

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5180845号
(P5180845)

(45) 発行日 平成25年4月10日 (2013. 4. 10)

(24) 登録日 平成25年1月18日 (2013.1.18)

(51) Int. Cl.

F I

C 1 2 N 15/09 (2006.01)

C 1 2 N 15/00 Z N A A

C 1 2 Q 1/68 (2006.01)

C 1 2 N 15/00 F

G O 1 N 33/543 (2006.01)

C 1 2 Q 1/68 A

G O 1 N 33/53 (2006.01)

G O 1 N 33/543 5 2 5 U

G O 1 N 33/543 5 2 5 W

請求項の数 8 (全 73 頁) 最終頁に続く

(21) 出願番号 特願2008-556583 (P2008-556583)
 (86) (22) 出願日 平成19年2月26日 (2007. 2. 26)
 (65) 公表番号 特表2010-500002 (P2010-500002A)
 (43) 公表日 平成22年1月7日 (2010. 1. 7)
 (86) 国際出願番号 PCT/US2007/062825
 (87) 国際公開番号 W02007/133831
 (87) 国際公開日 平成19年11月22日 (2007. 11. 22)
 審査請求日 平成20年10月27日 (2008. 10. 27)
 (31) 優先権主張番号 60/776, 415
 (32) 優先日 平成18年2月24日 (2006. 2. 24)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 507411419
 カリダ・ジェノミックス・インコーポレイ
 テッド
 CALLIDA GENOMICS, I
 NC.
 アメリカ合衆国94085カリフォルニア
 州サニーベイル、ノース・パストリア・ア
 ベニュー750番
 (74) 代理人 100081422
 弁理士 田中 光雄
 (74) 代理人 100084146
 弁理士 山崎 宏
 (74) 代理人 100106518
 弁理士 松谷 道子

最終頁に続く

(54) 【発明の名称】 DNAアレイ上でのハイスループットゲノム配列決定

(57) 【特許請求の範囲】

【請求項 1】

以下の工程を含む、ポリヌクレオチドの配列情報を決定する方法：

(a) 基体上にアレイ化された複数のコンカテマーを提供する工程；ここで、各コンカテマーは複数コピーのモノマー単位を含み、各モノマー単位は以下を含む：

i) 該ポリヌクレオチドの第一の標的配列；

ii) 該ポリヌクレオチドの第二の標的配列、ここで、該第一の標的配列と該第二の標的配列は該ポリヌクレオチドの近接するセグメントであったものである；

iii) 該第一の標的配列と該第二の標的配列の間に挿入された散在したアダプター、ここで、該散在したアダプターは該第一の標的配列に隣接する第一のプロープハイブリダイゼーション部位および該第二の標的配列に隣接する第二のプロープハイブリダイゼーション部位を含む；および、

(b) 該アレイ化されたコンカテマーの1以上において、該散在したアダプターに隣接する該第一の標的配列の1以上のヌクレオチドおよび該第二の標的配列の1以上のヌクレオチドを同定する工程；ここで、該第一の標的配列の1以上のヌクレオチドを同定する工程が以下の工程を含むか：

(i) 該散在したアダプター中の第一のプロープハイブリダイゼーション部位にアンカープロープをハイブリダイズさせる工程；

(ii) 該ハイブリダイズしたアンカープロープに隣接する標的配列に配列決定プロープをハイブリダイズさせる工程； 次いで

10

20

(iii) 該配列決定プローブと該アンカープローブをライゲーションする工程； および
 (iv) 該ライゲーションしたプローブを検出し、該隣接する標的配列のヌクレオチドを
同定する工程、または

該第二の標的配列の 1 以上のヌクレオチドを同定する工程が以下の工程を含む：

(i) 該散在したアダプター中の第二のプローブハイブリダイゼーション部位にアンカー
プローブをハイブリダイズさせる工程；

(ii) 該ハイブリダイズしたアンカープローブに隣接する標的配列に配列決定プローブ
をハイブリダイズさせる工程； 次いで

(iii) 該配列決定プローブと該アンカープローブをライゲーションする工程； および

(iv) 該ライゲーションしたプローブを検出し、該隣接する標的配列のヌクレオチドを
同定する工程。

10

【請求項 2】

以下の工程を含む、ポリヌクレオチドの配列情報を決定する方法：

(a) 基体上にアレイ化された複数のコンカテマーを提供する工程；ここで、各コンカテマ
ーは複数コピーのモノマー単位を含み、各モノマー単位は以下を含む：

i) 該ポリヌクレオチドの第一の標的配列；

ii) 該ポリヌクレオチドの第二の標的配列、ここで、該第一の標的配列と該第二の標的
配列は該ポリヌクレオチドの近接するセグメントであったものである；

iii) 該第一の標的配列と該第二の標的配列の間に挿入された散在したアダプター、こ
こで、該散在したアダプターは該第一の標的配列に隣接する第一のプローブハイブリダイ
ゼーション部位および該第二の標的配列に隣接する第二のプローブハイブリダイゼーシ
ョン部位を含む； および、

20

(b) 該アレイ化されたコンカテマーの 1 以上において、該散在したアダプターに隣接する
該第一の標的配列の 1 以上のヌクレオチドおよび該第二の標的配列の 1 以上のヌクレオチ
ドを同定する工程；ここで、該第一の標的配列の 1 以上のヌクレオチドを同定する工程が
以下の工程を含むか：

(i) 該散在したアダプター中の第一のプローブハイブリダイゼーション部位にアンカー
プローブをハイブリダイズさせる工程；

(ii) dNTPが第一の標的配列中の検出位置に完全に相補的であれば該 dNTP がアンカー
プローブに付加される条件下で、ポリメラーゼおよび標識を含む少なくとも 1 つのdNTPを
添加する工程； および

30

(iii) 該検出位置においてアンカープローブに付加された標識されたdNTPと相補的な、
第一の標的配列中のヌクレオチドを同定する工程、または

該第二の標的配列の 1 以上のヌクレオチドを同定する工程が以下の工程を含む：

(i) 該散在したアダプター中の第二のプローブハイブリダイゼーション部位にアンカー
プローブをハイブリダイズさせる工程；

(ii) dNTPが第二の標的配列中の検出位置に完全に相補的であれば該 dNTP がアンカー
プローブに付加される条件下で、ポリメラーゼおよび標識を含む少なくとも 1 つのdNTPを
添加する工程； および

(iii) 該検出位置においてアンカープローブに付加された標識されたdNTPと相補的な、
第二の標的配列中のヌクレオチドを同定する工程。

40

【請求項 3】

工程(i)から(iv)までを複数回繰り返して標的配列の複数のヌクレオチドを同定する、
請求項 1 の方法。

【請求項 4】

配列決定プローブが、該プローブ上の検出可能な特定の標識と結合した少なくとも 1 つ
のユニークな塩基を有する調査配列を含むものである、請求項 1 の方法。

【請求項 5】

第二の標的配列の第二のヌクレオチドを同定する工程および第三の標的配列のヌクレオ
チドを同定する工程をさらに含み、かつ、各モノマー単位が以下をさらに含むものである

50

、請求項 1 ~ 4 いずれかの方法：

(a) ポリヌクレオチドの第三の標的配列；

(b) 第二の標的配列と第三の標的配列の間に挿入された第二の散在したアダプター、ここで、第二の散在したアダプターは第三のプロープハイブリダイゼーション部位および第四のプロープハイブリダイゼーション部位を含む。

【請求項 6】

ポリヌクレオチドがゲノムDNAである、請求項 1 ~ 4 いずれかの方法。

【請求項 7】

コンカテマーが平均 1 ミクロンの最近隣距離で基体上にアレイ化されており、かつ、該アレイ上のコンカテマーの少なくとも 70 パーセントが光学的に分解可能なものである、請求項 1 ~ 4 いずれかの方法。

10

【請求項 8】

該複数のコンカテマーが直線パターンで基体上に存在し；

該コンカテマーが基体上の別々の空間的に離れた領域に結合しており； かつ、

該空間的に離れた領域の実質的に全てが単一のコンカテマーによって占有されている、請求項 1 ~ 4 いずれかの方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願との相互参照

20

本出願は引用によりその全体を本明細書に含める2006年2月24日出願の仮出願第60/776, 415号からの優先権を主張する。

【0002】

連邦支援研究に関する言及

本出願は、部分的には国立衛生研究の助成金番号1 U01 AI057315-01により連邦政府の資金援助を受けた。

【背景技術】

【0003】

発明の背景

ゲノム DNAのラージスケール配列分析は、ヒトおよび多くの経済的に重要な植物および動物における健康および疾患の状態に関する多様な生物学的現象の理解に中心的である。例えば、Collins et al (2003)、Nature、422: 835-847; Service、Science、311: 1544-1546 (2006); Hirschhorn et al (2005)、Nature Reviews Genetics、6: 95-108; National Cancer Institute、Report of Working Group on Biomedical Technology、"Recommendation for a Human Cancer Genome Project," (February、2005); Tring et al (2005)、Nature Reviews Genetics、6: 805-814。低コストハイスループット配列決定および再配列決定の必要性により、多くの標的 DNA 断片を同時に並行して分析するいくつかの新しいアプローチが開発された。例えば、Margulies et al、Nature、437: 376-380 (2005); Shendure et al (2005)、Science、309: 1728-1732; Metzker (2005)、Genome Research、15: 1767-1776; Shendure et al (2004)、Nature Reviews Genetics、5: 335-344; Lapidus et al、米国特許出願公開 US 2006/0024711; Drmanac et al、米国特許出願公開 US 2005/0191656; Brenner et al、Nature Biotechnology、18: 630-634 (2000); 等。かかるアプローチは、平面アレイにおける標的ポリヌクレオチド密度の上昇および特定の配列検出化学の各サイクルにおいて増えていく量の配列情報を得るための様々な解決手段を反映している。これらの新しいアプローチのほとんどはシグナルが顕著に悪化する前に数十のヌクレオチドを決定することに制限されており、それゆえ全体の配列決定効率に制限を与えている。

30

40

【0004】

伝統的なハイスループット配列決定技術のもう一つの制限は、多くの配列決定方法に用いられるアレイ表面上のDNA 標的のランダムな位置取りが、例えば、格子中のあらかじめ

50

規定された部位にDNAを付着させることによって可能なものからそれら標的の充填効率を低下させることである。

【発明の開示】

【発明が解決しようとする課題】

【0005】

かかる制限に鑑み、標的ポリヌクレオチドのアレイから得られうる配列決定情報の量を上昇させるさらなるアプローチが利用可能となれば当該分野にとって有益であろう。当該技術分野におけるさらなる要求はマイクロミリより小さいサイズおよび距離にて数十億の結合部位を有するアレイ支持体を調製する効率的かつ安価な方法である。

【課題を解決するための手段】

【0006】

発明の概要

したがって、一つの側面において、本発明はラージスケール DNA 配列決定に対する多くのアプローチによって生じる短い配列読み取り長に関連する問題、例えば、酵素サイクル当たり限定された配列情報しか得られない問題に取り組むものである。また、数十億の分子、例えばマイクロミリより小さいサイズおよび距離の分子を支持することが出来る操作された核酸分子のランダムアレイを調製するための方法および組成物も提供される。

【0007】

一つの側面において、本発明は、標的配列の検出位置における第一のヌクレオチドの実体を決定する方法を提供し、ここで標的配列は複数の検出位置を含む。好ましい側面において、方法は二工程を含む：複数のコンカテマーを提供する工程および第一のヌクレオチドを同定する工程。各コンカテマーは複数のモノマーを含み、各モノマーは以下を含む：

(i) 標的検出位置の第一のセットを含む標的配列の第一の標的ドメイン；(ii) IIs型エンドヌクレアーゼ制限部位を含む第一のアダプター；(iii) 標的検出位置の第二のセットを含む標的配列の第二の標的ドメイン；および (iv) IIs型エンドヌクレアーゼ制限部位を含む第二の散在したアダプター。好ましい態様において、標的配列コンカテマーは表面に固定化されている。さらに好ましい態様において、表面は官能化されている。

【0008】

一つの態様において、本発明は、標的配列の検出位置における第一のヌクレオチドの実体を決定する方法を提供し、ここで、同定する工程は、コンカテマーと配列決定プローブのセットとを接触させることを含む。例示的な態様において、配列決定プローブはそれぞれ、アダプターの一つに相補的な第一のドメイン、第一の調査位置における特有のヌクレオチド、および標識を含む。好ましい態様において、コンカテマーと配列決定プローブとの接触は、特有のヌクレオチドが第一のヌクレオチドに相補的であれば、配列決定プローブがコンカテマーにハイブリダイズする条件下で達成され、それによって第一のヌクレオチドが同定される。

【0009】

別の態様において、各アダプターはアンカープローブ、ハイブリダイゼーション部位および同定する工程を含む。例示的な態様において同定する工程は以下を含む：アンカープローブをアンカープローブハイブリダイゼーション部位にハイブリダイズさせること、配列決定プローブをアダプターに隣接する標的検出位置にハイブリダイズさせること、隣接するハイブリダイズした配列決定プローブおよびアンカープローブをライゲーションしてライゲーションしたプローブを形成すること、およびライゲーションしたプローブを検出して第一のヌクレオチドを同定すること。

【0010】

別の態様において、各アダプターはアンカープローブハイブリダイゼーション部位を含み、同定する工程は、アンカープローブをアンカープローブハイブリダイゼーション部位にハイブリダイズさせること、およびポリメラーゼおよび標識を含む少なくとも1つのdNTPを添加することを含む。ポリメラーゼおよび少なくとも1つのdNTPは、dNTPが検出位置に完全に相補的であれば、dNTPがアンカープローブに付加されて伸長したプローブを形

10

20

30

40

50

成する条件下で添加され、それにより伸長したプローブの調査位置が生じる。第一のヌクレオチドは、伸長したプローブの調査位置におけるヌクレオチドを決定することによって同定される。

【0011】

本発明のさらに好ましい態様において、第二の検出位置におけるヌクレオチドが同定される。本発明のさらなる態様において、第三の検出位置、第四の検出位置、第五の検出位置、および/または第六の検出位置におけるヌクレオチドが同定される。

【0012】

一つの態様において、本発明は、標的配列の検出位置における第一のヌクレオチドの実体を決定する方法を提供し、ここで標的配列、標的配列コンカテマーは表面に固定化され、表面はこれらに限定されないが、アミン、シランおよびヒドロキシルを含む官能部分を含む。さらに好ましい態様において、表面は該固定化コンカテマーを含む複数の空間個別的な領域を含む。さらなる態様において、コンカテマーは、捕捉プローブを用いて表面に固定化される。

【0013】

一つの側面において、本発明は、複数の固定化コンカテマーを含む基体を提供し、該コンカテマーの各モノマーは以下を含む：第一の標的配列、IIs型エンドヌクレアーゼ制限部位を含む第一のアダプター、第二の標的配列、およびIIs型エンドヌクレアーゼ制限部位を含む第二の散在したアダプター。第一のアダプターのIIs型エンドヌクレアーゼ制限部位は第二のアダプターのIIs型エンドヌクレアーゼ制限部位と同一であっても同一でなくてもよい。さらに好ましい態様において、各モノマーはさらに第三の標的配列およびIIs型エンドヌクレアーゼ制限部位を含む第三の散在したアダプターを含み、さらなる態様において、各モノマーはさらに、第四の標的配列およびIIs型エンドヌクレアーゼ制限部位を含む第四の散在したアダプターを含む。

【0014】

別の側面において、本発明は、複数のアダプターを標的配列に挿入する方法を提供する。好ましい側面において、該方法は以下の工程を含む：(i) 第一のアダプターを該標的配列の一つの末端にライゲーションする工程、ここでアダプターは制限酵素のための結合部位を含む；工程 (i)からの産物を環状化して第一の環状ポリヌクレオチドを作成する工程；環状ポリヌクレオチドを制限酵素で切断する工程、ここで制限酵素は第一のアダプター内の結合部位に結合することが出来る；第二のアダプターをライゲーションする工程、ここで該第二のアダプターは制限酵素のための結合部位を含む；および、工程 (iv)からの産物を環状化して第二の環状ポリヌクレオチドを作成する工程。ある態様において、工程 (ii) ~ (v)を反復して所望の数のアダプターを標的配列に挿入する。好ましい態様において、環状化工程はCircLigase^(商標) 酵素を添加することを含む。

【0015】

別の態様において、環状化工程は、環状化配列を標的配列の第二の末端に付加すること、架橋テンプレートをアダプターの少なくとも一部および環状化配列の一部とハイブリダイズさせること、および第一および第二の末端をライゲーションして一緒にし、標的配列を環状化することを含む。

【0016】

別の側面において、本発明は、標的配列のヌクレオチド配列を同定する方法を提供する。この方法において、複数の散在したアダプターが標的配列内に提供され、各散在したアダプターは少なくとも一つの標的配列との境界を有する。少なくとも二つの散在したアダプターの少なくとも一つの境界に隣接する少なくとも一つのヌクレオチドが同定され、それによって標的配列のヌクレオチド配列が同定される。

【0017】

さらに別の側面において、本発明は、ポリヌクレオチドのライブラリーを提供する。好ましい側面において、ライブラリーは二以上の核酸断片を含み、各断片はあらかじめ決められた順序で複数の散在したアダプターを含む。各散在したアダプターは複数のその他の

10

20

30

40

50

散在したアダプターのその他の配列とクロスハイブリダイズすることができない配列を含む少なくとも一つの末端を有する。さらに好ましい側面において、散在したアダプターのあらかじめ決められた順序は各核酸断片について同一である。

【0018】

一つの側面において、本発明は標的ポリヌクレオチドのヌクレオチド配列を同定する方法を提供し、該方法は、標的ポリヌクレオチドの各複数の断片からアンプリコンを作成してアンプリコンのランダムアレイを形成する工程、一以上の配列決定プローブをランダムアレイにハイブリダイズさせる工程、配列特異的反応にて一以上の配列決定プローブを伸長させることにより少なくとも一つの散在したアダプターに隣接する少なくとも一つのヌクレオチドの実体を決定する工程、およびハイブリダイゼーションおよび同定する工程を標的ポリヌクレオチドのヌクレオチド配列が同定されるまで反復する工程を含む。好ましい側面において、配列決定プローブは、一以上のプローブと散在したアダプター上の相補的配列との間の完全に一致(match)した二本鎖の形成を可能とする条件下でランダムアレイにハイブリダイズされる。好ましい側面において、各断片はあらかじめ決められた位置に複数の散在したアダプターを含む。さらなる態様において、各アンプリコンは断片の複数コピーを含み、その数は、断片が標的ポリヌクレオチドを実質的にカバーするような数である。さらなる側面において、ランダムアレイのアンプリコンはアンプリコンの少なくとも大部分が光学的に分解できる密度にて表面に固定化される。

10

【0019】

別の側面において、本発明は標的配列のヌクレオチド配列を同定する方法を提供し、該方法は以下の工程を含む：コンカテマーのランダムアレイを提供する工程、プローブの第一のセットからの一以上のプローブをランダムアレイにハイブリダイズさせる工程、プローブの第二のセットからの一以上のプローブをランダムアレイにハイブリダイズさせる工程、近接する部位にて標的コンカテマーにハイブリダイズした第一および第二のセットからのプローブをライゲーションする工程、ライゲーションした第一および第二のプローブの配列を同定する工程、および、ハイブリダイズ工程、ライゲーション工程および同定する工程を標的配列の配列が同定されるまで反復する工程。好ましい側面において、コンカテマーのランダムアレイは光学的に分解可能な別々の空間的に離れた領域のアレイを有する平面状表面に固定化されたコンカテマーを含み、各コンカテマーは標的ポリヌクレオチドの断片の複数コピーを含み、その数は異なるコンカテマーの数がそれぞれ断片が標的配列を実質的にカバーするような数である。さらなる態様において、各別々の空間的に離れた領域は $1\mu\text{m}^2$ 未満の面積を有し、それにより実質的にすべての別々の空間的に離れた領域は多くとも1つの結合したコンカテマーを有する。

20

30

【0020】

さらに別の側面において、本発明は標的配列のヌクレオチド配列を同定する方法を提供し、該方法は、標的配列の断片の複数コピーを含む複数のコンカテマーを作成する工程、コンカテマーの少なくとも大部分が光学的に分解可能な密度にて表面に固定化されたコンカテマーのランダムアレイを形成する工程、および、少なくとも一つのコンカテマーにおける少なくとも一つの散在したアダプターに隣接する各断片の少なくとも一部の配列を同定し、それにより標的配列のヌクレオチド配列を同定する工程を含む。

40

【0021】

図面の簡単な説明

図1A-1Gは、本発明およびその適用を図示する。

【0022】

図2A-2Gは、散在したアダプターを含む標的ポリヌクレオチドを作るために核酸断片にアダプターを挿入する様々な方法を図示する。

【0023】

図3A-3Eは、散在したアダプターを含む標的ポリヌクレオチドに対して実行されうるハイスループット配列決定方法を図示する。

【0024】

50

図4は、RCR 産物の付着により作られる構造化されたDNA アレイと標準的ランダム DNA アレイとの比較を提供する。

【0025】

図5は、配列されたアレイ上の参照パターンを図示する。

【0026】

図6は、rSBH 装置上に画像化されたランダムアレイを示す。

【0027】

図7は、より容易にみるためにわずかにシフトさせて重ね合わせた3つのアレイ画像を示す。

【0028】

図8は、わずかにシフトさせて重ね合わせた5つのアレイ画像を示す。

【0029】

図9は、わずかにシフトさせて重ね合わせた5つのアレイ画像を示す。

【0030】

図10は、カバーガラスの表面を横切る線状に並んだ捕捉プローブがDNBへの特異的結合に用いられたアレイの画像を示す。

【0031】

発明の詳細な説明

本発明の実施には、特に断りの無い限り、有機化学、ポリマー技術、分子生物学（組換え技術を含む）、細胞生物学、生化学および免疫学の常套の技術および記載を用いることが出来、それらは当業者の技術範囲内である。かかる常套の技術には、ポリマーアレイ合成、ハイブリダイゼーション、ライゲーション、および標識を用いたハイブリダイゼーションの検出が含まれる。適切な技術の具体的説明は以下の例を参照することにより得られる。しかし、その他の同等の常套の手順ももちろん用いることが出来る。かかる常套の技術および記載は、標準的研究室マニュアル、例えば、Genome Analysis: A Laboratory Manual Series (Vols. I-IV)、Using Antibodies: A Laboratory Manual、Cells: A Laboratory Manual、PCR Primer: A Laboratory Manual、and Molecular Cloning: A Laboratory Manual（すべてCold Spring Harbor Laboratory Press）、Stryer, L. (1995) Biochemistry (4th Ed.) Freeman, New York、Gait, "Oligonucleotide Synthesis: A Practical Approach" 1984、IRL Press, London、Nelson and Cox (2000)、Lehninger, Principles of Biochemistry 3rd Ed., W. H. Freeman Pub., New York, N.Y. およびBerg et al. (2002) Biochemistry, 5th Ed., W. H. Freeman Pub., New York, N.Y.にみることで、それらのすべてはあらゆる目的で引用によりその全体を本明細書に含める。

【0032】

概観

本発明は、標的配列(本明細書において「標的ポリヌクレオチド」とも称する)のヌクレオチド配列情報を標的ポリヌクレオチド中に散在したアダプターを用いて獲得するための方法および組成物に関する。配列情報は新規なものであってよく、例えば、未知核酸の配列決定であってよく、あるいは再配列決定、または遺伝子型同定であってよい。本発明は好ましくは、標的ポリヌクレオチドまたはポリヌクレオチドの断片内の分散した位置に複数のアダプターを挿入する方法を含む。かかるアダプターは本明細書において「散在したアダプター」と称され、様々な配列決定化学、例えば、プライマー伸長、プローブライゲーション等によりヌクレオチドを同定する方法を用いて隣接配列を調べるためのプラットフォームとして役立つ。即ち、本発明のいくつかの態様の一つの特有の構成成分は既知アダプター配列の標的配列への挿入であり、アダプターにより近接する標的配列に途切れが生じる。アダプターの「上流」および「下流」の両方の配列決定により完全な標的配列の配列情報が達成されうる。

【0033】

したがって、限定するものではないが、本発明は以下のように概説できる(ゲノム DNA が本明細書において例として用いられているが、これに限定する意図ではないことを注意

10

20

30

40

50

されたい)。あらゆる生物からのゲノム DNAが単離され、標準的技術を用いて標的配列へと断片化される。第一のアダプターは標的配列の一つの末端にライゲーションされる。アダプターは好ましくはIIs型制限エンドヌクレアーゼ部位を含み、これは認識配列の外側を切断する。酵素の結果「粘着」末端が生じると、オーバーハング部分は埋めても切断してもよい。

【0034】

一つの態様において、酵素を用いてアダプターおよび標的配列を含む直鎖状鎖の2つの末端をライゲーションして環状核酸を形成する。これは単一の工程の使用により行うことができる。あるいは、第二のアダプターを標的配列の別の末端(例えば、ポリAテイル)に付加してもよく、次いで架橋配列を2つのアダプターにハイブリダイズさせ、次いでライゲーションしてもよい。いずれの態様においても、環状配列が形成される。

10

【0035】

環状配列は次いで IIs型エンドヌクレアーゼにより切断される結果、直鎖状鎖が生じ、このプロセスが反復される。この結果、以前に近接していた標的配列内のよく規定された位置にアダプターが散在した環状配列が得られる。

【0036】

環状配列はローリングサークル複製(RCR)反応を用いて増幅され、元の標的配列のコンカテマー(例えばモノマーの多量体)が形成される。かかる長いコンカテマーは「DNA ナノボール」(「DNB」)を形成し、所望により以下に記載するような様々な方法で表面に固定化することが出来る。

20

【0037】

表面上で、既知のアダプター配列を用いて、介在する標的配列の配列決定が行われる。当該技術分野において知られているように、標的核酸の特定の位置の塩基の実体を検出または判定するために利用可能な多数の技術が存在し、例えば、これらに限定されないが、温度の使用、完全および不完全プローブの標的配列への競合的ハイブリダイゼーション、合成による配列決定、例えば、一塩基伸長技術を用いるもの(「ミニ配列決定」と称されることもある)、オリゴヌクレオチドリガーゼ増幅(OLA)反応、ローリングサークル複製(RCR)、対立遺伝子PCR、競合ハイブリダイゼーションおよびInvader(商標)技術が挙げられる。好ましい態様は、ライゲーションを伴うハイブリダイゼーションによる配列決定、およびハイブリダイゼーションによる配列決定を含む。

30

【0038】

配列情報はより大きい標的配列の配列の再構築、例えば、全ゲノムDNAの配列決定に利用できる。

【0039】

例えば、ゲノム分析、疫学的研究、および診断試験のような用途に必要な、多数の核酸の配列決定は一般に配列決定技術をハイスループット形式に適合させることを含む。しかし、伝統的なハイスループット配列決定技術には欠点があり、特に配列読み取り長さが短いという問題があり、言い換えると、多くのハイスループット配列決定アプローチには配列決定が成功しうる標的ポリヌクレオチドの長さやタイプに制限がある。この制限は主に一つの操作で一つの断片に対して決定すべき連続塩基の数に起因する。特定の配列決定化学を実施する、各標的ポリヌクレオチドまたは断片において複数の部位を設けることにより、本発明は隣接配列の読み取りの多重性を提供する。一つの側面において、これらの隣接する読み取りは連続であり、それゆえ大きなクラスの配列決定化学の期待される読み取り長が効率的に増幅される。

40

【0040】

本発明はしたがってより長い近接する、またはほとんど近接する標的配列の決定を、アダプターのそれぞれの側の配列を決定することにより可能とする。

【0041】

標的ポリヌクレオチドの組成/構造

したがって、本発明は、サンプルからの標的配列を利用する組成物および方法を提供す

50

る。当業者に理解されるように、サンプル溶液はあらゆる数のものを含んでいてもよく、例えば、これらに限定されないが、体液（例えば、これらに限定されないが、血液、尿、血清、リンパ液、唾液、肛門および膣分泌物、汗および精液）および実質的にあらゆる生物の細胞、哺乳類サンプルが好ましく、ヒトサンプルが特に好ましい；環境サンプル（例えば、これらに限定されないが、空気、農業、水系および土壌サンプル）；細菌兵器病原体サンプル；研究用サンプル（即ち核酸の場合、サンプルは増幅反応の産物であってもよく、標的およびシグナル増幅、例えば、PCR 増幅反応の両方が含まれる；精製サンプル、例えば、精製ゲノム DNA、RNA調製物、生のサンプル（細菌、ウイルス、ゲノム DNAなど）；が含まれ、当業者に理解されるように、実質的にあらゆる実験操作がサンプルに対して行われていてもよい。

10

【0042】

一般に、標的生物（動物、鳥類、哺乳類など）からの細胞が用いられる。ゲノム DNA が用いられる場合、本発明のアレイの構築に必要なゲノム DNAの量は大幅に変動しうる。一つの側面において、哺乳類の大きさのゲノムについて、断片は少なくとも約 10 ゲノムに匹敵する量のDNAから作られる；別の側面において、断片は少なくとも約 30 ゲノムに匹敵する量のDNAから作られる；別の側面において、断片は少なくとも約 60 ゲノムに匹敵する量のDNAから作られる。

【0043】

標的配列または標的ポリヌクレオチドは核酸である。「核酸」または「オリゴヌクレオチド」またはその文法的に同等な語は、本明細書において互いに共有結合した少なくとも二つのヌクレオチドを意味する。本発明の核酸は一般的にホスホジエステル結合を含むが、場合によっては、以下に記載するように（例えばプライマーおよびプローブ、例えば、標識プローブの構築において）、異なるバックボーンを有する核酸アナログも含まれ、例えば、ホスホラミド (Beaucage et al., *Tetrahedron* 49(10):1925 (1993) およびそのなかの引用文献；Letsinger, J. *Org. Chem.* 35:3800 (1970)；Sprinzl et al., *Eur. J. Biochem.* 81:579 (1977)；Letsinger et al., *Nucl. Acids Res.* 14:3487 (1986)；Sawai et al., *Chem. Lett.* 805 (1984)、Letsinger et al., *J. Am. Chem. Soc.* 110:4470 (1988)；および Pauwels et al., *Chemica Scripta* 26:141 (1986))、ホスホロチオエート (Mag et al., *Nucleic Acids Res.* 19:1437 (1991)；および米国特許5,644,048)、ホスホロジチオエート (Briu et al., *J. Am. Chem. Soc.* 111:2321 (1989)、O-メチルホスホロアミダイト結合 (Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press参照)、およびペプチド核酸バックボーンおよび結合(Egholm, *J. Am. Chem. Soc.* 114:1895 (1992)；Meier et al., *Chem. Int. Ed. Engl.* 31:1008 (1992)；Nielsen, *Nature*, 365:566 (1993)；Carlsson et al., *Nature* 380:207 (1996)を参照、それらのすべては引用により本明細書に含まれる)を含むものが挙げられる。その他のアナログ核酸には二環式構造のもの、例えば、ロックされた核酸 (Koshkin et al., *J. Am. Chem. Soc.* 120:13252 3 (1998))；ポジティブバックボーン (Denpcy et al., *Proc. Natl. Acad. Sci. USA* 92:6097 (1995))；非イオン性バックボーン (米国特許 5,386,023、5,637,684、5,602,240、5,216,141 および4,469,863；Kiedrowski et al., *Angew. Chem. Intl. Ed. English* 30:423 (1991)；Letsinger et al., *J. Am. Chem. Soc.* 110:4470 (1988)；Letsinger et al., *Nucleoside & Nucleotide* 13:1597 (1994)；Chapters 2 and 3, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y. S. Sanghui and P. Dan Cook；Mesmaeker et al., *Bioorganic & Medicinal Chem. Lett.* 4:395 (1994)；Jeffs et al., *J. Biomolecular NMR* 34:17 (1994)；*Tetrahedron Lett.* 37:743 (1996))および非リボースバックボーン、例えば、米国特許 5,235,033 および5,034,506、および Chapters 6 and 7, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y. S. Sanghui and P. Dan Cookに記載のものが挙げられる。一以上の炭素環の糖を含む核酸も核酸の定義に含まれる (Jenkins et al., *Chem. Soc. Rev.* (1995) pp 169 176参照)。いくつかの核酸アナログがRawls、*C & E News Jun. 2, 1997 page 35*に記載されている。これら参考文献のすべてを引用

20

30

40

50

により本明細書に明示的に含める。リボース-リン酸バックボーンのこれらの改変は生理的環境においてかかる分子の安定性および半減期を上昇させるために行いうる。例えば、PNA:DNA ハイブリッドはより高い安定性を示すことが出来、それゆえある態様において利用されうる。

【0044】

核酸は示すように一本鎖でも二本鎖でもよく、あるいは二本鎖または一本鎖の配列の部分の両方を含むものでもよい。核酸は、ゲノムおよびcDNAの両方のDNAであってもよく、RNAまたはハイブリッドであってもよく、ここで核酸はデオキシリボ-およびリボ-ヌクレオチドのあらゆる組合せ、塩基のあらゆる組合せ、例えば、ウラシル、アデニン、チミン、シトシン、グアニン、イノシン、キサンチン、ヒポキサンチン、イソシトシン、イソグアニン等の組合せを含む。

10

【0045】

「標的配列」または「標的核酸」という用語または文法的に同等の用語は本明細書において核酸の一本鎖の上の核酸配列を意味する。標的配列は、遺伝子、調節配列、ゲノムDNA、cDNA、RNA、例えば、mRNA およびrRNAまたはその他の一部であってもよい。本明細書に記載するように、標的配列はサンプルからの標的配列でもよいし、二次的標的、例えば、増幅反応等の産物であってもよい。それはいずれの長さのものでもよい。

【0046】

以下により詳細に記載するように、プローブはサンプル中の標的配列の存在または非存在を判定するために標的配列にハイブリダイズように設計される。一般に、この用語は当業者に理解されている。標的配列は異なる標的ドメインから構成されるものでもよい；例えば、サンプル標的配列の第一の標的ドメインは捕捉プローブにハイブリダイズし得、第二の標的ドメインは標識プローブにハイブリダイズし得るなどである。標的ドメインは示されるように隣接していても離れていてもよい。断りのない場合、「第一」および「第二」という語は標的配列の5'-3'配向に関して配列の方向性を付与する意味ではない。例えば、相補的標的配列の5'-3'配向を考慮して、第一の標的ドメインは第二のドメインの5'または第二のドメインの3'のいずれに位置していてもよい。

20

【0047】

一つの態様において、ゲノムDNA、特にヒトゲノムDNAが用いられる。ゲノムDNAは、例えば、Sambrook et al., 前掲、1999; Current Protocols in Molecular Biology, Ausubel et al., eds. (John Wiley and Sons, Inc., NY, 1999)、等の開示の常套の技術を用いて得られる。ゲノムDNAを単離するために重要な因子には以下が含まれる：1) DNAはDNAプロセッシング酵素および汚染する塩を含まない；2) 完全なゲノムが同等に表されている；および3) DNA断片は約5,000~100,000 bpの長さである。

30

【0048】

多くの場合、抽出したDNAの消化は必要ではない。というのは溶解および抽出の間に生じる剪断力により所望の範囲の断片が生じるからである。別の態様において、より短い断片(1-5 kb)を制限エンドヌクレアーゼを用いる酵素的断片化により作ることが出来る。一つの態様において、10-100 ゲノム相当量のDNAが断片の集団によりゲノム全体をカバーすることを保証する。ある場合には、キャリアDNA、例えば非関連環状合成二本鎖DNAをサンプルDNAと混合して用いるために提供するのが有利であり、かかる場合としては、少量のサンプルDNAしか手に入らない場合および、例えば容器の壁への非特異的結合等による損失の危険がある場合がある。一つの態様において、DNAは一本鎖断片を作るために断片化の後に変性される。

40

【0049】

標的ポリヌクレオチドはソース核酸、例えば、ゲノムDNAから、特定のサイズの断片を作るための断片化により作ることが出来る；一つの態様において、断片は50~600ヌクレオチド長である。別の態様において、断片は300~600または200~2000ヌクレオチド長である。さらに別の態様において、断片は10-100、50-100、50-300、100-200、200-300、50-400、100-400、200-400、400-500、400-600、500-600、50-1000、100-1000、200-

50

1000、300-1000、400-1000、500-1000、600-1000、700-1000、700-900、700-800、800-1000、900-1000、1500-2000、1750-2000、および50-2000ヌクレオチド長である。これらの断片はRCR反応、あるいはその他の生化学的プロセス、例えば、さらなるアダプターの挿入に用いるために環状化されうる。

【0050】

本発明のポリヌクレオチドは継続的または同時に複数の部位からの配列情報の獲得を可能とする散在したアダプターを有する。散在したアダプターは標的ポリヌクレオチドの内部領域中の分散した位置に挿入されたオリゴヌクレオチドである。一つの側面において、標的ポリヌクレオチドに言及する場合の「内部」とはプロセッシング、例えば、環状化および切断の前の標的ポリヌクレオチドの内部の部位を意味し、それは標的ポリヌクレオチド内のヌクレオチドの順序を破壊する配列逆位等の変換を導入し得る。

10

【0051】

一つの側面において、以下により詳細に説明するように、散在したアダプターは標的ポリヌクレオチドの近接領域内に間隔を開けて挿入される。ある場合において、かかる間隔はあらかじめ決められた長さであり、同一であってもそうでなくてもよい。別の場合において、散在したアダプターの間の空間は1から数ヌクレオチド(例えば1から15)、または1から数十ヌクレオチド(例えば、10~40)、または1から数百ヌクレオチド(例えば100~200)の正確さでのみ知られたものであり得る。好ましくは、各標的ポリヌクレオチド内の散在したアダプターの順序および数は知られている。本発明のある側面において、散在したアダプターは標的ポリヌクレオチドの末端に結合したアダプターと共に用いられる。

20

【0052】

一つの側面において、本発明は、標的ポリヌクレオチドまたは標的ポリヌクレオチドの断片の複数コピー(例えば、「複数モノマー」)を含むコンカテマーの形態における標的ポリヌクレオチドを提供する。常套の条件下(常套のDNAバッファー、例えばTE、SSC、SSPE、等、室温)でのDNAコンカテマーは直径約100~300nmを有する溶液中の球状体積をおよそ満たすランダムコイルを形成し、かかる直径はDNAのサイズおよびバッファー条件に当該技術分野において周知のように依存する。例えば、Edvinsson、"On the size and shape of Polymers and Polymer complexes," Dissertation 696 (University of Uppsala, 2002)。

30

【0053】

ランダムコイルポリマー、例えば、一本鎖DNAのサイズの一つの尺度は、ランダムコイル構造の直径の粗い尺度である末端間距離の二乗平均平方根である。本明細書において「ランダムコイル直径」と称するかかる直径は、例えば、Zetasizer Nano System (Malvern Instruments, UK)等の装置を用いて光散乱により測定することが出来る。本発明の高分子構造のさらなるサイズの尺度には、分子量、例えばダルトンにおけるもの、および総ポリマー長が含まれ、後者は分岐ポリマーの場合はすべてのその分岐の長さの和である。

【0054】

表面への結合の際に、結合化学、連結の密度、表面の性質等に応じて、一本鎖ポリヌクレオチドがランダムコイル立体配置におけるコンカテマーの直径におよそ同等である領域によって平均して結合される扁平な回転楕円体の体積を満たす。表面上で高分子構造のコンパクトな形態を保存することにより、プローブ、例えば、コンカテマーの成分に特異的に向けられた蛍光標識オリゴヌクレオチドによってより強いシグナルを生成させることが可能となる。

40

【0055】

ある態様において、ポリヌクレオチドのクラスが、異なるアンカープローブ結合部位を有するアダプターを提供することにより作ることができる。このタイプの「クラスターリング」によりポリヌクレオチドの配列情報の獲得における効率を上昇させることが出来る。

【0056】

断片化方法

50

有効なマッピングストラテジーが配列決定用途、例えば、複雑な二倍体ゲノムの配列決定、デノボ配列決定、およびゲノム混合物の配列決定に必要である。一つの態様において、階層的断片化手順がハプロタイプ情報の同定および二倍体ゲノムの親染色体のアセンブリのために提供される。かかる手順はタンパク質アレルの予測および短い読み取りをゲノム内の正しい位置にマッピングすることにも適用できる。かかる方法のさらなる用途は、複数の遺伝子の間に共有される DNA 配列の~100塩基内に起こる遺伝子ファミリー中の突然変異の正しい割当である。

【 0 0 5 7 】

図(1C-D) は本発明の一つの側面を図示し、ここでソース核酸 (1600) (一つまたは複数の標的ポリヌクレオチドであってもよいし、それらを含むものでもよい) が処理(1601)されて一本鎖断片 (1602)を形成し、それは好ましくは 50 ~ 600 ヌクレオチド、より好ましくは300 ~ 600 ヌクレオチドの範囲であり、次いでアダプターオリゴヌクレオチド (1604) にライゲーションされてアダプター-断片結合体 (1606)の集団を形成する。アダプター (1604) は通常最初のアダプターであり、それは配列中で元々隣接していた 2 つの配列を分けるという意味で、「散在」している必要はない。ソース核酸 (1600) は常套の技術を用いてサンプルから抽出されたゲノム DNAであってもよいし、常套の技術によって生じたcDNAまたは ゲノム ライブラリーまたは合成 DNA等であってもよい。処理 (1601) は通常、常套の技術による断片化、例えば、化学的断片化、酵素的断片化、または機械的断片化、次いで変性による一本鎖 DNA 断片の産生を伴う。

【 0 0 5 8 】

いずれの段階においても断片の作成において、断片はゲノム全体由来であってもゲノムの選択されたサブセット由来であってもよい。多くの技術がゲノムのサブセット由来の断片の単離または富化に利用でき、以下の参考文献に例証されており、これらは引用によりその全体を本明細書に含める: Kandpal et al (1990)、Nucleic Acids Research、18: 1 789-1795; Callow et al、米国特許出願公開 2005/0019776; Zabeau et al、米国特許 6, 045,994; Deugau et al、米国特許5,508,169; Sibson、米国特許 5,728,524; Guilfoyle et al、米国特許 5,994,068; Jones et al、米国特許出願公開 2005/0142577; Gullberg et al、米国特許出願公開 2005/0037356; Matsuzaki et al、米国特許出願公開 2004/00 67493; 等。

【 0 0 5 9 】

一つの態様において、溶解およびゲノム DNAの抽出の際の剪断力が所望の範囲の断片を生じる。制限エンドヌクレアーゼを利用する断片化方法も本発明に含まれる。

【 0 0 6 0 】

好ましい態様において、特に哺乳類のサイズのゲノムについて、断片化は少なくとも二段階で行われ、第一段階では約 100 キロベース(Kb)から約 250 キロベースのサイズ範囲の断片の集団を作成し、第二段階を各100-250 Kb 断片に別々に適用し、約 50 ~ 600 ヌクレオチドのサイズ範囲、より好ましくは約 300 ~ 600 ヌクレオチドのサイズ範囲の断片を、ランダムアレイのためのコンカテマーの作成のために生じさせる。本発明のある側面において、第一段階の断片化はかかる断片のあらかじめ決められたサブセット、例えばシグナル伝達経路のタンパク質をコードする遺伝子を含む断片等、を選択するために用い

【 0 0 6 1 】

一つの態様において、サンプル ゲノム DNAは、その内容全体を引用により本明細書に含める米国特許出願番号11/451,692に示す技術を用いて断片化される。この側面において、ゲノム DNAは30-300 kb サイズの断片として単離される。適切な希釈を介して、これら断片の小さいサブセットがランダムにマルチウェルプレートまたは類似の装備の別々のウェルに入れられる。例えば、96、384または1536 ウェルを有するプレートがこれら断片のサブセットのために利用できる。これらDNA アリコートを作る最適の方法は、自然に高分子量形態へと断片化する方法によりDNAを単離し、定量の後10-30 ゲノム相当量へと希釈し、次いで全調製物を384 ウェルに分けることである。これにより全ゲノム配列の表示が

提供され、DNA 単離を10-30 細胞に対して100 % 回収効率で行うことにより、すべての染色体領域が同じカバー度(coverage)で表示されることが保証される。アリコートをおの方法で提供することにより、染色体の同じ領域からの2つのオーバーラップする断片が同じプレートウェルに入る可能性が最小となる。10x カバー度にて表示される二倍体ゲノムについて、平均して20のオーバーラップする別々のウェルに分けられるべき断片が存在する。このサンプルを384 ウェルプレートに分配すると、各ウェルは平均して、1,562 断片を含む。標準的 384-ウェルプレートに384 画分を形成することにより、2つのオーバーラップする断片が同じウェルに入ってしまう確立はたった約1/400である。いくつかの一致する断片が同じウェルに入れられたとしても、各染色体領域からのその他のオーバーラップする断片は特有のマッピング情報を提供する。

10

【0062】

一つの態様において、長い断片の調製された群はさらに切断されて最終断片サイズ約 300 ~ 600 塩基とされる。群における各断片の十分な (例えば、10x) カバー度を得るために、各ウェルにおけるDNAはよく確立された全ゲノム増幅方法を用いて最終切断の前に増幅するとよい。

【0063】

1つのウェルからのすべての短い断片は次いでアレイとされ、1つの分離したユニットアレイ上またはより大きい連続するマトリックスの1つのセクションにおいて配列決定され得る。384 ユニットアレイの複合(composite)アレイがこれら断片の群の並行分析に理想的である。親染色体を表示する長い配列のアセンブリーにおいて、アルゴリズムが1つのユニットアレイにおいて検出される短い断片が、それぞれが1つの染色体の別々の部分を表示するより長い連続するセグメントの制限された数に属するという重要な情報を利用しうる。ほとんどすべての場合において、相同的染色体セグメントは異なるユニットアレイ上で分析されうる。長い (~100 Kb) 連続する最初のセグメントはテーリングパターンを形成し、以下に記載するように各親染色体を別々にアセンブリーする十分なマッピング情報を提供し、これは100 kbのDNA当たり約 100 多型部位であることに依存する。以下の例では点に対応するセグメントにおいて同一である100-1000 連続塩基を表す。

20

【化1】

ウェル 3	T.....C.....C.....G.....A.....	
ウェル 20	...C.....T.....T.....A.....G.....C...	
ウェル 157	T.....A.....G.....C.....A.....C...	
ウェル 258	...C.....C.....G.....A.....T.....G.....T.....	
ウェル 3 及び 258	親1の染色体1のアセンブル	: ...T.....C.....C.....G.....A.....T.....G.....T
ウェル 20 及び 157	親2の染色体1のアセンブル	: ...C.....T.....T.....A.....G.....C.....A.....C...

30

【0064】

一つの態様において、染色体分離手順において得られた単一の標的の増幅が全ゲノム増幅について当該技術分野に知られた方法を用いて達成される。好ましい態様において、10-100 倍増幅を生じる方法が用いられる。一つの態様において、これら手順は増幅すべき配列という点では識別しないが、そのかわりサンプル中のすべての配列を増幅する。かかる手順は全100 kb 断片のインタクトな増幅を必要とせず、より短い断片、例えば、1-10 kbの断片を利用できる。

40

【0065】

散在したアダプターの組成/構造

一つの側面において、散在したアダプターが標的ポリヌクレオチドの近接領域内に間隔をもって挿入される。散在したアダプターは長さにおいておおいに変動し得、それは望まれる機能的要素の数およびタイプに部分的には依存する。かかる機能的要素としては、これらに限定されないが、アンカー配列、捕捉プローブ配列に相補的な配列(例えば、表面への付着のため)、タギング配列、二次構造配列、標識プローブの結合/ハイブリダイゼーションのための配列、機能付与配列、プライマー結合部位、ヌクレアーゼ、例えば、ニッキング酵素、制限エンドヌクレアーゼ等のための認識部位が挙げられる。

50

【 0 0 6 6 】

一つの態様において、アダプターは当該技術分野において知られているように制限エンドヌクレアーゼ認識部位を含む。一つの態様において、かかる認識部位はニッキング酵素のためのものであってもよい。

【 0 0 6 7 】

一つの態様において、制限エンドヌクレアーゼ部位はIIs型制限エンドヌクレアーゼ部位である。IIs型エンドヌクレアーゼは一般に市販されており、当該技術分野に周知である。そのII型対応物と同様に、IIs型エンドヌクレアーゼは二本鎖ポリヌクレオチド配列内のヌクレオチド塩基対の特異的配列を認識する。その配列を認識すると、エンドヌクレアーゼはポリヌクレオチド配列を切断し、一般に配列の一方の鎖のオーバーハング、即ち、「粘着末端」を生じる。IIs型エンドヌクレアーゼはまた一般にその認識部位の外側を切断する；距離は認識部位から2~20ヌクレオチド離れている。ポリヌクレオチド配列の不明な部分内で切断が起こるので、それは本発明の方法の下で不明な配列を切断部位に捕捉することを可能とする。通常、IIs型制限エンドヌクレアーゼはその認識部位から少なくとも6ヌクレオチド(即ち、認識部位の末端と最も近い切断点の間のヌクレオチドの数)離れている切断部位を有するように選択される。例示的なIIs型制限エンドヌクレアーゼとしては、これらに限定されないが、Eco57M I、Mme I、Acu I、Bpm I、BceA I、Bbv I、BciV I、BpuE I、BseM II、BseR I、Bsg I、BsmF I、BtgZ I、Eci I、EcoP15 I、Eco57M I、Fok I、Hga I、Hph I、Mbo II、MnI I、SfaN I、TspDT I、TspDW I、Taq II、等が挙げられる。

【 0 0 6 8 】

ある態様において、各アダプターは同じIIs型制限エンドヌクレアーゼ部位を含む。代替的な態様において、異なるアダプターは異なる部位を含む。

【 0 0 6 9 】

一つの態様において、一以上のアダプターはアンカープローブハイブリダイゼーション部位を含む。以下に説明するように、アンカープローブは配列決定反応に用いられ、様々な形態をとりうる。一般に、アンカープローブハイブリダイゼーション部位の少なくとも一つの末端は標的配列とアダプターとの接合部である；即ち、配列決定反応は一般に標的配列の検出位置の直近のアンカープローブのハイブリダイゼーションに依存する。アンカーまたはプライマーは標的-アダプター接合部から左または右にシフトして1から約10またはそれ以上、好ましくは1から4の塩基を有するよう選択または設計され得る。本明細書において用いる場合、「検出位置」とは配列情報が望まれる標的配列の位置をいう。

【 0 0 7 0 】

多くの態様において、配列決定反応はアンカープローブの両方の末端に実行させることが出来る；したがって、ある態様において、アンカープローブハイブリダイゼーション部位は完全なアダプター配列を含む。あるいは、各アダプター内に2つのアンカープローブハイブリダイゼーション部位が存在してもよい；1つは標的配列の3'末端に隣接または近接し、1つは5'末端に隣接または近接する。当業者に理解されるように、アンカープローブの長さおよびアダプターの長さに応じて、2つのアンカープローブハイブリダイゼーション部位はアダプター内でオーバーラップしてもよく、それらは直接隣接していてもよく、あるいはそれらは介在する配列によって分離されていてよい。アンカープローブハイブリダイゼーション配列の長さはアッセイ条件に依存して変動する。

【 0 0 7 1 】

一つの態様において、一以上のアダプターはプライマー結合配列を含む。当該技術分野において知られているように、ポリメラーゼは一般に二本鎖核酸部分とともに一本鎖鋳型(例えばコンカテマー)を必要とする。本質的に、あらゆる二本鎖配列がポリメラーゼによって認識されるようにあらゆる配列はプライマー結合配列として作用し得、プライマーに結合する。一般に、プライマー結合配列は約3から約30ヌクレオチドの長さであり約15から約25が好ましい。プライマーオリゴヌクレオチドは通常6から25塩基の長さ

である。当業者に理解されるように、プライマー結合配列は別のアダプター配列のいずれの中に含まれていてもよい。

【 0 0 7 2 】

一つの態様において、一以上のアダプターは捕捉プローブ認識配列を含む。以下により詳細に説明するように、本発明の一つの態様はDNBを固定化するために基体表面上の捕捉プローブを利用する。この態様において、アダプターはドメインと捕捉プローブとのハイブリダイゼーションを可能とする一以上の 捕捉プローブに十分に相補的なドメインを含み、その結果表面上にDNBが固定化される。

【 0 0 7 3 】

一つの態様において、一以上のアダプターは二次構造配列を含む。例えば、コンカテマー内の複数の アダプターにおける回文構造配列は、アダプターの間のハイブリダイゼーションをもたらす(例えばコンカテマー中のコピーの間の分子内相互作用)、したがって DNA ナノボール (「DNB」)の三次元構造が「緊密化」する。これら回文構造配列単位は5、6、7、8、9、10 またはそれ以上のヌクレオチドの長さであってよく、様々な配列、例えば、特定の融解温度を提供するよう選択された配列であってよい。例えば、回文構造 AAA AAAATTTTTTTTは、隣接する2つユニットのレプリカの間で以下の形態にて14 塩基 dsDNA ハイブリッドを提供する。

AAAAAAATTTTTTTT
TTTTTTTAAAAAAA

【 0 0 7 4 】

一つの態様において、アダプターは標識プローブ結合配列を含む。ある態様において、例えば配列決定反応ではなく特定の配列の検出のために、標識 プローブをコンカテマーに添加して、特定の配列を検出することが出来る。標識 プローブは標識プローブ結合配列にハイブリダイズし、少なくとも一つの検出可能な標識を以下に説明するように含む。例えば、感染病原体、例えば、細菌またはウイルスの存在の検出はこのようにして行うことが出来る。

【 0 0 7 5 】

一つの態様において、アダプターはタギング配列を含む。この態様において、タギング配列は環状化標的配列、コンカテマー等を引き出すあるいは精製するために用いることが出来る。ある態様において、タギング配列はタグ付加サンプルの混合物中の標的配列の起源を同定するために利用されうる特有の核酸配列を含み得、あるいはリガンド結合対、例えば、ビオチン/ストレプトアビジン等の成分を含みうる。

【 0 0 7 6 】

一つの側面において、散在したアダプターのそれぞれは8から60 ヌクレオチドの範囲の長さを有する; 別の側面において、それらは8から32 ヌクレオチドの範囲の長さを有する; 別の側面において、それらは約 4 ~ 約 400 ヌクレオチド; 約 10 ~ 約 100 ヌクレオチド、約 400 ~ 約 4000 ヌクレオチド、約 10 ~ 約 80 ヌクレオチド、約 20 ~ 約 70 ヌクレオチド、約 30 ~ 約 60 ヌクレオチドおよび約 4 ~ 約 10 ヌクレオチドから選択される範囲の長さを有する。約 20 ~ 約 30 塩基の全長のアダプターを用いる態様がいくつかの態様で特に有用である。

【 0 0 7 7 】

標的ポリヌクレオチドに挿入される散在したアダプターの数は多数の因子に応じて大幅に変動し得、かかる因子としては、使用する配列決定/遺伝子型同定化学 (およびその読み取り長の能力)、特定のIIs型部位の切断部位の特定の長さ、各標的ポリヌクレオチド内で同定が望まれるヌクレオチドの数、挿入の間に増幅 工程が行われるか否か等が挙げられる。

【 0 0 7 8 】

一つの側面において、複数の散在したアダプターは標的ポリヌクレオチドの近接するセグメントにおける部位に挿入される; 標的ポリヌクレオチドの近接するセグメントにおける部位に挿入される散在したアダプターは2、3、4またはそれ以上を含みうる。あるい

10

20

30

40

50

は、標的ポリヌクレオチドに挿入される散在したアダプターの数の範囲は2 ~ 10; 2 ~ 4; 3 ~ 6; 3 ~ 4; および 4 ~ 6である。別の側面において、散在したアダプターは環状化操作において直接または間接的にライゲーションされて一緒になったより長いポリヌクレオチド(本明細書において「メイトペア」と称する)、例えば、0.4-4 Kbの長さのポリヌクレオチドセグメントの一方または両方に挿入され得る。一つの側面において、かかるポリヌクレオチドセグメントは4-400 (好ましくは 10-100) 塩基の長さであり得る。

【0079】

一般に、標的配列に結合した第一のアダプターは「散在」または「挿入」されていないことに注意されたい。即ち、第一のアダプターは一般に断片化された標的配列の一つの末端に結合しており、第二以降のアダプターは近接する標的配列内に散在している。

10

【0080】

一つの側面において、標的ポリヌクレオチドの群の各メンバーは同一のアンカープロンプ結合部位およびソース核酸からのDNA断片に結合したIIs型認識部位を有するアダプターを有する。別の態様において、ポリヌクレオチドのクラスが、異なるアンカープロンプ結合部位を有するアダプターを提供することによって作られうる。

【0081】

一つの側面において、アダプターは標的ポリヌクレオチドの近接領域内に間隔を開けて挿入され、ここで間隔はあらかじめ決められた長さを有する。これらあらかじめ決められた長さは同じであってもなくてもよい。ある態様において間隔の長さは約 1 ~ 200 ヌクレオチドの精度で知られたものであり、別の態様において約 1-15、10-40および100-200 ヌクレオチドである。

20

【0082】

本発明によると散在したアダプターは一本鎖でも二本鎖でもよい。

【0083】

一つの側面において、アダプターは標的ポリヌクレオチド内の分子内相互作用を起こさせる回文構造配列を含み、その結果「ナノボール」が生じる。

【0084】

複数のアダプターを挿入する方法

本発明の一つの側面は、散在したアダプターを有する標的ポリヌクレオチドを作る方法を提供し、これを図 (1A-1B)に模式的に示す。この方法において、標的ポリヌクレオチド (1002)が散在したアダプターであってもなくてもよいアダプター (1000)と一緒にされ、一本鎖でも二本鎖でもよい環 (1005)を形成する(1004)。標的ポリヌクレオチドは一般により大きい DNA片、例えば、染色体またはその他のゲノム DNAの断片化によって得られる。

30

【0085】

二本鎖 DNAを用いる場合、断片の末端は常套の技術を用いた「ポリッシュ」および任意のアダプターのライゲーションにより環状化のために調製されうる。例えば、常套のショットガン配列決定に用いられる方法等が使用できる。例えば、Bankier、Methods Mol. Biol.、167: 89-100 (2001); Roe、Methods Mol. Biol.、255: 171-185 (2004); 等。

【0086】

40

第二の散在したアダプターを挿入するための次の部位を作成するために、環 (1005)は典型的には少なくとも一時的に二本鎖とされる。アダプター (1000)は本発明のこの側面においてIIs型制限エンドヌクレアーゼの認識部位を含むよう設計され、それはその切断部位 (1006)が標的ポリヌクレオチドの内部、例えばアダプター (1000)の右側となるように配置され、それにより環 (1005)が開環(1008)する。好ましい態様において、散在したアダプターを挿入する方法は3' 突出鎖を切断後に残すIIs型制限エンドヌクレアーゼを用いる。正確さの劣る挿入のために、ニッキング酵素を用いてもよく、あるいは第一のアダプターの一本の鎖がライゲーションできないものであってもよく、それによって適当な距離で翻訳され得、ポリヌクレオチドの切断の開始に用いられるニックが生じる。

【0087】

50

ポリヌクレオチドが切断された後、散在したアダプター (1010) は常套の技術を用いて正しくライゲーションされ2つのアダプターを含む開いた環 (1012)が生じ、これはライゲーションにより閉環(1016)される。このプロセスが反復(1018)される: 標的ポリヌクレオチド (1002)に所望の数の散在したアダプター、例えば、3つが挿入 (1026)されるまでの切断、挿入、および閉環を図1Bに示す。散在したアダプターを含む最終的な環 (1024)は多数の方法で処理されて各散在したアダプターの少なくとも一つの境界に隣接する標的ポリヌクレオチドの部位での配列情報が得られる。

【0088】

典型的には、標的ポリヌクレオチドの配列は各散在したアダプターと標的ポリヌクレオチドとの間の境界 (例えば 1021)の一方または両方にてまたはそれに隣接して分析される。一つの側面において、最終的な環 (1024)、またはそのセグメントは、増幅されてアンプリコンを作成し得、それは選択された配列決定化学、例えば、ライゲーションに基づくものまたは合成による配列決定により分析される。一つの側面において、第一および最後の散在したアダプターは、散在したアダプターを含む最終的な環 (1024)の領域が環から切断(1038)され、その後ポリメラーゼ連鎖反応 (PCR)による増幅のためにアダプターがライゲーション(1040)されうるように選択されうる。環の切断はアダプター 1および3の外側の1または2の部位にて行うとよい。別の側面において、最終的な環 (1024)はローリングサークル複製 (RCR)によって直接アンプリコンの作成に用いられ得、これは以下により詳細に説明する。

【0089】

多くの異なる標的ポリヌクレオチドが並行して分析される用途のためには、散在したアダプターを有する標的ポリヌクレオチドをそれぞれ図(1C-1D)および図(1E-1G)に示すようにRCR またはエマルジョンPCR を用いて増幅するとよい。

【0090】

エマルジョンPCRにおいて、断片の混合物は、例えば Margulies et al、Nature、437: 376-380 (2005); Shendure et al (2005)、Science、309: 1728-1732; Berka et al、米国特許出願公開 2005/0079510; Church et al、PCT公開 WO 2005/082098; Nobile et al、米国特許出願公開 2005/0227264; Griffiths et al、米国特許 6,489,103; Tillett et al、PCT 公開 WO 03/106678; Kojima et al、Nucleic Acids Research、33 (17): e150 (2005); Dressman et al、Proc. Natl. Acad. Sci.、100: 8817-8822 (2003); Mitra et al、Anal. Biochem.、320: 55-65 (2003); Musyanovych et al、Biomacromolecules、6: 1824-1828 (2005); Li et al、Nature Methods、3: 95-97 (2006); 等に掲示のように増幅され得、これらはあらゆる目的のためにその内容を引用により本明細書に含める。

【0091】

簡単に説明すると、図(1E)に示すように、標的ポリヌクレオチドと散在したアダプターとを含むDNA 環 (1500)の単離の後、アダプターが、例えば図1Aに示されるように切り出され(1038)、切り出された配列の集団を形成し、それらはアダプター (1503)にライゲーションされる。アダプター付加された配列は、水油エマルジョン (1505)中で、切り出された配列の一方の末端にライゲーションされたアダプターに特異的なプライマー、切り出された配列の他方の末端にライゲーションされたアダプターに特異的なプライマーを結合して備えているビーズ、およびDNA ポリメラーゼと一緒にされる。条件は相当な数の (例えば 15-20パーセントを超える) の油(1506)中の水泡(1508)が単一のアダプター付加された配列 (1510)および少なくとも一つのビーズ(1512)を含むのを可能にするように選択される。泡(1508)中の水相はその他にはPCRの実施のための常套の 反応混合物を含み、その結果ビーズ(1518)はそれぞれ別々の アダプター付加された配列が結合したクローン集団を有することになる。

【0092】

本発明の一つの側面において、複数の散在したアダプターの単一のゲノム断片への導入が 以下を含む一連の工程を介して進行する: 1) IIs 制限酵素のための結合部位を有する最初の アダプター のライゲーションおよびDNA 環の閉環、次いで 2) プライマー伸長お

10

20

30

40

50

よびゲノム 配列の選択的制限切断による環の再開環；および 3) 第二のアダプターのライゲーションおよびDNA 環の開環。工程2および 3 を 第三のアダプターが ゲノム 配列に組み込まれるよう反復する (図2Bおよび2C)。第二のアダプターは第一のアダプターと同じ制限部位を利用してよく、それによってゲノム DNAの内部でのゲノム セグメントの切断が最小化される。一つの態様において、第二の アダプターの認識部位を用いるが第一のアダプターの認識部位は用いない制御された切断は第一のアダプター 制限部位における切断を当該技術分野に知られた技術により遮断することにより達成され、かかる技術としては例えば、 第二の部位での切断に先立つ第一の 制限部位のメチル化が挙げられる。

【0093】

異なる結合部位を有するアダプターをサンプルの2つのアリコートと共に用いて特定のゲノム断片の排除を防ぐことが出来る。一つの態様において、最後のアダプターの配列の一部をRCR プライミング部位として用い、アダプターの別の部分をガラス 表面に結合させたアンカー オリゴヌクレオチドのための結合部位として用いる。

【0094】

本発明の一つの側面において、アダプターをゲノム 断片に挿入する方法は第一のアダプターのライゲーション、次いで環形成によって始まる。100 ~ 300 (または300-600) 塩基の長さのゲノム 断片がDNase 断片化によって調製され得、 ライゲーションに好適な 5-プライムリン酸および 3-プライム OH基が生じる。高度に複雑なゲノム DNAは加熱(変性) および迅速な冷却により一本鎖(ss) DNAとして調製することが出来る。DNAは高度に複雑であるため、あらゆる断片のための相補的配列の局所濃度は無視することが出来、DNAがほとんど一本鎖状態である場合、次の手順を行うのに十分な時間が確保できる。ssDNAの使用は環形成を顕著に簡便にする。というのは各ssDNA 断片の5' および3' 末端が別々の極性を示すためである。第一段階はアダプター配列の各一本鎖ゲノム断片の末端へのライゲーションである。すべての可能な配列組合せがゲノム DNA中に表されているので、アダプターはすべての可能な配列について合成された架橋鑄型 分子の助けにより一方の末端にライゲーションできる(図2B)。ゲノム DNAと比較してこれらオリゴヌクレオチドは比較的高濃度であり得るため、ゲノム 断片の末端に相補的なオリゴヌクレオチド(または不一致(mismatch)を有する相補鎖)はハイブリダイズしうる。架橋はしたがってライゲーション部位に形成されて一本鎖ゲノム断片の5-プライム末端 のアダプターへのライゲーションを可能とする。一つの態様において、この構造編成はアダプターの断片の3-プライム末端へのライゲーションを可能としない。

【0095】

図2Bにおいて、別の例示的方法を複数の散在したアダプターのDNA 環への組込みについて図示する。かかる方法は以下の工程を含む: 1. アダプター (230) の一本鎖DNA (232)の5' および3' 末端へのライゲーション(アダプターは縮重 (6-9 塩基) 架橋テンプレート (234)を有する)、次いで3-塩基オーバーハング(236)を介するアダプターのライゲーション; 2. アダプター オリゴヌクレオチドからのポリメラーゼによる伸長 (238)によりIIs型制限酵素切断のための二本鎖 DNAを作る; 3. IIs型認識部位 (240) の12-16 塩基下流にて切断(242)し、環を開環する; 4. 加熱の結果、新しい鎖 (243)が失われる;そして 5. 断片に別のアダプター (230) が挿入でき、再び環を開環する。

【0096】

3'末端の環への捕捉には縮重塩基を用いて架橋構造がライゲーション部位のうえに形成されるよう調製されたオリゴヌクレオチド鑄型の使用が必要である。ゲノム 断片の3'末端の第二のアダプター セクションを使用して5'末端に結合したアダプターの末端に相補的な3-塩基 オーバーハングにより環を開環する。鑄型架橋のハイブリダイゼーションが起こる(しかし3 塩基 オーバーハングではない)のに好ましい温度でこのアダプター セグメントの結合を実施することにより、過剰の架橋分子がバッファー交換により除去できる。というのはゲノム/アダプター 分子は固体支持体に結合しているためである。3-塩基オーバーハングは環形成に十分であるが、温度が低下するまではそうではない。縮重塩基

10

20

30

40

50

を有する2つの架橋オリゴヌクレオチドの使用はゲノム DNAの多様な配列末端により生じるアーティファクトを排除しうる。好ましい態様において、両方の架橋オリゴヌクレオチドは互いに独立に結合し、縮重 オリゴヌクレオチドがその相補的配列に結合する自由を確保する。両方のアダプター成分が同じライゲーション反応においてそれぞれのDNA末端にライゲーションされ得、ライゲーション アーティファクトはブロックされた末端を有する架橋鋳型オリゴヌクレオチドの設計によりさらに防止されうる。

【0097】

捕捉機構、例えば、ビオチン/ストレプトアビジンの非-環 アダプター鎖への組込みが下流 クリーンアッププロセスにおいて利用できる。かかる態様において、ライゲーションしていない、およびライゲーションしたビオチン化アダプターの両方が存在するため、ライゲーションしていない過剰のアダプターは~200 塩基の長さであるアダプター-ゲノム断片のサイズ選択 により除去できる。アダプター-ゲノム 断片は次のクリーニング手順のためにストレプトアビジン被覆ビーズに結合され得る。さらなる選択肢は1つのライゲーションしたアダプターの部分に相補的な(おそらく PNAまたは LNAを組み込んでいる)捕捉オリゴヌクレオチドを有するビーズの使用である。第一のアダプター/鋳型のあらかじめアセンブリーした左側を有するビーズをプロセスのさらなる簡便化のために用いることが出来る。

【0098】

図2Cにおいて、散在したアダプターを導入する別の例示的な方法が図示される。その方法は以下の工程を含む: (1) 2つのアダプター セグメント(250 および252) を一本鎖DNA断片(254)に縮重塩基 (例えば、セグメント 256 および 258は 7の縮重塩基の使用を示すが、8の縮重塩基も使用できる)を含む鋳型オリゴヌクレオチド (250の二本鎖 セグメントは約 10 塩基長であり得、252の二本鎖 セグメントは8 -10 塩基長であり得る)を用いてライゲーションする。鋳型 オリゴヌクレオチドの両末端(250および256)を3'末端でのジデオキシ終止および5'末端でのOH-基またはビオチンによりライゲーションからブロックする。アダプター/鋳型 ハイブリッドを非常に高濃度、例えば、 $1\mu\text{M}$ およびゲノム DNAの濃度の1000倍過剰で用いる。(2) DNAを3' アダプター (250)の5' 末端の上のビオチンを介してストレプトアビジン 支持体 (260) 上に収集する。過剰の遊離の5' アダプターを上清から除く。(3) DNA を温度上昇によりストレプトアビジン 支持体から遊離させ、上清を収集する。(4) DNAをジデオキシ終止により3'末端がブロックされた長い捕捉オリゴヌクレオチド (262)を用いて固体支持体に再び捕捉する。後の手順において過剰の遊離のアダプターの除去を促進するために固体支持体にDNAを強く結合させるためにオリゴヌクレオチドはペプチド 核酸 (PNA)の形態であり得る。捕捉オリゴヌクレオチド (262)は安定性を上昇させるためにゲノム 部分に結合するための 1-10 縮重塩基の5' 末端 (264)での付加により伸長しうる。(5)架橋テンプレート (266、14-18 塩基長であり得る)を用いてアダプターの2つの末端を一緒にしてDNA 分子を 環状化する。それは5' 末端をアミド基でブロックされるが、3' -OH基は続く後での工程におけるDNA ポリメラーゼによる伸長に使用できる。キナーゼおよびリガーゼを5' アダプターの5' 末端のリン酸化反応およびDNA 分子の2つの末端のライゲーションのために提供する。

【0099】

複数のアダプターを挿入する別の例示的な捕捉手順において、2つのアダプターセグメントが縮重鋳型を用いてゲノム ssDNA 断片にライゲーションされる(図2C)。ゲノム DNAの5'末端 にライゲーションするアダプター セグメントの3' 末端はブロッキング相補鎖を有する。3' アダプター セグメントのための鋳型はビオチンを有する。アダプター/鋳型は非常に高濃度、例えば、 $1\mu\text{M}$ でありゲノム DNAよりも~1000x倍高い濃度を有する。DNAはストレプトアビジン 支持体上に収集され、溶液は過剰のアダプター成分とともに除去される。ゲノム DNA は温度上昇により遊離し、DNA 溶液が収集される。DNA はその他のすべての合成 DNAの除去とともに5'末端 アダプター セグメントに相補的な長いオリゴヌクレオチド (ブロックされた末端を有する)を有する第二の 固体支持体に再び収集される。架橋鋳型が添加され、プライマーとしても作用する。キナーゼおよびリガーゼ (お

よびポリメラーゼ)が添加されて環を閉環し、プライマーを約 30 塩基伸長する。伸長は時間またはddNTPの存在により制御される。酵素は熱不活性化され、DNAは次いで IIs型制限酵素により切断される。短い二本鎖部分は温度上昇により除かれ環が結合したオリゴヌクレオチドへの強いハイブリッドを介して固体支持体に結合する。このより強いハイブリッドはオリゴヌクレオチドにLNAまたはPNA 塩基を組み込むことにより維持される。第二のアダプターについての鑄型を有する2つのアダプター セグメントが添加され(上記と同じ設計)さらなる固体支持体結合は必要ではない。というのは環 DNAは 連続的に固体支持体にさらなる工程のために結合しているからである。温度上昇を使用して環状 DNAに結合した鑄型を除く。この工程を第三のアダプターの挿入まで反復する。さらなるアダプターが挿入されない場合、ポリメラーゼは添加されず、バッファー 交換の後、DNAが温度上昇によりRCR 反応のために遊離する。

10

【0100】

散在したアダプターを挿入する別の例示的な方法を図2Dに図示する。この方法では散在したアダプターに隣接するあらかじめ決められた長さの標的ポリヌクレオチドのセグメントを作成する。あらかじめ決められた長さは散在したアダプター内でIIs型制限エンドヌクレアーゼを選択し位置づけることにより選択される。この方法の一つの側面において、最初のアダプターから最後から2番目のアダプターまでの各異なる散在したアダプターは異なるIIs型制限エンドヌクレアーゼの認識部位を有する。二本鎖 DNA (dsDNA)が断片化されて傷ついた末端 (frayed ends) (269)を有する標的ポリヌクレオチド(270)を生じ、その後かかる末端は常套の技術を用いて修復され平滑末端を有する断片(271)が形成される。平滑末端断片(271)の3'末端に単一のヌクレオチド (273)例えば dAを、Taq ポリメラーゼまたはその他の酵素を用いて付加し、拡張した断片(272)を作る。拡張した断片(272)は 相補的 ヌクレオチド オーバーハング、例えば dTを有する散在したアダプター (274)と、リガーゼの存在下で一緒にされ、 複数のライゲーション 産物が形成され、それには単一の散在したアダプターおよび単一の断片を含む産物 (275)が含まれる。条件を産物 (275)の環状化 (276)を促進するよう調整し、dsDNA 環 (283)が形成される。その他の産物、例えば、散在したアダプターを両末端に有する結合体またはライゲーションしていない断片およびアダプターは、一般に環を形成する能力を有さず、一本鎖エキソヌクレアーゼでの消化により産物 (275)の環状化の後に除去できる。

20

【0101】

dsDNA 環(283)はアダプター (278)内の部位を認識するIIs型制限エンドヌクレアーゼで処理され、dsDNA 環(283)が切断されアダプター (278)に隣接する標的ポリヌクレオチド(270)のセグメント (277)が残る。この態様において、IIs型制限エンドヌクレアーゼによる切断は3' 目的末端を残し、これはDNA ポリメラーゼにより伸長されて平滑末端 (279)を形成し、その後この断片 (284) が単一ヌクレオチドをその3'末端に付加するよう上記のように処理される。断片 (284)に、相補的 オーバーハングを有する第二の 散在したアダプター (281)がライゲーションし、プロセスはさらなる散在したアダプターの組込みのため反復される。一つの態様において、散在したアダプター 組込みの各サイクルは所望の産物の増幅 工程を含み、それにより続くプロセッシング工程のために十分な材料が生じる。

30

40

【0102】

図2Eにおいて、標的ポリヌクレオチドのあらかじめ決められた位置に散在したアダプターを組み込むための別の例示的な方法を図示する。断片は図2Dにおけるように生じ、上記のようにIIs型認識部位を含む最初の 散在したアダプター (286)を有するdsDNA 環 (285)が生じ、dsDNA 環 (285) はあらかじめ決められた部位 (287)にて切断され、 3' オーバーハング (289)を有する断片 (288)が生じ、その長さは2よりも異なってもよい。断片 (288)の散在したアダプターはアダプターと断片の境界にニック (290)を含むか、あるいはそれはアダプター内部へのニック (291)の導入を可能とするニックングエンドヌクレアーゼのための認識部位を含む。いずれの場合においても、断片 (288)は DNA ポリメラーゼ (292)で処理され、それはニック (例えば291)から上側鎖を断片 (288)の下側鎖の末

50

端に伸長し得、一方の末端に3' オーバーハングを有し他方に平滑末端を有する断片を形成する。この断片に、縮重ヌクレオチド オーバーハングを一方の末端に有し、単一の3' ヌクレオチド (例えばdT) オーバーハングを他方の末端に有する散在したアダプター (294) をライゲーションし断片 (295)を形成し、これを処理して (例えば Taq ポリメラーゼにより) 3' dAをその平滑末端に付加し、断片 (296)を形成する。断片 (296)を次いで部位 (297)におけるライゲーションにより環状化しdsDNA 環 (298)を形成し、その他のライゲーション産物を上記のように消化する。このプロセスのさらなるサイクルを行ってさらなる散在したアダプターを同様に組み込むことができ、必要であれば、あるいは任意に増幅工程を各サイクルに加えることもできる。

【0103】

図2Fにおいて、散在したアダプターの間に様々な長さのセグメントを提供する、散在したアダプターを組み込むさらなる方法を図示する。即ち、散在したアダプターはあらかじめ決められた順序で組み込まれるが、間隔は正確には未知である。この方法は 既知の制限酵素により提供されるよりも長い距離でのアダプターの組込みを可能とする。上記のように、dsDNA 環(2000)がニッキング酵素のための認識部位 (2004)を含む最初のアダプター (2002) (これは散在したアダプターでもそうでなくてもよい)を有するよう調製される。ニック (2006)の作成後、dsDNA 環 (2000)はDNA ポリメラーゼ (2008)で処理され、これにより遊離の3'鎖が伸長され(2010)ニック部位 (2006)の遊離の5'末端を有する鎖が置換または分解される。反応はあらかじめ決められた 間隔の後で停止され、その間隔は数百塩基を超える合成に予測される時間よりも短くなるよう選択される。かかる伸長は様々な方法により停止させることが出来、例えば、反応条件、例えば、温度、塩濃度等の変化、用いるポリメラーゼの無効化などにより停止させることが出来る。これによりニックまたはその他のギャップ (2012)を有するdsDNA 環が残り、これはヌクレアーゼ活性を有する様々な酵素、例えば、DNA ポリメラーゼ、FEN-1 エンドヌクレアーゼ、S1 ヌクレアーゼ (2014)等により認識され切断され、これら酵素は単独で用いても組み合わせて用いてもよい。例えば、Lieber、BioEssays、19: 233-340 (1997)。ニックまたはギャップ (2012)での切断の後、標的ポリヌクレオチドの末端はショットガン配列決定に用いられる技術を用いて修復され得、その後、標的ポリヌクレオチド (2000)が交互のまたは粘着末端を残すIIs型制限エンドヌクレアーゼを用いて切断され (2017)、アダプター (2002)が残る。この平滑末端に、次の散在したアダプターが結合し、その後その結果得られたコンストラクトが常套の技術を用いて環状化され、さらに散在したアダプターが挿入される。一つの態様において、連続する散在したアダプター、例えば(2002)と (2018)との距離は、正確には未知であるが用いる切断酵素、用いるポリメラーゼ、合成に用いた時間間隔、合成の停止方法、反応条件、例えば、dNTP濃度等に依存する。

【0104】

一つの態様において、工程 (2010)において、ニックトランスレーションを鎖置換の代わりに用いることが出来る。一つの側面において、ポリヌクレオチド切断部 (2016)において、第二のアダプターを第一のアダプターに連結した側に単にライゲーションしてもよい。この方法は2つのセグメントが様々な長さ、例えば、(10-50) + (30-300) 塩基を有するメイトペア構造を作るための アダプターの逆側の第二の切断(2006)と組み合わせることが出来る。

【0105】

一つの側面において、本発明はCircLigase (商標) を用いてアダプターを挿入し、鋳型を用いずに一本鎖ポリヌクレオチド環を閉環する方法を提供する。この酵素はアダプターを単一のオリゴヌクレオチドとして使用する能力および1つのみの鋳型を用いる能力を提供する。この方法において、標準的リガーゼ、例えば、T4 DNA リガーゼを用いてアダプターが標的ポリヌクレオチドの5'末端にライゲーションされた後、過剰のアダプターおよび鋳型が除去される。CircLigase (商標) (およびアダプターが5'末端にてリン酸化されていなければキナーゼ)は、一本鎖ポリヌクレオチド環を閉環するのに利用できる。

【0106】

一つの態様において、最初のアダプターがポリヌクレオチドに挿入された後、一本鎖環を形成するためにはそれは支持体から遊離されなければならない場合がある。ポリヌクレオチドは支持体に再ハイブリダイズされうる；一つの態様において、この再ハイブリダイゼーションは支持体表面に結合した捕捉オリゴヌクレオチド上で起こる。プライマーが局所 dsDNA の作成のためのサイクルが終わった後、ポリメラーゼとともに添加され II S 型制限酵素による切断が可能となる：

【化 2】

GGGGGGGGGGGGGGG.UUUUUUUUUUUUUUUUUUUUUUUUUUUUU-5'OH 3'OH-GGGGGGGGGG....

10

【 0 1 0 7 】

複数のアダプターのライゲーションは、5' OHにより開始するかまたはまたはおそらくはヘアピンの形態をとる長いブロッキング鋳型を備えることにより阻止されうる:

【化 3】

$\text{[-N(N)N(N)N]UUCUU-}$ 固体
UU-P [-UUUUUUUU-]

ここで、U=共通塩基、N=縮重塩基、P= リン酸、G= 目的のゲノム または DNA。

【 0 1 0 8 】

一度環形成が起こると、アダプターに既にプレハイブリダイズしたプライマーがポリメラーゼにより伸長され、IIs型制限酵素切断のための十分な二本鎖 DNAが生じ、さらなるアダプターの正確な挿入が可能となる(図9)。ポリメラーゼ、例えば、Klenow が使用でき、ddNTPのレベルと共に伸長の長さを約 20-30 塩基に制御しうる。

20

【 0 1 0 9 】

2つのさらなるアダプターの挿入は本発明のある態様において各酵素工程が30分間未満で達成されるとすると2-3時間かかりうる。アダプター挿入プロセスに生じる孤発性のエラーは認容できる。というのは重複する数十のオーバーラップする配列が各塩基について生じており、プローブ-プローブデータがアダプター 挿入に供されない100 塩基を超える各DNA 断片について生じているからである。

【 0 1 1 0 】

一つの例示的方法において、50-100 塩基 +25 塩基 メイトペアを >1 Kb の距離にて有するdsDNA 環を調製することにより複数のアダプターを挿入することが出来る。この方法において、~1-3 Kb ゲノム 断片のdsDNA 環に A/T または平滑末端 ライゲーションを用いてアダプターが備えられる。一つの態様において、アダプターはニッキング酵素 結合部位を有するかまたはそれは1つのウラシルまたはその他の切断可能または光切断可能な塩基アナログまたは1つのライゲーションしていない3'末端および2種類のIIS 結合酵素のための認識部位を有する。

30

【 0 1 1 1 】

一つの態様において、DNA はニッキング酵素を用いてまたはウラシル部位にて切断され利用可能な3'末端が鎖置換酵素またはニクトランスレーション酵素により~75 塩基伸長される(あるいはアダプター ライゲーションがニックを残した場合にのみ伸長される); ライゲーションしていない 3' 部位を用いる場合、置換はアダプターを介するものであろう。例えば、長さは75 塩基+ アダプターの長さであろう。利用可能な3'末端はニクトランスレーションまたは鎖置換によるDNA 合成により除去できる。切断はニックまたは分岐構造にていくつかの酵素、例えば、一本鎖切断酵素、の一つによりなされうる。このプロセスの結果、最初のアダプターの一方の末端に隣接して30-110 塩基のdsDNA 断片が生じる。DNAは次いで長い切断距離を有するIIS型制限エンドヌクレアーゼにより切断されうる。一つの態様において、切断距離は18から 25 塩基である。環 はアダプター無しで閉環されうるし(ゲノム 断片の平滑末端 ライゲーション)あるいは第二のアダプターの方向性平滑末端 ライゲーションにより閉環されうる。両方のアダプターを異なるまたは同一

40

50

の酵素を用いたさらなるアダプターのさらなる挿入に利用できる。第一のアダプター部位が第二のアダプターの挿入の前にメチル化されている場合、第二のアダプターはアダプター末端から適切な距離に位置する同じ制限部位を利用することが出来、ゲノム DNA の特定の部位での切断が得られる。

【 0 1 1 2 】

環状化方法

様々な標準的 DNA 環形成手順を用いることが出来る。一例はアダプターの平滑末端ライゲーションである。このアプローチの問題点は複数の組み込まれるアダプターの配向およびライゲーションである。カセットの 1 つの鎖はライゲーションがブロックされた 5' および 3' 末端の両方を有しうる。カセットの配向はどちらの DNA 鎖が RCR を開始させる遊離の 3' 末端を有するかを決定する。これにより各鎖が約 50% の場合に複製されることになる。

【 化 4 】

```
DDDDDDDDDDXLLLLLLLLLLLLLXDDDDDDDDDDDD
DDDDDDDDDDOLLLLLLLLLLLLLODDDDDDDDDDDD
```

```
DDDDDDDDDDOLLLLLLLLLLLLLODDDDDDDDDDDD
DDDDDDDDDDXLLLLLLLLLLLLLXDDDDDDDDDDDD
```

D=DNA、L= アダプター、X= ブロックされたライゲーション部位、O=ライゲーションのために開かれている。

【 0 1 1 3 】

当業者に理解されるように、環状化したアダプター/標的配列成分を形成するいくつかの方法がある。一つの態様において、CircLigase (商 標) 酵素を使用して鋳型を用いずに一本鎖ポリヌクレオチド環を閉環する。あるいは、直鎖状鎖の 2 つの末端に相補的な架橋鋳型が用いられる。ある態様において、第一のアダプターの標的配列の 1 つの末端への付加は架橋鋳型の相補的部分の設計に用いられる。他方の末端はすべてのゲノム配列に結合するための縮重塩基を含むユニバーサル 鋳型 DNA であってよい。2 つの末端のハイブリダイゼーション、次いで、ライゲーションの結果、環状化した成分が得られる。あるいは、標的分子の 3' 末端をターミナルトランスフェラーゼを用いてポリ-dA テイルを付加することによって修飾してもよい。修飾された標的は次いでアダプター およびオリゴ-dA テイルに相補的な架橋鋳型を用いて環状化される。

【 0 1 1 4 】

別の態様において、ピオチンがライゲーションをガイドするのに用いる各鋳型 オリゴヌクレオチドに組み込まれる。これは、例えば、形成した環を除去せずに鋳型を除去する高温融解の適用により、鋳型の容易な除去を可能とする。これらのより長いオリゴヌクレオチドはRCRのためのプライマーとして作用しうるし、あるいは他の目的、例えば、さらなるカセットの挿入に用いられ得る。

【 0 1 1 5 】

別の態様において、標的 DNA は固体支持体、例えば、磁気ビーズまたはチューブ/プレートウェルの壁、に結合され得、標的 DNA に共有結合的にライゲーションしていないあらゆる鋳型またはアダプターの除去を可能とする。標的 ssDNA は支持体を用いてランダムプライマーと結合され得、それは伸長を行い約 20-80 塩基の dsDNA を作る。伸長の長さは時間または ddNTP の量により制御されうる。別のアプローチはアダプターを ssDNA の一方の末端にライゲーションすることであり、次いで ssDNA にライゲーションしたアダプターにより DNA をサイズ選択し、同時に遊離のアダプターを除去する。この場合、アダプターの部分に相補的な約 10-50 塩基長のアンカー配列が支持体に結合され得、DNA を捕捉し、それを次の工程に用いる。このアンカー分子はハイブリッド安定性を高めるさらなる成分を有し得、例えば、ペプチド核酸の組込みが挙げられる。一本鎖 DNA を結合させるさらなる方法は支持体に結合した一本鎖 DNA 結合タンパク質の利用による。

【 0 1 1 6 】

環状化の一つの方法において、図2Aに図示されるように、ゲノム DNA (200)が断片化され、変性(202)された後、一本鎖DNA 断片(204)がまずターミナルトランスフェラーゼ (206)で処理されて3-プライム末端にポリ dA テイル(208)が付加される。この後、遊離末端のライゲーション (212) が分子内で、一方の末端でポリ dA テイルに相補的であり他方の末端で縮重ヌクレオチドのセグメントによりあらゆる配列に相補的である架橋オリゴヌクレオチド (210) の助けにより起こる。架橋オリゴヌクレオチド (210)の二本鎖領域 (214) はRCRのための少なくとも1つのプライマー結合部位を含み、そしてある態様において、捕捉オリゴヌクレオチドの相補鎖を提供する配列を含み、これはプライマー結合部位配列と同じであっても異なってもよく、あるいはプライマー結合部位配列とオーバーラップしていてもよい。捕捉オリゴヌクレオチドの長さは大幅に変動し得、一つの側面において、捕捉オリゴヌクレオチドおよび架橋オリゴヌクレオチドにおけるその相補鎖は10~ 100 ヌクレオチドの範囲の長さを有し;より好ましくは、10~ 40 ヌクレオチドの範囲である。ある態様において、二本鎖領域 (214) はさらなる要素を含んでいてもよく、例えば、その随伴する DNA 断片が由来するソース核酸を同定するためのオリゴヌクレオチドタグが挙げられる。即ち、ある態様において、異なるソース核酸由来の環またはアダプター ライゲーションまたはコンカテマーが別々に調製され得、その間に特有のタグを含む架橋アダプターが用いられ、その後にそれらはコンカテマー調製または表面への適用のために混合され、ランダムアレイが作られる。結合した断片は、標識化タグ相補鎖のコンカテマーにおけるその対応するタグ配列へのハイブリダイズによって、あるいは完全なアダプターまたはアダプターのタグ領域の配列決定 によってかかるランダムアレイ上で同定されうる。環状産物 (218)は、常套の精製カラム、一以上の適当なエキソヌクレアーゼによる非環状 DNAの消化、またはその両方によって便宜に単離されうる。

【0117】

所望のサイズ範囲、例えば50 - 600 ヌクレオチドのDNA 断片は、環状化酵素、例えば、CircLigaseを鋳型の必要無しに一本鎖DNAを環状化する一本鎖DNA リガーゼとして用いることにより環状化することができる。DNA 断片および一以上のアダプターを含む一本鎖DNA 環を形成する好ましいプロトコールはアダプター をDNA 断片の一方の末端にライゲーションする標準的リガーゼ、例えば、T4 リガーゼの使用、次いで環を閉環するためのCircLigaseの適用である。

【0118】

例示的な方法において、アダプター オリゴヌクレオチドおよび標的配列を含むDNA 環は合成オリゴヌクレオチド T1N (配列 : 5' -NNNNNNNGCATANCACGANGTCATNATCGTNCAAACGT CAGTCCANGAATCAGATCCACTTAGANTGNCGNNNNNNNN-3') (配列番号1)である標的配列を利用するT4 リガーゼ の使用により作成される。アダプターは2つの別々のオリゴヌクレオチドによりできている。T1Nの5' 末端に連結されるアダプター オリゴヌクレオチドは BR2-ad (配列 : 5' -TATCATCTGGATGTTAGGAAGACAAAAGGAAGCTGAGGACATTAACGGAC-3') (配列番号2) でありT1N の3'末端に連結されるアダプター オリゴヌクレオチドは UR3-ext (配列 : 5' -ACCTTCAGACCAGAT-3') (配列番号3)である。

【0119】

UR3-ext はIIs型制限酵素部位 (Acl I : CTTCAG)を含み、第二のアダプターの挿入のためのDNA 環の直鎖化の方法を提供する。BR2-ad はBR2-temp (配列 5' -NNNNNNNGTCCGTAA TGTCTCAG-3') (配列番号4)にアニーリングして 二本鎖 アダプター BR2 アダプターを形成する。UR3-extはビオチン化UR3-temp (配列 5' -[ビオチン]ATCTGGTCTGAAGGTNNNNNNN-3') (配列番号5)にアニーリングして二本鎖 アダプター UR3 アダプターを形成する。1 p molの標的 T1Nが25 p molの BR2 アダプター および10 p molのUR3 アダプターと単一のライゲーション反応でライゲーションされ、反応は50mM Tris-Cl、pH7.8、10% PEG、1mM ATP、50 mg/L BSA、10mM MgCl₂、0.3 ユニット/μl T4 DNA リガーゼ (Epicentre Biotechnologies, WI) および 10 mM DTTを 最終体積 10 μlにて含む。ライゲーション反応を以下のサイクリングプログラムの温度にてインキュベートする: 15 11分、37 1分、18回繰り返し。反応を70 °Cで10分加熱することにより終結する。過剰のBR2 アダプターをライ

ゲーションした産物をストレプトアビジン 磁気ビーズ (New England Biolabs, MA)で捕捉することにより除去する。3.3 μ lの4x 結合バッファー (2M NaCl、80 mM Tris HCl pH 7.5) をライゲーション反応に添加し、これを次に 15 μ g のストレプトアビジン 磁気ビーズ と1x 結合 バッファー (0.5M NaCl、20 mM Tris HCl pH 7.5)中で混合する。15 分間の室温でのインキュベーションの後、ビーズを4 体積の 低塩 バッファー (0.15M NaCl、20 mM Tris HCl pH 7.5)で2 回洗浄する。溶出 バッファー (10 mM Tris HCl pH 7.5)をあらかじめ 70度に温めておき、その10 μ lをビーズに70 で5分添加する。磁気分離の後、上清を一次精製サンプルとして保持する。このサンプルは過剰のUR3 アダプターをビオチン化 オリゴヌクレオチド BR-rc-bio (配列 : 5' -[ビオチン]CTTTTGTCTTCCTAACAT CC-3') (配列番号6)とあらかじめ結合させた磁気ビーズにより除くことによってさらに精製でき、このオリゴヌクレオチドは上記と同様にBR2-adと逆相補的である。

10

【0120】

アダプター-標的ライゲーション産物の最終精製サンプル中の濃度はウレアポリアクリルアミドゲル電気泳動分析により評価できる。環状化はライゲーション 産物を0.2 ユニット/ μ l T4 ポリヌクレオチド キナーゼ (Epicentre Bio Biotechnologies)を用いて 1 mM ATP および製造御者により供給される標準的 バッファー中でリン酸化することにより行われ、10 倍モル過剰のスプリント オリゴヌクレオチド UR3-closing-88 (配列 5' -A GATGATAATCTGGTC-3') (配列番号7)により0.3 ユニット/ μ lのT4 DNA リガーゼ (Epicentre Bio Biotechnologies) および 1mM ATPを用いて環状化される。環状化産物はRCR 反応の実施により確認される。

20

【0121】

図2Aに図示する別の例示的な態様において、アダプター オリゴヌクレオチド(1604)を用いて図2A に示す方法によりDNA 環の集団 (1608)が形成(1608)される。一つの側面において、集団 (1608)の各メンバーはソース核酸 (1600)からのDNA 断片に結合した同一のアンカープローブ 結合部位およびIIs型認識部位を備えたアダプターを有する。アダプターはその他の機能的要素を有していてもよく、例えばこれらに限定されないが、タギング配列、固体表面への結合のための配列、制限部位、機能付与配列等が挙げられる。DNA 環のクラスが、異なるアンカープローブ結合部位を有するアダプターを提供することによって作成される。

【0122】

DNA 環(図(2A) 1608)が形成された後、さらなる散在したアダプターが一般に図(2A)に示されるように挿入されて散在したアダプターを含む環(1612)が形成される。これらの環に、プライマーおよびローリングサークル複製 (RCR) 試薬を添加すると常套の RCR 反応においてアダプター オリゴヌクレオチド および DNA 断片の相補鎖のコンカテマー(1617)の集団 (1616)が形成(1614)される。この集団は単離してもよいし、その他のプロセスに供してもよく(例えばサイズ選択)(1618)、それは常套の技術、例えば常套のスピンカラム等を用いて行われ、分析用の集団 (1620)が形成される。

30

【0123】

複数のアダプター DNA 環の形成が実施可能であることを示すために70 塩基の長さの合成標的 DNAおよび200-300 bpの長さの PCR 由来断片を得ることが出来る。一本鎖 PCR 断片はプライマーの一方のリン酸化およびラムダエキソヌクレアーゼによるリン酸化された鎖の除去のための処理により二本鎖 産物から簡単に得られる。一本鎖断片を環状化のためにアダプターにライゲーションする。重合、IIs型制限酵素消化および新しいアダプターとの再ライゲーションは本明細書に記載するようにして行うことが出来る。

40

【0124】

プロセスが成功したことの証明は最終的に得られた環のRCR 増幅によって行いうる。簡単に説明すると、DNA 環は 最後に挿入されたアダプターに相補的 なプライマーおよびph i29 ポリメラーゼ と1時間 30 でインキュベートされ元の DNA 環の反復コピーを数百含む単一のコンカテマー 分子が生じる。RCR 産物のカバーガラスの表面への結合は表面に結合したオリゴヌクレオチドに相補的なコンカテマーにおけるアダプター配列を用いて行

50

われる。アダプター特有プローブのハイブリダイゼーションを用いて個々のアダプターが環に組み込まれ最終的に RCR 産物となったかを調べることができる。環中の予測された位置にアダプターが組み込まれたかを調べるために、アダプターの末端配列を認識する非標識アンカープローブに対するライゲーションが起こるように配列特異的プローブ（標識化 5-mer）を合成またはPCR由来配列に用いることが出来る。クローニングおよび配列決定も DNAの統合性を確認するために用いることが出来る。

【0125】

一つの態様において、環形成に用いられる鋳型は局所dsDNAを作成するためのプライマーとしても利用できる。計画は各環切断の後にクリーンなssDNAを作ることによって簡単にでき、これにより各アダプター組込みに同じ環閉環化学を使用することができる。

10

【0126】

一つの態様において、粘着末端または平滑末端を有するDNA断片の溶液をDNA環の作成のために調製する。二以上のDNA分子を有する環の作成を避ける伝統的な方法は低濃度のDNA断片を用いて大容積のライゲーションを行うことであり分子間ライゲーションが避けられる。

【0127】

好ましい態様において、ライゲーション反応は大容積を必要としない。この態様にはDNA断片のアリコートの通常のサイズのライゲーション反応へのゆっくりした添加が含まれる。DNAアリコートと反応との迅速な混合は多量体形成を最小にする。DNA断片はライゲーション混合物中にリガーゼを用いずに調製できあるいは水またはTE-様バッファー中に調製できる。典型的には、DNA体積はライゲーション反応の最初の体積と同じかそれより小さい。DNAサンプルの添加の速度でライゲーション反応が蒸発するならばDNAは水または簡単なバッファー（例えば、TEバッファー）中で大容積であってもよい。蒸発は熱安定性リガーゼの使用により簡便に出来る。

20

【0128】

一つの態様において、環状化方法はDNAの小さいアリコートの通常のライゲーション反応への希釈（例えば、0.1-0.5 μl を10-50 μl に希釈し100倍以上の希釈が行われる）およびDNAの大部分が環を形成するよう十分な時間の待機、次いで、第二のアリコートの添加を含む。別の態様において、DNA断片はゆっくりと連続的に添加される。

【0129】

30

プロセスの様々な物理的実行が可能であり、例えば、特定の頻度での手動または自動ピペット操作、ドリッパー（重力または正圧）、ピエゾまたは音響学的スパイティングまたはナノドロッパー、30 nlほどの小ささの液滴を送達できるカプロポンプ(cuvro-pump)の使用が挙げられる。一つの態様において最大一時的濃度 1 fmol/ μl を有する100 μl 反応中の10 pmol が100 アリコートの連続添加を用いて処理される。別の態様において、10 pmolは30-50 μl アリコート中にある。1つのアリコートにおけるDNA断片の>70-80%の環状化時間はリガーゼ濃度、末端のタイプ（粘着1、2、または4塩基あるいは平滑）およびある程度温度（粘着末端の運動およびハイブリッド安定性）に依存する。好ましい態様において、反応の総時間はおよそ 4-16 時間である。

【0130】

40

一つの態様において、リガーゼ酵素が固体支持体、例えば、ビーズ上に固定化される。DNA断片は、当該技術分野に知られた方法を用いてゲルブロックまたはその他の孔のある容器からライゲーション反応中に拡散する。断片の間のライゲーション（環状化ではなく）を妨げるために、DNAを一時的にブロックする当該技術分野に知られた方法が用いられ、例えばこれらに限定されないが一致する粘着末端とライゲーション不可能なDNAまたはssDNA末端結合タンパク質の使用が挙げられる。

【0131】

小さい反応体積の流れの効率を上げるために、一つの態様において反応体積を非蒸発条件下で、例えば小滴の使用により分配する。非蒸発条件は湿度、支持体周囲の温度を制御することにより、また反応バッファーの組成の設計により確立できる。例示的な態様にお

50

いて、10 pI 液滴がピエゾスパイティング (~20 x 20 x 20 ミクロン)により分配される。拡散がなければこれは20 ミクロンの厚さの フローセルに匹敵する。拡散は体積の厚さをさらに低下させることを促進し得、約 5-10ミクロンとなる。10 pI 液滴を拡散なしで用いて 1 cm²を被覆するには、100 x 50 x 50 = 250,000 滴が用いられる。

【 0 1 3 2 】

ピエゾアプローチに加えて大きい表面あたりに低量のバッファーを送達するその他の形態を利用でき、例えば、支持体と反応 バッファーを満たした多孔性材料の接触または数個の10-30ミクロンの開口を有する長いスリットを表面に沿って動かすことによりバッファーが分配される。

【 0 1 3 3 】

環状化の1つの例示的な方法は単一のアダプターの dsDNAへの 2つのブロックした相補的鎖を用いるライゲーションを含む。この方法において、アダプターの2つの相補的鎖は独立に調製される。ウラシルを有し標的 DNAにライゲーションし得ないマッチングブロックオリゴを2つの相補的鎖のそれぞれについて作る。1つのアダプター 鎖と1つのブロックオリゴを含むdsDNA 産物をアセンブリーする。2つのアセンブリーしたdsDNA コンストラクトを設計し、ライゲーションまたはハイブリダイズが互いに起こらないようにする；コンストラクトは平滑末端でもよいし、DNA 標的へのライゲーションのためにT オーバーハングまたはその他のオーバーハングを有していてもよい。これら2つのコンストラクトの混合物 を平滑末端 dsDNA または対応する粘着末端を有するDNAとライゲーションする。約 50%のDNAは各コンストラクトの一方を含む；残りの50%は 同じコンストラクトを2つ含む。ブロックオリゴを次いで分解し、環を相補的鎖のハイブリダイゼーションおよびライゲーションにより閉環する。

【 0 1 3 4 】

一つの態様において、アダプターは配向の分離を避けるために回文構造であってもよい。かかるアプローチは平滑末端 ライゲーション効率およびA/T ライゲーション反応におけるDNA濃度に依存してA/T ライゲーションアプローチよりもよい収率を提供しうる。さらに好ましい態様において、2つではなく4つの ssDNA アダプター 成分が用いられる。

【 0 1 3 5 】

コンカテマーの作成方法

本発明の一つの側面において、ポリヌクレオチドのコンカテマーを含む単一の分子、通常 ポリヌクレオチド分析物、即ち、標的配列は、常套の ローリングサークル複製 (RCR) 反応により作られる。RCR 反応の条件および試薬を選択するガイダンスは当業者に入手可能な多くの文献から得られ、以下に示すようなものは引用により本明細書に含まれる：

Kool、米国特許 5,426,180；Lizardi、米国特許 5,854,033 および 6,143,495；Landegren、米国特許 5,871,921；等。一般に、RCR 反応成分は、一本鎖DNA 環、DNA 環にアニーリングする一以上の プライマー、DNA 環にアニーリングしたプライマーの3'末端を伸長する鎖置換活性を有するDNA ポリメラーゼ、ヌクレオシド三リン酸および常套の ポリメラーゼ 反応 バッファーを含む。かかる成分はプライマーがDNA 環にアニーリングし、DNA ポリメラーゼにより伸長されてDNA 環相補鎖のコンカテマーを形成する条件下で一緒にされる。例示的なRCR 反応 プロトコールは以下の通り： 50 μL 反応混合物中、以下の成分と一緒にする： 2-50 pmol 環状 DNA、0.5 ユニット/μL フェージ φ29 DNA ポリメラーゼ、0.2 μg/μL BSA、3 mM dNTP、1X φ29 DNA ポリメラーゼ 反応 バッファー (Amersham)。RCR 反応は 30 で12 時間行う。ある態様において、ポリメラーゼ 反応における環状 DNAの濃度は絡み合いおよびその他の分子間相互作用を避けるために 低く選択するとよい(およそ100-1000億環/ ml、または 10-100 環/pI)。

【 0 1 3 6 】

好ましくは、RCRによって生じるコンカテマーはおおよそ均一なサイズを有する；したがって、ある態様において、本発明のアレイを作成する方法はコンカテマーのサイズを選択する工程を含みうる。例えば、一つの側面において、コンカテマーは集団として分子量の変動係数が約 30%未満となるように選択する；別の態様において、約 20%未満である。一

つの側面において、サイズの均一性は低濃度の鎖ターミネーター、例えば、ddNTPをRCR反応混合物に添加することによりさらに改良され、ポリメラーゼによってより高速で合成されたDNA環によって生じたものなど非常に大きなコンカテマーの存在が低下する。一つの態様において、ddNTPの濃度は50-250 Kb、あるいは50-100 Kbの範囲の予測されるコンカテマーサイズが得られるように用いられる。別の側面において、コンカテマーは常套の分離技術、例えばサイズ排除クロマトグラフィー、膜ろ過等を用いて特定のサイズ範囲について富化されうる。

【0137】

コンカテマーを作る例示的な方法を図2Aに図示する。DNA環(1608)が形成された後、さらに散在したアダプターを一般に図(2A)に示すようにして挿入し散在したアダプターを含む環(1612)を形成する。これら環に、プライマーおよびローリングサークル複製(RCR)

10

試薬を添加して(1614)常套のRCR反応においてアダプターオリゴヌクレオチドおよびDNA断片の相補鎖のコンカテマー(1617)の集団(1616)が作られる。この集団は常套の技術、例えば、常套のスピンカラム等を用いて、単離されるか別の処理を受け(例えばサイズ選択)(1618)、分析のための集団(1620)が形成される。

【0138】

標的ポリヌクレオチドはソース核酸、例えば、ゲノムDNAから、断片化により作られ0.2-2 kbのサイズ、より好ましくは、0.3-0.6 kbのサイズの断片が得られ、RCR反応のために環状化されうる。

【0139】

20

別の側面において、本発明は、散在したアダプターを含む複数の標的ポリヌクレオチドのコンカテマーを作るための方法および組成物を提供する。一つの態様において、かかるコンカテマーは図1C-1Dに図示されるようにRCRにより調製されうる。

【0140】

ローリングサークル複製が本発明のコンカテマーの作成のために好ましい方法である。RCRプロセスはM13ゲノムの複数の連続するコピーを作成することが示されている(Blanco, et al., (1989) J Biol Chem 264:8935-8940)。このシステムにおいて、所望のDNA断片がDNAアダプターに「クローニング」され直鎖状コンカテマー化によって複製される。標的DNAはただちにハイブリダイゼーションおよび酵素的方法に好適な形態となり、細菌を介した継代は必要ではない。

30

【0141】

RCRプロセスは環状基体へとまず形成されるべき所望の標的分子に依存する。この直鎖状増幅は複製の複製ではない元のDNA分子を利用し、それによって配列の忠実度が確認される。環状であるため、分子は環の部分に相補的なプライマーを伸長する鎖置換ポリメラーゼにとって終わりのない鋳型として作用する。連続的鎖伸長により長い、一本鎖DNAが生じ、これは環に相補的な配列の複数コピーを含む数百のコンカテマーからなる。

【0142】

アレイの作成方法

一つの態様において、エマルジョンPCRがアレイ上への配置のためのアンプリコンの作成に用いられる。図(1B)に図示するようにエマルジョン(1505)を壊した後、アダプター付加された配列のクローンを含むビーズが配列分析のために固体表面(1522)上でアレイとされる(1520)。かかるビーズのアレイは図1Fに示すようにランダムであり得、ビーズの位置はアレイ作成の前には決定されておらず、あるいはアレイはあらかじめ決められた結合部位(1524)のパターンに、たとえビーズのかかる部位への分配がランダムに決定されるにしろ、従いうる。かかる分配の両方を本明細書において「ランダムアレイ」と称する。

40

【0143】

マイクロミリより小さいスポットの形態においてコンパクトで密なDNAの束を達成するために、ガラス表面に結合した捕捉プローブへのハイブリダイゼーションのために増幅した分子の領域が利用されうる。数百の捕捉プローブ分子(約10 nm 離れて配置している

50

）が直径500 nm未満のガラス表面領域に強固に結合した標的分子の数百の連鎖状のコピーを保持しうる。一つの態様において、アミン修飾捕捉オリゴヌクレオチドを結合するためのイソシアネート反応基の単層を生じるガラス活性化化学が適用される。

【0144】

一般に、単一の分子の密度は少なくとも20パーセント、または少なくとも30パーセント、または少なくとも40パーセント、または少なくともほとんどの分子が、シグナル生成および用いられる検出システムによって個別に分解されるように選択される。一つの側面において、密度は少なくとも70パーセントの単一の分子が別個に分解されるように選択される。一つの側面において、金ナノ粒子標識を有する分子-特異的プローブとともに走査型電子顕微鏡を用いると、例えば、引用により本明細書に含める、Nie et al (2006)、Anal. Chem., 78: 1528-1534のように、密度は少なくとも単一の分子のほとんどが最近隣距離50 nm以上を有するように選択される；別の側面において、かかる密度は単一の分子の少なくとも70パーセントが100 nm以上の最近隣距離を有するように選択される。別の側面において、例えば蛍光標識を有する分子-特異的プローブとともに光学顕微鏡を用いる場合、密度は少なくとも単一の分子のほとんどの最近隣距離が200 nm以上となるように選択される；別の側面において、かかる密度は単一の分子の少なくとも70パーセントが200 nm以上の最近隣距離を有するように選択される。さらに別の側面において、蛍光標識を有する分子-特異的プローブとともに光学顕微鏡を用いる場合、例えば、密度は少なくとも単一の分子のほとんどが300 nm以上の最近隣距離を有するように選択される；および別の側面において、かかる密度は単一の分子の少なくとも70パーセントが300 nm以上または400 nm以上、または500 nm以上または600 nm以上または700 nm以上または800 nm以上の最近隣距離を有するように選択される。さらに別の態様において、光学顕微鏡を用いる場合、密度は少なくとも単一の分子のほとんどが顕微鏡の最小特徴解像力の少なくとも2倍の最近隣距離を有するように選択される。別の側面において、本発明のポリマー分子は、別々に検出可能なポリマー分子の密度が少なくとも1000/ μm^2 、または少なくとも10,000/ μm^2 、または少なくとも100,000/ μm^2 となるように表面に配置される。

【0145】

本発明の別の側面において、所望の最近隣距離を確保するランダムに配置された単一の分子の密度の選択の要求は、表面上に単一の分子の結合について実質的に唯一の部位である別々の空間的に離れた領域を提供することによって得られる。すなわちかかる態様において、別々の空間的に離れた領域の間の表面の上の領域は、本明細書において「領域間領域」と称され、コンカテマーまたはその他の高分子構造がかかる領域に結合しないという意味で不活性である。ある態様において、かかる領域間領域はブロッキング剤、例えばコンカテマー DNAと関連のないDNA、その他のポリマー等によって処理せらる。一般に、別々の空間的に離れた領域の面積は、結合化学、用いた高分子構造等とともに、本発明の単一の分子のサイズに対応するよう選択され、単一の分子が表面に適用されると実質的にあらゆる領域が二以上の単一の分子によっては占有されないように選択される。別々の空間的に離れた領域当たり1つのみの単一の分子が備えられる確率は反応性官能基または捕捉オリゴヌクレオチドの密度の選択によって高くなり得、かかる部分が単一の分子上の対応する成分よりも少なくなるようにするとよい。したがって、単一の分子は特定の別々の空間的に離れた領域の表面に対するすべての結合を「占有」し、それにより第二の単一の分子が同じ領域に結合する確率が低減する。具体的には、一つの態様において、別々の空間的に離れた領域における実質的にすべての捕捉オリゴヌクレオチドが単一の高分子構造のアダプターオリゴヌクレオチドにハイブリダイズする。一つの側面において、別々の空間的に離れた領域は多数の反応性官能基または捕捉オリゴヌクレオチドを含み、それは単一の分子の相補的官能基またはアダプターオリゴヌクレオチドの数の約10パーセントから約50パーセントである。捕捉オリゴヌクレオチドの長さおよび配列は大幅に変動し得、周知の原理、例えば、Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26: 227-259 (1991); Britten and Davidson, chapter 1 in Hames et al, editors, Nucleic Acid Hybridization: A Practical Approach (IRL Press, Oxford

10

20

30

40

50

d、1985)にしたがって選択されうる。一つの側面において、捕捉オリゴヌクレオチドの長さは6から30ヌクレオチドの範囲であり、別の側面において、8から30ヌクレオチドまたは10～24ヌクレオチドの範囲である。捕捉オリゴヌクレオチドの長さおよび配列は以下のように選択される(i)高分子構造の表面への有効な結合を提供し、分析操作、例えば、洗浄などの工程の際の高分子構造の損失が最小となるようにする、および(ii)特に分析物分子がコンカテマーにおけるDNA断片である場合、分析物分子に対する分析操作に干渉することをさけるようにする。(i)に関して、一つの側面において、配列および長さを捕捉オリゴヌクレオチドおよびその相補鎖の間の二本鎖が十分に安定であり、ストリンジェントな洗浄において解離しないように選択する。(ii)に関して、DNA断片が特定の種の生物由来である場合、データベースが入手可能である場合それを用いて間違っまたは望ましくないDNA断片とのハイブリッドを形成する可能性のある捕捉配列を除くように選択する。捕捉オリゴヌクレオチドの配列の選択におけるその他の因子はプライマー、ハイブリダイゼーションプローブ、オリゴヌクレオチドタグ等の選択において考慮されているものと同様であり、それについては十分にガイダンスがあり、定義のセクションにおいて以下に引用する引用文献により示される。

【0146】

一つの側面において、別々の空間的に離れた領域の面積は $1\text{ }\mu\text{m}^2$ 未満である；別の側面において、別々の空間的に離れた領域の面積は $0.04\text{ }\mu\text{m}^2$ から $1\text{ }\mu\text{m}^2$ の範囲である；さらに別の側面において、別々の空間的に離れた領域の面積は $0.2\text{ }\mu\text{m}^2$ から $1\text{ }\mu\text{m}^2$ の範囲である。別の側面において、別々の空間的に離れた領域がおおよそ環状または正方形の形状である場合、それらのサイズは単一の線寸法により示すことができ、かかる領域のサイズは125 nmから250 nmの範囲、または200 nmから500 nmの範囲である。一つの側面において、直近のかかる領域の中心間距離は $0.25\text{ }\mu\text{m}$ から $20\text{ }\mu\text{m}$ の範囲である；別の側面において、かかる距離は $1\text{ }\mu\text{m}$ から $10\text{ }\mu\text{m}$ の範囲、または50～1000 nmの範囲である。好ましくは、コンカテマーを固定化するための空間的に離れた領域は直線または六角形のパターンにて配置される。

【0147】

一つの態様において、スペーサーDNBを試験DNBの結合のための表面の調製に用いる。表面はまず、合成DNBの2タイプの上に存在する結合部位に相補的な捕捉オリゴヌクレオチドによりまず被覆される；1つは捕捉DNBでありもう1つはスペーサーDNBである。スペーサーDNBは試験DNBの調製に使用するアダプターに相補的なDNAセグメントを有さず、捕捉DNBに対して約5-50、好ましくは10x過剰にて用いられる。捕捉オリゴヌクレオチドを備えた表面は(鎖ライゲーションまたはRCRにより作られた)合成DNBの混合物により「飽和」され、ここでスペーサーDNBは捕捉DNBに対して約10倍(または5から50倍)過剰にて用いられる。スペーサーと捕捉DNBとの比が $\sim 10:1$ であるので、捕捉DNBはスペーサーDNBの海にほとんど島のように孤立している。 $10:1$ の比により2つの捕捉DNBが平均して2つのスペーサーDNBにより分離される。DNBの直径が約200 nmであれば2つの捕捉DNBは約600 nmの中心間空間を有する。この表面を次いで捕捉DNBの領域に相補的であるがスペーサーDNBには存在しない結合部位を有する試験DNBまたはその他の分子構造の結合に用いる。

【0148】

捕捉DNBは捕捉DNBスポット当たり単一の試験DNBの結合を確実にするように試験DNBの結合部位の数よりも少ないコピーを有するよう調製するとよい。試験DNAは捕捉DNBのみと結合しうするため、試験DNBのアレイは集まることなく高い部位占有率を有するよう調製されうる。ランダムな結合により、表面上のいくつかの領域にはDNBが結合しないこともあり得るが、遊離捕捉オリゴヌクレオチドを有するこれらの領域は試験DNBに結合できないであろう。というのはそれらは捕捉オリゴヌクレオチドのための結合部位を有するようには設計されていないからである。本発明のアレイは格子パターンに配置されていてもそうでなくてもよい。

【0149】

一つの側面において、マイクロミリより小さいサイズの捕捉オリゴヌクレオチドスポットの高密度アレイが コアおよびクラッド材を有する約 100億～ 1億光ファイバーの束状構造、または束の束状構造から調製されたプリンティングヘッドまたはインプリントマスター を用いて調製される。ファイバーを適切に引き、融合させることにより、同程度または2-5 倍小さいまたは大きいサイズのクラッド材により分離された約 50-1000 nmのコアを有する特有の材料が作られうる。一つの態様において、異なるクラッド材のエッチング(溶解)はナノサイズのポストを大量に有するナノプリンティングヘッドを提供する。このプリンティングヘッドを、オリゴヌクレオチドまたはその他の生物学的(タンパク質、オリゴペプチド、DNA、アプタマー) または化学的物質、例えば、様々な活性基を有するシランの沈着のために用いることが出来る。

10

【0150】

一つの態様において、グラスファイバー手段をパターン化された支持体としてオリゴヌクレオチドまたはその他の生物学的または化学的物質の沈着に用いることが出来る。この場合エッチングにより作られたポストのみが沈着すべき物質と接触されうる。別の態様において、融合した繊維束状構造の平坦なカットを用いて光をコアを通るようにガイドし、光誘導性化学がコアのチップ表面のみで起こるようにすることができ、したがってエッチングの必要がなくなる。両方の態様において、同じ支持体がタグオリゴヌクレオチドまたはその他の反応物に用いられる蛍光標識のイメージングのための光ガイド/収集装置として用いられる。この装置は大きい開口数 (可能性としては >1)を有する大きな視野を提供する。

20

【0151】

活性材料またはオリゴヌクレオチド沈着を行うためのスタンプまたはプリント手段を2から100種類のオリゴヌクレオチドを交互パターンにてプリントするのに用いることが出来る。このタイプのオリゴヌクレオチド アレイ は2から100種類の DNA 集団、例えば、異なるソースDNA由来の集団を結合させるために用いることができる。それらはまた垂光分解スポットからの並行読み取りに用いることが出来、それは DNA 特異的アンカーまたはタグの使用によって行われる。情報は、DNA 特異的タグ、例えば、16 DNAにつき16 の特異的アンカーおよび5-6色の組合せによる2 塩基の読み取り、および16 ライゲーションサイクルまたは1 ライゲーション サイクルと16 デコーディングサイクルの使用によりアクセスできる。

30

【0152】

本発明の態様において、フォトリソグラフィー、電子ビームリソグラフィー、ナノインプリントリソグラフィー、およびナノプリンティングを多様な表面上でのかかるパターンの作成に用いることが出来る。例えば、Pirrung et al、米国特許 5,143,854; Fodor et al、米国特許 5,774,305; Guo、(2004) Journal of Physics D: Applied Physics、37: R123-141; これらは引用により本明細書に含まれる。これら技術を1ミクロンの1/10のオーダーの特徴のパターンを作成するのに利用でき、これらは半導体産業での使用のために開発されたものである。好ましい態様において、単一の「マスキング」操作がDNA アレイ基体に行われ、これに対し一つの半導体を作るのにさえ典型的には20～ 30 マスキング操作が必要である。単一のマスキング操作を用いることにより、同じ基体に多くのマスクを正しく並べる必要が排除される。これにはまた材料のドーピングの必要がない。パターンにおけるわずかな欠陥はアレイの利用可能性にほとんどまたはまったく影響を及ぼさず、したがって生産収率は100%に近い。

40

【0153】

一つの態様において、高密度に構造化されたランダムDNA アレイチップは捕捉オリゴヌクレオチドを濃縮して小さくして有し、分離した捕捉セルを正方形の格子を形成するように並べる(図4)。好ましくは、各捕捉セルまたは結合部位は不活性表面により囲まれ、十分ではあるが限られた数(100-400)の捕捉分子を有しうる。各捕捉分子はRCRにより生じたDNA コンカテマー上的一致(matching)するアダプター配列の1コピーに結合しうる。各コンカテマーは1000コピーを超えるアダプター配列を含むため、接触すると結合部位が迅速

50

に飽和されることが可能であり、その他のコンカテマーが結合するのを妨げ、その結果、1つの結合部位またはスポット当たり1つのRCR産物のみが結合できる。十分なRCR産物を提供することにより、アレイ上のほぼすべてのスポットが1つであって唯一の特有のDNA標的を含むことが出来る。

【0154】

RCR「分子クローニング」によりランダムアレイの作成における飽和/排除(単一の占有)原理の適用が可能となる。インサイチュ増幅が代わりに適用されると単一の分子アレイの作成における排除プロセスは実現不可能である。RCRコンカテマーは小さい非混合DNAスポットの形成に最適なサイズを提供する。約100 kbの各コンカテマーは約 $0.1 \times 0.1 \times 0.1 \mu\text{m}$ の空間を占有すると予測され、それによりRCR産物が100 nm 捕捉セルに適合する。RCR産物の1つの利点は、一本鎖DNAはハイブリダイゼーションに容易であり、DNAのランダムコイルのボールを形成するために非常に柔軟であることである。RCRにより作られる1000コピーのDNA標的は単一の分子の分析により可能なものよりも非常に高い特異性を提供する。

【0155】

パターン化されたDNAチップを作成する当該技術分野に知られた方法が存在する。好ましい態様において、チップ上のすべてのスポットは同じ捕捉オリゴヌクレオチドを有し、0.5 ミクロンピッチにて0.2-0.3 ミクロンのスポットサイズを有する。新しいオリゴヌクレオチド結合化学の開発を必要としないため、ナノプリンティングアプローチをかけるパターンの生産に用いることが出来る。

【0156】

ナノインプリント技術はマスターモールドを作る古典的なフォトリソグラフィ技術に依存する。マスターモールドはポリマー、例えば、PMMAまたはPDMSを用いて複製される。これらポリマーは、硬化すると、マスターのネガティブモールドを形成する。モールドは基体上に材料のパターンを「プリント」するのに用いられる。ナノインプリント技術はガラス、シリコン、および金表面にタンパク質特徴部を作るのに利用できる。例示的な態様において、マスターモールドを用いて多くのスタンプ装置が作られ、各スタンプ装置は化学物質(例えば、オリゴヌクレオチド溶液、オリゴヌクレオチド結合またはガラス活性化化学物質)の多くのプリントを作りうる。進化したナノプリンティング技術では10 nmの小ささの特徴部を作ることが出来、したがって蛍光検出に適当な $>200 \text{ nm}$ のサイズの特徴部、例えば1000 ミクロンの中心間距離の300-500 nmの特徴部が常套的に作成できる。

【0157】

様々な化学的修飾を用いて表面特性を改変することが出来、マスターモールドの広範な材料との適合性が上昇し、それによって小さい特徴部、低密度モールドの使用が可能となり、高密度アレイが作成できるようになる。一つの態様において、特徴部ピッチが $4 \mu\text{m}$ のモールドを用いて同じ物質を 4×4 の格子に16回プリントすることにより基体上に $1 \mu\text{m}$ の特徴部ピッチを作ることが出来る。

【0158】

一つの側面において、DNAアレイの作成方法はフォトレジストの薄層を用いて機能付与プロセスの際に基体表面の部分を保護することを含む。パターン化されたフォトレジストは機能付与の後に除かれ、活性化した領域のアレイが残る。第二のアプローチは修飾オリゴヌクレオチドの単層を基体に結合させることを含む。オリゴヌクレオチドは光切断可能な保護基により修飾される。かかる保護基は光源への露光により除かれ、ハイブリダイゼーションによるDNBの結合のための捕捉オリゴヌクレオチドのパターン化されたライゲーションが可能となる。

【0159】

別の態様において、市販の、光学的に平面の、石英ウエハが100-500 nmの厚さのフォトレジストの層によりスピンコートされる。フォトレジストは石英ウエハ上で焼かれ、ステッパーと一般に称される機械を用いて活性化すべきスポットパターンを有するレチクルのイメージがフォトレジストの表面に投影される。露光の後、フォトレジストは現像され

、UV光源に露光された投影されたパターンの領域が除かれる。これは非常に微細な詳細部を作ることが出来る乾式現像技術であるプラズマエッチングにより達成される。ウエハは次いで残りのフォトレジストを強化するために焼かれる。

【0160】

焼いた後、石英ウエハはすぐに機能付与に使用できる。ウエハは次いで進行中の機能付与プロセスに用いられるのと同じモノマーである3-アミノプロピルジメチルエトキシシランの蒸着に供される。アミノ官能化モノマーの密度はモノマーの濃度および基体の曝露時間の変化により厳密に制御できる。プラズマエッチングプロセスに曝露された石英の領域のみがモノマーと反応でき、モノマーを捕捉できる。ウエハは次いで再び焼かれてアミノ官能化モノマーの単層を曝露された石英に硬化させる。焼いた後、残っているフォトレジストはアセトンを用いて除くことが出来る。レジストとシランとの結合化学における相違により、基体上のアミノシラン-官能化領域はアセトンによるすすぎの間もインタクトなまま残る。これら領域はさらにピリジンおよびN-N-ジメチルホルムアミドの溶液中のp-フェニレンジイソチオシアネートとの反応により官能化されうる。基体はアミン-修飾オリゴヌクレオチドと適合性になりうる。あるいは、オリゴヌクレオチドは5'-カルボキシ-修飾因子-c10 (Glen Research: <http://www.glenres.com/ProductFiles/10-1935.html>)によっても調製することが出来る。この技術はアミン修飾支持体に直接オリゴヌクレオチドを結合可能とし、それによりさらなる機能付与工程が避けられる。

【0161】

別の態様において、マスターインプリントツールの生産によって開始するナノインプリントリソグラフィー (NIL) プロセスを用いる。このツールは高分解能電子線(e-beam) リソグラフィーを用いて生産され、使用するNIL ポリマーに応じて多数のインプリントを作成するのに利用できる。DNA アレイ生産のために、石英基体をレジストの層でスピンコートし、この層は一般にトランスファー層と称される。第二のタイプのレジストを次いでトランスファー層の上に適用し、この層は一般にインプリント層と称される。マスターインプリントツールは次いでインプリント層の上に印象を作る。インプリント層の全厚さをプラズマエッチングにより減らしてインプリントの低い領域がトランスファー層に達するようにする。トランスファー層はインプリント層よりも取り除きにくいいため、それはほぼ未変化のまま残る。インプリントおよびトランスファー層を加熱により硬化させる。基体を次いでプラズマエッチャーに戻し、インプリントの低い領域が石英に達するようにする。基体を次いで方法 1aに記載のように蒸着により誘導体化する。

【0162】

別の態様において、ナノプリンティング方法が用いられる。かかるプロセスは、フォト、インプリント、または電子線リソグラフィーを用いてマスターモールドを作る。ナノインプリントツールの製造に用いられる技術に対して様々な改変が存在する。一つの例示的方法において、マスターモールドをプリントヘッド上に必要な特徴部のネガティブイメージとして作成する。プリントヘッドは通常、軟らかい柔軟なポリマー 例えば、ポリジメチルシロキサン (PDMS)からできている。様々な特徴を有するこの材料、または材料の層が、石英基体上にスピンコートされる。制御された温度および圧力条件下でモールドを用いて特徴部をレジスト材料の最上層にエンボス加工する。プリントヘッドは次いでプラズマに基づくエッチングプロセスに供されプリントヘッドの縦横比が改善され、エンボス加工した材料の緩和時間に起因するプリントヘッドの歪曲が除かれる。プリントヘッドを用いてアミン修飾オリゴヌクレオチドのパターンを均一に誘導体化した表面に沈着させる。これらオリゴ-ヌクレオチドはDNBのための捕捉プローブとして作用する。ナノプリンティングの1つの利点は異なる捕捉プローブの交互配置されたパターンをランダムアレイ支持体上にプリントする能力である。これは複数のプリントヘッドによる連続的プリンティングにより達成され得、それぞれのヘッドが異なるパターンを有し、すべてのパターンが互いに適合して最終的な構造化された支持体パターンを形成する。かかる方法はランダムアレイ内のDNA 要素の位置的コード化を可能とする。例えば、特異的アンカー配列を含む対照DNBは規則的間隔でランダムアレイに結合されうる。

【 0 1 6 3 】

電子ビームリソグラフィーは、基体の作成にも利用できる。このプロセスはフォトリソグラフィーと非常によく似ているが、ただし、パターンは特別のレジスト材料に電子ビーム銃を用いて直接描かれる。このプロセスの利点は UV フォトリソグラフィー方法によるものと比較して特徴部 サイズがかなり小さく、より正確となりうることである。可能性のある欠点はパターンを作成するのに必要な時間の量が基体当たり時間のオーダーであることであり、一方、フォトリソグラフィー方法を用いると数秒であり、NILについては 1 分未満である。

【 0 1 6 4 】

一つの態様において、アレイは保護基とも称される光切断可能な修飾因子を用いて作成される。かかる方法において、捕捉セルはオリゴヌクレオチドに対する市販の光切断可能な修飾因子、例えば、PC Linker Phosphoramidite (Glen Researchから市販)を用いて作成することが出来る。5プライム光切断可能保護基、この場合はDMTOを有するオリゴヌクレオチドは、3'末端にて石英の完全に官能化されたピースに結合される。露光された領域は保護基を失い、5'リン酸が残る。鋳型オリゴヌクレオチドが以下に示すように提供された場合、オリゴヌクレオチドライゲーションを用いて、RCR 産物のアダプター領域に相補的な捕捉オリゴヌクレオチドが露光されたリン酸基にライゲーションする：

【 化 5 】

(表面上のオリゴヌクレオチド)

|------cttactgtgc-P OH-ggactaccgtttagg..cccgtgg (捕捉オリゴヌクレオチド)
gaatgacacg.....cctgatggca (単一鋳型オリゴヌクレオチド)

【 0 1 6 5 】

捕捉オリゴヌクレオチドの脱保護された表面オリゴヌクレオチドへのライゲーションの後、基体全体がUV光源に露光されて残りの保護基が除かれうる。遊離のリン酸基はヘアピン様オリゴヌクレオチドのライゲーションによりブロックされ得、配列決定プロセスに用いられる標識化プローブの支持体オリゴヌクレオチドへのライゲーションが妨げられる。

【 0 1 6 6 】

製造方法に用いられるフォトレジスト材料は一般に極めて疎水性で、かかる材料に作られるパターンは非常に小さな孔からなる。フォトレジストの疎水性作用により石英の露光される表面がアミノ官能化モノマーの水溶液に接触しないようにされることがある。この問題を回避するために、本発明の一つの態様では、超音波を使用して液体をマスクの小さい開口部に通す力を加える。少量の界面活性剤、アセトン、またはその他の添加剤を溶液に加えて水の表面張力を破壊することも可能である。このような溶媒の使用はマスク材料をわずかに膨潤させる可能性があるが溶解することはない。レジスト除去プロセスの間にレジスト材料がアミノ官能化表面と非適合性である場合、例えばアミンと反応して破壊する可能性がある場合、強いアクリル系に基づく接着剤をポリマーシートに用いてレジスト材料を機械的に剥離することが可能である。

【 0 1 6 7 】

DNA アレイ基体の各バッチが作られた後、バッチが規格に耐えうるかを判定することが重要であり得る。規格はマスク設計および生化学最適化相の間に決定されうる。基体の各バッチの品質コントロールはFITC または蛍光標識を有するアミン-修飾オリゴヌクレオチドを反応性表面に結合させ、基体表面での蛍光の強度とパターンを観察することにより行うことが出来る。活性の領域の総強度は捕捉セルにおける反応性部位の密度に比例しうる。現在の顕微鏡システムは100x、1.4NA レンズを有し、理論的分解能は約 180 nmである。現在のイメージ 獲得システムの感度はピクセル当たり約 3 色素分子であり、各ピクセルは基体の60x60 nmの領域を画像化する。60 nm 正方形領域当たり10-50 捕捉オリゴヌクレオチドを結合させることが可能であると予測される。これにより高精度で基体の結合効率および格子特性を直接的に測定することが可能となる。各捕捉セルはおよそ10 ピク

10

20

30

40

50

セルにより画像化されうる。

【0168】

QCデータを用いると、どの基体調製工程に改良が必要かを判定することが出来る。プロセスのこの時点での捕捉セルの間の強度の変動は機能付与プロセスの際の不均一な反応条件またはフォトレジスト層の不均一な現象を指摘するであろう。セルの間に架橋が存在すれば、フォトレジスト材料が石英表面から剥がれたか、または露光プロセスの間に何かがあるまいかなかったことが示唆される。シグナル強度に関する問題は機能付与工程の制御不良を指摘するであろう。さらなる測定基準がプロセスが成熟するにつれて開発される必要があり得る。

【0169】

レプリカアレイ

本発明の一つの側面において、マスターアレイ上に合成された相補的ポリヌクレオチドがレプリカアレイに移される。かかる移動を達成するために、2つの表面はdsDNAを変性させ、新たに作られたDNA鎖を遊離させるために加熱の存在下で接触されうる。別の態様において、移動はプライマーよりも約5-50倍大きな電荷を有するレプリカのDNAのみを識別して移動させるよう電場を与えることにより達成される。さらに好ましい態様において、移動された鎖のハイブリダイズの後、逆電場が温度低下と組み合わせられてプライマーをマスターアレイに戻す。移動が電場の付与により達成される態様において、多孔性ガラスが電場の付与のために好ましくは用いられる。

【0170】

一つの態様において、捕捉オリゴヌクレオチドはプライミング部位の逆のアンプリコンの末端に対応するように設計され、全長コピーの排他的な保持が確認される。9以上の異なる捕捉オリゴヌクレオチドのパターンを有することによりマスターアレイからのDNA移動の際の「クロストーク」の確率が最小化される。一つの態様において、移動はレプリカアレイ上でのDNAのさらなる増幅なしに達成される；同じレプリカへの複数の移動はより強いシグナルの生成にも利用できる。別の態様において、複数のレプリカはマスターアレイからの部分的移動により作られ得、各レプリカアレイにおいてDNA増幅が行われる。

【0171】

例示的な態様において、レプリカアレイのための基体は第一のアレイに結合した鋳型DNAを用いるDNA合成の開始のためのプライマーを含む。マスターアレイの表面と「形成すべき」レプリカアレイの支持体とのDNAポリメラーゼ、dNTPおよび好適なバッファの存在下での最適温度での接触の後、プライマー分子はマスターアレイ上の鋳型DNAにハイブリダイズし、ポリメラーゼによって伸長されるようになる。停止剤、例えば、dsDNAを1コピーの末端でDNAを停止させるのに用いることが出来る。温度の上昇により、またはその他のDNA変性剤の使用により、DNA鎖は分離してレプリカアレイは離れて第一のアレイを形成しうる。元のDNAがマスターアレイから除かれるのを防ぐために、元のDNAを直接的に(または捕捉オリゴヌクレオチドを介して間接的に)マスターアレイ支持体に共有結合させうる。

【0172】

レプリカアレイに結合したあらゆる不完全なDNAは様々な当該技術分野に知られた方法、例えば、特定の末端を有する完全な分子の保護ライゲーションを用いて複製反応の完了後に特異的に除くことが出来る。不完全な分子は完全な分子を失うことなく除去することが出来る。

【0173】

一つの態様において、プライマーはアレイ調製のための基体表面全体をカバーする。ミクロン平方当たり10,000のプライマー密度は2つの支持体の間の1ミクロンにおける局所濃度をPCRに用いたものと同様または約10倍高い濃度とする。プライマーは第一のアレイの支持体上のDNA鋳型に到達することが出来るように非常に長い結合リンカーを有しうる。このプロセスにおいてDNA拡散の可能性はなく、レプリカDNAスポットは元のスポットよりわずかに大きくなりうるのみである。非常に平坦な表面を2つの表面の近接の確保

10

20

30

40

50

のために用いることが出来る。一つの態様において、DNBは約 300-500 nm の十分なDNA ループを提供し、100 nm プライマーリンカーと組み合わせられると、表面の欠陥の認容を可能とする。

【 0 1 7 4 】

レプリカアレイはさらなるレプリカの作成に利用できる。次の世代のレプリカは元のアレイと同じDNA 鎖を有するだろう。

【 0 1 7 5 】

レプリカアレイはDNA 断片の同じセットの並行分析、例えば、多数のプロープまたはプローププールによるハイブリダイゼーションに利用できる。別の態様において、自己アセンブリーしたゲノム断片を含むDNA マスターチップを複製してデコードが必要でない多くの検出アレイを作ることが出来る。というのはそれらは既にデコードされた同じマスターチップと一致するからである。したがってアレイの複製により自己アセンブリーした DNA アレイの調製がデコードのコストを最小にしつつ可能となる。というのは1つのマスターとそのレプリカが数千の最終アレイの作成に利用できるからである。

【 0 1 7 6 】

捕捉オリゴの構造

一つの態様において、表面 (図1C & D -- 1622) はコンカテマー中のアダプターオリゴヌクレオチドのセグメント、例えば、アンカー 結合部位 またはその他の要素との複合体、例えば、二本鎖の二本鎖を形成する捕捉オリゴヌクレオチドを結合して有し得る。別の態様において、捕捉オリゴヌクレオチドはオリゴヌクレオチドクランプまたは同様の構造を含み得、それは、アダプターオリゴヌクレオチドと三重鎖を形成する。例えば、Gryaznov et al、米国特許 5,473,060。別の態様において、表面 (1622) はコンカテマー上の相補的官能基と反応する反応性官能基を有し得、例えば、cDNAのマイクロアレイへの結合に用いられるものと同じ技術により共有結合を形成する。例えば、Smirnov et al (2004)、Genes、Chromosomes & Cancer、40: 72-77; Beaucage (2001)、Current Medicinal Chemistry、8: 1213-1244、これらは引用により本明細書に含まれる。

【 0 1 7 7 】

一つの側面において、酵素的プロセッシングが必要でない場合、捕捉オリゴヌクレオチドは好ましい特性、例えば、二本鎖安定性の上昇を付与する非天然ヌクレオシド性ユニットおよび/または結合を含み得る；かかる化合物としては、これらに限定されないが、ペプチド 核酸(PNA)、ロックされた核酸(LNA)、オリゴヌクレオチド N3' → P5' ホスホラミダート、オリゴ-2'-O-アルキルリボヌクレオチド等が挙げられる。

【 0 1 7 8 】

ランダムアレイの構造

一つの側面において、コンカテマー (1620 図1C & D) は共有結合および非共有結合を含む様々な技術によって表面 (1622) に固定され得る。一つの態様において、表面 (1622) は捕捉オリゴヌクレオチドを結合して有しており、捕捉オリゴヌクレオチドはコンカテマー中のアダプター オリゴヌクレオチドのセグメント、例えば、アンカー 結合部位またはその他の要素と複合体、例えば 二本鎖の二本鎖を形成する。別の態様において、捕捉オリゴヌクレオチドはオリゴヌクレオチド クランプまたは同様の構造を有し得、それはアダプターオリゴヌクレオチドと三重鎖を形成する。例えば、Gryaznov et al、米国特許 5,473,060。別の態様において、表面 (1622) はコンカテマー上の相補的官能基と反応する反応性官能基を有し得、cDNAのマイクロアレイへの結合に用いられるのと同じ技術により共有結合を形成する。例えば、Smirnov et al (2004)、Genes、Chromosomes & Cancer、40: 72-77; Beaucage (2001)、Current Medicinal Chemistry、8: 1213-1244、これらは引用により本明細書に含まれる。長いDNA 分子、例えば数百またはそれ以上のヌクレオチドも疎水性表面、例えば、低濃度の様々な反応性官能基、例えば、-OH基を有するクリーンなガラス表面に効率的に結合させることが出来る。

【 0 1 7 9 】

一つの態様において、完全なゲノム配列決定は分析されるポリヌクレオチド断片の50 ~

10

20

30

40

50

200x ゲノムカバード度を含むアレイを用いる。例えば平均断片長さ100 塩基を有する60億のDNBは100x ゲノムカバード度を表す6000億塩基を含むであろう。一つの態様において、アレイは300-600 塩基長のDNA 断片から構成される60億の DNBを含む。DNBは1 ミクロンのピッチで四角に圧縮された配置にてアレイ基体に結合し得、アレイ基体は16 セグメントに沿って分かれる。さらに好ましい態様において、各セグメントは24 ユニットのサブアレイを含み、各ユニットのサブアレイは2x2 mm²面積上に1600万の結合したDNBを含む。

【0180】

8 セグメントおよびDNBの250 塩基長を用いる配列決定アッセイは配列決定のために350のプローププールを必要としうる。断片長、DNB 数、プールセット、およびオーバーラップの間の様々な交換を用いて画像化時間に対する配列の質を最適にすることが出来る。例えば、16 セグメントにセグメント化した同じランダムアレイは配列決定のために225プローププールを必要とするであろう。これにはより少ないプローププールサイクル、画像化時間の低下が要求される。さらに、DNBは500 塩基長断片から構成され得、その場合16 反応チャンバにおいて試験される16 セグメントを用いる350 プローププールに対してアッセイされるべき30億の DNBが必要である。この形式は256x ゲノムカバード度のランダムアレイを作ることになり、これによりユニットアレイサイズが2mm²に低下する。一つの態様において、各プローププールは6 フルオロフォアのうち2を用いてコンビナトリアルに 標識され、それにより21の可能な蛍光標識の組合せが得られる。この標識化スキームは多くのプロープに対して同時にアッセイすることを可能とし、ハイブリダイゼーション時間を一桁減らすことが出来る。

【0181】

多種多様な支持体が本発明のアレイのために利用可能である。一つの側面において、支持体は、表面、好ましくは実質的に平面状の表面を有する堅固な固体であり、それにより調べるべき単一の分子が同じ平面上になる。後者の特徴部により検出光学により効率的なシグナル収集が可能となる。

【0182】

別の側面において、本発明の固体支持体は非多孔性であり、特に単一の分子のランダムアレイが小体積を必要とするハイブリダイゼーション反応によって分析される場合は非多孔性である。好適な固体支持体材料としては、例えば、 ガラス、ポリアクリルアミド-被覆ガラス、セラミクス、シリカ、シリコン、石英、様々なプラスチック等の材料が挙げられる。

【0183】

一つの側面において、平面状表面の面積は 0.5 ~ 4 cm²の範囲であり得る。一つの側面において、固体支持体はガラスまたは石英であり、例えば、表面が均一にシラン化された顕微鏡スライドである。これは常套のプロトコルを用いて達成でき、例えば、 酸処理、次いで3-グリシドキシプロピル トリメトキシシラン、 N,N-ジイソプロピルエチルアミン、および無水キシレン (8:1:24 v/v)の溶液への80 での浸漬により、エポキシシラン化表面が形成される。例えば、Beattie et al (1995)、Molecular Biotechnology、4: 213。かかる表面は容易に処理されて捕捉オリゴヌクレオチドが末端に結合され、それは例えば、表面への適用前に3'または 5' トリエチレン グリコール ホスホリル スペースを捕捉オリゴヌクレオチドに提供することによりなされる。多くのその他のプロトコルがガラス およびその他の表面に反応性官能基を結合させるために利用でき、Beaucage (先に引用)の開示に示されている通りである。

【0184】

散在したアダプターを有するDNA 標的のアレイは単一の分子またはコンカテマーに限定されず、インサイチュで増幅された DNA スポットのアレイまたは粒子のアレイを含み得、そのそれぞれは標的核酸の複数コピーを含む (例えばエマルジョンPCRにおいて用いられたビーズ)。さらに、本明細書に記載する、示差的に除去できるかその他の方法により識別できる複数のアンカーまたはプライマーを利用する方法は散在したアダプターに限定されず、即ちそれは2つの「標準」、即ち全部で4つのアンカー部位を有する末端ライゲ

ーションしたアダプターを有するサンプルに対して達成することが出来る。

【 0 1 8 5 】

プローブの構造

「プローブ」という用語は直接ハイブリダイゼーション、または2つのプローブのライゲーションまたはアンカーを備えるプローブ またはアンカープローブを備えるプローブにおいて用いられるオリゴヌクレオチドの広義の意味で用いられる。 プローブは数個のみの特異的塩基および多くの縮重塩基を有しうる：例えば BNNNNNNN または BBNNNNNN または NNBBNNNN等である。アンカープローブはアンカー U5-10 配列に相補的なアダプター配列に隣接する1-4 塩基を読み取るようU5-10B1-4として設計されうる。

【 0 1 8 6 】

本発明のオリゴヌクレオチド プローブは様々な方法で標識でき、例えば、放射性部分、蛍光部分、比色分析部分、化学発光部分等の直接的または間接的結合などの方法が挙げられる。DNA の標識およびDNA アダプターの構築についての方法の多くの包括的な総論が本発明のオリゴヌクレオチド プローブの構築に適用可能なガイダンスを提供する。かかる総論としては、Kricka、Ann. Clin. Biochem.、39: 114-129 (2002); Schaferling et al、Anal. Bioanal. Chem.、(April 12、2006); Matthews et al、Anal. Biochem.、Vol 169、pgs. 1-25 (1988); Haugland、Handbook of Fluorescence Probes and Research Chemicals、Tenth Edition (Invitrogen/Molecular Probes、Inc.、Eugene、2006); Keller and Manak、DNA Probes、2nd Edition (Stockton Press、New York、1993);および Eckstein、editor、Oligonucleotides and Analogues: A Practical Approach (IRL Press、Oxford、1991); Wetmur、Critical Reviews in Biochemistry and Molecular Biology、26: 227-259 (1991); Hermanson、Bioconjugate Techniques (Academic Press、New York、1996); 等が挙げられる。本発明に適用可能な多くのより具体的な方法が以下の引用文献に例示されている： Fung et al、米国特許 4,757,141; Hobbs、Jr.、et al 米国特許 5,151,507; Cruickshank、米国特許 5,091,519; (Synthesis of functionalized Oligonucleotides for attachment of reporter groups); Jablonski et al、Nucleic Acids Research、14: 6115-6128 (1986)(enzyme-oligonucleotide conjugates); Ju et al、Nature Medicine、2: 246-249 (1996); Bawendi et al、米国特許 6,326,144 (derivatized fluorescent nanocrystals); Bruchez et al、米国特許 6,274,323 (derivatized fluorescent nanocrystals); 等。

【 0 1 8 7 】

一つの側面において、一以上の蛍光色素がオリゴヌクレオチドプローブの標識として用いられ、例えば、Menchen et al、米国特許 5,188,934 (4,7-ジクロロフルオレセイン色素); Begot et al、米国特許 5,366,860 (スペクトル的に分離可能なローダミン色素); Lee et al、米国特許 5,847,162 (4,7-ジクロロローダミン色素); Khanna et al、米国特許 4,318,846 (エーテル-置換フルオレセイン色素); Lee et al、米国特許 5,800,996 (エネルギー移動色素); Lee et al、米国特許 5,066,580 (キサテン色素); Mathies et al、米国特許 5,688,648 (エネルギー移動色素); 等が開示されている。標識は量子ドットによっても行うことが出来、以下の特許および特許公報が開示されているとおりであり、これらは引用により本明細書に含める： 6,322,901; 6,576,291; 6,423,551; 6,251,303; 6,319,426; 6,426,513; 6,444,143; 5,990,479; 6,207,392; 2002/0045045; 2003/017264; 等。本明細書において用いる場合、「蛍光シグナル生成部分」という用語は一以上の分子の蛍光吸収および/または発光特性を介して情報を運ぶシグナル伝達手段を意味する。かかる蛍光特性としては、蛍光強度、蛍光寿命、発光スペクトル特性、エネルギー移動等が挙げられる。

【 0 1 8 8 】

オリゴヌクレオチドの標識に容易に組み込まれる市販の 蛍光 ヌクレオチド アナログとしては、例えば、Cy3-dCTP、Cy3-dUTP、Cy5-dCTP、Cy5-dUTP (Amersham Biosciences、Piscataway、New Jersey、USA)、フルオレセイン-12-dUTP、テトラメチルローダミン-6-dUTP、テキサスレッド(登録商標)-5-dUTP、カスケードブルー(登録商標)-7-dUTP、BOD

10

20

30

40

50

IPY (登録商標) FL-14-dUTP、BODIPY (登録商標) R-14-dUTP、BODIPY (登録商標) TR-14-dUTP、ローダミンググリーン (商標) -5-dUTP、オレゴングリーン (登録商標) 488-5-dUTP、テキサスレッド (登録商標) -12-dUTP、BODIPY (登録商標) 630/650-14-dUTP、BODIPY (登録商標) 650/665-14-dUTP、Alexa Fluor (登録商標) 488-5-dUTP、Alexa Fluor (登録商標) 532-5-dUTP、Alexa Fluor (登録商標) 568-5-dUTP、Alexa Fluor (登録商標) 594-5-dUTP、Alexa Fluor (登録商標) 546-14-dUTP、フルオレセイン-12-UTP、テトラメチルローダミン-6-UTP、テキサスレッド (登録商標) -5-UTP、カスケードブルー (登録商標) -7-UTP、BODIPY (登録商標) -FL-14-UTP、BODIPY (登録商標) -TMR-14-UTP、BODIPY (登録商標) TR-14-UTP、ローダミンググリーン (商標) -5-UTP、Alexa Fluor (登録商標) 488-5-UTP、Alexa Fluor (登録商標) 546-14-UTP (Molecular Probes、Inc. Eugene、OR、USA) が挙げられる。合成後の結合に利用可能なその他のフルオロフォアとしては、とりわけ、Alexa Fluor (登録商標) 350、Alexa Fluor (登録商標) 532、Alexa Fluor (登録商標) 546、Alexa Fluor (登録商標) 568、Alexa Fluor (登録商標) 594、Alexa Fluor (登録商標) 647、BODIPY 493/503、BODIPY FL、BODIPY R6G、BODIPY 530/550、BODIPY TMR、BODIPY 558/568、BODIPY 558/568、BODIPY 564/570、BODIPY 576/589、BODIPY 581/591、BODIPY 630/650、BODIPY 650/665、カスケードブルー、カスケードイエロー、ダンシル、リサミン ローダミン B、マリナブルー、オレゴングリーン 488、オレゴングリーン 514、パシフィックブルー、ローダミン 6G、ローダミンググリーン、ローダミンレッド、テトラメチルローダミン、テキサスレッド (Molecular Probes、Inc.、Eugene、OR、USAから市販)、および Cy2、Cy3.5、Cy5.5、および Cy7 (Amersham Biosciences、Piscataway、NJ USA等) が挙げられる。FRET タンデムフルオロフォアも利用でき、例えば、PerCP-Cy5.5、PE-Cy5、PE-Cy5.5、PE-Cy7、PE-テキサスレッド、およびAPC-Cy7が挙げられる; さらに、PE-Alexa 色素 (610、647、680) およびAPC-Alexa 色素もある。ビオチンまたはその誘導体も検出オリゴヌクレオチドの標識として利用でき、後で検出可能に標識したアビジン/ストレプトアビジン 誘導体 (例えば、フィコエリトリン-結合ストレプトアビジン)、または検出可能に標識した抗-ビオチン抗体により結合される。ジゴキシゲニンも標識として組み込むことができ、後で検出可能に標識した抗-ジゴキシゲニン抗体 (例えばフルオレセイン化抗-ジゴキシゲニン) によって結合される。アミノアリル-dUTP 残基も検出オリゴヌクレオチドに組み込むことができ、後でN-ヒドロキシスクシンイミド (NHS) 誘導体蛍光色素、例えば、前記のものに結合される。一般に、結合体対のいずれのメンバーも検出オリゴヌクレオチドに組み込むことができるが、ただし、検出可能に標識した 結合体 パートナーが検出を行うために結合されうる必要がある。本明細書において用いる場合、抗体という用語はあらゆるクラスの抗体分子またはその亜断片、例えば、Fabをいう。検出オリゴヌクレオチドのためのその他の好適な標識としては、フルオレセイン (FAM)、ジゴキシゲニン、ジニトロフェノール (DNP)、ダンシル、ビオチン、プロモデオキシウリジン (BrdU)、ヘキサヒスチジン (6xHis)、リン-アミノ酸 (例えば、P-tyr、P-ser、P-thr)、またはその他の好適な標識が挙げられる。一つの態様において以下のハプテン/抗体対が検出に用いられ、ここで各抗体は検出可能な標識により誘導体化されている: ビオチン/ α -ビオチン、ジゴキシゲニン/ α -ジゴキシゲニン、ジニトロフェノール (DNP)/ α -DNP、5-カルボキシフルオレセイン (FAM)/ α -FAM。以下のスキームに記載するように、プローブは間接的に標識されていてもよく、特に捕捉剤によって結合されるハプテンによって標識されていてもよく、例えば、Holtke et al、米国特許 5,344,757; 5,702,888; および 5,354,657; Huber et al、米国特許 5,198,537; Miyoshi、米国特許 4,849,336; Misiura and Gait、PCT公開 WO 91/17160; 等に開示されている。多様なハプテン-捕捉剤対が本発明に利用できる。例示的なハプテンとしては、ビオチン、デス-ビオチンおよびその他の誘導体、ジニトロフェノール、ダンシル、フルオレセイン、CY5、およびその他の色素、ジゴキシゲニン等が挙げられる。ビオチンについては、捕捉剤はアビジン、ストレプトアビジン、または抗体であり得る。抗体はその他のハプテンのための捕捉剤として利用できる (多くの色素-抗体対が市販されている、例えば、Molecular Probes)。

10

20

30

40

50

【0189】

一つの側面において、プローブのプールが提供され、それは好ましくは約 1 から約 3 塩基を有し、縮重位置において異なる配列について均一かつ最適化されたシグナルを可能とする。一つの態様において、濃度が調整された3-mer 構成要素の混合物がプローブ合成に用いられる。

【0190】

プローブは直接標識されるかわりに核酸タグテイルを有するよう調製してもよい。テイルは好ましくは試験DNAと相互作用しないものである。これらテイルは天然塩基またはそれ自体の間でのみ対合する修飾塩基、例えば、isoCおよびisoGから調製されうる。isoC およびisoG ヌクレオチドを用いる場合、配列は5' アミノ-リンカーを有するよう別々に合成してもよく、それにより各タグ付加プローブ上に合成される5' カルボキシ 修飾リンカーへの結合が可能となる。これにより別々に合成されたタグ配列が既知のプローブと合わさることが可能となる一方、それらはカラムに結合したままである。一つの態様において、21のタグ付加配列を1024 の既知のプローブと合わせて用いる。

【0191】

テイルは1-3 またはそれ以上の縮重塩基、脱塩基部位またはその他のリンカーによりプローブから離れていてもよい。テイルと標的 DNAとの相互作用を最小化する1つのアプローチは標的 DNAにおいて非常に稀な配列を使用することである。例えば、CGCGATATCGCGAT ATまたは CGATCGATCGATは哺乳類ゲノム中で稀なものであると予測される。1つの選択肢は非標識化タグとプレ-ハイブリダイズしたテイルを有するプローブの使用であり、非標識化タグはライゲーションの後、標識化タグとのハイブリダイゼーションの前に変性され洗い流されうるであろう。ウラシルを分解可能なテイル/タグの作成および温度による除去の使用のかわりに新規なサイクルの実行の前のそれらの除去に使用することができる。

【0192】

一つの側面においてプローブのハイプレックス多重ライゲーション アッセイを用い、ここでプローブは蛍光色素により標識されていないため、バックグラウンドおよびアッセイ費用を低減できる。例えば8 色について、 $4 \times 8 = 32$ の異なるコード化テイルを調製し、32のプールとしてのプローブをハイブリダイゼーション/ライゲーションに用いるとよい。デコード化(解読)プロセスにおいて、それぞれ8 タグを有する4サイクルが用いられる。したがって各色は4のデコード化サイクルに用いられる4つのタグのために用いられる。各サイクルの後、タグは除去され得、色素は光退色されうる。このプロセスはデコードすべきプローブの最後のセットが4のデコード化サイクルにわたってハイブリダイズしたままであることを必要とする。

【0193】

一つの態様において、さらなる特性が同じ色を用いる異なるプローブの識別能力を提供するために含まれ、例えば T_m /安定性、組み込まれたウラシル塩基およびUDG 酵素による分解可能性、および化学的または光化学的切断可能な結合が挙げられる。直接にまたは同じ色のための8の別々のタグを提供するための安定化剤を切断または除去した後；二以上の切断タイプを用いて 3以上の群を作ることが出来る；同じ色のこの4-8 または6-12の曝露を実行するために、2つの特性、例えば、温度安定性の組合せが要求されることがあり得、低い光退色条件、例えば、増感CCD (ICCD)により検出されうる低い強度の照明などが要求される。例えば1つの特性が融解温度 (T_m)であり、別々の T_m を有する4つのタグ-オリゴまたはアンカーまたはプライマーが存在する場合、第一の 4のプローブを有するさらなるセットの4つのオリゴを調製することが出来、第一の4のプローブは安定化剤を有さない第一の群のもっとも安定なオリゴよりもこれら4 オリゴの T_m を高くする安定化剤と連結または相互作用可能とされる。第一の群からの4 オリゴを連続する融解により分離した後、温度を最初の低いレベルに下げるとよく、安定化剤は切断または除去され得、4のタグ付加-オリゴまたはアンカーまたはプライマーは第一の群についてのものと同じ温度の点を用いて示差的に融解されうる。

【0194】

一つの側面において、プローブ-プローブ ハイブリッドは別の非標識化オリゴヌクレオチドへのライゲーションを介して安定化される。

【0195】

散在したアダプターを用いる配列決定方法

一つの側面において、本発明は標的ポリヌクレオチドのヌクレオチド配列を決定する方法を含み、該方法は以下の工程を含む： (a) 標的ポリヌクレオチドの中に複数の散在したアダプターを作成する工程、各散在したアダプターは標的ポリヌクレオチドとの少なくとも一つの境界を有する；および(b) 少なくとも二つの散在したアダプターの少なくとも一つの境界に隣接する少なくとも一つのヌクレオチドの実体を判定し、それにより標的ポリヌクレオチドのヌクレオチド配列を決定する工程。以下により詳細に説明するように、標的配列は配列情報が望まれる位置を含み、一般にそれは本明細書において「検出位置」と称される。一般に、配列情報（例えば、特定の検出位置でのヌクレオチドの同定）が複数の検出位置について望まれる。本明細書において用いる場合、「複数」とは少なくとも二つを意味する。しかし場合によっては、例えば 一ヌクレオチド多型 (SNP) 検出における場合などは、情報は特定の標的配列内の単一の検出位置についてのみ望まれうる。本明細書において用いる場合、ハイブリッドにおける検出位置塩基と塩基対形成する塩基は「調査位置」と称される。

10

【0196】

本発明の重要な特徴は標的ポリヌクレオチドに関連する配列情報の獲得のための標的ポリヌクレオチド アンプリコンにおける散在したアダプターの使用である。様々な配列決定方法を散在したアダプターに関して用いることが出来、例えば、これらに限定されないが、ハイブリダイゼーションに基づく方法、例えば、Drmanac、米国特許 6,864,052；6,309,824；および 6,401,267；およびDrmanac et al、米国特許出願公開 2005/0191656に開示のもの、ならびに合成方法による配列決定、例えば、Nyren et al、米国特許 6,210,891；Ronaghi、米国特許 6,828,100；Ronaghi et al (1998)、Science、281： 363-365；Balasubramanian、米国特許 6,833,246；Quake、米国特許 6,911,345；Li et al、Proc. Natl. Acad. Sci.、100： 414-419 (2003)；Smith et al、PCT 公報 WO 2006/074351に開示のもの；およびライゲーションに基づく方法、例えば、Shendure et al (2005)、Science、309： 1728-1739、Macevicz、米国特許 6,306,597に開示のものが挙げられる；これら文献は引用により本明細書に含まれる。

20

30

【0197】

一つの側面において、本発明による標的ポリヌクレオチドのヌクレオチド配列を決定する方法は以下の工程を含む： (a) 標的ポリヌクレオチドから複数の標的 コンカテマーを作成する工程、各標的 コンカテマーは標的ポリヌクレオチドの断片の複数コピーを含み、複数の標的 コンカテマーは標的ポリヌクレオチドを実質的にカバーする多数の断片を含む；(b) 少なくとも大部分の標的 コンカテマーが光学的に分解可能な密度で表面に固定した標的 コンカテマーのランダムアレイを形成する工程；(c) 各標的 コンカテマーにおける各断片の少なくとも部分の配列を同定する工程；および (d) コンカテマーの断片の部分の配列の実体から標的ポリヌクレオチドのヌクレオチド配列を再構築する工程。通常、「実質的にカバーする」とは分析すべきDNAの量が少なくとも2コピーの標的ポリヌクレオチドの相当量を含むことをいい、あるいは別の側面において、少なくとも10コピー、または別の側面において、少なくとも20コピー、あるいは別の側面において、少なくとも100コピーの相当量を含むことをいう。標的ポリヌクレオチドはDNA断片を含んでいてもよく、例えばゲノム DNA断片およびcDNA断片およびRNA断片が挙げられる。標的ポリヌクレオチド配列の再構築工程についてのガイダンスは以下の文献にみることができ、それらは引用により本明細書に含まれる： Lander et al、Genomics、2： 231-239 (1988)；Vingron et al、J. Mol. Biol.、235： 1-12 (1994)；等の文献。

40

【0198】

本発明の一つの側面において、ライゲーションに基づく配列決定方法を図3A-3Eに示すように用いることが出来る。この配列決定アプローチの多様な改変が以下のような因子に

50

依存して当業者により選択されうる：例えば、望まれる配列決定の量、用いる標識のタイプ、用いる標的ポリヌクレオチドアンプリコンのタイプおよびその表面への結合方法、配列決定操作の望まれる速度、シグナル検出アプローチ等。 変更は単に例示のために図3A-3Eに示される。

【0199】

本発明の一つの側面において、標識化プローブは対合するプローブへのライゲーションの後にのみ安定なハイブリッドを形成することが出来る。プローブライゲーションの使用はハイブリダイゼーション方法による標準的配列決定よりもデータ特異性を向上させる。プローブライゲーションは位置特異的塩基同定（例えば DNA末端）または全配列走査方法（例えばすべての内部オーバーラップ配列）にも適用される。

10

【0200】

未知の配列の特定の部位における配列、例えば、配列の末端の配列を同定するために、標識化プローブはアンカープローブへのライゲーションが可能ないように設計されうる。より長いアンカープローブが決定すべき未知配列、例えば、検出位置の末端に隣接する既知のアダプター配列にハイブリダイズされる。標識化プローブは様々な数の特異的および縮重塩基を有し得る。例えば、2の末端塩基はプローブ BBNNNNNN (A = アンカー、D = アダプター、G = ゲノム、B = 塩基を規定するプローブ、N = 縮重塩基、* = 標識)を用いて決定されうる：

【化6】

AAAAAAAAA.BBNNNNNN*
DDDDDDDDDDDDDDGGGGGGGGGGGGGGGGGG

20

【0201】

かかるプローブ構造のために16の配列-読み取りプローブが存在し、それぞれ5-プライム末端における2の特異的塩基からなる。16のすべてのプローブが試験されると、1つのみが効率的にアンカープローブにライゲーションしてアンカープローブにライゲーションしなかったプローブの除去の後、強いシグナルを生じる。このようなポジティブプローブはゲノム DNA 断片の末端の2塩基を高い特異性にて検出し、特異性はライゲーション部位に近い相補的塩基に対するT4 DNA リガーゼの強い優先性により提供される。

【0202】

本発明の一つの側面において、一本鎖標的ポリヌクレオチドは複数の散在したアダプターを含むように提供される。図3Aにおいて、3つの散在したアダプター(3002、3004、および 3006)が示され、これはアンプリコン、例えば、標的ポリヌクレオチド (3000)の複数コピーを含むコンカテマーの一部であり得る。各散在したアダプターは特有の配列を有する領域（例えば3008 および3012）を各末端に有し（この例において3つの散在したアダプターの間に全部で6つのかかる特有の配列がある）、これは配列決定プローブがライゲーションするオリゴヌクレオチド（これは標識を有していても有していなくてもよい）である対応するアンカープローブに対する結合部位として設計される。かかる末端領域は6 から14 ヌクレオチドの範囲、より通常には、8 から12 ヌクレオチドの範囲の長さを有しうる。散在したアダプターは中心領域 (3010)を有していてもよく、これはさらなる要素、例えば、様々な酵素の認識部位（二本鎖形態の場合）または標的ポリヌクレオチドアンプリコンを表面に固定化する捕捉オリゴヌクレオチドの結合部位等を含んでいてもよい。一つの側面において、散在したアダプター (3002-3006)を用いる配列決定操作は異なる特有のアンカープローブ結合部位のそれぞれへアンカープローブをハイブリダイズさせる6つの連続するルーチンを含む。各ルーチンのそれぞれは、アンカープローブの散在したアダプターのその末端部位へのハイブリダイズ工程、完全に一致したプローブのみのハイブリダイゼーションを可能とする条件下での配列決定プローブとの混合工程、完全に一致した配列決定プローブの並置したアンカープローブとのライゲーション工程、ライゲーションした配列決定プローブの検出工程、配列決定プローブにより生成するシグナルによるアンカープローブに隣接する一以上の塩基の同定工程、および標的ポリヌクレオチドアンプリコンからの配列決定プローブおよびアンカープローブの除去工程のサイクルを含

30

40

50

む。

【 0 2 0 3 】

さらなる態様は300-3000 塩基長のDNA 環を作成することおよび最初のアダプターの各側に2-3のアダプターを挿入することを含む。このようにすると、2つの、20-60 塩基長配列の、300-3000 塩基離れたメイトペアが作成される。配列データのレベルを2倍にすることに加えて、この方法は価値あるマッピング情報を提供する。メイトペアはデノボ配列 アセンブリーにおける反復を架橋し得、ゲノム 再配列決定において20-50 塩基より長い反復における突然変異を正確に位置づけるのに用いることが出来る。1つまたは2つの~20-50 塩基 配列のメイトペアは、プローブハイブリダイゼーションまたはプローブ-プローブライゲーションデータにより補完されうる。全部で 5-mer、6-mer、7-merまたは 8-merの1/8から 1/16の部分的セットをスコア付けして200-4000 塩基長断片のマッピング情報を提供することが出来る。さらに、所与の長さすべてのプローブ(例えば、すべての6-mer)を所与のゲノムについての全DNA アレイの4-16 セクションを含む4-16 反応 チャンバ中でスコア付けすることができる。各チャンバにおいてすべてのプローブの1/4~1/16がスコア付けされうる。個々のDNA 断片のマッピングの後、すべてのプローブを編集するとオーバーラップした断片におけるオーバーラップした プローブにおいて塩基あたり100~ 1000の読み取りを提供することが出来る。

【 0 2 0 4 】

一つの態様において、6の連続するルーチンを1から4回、好ましくは2から3回繰り返すことにより散在したアダプターから異なる距離にあるヌクレオチドを同定することが出来る。別の態様において、6の連続するルーチンを1回行うが、アンカープローブハイブリダイゼーション、配列決定プローブハイブリダイゼーション、ライゲーション等の各サイクルを、1から4回、または2から3回繰り返す。前者を図3Aに図示し、ここではアンカープローブ (3015) が散在したアダプター (3002)における結合部位にハイブリダイズした後、標識化配列決定プローブ(3016)が完全に一致する二本鎖が形成されるとアンカープローブ (3015)とのライゲーションを可能とする条件下で反応混合物に添加される。

【 0 2 0 5 】

配列決定プローブは様々な異なる構造を有しうる。典型的には、それらは縮重配列を含み直接的または間接的に標識されている。図3Aの例において、配列決定プローブは直接的に例えば蛍光色素 F1、F2、F3、およびF4により標識されており、これらは相互に識別可能なシグナルを生成するものであり、また、蛍光色素 G1、G2、G3、および G4により標識されており、それらも相互に識別可能なシグナルを生成するものである。この例において、各セットの色素、即ち、FおよびGが、異なるサイクルにおいて検出されるため、それらは同一の色素であってもよい。8-mer 配列決定プローブを用いる場合、散在したアダプターに直近する塩基を同定するためのF-標識化プローブのセットは 以下の構造を有しうる： 3'-F1-NNNNNNNp、3'-F2-NNNNNNNCp、3'-F3-NNNNNNNGp、3'-F4-NNNNNNNT。ここで配列 (3000)は左から右に5'→3' 配向と想定し 常套の リガーゼ媒介ライゲーション反応を用いる場合F-標識化プローブは5'末端にリン酸基を担持していなければならない。同様に、G-標識化プローブの対応するセットは以下の構造を有しうる： 3'-ANNNNNNN-G1、3'-CNNNNNNN-G2、3'-GNNNNNNN-G3、3'-TNNNNNNN-G4、そしてこれらプローブのライゲーションのためにそれらが結合するアンカープローブは5'-リン酸基を有していなければならない。連続サイクルにおける F-標識化プローブは以下の構造を有しうる： 3'-F1-NNNNNNANp、3'-F2-NNNNNNNCnp、3'-F3-NNNNNNNGnp、3'-F4-NNNNNNNTN、および 3'-F1-NNNNNNANp、3'-F2-NNNNNNCnp、3'-F3-NNNNNNGnp、3'-F4-NNNNNTNN等。

【 0 2 0 6 】

図3Aに戻ると、ライゲーションしたプローブ (3018)が同定された後、それは標的ポリヌクレオチド アンプリコン から除去され(3020)、次のアンカープローブ (3022)が対応する結合部位にハイブリダイズする。 G-標識化配列決定プローブが標的ポリヌクレオチドにハイブリダイズし、アンカープローブに並置される完全に一致する二本鎖を形成したものがライゲーションされて同定される。このプロセスは最後のライゲーションしたプロ

ープ (3028) が同定されるまで各アンカープローブ 結合部位について継続される。サイクルの全順序が反復され(3030)、対応するアンカープローブに隣接する異なる塩基を同定するように設計されたF-標識化配列決定プローブおよびG-標識化配列決定プローブが用いられる。

【 0 2 0 7 】

図3Bは図3Aの方法の改変法であり、ここではアンカープローブが2つ同時にそれぞれの結合部位にハイブリダイズする。アンカープローブのあらゆる対が、対の1つのメンバーが散在したアダプターの3' 結合部位に結合し、対のもう1つのメンバーが散在したアダプターの5' 結合部位に結合する限り用いられ得る。直接的に標識した配列決定プローブについて、示されるようにこの態様は8つの識別可能な標識の使用を必要とする；即ち標識F1-F4とG1-G4はそれぞれ互いに識別できなければならない。図3Bにおいて、アンカープローブ(3100 および3102) は散在したアダプター (3002)におけるそれぞれの結合部位にハイブリダイズし、その後配列決定プローブのセット(3104)がストリンジェントなハイブリダイゼーション条件下で添加される。完全に一致した二本鎖を形成するプローブがライゲーションされ、ライゲーションしていないプローブは洗い流され、その後ライゲーションしたプローブが同定される。散在したアダプター (3002)に隣接する異なる部位での塩基を同定するように設計された配列決定プローブのセットを用いてかかるハイブリダイゼーション、ライゲーションおよび洗浄のサイクルが繰り返される(3110)。プロセスは各散在したアダプターについて繰り返される。

【 0 2 0 8 】

図3Cは 図3Aの態様のさらなる改変法を図示し、ここでアンカープローブに隣接する各位置での塩基を同定するための配列決定プローブが、別の散在したアダプターのためのアンカープローブが用いられる前に、完全に実施される。簡単に説明すると、各破線の箱(3200)内の工程は各アンカープローブ 結合部位について1回で行われる；したがって、各破線の箱は異なるアンカープローブ結合部位に対応する。各箱の中で、連続するサイクルが行われ、それはアンカープローブのハイブリダイズ工程、配列決定プローブとのライゲーション工程、ライゲーションした配列決定プローブの同定工程をそれぞれ含む。

【 0 2 0 9 】

図3Dはコード化標識を用いる態様を図示し、これは引用により本明細書に含めるAlbrecht et al、米国特許 6,013,445に開示のコード化アダプターを用いるものと類似している。プロセス は図3Cに記載のものと類似しているが、ただし直接的に標識した配列決定プローブのかわりに、かかるプローブはオリゴヌクレオチドタグにより間接的に標識されている。かかるタグの使用により、ライゲーション工程の数が低減できる。というのは 各配列決定プローブ混合物は4を上回る塩基を同定する配列を含みうるからである。例えば、クロスハイブリダイズしないオリゴヌクレオチドタグを塩基の16の対のそれぞれに対応するように選択し、ライゲーションの後、ライゲーションした配列決定プローブを各2塩基配列が同定されるまで標識化抗タグのセットにより調べることができるようにするとよい。したがって、アンカープローブに隣接する標的ポリヌクレオチドの配列は一度に2つまたは一度に3つまたはそれ以上、コード化された配列決定プローブを用いて同定することが出来る。図3Dによると、アンカープローブ (352)がアンカー結合部位 (381)にハイブリダイズし、その後コード化された配列決定プローブが、完全に相補的な配列決定プローブ (354)のみがアンカープローブ(352)にライゲーションすることを可能とする条件下で添加される。かかるライゲーションおよびライゲーションしていない配列決定プローブの洗浄除去の後、標識化された抗-タグ(358) が配列決定プローブのオリゴヌクレオチドタグに ストリンジェントな条件下で連続的にハイブリダイズし、完全に一致した二本鎖を形成する標識化された抗-タグのみが検出される。様々な異なる標識化スキームを 抗-タグとともに用いることが出来る。単一の標識をすべての抗-タグに用いて各抗-タグが別々にコード化された配列決定タグにハイブリダイズするようにしてもよい。あるいは、抗-タグのセットを用いて行わなければならないハイブリダイゼーションおよび洗浄の数を低減することも出来る。例えば、各配列決定プローブが2塩基を同定する場合、4つの抗

-タグそれぞれの2セットを適用すればよく、ここで所与のセットの各タグは配列決定プローブによって同定される2つの塩基の1つの実体に応じて別々の 標識を担持する。同様に、配列決定プローブが3塩基を同定する場合、4つの抗-タグそれぞれの3つのセットがデコード化に使用されうる。デコード化のかかるサイクルは各散在したアダプターについて行うとよく、その後さらなるサイクルを異なる部位での塩基を同定する配列決定プローブを用いて行うとよい。

【0210】

図3Eは図3Bに記載のものと類似の態様を図示するが、ただし、コード化された配列決定プローブが用いられる。したがって一度に2つのアンカープローブが標的ポリヌクレオチドにハイブリダイズし、対応する配列決定プローブが標識化された抗-タグによるデコード化によって同定される。示されるように、アンカープローブ (316および318)は散在したアダプター (3002)のそれぞれの結合部位にハイブリダイズし、その後コード化された配列決定プローブ (327)の2セットが、完全に一致した二本鎖を形成するプローブのみがライゲーションする条件下で添加される。ライゲーションしていないプローブの除去の後、ライゲーションしたプローブのオリゴヌクレオチドタグが標識化された抗-タグによりデコードされる。上記のように、様々なスキームがライゲーションした配列決定プローブのデコード化に使用できる。

【0211】

別の側面において、複数の DNAまたはRNA 断片における配列の決定のための本発明に使用される配列決定方法は以下の工程を含む: (a) DNAまたはRNA 断片のコンカテマーをそれぞれ含む複数の ポリヌクレオチド 分子を作成する工程; (b)少なくとも大部分の標的コンカテマーが光学的に分解可能な密度で表面に固定されたポリヌクレオチド 分子のランダムアレイを形成する工程;および (c) 光学的に検出可能な反応物の少なくとも一つの化学的反応を用いて分離可能なポリヌクレオチドにおける各DNA またはRNA 断片の少なくとも部分配列を同定する工程。一つの態様において、かかる光学的に検出可能な反応物はオリゴヌクレオチドである。別の態様において、かかる光学的に検出可能な反応物はヌクレオシド三リン酸、例えば蛍光標識されたヌクレオシド三リン酸であり、それはコンカテマーにハイブリダイズしたオリゴヌクレオチドの伸長に使用可能である。別の態様において、かかる光学的に検出可能な試薬は第一および第二の オリゴヌクレオチドをライゲーションして形成され、コンカテマー上の隣接する二本鎖を形成するオリゴヌクレオチドである。別の態様において、かかる化学的反応は DNA またはRNAの合成であり、例えば、コンカテマーにハイブリダイズしたプライマーの伸長によるものである。

【0212】

一つの側面において、ランダムアレイ上の標的ポリヌクレオチドのコンカテマーの並行配列決定がコンビナトリアル SBH (cSBH)により達成され、これは上記特許において Drmanacにより開示されている。一つの側面において、オリゴヌクレオチドプローブの第一および第二のセットが提供され、ここで各セットは、セットにおけるプローブの規定された長さについて可能なあらゆる配列を有するオリゴヌクレオチドを含むメンバー プローブを有する。例えば、セットが長さ6のプローブを含む場合、それは 4^{06} のプローブを含む。別の側面において、オリゴヌクレオチド プローブの第一および第二のセットは標的ポリヌクレオチドの選択されたセットを検出するよう設計された選択されたヌクレオチド配列を有するプローブを含む。配列は、プローブまたはプローブのプールをハイブリダイズさせること、第二の プローブ または第二のプローブのプールをハイブリダイズさせること、完全に一致した二本鎖を標的配列上に形成するプローブをライゲーションすること、ライゲーションしたプローブを同定し、標的配列に関する配列情報を得ること、すべてのプローブまたはプローブのプールがハイブリダイズするまで工程を繰り返すこと、およびハイブリダイゼーションおよび同定工程において蓄積された配列情報から標的ヌクレオチド配列を決定することにより決定される。

【0213】

配列決定操作のために、ある態様において、米国特許 6,864,052に開示のようにセット

10

20

30

40

50

はプールにおいて一緒に用いられるサブセットに分割してもよい。第一および第二のセットからのプローブは一緒にまたは順番に、全セットまたはサブセットあるいはプールとして標的配列にハイブリダイズしうる。一つの側面において、第一または第二のセットにおけるプローブの長さは 5 から 10 ヌクレオチドの範囲であり、別の側面において、5 から 7 ヌクレオチドの範囲であり、それによりライゲーションした場合、それぞれ 10 ~ 20 および 10 ~ 14 の範囲の長さのライゲーション産物が形成される。

【 0 2 1 4 】

別の側面において、かかる技術を用いて、各結合した DNA コンカテマーの配列同一性が「サイン」アプローチにより決定されうる。約 50 ~ 100 またはおそらくは 200 プローブが用いられて約 25-50% あるいは用途によっては 10-30% の結合したコンカテマーが各プローブについて完全に一致する配列を有する。このタイプのデータによりコンカテマー内の各増幅された DNA 断片が参照配列に対してマッピング可能となる。例えば、かかるプロセスにより 4 色標識スキームにおいて 16 のハイブリダイゼーション/ストリップオフ サイクルを用いて 64 の 4-mer (即ち可能性のあるすべての 256 の 4-mer の 25%) をスコア付けすることが可能である。60-70 塩基断片がコンカテマーにおいて増幅されると 64 プローブの約 16 がポジティブとなる。というのは 64 塩基長配列には 64 の可能性のある 4-mer が存在するからである (即ちすべての可能な 4-mer の 4 分の 1)。関連のない 60-70 塩基断片は約 16 のポジティブデコード化 プローブの非常に異なるセットを有するであろう。64 プローブのなかから 16 プローブの組合せは 10 億断片のうち 1 の確率のランダムな確率を有し、それは実際的にはそのコンカテマーについての特有のサインを提供する。20 サイクルにおいて 80 プローブをスコアリングし、20 のポジティブプローブを作成することにより、より特有である可能性が高いサインが生じる: 10 億 × 10 億に付き 1 の確率である。以前には、「サイン」アプローチは cDNA ライブラリーからの新規遺伝子の選択に用いられていた。サインアプローチの実行はすべての試験プローブの得られた強度をソートし、ポジティブプローブ閾値を満足するプローブのあらかじめ決められた (予測された) 数まで選択することである。これらのプローブはアレイに存在すると予測されるすべての DNA 断片の配列に対してマッピングされる (より長い参照配列のウィンドウをずらすことを利用する)。選択したポジティブプローブのすべてまたは統計的に十分な数を有する配列が所与のコンカテマーにおける DNA 断片の配列として割り当てられる。別のアプローチにおいて、予測されるシグナルをあらかじめ測定された完全な一致および不一致ハイブリダイゼーション/ライゲーション効率を用いて使用されたすべてのプローブについて規定することができる。この場合、相関係数に類似の尺度を計算することができる。

【 0 2 1 5 】

4-mer のスコア付けの好ましい方法はプローブの対のライゲーションであり、例えば: $N_{(5-7)}BBB$ と $BN_{(7-9)}$ であり、ここで B は規定された塩基であり N は 縮重塩基である。より長い DNA コンカテマープローブ上のサインの作成のために、より特有の塩基が使用される。例えば、長さ 1000 塩基の断片において 25% がポジティブである比は $N_{(4-6)}BBBB$ と $BN_{(6-8)}$ により達成されるであろう。より長い断片であっても同じ数の約 60-80 プローブを必要とすることに注意されたい (4 色を用いる 15-20 ライゲーション サイクル)。

【 0 2 1 6 】

一つの態様において、所与の長さのすべてのプローブ (例えば、 $4096 N_{2-4}BBBBBN_{2-4}$) またはすべてのライゲーション対を用いてコンカテマーの DNA の完全な配列を決定することが出来る。例えば、 $N_{(5-7)}B_3$ と $BBN_{(6-8)}$ との 1024 の組合せをスコア付けして (4 色を用いる場合 256 サイクル)、約 250 塩基、好ましくは約 100 塩基までの DNA 断片の配列を決定することが出来る。

【 0 2 1 7 】

配列決定プローブの多数の N によるデコード化は、効率における相違を最小化するために縮重塩基における配列のサブセットの多重合成から調製できる。各サブセットが適当な濃度にて混合物に添加される。また、いくつかのサブセットはその他と比較してより多くの縮重位置を有しうる。例えば、セット $N_{(5-7)}BBB$ からの 64 プローブのそれぞれは 4

10

20

30

40

50

種類の合成において調製されうる。第一は通常のすべての完全に縮重されるべき5-7 塩基であり；第二は N0-3(A,T)5BBBであり；第三は N0-2(A,T)(G,C)(A,T)(G,C)(A,T)BBBであり第四は N0-2(G,C)(A,T)(G,C)(A,T)(G,C)BBBである。

【0218】

3つの特定の合成からのオリゴヌクレオチド 調製物が実験的に決定された量にて通常の合成に添加され、BBB 配列の前に ATリッチ (例えば AATAT)または(AまたはT)と(GまたはC)との交互配列 (例えばACAGTまたはGAGAC)を有する標的配列とのハイブリッド生成を増加させる。これら配列はハイブリッドの形成においてより効率が低いと期待される。すべての1024 標的配列を N_{0-3} NNNNBBB プローブとのハイブリッド形成の効率について試験でき、最も弱い結合を与えるタイプが約 1-10 のさらなる合成において調製され得、基本的プローブ調製物に添加される。

10

【0219】

別の態様において、より少ない数のプローブを少ない数の別々のサンプルについて用いる；例えば、20 プローブのなかからの5-7 のポジティブ (4 色を用いる5 サイクル)は約 1万-10万の別々の断片を識別する能力を有する。

【0220】

一つの側面において、8-20-mer RCR 産物はDNA コンカテマーの形態において特有の8から20 塩基の認識配列のランダムな分布として形成されたアレイの提供によりデコードされる。プローブは多数の可能な方法を用いる8-20 塩基 プローブ 領域 の配列の決定のためにデコードされる。例示的方法において、配列の半分を短いプローブのハイブリダイゼーション特異性および完全に一致するハイブリッドのライゲーション特異性を用いて決定する。12 merに隣接する6 ~ 10の塩基はあらかじめ規定されており6mer から10-mer オリゴヌクレオチドのための支持体として作用する。この短い6merはその3-プライム末端にて4 標識化6-merから10-merのいずれかにライゲーションする。これらデコード化 プローブは4 オリゴヌクレオチドのプールからなり、ここで各オリゴヌクレオチドは 4-9 の縮重塩基と 1の規定された塩基とからなる。このオリゴヌクレオチドは4つの蛍光 標識の1つにより標識されている。4つの可能な塩基、A、C、G、またはTのそれぞれはそれゆえ蛍光色素によって表される。例えば これら 5群の4 オリゴヌクレオチドおよび1ユニバーサル オリゴヌクレオチド(U)は 12-merの最初の5 塩基の配列決定のためのライゲーション アッセイにおいて利用される： B=末端において特異的色素またはタグを結合された4 塩基のそれぞれ：

20

30

【化7】

```
UUUUUUUU.BNNNNNNN*
UUUUUUUU.NBNNNNNN
UUUUUUUU.NNBNNNNN
UUUUUUUU.NNNBNNNN
UUUUUUUU.NNNNBNNN
```

【0221】

6以上の塩基をさらなるプローブプールにより配列決定できる。12-merの中心近くの位置の識別を改良するために 6-mer オリゴヌクレオチドを12-mer 配列中にさらに位置させるとよい。これはシフトを収容するために縮重塩基の非標識化オリゴヌクレオチドの3'末端への組込みを必要とする。以下は12-merにおける位置6および7のためのデコード化プローブの例である：

40

【化8】

```
UUUUUUUN.NNNNBNNN
UUUUUUUN.NNNNBNNN
```

【0222】

50

同様にして 12-merの右側からの6 塩基を固定されたオリゴヌクレオチドおよび5-プライム標識化プローブを用いてデコードすることができる。上記システムにおいて6 サイクルが12-merの1つの側の6 塩基の規定に必要である。ライゲーション部位から遠い塩基の余分なサイクル 分析によりこれは7または8 サイクルに増えることがある。12-merの完全な配列決定はしたがって12-16 サイクルのライゲーションにより達成できる。

【0223】

一つの態様において、本発明は2つの別々のタイプのディテクタープローブのライブラリーを合わせるによりアレイとされた DNAの部分的または完全な配列決定のための方法を提供する。このアプローチにおいて1セットは一般的タイプ $N_{3-8}B_{4-6}$ (アンカー)のプローブを有し、それらはセット BN_{6-8} 、 NBN_{5-7} 、 N_2BN_{4-6} 、および N_3BN_{3-5} からの第一の2または3または4 プローブ/プローブプールとライゲーションされる。例示的方法において、1-4またはそれ以上の4-merは5-mer アンカーにハイブリダイズして分子の約70%-80%についてDNA当たり1または2のアンカーが得られる。一つの態様において、ポジティブ アンカー が別々の ハイブリッド 安定性 (付加される Nの数が異なることにおそらくは起因する) を有する特異的プローブの混合により決定される。アンカーはタグ付加されてプールからのどのアンカーがスポットにハイブリダイズするのかを決定することが出来る。付加的DNA セグメントとしてのタグを、検出方法としての調節可能な置換のために利用できる。例えば、EEEEEEENNNAAAAAおよび FFFFFFFFNNNCCCC プローブはハイブリダイゼーションまたはハイブリダイゼーションおよびライゲーションの後、示差的に2つの対応するディスプレイサ: EEEEEENNNNNN および FFFFFFFFNNNNNNNN により除去され得、ここで第二の方がより有効である。別の態様において、別々のサイクルを用いてどのアンカーがポジティブであるかを判定することが出来る。この目的のために複数の色により標識またはタグ付加されたアンカーが非標識化 N7-N10 サポーター オリゴヌクレオチドにライゲーションすることが出来る。

【0224】

BNNNNNNNN プローブは次いで4 塩基に対応する4 色とハイブリダイズする。識別的洗浄またはタグの相補鎖による置換を用いて2つのスコア付けされた塩基のいずれが2つのアンカーが1つのDNAにおいてポジティブである場合アンカーに結合するのかを読み取る。したがって、2つの7-10 塩基 配列が同時にスコア付けされうる。2-4 サイクルを用いて4-6 塩基 アンカーを各アレイあたり16種類のアンカーの2-4 塩基のさらなるランのために伸長でき(4 色を用いる場合32-64物理的サイクル)、各断片当たり約 16 の可能な8-mer (合計~100 塩基)が決定される。これは参照可能性にそれをマッピングするのに十分であり、参照可能性は100-merが10の8-merのセットを有するのが一兆×一兆あたり 1未満; ($10e^{-28}$)である。別のアレイにおける同じ断片で並行にスコア付けされた異なるアンカーからのデータをまとめることにより、その断片の配列決定が完了し、伸長により 全ゲノムがオーバーラップする 7-10-merから作成されうる。

【0225】

一つの側面において、本発明は、より多重のデコード化または配列決定プローブのためにDNA タグによりプローブにタグ付加する方法を提供する。直接の標識の代わりに、 プローブは天然塩基または新規合成塩基 (例えば、isoGおよびisoC)からなる様々なオリゴヌクレオチド配列によりタグ付加されうる。タグは異なるオリゴヌクレオチド長(約 6-24 塩基) および/またはGC含量を含む配列を用いてその抗-タグによる非常に正確な結合効率を有するよう設計されうる。例えば 4の異なるタグが 4 連続サイクルまたは1のハイブリダイゼーション サイクルとその後の識別的洗浄において特異的抗-タグにより識別できるように設計されうる。識別的洗浄において、最初のシグナルはそれぞれ各タグにつき95-99%、30-40%、10-20% および0-5%に低下する。この場合、異なるタグを有するプローブはめったに同じドットにハイブリダイズしないと仮定すると、2つのイメージを得ることにより、4の測定が得られる。継続的にデコードされる(即ち2-16の別々の色により一度に標識される2-16)としても多くの異なるタグを有するさらなる利点は、1つのアッセイ反応において個々に認識可能な多数のプローブを使用する能力である。このようにプロー

ブが 短いインキュベーションおよび除去反応においてデコードされると4-64 倍長いアクセス時間(より特異的または強いシグナルを提供しうる)を与えることが出来る。

【 0 2 2 6 】

デコード化プロセスには 48-96 またはそれ以上のデコード化 プローブの使用が必要である。これらプールはさらにそれらをそれぞれ異なる発光スペクトルを有する4のフルオロフォアでコード化することにより、12-24 またはそれ以上のプールへと一緒にされる。20x 対物レンズを用いると、10 メガピクセルカメラの使用により各6 mm x 6 mm アレイは全カバーのためにおよそ30 イメージを要求しうる。各1マイクロメーターアレイ領域は約 8 ピクセルによって読み取られる。各イメージは250 ミリ秒において獲得できる：150 msは露光のためであり、100 msはステージの移動のためである。この速い獲得を使用すると各アレイのイメージングに~7.5秒かかり、即ち各基体上の96 アレイの完全なセットのイメージングに12 分間かかる。

10

【 0 2 2 7 】

イメージングシステムの一つの態様において、高いイメージ 獲得速度が4つの10メガピクセルカメラの使用により達成され、そのそれぞれは異なるフルオロフォアの発光スペクトルをイメージングする。カメラは一連のダイクロイック・ビームスプリッタを介して顕微鏡に接続されている。さらなる時間がかかるオートフォーカスルーチンは獲得したイメージの焦点がずれている場合にのみ作動する。それは Z 軸位置情報を保存し次のイメージングサイクルの際にそのアレイのそのセクションへと戻るのに用いられる。基体の上の各位置についてのオートフォーカス位置をマッピングすることによりイメージ 獲得に必要な時間を劇的に低減することが出来る。

20

【 0 2 2 8 】

典型的には、各アレイはデコードのために約 12-24 サイクル必要とする。各サイクルは、ハイブリダイゼーション、洗浄、アレイ イメージング、およびストリップオフ 工程からなる。これら 工程は、それぞれの順序において、上記例についてそれぞれ5、2、12、および 5 分間必要とし得、操作を線形に行う場合、各サイクルに全部で24 分間かかり、各アレイについて およそ5-10 時間かかることになる。各アレイをデコードする時間はシステムがイメージングを一定にするようにすることにより2桁低減することが出来る。これを達成するために、各顕微鏡上の2つの別々の基体のイメージングを一緒に行い、即ち、1つの基体の反応中に、もう1つの基体のイメージングを行う。

30

【 0 2 2 9 】

cSBHを用いる例示的なデコード化 サイクルは以下の工程を含む：(i) アレイ温度をハイブリダイゼーション 温度 (通常5-25 の範囲)に設定する工程；(ii)少量の デコード化 プローブ と適量のハイブリダイゼーション バッファーを予備混合するためにロボットピペッターを使用する工程；(iii) 混合した試薬をハイブリダイゼーション チャンバにピペットにより入れる工程；(iv)あらかじめ決められた時間ハイブリダイズさせる工程；(v)ポンプ(シリンジまたはその他)を用いて試薬をチャンバから排出する工程；(vi)非-ハイブリッドの不一致(mismatch)を洗浄するためのバッファーを添加する工程；(vii)チャンバ 温度を適当な洗浄温度に調節する工程 (約 10-40)；(viii)チャンバから排出する工程；(ix)イメージングの改良に必要であればさらなる洗浄バッファーを添加する工程；(x)好ましくは1以上の高ピクセル数高感度CCD カメラに光学的に連結した中程度の(20x) 顕微鏡対物レンズを用いて各アレイをイメージングする工程；プレートステージがチャンバ (またはおそらくは入力漏斗を有するフローセル)を対象の上で移動させ、または対物レンズ光学アセンブリーをチャンバの下で移動させる；特定の光学的配置を、二色ミラー/ビームスプリッタを用いて利用して多スペクトル イメージを同時に獲得することが出来、それによりイメージ 獲得時間が低下する；アレイはセクションまたは全体にてイメージング可能であり、それはアレイ/画像 サイズ/ピクセル 密度に依存する；セクションは基体 にあらかじめ(活性部位作成の際)コードされた統計的に有意な空の領域を用いてイメージをアラインすることによりアセンブリーでき、あるいは多工程 ナノプリンティング 技術を用いて作ることが出来、例えば部位 (活性化部位の格子)を特異的捕捉プロ

40

50

ープを用いてプリントし得、空の領域を格子に残す；別々のプリントヘッドを用いてその領域に異なるパターンまたは捕捉プローブをプリントする；(xi)チャンバ から排出し、プローブストリップバッファ ー により交換し(またはすでに入れられたバッファ ー を使用する)、チャンバをプローブストリップオフ 温度 (60-90)まで加熱する工程；高 pH バッファ ー をストリップオフ 工程に用いてストリップオフ 温度を低下させることが出来る；特定時間待機する；(xii)バッファ ー を除去する工程；(xiii)セットにおける次のデコード化 プローブプールを用いて次のサイクルを開始する工程。

【 0 2 3 0 】

ハイブリダイゼーションによる配列決定のためのコンビナトリアルプローブライゲーション

10

本発明の好ましい側面において、標的ポリヌクレオチドの配列に関する情報は、コンビナトリアルプローブライゲーションを利用するハイブリダイゼーション 方法による配列決定を介して得られる。本発明のこの側面において、2つの完全な、短いプローブのユニバーサルセットがDNA リガーゼ の存在下で標的 DNA に曝される(R. Drmanac、米国特許 6,401,267、2002)。典型的には1つのプローブセットは固体支持体、例えば、ガラススライドに結合しており、フルオロフォアにより標識された他方のセットは、溶液中で可動性である。結合した標識化プローブが標的に正確に隣接する位置でハイブリダイズするとそれらはライゲーションされ、スライド表面に共有結合した長い標識化プローブが生成する。所与の位置でのポジティブ シグナルは標的内の配列の存在を示し、それはシグナルを生成するために一緒にされた2つのプローブを補完する。

20

【 0 2 3 1 】

好ましい態様において、ユニバーサル 配列決定 チップ、例えば、Complete Genomics により開発されたHyChip (商標) スライドが本発明のハイブリダイゼーション 方法によるコンビナトリアル 配列決定に使用される。一つの態様において、各HyChip (商標) は通常の顕微鏡ガラススライドを含み、スライドは結合した6-merの8つのレプリカアレイを含み、4096のアレイ化した 6-merおよび1024の 標識化5-mer プローブを用いてサンプル当たり400万を超える11-mer プローブの完全なセットを用いる分析が可能となる。好ましい態様において、HyChip (商標) システムを利用する配列決定方法は別々の非関連の DNA 断片の混合物の配列決定に用いられる。

【 0 2 3 2 】

30

本発明の配列決定方法に利用するためのDNA サンプルは PCRによって調製することが出来る。

【 0 2 3 3 】

好ましい側面において、本発明は1平方ミクロン当たり約1スポットの密度にて光学的に清浄な 表面上にランダムに配置された数100万の個々の ポリヌクレオチド 分子のアレイを提供する。これらのポリヌクレオチド 分子は蛍光-タグ付加プローブプールのハイブリダイゼーションおよびライゲーションのための鋳型として作用する。一つの態様において、プローブプールはDNA リガーゼと混合されランダムアレイに提示される。 プローブが標的断片上の隣接部位にハイブリダイズすると、それらは互いにライゲーションして安定なハイブリッドを形成する。発達した光学を用いる高感度メガピクセル CCD カメラを利用して同時に全アレイ上の数100万のこれら個々のハイブリダイゼーション/ライゲーション事象が検出できる。第一のプール対からのシグナルがいったん検出されるとプローブは除かれ、連続的ライゲーション サイクルを用いて異なるプローブ組合せが試験される。本発明の好ましい側面において3.2 x 3.2 mm アレイ は1000万の断片、またはおよそ 10-100 億のDNA 塩基を保持する能力を有する。

40

【 0 2 3 4 】

標識化タグを用いるコンビナトリアル標識化

一つの側面において、単一のハイブリダイゼーション/ライゲーションサイクルは、16 蛍光色の使用によりすべての16の可能なプローブを試験するのに用いることが出来る。かかる試験はより少ない蛍光色からの蛍光サインを作成する方法を用いて達成することも出

50

来る。蛍光インサイチュ ハイブリダイゼーション (FISH) 染色体「ペインティング」において、蛍光プローブの組合せを用いてプローブのかかる組合せについての新しい蛍光サインを作ることが出来る。例えば、4のセットからの2つのプローブの組合せは10の可能なサイン蛍光シグナルを作ることが出来、5は15、6は21を作ることができ、以下同様である。それゆえ、単一のハイブリダイゼーションサイクルにおいて 16 プローブのうちどれがアンカープローブにハイブリダイズしたかを識別するのが可能であろう。

【0235】

あるいは、BBNNNNNN プローブの1つが非標識化のままであれば (それはその他のすべてのプローブについてのシグナルの欠如により推測される)、5 色が残りの 15 ジヌクレオチドのすべての標識に十分であろう。4 色を単一の塩基を読み取る4 プローブの標識に用いることが出来、あるいは2 塩基を読み取る(16のすべての必要なプローブのなかから) 8 プローブの標識に用いることが出来る。後者の場合すべての16のプローブは2サイクルにおいてスコア付けすることができる (以下参照)。したがって5または6 色 システムは非コンビナトリアル標識化に必要である 16 色と比較してより実行が容易であり得る。

【0236】

効率的なコンビナトリアル標識化のために、2-mer プローブをタグ結合部位を含むテイル配列を有するよう調製してもよい。テイル配列は 5 (または 6) 標識化オリゴヌクレオチドタグのなかから2に結合するようコンビナトリアルに設計され得、あるいは1または2の 蛍光色素を有する16 タグが16 テイルのそれぞれについて合成されうる。直接的に標識化されたプローブではなく標識化タグを使用することはさらなる利点を有する。すべての16 BBNNNNNN プローブを試験するには単一のプローブと比べて約 1024-倍 多くのプローブが必要であるであろう(ライゲーション部位から遠い位置での低い識別を仮定して)。例えば、BBNNNNNNのプローブ混合物内に1 μM 濃度にてプローブ AGCTANNNを有するには、混合物は1024 μM でなければならない。標識化プローブは非標識化プローブよりも合成にコストがかかるので、テイル配列により非標識化プローブが検出され得、標識化タグプローブは低濃度で利用できる。というのはそれはテイル配列に完全に相補的であり得るからである。さらに、非標識化テイル化プローブの使用はより低いバックグラウンドの維持にも有用である。というのはフルオロフォアは低濃度であるからである。全部で100-倍のコスト削減が、6 標識化タグ(縮重塩基を有さない)を同等の1024 標識化プローブのかわりに用いることによって期待される。

【0237】

タグは単一のライゲーション反応においてすべての16 ジヌクレオチドを読み取るために4 色のみを使用する効率的な 選択肢を提供する。かかる態様において、4の別々のタグの2セットがそれぞれ8つの2-merをデコード化ために設計されうる。すべての16の2-merは2のデコード化サイクルにおいてデコードされうる。この戦略はアダプターの各末端の2 塩基を読み取るための同じ4 色の使用に拡張されうる。この場合、4群の4タグを、4 塩基を読み取る各ライゲーション サイクルのための4 デコード化 工程に用いる。複数ライゲーションサイクルの代わりに複数デコード化 サイクルを行うことは より安価であり(酵素使用量がより少ない)、およびライゲーションサイクルをより長期に延長でき、プローブ濃度を低くでき、不一致 ライゲーションが低減されうる。

【0238】

タグはまた例えば天然塩基と対形成しないisoC およびisoG 塩基対の使用により分析されるDNAによる干渉を最小化するように設計され得る。別の選択肢は標準的 DNA 化学の使用であるが、ヒト ゲノムにおいて非常に稀な配列を設計 することである。 さらに別の選択肢はライゲーションの後および標識化タグとのハイブリダイゼーションの前に除去される非標識化タグとプレ-ハイブリダイズしたテイルを有するプローブの使用である。

【0239】

デコードされうる塩基の数の拡張

アンカープローブからの2を超えるヌクレオチドを読み取るために、本発明のいくつかの側面においてさらなるラウンドのプローブ-アンカー ライゲーションを使用し、次のサ

10

20

30

40

50

イクルの開始の前に標的からアンカー/標識 プローブを除去することができる。ライゲーションしたプローブ-アンカーは多数の当該技術分野に知られた方法を用いて除去でき、方法としては、加熱、またはアンカープローブにおける温度 または光切断可能な結合によるものが挙げられ、加熱 工程においてアンカーが断片化され不安定化される。配列決定すべき塩基はアダプターから3および4 塩基であるので、修飾をアンカープローブまたは標識化プローブに対して行う必要がある。アンカープローブの場合、それは本発明の一つの態様においてライゲーション末端に2 のさらなる縮重塩基を有するよう調製されうる。次のライゲーション効率の維持を保証するために、一つの態様において、アンカーは鋳型 DNA上の2つのより短いオリゴヌクレオチドのライゲーションを介して設計される。あるいは、配列決定プローブはライゲーション末端の2つの縮重塩基を用いて以下のように調製されうる：NNBBNNNN-タグ。本発明の別の側面において、アッセイは16 アンカープローブを用いてさらなる2 塩基を読み取るよう設計され得る。

10

【0240】

プローブ-アンカー ライゲーションの特異性は非常に高い。というのはライゲーション部位のまわりの2-4 塩基のみが試験されるからである。かかる塩基についての平均識別度は50-100 倍である。いくらかの不一致、例えば、GTがかなり強いが、識別度は たった5-20 倍である。本発明のある態様において、特定の不一致の識別における相違を考慮するソフトウェアが提供される。

【0241】

本発明のある側面において、各プローブ、アンカー およびタグは最適化され（例えば、濃度、縮重塩基の数、タグの配列および長さにより）、完全一致シグナルが最大に同等とされる。プローブおよびアンカーのオーバーラップし、シフトした対が本発明の一つの態様において設計され得、塩基呼び出し精度を高めるために各塩基を2-3回読み取るようにされる。

20

【0242】

正確な短い距離におけるアンカー 領域内のさらなる内部アダプターの挿入はゲノム 断片内の規定された位置での塩基の配列決定能力を拡張する。例えば、8 塩基空間離れて元のもの+2のさらなるアダプターを有することにより、10 サイクルにおいて20の連続的塩基を決定することが可能となり、これは5の連続する アダプター末端からの4 塩基の読み取りによる。

30

【化9】

最初のアダプター 最初の8塩基 アダプター2 第二の8塩基 アダプター3 さらに~200塩基
 DDDDDDDDDDDCGGGGGGGGDDDDDDDDDDGGGGGGGGDDDDDDDDDDGGGGGGGGGG
 AAAAAAA.BBNNNNNN-テイル AAAAAAA.BBNNNNNN-テイル AAAAAAA.BBNNNNNN-テイル
 AAAAAAA.NNBBNNNN-テイル AAAAAAA.NNBBNNNN-テイル AAAAAAA.NNBBNNNN-テイル
 テイル-NNNNBBNN.AAAAAAA テイル-NNNNBBNN.AAAAAAA
 テイル-NNNNNNBB.AAAAAAA テイル-NNNNNNBB.AAAAAAA

D= アダプター、G= ゲノム DNA、A = アンカー、B= 特定のプローブ塩基、N= 縮重プローブ塩基

【0243】

40

複数のアダプターはまた、読み取り能力をさらに高める機会およびサイクル当たり2を超える塩基を決定する能力を提供する。一つの態様において、サイクル当たり4-12 塩基が同定される。別の態様において、サイクル当たり4-8 塩基が同定される。さらに別の態様において、サイクル当たり12-16またはそれを超える塩基が決定される。

【0244】

一つの態様において、3 アダプターを12 塩基離れて位置させることにより、5 末端のそれぞれにおいて6 塩基を読み取ることにより得られる連続的配列は30 塩基となる。別の態様において、全部で 4 アダプターおよび 2つのアダプターの間の16 塩基の読み取りにより28 サイクルにおいて56 塩基の連続的配列が生じる。別の態様において、24 塩基を読み取る2つの(最初+1のさらなる) 16 塩基離れたアダプターが用いられる。

50

【 0 2 4 5 】

一つの態様において、各アンカー部位に用いられる16 ジヌクレオチド プローブの同じセットを有するが、各アンカー部位を別々に読み取る複数またはすべてのアンカー部位へのプローブの同時のハイブリダイズによりサイクルあたり複数塩基が同定される。一つの態様において、この同時のプローブライゲーションは異なる融解温度を有するアンカーを設計し、複数のあらかじめ規定された温度にて色強度を測定することにより達成される。

【 0 2 4 6 】

別の態様において、複数のアダプターがサイクル的プライマー伸長に用いられ、各個々のプライマーからより少ないサイクルにおいてより長い読み取りを提供する。

【 0 2 4 7 】

一つの態様において、マッピング情報は全 DNA 断片に分布する十分な数の 短い配列をスコア付けすることにより正確な位置でのより少数の短い配列から得られ、位置情報は用いられない。このプロセスの改変法は「ハイブリダイゼーション サイン」と称され、ここで予測強度と観察強度が比較される。別の態様において、短い配列は局所(間欠的または連続的) 配列情報を提供するように設計されてもよい。かかる短い配列の3つの例を以下に模式的に表示する：

【 化 1 0 】

a. (X)aBB(X)bBB(X)cBB(X)dBB(X)cBB(X)f...

b1. BBX6BBX4BBX6BBX4BBXa....

b2. B16Xa

【 0 2 4 8 】

完全なマッピング情報に必要なオリゴヌクレオチド配列の数は標的配列のサイズ、用いるDNA 断片のサイズおよび ソース DNAの複雑性に依存する。ヒトおよびその他の同様に複雑なゲノムについて 約 5の ポジティブ 8-merまたは10のポジティブ 6-merが100 塩基 DNA 断片に十分であり得る。2 サイクルにおいて1つのポジティブ 8-merをスコア付けするために、全部で約 10 サイクルを用いることが出来、これはアンカー配列決定よりも3-倍多くのサイクルの使用による。一つの態様において、このプロセスは2つのアンカーの挿入を利用せず、直接のハイブリダイゼーションを用いて酵素なしに行われうる。かかる態様において、3000の8-merが利用できる。

【 0 2 4 9 】

一つの態様において同一のセットのプローブが異なる群の組合せ(コンビナトリアル プール)において同一の標識を有するプローブのプールからのどのプローブがポジティブであるかをデコードするために用いられ得る。例えば、300の別々の標識により標識されたすべての 3000のプローブが同じプローブ 組合せにより標識された5 のプローブを有することにより2つの反応においてスコア付けされうる。6の真のポジティブに加えて、その他の30またはそれを超えるプールに関連する偽のポジティブがこれら2つの反応において見いだされる。プローブが別々にグループ化されるさらに2のハイブリダイゼーション サイクルを行うことにより、真のポジティブプローブのみがデコードされる。というのはそれらは2つのデータセット間でポジティブが共有されており、1未満の偽のポジティブプローブが共有されているからである。ポジティブプローブを見いだすことは各プローブについての2つのスコア付けの低い方を使用して行われ得る。真のポジティブプローブについては低い方のスコア付けは高いと予測される。ほとんどのネガティブプローブについては少なくとも一つの スコア付けは非常に低く、それは1つの偽のポジティブ スコア付けをキャンセルする。このプロセスはサイクルの数または要求される標識の数の低減を助け、コンビナトリアル標識化を使用する必要なく多くの用途に十分な力を与えうる。

【 0 2 5 0 】

別の態様において、分析される断片の高度にオーバーラップした セットは異なるプローブのサブセットを有する異なるサブアレイ上の2-16 サブセットの形態であり、大量のマッピング情報を提供する。例えば、平均して各塩基から出発する250 塩基 断片がプロ

10

20

30

40

50

ープの2-16の異なるサブセットを有する2-16 サブセットとして分析されうる。2-26 塩基のみシフトしたDNA 断片は少なくとも一つのプローブ サブセットを有する特有の染色体同定を提供するすべてではなくても数個の使用したプローブ サブセットにより分析される。

【 0 2 5 1 】

典型的には、20の特異的塩基はほとんどの特有の配列に必要な情報を提供する。一つの態様において、この情報は5 サイクルにおいて5x4 塩基の読み取りについての256 タグを有する2つのアンカー、またはサイクル 当たり8 塩基を読み取ることにより24 塩基についての3 サイクル(512 タグ付加組合せ)を用いてえられうる。別の態様において、3 サイクルx 6 塩基=18 塩基(5x3+3、距離 20-30 塩基)、およびさらに別の態様において 3-mer について4分の1 少ないタグは、3 アンカーを必要としうる(3x6 + 3+3 塩基)。

【 0 2 5 2 】

一つの側面において、高用量DNA アレイ プラットフォームが100の患者 またはその他の DNA サンプルを同時に分析するのに利用できる。マッピングの直接のハイブリダイゼーション (またはコンビナトリアル ライゲーション) アプローチにおいては、プローブのサブセットのみが使用され、自動的にタグ配列を提供しない。4-塩基 タグについてはすべての 256 プローブ(例えば、NxUxBBBBUXNx)がマッピングに利用され得、または さらなるプローブとして利用されうる。これらプローブがマッピングに利用されると、256 のシフトしたプローブの複数のセットが、タグ配列の同定に必要であり得る。

【 0 2 5 3 】

一つの側面において、5-6 色が 16のすべてのジヌクレオチドのデコード および1つのデコード化 サイクルにおける 2-12 塩基の読み取りに用いられる。一つの態様において、4 タグのセットが用いられる; 別の態様において、セットは6 タグに拡張される。複数デコード化 サイクルは単独でまたは異なる融解温度を有するアンカーと組合せて、単一の デコード化 サイクルにて読み取られうる塩基の数の増加に利用されうる。

【 0 2 5 4 】

一つの側面において、ライゲーション サイクル当たり4 塩基がアダプターの各末端の2 塩基の試験および2つの対応する アンカーの使用により読み取られる。プローブの両方のタイプ、 B2N6-テイル およびテイル-N6B2を同時に利用できる。各プローブ タイプは特有のテイルおよび6の特有のタグの一致するセットを有しうる。6 タグの2 セットを用いる2のデコード化 サイクルにより4 塩基が同定される。11 ライゲーション サイクルにおいて、42 連続および 2 重複塩基が決定される。42+18=60 塩基のメイトペアを読み取るために、15 ライゲーション サイクルが必要とされる。

【 0 2 5 5 】

別の側面において、ライゲーション サイクル当たり8 塩基が読み取られる。全部で 4 アンカーが利用できる(2つのアダプターの2つの側のそれぞれ)。プローブおよびタグは第一の選択肢におけるものと同ーであってよい。したがって2のデコード化 サイクルにおいて1のアダプターの各側の2 塩基が決定されうる。さらなる2 アンカーを第二のアダプターに用いることが出来るので、さらなる情報が2つのポジティブ 2-merのいずれがアンカー/アダプター末端に属するのかを識別するために必要である。これはより高い融解温度 (Tm)を有する第二のアダプターについての2つのアンカーの設計により達成されうる。したがって、模式的に、4 アンカーは以下の通り:

【 化 1 1 】

アダプター 1	アダプター 2
...GGGGDDDDDDDDDDDDDDDDGGGGGGGGGGGGDDDDDDDDDDDDDDDDDDDDDDDDGGG	
GGG...	
AAAAAAAAA AAAAAAAAA	AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA

D= アダプター 塩基、G = ゲノム 塩基、A = アンカー 塩基。

【 0 2 5 6 】

デコード化および5-6 色素のイメージングの2の標準的 サイクルの後、ストリンジェ

10

20

30

40

(サイクル 3)

50

ン サイクルにて 2 から8または12 塩基になる。 さらに好ましい態様において、隣接するアダプター間の16 塩基が読み取られ、最初 + 2 挿入アダプターのための使用により、連続的配列の40 (2x16+8) 塩基の決定能力が導かれる。

【0261】

多重プローブ-アンカー ライゲーションアッセイ

一つの側面において、テイルがおよそ15から 20 塩基の長さでありフルオロフォアにより標識されたテイル にタグ配列が相補的である、構造 BBNNNNNN-テイルの16 プローブを含むプローブセットが調製される。テイルおよびタグは分析する DNAによる干渉を最小化するように設計される。一つの態様において、テイルおよびタグ配列はiso-cおよびiso-gヌクレオチドから調製され、タグ配列が鋳型 DNAと相互作用するのを妨げる。

10

【0262】

異なるテイルおよびタグ配列を有する異なるBBNNNNNN-テイル プローブの効率を試験することが可能である。16テイル配列が要求されうるが、16 プローブ (16の異なるテイルを有する) のうち8つのみが各デコード化 サイクルにおいて分析されうる。というのは、4-色混合物の最大能力は 2の可能性のある組合せの10 であるからである (可能なプローブ指標としてヌルシグナルを含めない)。各テイル配列は2のタグに結合する能力を有し得、この設計における各タグは 1つのみのフルオロフォアを結合して有し得る。4 タグのセットの最初の 設計は、各色について1が実行されうる。これらタグの相補的配列を組み合わせる8 テイル(全部で 10の可能な組合せのなかから8)を作ることができる。16 テイルの残りの 8は4 タグのさらなるセットを必要としうるが それらは、4 タグの第一のセットに用いられたものと同じフルオロフォアを担持しうる。

20

【0263】

一つの側面において、プローブは単一の フルオロフォア (例えば、TAMRA)を有するように調製され得、異なるタグ組合せの相対的強度(すなわち ハイブリッド強度)が測定される。いったんこの情報が得られると、フルオロフォアをタグに一致させ強度を規準化することが可能である。タグの単一のフルオロフォアセットを用いて共通テイル 構造を有するプローブのBBNNNNNN 領域の相対的効率を測定することもできる。かかるパラメーターがいったん測定されると、16の BBNNNNNN-テイル プローブのセットが調製されうる。このプローブセットは PCR由来のRCR 産物および合成標的 環または複雑なゲノムサンプルへの ハイブリダイズに利用できる。

30

【0264】

一つの態様において、アレイとされた RCR 標的はまずアダプター プローブとハイブリダイズされDNB 位置および相対強度が判定される。このプローブを標準的 技術、例えば、温度上昇により除き、プローブの第二のセットをアレイにハイブリダイズさせうる。第二の プローブセットはアンカープローブおよび16の BBNNNNNN-テイル プローブをライゲーション混合物中に含む。反応は 十分な時間、好ましくは 約 30 分間進行し、ライゲーションしていない、ハイブリダイズしていない プローブは洗浄除去される。チャンバへの次の付加は、ライゲーションし、 ハイブリダイズした BBNNNNNN プローブのテイルにハイブリダイズする4 タグ プローブを含みうる。このハイブリダイゼーションはある態様において5 分間の短さでよく、 高いシグナル 強度が達成される。チャンバを再び洗浄し、イメージングを所望の波長で行う。チャンバは次いで加熱を経てタグが除去されるが、アンカー-BBNNNNNN-テイル プローブはハイブリッドに維持される。第二の群の4 タグは第二の 群の8の BBNNNNNN プローブの存在のスコア付けのためにハイブリダイズされうる。一致BBNNNNNN プローブ 及びその他の15 不一致 BBNNNNNN プローブの間の識別のレベル はシグナル強度のレベルおよび組合せを介して判定されうる。

40

【0265】

一つの態様において、プローブ-アンカー ライゲーション アッセイを確立するために、プローブ、例えば構造 AANNNNNNのプローブが提供され、TATA 配列について低いΔGを有するAATATANN DNA スポットのための十分なシグナルが生じる。最適条件のためのシグナルがある DNA 配列について低ければ、一致プローブを独立に調製し、これらプローブ

50

についてのみ選択的に濃度を高めるように混合物に添加するとよい。第一の 4 縮重位置での256のうち20 配列が調整の必要があれば、16x20のさらなるプローブを調製すればよい。

【0266】

一つの態様において2のアダプターの間ゲノム セグメントの他方の側からの2-塩基配列を読み取るための16 プローブの開発および試験が達成される。これらプローブについてのテイル および縮重塩基は5' 末端であり得、例えば、 テイル-NNNNNNBBである。

【0267】

本発明の一つの側面において、識別されうる色素の数を多重特異的励起 パターンおよび最大数の各励起パターンのためのフィルターの利用により最大化する。例えば、それぞれ4の異なる波長を有する2-4 励起(全部で16 波長)を各励起に付き8-16 フィルターと組合せて利用できる。アルゴリズムおよびソフトウェアを用いて強度パターンを分析し、8-24 色素のそれぞれからのシグナル量を推定する。

【0268】

一つの態様において、色素による直接標識化をハプテン (例えば、 ビオチン)を用いる間接 標識化と組み合わせて、多重プローブを特異的に染色する。直接的に結合させた色素は光退色してもよく、強度における相違は染色の前後に計算するとよい。

【0269】

一つの態様において、利用可能な色標識の数をクエンチャーの光または化学的脱保護または所与の色素の吸収をシフトさせる化学的修飾によって拡張させる。色強度は脱保護処理の前後に測定される。第一の イメージングが行われた後、色素は所与の波長についてのシグナル上昇が測定される前に光退色されうる。 多重 タイプのクエンチャーまたは修飾因子(3-4-6)および8 色により、全部で 24-48 非コンビナトリアル 標識が作成できる。24-48 標識のなかの2によるコンビナトリアル 標識化は 276-1128の 2 -標識 組合せの可能性を与える。

【0270】

長い安定なアンカーは異なる標的に対するプローブハイブリダイゼーション およびライゲーションを改善しうる。一つの態様において、縮重塩基の数を増加させて不安定ハイブリッドを形成する標的配列、 例えば、 5' TATA3'の影響を最小化する。これはプローブ/標的 ハイブリッドの安定性を高めうるがライゲーション部位に近い第一の 2-4 位置に完全に一致しないプローブが標的にハイブリダイズし得、ライゲーションを妨げる。このネガティブな影響を最小化するために、 1つの態様はより高い出発温度 および/または温度 サイクリングを提供し、アンカーに隣接してハイブリダイズしたライゲーション可能なプローブの数を増加させる。

【0271】

プライマー伸長を用いる配列決定

末端配列決定を特異的に標識されたヌクレオチドを用いる単一の塩基 伸長の多くの連続する サイクルにより1つのアンカー/プライマー末端から行うことができる。一つの態様において、プロセスは色素またはブロッカーが除去されて伸長を繰り返す工程を含む。複数のアダプターはこのプロセスの増加した柔軟性を提供する。一つの態様において、2-6またはそれより多い塩基が連続反応においてシフトしたプライマーを使用することにより単一の塩基 プライマー伸長により読み取られる。1つのアダプター上の複数の同時のシフトした0+1または1+1 プライマーフレームまたは複数のアダプター上の単一のフレームまたはその両方を利用できる。

【0272】

一つの態様において、最初 + 3 のさらなるアンカーの使用は、4 プライマーを提供する。各プライマーの4 塩基を読み取ることにより、16 塩基が16 サイクルで4 標準的色を用いて決定され、これはコンビナトリアル 標識化またはタグ付加を行わずに達成できる。この態様において、プライマー伸長は標識化成分上に縮重塩基を有さず、したがって使用される色素濃度が低減される。16 塩基はマッピングに十分ではないことがあるので、2

10

20

30

40

50

0-24 サイクルにおける4 プライマー x 5-6 塩基の伸長を利用するとよい。

【0273】

多重プライマー伸長はプライマーの弁別的 除去により可能である。数種の方法が以下を含む因子に基づく除去に利用できる：プライマー長、GC含量、塩基またはバックボーン修飾、例えば、LNAまたはPNA、ウラシル組込み、あるいは選択された塩基間の光感受性結合。1の群の2～8の安定性レベルが設計され得る。また異なる安定化剤または保護剤を有する2から4の別々の群が利用できる。これら標識化方法の適用により、20-24塩基が3-5 酵素サイクルといった少ないサイクルにて決定されうる。別の態様において1度に1塩基の多重プライマー伸長のためのプライマー 保護 アッセイが用いられる。かかる態様において、プライマー、例えば UUUUUUUNNN、が 第四の伸長に利用され、十分なシグナルを提供する。というのはNNNでの不一致は標的の50% または90%より多くを占有し得、効率的に伸長されないからである。より高い特異性を有するプライマーはUUUUUUU.UUUUNNN または UUUUUUU.UUNNNNNをライゲーションすることにより作ることが出来る。

10

【0274】

一つの側面において、アンカーの各側の配列決定を可能とするために、結合した ssDNAを結合したプライマーおよび 元の鎖の除去または プライマー 侵入技術を用いてdsDNAに変換するとよい。元の 鎖を除く1つのアプローチは、1つの鎖のみを切断する制限酵素のための挿入されたアダプター 結合部位を組み込むことである。断片化された鎖は変性され洗浄除去される。

【0275】

20

連続またはオーバーラップするフレームの実行または2-3 塩基の読み取りのために、異なるアンカー および/または プローブ 設計を利用できる。例えば：

【化13】

サイクル 1: UUUUUUUUUUUU.BBNNNNNN
 サイクル 2: UUUUUUUUUUNN.BBNNNNNN または
 UUUUUUUUUUUU.NNBBNNNN
 サイクル 3: UUUUUUUUUUNN.NNBBNNNN

ここでUは共通のあらかじめ規定された 塩基、特異的塩基 および Nは縮重塩基を示す。

【0276】

30

縮重塩基を有するアンカーは縮重位置にて一致する塩基を有するアンカーの優先的結合を確実にするために二部にて設計され得る。オーバーラップまたはシフトしたフレームを用いて同じ標的において各塩基を複数回読み取ることが出来る。アンカーの後の第一の4塩基の複数の読み取りの2例を以下に示す：

【化14】

UUUUUUUUUUU.UBBNNNNN
 UUUUUUUUUUUU.BBNNNNNN
 UUUUUUUUUUN.BBNNNNNN

UUUUUUUUUUUU.NNBBNNNN
 UUUUUUUUUUNN.BBNNNNNN
 UUUUUUUUUUN.BBNNNNNN

40

ここで Uは共通の あらかじめ規定された 塩基、Bは特異的 塩基およびNは縮重塩基を表す。ライゲーション部位は点(.)を付して示す。

【0277】

検出装置

本発明の一つの側面において、ハードウェア が配列決定方法のライゲーションおよびハイブリダイゼーション事象の検出を可能とするために提供される。一つの態様において、システム ハードウェアは3つの主な成分を含む；照明システム、反応 チャンバ、および検出システム。検出装置はいくつかの 特徴、例えば、以下のものを含みうる：調整可

50

能な レーザー出力、電子的シャッター、オートフォーカス、および実行ソフトウェア。

【0278】

本発明により作られたランダムアレイ上の単一の分子からシグナルが生じ得、多数の検出システムにより検出され得、検出システムとしては、これらに限定されないが、走査型電子顕微鏡、近接場走査型光学顕微鏡 (NSOM)、全内部反射蛍光 顕微鏡 (TIRFM)等が挙げられる。豊富なガイダンスがナノスケール構造を表面上で分析および検出するためのかかる技術の適用についての文献にみられ、引用により本明細書に含まれる以下の参考文献に例証されている: Reimer et al, editors, Scanning Electron Microscopy: Physics of Image Formation and Microanalysis, 2nd Edition (Springer, 1998); Nie et al, Anal. Chem., 78: 1528-1534 (2006); Hecht et al, Journal Chemical Physics, 112: 7761-7774 (2000); Zhu et al, editors, Near-Field Optics: Principles and Applications (World Scientific Publishing, Singapore, 1999); Drmanac, 国際公報 WO 2004/076683; Lehr et al, Anal. Chem., 75: 2414-2420 (2003); Neuschafer et al, Biosensors & Bioelectronics, 18: 489-497 (2003); Neuschafer et al, 米国特許 6,289,144; 等。特に興味深いのはTIRFMであり、例えば、 Neuschafer et al, 米国特許 6,289,144; Lehr et al (上記に引用); およびDrmanac, 国際公報 WO 2004/076683に開示されている。

【0279】

一つの側面において、本発明のアレイの使用のための 装置は3つの基本成分を含む:(i) 検出およびプロセッシング試薬、例えば、プローブ、洗浄溶液等を保存し、アレイに移動させる流体工学システム; (ii) 流れおよび温度制御能力を有するアレイを担持または含む反応 チャンバ、またはフローセル;および (iii)照明および検出システム。一つの態様において、フローセルは温度 を約 5-95 の範囲、より具体的には10-85 の範囲に維持する能力を有し、温度を約 0.5-2 /秒の速度で変化させることが出来る温度制御サブシステムを有する。

【0280】

一つの側面において、1" 平方の170 μm の厚さのカバーガラスのためのフローセルであって本発明の高分子構造を結合するように誘導体化したものを利用することが出来る。セルはガラスを挟むことにより「アレイ」を封入し、 2つの平面の間のガasketを含む。1つの平面はイメージングを可能とする十分なサイズの開口部を有し、カバーガラスのためのインデックスポケットを有する。もう1つの平面はガasketのためのインデックスポケット、流体ポートおよび温度制御 システムを有する。1つの流体ポートはシリンジポンプに連結しており、それは流体 をフローセルから「引き」または「押す」ものであり、もう1つのポートは 漏斗様混合チャンバに連結している。チャンバは、液体レベルセンサーを備えている。溶液は漏斗を通して分配され、必要により混合され、フローセルへと流される。レベルセンサーがフローセルにつながった漏斗内の空気を検知するとポンプが逆となり既知の量の流体が漏斗へと戻される。これにより空気がフローセルに入るのが妨げられる。カバーガラス表面は切断され、 ストリップに分かたれていてもよく、ストリップは流体の流れ/挟むことにより生じた毛細管作用を収納する。かかる基体は「オープンエア」 / 「片側ガラスの」チャンバに収納され得、毛細管流動作用を排除することによりバッファが基体の上を均一に流れることを促進する。イメージングは100x対物レンズにより達成され得、 TIRFまたは落射照明およびZeiss axiovert 200上の1.3メガピクセル Hamamatsu orca-er-agなどの システムを用いる。この配置は基体 (配列されていないアレイ)にランダムに結合されたRCR コンカテマーを画像化する。イメージング速度は、対物レンズ拡大率の低下、格子パターン化アレイの使用、および各イメージに収集されるデータのピクセルの数の上昇により改善できる。

【0281】

一つの態様において、4または5以上のカメラを用いることが出来、好ましくは 10-16メガピクセルの範囲である。 多重バンド通過フィルターおよび二色ミラーを用いて4または5以上の発光スペクトルにわたってピクセルデータを収集することが出来る。拡大率の低下による低い集光力を補償するために、励起光源の出力を上昇させるとよい。スルー

ブットは一以上の流体チャンバを各カメラとともに使用することにより上昇できそれによりイメージング システムは空回りせず、サンプルはハイブリダイズ/反応される。アレイのプロービングは非連続的でありうるので、二以上のイメージング システムを用いてアレイのセットからデータを集めるとよく、さらにアッセイ時間が削減できる。

【 0 2 8 2 】

イメージング プロセスの際に、基体は焦点があっているままであるのが好ましい。焦点を維持するいくつかの重要な因子は基体の平坦性、基体の焦点面に対する直交性、およびそれを変形させうる基体への機械的力である。基体の平坦性は十分に制御され得、1/4波平坦度より良好な平坦度を有するガラスプレートが容易に入手できる。基体に対する不均一な機械的力はハイブリダイゼーション チャンバの適切な設計を介して最小化できる。焦点面との直交性はよく調整された、高精度ステージにより達成することができる。オートフォーカスルーチンは一般にさらなる作動時間を有するので、必要な場合にのみ用いるのが望ましい。好ましい態様において、各イメージが獲得され、高速アルゴリズムを用いて分析され、イメージの焦点があっているかが判定される。イメージの焦点があってなければ、オートフォーカスルーチンを作動させる。システムは次いで目的物の Z 位置情報を保存し、それを用いて 次のイメージング サイクルの際にそのアレイのそのセクションに戻らせる。目的物のZ 位置を基体上の様々な位置でマッピングすることにより、基体のイメージ 獲得に必要な時間を削減することができる。

10

【 0 2 8 3 】

一つの側面において、蛍光に基づくシグナルのための好適な照明および検出 システムは 80 ミリワット532 nm固相レーザーに共役したTIRFスライダーを備えたZeiss Axiovert 200である。スライダーは基体を対物レンズを介して正しい TIRF 照明角度で照明する。TIRFは基体に光学的に結合したプリズムを介して基体を照射することにより対物レンズを使用しなくても達成することが出来る。平面波ガイドも基体上のTIRF の実行に利用できる。落射照明も利用できる。光源はラスターされ得、位相のそろった、位相のそろっていない、単一または複数スペクトルソースから生ずるビームを発する。

20

【 0 2 8 4 】

イメージング システムの1つの態様は1.25 mm 視野の20x レンズを含む。 10 メガピクセル カメラが 検出に用いられる。かかるシステムはパターン化されたアレイ に1 ミクロン ピッチで結合したおよそ150万のコンカテマーをイメージングすることができる。かかる配置の下で、コンカテマー当たりおよそ 6.4 ピクセル存在する。コンカテマー当たりのピクセルの数は対物レンズの視野を上昇または低下させることにより調整することが出来る。例えば、1 mm 視野は10 ピクセル/コンカテマーの値を与え、 2 mm 視野は2.5 ピクセル/コンカテマーの値を与える。視野は対物レンズの拡大率および開口数に応じて調整でき、最小のピクセル数/コンカテマーが得られ、それは光学およびイメージ分析 ソフトウェアにより分解することが出来る。

30

【 0 2 8 5 】

TIRFおよびEPI 照明の両方においてほとんどあらゆる光源を用いることが出来る。1つの照明スキームは 共通のセットの単色光源 (6-8 色につき約 4レーザー)を提供し、これはイメージャー間で共有される。各イメージャーは異なる波長で所与の時間データを収集し、光源は光学スイッチシステムを介してイメージャーへとスイッチされる。かかる態様において、光源は好ましくは少なくとも 6、より 好ましくは 8 種類の波長を生じる。かかるソースには、ガスレーザー、ファイバー結合器を介して組み合わせられた多重ダイオードポンプ固相レーザー、フィルターされたキセノンアルゴンクランプ、波長可変レーザー、またはより新規なSpectralum Light Engine (Tidal Photonicsにより提供予定) が挙げられる。Spectralum Light Engine はプリズムを用いて光をスペクトル的に分離する。スペクトルはTexas Instruments Digital Light Processor上に投影され、それはスペクトルのあらゆる部分をファイバーまたは光コネクタに選択的に反射することができる。このシステムは個々の波長を横切る 出力をモニターおよび較正することができ、それを一定に維持し、強度相違を電球の古さとしてまたは電球交換の間に自動的に補償する。以

40

50

下の表は可能なレーザー、色素 およびフィルターの例を表す：

【表 1】

レーザー	励起フィルター	発光フィルター	色素	
407nm	405/12	436/12	Alexa-405	401/421
407nm	405/12	546/10	カスケードイエロー	409/558
488nm	488/10	514/11	Alexa-488	492/517
543nm	546/10	540/565	Tamra	540/565
543nm	546/10	620/12	Bodipy 577/618	577/618
	546/10	620/12	Alexa-594	594/613
635nm	635/11	650/11	Alexa-635	632/647
635nm	635/11		Alexa700	702/723

10

【0286】

一つの側面において、イメージングは100x 対物レンズを介して達成される。励起光源は 80 ミリワット ダイオードポンプ固相 レーザーである。この光源はTIRFM および落射照明技術との使用が成功している。イメージは1.3 メガピクセル Hamamatsu orca-er-ag カメラ および Ziess axiovert 200 倒立顕微鏡を用いて獲得される。この配置は現在、基体にランダムに結合したDNBを0.5 秒の露光時間でイメージする。

【0287】

20

複数ハイブリダイゼーションサイクルを扱うために、反応 チャンバと検出 システムとの両方に完全に統合されたロボットステーションが本発明での使用に用いられ得る。落射蛍光が標的部位当たり10-20を超える蛍光分子の検出に利用できる。落射蛍光を用いる利点は複数色のプローブを標準的顕微鏡により使用することを可能とすることである。

【0288】

一つの側面において1" 平方、170 μm 厚さカバーガラスを収容し、DNBを結合するよう誘導体化および活性化された2ピースフローセルが使用される。側方ポートは流体をフローセルから「引き」または「押す」シリンジポンプと連結している。第二のポートは液体レベルセンサーを備えている漏斗様混合チャンバに連結している。溶液は混合チャンバへと分配され、必要により混合され、フローセルへと流される。レベルセンサーがフローセルにつながった漏斗内の空気を検知するとポンプが逆となり既知の量の流体が漏斗へと戻される。これにより空気がフローセルに入るのが妨げられる。このチャンバはカバーガラスの大きさの基体のために良好に作動し、より大きい基体のために改変形態で用いることも出来る。ハイブリダイゼーション チャンバ およびイメージングサブシステムと統合されたかかる三軸ロボットガントリーピペット操作 システムを完全に自動化されたプローブ ピペット操作のために機能化することができる。

30

【0289】

基準

一つの態様において、規則的パターンの捕捉セルは各獲得イメージへと位置 情報をコード化するよう解釈される。およそ 1000 セル/イメージが パターンから除かれて10 ビットコードを作り、それは各基体上の1024までの名付けられた位置を表しうる(図5)。

40

【0290】

コード化領域の物理的 特徴をイメージ 分析の際にすべてのピクセルをイメージに位置づけるための参照として用いることが出来、コードそれ自体は装置が基体の正しい領域をイメージングしたことを確認するために用いられる。コード化領域の重要な特徴は各要素が非結合スポットである「空の領域」ブロックによって表されていることである。これにより蛍光 マーカーの 基体上での必要性が排除される。所与のプローブセットについてポジティブなRCR 産物は各要素の境界を規定する。これは領域 が表面に結合したRCR 産物の5%から10% のみが 所与のプローブプールについてポジティブである場合でも認識可能であることを意味する。一つの態様において、コードは各コード要素 が 50の捕捉セルを

50

表す場合に読み取り可能である。

【0291】

本発明のキット

本明細書に記載する方法の商業化において、本発明のランダムアレイの構築およびそれを様々な用途に用いるための特定のキットが特に有用である。本発明のランダムアレイの適用のためのキットは、これらに限定されないが、標的ポリヌクレオチドのヌクレオチド配列の決定のためのキットを含む。キットは典型的には少なくとも一つの表面を有する支持体および本発明のランダムアレイの構築またはその適用の実施に必要なまたは有用な一以上の試薬を含む。かかる試薬には、限定するものではないが、核酸 プライマー、プローブ、アダプター、酵素等が含まれ、それぞれ容器、例えば、限定するものではないが、バイアル、チューブまたはボトルにパックされ、販売に好適なパッケージ、例えば、限定するものではないが、箱、シールされたパウチ、プリスターパックおよび大型容器に梱包される。パッケージは典型的にはパッケージされた材料の使用を示す標識またはパッケージング挿入物を含む。本明細書において用いる場合、「パッケージ材料」はキットにおける試薬の分配のための梱包に使用されるあらゆるものを含み、限定するものではないが、容器、バイアル、チューブ、ボトル、パウチ、プリスターパッケージング、標識、タグ、指示用シートおよびパッケージ挿入物が含まれる。

【0292】

別の側面において、本発明は以下の成分を含む標的ポリヌクレオチドの配列決定のためのキットを提供する： (i) 光学的に分解可能な別々の空間的に離れた領域のアレイを有する平面状表面を有する支持体、ここで各別々の空間的に離れた領域の面積は1 μm^2 未満である； (ii) 別々の空間的に離れた領域にランダムに配置された複数のコンカテマーにハイブリダイズするためのプローブの第一のセット、各コンカテマーは標的ポリヌクレオチドのDNA 断片の複数コピーを含む； および (iii) 複数のコンカテマーにハイブリダイズするためのプローブの第二のセット、第一のセットからのプローブが第二のセットからのプローブに連続的にハイブリダイズする場合、それらプローブはライゲーションされる。かかるキットはさらに、リガーゼ、リガーゼ バッファー、およびハイブリダイゼーション バッファーを含んでいてもよい。ある態様において、別々の空間的に離れた領域は捕捉オリゴヌクレオチドを結合して有し得、コンカテマーはそれぞれ捕捉オリゴヌクレオチドに対して相補的な領域を有し得、それにより該コンカテマーは捕捉オリゴヌクレオチドと該コンカテマーの相補的領域の間の複合体形成により別々の空間的に離れた領域に結合されることができる。

【0293】

別の側面において、本発明はDNA 断片を環状化するためのキットを含む。例示的な態様において、かかるキットは以下の成分を含む： (a) 一以上の DNA 断片へのライゲーションおよびそれとDNA 環を形成するための少なくとも一つのアダプター オリゴヌクレオチド、 (b) ホモポリマー テイルを該 DNA 断片に結合させ、該 アダプター オリゴヌクレオチドの第一の末端のための結合部位を提供するターミナルトランスフェラーゼ、 (c) 該 アダプター オリゴヌクレオチドの鎖を該 DNA 断片の末端にライゲーションし該 DNA 環を形成させるリガーゼ、 (d) 該 アダプター オリゴヌクレオチドの鎖の領域にアニーリングするためのプライマー、および (e) ローリングサークル複製 反応において鎖にアニーリングしたプライマーを伸長させるためのDNA ポリメラーゼ。さらに好ましい態様において、上記アダプター オリゴヌクレオチドは 4 から12の範囲の多数の縮重塩基を有する第二の末端を有していてもよい。上記キットはさらにターミナルトランスフェラーゼ、リガーゼおよびDNA ポリメラーゼのための反応 バッファーを含んでいてもよい。

【0294】

さらに別の側面において、本発明はCircLigase (商標) 酵素 (Epicentre Biotechnologies, Madison, WI) を用いてDNA 断片を環状化するためのキットを含み、このキットは体積 排除 ポリマーを含む。さらに好ましい態様において、キットは以下の成分を含む： (a) pH制御およびCircLigaseのための最適化塩組成を提供するための反応 バッファー、およ

び(b) CircLigase 補因子。別の側面において、かかるキットのための反応バッファは、0.5 M MOPS (pH 7.5)、0.1 M KCl、50 mM MgCl₂、および 10 mM DTTを含む。別の側面において、かかるキットは CircLigase、例えば、10-100 μ L CircLigase 溶液 (100 ユニット/ μ L)を含む。例示的な体積排除 ポリマーは引用により本明細書に含める米国特許 4,886,741に開示されており、ポリエチレングリコール、ポリビニルピロリドン、デキストラン硫酸および類似のポリマーを含む。一つの側面において、ポリエチレングリコール (PEG)は50% PEG4000である。一つの側面において、環形成のためのキットは以下を含む：

【表 2】

量	成分	終濃度
2 μ L	CircLigase (商標) 10X 反応 バッファ	1X
0.5 μ L	1 mM ATP	25 μ M
0.5 μ L	50 mM MnCl ₂	1.25 mM
4 μ L	50% PEG4000	10%
2 μ L	CircLigase (商標) ssDNA リガーゼ (100 ユニット/ μ L)	10 ユニット/ μ L
	一本鎖DNA 鋳型	0.5-10pmol/ μ L
	滅菌水	

10

最終反応体積： 20 μ L

20

【0295】

上記成分は当該技術分野に知られた多数の異なるプロトコールにおいて利用でき、例えば以下のプロトコールが挙げられる：(1) DNAの長さに応じてDNAを60- 96 で加熱する (5'-リン酸および 3'-ヒドロキシル基を有するssDNA 鋳型)；(2) 2.2X 反応混合物を60 で 約 5-10 分間予備加熱する；(3) DNAが96 まで予備加熱されたら60 まで冷却する。DNAと バッファを 60 で混合し、冷却せず、2-3時間インキュベートする；(4) 酵素を加熱により不活性化しライゲーション反応を終結する。

【0296】

本発明は以下の非限定的な実施例を参照するとより良好に理解されうる。実施例は、本発明の例示のために提供される。以下の実施例は本発明の好ましい態様をより完全に説明するためのものであり、本発明の広い範囲を制限するよう解釈してはならない。

30

【実施例】

【0297】

実施例 1: DNBのRCRに基づく形成および結合

2つの合成標的を共増幅した。約100万分子をガラス表面上に捕捉し、標的の1つについてプローブした。イメージングおよび第一のプローブの光退色の後、第二の標的をプローブした。アンプリコン特異的プローブによる連続するハイブリダイゼーションは、アレイ上の各スポットが2つのアンプリコン配列のいずれか1つに特有に対応していることを示した。70 への加熱によりプローブが除去され、再ハイブリダイズして同等に強いシグナルを生じること確認された。

40

【0298】

実施例 2: 環形成および増幅の確証

環形成および増幅プロセスを大腸菌 DNAを用いて確証した(図6)。捕捉プローブの結合部位およびRCR プライマーとしても作用するユニバーサルアダプターを、すべてのゲノム配列への結合のための縮重塩基を含むユニバーサル 鋳型 DNAを用いて標的分子の5' 末端にライゲーションした。標的分子 の3'末端をターミナルトランスフェラーゼを用いてポリ-dA テイルの付加により改変した。改変した標的を次いでアダプターおよびオリゴ-dA テイルに相補的な架橋鋳型を用いて環状化した。

【0299】

実施例 3: 凝集したコンカテマーとのライゲーションの確証

50

凝集したコンカテマーとのプローブライゲーションが起こる能力を試験した。反応はリガーゼを用いて20 で10分行い、次いでチャンバを簡単に洗浄して過剰のプローブを除いた。6-merおよび標識化 5-merのライゲーションにより11-merと匹敵するシグナルレベルが生じた。ランダムアレイのイメージ分析を含むソフトウェアモジュールを、全ゲノム配列再構築のためのシュミレーションデータに対して試験した。

【 0 3 0 0 】

実施例 4: 単一のアレイを用いる複数の病原体からの標的の同定

*Bacillus anthracis*および*Yersinia pestis*の診断領域からのPCR 産物を一本鎖DNAに変換し、ユニバーサルアダプターに結合させた。これら2 サンプルを混合し、RCRを用いて共に複製し、チップ表面上にランダムアレイとして沈着させた。アンブリコン 特異的プローブによる連続的ハイブリダイゼーションは、アレイ上の各スポットが2つのアンブリコン配列のいずれか1つに特有に対応していること、および、プローブによりそれらが特異的に同定できることを示し(図7)、それによりRCR 反応によって生じたDNA 断片の約 10 0-1000 コピーを有するマイクロミリより小さいサイズの DNA ナノボールに存在するDNA を同定するための感受性および特異性が実証された。

【 0 3 0 1 】

*B. anthracis*からの155 bp アンブリコン配列および*Y. pestis* からの275 bp アンブリコン配列を標準的 PCR 技術を用いて増幅し、ここで対の1つのプライマーがリン酸化されたPCR プライマーを用いた。PCR 産物の一本鎖形態をリン酸化した鎖のラムダエキソヌクレアーゼを用いる分解により作成した。残っている鎖の5'末端をT4 DNA ポリヌクレオチド キナーゼを用いてリン酸化し、一本鎖産物のユニバーサルアダプターへのライゲーションを可能とした。ユニバーサル アダプターをT4 DNA リガーゼを用いて標的分子の5'末端にライゲーションし、これを標的の5'末端 およびユニバーサル アダプターの3'末端に相補的な鋳型 オリゴヌクレオチドにより補助した。標的にライゲーションしたアダプターを次いでアダプターおよび標的の3'末端に相補的な塩基を有する架橋オリゴヌクレオチドを用いて環状化した。直鎖状 DNA 分子をエキソヌクレアーゼ Iによる処理により除去した。RCRを一本鎖サンプルの混合およびPhi29 ポリメラーゼの使用により行い開始プライマーとして架橋オリゴヌクレオチドを用いて環状化アダプター-標的分子の周囲を複製した。RCR 産物を、誘導体化ガラスカバーガラスに結合し、ユニバーサル アダプター配列に相補的な捕捉オリゴヌクレオチドを介してガラス スライド上に捕捉した。

【 0 3 0 2 】

B. anthracis および*Y. pestis* PCR アンブリコン由来のアレイとした標的ナノボール分子を連続的にユニバーサルアダプター配列に相補的なTAMRA-標識化 11-mer プローブまたは2つのアンブリコン配列の1つに相補的な11-mer プローブによりプローブした。3つのプローブの連続的ハイブリダイゼーションから得られたイメージを重ねることにより(図7)、アダプタープローブにハイブリダイズしたアレイとした分子のほとんど(青色スポット)はアンブリコン 1 プローブ (赤色スポット)またはアンブリコン 2 プローブ (緑色スポット)のいずれかにのみハイブリダイズし、両方にハイブリダイズするものはほとんどないことが理解できる。この特異的ハイブリダイゼーション パターンは、アレイ上の各スポットが*B. anthracis* アンブリコンまたは *Y. pestis* アンブリコンのいずれかの1タイプの配列のみを含むことを示す。それはまたrSBH プロセスが配列特異的プローブの使用によりアレイに沈着した異なる配列の標的分子を識別することが出来ることを示した。

【 0 3 0 3 】

実施例 5:縮重塩基を有する80-mer オリゴヌクレオチドから生じたアレイ化DNBにおける塩基位置のデコード化

縮重塩基を含む合成オリゴヌクレオチドの個々の分子を4の亜集団に分け、そのそれぞれは特定の位置にA、C、G またはT 塩基のいずれかを有する。この合成DNAから作成したDNBのアレイは塩基のそれぞれを有する約 25%のスポットを有しうる。4の連続する4 塩基のそれぞれに特異的なプローブの対のハイブリダイゼーションおよびライゲーションによ

り亜集団が同定された(図8)。

【0304】

5'リン酸化された3'TAMRA-標識化ペンタマーオリゴヌクレオチドは4つのヘキサマーオリゴヌクレオチドの1つと対形成した。これら4つのライゲーションプローブ対のそれぞれはA、C、GまたはT-含有バージョンの標的のいずれかにハイブリダイズする。3を超える識別スコアがほとんどの標的から得られ、ナノボール標的の間の単一の塩基の相違を同定する能力が示された。識別スコアは同じスポットのその他の3塩基-特異的シグナルの平均により割った最高のスポットスコアである。アッセイ条件(バッファー組成、すべての成分の濃度、サイクルにおける各工程の時間および温度)の調整により、バックグラウンドに対しより高いシグナルが得られ、完全一致の不一致に対する比の計算が可能となった。

10

【0305】

同様のライゲーションアッセイを6-mer プローブのスポットしたアレイに対して行った。この場合 完全一致/バックグラウンド比は約 50であり、平均完全一致/不一致比は30であった。結果はさらに連続する プローブ サイクルの数の増加または各サイクル当たり異なる色素で標識された4またはそれ以上のプローブの使用により、DNBに存在するDNAの部分的または完全 配列を決定する能力を示した。

【0306】

亜集団を同定するために、4塩基のそれぞれに特異的な4ライゲーションプローブのセットを用いた。配列 CAAAC (プローブ T1A9b)を有するT1Aの位置 33-37に対応する5'リン酸化された3'TAMRA-標識化ペンタマーオリゴヌクレオチドは位置 27-32に対応する以下のヘキサマーオリゴヌクレオチドの1つと対形成した: ACTGTA (プローブ T1A9a)、ACTGTC (プローブ T1A10a)、ACTGTG (プローブ T1A11a)、ACTGTT (プローブ T1A12a)。これら4ライゲーションプローブ対のそれぞれはA、C、GまたはT含有バージョンのT1Aのいずれかとハイブリダイズするはずである。各ハイブリダイゼーションサイクルについて、プローブをT4 DNA リガーゼを含むライゲーション/ハイブリダイゼーションバッファー中でアレイとともに20で5分間インキュベートした。過剰のプローブを20で洗浄除去し、イメージをTIRF顕微鏡で取得した。結合したプローブは次のラウンドのハイブリダイゼーションの準備のために剥がした。

20

【0307】

アダプター 特異的プローブ (BrPrb3)がアレイにハイブリダイズし、すべてのスポットの位置を確立した(図8)。0.4 μMの4ライゲーションプローブ対を次いで連続的にアレイにハイブリダイズさせた: A-特異的ライゲーションプローブ対にハイブリダイズしたスポットを図5で赤色で示し、C-特異的スポットは緑色、G-特異的スポットは黄色およびT-特異的スポットはシアンで示す。図5において、環AはアダプタープローブおよびA-特異的ライゲーションプローブ対の両方にハイブリダイズしたスポットの1つの位置を示し、このスポットにアレイとされたDNAが位置32にAを含むT1Aの分子由来であることが示唆される。明らかにほとんどのスポットは4ライゲーションプローブ対のいずれか1つのみに結合し、位置32における塩基の実体を特異的に決定することが可能となった。

30

【0308】

社内(in-house)イメージ分析プログラムを用いて、スポットをアダプタープローブを用いてハイブリダイゼーションサイクルから取得したイメージを用いて同定した。塩基-特異的ライゲーションプローブを用いて同じスポットを同定し、蛍光シグナルを次のサイクルについて定量した。識別スコアを各スポットに対する各塩基-特異的シグナルのそれぞれのシグナルについて計算した。識別スコアは、同じスポットのその他の3塩基-特異的シグナルの平均により割ったスポットスコアである。各スポットについて、最高の4塩基-特異的識別スコアを次に高いスコアと比較した。2つの比が1.8を超えると、最大識別スコアに対応する塩基を塩基呼び出しのために選択した。この分析において500を超えるスポットの塩基-呼び出しが成功し、平均識別スコアは3.34であった。平均完全一致シグナルは272であり、平均単一不一致シグナル(非選択塩基からのシグナル)は83.2で

40

50

あった。したがって完全一致 / 不一致 比は3.27であった。イメージ バックグラウンドノイズをランダムに選択した空のスポットからのシグナルを定量することによって計算し、これら空のスポットの平均 シグナル は82.9であった。したがって完全一致 / バックグラウンドノイズ比は3.28であった。これらの実験において不一致識別は完全一致シグナルがバックグラウンドに対して低いことにより制限された。

【 0 3 0 9 】

実施例 6: プローブ-アンカー ライゲーション アッセイを用いる合成 80-mer オリゴヌクレオチドの末端の2 縮重塩基のデコード化

5' 末端に8 縮重塩基を含む合成オリゴヌクレオチドを用いてランダムゲノム DNA末端をシュミレーションした。このオリゴヌクレオチドから作られたDNA-ナノボールはアダプター配列に直近して配置したこれら8 縮重塩基を有する。既知のアダプター配列に隣接する2の未知の塩基のプローブ-アンカー ライゲーション アプローチを用いた配列決定の実行可能性を示すために、アダプター配列の3'末端にハイブリダイズする特異的配列を有する12-mer オリゴヌクレオチドをアンカーとして使い、BBNNNNNNの形態の16 TAMRA-標識化オリゴヌクレオチドのセットを配列読み取りプローブとして用いた。

【 0 3 1 0 】

BBNNNNNN プローブセットのサブセット (即ちBBの位置にGA、GC、GG およびGT)を用いて、これら4 プローブの1つに特異的に結合する標的から作られたナノボール アレイ上にスポットを同定することが出来、平均完全一致/不一致 比は20を超えた(図9)。

【 0 3 1 1 】

実施例 7: 構造化されたナノボールアレイの作成

平均して5 μ m離れた捕捉プローブの配列されたアレイの線を調製した。線は5 μ mのチップサイズに対して45度の角度の引っ張られたガラスキャピラリーを用いて作り、キャピラリーには水中の1 μ lの5 μ M 捕捉プローブを充填し、精密なガントリーロボットによりガラススライド上に線を引いた。DNBはカバーガラスの表面に結合し、アダプターに特異的なプローブにより検出された。図10は領域に対する高密度結合を示し、ここで捕捉プローブが表面上に沈着されており、マイクロミリより小さい 結合部位を有する基体が調製されたらDNBを格子状に配置できることが示された。

【 0 3 1 2 】

実施例 8: 複数のアダプターを有する環形成の実証

長さ70 塩基の合成標的 DNAおよび長さ200-300 bp のPCR由来断片をプライマーの1つのリン酸化により二本鎖 産物から得、ラムダエキソヌクレアーゼによる処理によりリン酸化した鎖を除去した。一本鎖断片を環状化のためにアダプターとライゲーションした。重合、Hls型制限酵素消化および新しいアダプターとの再-ライゲーションを本明細書に記載するように行った。

【 0 3 1 3 】

プロセスが成功したことの实証は最終的に得られた環のRCR 増幅を用いて達成した。簡単に説明すると、DNA 環を最後に導入されたアダプター に相補的なプライマーおよびPhi 29 ポリメラーゼとともに1 時間30 でインキュベートし、元の DNA 環の数百の反復コピーからなる単一のコンカテマー 分子を作成した。RCR 産物のカバーガラス表面への結合は、表面に結合したオリゴヌクレオチドに相補的なコンカテマー中のアダプター配列の利用により達成することが出来た。アダプター特有プローブのハイブリダイゼーションを用いて個々の アダプターが環に、そして最終的にRCR 産物に組み込まれたことを実証した。アダプターが環の中の予測した位置に組み込まれたことを実証するために、配列特異的プローブ (標識化5-mer)を合成またはPCR 由来配列のために使い、アダプターの末端配列を認識する非標識化 アンカープローブとのライゲーションが起こるようにした。クローニングおよび配列決定もDNA 完全性の確認に用いた。プロセスを清浄な ssDNAを各環切断の後に作成することにより簡便にし、同じ閉環化学を各アダプター組込みに使用することを可能とした。

【図面の簡単な説明】

【 0 3 1 4 】

【図 1】図1A-1Gは、本発明およびその適用を図示する。

【図 2】図2A-2Gは、散在したアダプターを含む標的ポリヌクレオチドを作るために核酸断片にアダプターを挿入する様々な方法を図示する。

【図 3】図3A-3Eは、散在したアダプターを含む標的ポリヌクレオチドに対して実行されるハイスループット配列決定方法を図示する。

【図 4】図4は、RCR 産物の付着により作られる構造化されたDNA アレイと標準的ランダム DNA アレイとの比較を提供する。

【図 5】図5は、配列されたアレイ上の参照パターンを図示する。

【図 6】図6は、rSBH 装置上に画像化されたランダムアレイを示す。

【図 7】図7は、より容易にみるためにわずかにシフトさせて重ね合わせた3つのアレイ画像を示す。

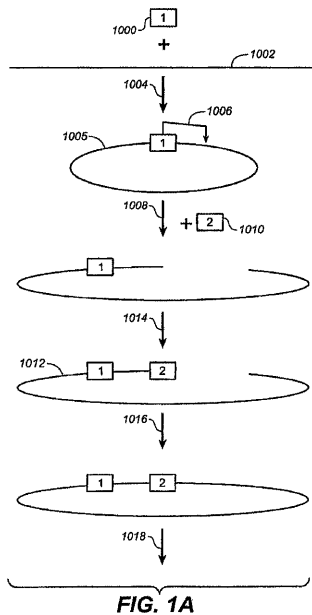
【図 8】図8は、わずかにシフトさせて重ね合わせた5つのアレイ画像を示す。

【図 9】図9は、わずかにシフトさせて重ね合わせた5つのアレイ画像を示す。

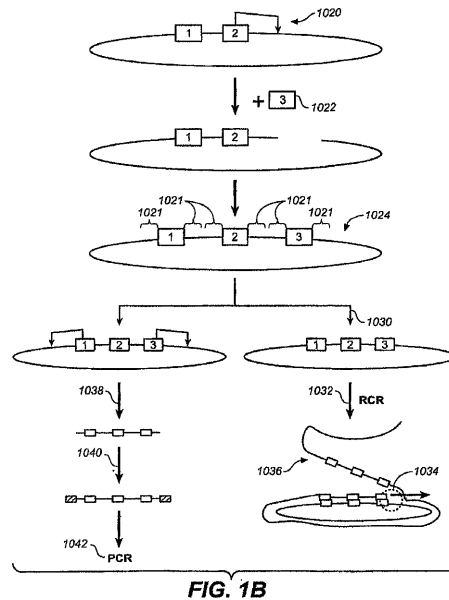
【図 10】図10は、カバーガラスの表面を横切る線状に並んだ捕捉プローブがDNBへの特異的結合に用いられたアレイの画像を示す。

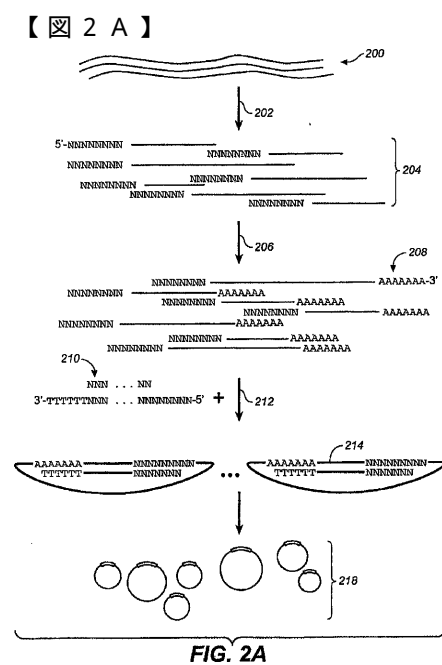
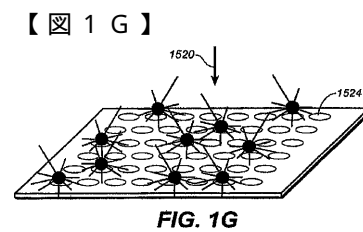
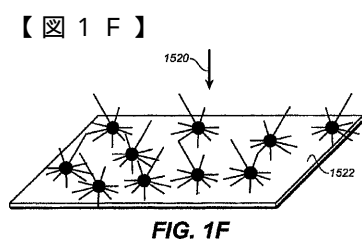
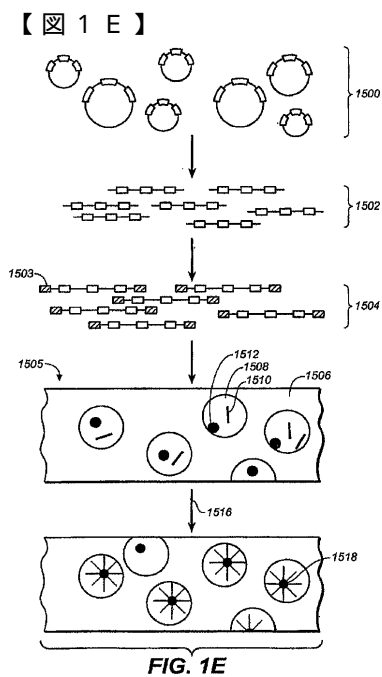
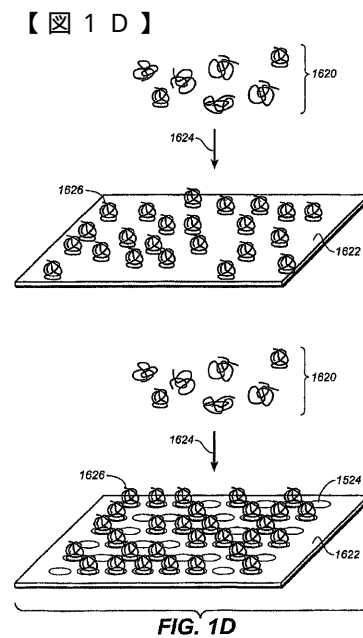
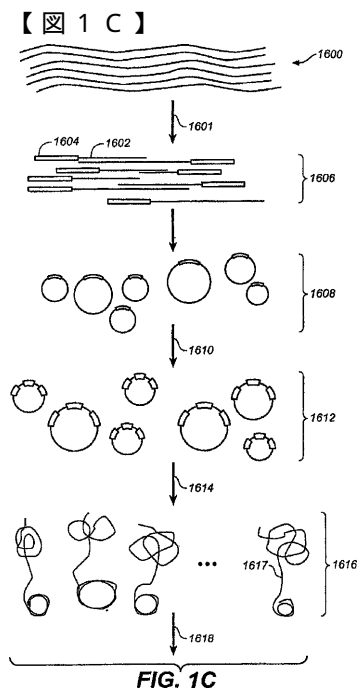
10

【図 1 A】

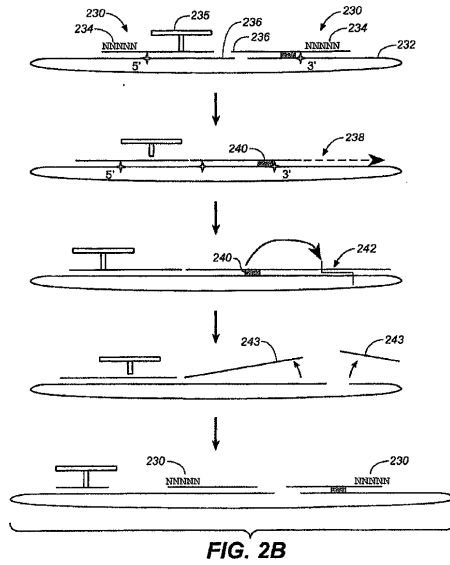


【図 1 B】

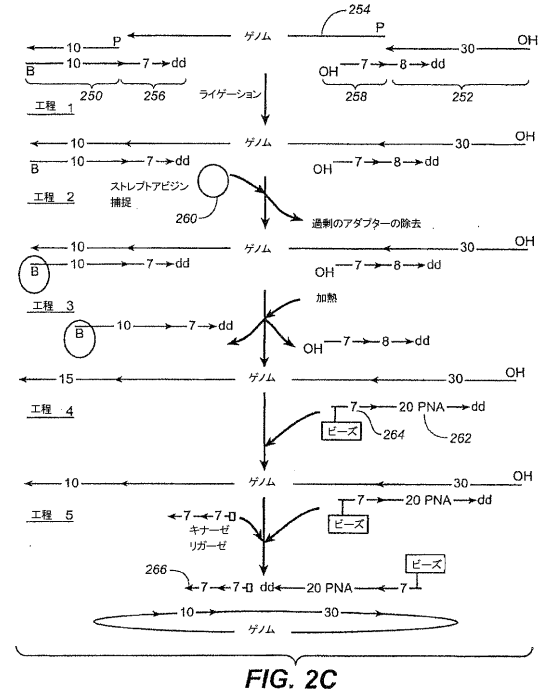




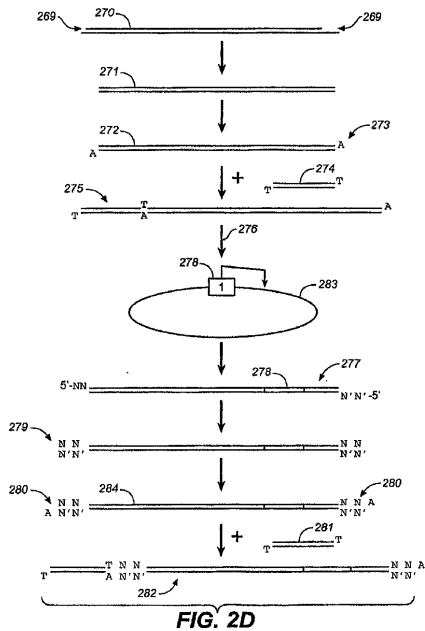
【図 2 B】



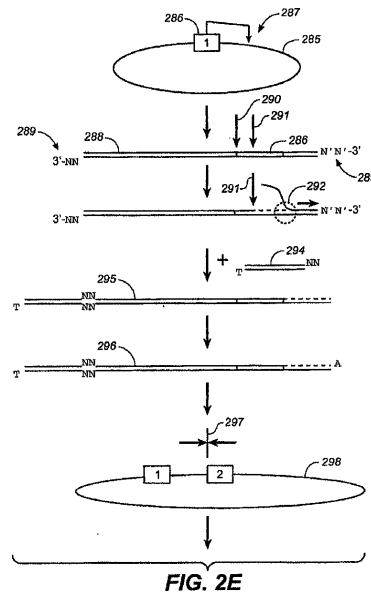
【図 2 - C】

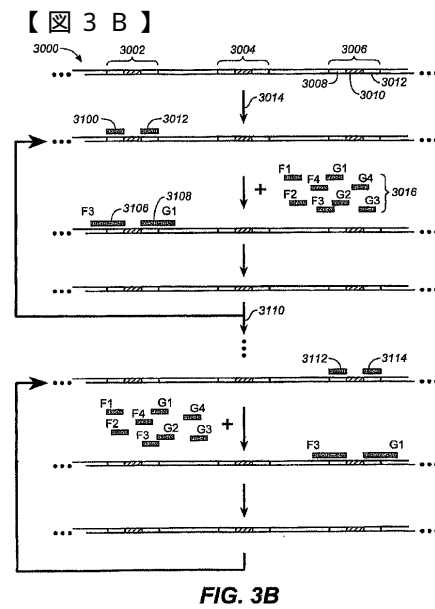
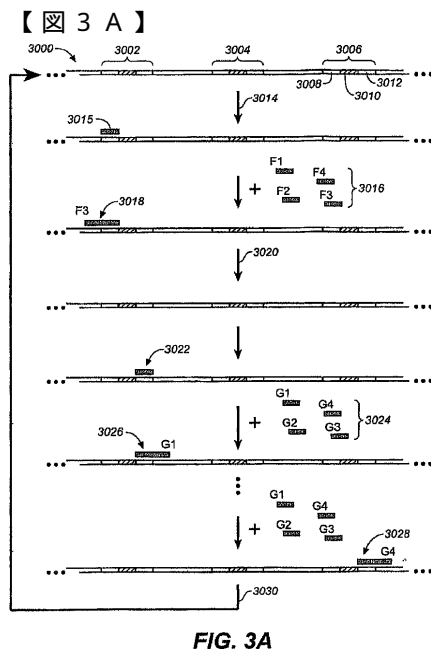
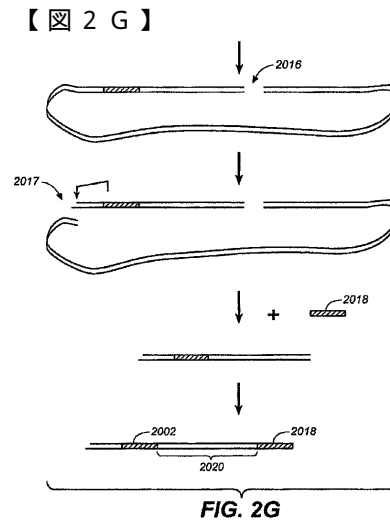
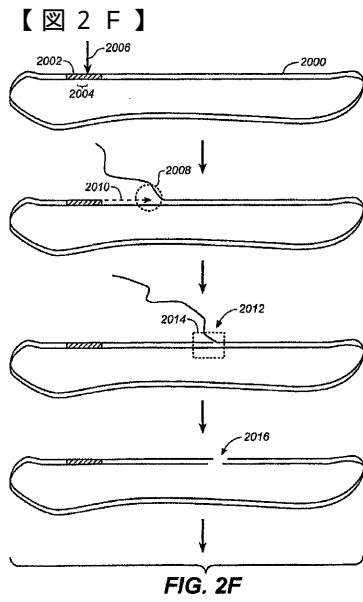


【図 2 D】



【図 2 E】





【図 3 C】

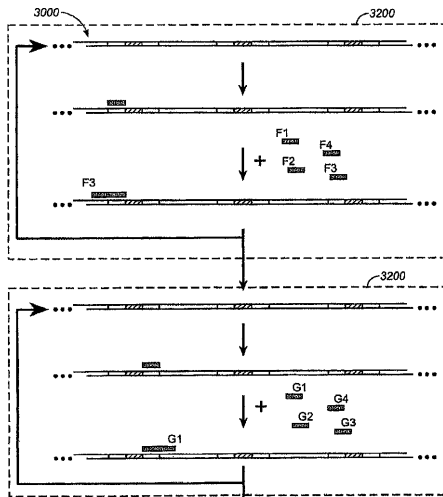


FIG. 3C

【図 3 - D】

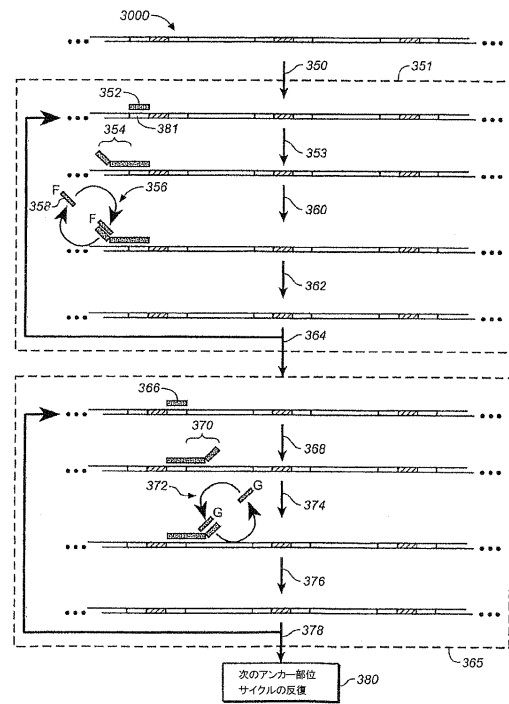


FIG. 3D

【図 3 E】

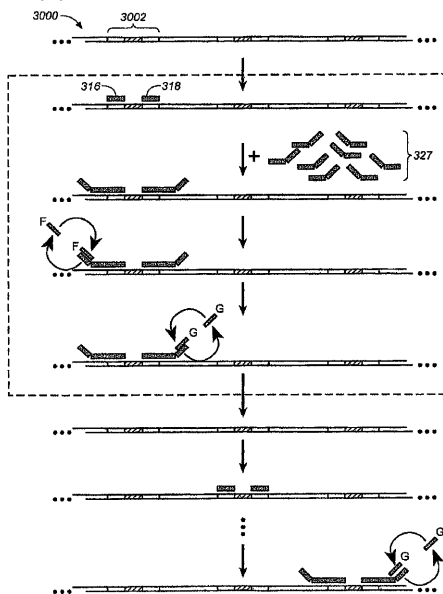


FIG. 3E

【図 4】

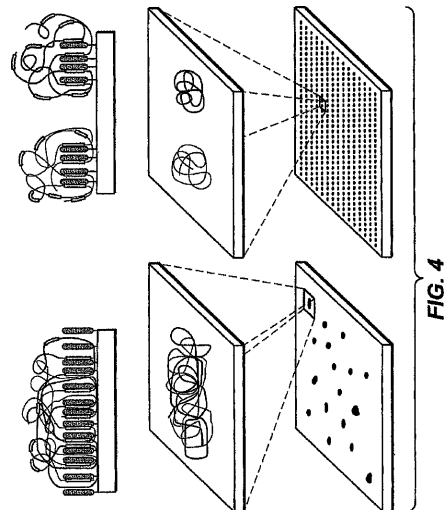


FIG. 4

【図 5】

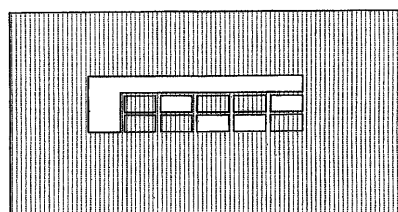


FIG. 5

【図 6】

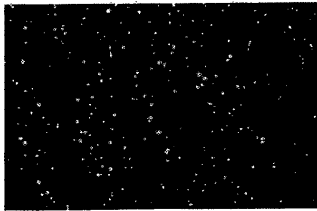


FIG. 6

【図 7】

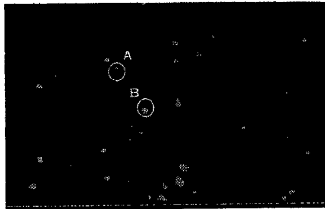


FIG. 7

【図 8】

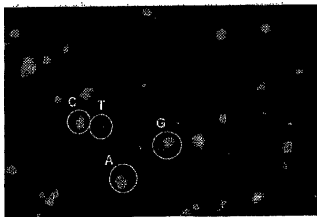


FIG. 8

【配列表】

0005180845000001.app

【図 9】

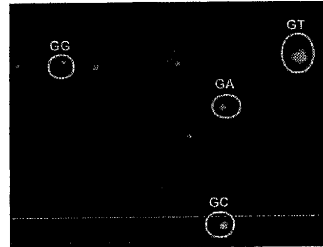


FIG. 9

【図 10】

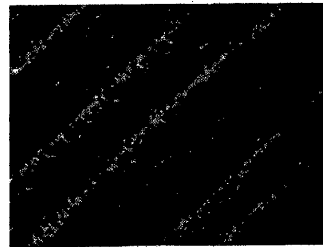


FIG. 10

フロントページの続き

(51)Int.Cl. F I
G 0 1 N 33/53 M

(74)代理人 100127638

弁理士 志賀 美苗

(74)代理人 100138911

弁理士 櫻井 陽子

(74)代理人 100146259

弁理士 橋本 諭志

(72)発明者 ラドジェ・ドラマナック

アメリカ合衆国 9 4 0 2 2 カリフォルニア州ロス・アルトス・ヒルズ、レッド・ロック・ロード 2
7 6 3 5 番

(72)発明者 マシュー・ジェイ・カロウ

アメリカ合衆国 9 4 0 6 5 カリフォルニア州レッドウッド・シティ、バッテン・レイン 3 番

(72)発明者 スネザナ・ドラマナック

アメリカ合衆国 9 4 0 2 2 カリフォルニア州ロス・アルトス・ヒルズ、レッド・ロック・ロード 2
7 6 3 5 番

審査官 引地 進

(56)参考文献 特開 2 0 0 2 - 0 8 5 0 9 7 (J P , A)

国際公開第 2 0 0 4 / 0 8 5 6 7 8 (W O , A 1)

米国特許出願公開第 2 0 0 4 / 0 2 4 8 1 6 1 (U S , A 1)

国際公開第 2 0 0 6 / 0 4 0 5 4 9 (W O , A 1)

特開 2 0 0 3 - 3 7 2 1 8 8 (J P , A)

Proc. Natl. Acad. Sci. USA, vol. 88, pages 1247-1250 (1991)

(58)調査した分野(Int.Cl. , D B 名)

C12N 15/00-15/90

PubMed

JSTPlus(JDreamII)

CAPLUS/BIOSIS/WPI(STN)