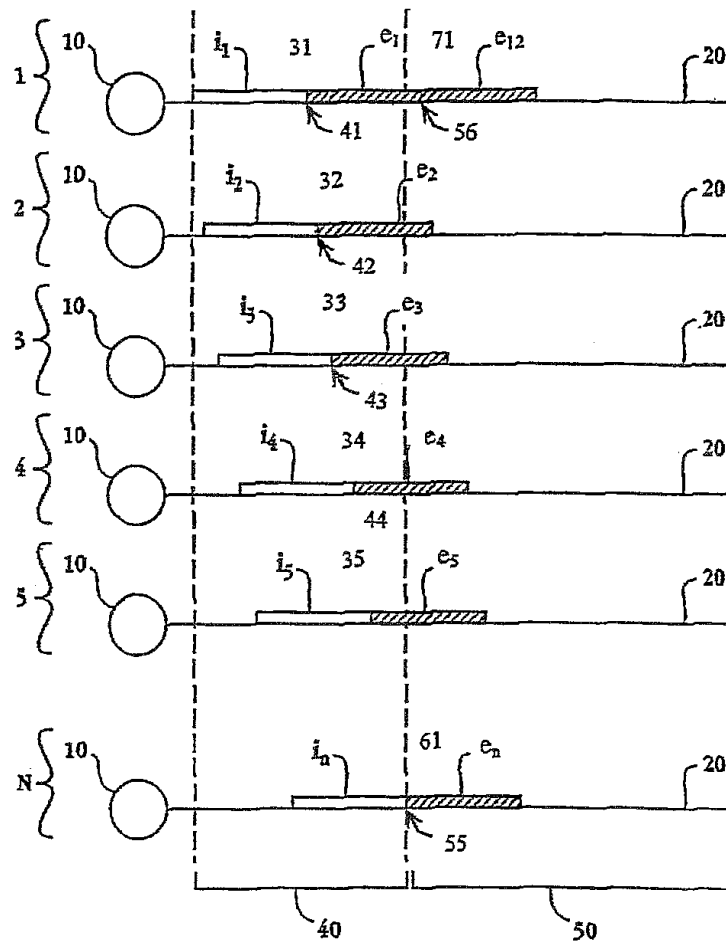
도면1B



도면2

색	5'					3'
레드	N	N	N	N	N	C
옐로우	N	N	N	N	N	A
그린	N	N	N	N	N	G
블루	N	N	N	N	N	T

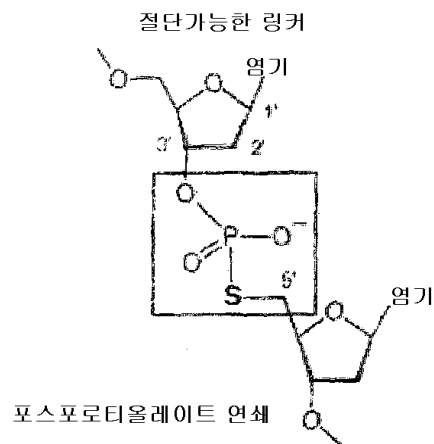
도면3A



도면3B

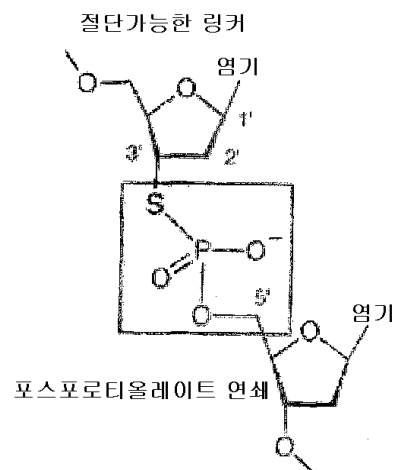
주형
 프라이머 1 ACCAAATGGCAGCCCAATTTTGCATCCCAGGGGCATTACCGAATGGAGCCGTATC
 프라이머 2 (n-1) A A A T A G T A G C T C
 프라이머 3 (n-2) C A G C A A T G C A A T G A T A T C
 프라이머 4 (n-3) C C T G C A T T G C C C G C A T T A C C G A A T G G A G C C G T A T C
 프라이머 5 (n-4) A T G C C A T T G C C C A G G G C A T T A C C G A A T G G A G C C G T A T C
 프라이머 6 (n-5) A T G C C A T T G C C C A G G G C A T T A C C G A A T G G A G C C G T A T C

도면4A



적절한 조건하에서 Ag^+ 에 의해 절단된 S-P결합

도면4B

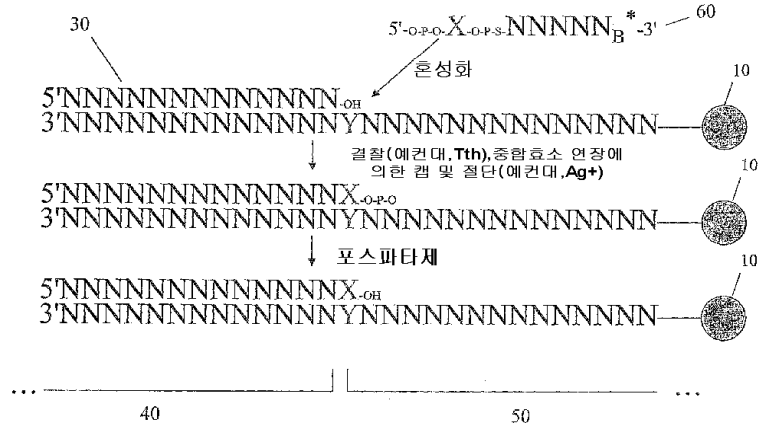


적절한 조건하에서 Ag^+ 에 의해 절단된 S-P결합

도면5A

5'S 도식

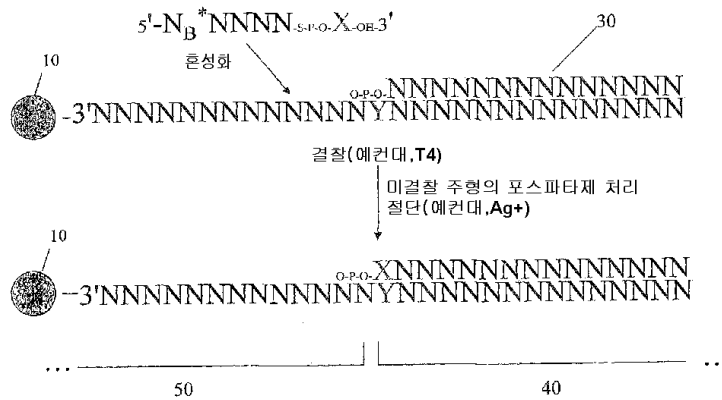
형태: 5'-O-P-O-X-O-P-S-NNNNN_B*-3'의 6량체



도면5B

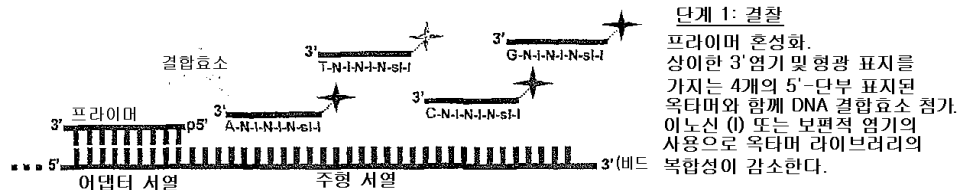
3'S 도식

형태: 5'-N_B*NNNN-S-P-O-X-OH-3'의 6량체



도면6A

주기적인 올리고뉴클레오타이드 결찰 및 절단에 의한 서열화

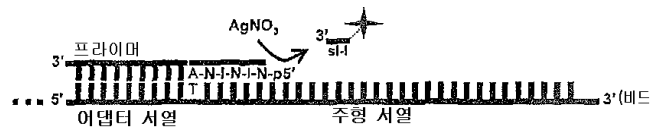


도면 6B



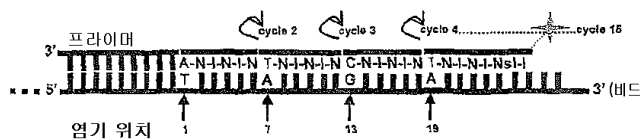
단계 2: 영상
정확한 옥타머가 걸찰된 후
비드를 영상화하고 프라이머
아래의 첫 번째 염기를 기록한다.

도면 6C



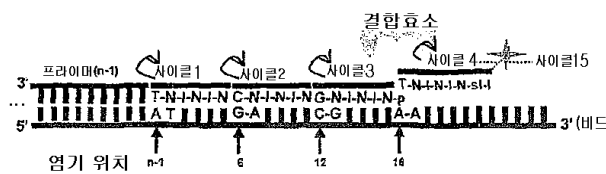
단계 3: 절단
플루오로포어를 방출하고
5' 포스페이트를 노출하기
위해 질산은으로
가교 P-S를 절단한다.

도면 6D



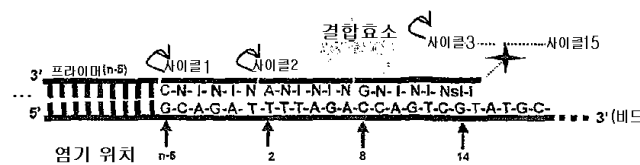
단계 4: 결찰/절단의 반복
동일한 세트의 표지된 옥타머를 사용하여 결찰과 절단의 추가 사이클을 수행함으로써
가닥을 연장한다. 매 여섯 번째의
염기의 서열을 결정한다
(예컨대 15 사이클 후에
1, 7, 13, 19, 25, 31, 37, 43,
49, 55, 61, 67, 73, 79, 85).

도면 6E



단계 5: 프라이머 재설정, 사이클 반복
열-스트립 연장된 가닥. n-1 위치에서 끝나는 새로운 프라이머를 혼성화하고 겹칠/절단 사이클을 반복하여 다음 염기 서열을 서열화한다
(예컨대 15 사이클 후에 염기-1, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84).

도면 6F

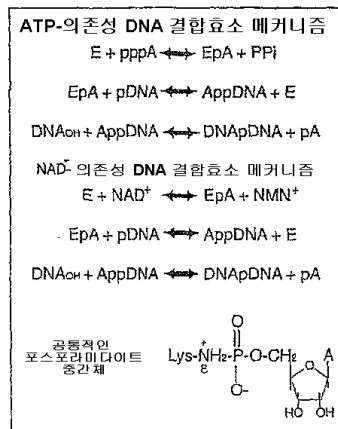


단계 6-9: 재설정/사이클 반복

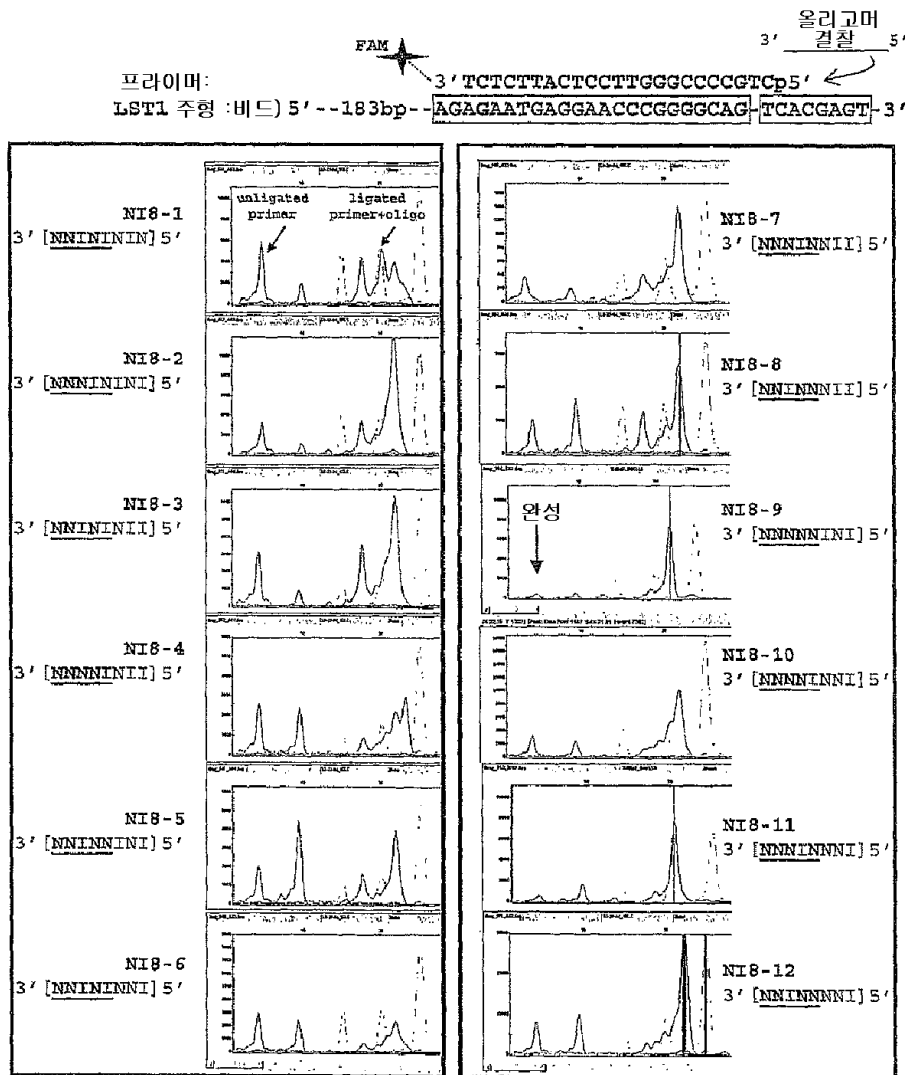
단계 5를 $n-2$, $n-3$, $n-4$ 및 $n-5$ 에서 끝나는 프라이머를 사용하여 반복하여 사이에 낀 염기를 판독한다.

판독 길이 = 85염기
(2번째 판독=170 염기)

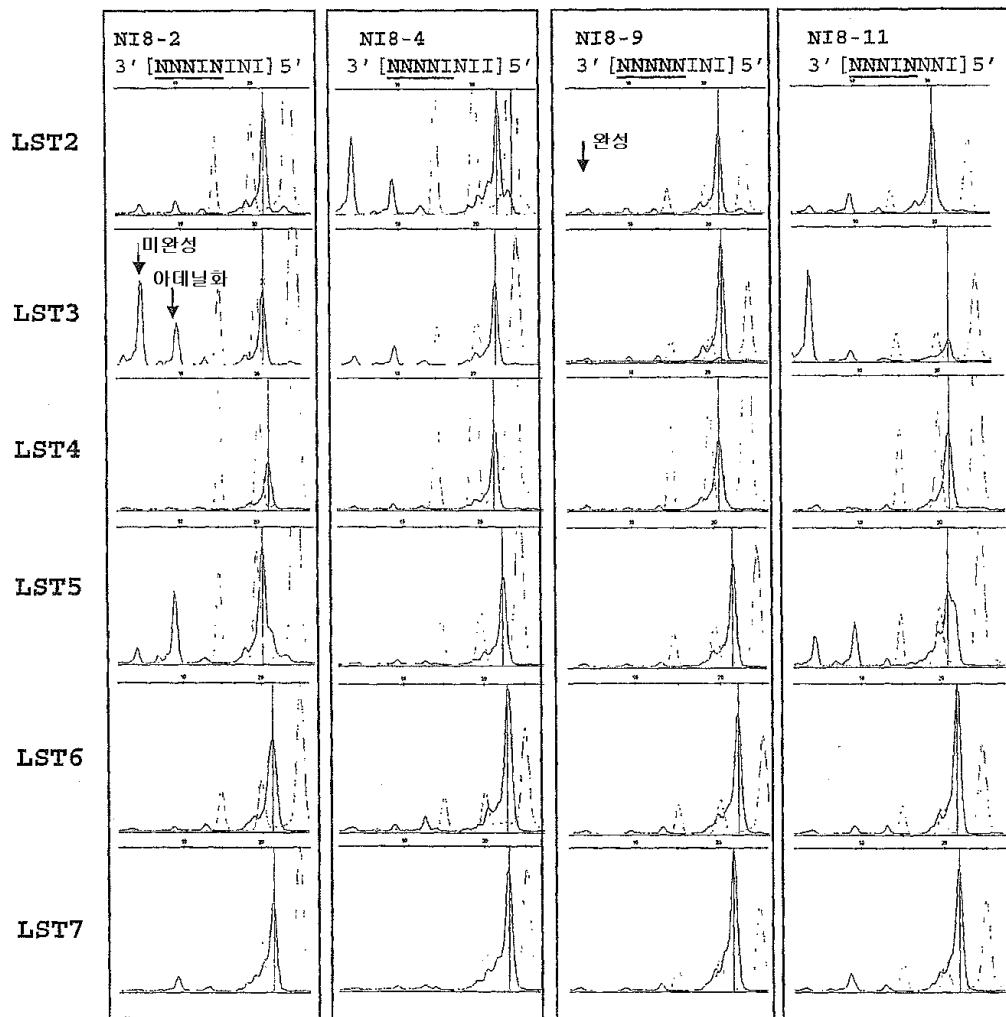
도면8F



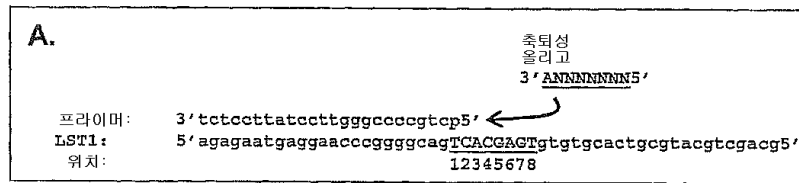
도면9



도면10



도면11



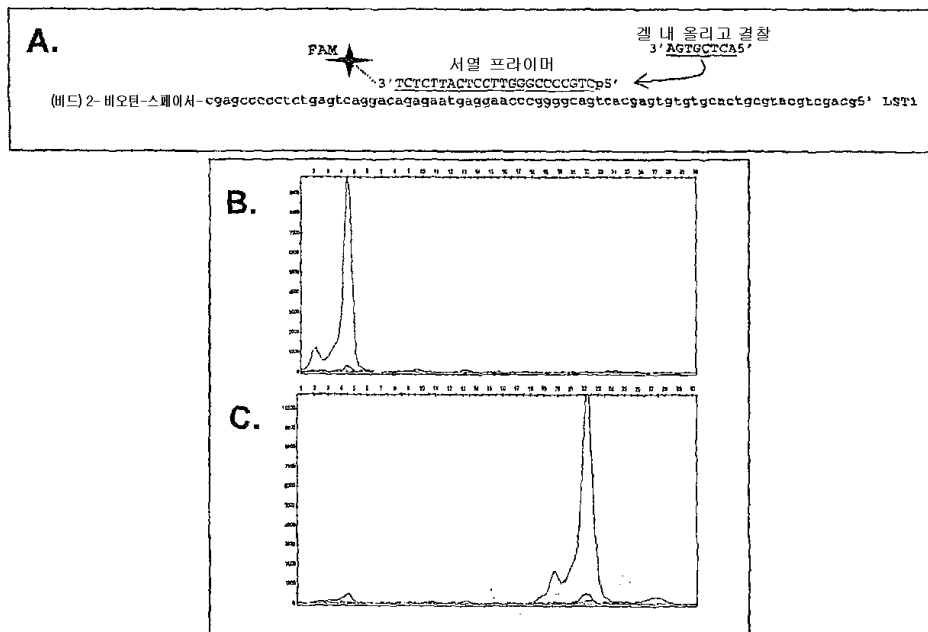
B. Taq DNA 결함 효소

	1	2	3	4	5	6	7	8
	T	C	A	C	G	A	G	T
T	100% (52)	0	2% (1)	0	0	0	0	92% (48)
C	0	100% (52)	2% (1)	100% (52)	0	6% (3)	8% (4)	2% (1)
A	0	0	94% (49)	0	2% (1)	94% (49)	0	6% (3)
G	0	0	2% (1)	0	98% (51%)	0	92% (48)	0

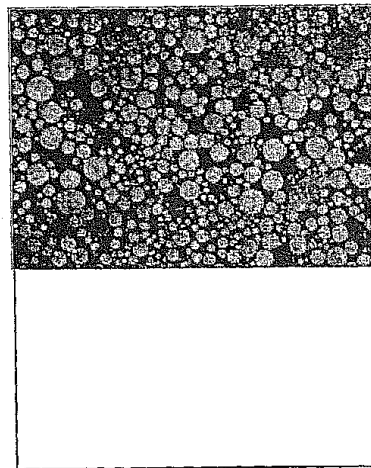
C. T4 DNA 결함 효소

	1	2	3	4	5	6	7	8
	T	C	A	C	G	A	G	T
T	100% (38)	0	5% (2)	0	0	13% (5)	21% (8)	76% (29)
C	0	100% (38)	5% (2)	100% (38)	0	2.5% (1)	13% (5)	8% (3)
A	0	0	90% (34)	0	2% (1)	68% (26)	8% (3)	8% (3)
G	0	0	0	0	100% (38)	16% (6)	58% (22)	8% (3)

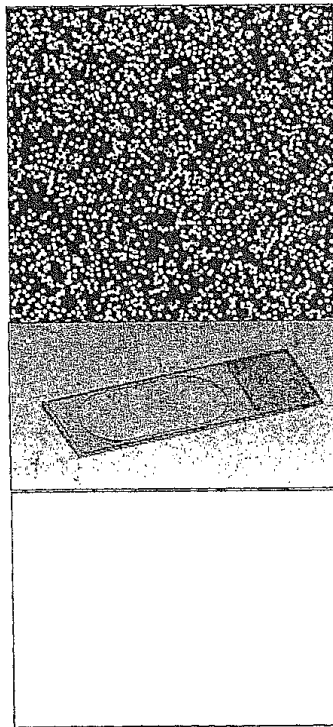
도면13



도면14A



도면14B

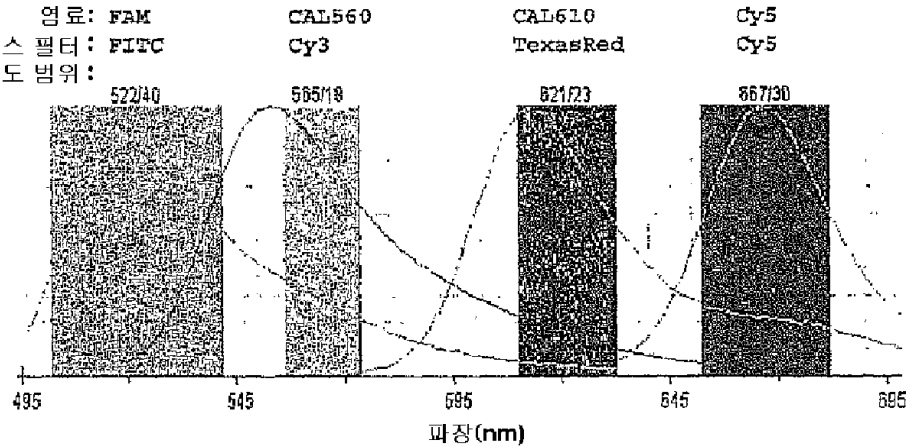


도면15

표 1. 4-색 프로브 검출 및 디자인

4-색 프로브 검출

염료: FAM
밴드패스 필터: FITC
흡광도 범위:



올리고뉴클레오타이드 프로브 디자인

원래의 주형	염료(필터)	Ex ¹	Em	독특한 ²	축퇴성
LST1.A	FAM ³ (FITC)	495	520	AGTGCTCA ^{†,‡}	ANNNNNINI [‡]
LST1.G	Cy5 ⁴ (Cy5)	643	667	GGTGCTCA [‡]	
LST1.C	CAL610 ⁵ (TxRd)	590	610	CGTGCTCA [‡]	
LST1.T	CAL560 ⁶ (Cy3)	540	561	TGTGCTCA [‡]	
LST2	Cy5	643	667	GCCACGAA [†]	GNNNNINI [‡]
LST3	CAL610	590	610	CCACATAT [†]	CNNNNINI [‡]
LST4	CAL560	540	561	TAAGAGCT [†]	TNNNNINI [‡]

¹ 여기 및 방출 흡광도 값(nm)

² 5'-플루오로포어 표지가 있는 올리고뉴클레오타이드 서열(3'-5')

³ (6)-카복시 플루오레신(FAM)

⁴ CAL 형광 오렌지(CAL560)(특허 계류중,바이오서치 테크놀로지스,노바토,CA)

⁵ CAL 형광 레드(CAL610)(특허 계류중,바이오서치 테크놀로지스,노바토,CA)

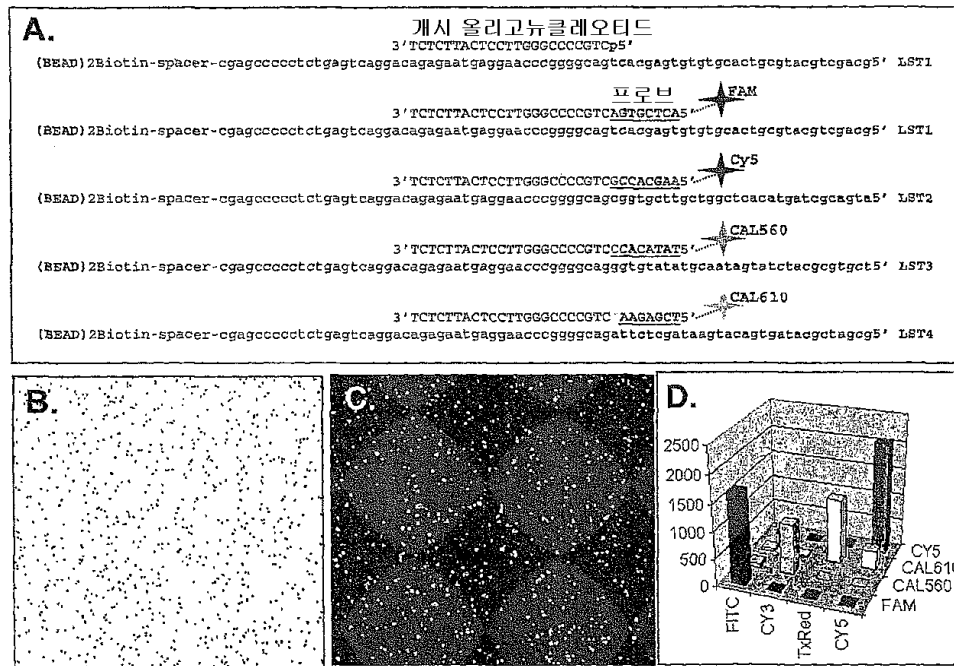
⁶ 시아닌 5(Cy5)

[†] 프로브 세트 #1

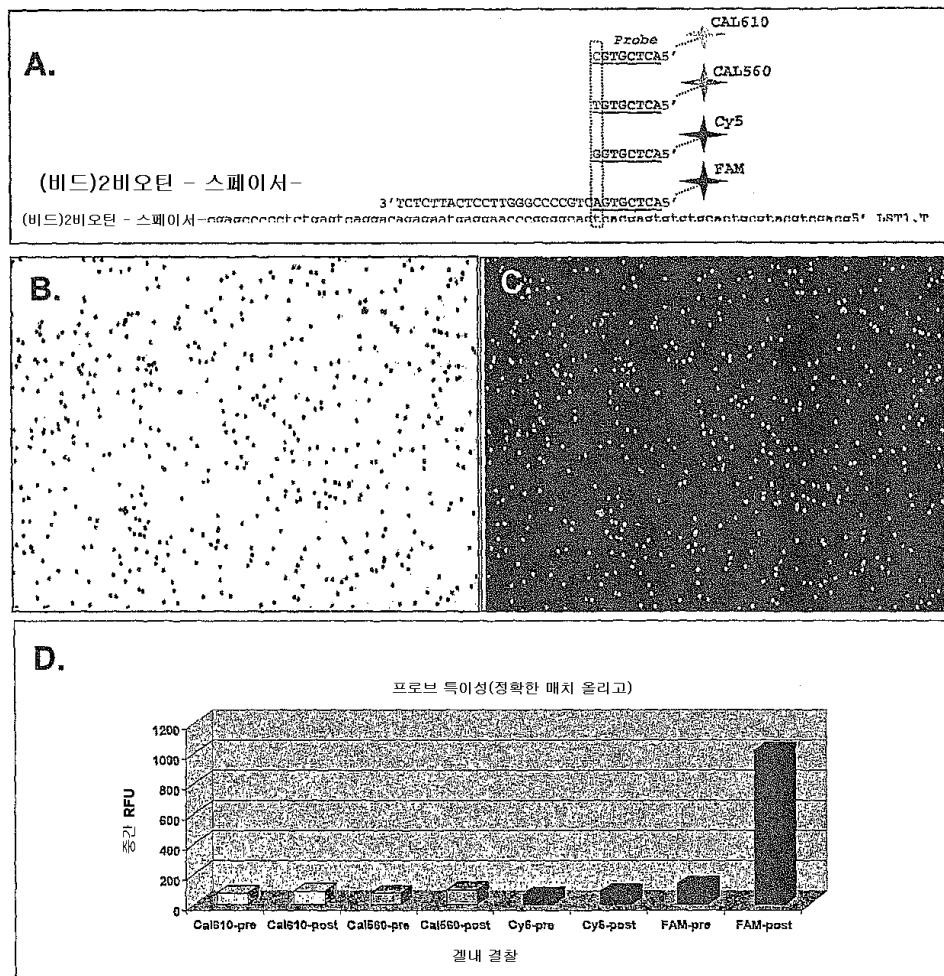
[‡] 프로브 세트 #2

[‡] 프로브 세트 #3

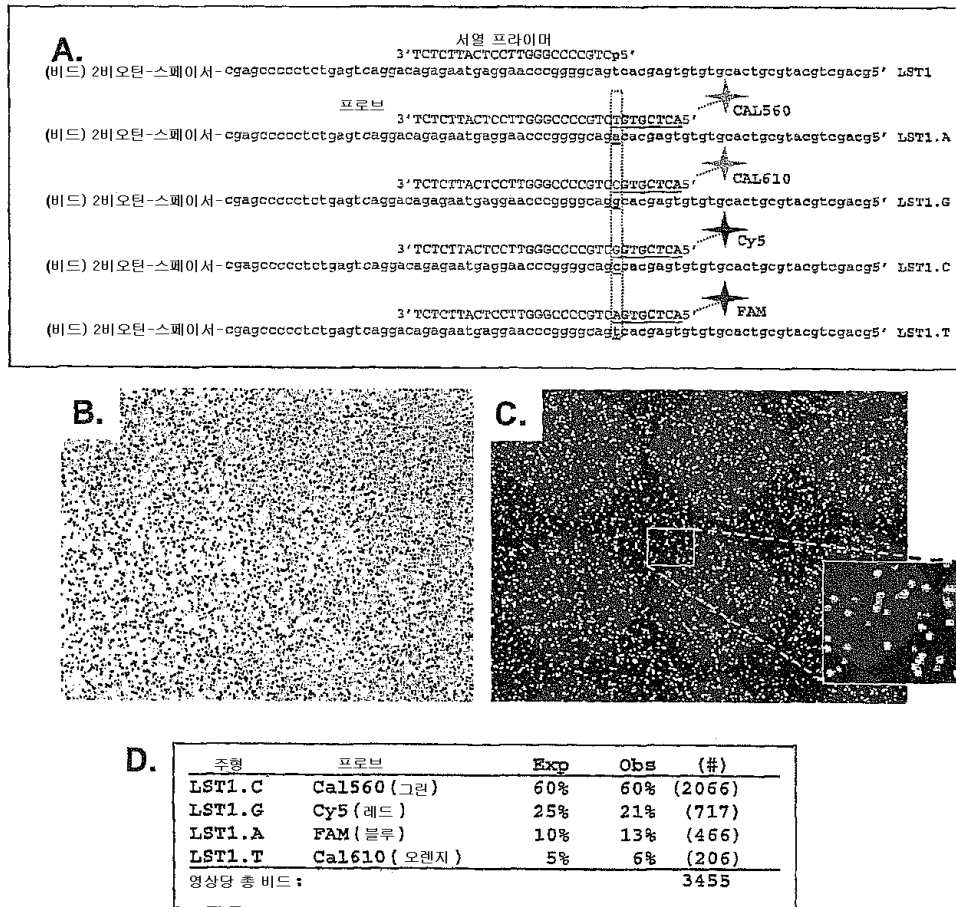
도면16



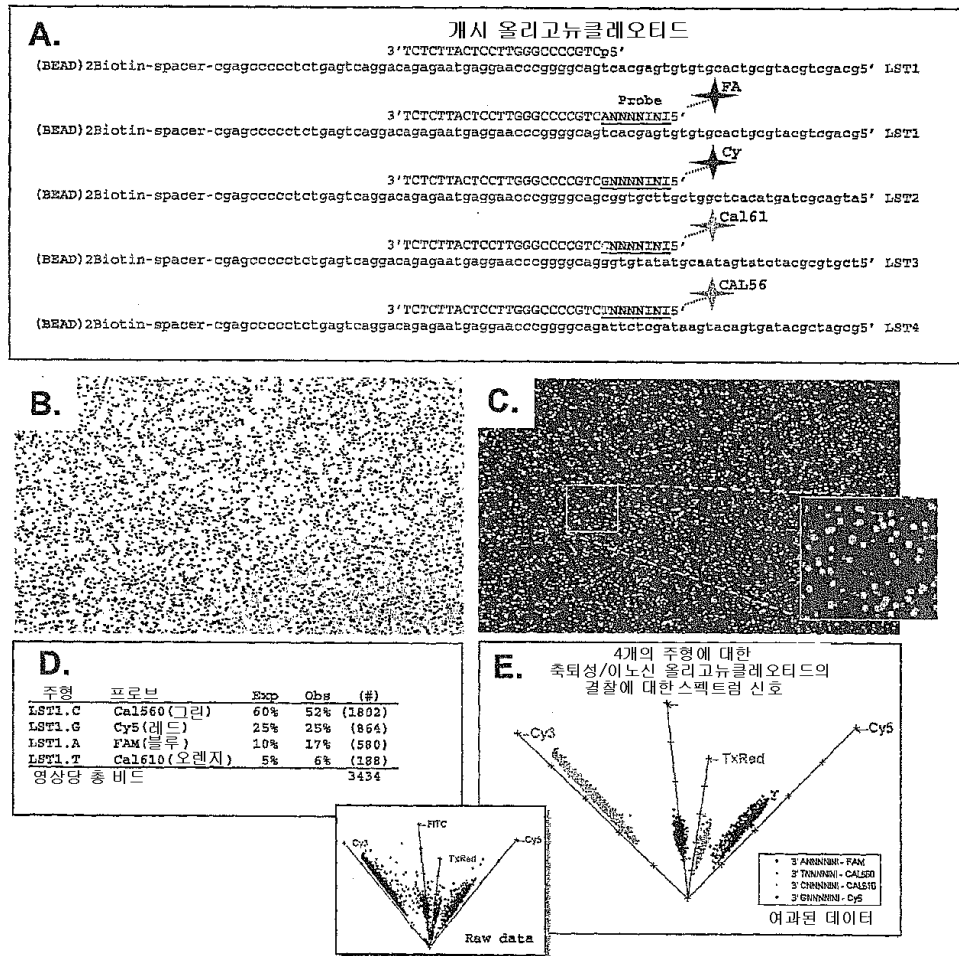
도면17



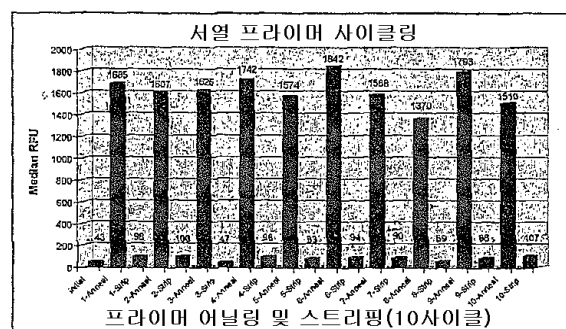
도면18



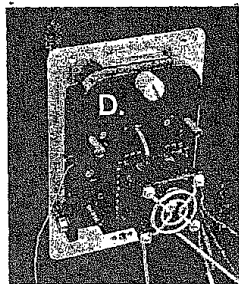
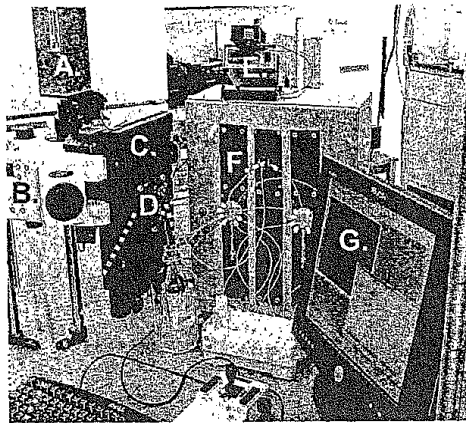
도면19



도면20

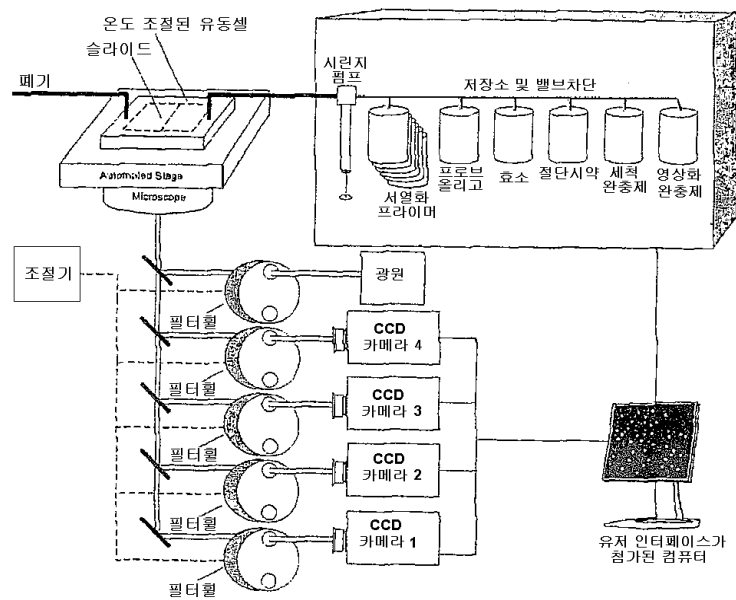


도면21

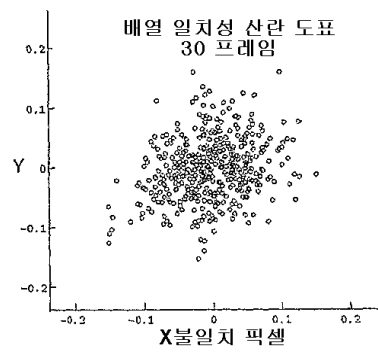


- A. CCD 카메라
B. 형광 현미경
C. 이동성 스테이지
D. 펄스에 유동 셀
E. 주형 조절
F. 유체 조작
G. 컴퓨터

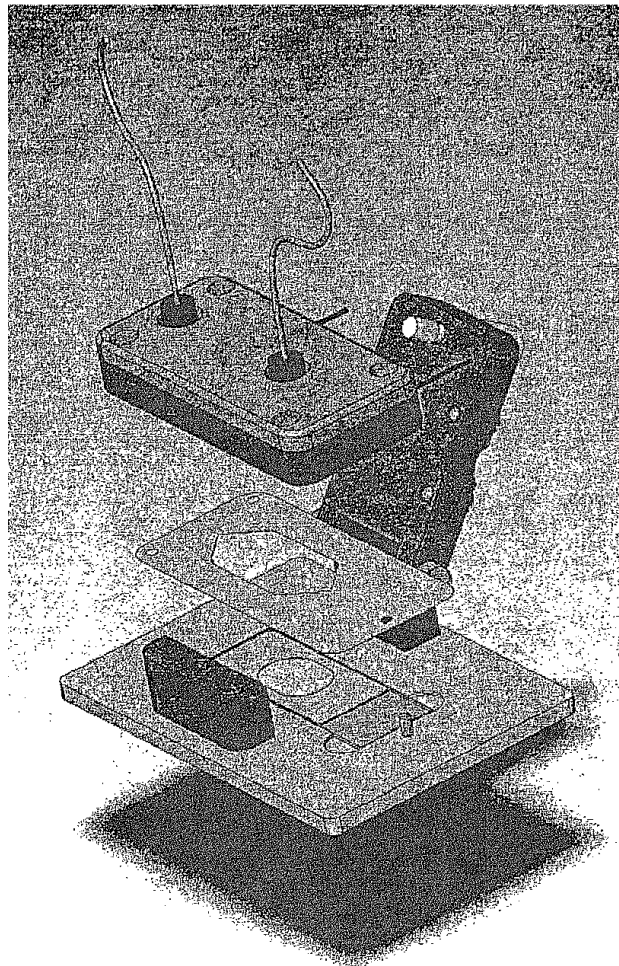
도면22



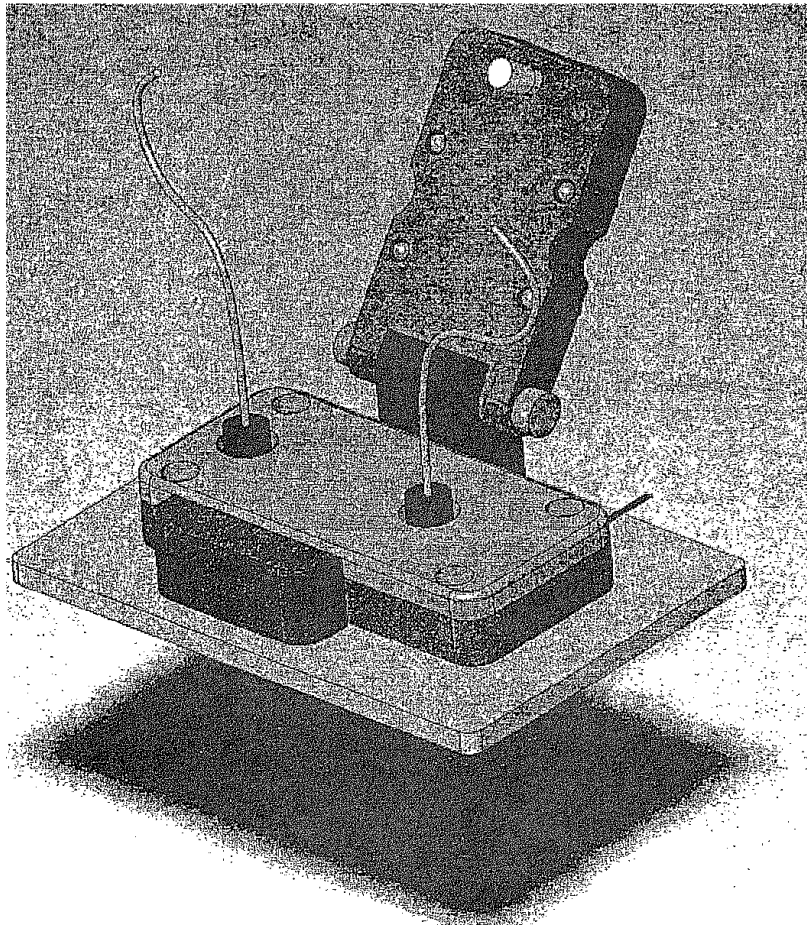
도면23



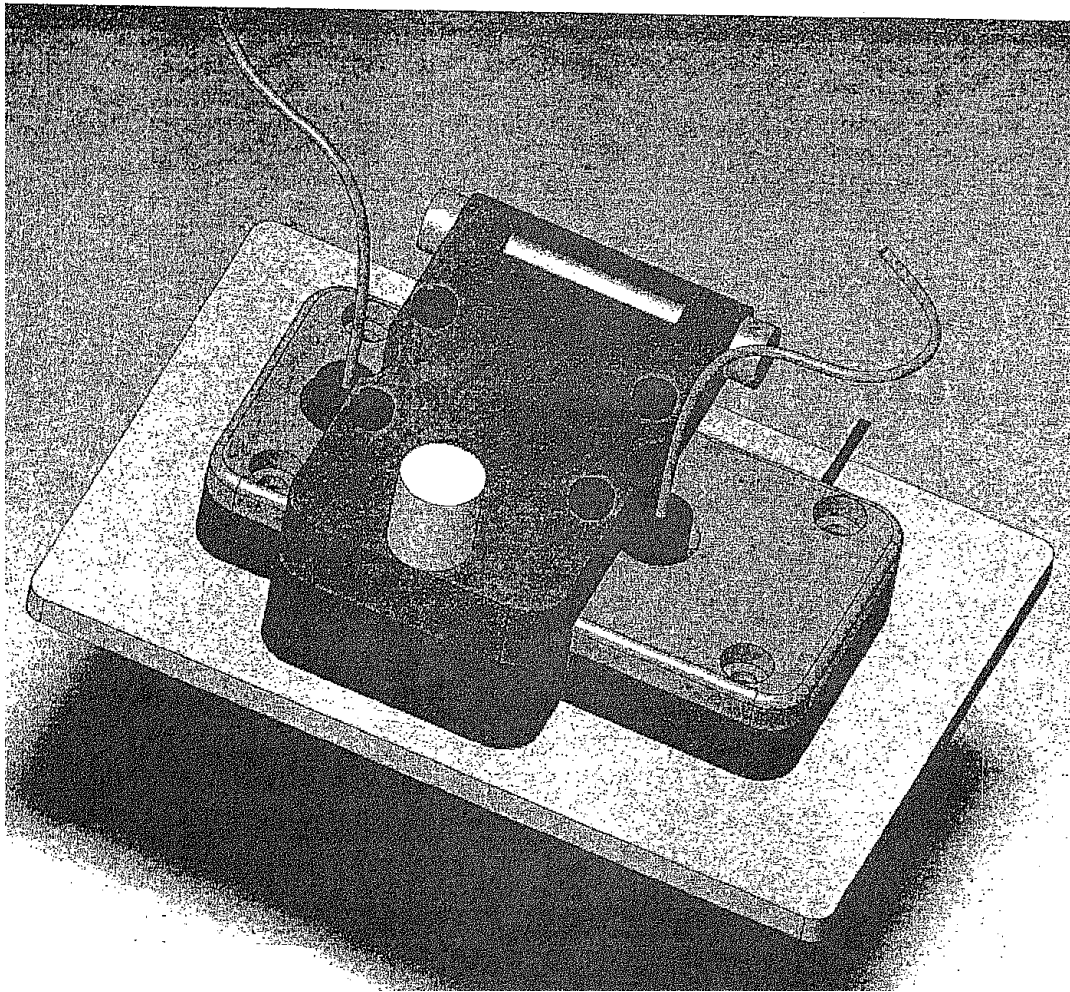
도면24A



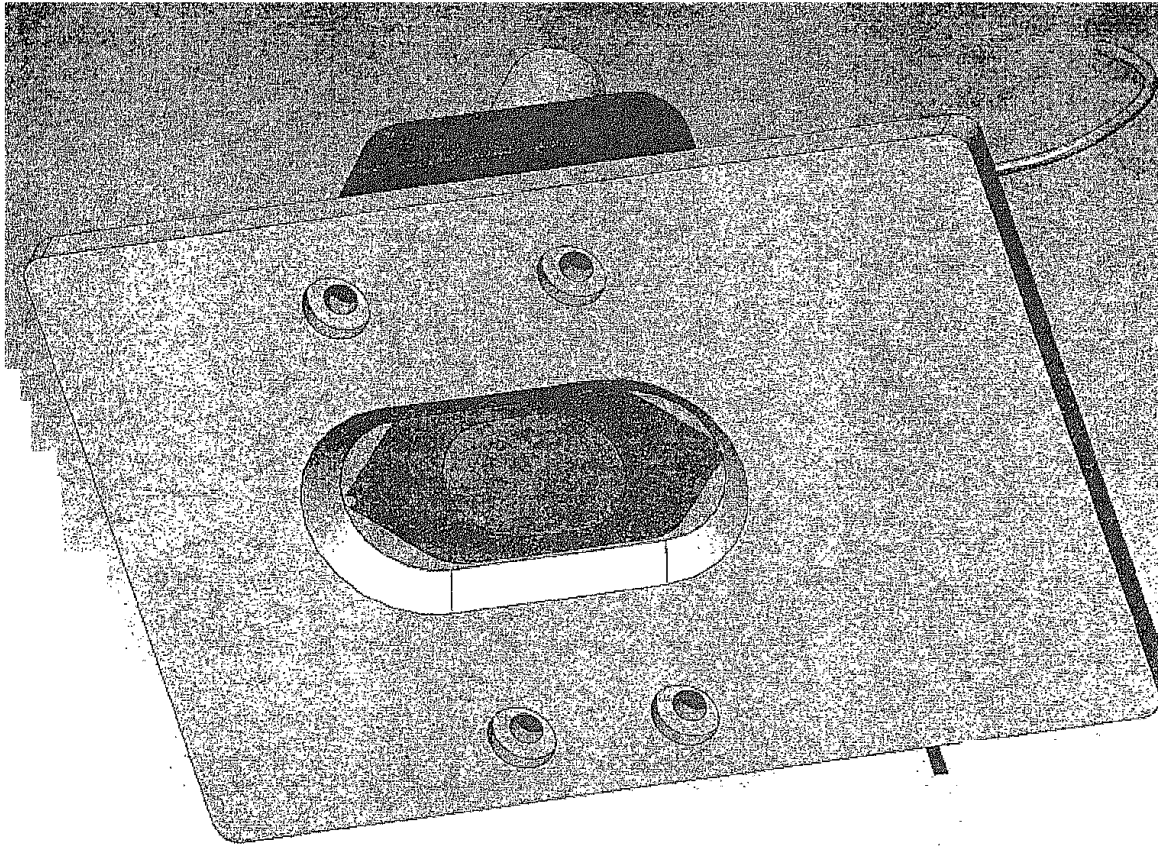
도면24B



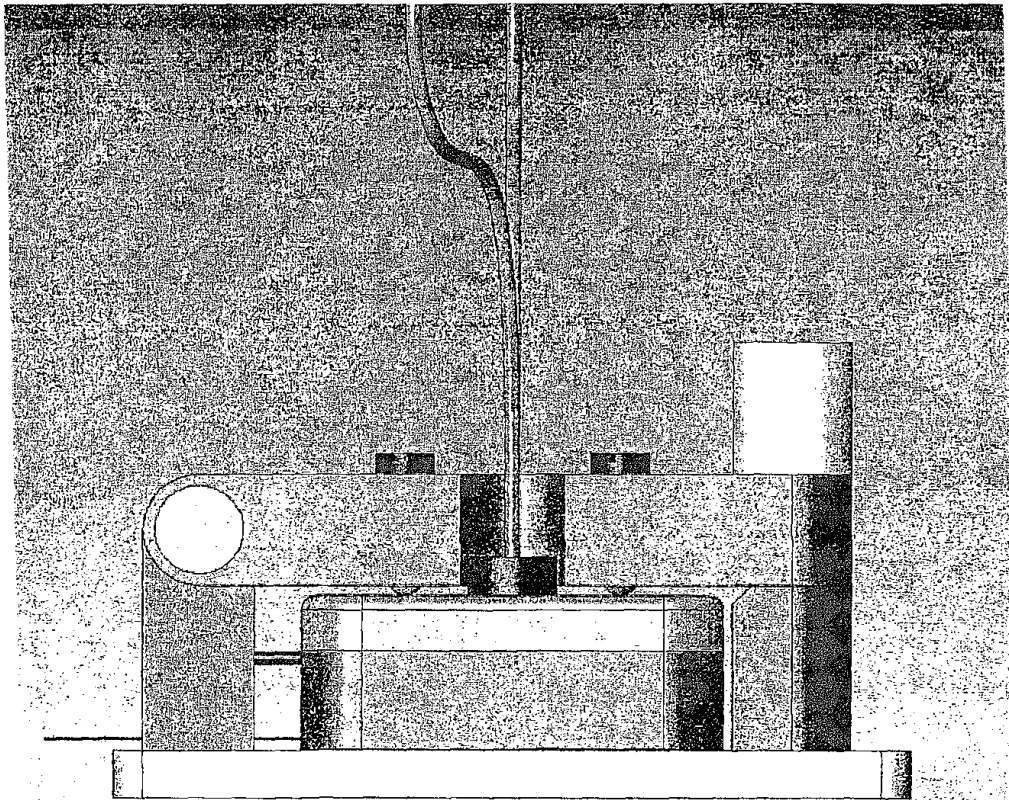
도면24C



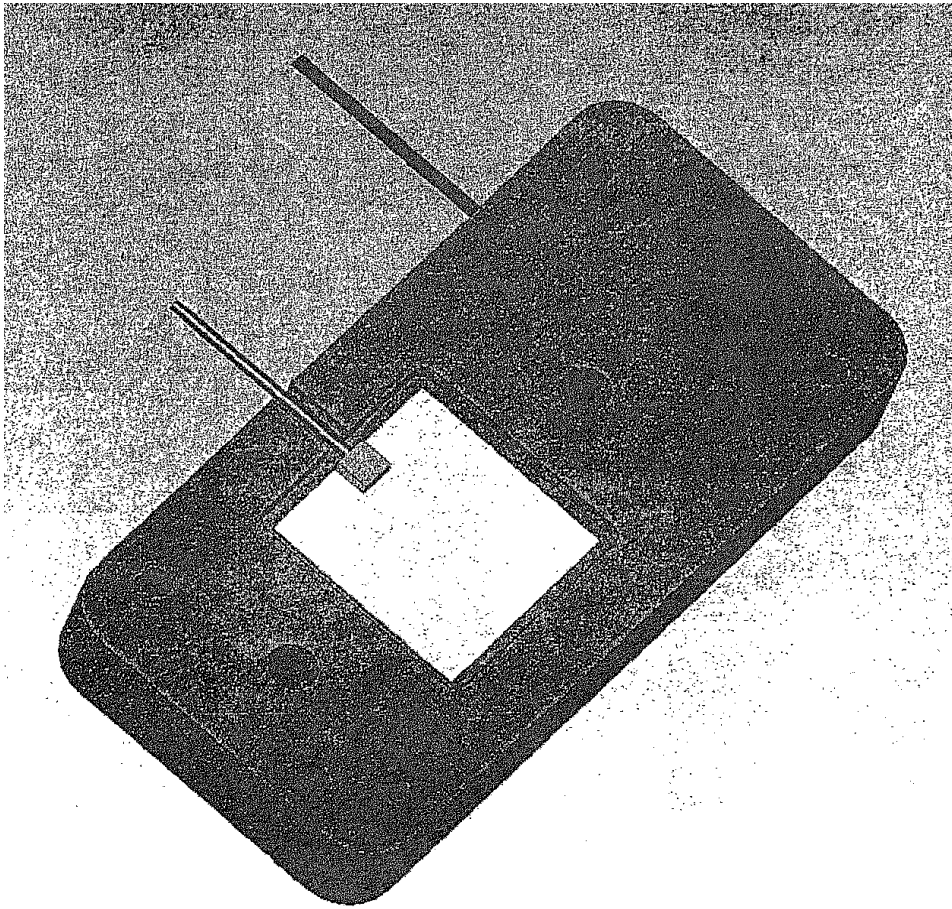
도면24D



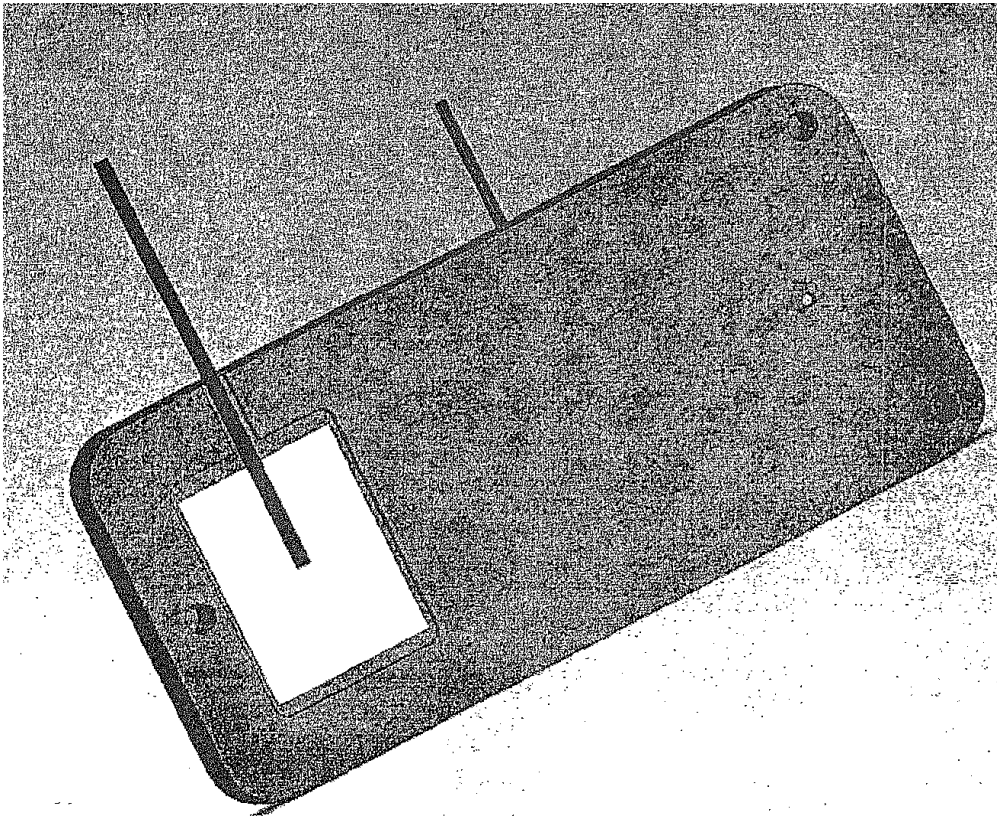
도면24E



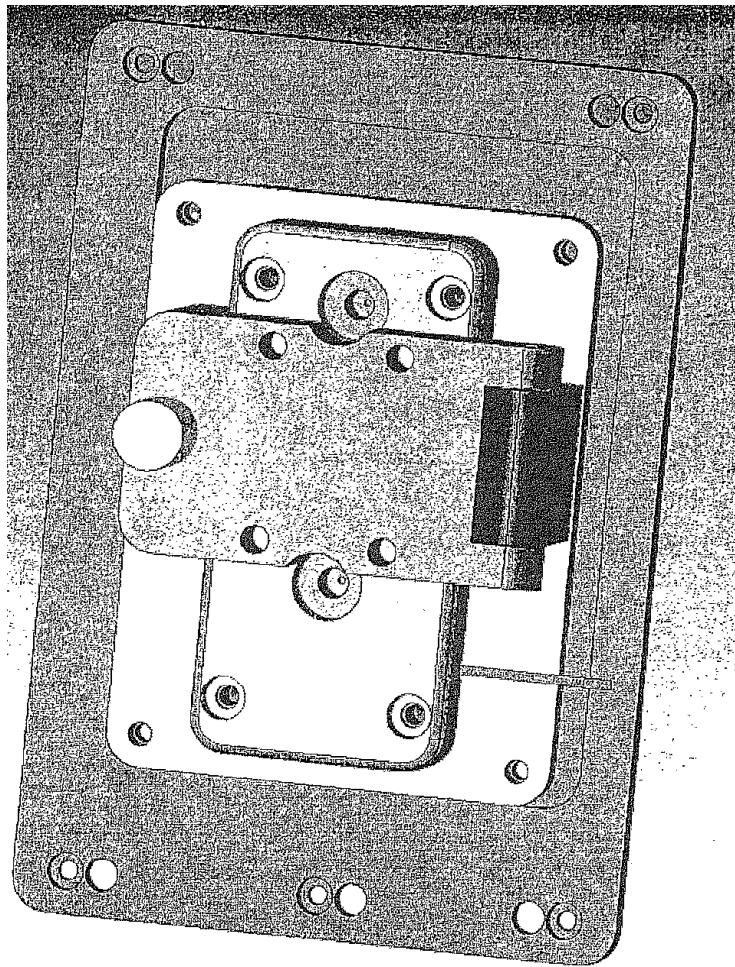
도면24F



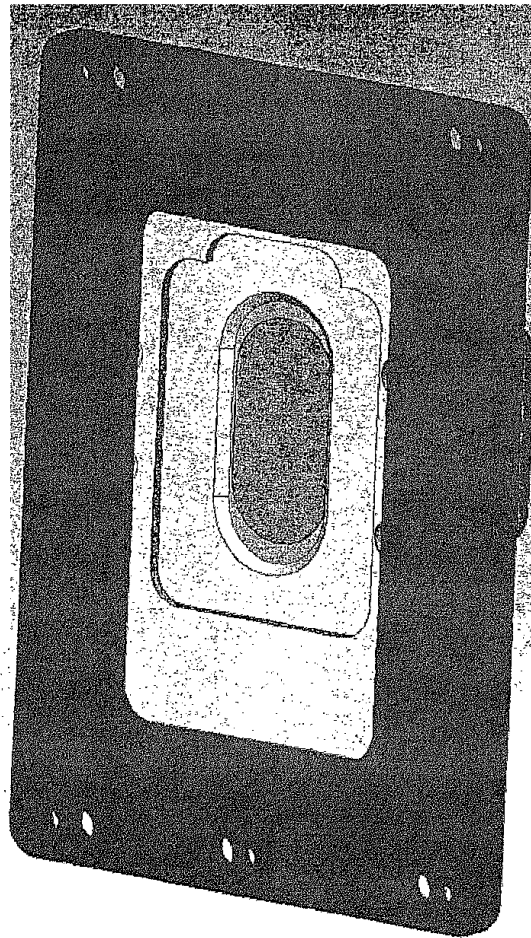
도면24G



도면24H



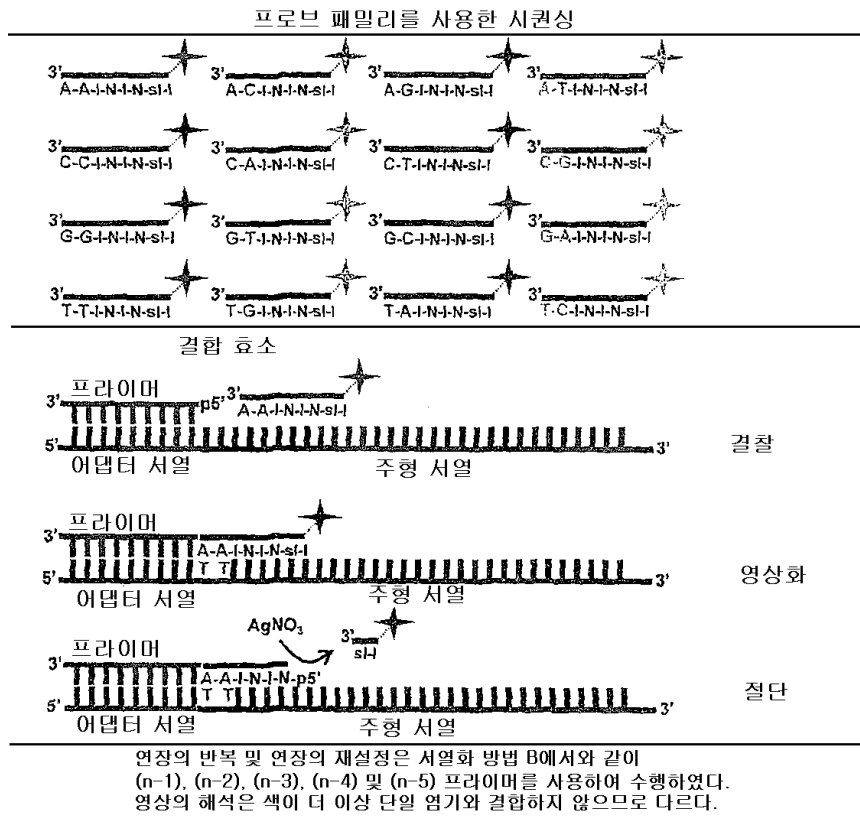
도면24I



도면25A

표지	5'					3'
레드	N	N	N	N	T	C
레드	N	N	N	N	G	A
레드	N	N	N	N	A	G
레드	N	N	N	N	C	T
옐로우	N	N	N	N	C	C
옐로우	N	N	N	N	T	A
옐로우	N	N	N	N	G	G
옐로우	N	N	N	N	A	T
그린	N	N	N	N	A	C
그린	N	N	N	N	C	A
그린	N	N	N	N	T	G
그린	N	N	N	N	G	T
블루	N	N	N	N	G	C
블루	N	N	N	N	A	A
블루	N	N	N	N	C	G
블루	N	N	N	N	T	T

도면25B



도면26

표지	5'					3'
레드	N	N	N	N	A	A
레드	N	N	N	N	C	G
레드	N	N	N	N	G	T
레드	N	N	N	N	T	C
옐로우	N	N	N	N	A	C
옐로우	N	N	N	N	C	A
옐로우	N	N	N	N	G	G
옐로우	N	N	N	N	T	T
그린	N	N	N	N	A	T
그린	N	N	N	N	C	C
그린	N	N	N	N	G	A
그린	N	N	N	N	T	G
블루	N	N	N	N	A	G
블루	N	N	N	N	C	T
블루	N	N	N	N	G	C
블루	N	N	N	N	A	T

도면27A

첫번째 염기	두번째 염기			
	표지 1	표지 2	표지 3	표지 4
A	A	C	G	T
C	C	A	T	G
G	G	T	C	A
T	T	G	A	C

도면27B

첫번째 염기	두번째 염기	표지
A	A	1
C	C	1
G	G	1
T	T	1
A	C	2
C	A	2
G	T	2
T	G	2
A	G	3
C	T	3
G	C	3
T	A	3
A	T	4
C	G	4
G	A	4
T	C	4

도면27C

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	C	A	T	G	C	C	G	T	A	C	C	T	A	G	C	C	A	T	G
G	G	T	C	A	G	G	T	A	C	G	G	A	T	C	G	G	T	A	C
T	T	G	A	C	T	T	A	C	G	A	T	G	C	A	T	T	G	C	A

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	C	A	T	G	C	C	G	T	A	C	C	T	A	G	C	C	A	T	G
G	T	G	A	C	G	T	A	C	G	G	T	G	C	A	G	T	G	C	A
T	G	T	C	A	T	G	T	A	C	T	G	A	T	C	T	G	T	A	C

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	G	A	T	C	C	G	T	A	C	C	G	T	C	A	C	C	G	T	A
G	C	T	A	G	G	C	A	T	G	G	C	A	T	G	G	C	G	T	A
T	T	G	C	A	T	T	G	C	A	T	T	G	A	C	T	T	A	C	G

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	G	A	T	C	C	G	T	A	C	C	G	T	A	C	C	G	T	C	A
G	T	G	C	A	G	T	A	C	G	G	T	G	C	A	G	G	T	G	A
T	C	T	A	G	T	C	G	T	A	T	C	A	T	G	T	C	A	T	G

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	T	A	C	G	C	T	G	A	C	C	T	G	C	A	C	C	T	G	C
G	C	G	T	A	G	G	C	A	T	G	G	C	T	A	G	G	C	A	T
T	G	T	A	C	T	G	T	C	A	T	G	A	T	C	T	G	T	A	C

A	A	C	G	T	A	A	C	G	T	A	A	C	G	T	A	A	C	G	T
C	T	A	C	G	C	T	G	C	A	C	C	T	G	A	C	C	T	G	C
G	G	T	A	C	G	G	A	T	C	G	G	T	C	A	G	G	T	A	C
T	C	G	T	A	T	C	T	A	G	T	C	A	T	G	T	C	A	T	G

도면28

염기 1	염기 2	표지
A	A	1
A	C	1
A	G	3
A	T	3
C	A	2
C	C	2
C	G	4
C	T	4
G	A	1
G	C	1
G	G	3
G	T	3
T	A	2
T	C	2
T	G	4
T	T	4

도면29A

	1	2	3	4	
A	A	T	G	C	A
C	C	G	T	A	
G	G	C	A	T	
T	T	A	C	G	
A	C	A	T	G	C
C	A	C	G	T	
G	T	G	C	A	
T	G	T	A	C	
A	G	C	A	T	G
C	T	A	C	G	
G	A	T	G	C	
T	C	G	T	A	
A	T	G	C	A	T
C	G	T	A	C	
G	C	A	T	G	
T	A	C	G	T	

도면29B

A	A	C	G	T
C	C	A	T	G
G	G	T	A	C
T	T	G	C	A

	1	2	3	4	
A	A	T	G	C	A
C	C	G	T	A	
G	G	C	A	T	
T	T	A	C	G	
A	C	A	T	G	C
C	A	C	G	T	
G	T	G	C	A	
T	G		A	C	
A	G	C	A	T	G
C	T	A	C	G	
G	A	T	G	C	
T	C	G	T	A	
A	T	G	C	A	T
C	G	T	A	C	
G	C	A	T	G	
T	A	C	G	T	

도면30

관찰된 표지	서열화 가능성				
	5	4	3	2	1
옐로우				C	C
				T	A
				G	G
				A	T
그린			A	C	
			C	A	
			T	G	
			G	T	
레드		T	C		
		G	A		
		A	G		
		C	T		
블루	G	C			
	A	A			
	C	G			
	T	T			
4가지 가능한 서열	C	G	A	C	C
	A	A	G	T	A
	G	C	T	G	G
	T	T	C	A	T

도면31A

표지	5'					3'
레드	N	N	N	N	A	A
레드	N	N	N	N	C	C
레드	N	N	N	N	G	G
레드	N	N	N	N	T	T
엘로우	N	N	N	N	A	C
엘로우	N	N	N	N	C	A
엘로우	N	N	N	N	G	T
엘로우	N	N	N	N	T	G
그린	N	N	N	N	A	G
그린	N	N	N	N	C	T
그린	N	N	N	N	G	A
그린	N	N	N	N	T	C
블루	N	N	N	N	A	T
블루	N	N	N	N	C	G
블루	N	N	N	N	G	C
블루	N	N	N	N	T	A

도면31B

관찰된 표지 서열화 가능성	서열화 가능성				
	5	4	3	2	1
엘로우				A	C
				C	A
				G	T
				T	G
그린			A	G	
			C	T	
			G	A	
			T	C	
레드		A	A		
		C	C		
		G	G		
		T	T		
블루	A	T			
	C	G			
	G	C			
	T	A			
4가지 가능한 서열	C	G	G	A	C
	A	T	T	C	A
	G	C	C	T	G
	T	A	A	G	T

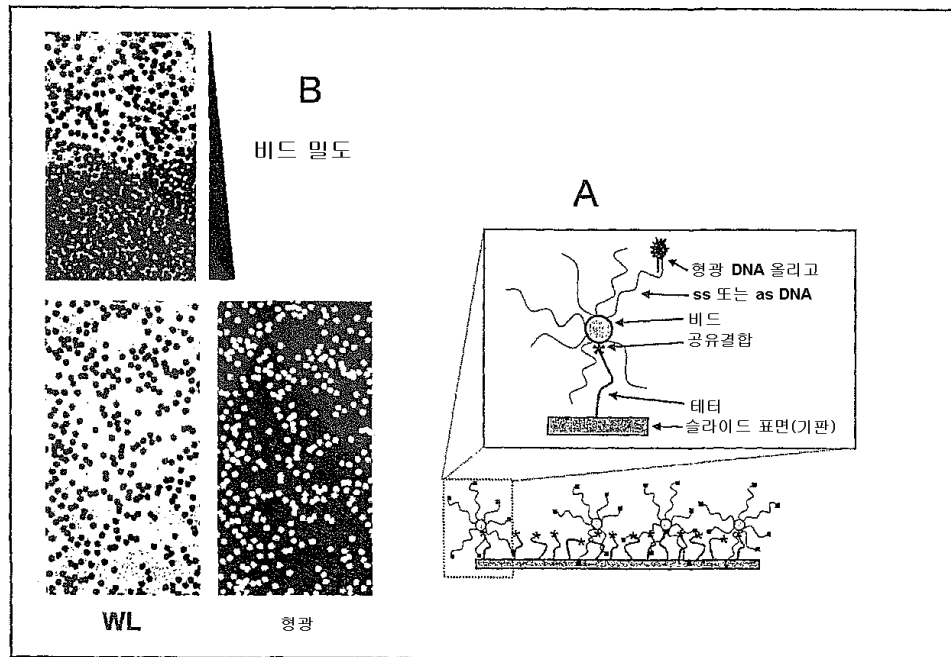
도면31C

표지	5'					3'
레드	N	N	N	N	A	A
레드	N	N	N	N	C	G
레드	N	N	N	N	G	T
레드	N	N	N	N	T	C
옐로우	N	N	N	N	A	C
옐로우	N	N	N	N	C	T
옐로우	N	N	N	N	G	G
옐로우	N	N	N	N	T	A
그린	N	N	N	N	A	G
그린	N	N	N	N	C	A
그린	N	N	N	N	G	C
그린	N	N	N	N	T	T
블루	N	N	N	N	A	T
블루	N	N	N	N	C	C
블루	N	N	N	N	G	A
블루	N	N	N	N	T	G

도면32

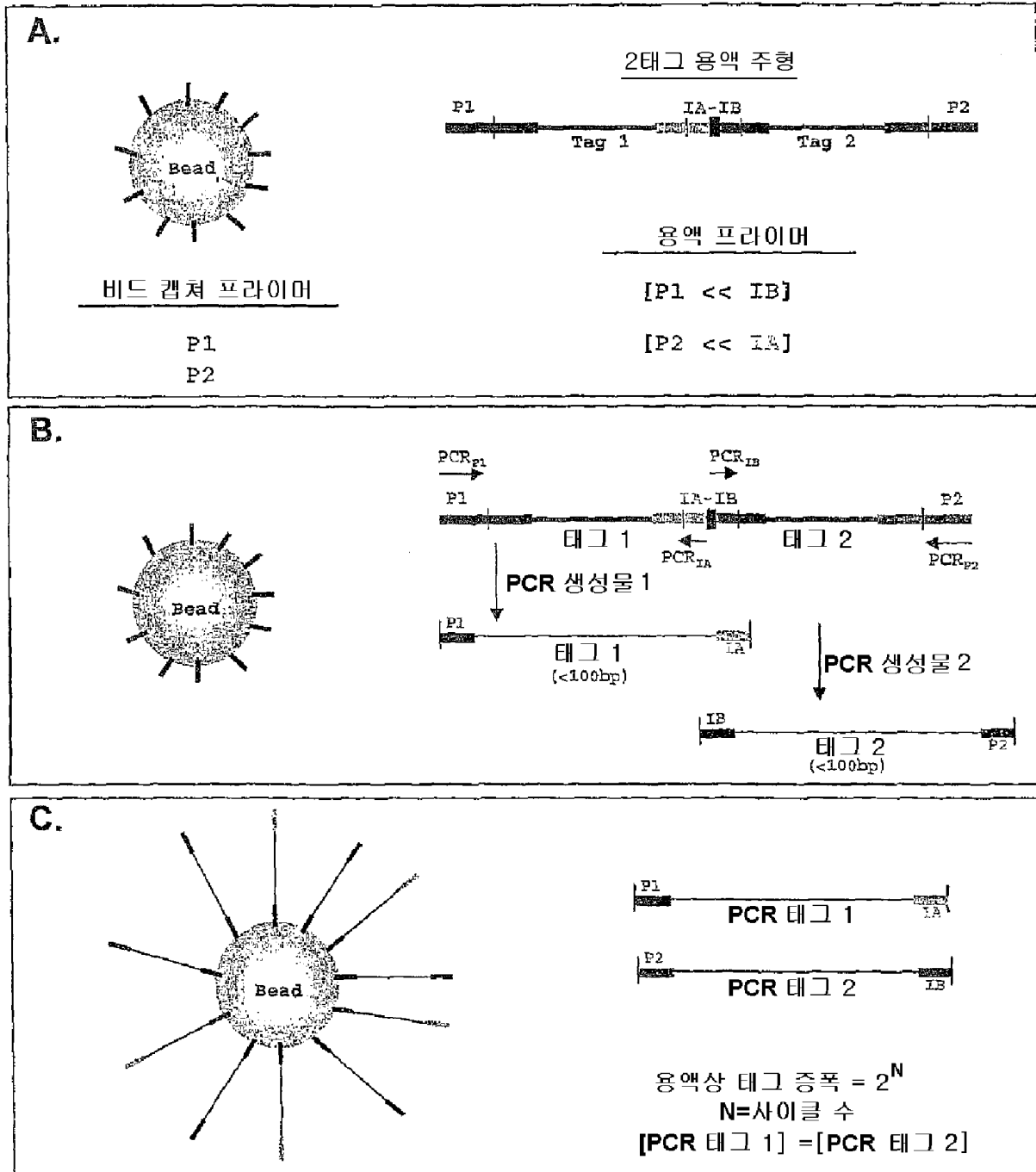
염기						표지
1	2	3	4	5	6	
A	A					1
A	C					
G	A					
G	C					
	C	A				2
	C	C				
	T	A				
	T	C				
		A	G			3
		A	T			
		G	G			
		G	T			
			C	G		4
			C	T		
			T	G		
			T	T		
				A	A	1
				A	C	
				G	A	
				G	C	
A	C	A	T	G	A	4가지 가능한 서열
G	C	A	T	G	A	
A	C	A	T	G	C	
G	C	A	T	G	C	

도면33



도면34

2태그 증폭 도식 :
라이브러리 단편(~200-bp)
쌍을 이룬 단부 디자인



도면35

2태크 증폭 도식:
프라이머 디자인

