

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-204302

(P2006-204302A)

(43) 公開日 平成18年8月10日(2006.8.10)

(51) Int. Cl.		F I		テーマコード (参考)
<b>C 1 2 Q</b> 1/68 (2006.01)		C 1 2 Q	1/68 Z N A A	4 B O 2 4
<b>C 1 2 N</b> 15/09 (2006.01)		C 1 2 N	15/00 A	4 B O 6 3

審査請求 有 請求項の数 1 O L (全 34 頁)

(21) 出願番号	特願2006-44556 (P2006-44556)	(71) 出願人	399125757 アフィメトリックス インコーポレイテッド
(22) 出願日	平成18年2月21日 (2006.2.21)		アメリカ合衆国 カリフォルニア州 95051 サンタ クララ セントラル エクスプレス ウェイ 3420
(62) 分割の表示	特願平11-261280の分割	(74) 代理人	100078282 弁理士 山本 秀策
原出願日	平成11年9月14日 (1999.9.14)	(74) 代理人	100062409 弁理士 安村 高明
(31) 優先権主張番号	60/100.393	(74) 代理人	100113413 弁理士 森下 夏樹
(32) 優先日	平成10年9月15日 (1998.9.15)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 完全N量体アレイを用いる核酸分析

(57) 【要約】 (修正有)

【課題】 ポリヌクレオチドの変異を特徴付ける方法の提供。

【解決手段】 参照ポリヌクレオチドのフラグメントおよび標的ポリヌクレオチドのフラグメントを完全もしくは実質的に完全なn量体セットを構成する一本鎖のオーバーハングを有する同一または実質的に同一なポリヌクレオチドプローブのアレイに、別々にライゲーションし、そしてそれらのハイブリダイゼーションパターンを比較する工程を包含する。この方法はまた、多型性の存在を検出するために使用され得る。また、未知の配列の2つ以上のポリヌクレオチドが同一であるかどうかを決定することができる。さらに、末端配列に基づいてポリヌクレオチドのフラグメントを列挙し、そして区別することもできる。これらの方法は、医学、薬理遺伝学、生化学、および法医学に有用である。

【選択図】 なし

## 【特許請求の範囲】

## 【請求項1】

標的ポリヌクレオチドの変異の存在を決定する方法であって、以下の工程：

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全なn量体のセットを構成するように、各プローブが、二本鎖領域および一本鎖n量体オーバーハング領域を含む、工程；

(b) 該標的ポリヌクレオチドを、1つのアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、標的ハイブリダイゼーションパターンを生成する工程；

(c) 参照ポリヌクレオチドを、第2のアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、参照ハイブリダイゼーションパターンを生成する工程；および

(d) 該参照ハイブリダイゼーションパターンと該標的ハイブリダイゼーションパターンとを比較することによって、該標的ポリヌクレオチド中の変異の存在を決定する工程、を包含する方法。

10

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

(関連出願)

本願は、1998年9月15日に出願された米国仮特許出願第60/100,393号からの優先権を主張し、その内容は本明細書中に参考として援用される。

20

## 【0002】

(発明の分野)

本発明は、ポリヌクレオチド配列の分析に関する。それゆえ、本発明は、分子生物学、生化学、化学、医学、および医学診断学などの多様な分野に関する。

## 【背景技術】

## 【0003】

ゲノム配列決定プロジェクトの目的は、参照ゲノムの完全配列を得ることである。次の、そしておそらくより困難な段階は、配列のバリエーションを分析すること、およびこの情報を重要な表現型に関連付けることである。一本鎖ポリヌクレオチドプローブの高密度アレイ(非特許文献1)は、大規模ポリヌクレオチド配列決定分析に高度に匹敵し、そして大規模に実現可能な(scalable)アプローチを提供する(非特許文献3)。遺伝学的分析についてのDNAアレイの有用性は、変異の検出、遺伝子型決定、物理的マッピングおよび遺伝子発現モニタリングを含む、多数の適用において以前に確認されている(非特許文献4~8)。

30

## 【0004】

いくつかのアッセイ形式において、固体支持体に固定された一本鎖ポリヌクレオチドプローブアレイが、標的ポリヌクレオチド中の特定のポリヌクレオチド配列を検出するために使用されてきた。例えば、特許文献1および特許文献2を参照のこと。VLSIPS<sup>TM</sup>技術の開発は、非常に小さい領域における一本鎖のポリヌクレオチドプローブの非常に大きいアレイの作製の方法を提供した。特許文献3および特許文献4および特許文献5(これらの各々は本明細書中で援用される)を参照のこと。特許文献6は、「DNAチップ」と呼ばれる一本鎖ポリヌクレオチドプローブの多数のアレイの微小製作の方法を記載し、そこではプローブはアレイのあらかじめ規定された位置に特定の配列を有し、特定の標的ポリヌクレオチドの同定を容易にするか、またはサンプル中の標的ポリヌクレオチドの1つ以上の特定の配列が、先に特徴付けた配列から変化しているかどうかを検出する。

40

## 【0005】

今日まで首尾よく使用されてきた一本鎖ポリヌクレオチドプローブのアレイの大部分は、あらかじめ規定された参照配列に基づいて特別に調製されたカスタマイズされたプローブのセットを使用してきた。より一般的には、アレイに基づく配列決定および遺伝的分析

50

(すなわち、変異および多型性分析)へのより効率よく、より迅速なアプローチは、所定の長さの一本鎖ポリヌクレオチドプローブの完全セット(「n量体セット」)を使用することであり得、その結果、この技術に存在する問題が打破され得る場合は、原理的にはいかなる標的配列も分析され得る。

【0006】

一本鎖ポリヌクレオチドの完全n量体プローブセットの開発において主要な問題が存在する。第1に、n量体( $4^n$ )の完全セットのプローブの総数が、「n」が増加するにつれて非常に大きくなり得る。技術的進歩がなされ、多数の核酸プローブアレイの合成が容易になったのはようやく最近のことであった。

【0007】

さらに、n量体プローブの完全セットへのハイブリダイゼーションに基づくDNA配列再構築を記載するいくつかの理論的な提案が提出されたが(非特許文献9~11)、これらの方法は成功しなかった。配列再構成は、複合ポリヌクレオチドにおける配列反復の存在、およびn量体の独自性の欠如から生ずる不明確さによって妨害されてきた。さらなる問題には、必然的に不完全であるハイブリダイゼーションデータを取り扱うためのコンピュータ化された方法、および分析的方法の不足が含まれる。

【0008】

特許文献7は、配列決定および分析のための定常的な二本鎖領域および可変的な一本鎖領域を有するポリヌクレオチドのアレイを記載する。特許文献7の方法はプローブアレイへの標的ポリヌクレオチドのハイブリダイゼーションを使用した。ハイブリダイゼーションに続いてライゲーションを行った。しかし、多くの工程を含む'980特許の方法論は煩わしく、そしてプローブの単一のアレイ上で多くの異なったサンプルを試験することを可能にするための標準化はされていない。

【0009】

例えば、特許文献7特許の方法は、プローブアレイをカスタマイズするために参照ポリヌクレオチドを使用すること、次いで標的ポリヌクレオチド中の変異を検出するためにそのカスタマイズされたプローブアレイを使用することを必要とした。結果として、特許文献7の方法論は、そのために設計された標的ポリヌクレオチドのみにおける変異を検出するために有用であるプローブアレイを産生する。さらに、この方法論は、単一の置換変異のみ(挿入または欠失ではなく)を検出し得、そしてそれが検出する変異を特徴付けし得ないし、またヌクレオチド配列中の変異の位置を同定し得ない。

【0010】

この出願の発明に関連する先行技術文献情報としては次のものがある。

【特許文献1】国際公開第WO89/10977パンフレット

【特許文献2】国際公開第WO89/11548パンフレット

【特許文献3】米国特許第5,143,854号明細書

【特許文献4】国際公開第WO90/15070パンフレット

【特許文献5】国際公開第WO92/10092パンフレット

【特許文献6】米国特許第5,837,832号明細書

【特許文献7】米国特許第5,503,980号明細書

【非特許文献1】Fodorら著、「Nature」、(英国)、1993年、364:555-6

【非特許文献2】Peaseら著、「Proceedings of the National Academy of Sciences of the United States of America」、(米国)、1994年、91:5022-6

【非特許文献3】Lipshutz著、「Journal of biomolecular structure & dynamics」、(米国)、1993年、11:637-53

【非特許文献4】Cheeら著、「Science」、(米国)、1996年、274:610-14

10

20

30

40

50

- 【非特許文献5】Lockhartら著、「Nature Biotechnology」、(英国)、1996年、14: 1675-80
- 【非特許文献6】Croninら著、「Human mutation」、(米国)、1996年、7: 244-55
- 【非特許文献7】Kozalら著、「Nature Medicine」、(英国)、1996年、2: 753-59
- 【非特許文献8】Wodickaら著、「Nature Biotechnology」、(英国)、1997年、15: 1359-72
- 【非特許文献9】BainsおよびSmith著、「Journal of theoretical biology」、(オランダ国)、1988年、135: 303-7 10
- 【非特許文献10】Lysov Iuら著、「Doklady Akademii nauk SSSR.」、(ロシア)、1988年、303: 1508-11
- 【非特許文献11】Drmanacら著、「Genomics」、(米国)、1989年、4: 114-28,

【発明の開示】

【発明が解決しようとする課題】

【0011】

ハイブリダイゼーションによる配列決定に固有の困難さを避ける一方、完全なn量体アレイの利点を保持するような配列分析の方法(例えば、変異の特徴付け、多型性配列の検出 20 など)が所望される。特に、所定の長さのポリヌクレオチドの標準化されたアレイの完全セットは、各々の標的に対する特異的なアレイを構築する必要性なしで、高度に所望される。このような標準化されたアレイ方法論は、さまざまな供給源由来の標的ポリヌクレオチド中における変異の特徴付けおよび多型性配列のバリエーションの検出を含む、直接的、正確、そして効率的な配列分析を可能にする。本発明はこれらの必要性を満たす。

【課題を解決するための手段】

【0012】

本発明は、以下を提供する:

1. 標的ポリヌクレオチドの変異の存在を決定する方法であって、以下の工程:

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全なn量体のセットを構成するように、各プローブが、二本鎖領域および一本鎖n量体オーバーハング領域を含む、工程; 30

(b) 該標的ポリヌクレオチドを、1つのアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、標的ハイブリダイゼーションパターンを生成する工程;

(c) 参照ポリヌクレオチドを、第2のアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、参照ハイブリダイゼーションパターンを生成する工程; および

(d) 該参照ハイブリダイゼーションパターンと該標的ハイブリダイゼーションパターンとを比較することによって、該標的ポリヌクレオチド中の変異の存在を決定する工程、 40 を包含する方法。

【0013】

2. 前記工程b)において、前記ハイブリダイズした標的ポリヌクレオチドが前記プローブにライゲートされる、項1に記載の方法。

【0014】

3. 前記工程c)において、前記ハイブリダイズした参照ポリヌクレオチドが前記プローブにライゲートされる、項1に記載の方法。

【0015】

4. 前記オーバーハングが遊離の5'末端を有する、項1に記載の方法。

【0016】

5. 前記オーバーハングが遊離の3'末端を有する、項1に記載の方法。

【0017】

6. 前記n量体が約4~約50ヌクレオチドを含む、項1に記載の方法。

【0018】

7. 前記変異が置換変異である、項1に記載の方法。

【0019】

8. 前記変異が欠失変異である、項1に記載の方法。

【0020】

9. 前記変異が挿入変異である、項1に記載の方法。

【0021】

10. 前記標的ポリヌクレオチドが、囊胞性線維症膜貫通伝導性調節遺伝子、p53遺伝子、ミトコンドリアDNA、またはHIV遺伝子からなる群から選択される、項1に記載の方法。

【0022】

11. 前記アレイが平行に配置されている、項1に記載の方法。

【0023】

12. 2つ以上の標的ポリヌクレオチドが同一かどうかを決定する方法であって、以下の工程：

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全なn量体のセットを構成するように、各プローブが、二本鎖領域および一本鎖n量体オーバーハング領域を含む、工程；

(b) 第1の標的ポリヌクレオチドを、1つのアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、第1のハイブリダイゼーションパターンを生成する工程；

(c) 第2の標的ポリヌクレオチドを、第2のアレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせ、第2のハイブリダイゼーションパターンを生成する工程；および

(d) 第1のハイブリダイゼーションパターンと第2のハイブリダイゼーションパターンとを比較する工程、

を包含する方法。

【0024】

13. 前記工程b)において、前記ハイブリダイズした標的ポリヌクレオチドが前記プローブにライゲートされる、項12に記載の方法。

【0025】

14. 前記工程c)において、前記ハイブリダイズした参照ポリヌクレオチドが前記プローブにライゲートされる、項12に記載の方法。

【0026】

15. 前記オーバーハングが遊離の5'末端を有する、項12に記載の方法。

【0027】

16. 前記オーバーハングが遊離の3'末端を有する、項12に記載の方法。

【0028】

17. 前記n量体が約4~約50ヌクレオチドを含む、項12に記載の方法。

【0029】

18. 前記アレイが平行に配置される、項12に記載の方法。

【0030】

19. ポリヌクレオチドの混合物中の個々のポリヌクレオチドを区別する方法であって、以下の工程：

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全なn

10

20

30

40

50

量体のセットを構成するように、各プローブが、二本鎖領域および一本鎖 n 量体オーバーハング領域を含む、工程；

(b) 該混合物中のポリヌクレオチドを、該アレイのプローブポリヌクレオチドの該オーバーハングにハイブリダイズさせる工程；および

(c) ハイブリダイゼーションパターンを決定する工程、  
を包含する工程。

【0031】

20. 前記区別された個々のポリヌクレオチドを数え上げる工程をさらに含む、項19に記載の方法。

【0032】

21. 前記工程b)において、前記ハイブリダイズした標的ポリヌクレオチドが前記プローブにライゲートされる、項19に記載の方法。

【0033】

22. 前記オーバーハングが遊離の5'末端を有する、項19に記載の方法。

【0034】

23. 前記オーバーハングが遊離の3'末端を有する、項19に記載の方法。

【0035】

24. 前記 n 量体が約4~約50ヌクレオチドを含む、項19に記載の方法。

【0036】

25. 前記アレイが平行に配置される、項19に記載の方法。 標的ポリヌクレオチドにおける変異の存在を決定する方法が提供される。この方法は以下の工程を含む： 20

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全な n 量体のセットを構成するように、各プローブが、二本鎖領域および一本鎖 n 量体オーバーハング領域を含む、工程；

(b) 標的ポリヌクレオチドを、1つのアレイ中のプローブポリヌクレオチドの前記オーバーハングにハイブリダイズさせ、標的ハイブリダイゼーションパターンを生成する工程；

(c) 参照ポリヌクレオチドを、第2のアレイ中のプローブポリヌクレオチドの前記オーバーハングにハイブリダイズさせ、参照ハイブリダイゼーションパターンを生成する工程 30  
；および

(d) 参照ハイブリダイゼーションパターンおよび標的ハイブリダイゼーションパターンを比較することによって、標的ポリヌクレオチド中の変異の存在を決定する工程。

【0037】

以下の工程を包含する方法がまた、2つ以上の標的ポリヌクレオチドが同一であるかどうかを決定するために提供される：

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全な n 量体のセットを構成するように、各プローブが二本鎖領域、および一本鎖 n 量体オーバーハング領域を含む、工程； 40

(b) 第1の標的ポリヌクレオチドを、1つのアレイのプローブポリヌクレオチドの前記オーバーハングにハイブリダイズさせ、第1のハイブリダイゼーションパターンを生成する工程；

(c) 第2の標的ポリヌクレオチドを、第2のアレイのプローブポリヌクレオチドの前記オーバーハングにハイブリダイズさせ、第2のハイブリダイゼーションパターンを生成する工程；および

(d) 第1および第2のハイブリダイゼーションパターンを比較する工程。

【0038】

さらに、以下の工程を包含する、ポリヌクレオチドの混合物中の個々のポリヌクレオチドを区別および列挙する方法が提供される：

10

20

30

40

50

(a) 少なくとも2つの同一または実質的に同一なポリヌクレオチドプローブアレイを提供する工程であって、ここで各アレイのオーバーハングが完全または実質的に完全なn量体のセットを構成するように、各プローブが二本鎖領域および一本鎖n量体オーバーハング領域を含む、工程；

(b) 混合物中のポリヌクレオチドを、アレイのプローブポリヌクレオチドの前記オーバーハングにハイブリダイズさせる工程；および

(c) ハイブリダイゼーションパターンを決定する工程。

【発明を実施するための最良の形態】

【0039】

本開示を通して、種々の刊行物、特許、および公開特許明細書は、確認された引用によって援用される。これらの刊行物、特許、および公開特許明細書の開示は、本明細書中にその全体が本開示に対する参考として援用される。

【0040】

(A. 一般的技術)

本発明の実施は、他に断らない限り、有機化学、分子生物学(組換え技術を含む)、細胞生物学、生化学、および免疫学の従来技術を利用し得、これらは当該分野の範囲内である。このような従来技術は、ポリマーアレイ合成、ハイブリダイゼーション、ライゲーシオン、標識を用いたハイブリダイゼーションの検出を含む。適切な技術の特定の例示は、以下の実施例に参考として含まれ得る。しかし、他の等価な従来手順もまた、当然ながら使用され得る。このような従来技術は、例えば、Genome Analysis: A Laboratory Manual Series (I~IV巻)、Using Antibodies: A Laboratory Manual、Cells: A Laboratory Manual、PCR Primer: A Laboratory Manual、およびMolecular Cloning: A Laboratory Manual(すべてCold Spring Harbor Laboratory Pressより)のような標準的な実験室マニュアルに見出され得る。

【0041】

(B. 定義)

本明細書で使用されるように、特定の用語は以下に定義する意味を有し得る。

【0042】

明細書および請求項で使用するように、単数形「a」「an」および「the」は、文脈が明確に別に指示しない限りは、複数の引用を含み得る。例えば、用語「an array」は、文脈が明確に別に指示しない限りは、複数のアレイを含み得る。

【0043】

「アレイ」は、意図的に作製した分子の収集物を表し、これは合成的または生合成的に調製され得る。特に、本明細書中で用語「アレイ」は、支持体上に意図的に作製したポリヌクレオチドの収集物を意味し、そこでは所定のあらかじめ定義した領域において各々のポリヌクレオチドの同一性が知られる。従って、用語「アレイ」および「生物学的チップ」は、相互に交換可能に用いられる。

【0044】

分子アレイは、種々の異なった形式(例えば、可溶性分子のライブラリー；および樹脂ビーズ、シリカチップ、または他の固体支持体につないだ化合物のライブラリー)において生物学的活性についてスクリーニングされ得る。固体基板上におけるポリヌクレオチドアレイの作製、および異なったアッセイにおけるアレイの使用の方法は、以下：米国特許第5,677,195号、同第5,624,711号、同第5,599,695号、同第5,445,934号、同第5,451,683号、同第5,424,186号、同第5,412,087号、同第5,405,783号、同第5,384,261号、同第5,252,743号、同第5,744,101号、同第5,143,854号、およびPCT WO 92/10092に記載され、これらの各々の開示は本明細書中に援用される。

## 【0045】

「固体支持体」「支持体」および「基板」は、強固な、または半強固な1つまたは複数の表面を有する物質または物質群をいう。多くの実施態様において、少なくとも1つの固体支持体の表面は実質的に平坦であり、しかしいくつかの実施態様においては、異なった化合物のための合成領域を、例えば、ウェル、上昇させた領域、ピン、刻印した溝などで物理的に分離することが所望され得る。他の実施態様に従って、固体支持体は、ビーズ、樹脂、ゲル、微粒子または他の幾何学的な構成をとる。

## 【0046】

「あらかじめ規定された領域」は、固体支持体上の局在化した領域をいい、それは選択された分子の形成に使用されるか、使用されたか、または使用されるように意図され、そしてさもなくば、本明細書中で代替的に「選択された」領域といわれる。あらかじめ規定された領域は、例えば円形、長方形、楕円形、くさび型などの任意の便利な形状を有し得る。本明細書では、簡略化のために、「あらかじめ規定された領域」は、時々単に「領域」といわれる。いくつかの実施態様において、あらかじめ規定された領域、およびそれゆえに、各々の別個の化合物が合成される領域は、約 $1\text{ cm}^2$ より小さいか、または $1\text{ mm}^2$ 未満でさえある。さらなる実施態様において、あらかじめ規定された領域は、領域(すなわち、ビーズ、樹脂、ゲルなど)をウェル、トレイなどへと物理的に分離することによって達成され得る。

10

## 【0047】

本明細書で使用されるように、「ポリヌクレオチド」は、2つ以上のヌクレオチドの配列である。本発明のポリヌクレオチドは、デオキシリボ核酸(DNA)またはリボ核酸(RNA)の配列を含み、これらは天然の供給源から単離され得るか、組換え的に産生されるか、または人工的に合成され得る。本発明のポリヌクレオチドのさらなる例は、ポリアミドポリヌクレオチドであり得る。本発明はまた、特定のtRNA分子中で同定されそして三重らせん中に存在することが仮定されている、H o o g s t e e n塩基対のような非従来の塩基対が存在する状況を包含する。

20

## 【0048】

用語「ヌクレオチド」は、デオキシヌクレオチドおよびそのアナログを含む。これらのアナログは、ポリヌクレオチド配列に組み込まれた場合、それらが溶液中で相補的なポリヌクレオチドとのハイブリダイゼーションを可能にするような、天然に生じるヌクレオチドと共通したいくつかの構造的な特徴を有する分子である。代表的には、これらのアナログは1つまたは複数の修飾した塩基、ならびにリボースおよびホスホジエステル部分の修飾形態を有し得る。変化は、ハイブリッド形成の安定化もしくは不安定化のために、または所望するように相補的ポリヌクレオチド配列とのハイブリダイゼーションの特異性を増強するために、またはポリヌクレオチドの安定性を増強するために、あつらえ仕立てに作製され得る。

30

## 【0049】

アナログはまた、ポリヌクレオチド合成において便利に使用されるように、保護および/または修飾されたモノマーを含む。当業者がよく気付いているように、ポリヌクレオチド合成は種々の塩基保護したヌクレオシド誘導體を使用し、そこではプリンおよびピリミジン部分の1つ以上の窒素が、ジメトキシトリチル、ベンジル、tert-ブチル、イソブチル、などといった基によって保護される。

40

## 【0050】

例えば、構造的な基が、必要に応じてポリヌクレオチドへの組み込みのためにヌクレオシドのリボースまたは塩基に添加される(例えば、リボースの2'-O位でのメチル基、プロピル基、またはアリル基、または2'-O基を置換するフルオロ基、またはリボヌクレオシド塩基のプロモ基)。2'-O-メチルオリゴリボヌクレオチド(2'-O-MeORN)は、相補的ポリヌクレオチド(特にRNA)に対して、それらの修飾していない対応物よりも高いアフィニティを有する。2'-O-MeORNホスホラミダイトモノマーは市販されている(例えば、Chem Genes Corp. またはGlen

50

Research, Inc. より)。あるいは、プリンまたはピリミジン複素環の1つ以上のN原子がC原子によって置換されているデアザプリンおよびデアザピリミジンがまた、使用され得る。

【0051】

ポリヌクレオチドのホスホジエステル結合、または「糖-リン酸骨格」はまた、例えば、メチルホスホン酸、O-メチルリン酸またはホスホロチオネートで置換または修飾され得る。本開示の目的のためのこのような修飾した結合を含むポリヌクレオチドの別の例は、「ペプチドポリヌクレオチド」を含有し、ここではポリアミド骨格がポリヌクレオチド塩基または修飾したポリヌクレオチド塩基に接続されている。天然に生じるヌクレオチドに見出されるポリアミド骨格および塩基を含むペプチドポリヌクレオチドは、例えばBiosearch, Inc. (Bedford, MA) から市販されている。米国特許第5,773,571号および同第5,786,461号もまた参照のこと。

10

【0052】

修飾した塩基を有するヌクレオチドはまた、本発明で使用され得る。塩基修飾のいくつかの例は、2-アミノアデニン、5-メチルシトシン、5-(プロピン-1-イル)シトシン、5-(プロピン-1-イル)ウラシル、5-プロモウラシル、5-プロモシトシン、ヒドロキシメチルシトシン、メチルウラシル、ヒドロキシメチルウラシル、およびジヒドロキシペンチルウラシルを含み、これらは相補的なポリヌクレオチドの結合アフィニティを修飾するためにポリヌクレオチドに組み込まれ得る。

【0053】

官能基はまた、ヌクレオシド糖の環またはプリン環もしくはピリミジン環の種々の位置にライゲートされ得、これは負に荷電したリン酸骨格との静電的な相互作用によって、または主溝および副溝における相互作用を通して二重鎖を安定化し得る。例えば、アデノシンおよびグアノシンヌクレオチドは、イミダゾールイルプロピル基を有するN<sup>2</sup>位で置換され得、二重鎖の安定性を増加する。3-ニトロピロールおよび5-ニトロインドールのような汎用の塩基アナログもまた、含まれ得る。本発明における使用に適する種々の修飾したポリヌクレオチドは、例えば、「Antisense Research and Application」、S.T. CrookeおよびB. LeBleu (編) (CRC Press, 1993) および「Carbohydrate Modifications in Antisense Research」ACS Symp. Ser. #580, Y.S. Sanghvi および P.D. Cook (編) ACS, Washington, D.C. 1994に記載される。

20

30

【0054】

「標的」ポリヌクレオチドは、例えば、その変異によって特徴付けられるポリヌクレオチドである。一般的に、標的ヌクレオチドの配列は未知である。しかし、いくつかの局面において、例えば、複数の標的ポリヌクレオチドが同一であるかどうかを決定する場合、1つのポリヌクレオチドの配列が知られ得る。標的は、任意の生物学的供給源に由来し得、さらに部分的または全体的に合成され得る。

【0055】

「参照」ポリヌクレオチドは、そのプローブアレイに対するハイブリダイゼーションパターンが、別の同一なまたは実質的に同一なプローブアレイに対する標的ポリヌクレオチドのハイブリダイゼーションパターンと比較されるポリヌクレオチドである。参照ポリヌクレオチドの配列は既知であり得るが、配列が既知であることは必須ではない。

40

【0056】

ポリヌクレオチドの「フラグメント」は、任意の長さであり得るポリヌクレオチドの部分である。このようなフラグメントは、一本鎖または二本鎖ポリヌクレオチドであり得、あるいは一本鎖および二本鎖部分の両方を含有し得る。

【0057】

「ポリヌクレオチドプローブ」は、本明細書中に記載されるハイブリダイゼーション工程およびライゲーション工程に使用され得る一本鎖および二本鎖ポリヌクレオチドの任意

50

の配列である。最も一般的には、ポリヌクレオチドプローブは一本鎖ポリヌクレオチド領域に共有結合した二本鎖ポリヌクレオチド領域からなる。各プローブの二本鎖領域についてのポリヌクレオチド配列は、同一であり得るか、または異なり得る。作製の簡便さのため、二本鎖領域は各プローブにおいて同一であり得、そしてこのような同一領域はまた、「定常」領域として公知であり得る。

#### 【0058】

一本鎖領域におけるヌクレオチドは好ましくは可変性である。この一本鎖可変領域は「オーバーハング(overhang)」として公知であり、そしてこのオーバーハングは5'配向性または3'配向性であり得る(すなわち、それぞれフリーの5'末端またはフリーの3'末端を所有し得る)。さらに、オーバーハングはいくつかのヌクレオチド(「n量体」)のものであり得、例えば、約4~約50ヌクレオチドの範囲であり得る。好ましくは、オーバーハングは約5~約20ヌクレオチドであり、より好ましくは約6~約12ヌクレオチドであり、そして最も好ましくは8~9ヌクレオチドである。

10

#### 【0059】

用語「ハイブリダイゼーション」は、2つの一本鎖ポリヌクレオチドが非共有的に結合して安定な二本鎖ポリヌクレオチドを形成する工程をいい；三本鎖ハイブリダイゼーションもまた、理論的に可能である。得られる(通常の)二本鎖ポリヌクレオチドは「ハイブリッド」である。安定なハイブリッドを形成するポリヌクレオチドの集団の割合を、本明細書中では「ハイブリダイゼーションの程度」という。

#### 【0060】

ポリヌクレオチドハイブリダイゼーションアッセイを行う方法は、当該分野において十分開発されている。ハイブリダイゼーションアッセイ手順および条件は、適用に依存して変化し、そして以下に記載される方法を含む、公知の一般的結合方法に従って選択される：Maniatisら、「Molecular Cloning: A Laboratory Manual」第2版、Cold Spring Harbor, N.Y., 1989；BergerおよびKimmel、「Methods in Enzymology」152巻、「Guide to Molecular Cloning Techniques」, Academic Press, Inc., San Diego, CA., 1987；YoungおよびDavis, Proc. Natl. Acad. Sci., U.S.A., 80:1994(1983)、これらはそれぞれ本明細書中に、参考として

20

30

#### 【0061】

2つの一本鎖ポリヌクレオチドのハイブリダイズする能力は、その相補性の程度およびハイブリダイゼーション反応条件のストリンジェンシーのような因子に依存することが理解される。

#### 【0062】

本明細書中で使用される場合、「ストリンジェンシー」とは、ポリヌクレオチドがハイブリダイズする程度に影響を与えるハイブリダイゼーション反応の条件をいう。ストリンジェンシーな条件は、ポリヌクレオチド二重鎖がそのミスマッチの程度に基づいて区別されることを可能にする条件が選択され得る。高いストリンジェンシーは、ミスマッチした塩基を含有する二重鎖の形成についての低い可能性と、相関関係にある。従って、ストリンジェンシーが高くなるほど、ミスマッチした二重鎖を形成し得る2つの一本鎖ポリヌクレオチドが一本鎖のまま残る可能性が高くなる。逆に、より低いストリンジェンシーにおいては、ミスマッチした二重鎖を形成する可能性は増加する。

40

#### 【0063】

1つ以上のミスマッチを含有する二重鎖と比較して、完全にマッチした二重鎖の選択を可能にする適切なストリンジェンシー(または、高い程度のミスマッチを含有する二重鎖と比較して、特定のミスマッチした二重鎖の選択を可能にする適切なストリンジェンシー)は、一般に経験的に決定される。ハイブリダイゼーション反応のストリンジェンシーを調整する方法は、当業者に周知である。例えば、Sambrookら、「Molecul

50

ar Cloning: A Laboratory Manual」第2版、Cold Spring Harbor Laboratory Press, 1989; Ausubelら、「Current Protocols In Molecular Biology」John WileyおよびSons、1987、1988、1989、1990、1991、1992、1993、1994、1995、1996および定期的な最新刊まで; ならびにHamesら、「Nucleic Acid Hybridization: A Practical Approach」IRL Press, Ltd., 1985を参照のこと。

【0064】

一般に、ストリンジェンシーを増加させる(すなわち、より密接にマッチした二重鎖の形成を選択する)条件は、より高い温度、より低いイオン強度、および溶媒の存在または非存在を含み; より低いストリンジェンシーは、より低い温度、より高いイオン強度およびより低い濃度またはより高い濃度の溶媒(例えば、より低い濃度のホルムアミドまたはジメチルスルホキシド)により好都合とされる。ハイブリダイゼーション反応の時間および反応物(すなわち、一本鎖ポリヌクレオチド)の濃度もまたストリンジェンシーに影響し、短い反応時間および低い反応物濃度はより高いストリンジェンシーを好む。

【0065】

用語「ライゲーション」とは、2つの別々のポリヌクレオチド鎖上の2つの末端ヌクレオチドを共有結合させ、それによって2つの鎖が1つの一本鎖に連結されるプロセスをいう。本発明の方法論において、いくつかの局面において、ハイブリダイゼーションおよびライゲーションは、1つの工程において達成される。

【0066】

「n量体」とは、「n」数のヌクレオチドの一本鎖ポリヌクレオチドである。n量体の「完全なセット」とは、「n」数のヌクレオチドの一本鎖ポリヌクレオチドのセットをいい、ここでセットとは「n」個のヌクレオチドの全ての可能な組合せを示す。例えば、「n」が8を示す場合、完全なn量体のセットとは、 $4^8$ の一本鎖ポリヌクレオチドを示す。n量体の「実質的に完全なセット」とは、「n」個のヌクレオチドの全ての可能な組合せを有さなくてもよいが、標的ポリヌクレオチド中の変異が特徴付けられるようにプローブとして使用されるような一本鎖の十分な数を提供するセットを表す。従って、いくつかの局面において、このような実質的に完全なセットは一本鎖ポリヌクレオチドプローブのサブセットを含み得る。

【0067】

アレイの記載の状況において使用される場合、用語「実質的に同一」とは、各プローブの構成に関して同一でなくてもよいアレイをいう。しかし、このようなアレイは、標的ポリヌクレオチドにおける変異が特徴付けられ得るような本発明のハイブリダイゼーションおよびライゲーション工程に参加することに関しては、機能的である。例えば、1つのアレイが特定の位置の特定の塩基を有するプローブを提供し、そして他方のアレイが対応する位置の塩基のアナログを提供する場合、2つのアレイは実質的に同一であり得る。

【0068】

用語「塩基呼び出し(base call)」とは、標的ポリヌクレオチド中の未知の塩基の同定の決定をいう。塩基呼び出し方法は、標的ポリヌクレオチドとプローブポリヌクレオチドとの間のハイブリダイゼーションの程度を、参照ポリヌクレオチドとプローブポリヌクレオチドとの間のハイブリダイゼーションの程度と比較することによりなされ得る。所定の場合、塩基呼び出しはハイブリダイゼーションの限界の程度を選択することによりなされ得るかまたは抑制され得る。

【0069】

「塩基呼び出しセット」は、特定のヌクレオチド、または標的ポリヌクレオチド中のヌクレオチドの配列のための塩基呼び出しを生じるポリヌクレオチドの集団のセットである。従って、各塩基呼び出しセットは単一の置換位置(図3において文字Nによって示す)で異なる4つのプローブからなり、特定の標的塩基を問いただすように配置される。

10

20

30

40

50

## 【0070】

「変異」は、正常な（「野生型」）集団のポリヌクレオチドからのポリヌクレオチドのヌクレオチド配列中の変化である。「置換」変異とは、あるヌクレオチドが別のヌクレオチドにより置換されたものである。「欠失」変異とは、1つ以上のヌクレオチドが正常な配列から欠失され、そして「挿入」変異とは、1つ以上のヌクレオチドが正常な配列中に挿入されたものである。「ホモ接合体」変異とは、2倍体生物中の両方の染色体に現れる変異である。「ヘテロ接合体」変異とは、2倍体生物中の1つの染色体のみに現れる変異であり、他方の染色体は正常である。変異は、例えば、435位でのA G置換として特徴付けられ得る。

## 【0071】

「多型性」とは、同じ集団における、複数の別々の対立遺伝子状態の発生であって、その状態の少なくとも1つが高い頻度（通常は>1%以上）を有する状態である、発生である。当業者にとって、用語「多型性」が変異に関することは公知である。

## 【0072】

本開示にわたって、本発明の種々の局面はある範囲の形式で存在する。ある範囲の形式における記載は単に簡便さおよび簡潔さのためのものであり、ならびに本発明の範囲における確固たる制限として構成されるべきではないことが理解されるべきである。従って、ある範囲の記載は、可能な部分集合ならびにその範囲内の個々の数値を全て詳細に開示したと見なされるべきである。例えば、4~50のような範囲の記載は4~10、4~20、4~30、4~40、4~50、5~10、5~20など、ならびに例えば、6、8、15、20、32、39、43、48などの範囲内の個々の数のような全ての部分集合を詳細に開示したと見なされるべきである。この適応は範囲の広さとは見なされない。

## 【0073】

(C. 方法)

## 1. 概要

本発明は、標的ポリヌクレオチドにおける変異の存在を、「n量体オーバーハング」として公知の変一本鎖領域を有する二本鎖ポリヌクレオチド領域を有する完全なまたは実質的に完全なn量体アレイのポリヌクレオチドプローブを用いて、決定する方法を提供する。1つの方法は標的ポリヌクレオチドのフラグメントを、アレイのポリヌクレオチドプローブとハイブリダイズさせる工程を包含する。別の同一なまたは実質的に同一なアレイのポリヌクレオチドプローブは、参照ポリヌクレオチドのフラグメントと同様にハイブリダイズする。アレイ上の標的ポリヌクレオチドと参照ポリヌクレオチドのハイブリダイゼーションパターンは直接比較され、変異の存在の決定を可能にし得る。

## 【0074】

方法論の感度および精度は、ハイブリダイゼーションパターンから「仮想タイリング(virtual tiling)」(以下を参照のこと)を作製し、そして仮想タイリングを直接比較することによって増加され得る。このような直接的な比較によって、全体のおよび再生産可能な交差ハイブリダイゼーションならびに複数のプローブ標的相互作用から得られる、潜在的に混乱させるシグナルを減算する。この方法は、2.5kbpの標的において90%より多くの変異検出感度を与え(8量体のアレイ上で)、非常に正確性が高い(すなわち、低い数の偽陽性、1塩基当たり<0.02%)。仮想タイリング工程は、適切な照会プローブを電子的に組み立てることによって標的内の、予測できない変異(挿入、欠失、複数の点変異)の検出の焦点、を絞った検索を可能にする。

## 【0075】

本明細書中に開示される場合、本発明の方法は変異の存在の決定のみではなく、変異の位置および型の同定(すなわち、変異が単一または複数の塩基の置換、欠失、もしくは挿入であるかどうか)にも関する。本方法はまた、ホモ接合体、およびヘテロ接合体変異の両方ならびに多型性の検出にも使用され得る。さらに、本方法は2つ以上の未知のポリヌクレオチド配列を比較してこれらが同一であるかどうかを決定するためにも使用され得、これはポリヌクレオチドの1つの配列を参照として作用させるために知る必要がない。こ

10

20

30

40

50

これらの方法はこのような分析を可能にする。なぜならば、配列を比較する伝統的な技術と異なり、これらの方法はハイブリダイゼーションパターンを比較するからである。

【0076】

従って、複数のポリヌクレオチドを迅速におよび正確に比較し、これらが配列において同一であるかどうかを決定し得る。これらの迅速なおよび正確な方法は、配列決定を行う必要性または任意のポリヌクレオチドの配列が公知である必要性を伴わずに種々の供給源からのポリヌクレオチドを比較する、法医学的な目的に特に有用である。これらの方法は、それらの正確さ、感度、それゆえに法廷における広範な採用を得ることが期待され得るという点で、従来の方法を越える利点を提供し得る。さらに、本方法は、曝露の前後のハイブリダイゼーションパターンを比較することによって、薬剤、化学物質、放射能、ストレス、または環境因子の遺伝子に対する変異性の影響を迅速にスクリーニングするために使用され得る。本方法の汎用性は、さらに、以下に記載される技術および米国特許第5,545,531号に記載される技術を並行して処理することによって、利用され得、そしてそれによって複数のポリヌクレオチド配列が同時に分析され得る。

10

【0077】

標準化された、完全なまたは実質的に完全なアレイの使用は、分析されるべき各ポリヌクレオチドに対してカスタマイズされたアレイを構築する必要性を排除し、そしてこのことは「1つのサイズが全てに適合する(one-size-fits-all)」方法論(これは実質的にいずれのポリヌクレオチド分析にも使用され得る)を提供する。従って、本方法は従来ゲルに基づいた複雑なポリヌクレオチドサンプルの分析に代わる、技術

20

【0078】

本発明を実施する1つの方法は、参照ポリヌクレオチドおよび標的ポリヌクレオチドとハイブリダイズする同一なプローブアレイまたは実質的に同一なプローブアレイの使用に関する。本発明はまた、参照ポリヌクレオチドおよび標的ポリヌクレオチドが首尾良くハイブリダイズする単一のアレイを用いて、各々分析され、そして保存されたハイブリダイゼーションパターンを用いて、実施され得る。あるいは、2つのセットのプローブが1つの基板(substrate)上に提供され、ここで2つのプローブセットは明確に着色されるかまたは標識化され、その結果として、標的および参照物の両方について同時に本発明が行われ得る。

30

【0079】

本明細書中に記載される方法は、2つ以上のアレイの使用がそれぞれ意図され、ここで各アレイは完全なn量体のセットのプローブのポリヌクレオチドを構築する。これは完全に当業者の範囲内であり、本方法(ハイブリダイゼーション、ライゲーションおよび特徴付け)は並行して行われ、ここで複数のアレイは、複数の試験ウェルを有し、そして各試験ウェルは生物学的チップまたはアレイを包含する1つのプレート(生物学的チッププレート)上に提供され得る。このような分析の並行法は、米国特許5,545,531号に記載され、これは本明細書中においてその全体が参照として援用される。

【0080】

簡便には、標的ポリヌクレオチドおよび参照ポリヌクレオチドは試験ウェル中に導入され得る。体液操作デバイスは、例えば、体液をウェルに添加すること、またはウェルから除去することによって、ウェル中の液体を所定の温度に維持することによって、ならびに必要な場合ウェルを攪拌し、それによって種々のハイブリダイゼーション、ライゲーションおよび洗浄工程を行うことにより、試験ウェルを選択されたセットの反応条件に曝露し得る。次いでアレイの読み取りは試験ウェル中のプローブアレイを審問し得、それによりハイブリダイゼーションデータを得る。適切なプログラムを備えたコンピューターは、さらに本明細書中に記載されるようなハイブリダイゼーションパターンを分析し得る。このような並行処理は効率を非常に改善させる。

40

【0081】

本発明の方法は任意のポリヌクレオチドに関する変異を特徴付けるために使用され得る

50

。例えば、嚢胞性線維症に関する変異（CFTR遺伝子、腫瘍抑制因子p53遺伝子およびHIV-1pol遺伝子に関する変異を含む）が特徴付けられ得る。本方法はまた、DNAフィンガープリンティング、配列特異的5'末端サンプリング、および異なる遺伝子発現アッセイに有用な技術、酵素的および化学的なミスマッチの切断ならびにポリヌクレオチドスキニングのような、法医学的な同定目的のための、ミトコンドリアDNAの特異的な配列を検出する技術を提供する。

【0082】

2. 完全なまたは実質的に完全なn量体アレイ

a) アレイの特徴

本発明のポリヌクレオチドプローブのアレイは予め規定された領域で基板に付着されたポリヌクレオチドプローブを包含する。ポリヌクレオチドプローブの1つの例が図1に示されている。 10

【0083】

各プローブは二本鎖領域および一本鎖領域を含む。一つの局面において、プローブの二本鎖部分は固体基板上に共有結合されている。必要に応じて二本鎖部分はソラレンのような架橋剤を用いて共有結合される。プローブの二本鎖部分は約4~約200以上のヌクレオチド長であり得る。一つの局面において、ポリヌクレオチドプローブの二本鎖部分は、定常である（すなわち、所定のアレイにおいて、あるプローブから別のものに変化しない）ヌクレオチド配列からなる。しかし、プローブからプローブへと異なる二本鎖部分を用いて本発明を実施することは可能であり、いくつかの場合において所望され得る。二本鎖領域の存在は、本方法におけるハイブリダイゼーション工程の後に行われるべきライゲーション工程を可能にする。 20

【0084】

ポリヌクレオチドプローブの一本鎖領域は「オーバーハング」として公知であり、そしてこれらのオーバーハングは5'または3'配向性であり得る（すなわち、それぞれフリーの5'末端またはフリーの3'末端を有し得る）。図1は5'配向性であるオーバーハングを示す。オーバーハングは任意の長さ（「n」）の異なるヌクレオチドであり得、「n量体」として公知である。本明細書中に記載される場合、完全な「n量体アレイ」とは、プローブの一本鎖部分の「n」長の配列の全ての可能な組合せを包含する。4つの塩基のいずれもオーバーハングの任意の部分に存在し得るので、完全なアレイのプローブの一本鎖部分に対する可能なヌクレオチド配列の総数は、 $4^n$ である。例えば、 $n=8$ の場合、プローブ配列の総数は $4^8=65,536$ である。オーバーハングの長さは、例えば、約4~約50以上のヌクレオチドであるが、より短いおよびより長い鎖もまた、使用され得る。好ましくは、オーバーハングは約6~約12ヌクレオチドそしてより好ましくは約8~約9ヌクレオチドである。 30

【0085】

完全なアレイが好ましいが、本発明の実施のためにアレイが完全である必要はない。アレイは実質的に完全であり得、そして本発明はなお実施され得る。n量体の「実質的に完全なセット」は、「n」ヌクレオチドの全ての可能な配列を有さないセットを示し得るが、標的ポリヌクレオチド中の変異を特徴付けるためのプローブとして使用され得る十分な数の異なる配列が提供される。 40

【0086】

b) アレイ合成

ポリヌクレオチドプローブの完全なまたは実質的に完全なアレイは当該分野において公知の方法に従って合成され得る。DMT-ヘキサエチレングリコール-(2-シアノエチル-N,N-ジイソプロピル)ホスホルアミダイトのようなリンカーは、ガラスまたはシリカ基板のような誘導体化された固体支持体に付着され得る。二本鎖アンカー領域の一本鎖のヌクレオチドのすべて（しかし最後のものは、従来の化学合成法を用いて、例えば、DMT-PACホスホルアミダイトを用いて合成される。このような技術は当該分野において周知である。例えば、米国特許第4,458,066号、同第4,500,707 50

号、同第4, 725, 677号、ならびにJones、第2章、およびAtkinsonら、第3章、Gait編、Oligonucleotide Synthesis: A Practical Approach (IRL Press, Washington, D.C., 1984); Froehlerら、Tetrahedron Letters, 27: 469-472 (1986)を参照のこと。

【0087】

二本鎖領域の一本鎖の最後のヌクレオチドは、一本鎖に相補的な鎖(一本鎖n量体のオーバーハングになる「n」個のヌクレオチドに従う)は、例えば、「光指向性」の写真平板技術によって、合成され得る。プローブ上において、一方の鎖は他方よりも長く、その結果として2つの鎖は短い方の鎖の全長にわたり相補的であり(従って、プローブの二本鎖部分を形成する)、そしてオーバーハングはハイブリダイゼーション部分を超えて伸長するより長い鎖の部分であることが理解される。図1を参照のこと。

10

【0088】

光指向性核酸合成技術は当該分野において周知である。例えば、米国特許第5, 143, 854号、同第5, 424, 186号、同第5, 445, 934号、同第5, 527, 681号、同第5, 677, 195号および同第5, 770, 722号を参照のこと、これらの開示は本明細書中において援用される。

【0089】

簡潔には、光指向性技術において、選択された領域の基板が光に曝露されてこれらの領域を活性化し、そして光除去可能な基によりヌクレオチドモノマーは活性化された領域に付着される。モノマーは感光性保護基により5'ヒドロキシルで保護されたホスホルアミダイト活性化ヌクレオチドであり得る。活性化および付着の工程は、所望の長さおよび配列のポリヌクレオチドが合成されるまで繰り返される。次いで、感光性の基は必要に応じて除去され、その後、配列は必要に応じてキャップ化される。側鎖保護基もまた、存在する場合は除去される。

20

【0090】

これらの写真平板技術および製造技術によって、各プローブ配列が支持体上の非常に小さな領域を占有することを可能にする。例えば、 $0.25\text{ mm}^2$ のプローブアレイは $10^0 \sim 10^4$ または $10^0 \sim 10^5$ または $10^0 \sim 10^6$ の特徴を有し得る。さらに所望のプローブ密度の達成における詳細は、米国特許第5, 510, 270号に見出され得

30

【0091】

コンピューターツールは、写真平板技術を使用するアレイの形成に使用され得る。例えば、コンピューターシステムは基板上のポリヌクレオチドまたは他のポリマープローブを選択するため、ならびにアレイのレイアウトを設計するために使用され得る。米国特許第5, 571, 639号を参照のこと、これらの開示は本明細書中において援用される。あるいは、アレイは本明細書中において参照として援用される米国特許第5, 384, 261号に開示される機械的技術を使用して合成され得る。

【0092】

基板は好ましくは平面であるが、種々の代替的な表面配置を取り得る。基板およびその表面はまた、適切な光吸収特徴を提供するように選択され得る。例えば、基板は機能化されたガラス、Si、Ge、GaAs、GaP、 $\text{SiO}_2$ 、 $\text{SiN}_4$ 、改変型シリコン、または任意の広範な種々のゲルもしくは(ポリ)テトラフルオロエチレン、(ポリ)ビニリデンジフルオライド、ポリスチレン、ポリカーボネート、ポリプロピレン、またはそれらの組み合わせのようなポリマーであり得る。本開示を考慮すると、他の基板物質が容易に当業者に明らかになり得る。好ましい実施態様において、基板は平面ガラスかまたは平面シリカである。

40

【0093】

これらの使用のための種々の光開裂可能な保護基および方法は、当該分野において公知である。例えば、米国特許第5, 489, 678号、同第5, 744, 101号、および

50

同第5, 753, 788号を参照のこと。保護基は代表的には光分解性、すなわち、約340nm、好ましくは約365nmより大きい波長での照射により除去される。光分解は、通常、水、アルコールもしくは混合型水-アルコールまたは混合型水-有機溶媒混合物のような水酸化性溶媒あるいはプロトン性溶媒の存在下で行われる。頻繁に使用されるアルコール性溶媒には、メタノールおよびエタノールが挙げられる。光分解は頻繁に中性pHまたは塩基性pHで行われる。あるいは、酸性または塩基性触媒化開裂条件の使用が有利であり得る。このような化学的開裂方法は、例えば、米国特許第5, 599, 695号に記載される。

#### 【0094】

いくつかの局面において、二本鎖領域およびn量体オーバーハング領域の合成に続いて、2塩基のイノシンが、例えば、5'末端に、従来のDMTホスホルアミダイト化学を使用して付加され得る。2つのイノシンの付加は識別に大きく影響を与えることなく全体的なアレイングナル強度を改善する。

10

#### 【0095】

アンカー二本鎖ポリヌクレオチドは適切な塩基の組合せおよび/または適切な共有結合架橋(例えば、5'末端でのATジヌクレオチド対とソラレンC6部分)を可能にする化学的構成成分を含み得る。一般に、AおよびG(プリン、好ましくはG)の3'塩基を有するプローブのアンカー二本鎖ポリヌクレオチド部分は、CおよびT(ピリミジン)3'塩基を有するポリヌクレオチドよりも高いシグナル強度を生成する。次いで、一本鎖n量体をアレイと接触し、そして相補的なアンカー配列とハイブリダイズさせて一本鎖オーバーハングを有する二本鎖部分を形成する。オーバーハングの配向は、所望のように3'または5'へと変化され得る。これは、オーバーハングがフリーの3'または5'末端を有し得ることを意味する。3'ヒドロキシル末端は酵素的ライゲーションに好ましいが、一方、3'ホスホリルバージョンは化学的ライゲーションに好ましい。

20

#### 【0096】

##### 3. フラグメント化、ハイブリダイゼーション、およびライゲーション

標的およびまたは参照ポリヌクレオチドはフラグメント化され、そして必要な場合には一本鎖フラグメントに変性される。フラグメントはランダムまたは偽ランダムに種々の長さのものであり得、好ましくはn量体オーバーハングよりも短くはない。いくつかの局面において、フラグメントは約100塩基対までかまたはそれより長く、一方、約20~約60塩基対のフラグメントが好ましくあり得る。当業者は、フラグメントサイズが容易に所望のように最適化され得ることを理解する。フラグメント化は、化学的に、酵素的に、または当該分野で公知の他の手段によって達成され得る。1つの局面において、ヌクレアーゼ、例えば、標的ポリヌクレオチドがDNAの場合はDNase Iが使用され得る。

30

#### 【0097】

フラグメント化に続いて、ポリヌクレオチドが二本鎖の場合、例えば、95で15分間加熱することによって変性される。この手順によって、ポリヌクレオチドはその配列全体を示すホスホリル末端一本鎖のサブ配列のランダムなまたは偽ランダムな集団へと減少される。ホスホリル末端一本鎖のサブ配列は3'または5'配向型であり得るが、5'配向型であることが好ましい。相補鎖は両方とも示される。必要に応じて、フラグメントは、末端ジデオキシヌクレオチジルトランスフェラーゼおよびddATP-N6-ビオチンまたはポリヌクレオチドキナーゼおよび垂リン酸標識化ATPのような、当該分野において公知の末端標識化薬剤で(その3'末端またはその5'末端のいずれかで)末端標識され得る。

40

#### 【0098】

フラグメント化参照および標的ポリヌクレオチドのプローブオーバーハングに対するハイブリダイゼーションは当該分野において周知の条件下において行われ得る。最適なストリンジェンシーの条件下でのハイブリダイゼーションはミスマッチを減少させることが当該分野において周知である。ハイブリダイゼーション反応のストリンジェンシーを調節する手段は当業者に周知である。例えば、Sambrookら、「Molecular C

50

loning: A Laboratory Manual」第二版、Cold Spring Harbor Laboratory Press, 1989; Ausubelら、「Current Protocols In Molecular Biology」John WileyおよびSons, 1996および最新刊までの定期刊行物; ならびにHamesら、「Nucleic Acid Hybridization: A Practical Approach」IRL Press, Ltd., 1985を参照のこと。

**【0099】**

1つの局面において、ハイブリダイゼーションは6×SSPE緩衝液のような緩衝液中で約7.0~約8.0、好ましくは約7.4のpHで、約25~45の温度で、好ましくは約40で、そして約1~2時間の間、行われ得る。別の局面において、本発明のハイブリダイゼーションは、シグナルエンハンサーとしてベタインを使用して行われ得る。ハイブリダイゼーションシグナルを改善するベタインの使用は、1996年5月16日に出願された同時継続出願の米国出願番号08/648,709号に記載される(現在許可されている)。ハイブリダイゼーションに続いて、ライゲーションの調製において、例えば、GeneChip™ Fluidics Station (Affymetrix, Inc. から)上で6×SSPEを用いて22でアレイを洗浄する。

10

**【0100】**

全てのフラグメントの適切な末端での「n」のヌクレオチドにより、アレイが完全または実質的に完全である場合にハイブリダイズするプローブの一本鎖部分上に正確に相補的なn量体が見出される。フラグメントは「n」のヌクレオチドより長くあり得るが、ハイブリダイゼーションから得られるべき情報は、プローブ上のその正確に相補的な配列とハイブリダイズするフラグメントの「n」のヌクレオチドから生じる。それゆえ、「n」の長さの標的ポリヌクレオチドにおける全ての末端ヌクレオチド配列はアレイ中でハイブリダイズしたプローブによって表される。

20

**【0101】**

本方法はまた、プローブのライゲーションのさらなる工程を包含し得、一方いくつかの局面において、ハイブリダイゼーションおよびライゲーションの両方を、単一の工程で達成し得る。用語「プローブとのライゲーション」はハイブリダイズした標的または参照ポリヌクレオチドフラグメントがプローブのより短い鎖に連結されるプロセスをいうことが理解される。図1を参照のこと。鎖の末端ヌクレオチドのみが連結可能である。さらに、標的および参照フラグメントは、異なるが、同一または実質的に同一であるアレイに、別々にハイブリダイズされることが理解される。本方法を使用して2つ以上の未知のポリヌクレオチドを分析する場合、各々の未知のポリヌクレオチドのフラグメントは、異なるが同一または実質的に同一であるアレイに別々にハイブリダイズする。

30

**【0102】**

ライゲーションは、アレイの洗浄により、ハイブリダイズしていないポリヌクレオチドフラグメントを除去し得る。従って、1つの局面において、フラグメント化した、または変性したポリヌクレオチドは氷上で冷却されてライゲーションミックスに添加され、マイクロチューブ(microtube)中で遠心されて任意の沈殿物の存在にペレット化され、そしてポリヌクレオチドプローブアレイに適用される。

40

**【0103】**

酵素的および化学的ライゲーション手順の両方が満足の結果を生成する。酵素的ライゲーションは、T4 DNAリガーゼ、E. coli (NAD依存型) DNAリガーゼ、またはTaq DNAリガーゼのようなDNAリガーゼを使用し得る。これらの酵素の各々に適切な標準ライゲーション緩衝液および温度は、当該分野において周知である。ライゲーション混合物は代表的には一晚(14~16時間)インキュベートされるが、2時間の短さのインキュベーションでも十分であり得る。

**【0104】**

化学的ライゲーション反応は、1-エチル-3-(3-ジメチルアミノプロピル)カル

50

ボジイミドヒドロクロリドのようなカルボジイミドを使用してフラグメントのホスホリル部分とアンカーポリヌクレオチドプローブ上のホスホリル部分との間にピロホスフェート結合を生成し得る。例えば、フラグメントのホスホリル部分は5'配向型であり得、一方、アンカープローブのホスホリル部分は3'配向型である。代表的には、化学的ライゲーションは、30分で14時間の間進行される。1つの局面において、3~4Mまでのテトラメチルアンモニウムクロライド(TMACL)が、A/T-リッチおよびG/C-リッチなハイブリッドの安定性の標準化を助けるためにライゲーション緩衝液中で使用され得る。

#### 【0105】

ライゲーションに続いて、アレイは、例えば、GeneChip<sup>TM</sup> fluidics station上で上述のように洗浄される。必要に応じて、アレイは当該分野で公知の染料で染色され、次いで洗浄され得る。このような染色の1つの例は、ストレプトアビジン-フィコエリトリン結合体である。

10

#### 【0106】

##### 4. 変異の存在の決定

##### a) ハイブリダイゼーションパターン

ハイブリダイゼーションパターンが評価され得ることによって任意の方法が使用され得る。いくつかのこのような方法が当該分野において公知である。例えば、米国特許第5,599,695号を参照のこと。1つの好ましい方法において、標的および参照ポリヌクレオチドのフラグメントが標識化される。より好ましい局面において、標識は蛍光標識であり、これはビオチン標識化に続くストレプトアビジン-フィコエリトリン結合体形成によって得られる。アレイからの蛍光シグナルは特別に設計された共焦点スキャナーを用いて検出され得、これは、例えば、8量体のアレイを6.8~7.5μmの解像度で画像化し、そして9量体のアレイを3.5μmの解像度で画像化する。画像は、例えば、GeneChip<sup>TM</sup> software (Affymetrix, Inc., Santa Clara, CA)で処理され、そして得られた強度のデータはプローブ配列のアルファベット順のリストおよび対応する強度を含むテキストファイルに変換され得る。次いで、これらのデータは、例えば、Visual Basic<sup>TM</sup> (v.4.0, Microsoft)に書き込まれたカスタムソフトウェアにより分析される。

20

#### 【0107】

蛍光強度はバックグラウンドが減算されて、次いで各プローブ(総数4<sup>n</sup>プローブ)の強度を、全ての関連する単一塩基ミスマッチ(MM)プローブの強度の合計(8量体のアレイ上の各プローブについては24MMプローブ)およびプローブ自身の強度で除することによって標準化する。式は： $I_{標準化} = I / (I_{MM} + I)$ である。この標準化アルゴリズムは、プローブの配列依存的強度変動を和らげることにより、塩基呼び出し法の精度を改善する(塩基呼び出し法の記載については以下を参照のこと)。標準化はまた、塩基呼び出しセットのレベルで行われ、その結果として塩基呼び出しセット中の4つのプローブの強度の合計は常に一貫している。このことは、任意の1つの塩基呼び出しセットが合成(composite)平均を優性に占めることを防ぐ。

30

#### 【0108】

##### b) 塩基呼び出し

「塩基呼び出し」とは、ヌクレオチド配列における特定の位置でのヌクレオチド中の塩基を同定するプロセスである。上記のように、各標的または参照ポリヌクレオチドフラグメントの末端での「n」ヌクレオチドは、正確に相補的なn量体オーバーハングとハイブリダイズする。さらに、フラグメントの末端のnヌクレオチドにおける所定の塩基について、その1つの塩基の同一性においてのみ正確に相補的であるプローブとは異なるアレー上に3つのプローブもまた存在する。例えば、塩基がアデニンである場合、その相補塩基はチミンであり、そして相補的ではない3つの他のプローブは、その位置にアデニン、グアニン、またはシトシンを含む。それゆえ、「n」長サブ配列における所定の塩基について、1つの完全に相補的なプローブおよび3つの相補的でなく、その1つの塩基でのみ異なる

40

50

プローブが存在する。プローブのこのセットは、「塩基呼び出し」セットと呼ばれる。米国特許第5,837,832号を参照のこと。これは、本明細書において参考として援用される。

#### 【0109】

ポリヌクレオチドサンプルにおける各ヌクレオチドは、塩基呼び出しセットの「p」数を生じる。ここで「p」は、ポリヌクレオチドにおけるヌクレオチドの数を表す。塩基呼び出しセットの例については図3aを参照のこと。両方の相補鎖が分析される場合、各ヌクレオチドについての塩基呼び出しセットの数は、 $2p$ へと増加する。従って、各塩基は、縮重的に検索され、これは、サンプルの比較の正確さを増加させる。所定の4つのヌクレオチド塩基呼び出しセットにおいて、ハイブリダイゼーションの程度は、ポリヌクレオチドサンプルにおけるヌクレオチド配列に正確に相補的であるプローブについて最大である。図3bの例示を参照のこと。所定のヌクレオチドについての塩基呼び出しセットの縮重性のため、ヌクレオチドについてのすべての塩基呼び出しセットを横切るハイブリダイゼーションの程度を合計すると、ヌクレオチドを同定する非常に正確な方法が与えられる。こでを、図3cに示す。この論法は、通常のア、T、G、およびC以外のヌクレオチド（例えば、ヌクレオチドアナログ）にも拡張され得る。

10

#### 【0110】

このようにして得られた情報は、「トレース」において連続的なヌクレオチド位置について表示され得る。トレースは、配列決定における各位置について最も確からしい塩基呼び出しを示す。図4を参照のこと。配列決定トレースは、配列位置の関数として塩基呼び出しセットの各メンバーについて強度の値をプロットすることによって構築される。図4において、66位～68位を、540塩基対の嚢胞性線維症遺伝子フラグメントについてプロットする。強度を、各塩基対呼び出しセット内の4つのプローブの強度が合計1となるように基準化する。

20

#### 【0111】

ほとんどの場合、塩基呼び出しは、塩基呼び出しセットの異なるメンバーに対するハイブリダイゼーションの程度の比較から明らかであるが、いくつかの場合、ハイブリダイゼーションの程度の差異が呼び出しを行うのに十分に大きくない。呼び出しまたは非呼び出しについてのカットオフ点は、信頼度の使用を通じて見い出され得る。塩基呼び出しの信頼度（質）は、以下のように規定される：塩基呼び出しセット内のプローブについて  $I_{m, st-intense} / I_{next-most-intense}$ 。この値が閾値1.15未満である場合、塩基は、呼び出しなしとして分類される。この手段を用いて、ほとんどの正確ではない呼び出しは低い信頼度によって特徴付けられる。図5bを参照のこと。

30

#### 【0112】

##### （c）仮想タイリング

所定の変異に対して正確に相補的であるプローブは、その変異を含む標的ポリヌクレオチド由来のフラグメントに対して、参照配列に対するよりも高い程度でハイブリダイズする。これらのプローブは、さらなる分析について選択され得る。所定のポリヌクレオチドにおける変異を検査するための適切なプローブを選択するためのこの戦略は、「仮想タイリング」と呼ばれ、これは、アレイ中のプローブに実際に存在する物理的なタイリングに類似する概念である。タイリングの基本は、Cheerら、Science、274:610-14（1996）においてさらに説明される。これは、本明細書において参考として援用する。

40

#### 【0113】

物理的タイリングにおいて、プローブのアレイは、特定の遺伝子における変異に対してのみ対応するプローブのセットを含むように構築される。すなわち、尋問（interrogation）プローブは、分析される特定のヌクレオチド配列に基づくアレイ上での合成の前に、選択される。対照的に、仮想タイリングは、完全なn量体プローブのセットに由来する適切な尋問プローブを「電氣的に」組立てることを含む。この設計は、各標的塩基について複数の尋問プローブを生成し、これは、単純な基準化合計（平均化）プロセ

50

スを通して統合され得る縮重塩基呼び出しセットを生成する。従って、ポリヌクレオチド配列における各塩基は、完全なまたは実質的に完全なポリヌクレオチドプローブアレイから抽出した「n」塩基呼び出しセットによって縮重的にプロービングされ得る。

#### 【0114】

さらに、仮想タイリングの方法は、プローブセットが参照ポリヌクレオチドと比較した標的ポリヌクレオチドにおける欠失または挿入について、そして高度の正確さをもって試験するように構築することを可能にする。このような直接の比較は、系統的かつ再現性のよい交叉ハイブリダイゼーションならびに複数のプローブ標的相互作用から生じる潜在的に混乱性のシグナルを減少させる。従って、完全なまたは実質的に完全なn量体アレイおよび仮想タイリングの方法を用いて、任意の仮想ポリヌクレオチドについての変異を、そのポリヌクレオチドについてカスタム化されたアレイを構築することなく特徴付け得る。さらに、抽出された情報はまた、ポリヌクレオチド（例えば、ゲノムDNAフラグメント）についてハイブリダイゼーションによる一次配列決定のために、または他の手段によって達成される配列番号の正確さを評価するために使用され得る。

10

#### 【0115】

##### (d) 変異の存在を決定すること

本発明は、配列決定の目的のために使用され得るが、その最も強力な形態において、本発明は、標的ポリヌクレオチドの配列の決定を必要とせず、また参照ポリヌクレオチドの配列が公知であることも必要としない。むしろ、参照ポリヌクレオチドおよび標的ポリヌクレオチドのハイブリダイゼーションパターンが直接比較される。この比較から、標的ポリヌクレオチドを配列決定することなく、高度の正確さをもって標的ポリヌクレオチドにおける変異を決定し得る。

20

#### 【0116】

##### (i) 置換変異の特徴付け

置換変異について参照ポリヌクレオチドと標的ポリヌクレオチドとを比較するために、参照サンプルの公知の配列についての塩基呼び出しセットの完全なグループを表すプローブのグループを、プローブの各アレイから（好ましくは電氣的に）抽出する。各塩基呼び出しセットの各メンバーについて、標識の強度によって表される場合、標的ポリヌクレオチドフラグメントのハイブリダイゼーションの程度は、算術的な差異をとることによって、参照フラグメントのハイブリダイゼーションの程度を比較する。配列における所定の位置での塩基は、標的および参照の両方において同一である場合、塩基呼び出しセットにおける4つすべての塩基についてのハイブリダイゼーションの程度の差異がほぼゼロである。

30

#### 【0117】

しかし、標的ポリヌクレオチドが塩基において置換変異を含む場合、ハイブリダイゼーションの程度の差異は、塩基呼び出しセットの4つのメンバーの2つについて見られる。これは、もとの塩基を含むフラグメントに相補的なプローブおよび変異に存在する塩基を含むフラグメントに相補的なプローブを表す。図6（これは、HIV遺伝子の435位でのA G置換の特徴付けを示す）を参照のこと。塩基呼び出しセットの2つのメンバーの差異は、変異を含むサブ配列にハイブリダイズするすべての「p」塩基呼び出しセット（正方向および逆方向の両方の鎖が分析される場合は2nである）において見られる。標的における変異の存在は、変異部位を含む、広い（10～14塩基）フットプリントとして容易に視覚化される。図7cに示す配列トレースにおける変異スキンの差異を参照のこと。このフットプリントは、変異部位の上流および下流の重複するプローブに影響を与える単一の塩基変化から生じる（両方の鎖が使用される場合）。方法論はさらに、以下の実施例2で議論する。

40

#### 【0118】

さらなる視覚単純化は、差異のトレースの陽性および陰性エンベロープを単にプロットすることならびに関連する最大陽性値および最大陰性値を作成する対応する塩基を記録することによって実現され得る。この型のプロットは、「置換」変異スキンと言われる。

50

これを、図 8 に示す。図は、この方法によって正確に同定された 11 個の一塩基置換を明らかに示す。同定された置換は：G 1 4 A、A 1 2 8 G、C 2 0 6 T、A 4 3 5 G、T 6 1 4 C、A 6 7 9 G、G 6 9 3 A、A 8 3 2 T、A 8 6 5 G、A 9 2 7 G、T 9 5 5 C。

#### 【0119】

##### (ii) 挿入および欠失変異の特徴付け

上記の原理を用いて、挿入または欠失変異についての試験をするための塩基呼び出しセットを、置換変異について同時に試験するように構築し得る。検索される変異のセットは、挿入および欠失のタイリングプローブの塩基呼び出しセットを形成することによって拡張され得る。差異分析は、挿入および欠失について、置換の分析と同様に行い得る。図 9 (これは、p 53 遺伝子における挿入の検出を示す) を参照のこと。変異スキャンを、置換、欠失および挿入プローブセットについて同時に行い得る。これらのスキャンは、ある標的において 1 つの置換、および別の標的においては挿入を容易に同定し得る。図 9 を参照のこと。

#### 【0120】

##### (iii) ホモ接合およびヘテロ接合変異の特徴付け

多くの変異がホモ接合形態またはヘテロ接合形態のいずれかで存在し得る。本発明の方法は、ホモ接合変異およびヘテロ接合変異の両方を検出し得る。通常、変異染色体からのポリヌクレオチドは常に、それがハイブリダイズし得る完全なまたは実質的に完全なアレイにおいてプローブに遭遇する。従って、これは、常に参照ポリヌクレオチドの全体パターンとは異なる全体パターンを生成する。しかし、ヘテロ接合変異において、標的の非変異染色体からのポリヌクレオチドは、参照ポリヌクレオチドにマッチし得る。これらの差異は「完全マッチ (perfect match)」(PM) プロットを用いて検出し得る。PM プロットにおいて、参照ポリヌクレオチドについて完全にマッチする塩基呼び出しセットにおいて、標的ポリヌクレオチドおよび参照ポリヌクレオチドの両方のハイブリダイゼーションの程度が比較される。

#### 【0121】

図 10 は、CFTR 遺伝子の 1079 bp 領域におけるホモ接合シナリオおよびヘテロ接合シナリオを比較する PM プロットである。図 10 a は、野生型 CFTR ポリヌクレオチド配列を、3 bp DF508 (TTT) ホモ接合欠失を含む標的と比較する、170 bp 領域の変異スキャンを示す。完全マッチ (PM) 差異プロットは、2 つのアレイからの参照についての完全マッチの間の差異を基準化した強度を表示する。変異スキャンはまた、置換 (陽性エンベロープのみ)、挿入、および欠失プローブについて示される。興味深いことに、3 bp 欠失 (分析ソフトウェアで「TTT」と同定される) は、挿入スキャンによって容易に検出された。

#### 【0122】

ホモ接合変異と対照的に、差異は、ヘテロ接合変異を含む PM プロットにおいては明らかではない。これは、上記で議論したように、ヘテロ接合変異において、参照染色体となお同一な染色体が 1 つ存在し、これが PM 塩基呼び出しセットとハイブリダイズするポリヌクレオチドフラグメントを提供するからである；従って、参照と標的との間には明らかな差異が存在しない。これを、図 10 b に示す。フットプリントは、PM 差異スキャンにおいて検出可能ではなかった。なぜなら、野生型参照配列を有する DNA が両方のサンプルに存在するからである。しかし、興味深いことに、欠失スキャンは、ヘテロ接合サンプルにおいて 3 bp 「TTT」欠失を正確に同定した。

#### 【0123】

同様に、図 11 は、2.5 kb p ミトコンドリアアンプリコンのホモ接合およびヘテロ接合変異スキャンを示す。図 11 a は、ホモ接合サンプルの 300 bp 変異スキャンを示す。このスキャンは、3 つの 1 塩基置換を示した。8 つの異なるサンプルの中で、合計 176 個の配列差異が検査された。ホモ接合ポリヌクレオチドの変異スキャンは、これらの配列差異の 90% より多く正確に同定された。さらに、偽陽性率は、非常に低かった (スクリーニングした 3900 bp のうち < 1)。図 11 b は、図 11 a と同じ 300 bp ス

10

20

30

40

50

キャンを示すがサンプルはヘテロ接合である。このスキャンは、3つの1塩基置換を示すが、シグナル対ノイズ比は減少した。一般に、2.5 kbp アンプリコンにおけるヘテロ接合変異は、減少した感度(70%~80%)で検出された。

#### 【0124】

まとめると、本発明の方法は、例えば、ホモ接合またはヘテロ接合のような変異の決定、ならびに変異の位置および型の同定を可能にする。

#### 【0125】

(e) 未知の配列の2つ以上のポリヌクレオチドが同一であるか否かの決定

上記の方法は、未知の配列の2つ以上のポリヌクレオチドが同一であるか否かを決定するように容易に適合され得る。例えば、未知の配列の2つのポリヌクレオチドは、2つの異なる同一のまたは実質的に同一の、完全なまたは実質的に完全なn量体プローブポリヌクレオチドのアレイに別個にハイブリダイズされ得る。それらのハイブリダイゼーションパターンを、上記の方法を用いて比較し得る。2つの配列が同一である場合、PMプロットにおける完全なマッチが予測される。上記のように、この方法論は、見かけ上種々の供給源由来であるポリヌクレオチドが実は単一の供給源由来であるか否かを、いずれのポリヌクレオチドの配列を知ることなくして、迅速にかつより正確に決定するのに使用し得る。

10

#### 【0126】

(f) 配列特異的末端サンプリング

ポリヌクレオチドの混合物における個々のポリヌクレオチドを識別する方法もまた提供される。ポリヌクレオチドのこのような混合物は、例えば、部位特異的DNAヌクレアーゼを用いるDNA標的のフラグメント化から生じ得る。このようなフラグメント化は、すべての可能な5'末端サブ配列または3'末端サブ配列の小下位集団を代表するのみであるフラグメントの集団を生成し得る。n量体アレイは、すでに上記に記載した方法論を用いてこのような複合体サンプルにおける所望の末端を同定するのに使用され得る。完全なまたは実質的に完全なn量体アレイに対して、ポリヌクレオチドまたはそのフラグメントの複合体サンプルをハイブリダイズさせることによって観察されるハイブリダイゼーションパターンは、それ自身を、目的のポリヌクレオチドまたはそのフラグメントの数および型を決定するのに役立たせる。複合体集団における特定のポリヌクレオチドを識別する方法は、ゲルベースまたは電荷ベースの方法のような伝統的な方法と比較してより迅速かつより正確である。潜在的な適用には、DNAフィンガープリントまたは差示的ディスプレイのようなランダムまたは指向性ゲノムサンプリング戦略によって生成されるDNAフラグメントの分析が挙げられる。

20

30

#### 【実施例】

#### 【0127】

以下の実施例は、アレイの合成、標的ポリヌクレオチドおよび参照ポリヌクレオチドのフラグメント化、ハイブリダイゼーションおよび連結、アレイの洗浄および染色、ならびにデータ分析を詳細に記載する。これらの実施例は、本明細書に記載される方法をさらに明らかにするが、それらは、本発明の範囲を限定するものとみなされるものではない。

#### 【0128】

(一般的技術)

完全n量体デオキシポリヌクレオチドを、Peaseら(1994)、Proc Natl Acad Sci USA 91, 5022-6(その全体を本明細書において参考として援用する。)によって以前に記載されるように、光指向性光化学を用いて、誘導体化したガラス基板上で合成した。8量体アレイを、50x50μmプローブフィーチャーサイズを用いて、および9量体を25x25μmプローブフィーチャーサイズを用いて合成する。光化学合成の前に、DMTヘキサエチレングリコール-(2-シアノエチル-N,N-ジイソプロピル)ホスホルアミダイトリンカー(例えば、ChemGenesからのもの、アセトニトリル中に25mM)および19塩基の定常アンカーポリヌクレオチド(3'ATACGTAGACACTGCTGGAC5')を従来のDNA PACホス

40

50

ホルアミダイト (アセトニトリル中 25 mM) ; T シアノエチル、IBU d C シアノエチル、PAC d A - シアノエチル、および iPr - PAC d G - gal シアノエチル ホスホルアミダイト (例えば、Pharmacia Biotech から入手可能なもの) を用いてアレイ上で合成した。

#### 【0129】

アンカーポリヌクレオチドの最後の塩基および続きの n (本実施態様において n = 8 または 9) の組合せ塩基を、光指向性光化学 (Pease ら、1994) を用いて合成した。これらの n 塩基の合成後、2 塩基のイノシン (例えば、Glen Research からの 2' - デオキシイノシン) を従来の DMT ホスホルアミダイト化学を用いて 5' 末端に付加した。2 つのイノシンの付加は、識別に多大な影響を与えることなくアレイ全体のシグナル強度を改善する。合成後、アレイ上の DNA 分子を 50% エチレンジアミン (EDA) / 50% エタノール浴中で約 10 時間脱保護化した。

10

#### 【0130】

プローブの定常アンカー配列に対して相補的な 20 量体アンカーポリヌクレオチド (5' ソラレン - TATGCATCTGTGACGACCTG - 3') を、標準的なポリヌクレオチド合成機を用いて合成した。2 つのバージョンのポリヌクレオチドを異なる様式の連結 (酵素連結のための 3' ヒドロキシルバージョンおよび化学連結のための 3' ホスホリルバージョン) のために合成した。アンカーポリヌクレオチドはまた、必要に応じて、交叉連結プロトコルにおける使用のために、5' 末端に AT ジヌクレオチド対およびソラレン C6 (例えば、Glen Research から) 部分を含んだ。ポリヌクレオチドを疎水性 5' ソラレン部分を介してカートリッジ精製した。末端 G 3' 塩基を有するアンカーポリヌクレオチドを使用した。実際のアンカーベースの二重鎖プローブを、相補的 20 量体アンカーポリヌクレオチド (500 nM) を、アレイ結合プローブの定常部分に対してハイブリダイズさせることによって作製した。

20

#### 【0131】

ハイブリダイゼーション条件は以下のとおりである: 6 x SSPE 緩衝液 (pH 7.4)、40 で 1 ~ 2 時間、続いて GeneChip 液性ステーション (例えば、Affimetrix, Inc. からの型のもの) 上での 6 x SSPE で 22 での洗浄。この実施態様において使用した連結条件下 (30 ~ 40) では、ソラレン交叉連結は、二重鎖安定性の維持のためには不必要であった。

30

#### 【0132】

二本鎖 DNA 標的 (約 50 ~ 100 fmol、約 80 ~ 160 ng の 2.5 kb 標的) を、緩衝液 (10 mM Tris 酢酸、pH 7.5、10 mM 酢酸マグネシウム、50 mM 酢酸カリウム) 中の DNase I (0.5 U / 20 μl) で 15 分間、37 で、ランダムに断片化した。DNase I を熱不活化し、そしてサンプルを、95 で 15 分間でのインキュベーションにより変性させた。サンプルを氷上で冷却し、そしてターミナルジデオキシヌクレオチジルトランスフェラーゼ (5 U / 20 μl) (TdT) (Gibco BRL, Gaithersburg MD) および ddATP - N6 - ビオチン (25 μM) (NEN Life Sciences, Boston, USA) を、37 で 60 分間用いて 3' 末端標識した。TdT を熱不活化し、そして DNA を、95 で 10 分間インキュベートすることによって変性させた。変性させた DNA を氷上で冷却し、連結混合物 (総量 200 μl、DNA 標的の最終濃度約 500 pM、以下を参照のこと) に添加し、微量遠心分離機中で 3 分間、14000 g でスピンドして、存在する任意の沈澱物をペレット化し DNA アレイに適用した。連結溶液を、所望の温度に設定したインキュベーターに配置した回転器 (rotisserie) での回転 (40 ~ 50 r.p.m.) によってアレイの表面を横切って連続的に混合した。

40

#### 【0133】

酵素連結または化学連結のいずれかを使用した。各々が、特定のプロトコルに従って使用した場合に、満足のいく結果を生成した。酵素連結条件を、以下の酵素について与える: T4 DNA リガーゼ、E. coli (NAD 依存性) DNA リガーゼ、および Taq

50

DNAリガーゼ。標準的なT4連結緩衝液は、以下からなった：50mM Tris-HCl (pH7.8)、10mM MgCl<sub>2</sub>、10mM DTT、1mM ATP、50μg/ml BSA、100mM NaCl、0.1% TX-100、および2.0U/μl T4 DNA (pH8.0)、10mM DTT、1mM ATP、50μg/ml BDA、100mM NaCl、0.1% TX-100、および2.0U/μl T4 DNAリガーゼ (New England Biolabs, Inc.) E. coli DNAリガーゼ緩衝液は、以下からなった：40mM Tris-HCl (pH8.0)、10mM MgCl<sub>2</sub>、5mM DTT、0.5mM NADH、0.5μg/ml BSA、0.1% TX-100、および0.025U/μl E. coli DNAリガーゼ (Amersham)。

10

## 【0134】

Taq DNA連結緩衝液は以下からなる：20mM Tris-HCl (pH7.6)、25mM 酢酸カリウム、10mM 酢酸マグネシウム、10mM DTT、1mM NADH、50μg/ml BSA、0.1% Triton X-100、10% PEG、100mM NaCl、および1.0U/μl Taq DNAリガーゼ (New England Biolabs)。8量体および9量体の両方について、T4 DNAリガーゼ反応、およびE. coli DNAリガーゼ反応を30で行い、そしてTaq DNAリガーゼ反応を37~40 (30でのTaqの低活性に起因する)で行った。連結反応物を、代表的に、一晚(14~16時間)インキュベートしたが、2時間だけのインキュベーションもまた、満足のいく結果を与えた。

20

## 【0135】

化学連結の間、3~4M 塩化テトラメチルアンモニウム (TMACl) を、A/TリッチハイブリッドおよびG/Cリッチハイブリッドの安定性を標準化するのを補佐するために、連結緩衝液中で使用した。Woodら、Proc. Natl. Acad. Sci. U.S.A. 82、1585-1588 (1985)を参照のこと。その全体を本明細書において参考として援用する。新たに溶解した1-エチル-3-(3-ジメチルアミノプロピル)カルボジイミドジヒドロクロリド (EDC, Pierce Biochemicals) (10倍ストックとしてH<sub>2</sub>O中に2M) を使用して、標的の5'ホスホリル部分と、相補的なアンカーオリゴヌクレオチド (合成中に添加した) 上の3'ホスホリル部分との間のピロリン酸結合を生成した。例えば、Hermanson, Bioconjugate Techniques: Academic Press (1996); Shabarova, Biochimie 70:1323-1334 (1988); Kuznetsovaら、Mol. Biol. (Mosk.) 28:290-299 (1994)を参照のこと。これらはすべて、その全体を本明細書において参考として援用する。化学連結条件は：50mM 2-[N-モルホリノ]エタンスルホン酸 (MES) (KOHを用いてpH6.0)、10mM MgCl<sub>2</sub>、0.001% SDS、200mM EDC、50mM イミダゾール (HClを用いてpH6.0)、および3.0~4.0M 塩化テトラメチルアンモニウム (TMACl) (Sigma)、30で14時間であった。

30

## 【0136】

連結後、アレイを5~10回、1xSSPE (pH7.4、22) で、GeneChip (登録商標) 液性ステーション上で洗浄し、回転する回転器で22で5分間ストレプトアビジン-フィコエリトリン結合体 (Molecular Probe、1xSSPE、50μg/ml BSA中に2ng/μl) を用いて染色し、そしてさらに5~10回、1xSSPEで洗浄した。

40

## 【0137】

アレイからの蛍光シグナルを、特別に設計した共焦点スキャナー (Affymetrix, Inc., Santa Clara, CA) を用いて検出した。Cheeら、Science 274、610-614 (1996)を参照のこと。これは、本明細書において参考として援用する。スキャナーは、8量体アレイを6.8~7.5μmの解像度で画

50

像化し、そして9量体アレイを3.5 μmの解像度で画像化した。イメージを、GeneChip (登録商標)ソフトウェア (Affymetrix, Inc., Santa Clara, CA)を用いて処理し、そして得られた強度データを、プローブ配列および対応する強度のアルファベット化したリストを含むテキストファイルに変換した。次いで、これらのデータを、Visual Basic (登録商標) (v. 4.0, Microsoft)で書かれた特注のソフトウェアにより分析した。

**【0138】**

(実施例1:塩基呼び出しの正確さと、プローブ長および標的長との間の関係)

閾値に関して、標的長またはプローブ長のいずれかに対する塩基呼び出しの正確さの依存度を決定すべきである。この方法を使用して、種々の長さのポリヌクレオチドを配列決定し、次いで呼び出された配列を、実際の公知の配列と比較した。異なる長さのDNA配列を、8量体プローブアレイおよび9量体プローブアレイの両方にハイブリダイズさせ、そして酵素連結および化学連結を用いて連結させた。ヒトCFTR遺伝子からの535 bp PCR産物; HIV-1 pol遺伝子からの1.2 kb PCR産物; ミトコンドリアのシトクロームb遺伝子および制御領域を跨ぐ2.5 kb PCRフラグメント; ならびにバクテリオファージ X174の5.4 kb ゲノムを含む種々の配列を分析した。

10

**【0139】**

結果は、標的長およびプローブ長が、塩基呼び出しの正確度に有意に影響することを示す。表1を参照のこと。標的長が増大するにつれて、塩基呼び出し性格度は、535 bpにおける100%から(9量体に関して)5.4 kbにおける89%にまで減少した。信頼性閾値を使用することにより、不正確な塩基呼び出しの数を、最小限に抑えた。89%の正確度で呼び出された5.4 kb 標的の場合において、残りの11%の塩基呼び出しは、10.6%の呼び出しなし、およびほんの0.4%の不正確な呼び出し(偽陽性)からなった。さらに、89.4%(89.0%+0.4%)の実際に呼び出された塩基について、正確な呼び出し率は、99.4%(100%×89.0/89.4)である。従って、塩基呼び出しの呼び出し/非呼び出しカテゴリーへの分類は、この塩基呼び出し情報の有用性を大いに増大させた。

20

**【0140】**

一般に、標的長の増大に伴う塩基呼び出し正確度における減少は、プローブの独特性の欠如から生じる呼び出しなしに主に起因している。プローブ長の増大は、呼び出しなしの数を劇的に減少させ、そしてはるかにより高い塩基呼び出し正確度をもたらす。以下の表1に示すように、5.4 kb 標的についての8量体対9量体のプローブでの塩基呼び出し正確度における対比に特に注意されたい。他方、連結の様式は、ほんのわずかの効果を有していた。酵素連結は、ほとんどの場合、化学連結よりもわずかにより高い塩基呼び出し正確度を有していた。

30

**【0141】**

【表 1】

標的の複雑性、プローブ長、および連結方法の関数としての塩基呼び出し正確度

		540bp CFTR	1.2kb HIV	2.5kbp Mito	5.4kbp $\phi$ X174
酵素連結	8マー(%)	99.0(0.2)	93.2(0.1)	88.9(1.3)	73.9(4.1)
	9マー(%)	99.8(0.0)	99.9(0.0)	94.5(0.3)	85.2(0.5)
化学連結	8マー(%)	99.0(0.4)	96.1(0.5)	87.8(1.5)	69.0(2.8)
	9マー(%)	100.0(0.0)	*99.5(0.0)	93.4(0.)	89.0(0.4)

10

不正確な呼び出し率を、括弧内に示す。酵素連結は、標準的なT4 DNAリガーゼ条件を使用した。1.15の信頼性閾値を塩基呼び出しについて使用した。非呼び出し率+不正確な呼び出し率+塩基呼び出し正確度=100%。\*Standard deviation=0.1% (n=4回の実験)。

(実施例2: HIV-1 pol 遺伝子における置換変異の特徴付け)

HIV-1 pol 遺伝子を、置換変異に供される遺伝子の例として使用した。例示の目的で、2つの1.2kbpの遺伝子のアンプリコンを分析した。

20

## 【0142】

第一に、単一置換変異を分析した。トレースを、上記に記載のように、2つの異なる1.2kbp HIVサンプル(参照サンプルおよび標的サンプル)におけるA-G塩基置換について、作製し、別個の8量体アレイに連結し、構築した。図6を参照のこと。サンプルを断片化し、そして上記に記載の技術を用いて別個の8量体アレイに連結し、そして参照配列内の435位で問い合わせを行う8塩基呼び出しセットを分析した。(参照と比較して標的は、)G尋問プローブの強度が増加し、そしてA尋問プローブ(の強度)が減少した。このことは、A-G塩基置換を示す。

## 【0143】

この実施例において、最初の2つの塩基呼び出しセットは、差を示さなかった；これは、おそらく、これらのプローブに対するサンプル中の他のサブ配列の結合に起因する。しかし、各塩基についての高度の縮重度のために、435位についての他の塩基呼び出しセットは、最初の2つの塩基呼び出しセットによって生成される系統席な塩基呼び出し「誤差」を除去する。置換変異は、変異部位の周りの12から14塩基を検査することによって容易に特徴付けられる。なぜなら、変異は、変異部位(A435G)の周りに中心を置いた特徴的な「フットプリント」を残すからである。これを、図7cに示す。

30

## 【0144】

さらなる実験において、2つの1.2kbp HIVポリヌクレオチド(参照および標的)を断片化し、そして標準的なプローブアレイに連結した。図8に示すように、11/11の変異を、偽陽性を伴うことなく変異スキャンによって正確に同定した。潜在的な偽陽性変異は、770位(図8でアスタリスクによって示す)での狭いフットプリントとして観察された。しかし、このフットプリントは、予め選択した閾値内に収まっており、そしてより重要なことに、このフットプリントは、実際の変異の広範なプロフィール特徴を示さなかった。

40

## 【0145】

HIV遺伝子を用いる第三の例は、本発明がまた、より多様なセットの標的の間の配列差異を検出するのに適していることを実証する。27の一塩基置換を含む第二の1.2kbp HIV改変体サンプルの変異スキャンを行った。8量体アレイにおいて、96.3%(26/27、1偽陽性)の変異を検出し、そして9量体アレイにおいて、100%(27/27、1偽陽性)の変異を検出した。問い合わせされた標的配列が参照配列から分岐

50

しているので、偽陽性の確率が塩基変化によって影響される多数のプローブの数に起因して増加することに注意するべきである。これらの結果は、40塩基中多くて1塩基違うキロベースサイズの標的が、このアプローチを用いて正確に特徴付けされ得ることを示す。

#### 【0146】

(実施例3：より長いポリヌクレオチド配列(ミトコンドリアDNA)における置換変異の特徴付け)

176の公知の配列差異について8つの異なる2.5kbpのヒトミトコンドリアシトクロームbアンプリコンをスクリーニングした。8量体アレイを用いて、90%を超える(160/176)配列差異を、3900塩基あたり1塩基未満の低い偽陽性率(<0.02%未満)をともなって正確に同定した。比較アプローチ(すなわち、別個に決定した配列と比較すること)をしなければ、偽陽性率は、かなりより高かった(約1%~2%)。偽陰性(検出されない変異は、通常、低いプローブシグナル強度を生成する標的領域(例えば、ATリッチ領域)か、または反復配列を含む領域に生じた。このアプローチを用いて有効にスクリーニングされ得る最大の標的長は、寛容される偽陽性(不正確に変異として同定される)数と偽陰性(検出されない変異)の数との間のトレードオフ(訳注、すなわち、両者のどちらをとるか)に依存する。従って、変異の見出しの目的のために、8量体アレイを使用して、高感度(代表的には、90%を超える)および最少数の偽陽性(3900塩基あたり1未満)を伴って、標的2.5kbp長をスキャンし得る。

#### 【0147】

さらにより長い片のDNAをスクリーニングを、16.6kbpヒトミトコンドリアゲノムの全体を用いて行った。2つの異なる個体からの3つの重複するPCRアンプリコンからなる2つの16.6kbp標的を使用した。配列の解像度を改良するために、調製した標的を9量体アレイに連結した。変異スクリーニングは、16,600塩基あたり4塩基(0.024%)の低い偽陽性率を伴って、66%を超える(16/24)の公知の配列差異を同定した。これらの結果は、完全なn量体DNAアレイにおける比較再配列決定を使用して、変異について非常に複雑なDNAサンプルを迅速にスクリーニングし得ることを示す。

#### 【0148】

(実施例4：p53遺伝子における挿入変異および欠失変異の特徴付け)

本明細書において記載される塩基呼び出し法および仮想タイリング法は、置換変異についてのみならず、挿入変異および欠失変異についてもまた遺伝子を検査することを可能にする。この局面を示すために、尋問変異のセットを、挿入タイリングプローブおよび欠失タイリングプローブの塩基呼び出しセットを形成することによって、拡張した。差異分析を、置換の分析と同様に、挿入および欠失について行った。未知の相同p53(11エキソン)標的(1つのアレイに適用される)を、野生型p53参照(第二のアレイに適用される)と比較した。変異スキャンを、置換プローブセット、欠失プローブセット、および挿入プローブセットについて行った。図9は、変異スキャンを示す。これらのスキャンは、エキソン4におけるG C置換およびエキソン5におけるC挿入を容易に同定した。図9aを参照のこと。

#### 【0149】

他方、エキソン5における変異スキャンは、置換スキャンおよび挿入スキャンの両方におけるフットプリントを示した。図9bを参照のこと。エキソン5の置換スキャンおよび挿入スキャンの両方におけるフットプリントの存在は、どの型の変異が存在していたかを決定する際に困難をもたらす。しかし、一塩基置換を含むHIVサンプルの以前の分析は、真正の置換変異が、代表的には、挿入スキャンまたは欠失スキャンにおけるフットプリントを表示しないことを示唆する。従って、エキソン5スキャンにおける挿入フットプリントの観察は、従来配列決定によって検証された挿入呼び出しを強く支持する。

#### 【0150】

(実施例5：CFTR遺伝子におけるホモ接合型変異およびヘテロ接合型変異の特徴付け)

10

20

30

40

50

8量体アレイの、公知の F508 欠失 (3bp) を含む嚢胞性線維症膜貫通伝達レギュレーター (CFTR) アンプリコン (約 1.1 kbp) におけるホモ接合型変異およびヘテロ接合型変異を検出する能力を調べた。野生型 (正常) CFTR サンプルを参照として用いた。変異スキャンを図 10 に示す。F508 欠失を、PM 参照プローブスキャン (完全一致参照プローブにおける差異のスキャン)、置換スキャン、および欠失スキャンにおけるフットプリントの存在によってホモ接合型サンプルにおいて容易に検出した。図 10 a を参照のこと。図 10 b は、ヘテロ接合型サンプルについての変異スキャンを示す。ヘテロ接合型欠失の場合において、フットプリントは PM 参照プローブスキャンにおいて見られなかった。なぜなら、参照配列を有する DNA が、比較しているサンプル両方に存在するからである。しかし、欠失スキャンは、良好に解明されたフットプリントをなお示す。なぜなら、欠失された配列を含む標的がヘテロ接合型サンプルに存在するが、参照サンプルからは完全に欠如しているからである。従って、変異は、ヘテロ接合型欠失として特徴付けられ得る。

10

#### 【0151】

方法を、2.5 kbp ミトコンドリアアンプリコンのより長い「ヘテロ接合型」サンプルへと拡張した。図 11 は、2.5 kbp ミトコンドリアアンプリコン (ホモ接合型対ヘテロ接合型) の変異スキャンを示す。ホモ接合型サンプル (野生型とホモ接合型 DNA との 50:50 混合物) の 300bp 変異スキャンによって、3つの一塩基置換が明らかになった。図 11 a を参照のこと。8つの異なるサンプルのうち合計 176 の配列差異を検査した。ホモ接合型サンプルの変異スキャンは、これらの配列差異の 90% を超える正確さで同定した。さらに、偽陽性率は、非常に低かった (スクリーニングした 3900bp あたり 1 未満)。

20

#### 【0152】

同じ 300bp スキャンを、「ヘテロ接合型」サンプル (野生型およびヘテロ接合型 DNA の 50:50 混合物) を用いたことを除いて、反復した。変異スキャンを図 11 b に示す。このスキャンによって、3つの一塩基置換が明らかになったが、解像度は減少していた。一般に、2.5 kbp アンプリコンにおける接合ヘテロ型変異は、感度を減少して検出された。

#### 【0153】

(実施例 6: バクテリオファージ X174 の部位特異的末端サンプリング)

30

部位特異的末端サンプリングについての本発明の使用の例として、バクテリオファージ X174 (約 5.4 kbp) を、制限エンドヌクレアーゼ P a l I (G G C C) および N l a I I I (C A T G) を用いて消化し、そして得られたフラグメント集団を、8量体アレイに対して連結した。結果は、33の制限部位のうち32を検出したことを示し、このことは、n量体アレイを、複合サンプルにおける5'末端サブ配列を同定するのに使用し得ることを示す。

#### 【0154】

上記の記載は、例示的であり、そして制限的ではない。本発明の多くの改変は、本開示を参照すれば、当業者には明らかとなる。例のためのみに、種々の基質、ポリマー、連結基、合成開始部位、および他の材料が本発明の範囲を逸脱することなく使用され得る。従って、本発明の範囲は、上記の記載を参照して決定されるべきではなく、代わりに、添付の請求の範囲を、それらの等価物の全範囲とともに参照して決定されるべきである。

40

#### 【図面の簡単な説明】

#### 【0155】

【図1】 n量体アレイの設計の基礎を示す模式図である。図解のために、5'一本鎖オーバーハングを有する二本鎖定常アンカー領域が示される。この図はまた、必要に応じた共有結合的架橋工程 (ここではソラレンを用いて例示した) を示す。

【図2】 1.2 kb HIV アンプリコンの8量体アレイへのライゲーションの効果を示す図である。図 2 a はリガーゼの非存在下でのハイブリダイゼーションを示す。図 2 b は、T4 DNA リガーゼの存在下でのハイブリダイゼーションおよびライゲーションを示す

50

。図 2 c は、個々のプローブの特徴を例示する図 2 b の拡大図を示す。

【図 3】参照配列を使用した、塩基呼び出しセットの構築を示す図である。図 3 a は、540 bp の DNA 標的の参照配列（順方向鎖）の部分がそこにライゲートしている、8 量体アレイを示す。配列内の 77 位を尋問する 8 塩基呼び出しセットのプローブ配列が示される。各々の塩基呼び出しセットは、単一置換位置（文字 N で示す）が異なる 4 つのプローブから成り、特定の標的塩基を尋問するように位置する。図 3 b は、アレイデータから抽出され、そしてプロットされた、各々の塩基呼び出しセット内のプローブの強度を示す。プローブは、A - プローブは実際に相補的な「T」塩基置換（順方向鎖照会）または「A」塩基置換（逆方向鎖照会）を含むように、列挙される。図 3 c は、すべての 8 塩基呼び出しセットを横断する尋問プローブ（A、C、G、T）の合計を示す。この合計は、順方向鎖および逆方向鎖の両方について合計した塩基呼び出しセットを生じる。これらの合計のセットに由来する複合性の塩基呼び出しセットは、最終的な高度に正確な塩基呼び出しが作製されることを可能にする。

10

【図 4】「配列決定トレース」における塩基呼び出し情報の表示を示す図である。

【図 5】呼び出しのないカットオフ点を提供するための、塩基呼び出し品質の使用法を示す図である。図 5 a は、HIV 標的配列内の AT リッチな領域の配列トレースによって実証された、低品質塩基呼び出し（強度値間の不十分な分離）を示す標的領域の例である。図 5 b は、カットオフの閾値がいくつかの不正確な塩基呼び出しを排除したことを図示する。

【図 6】別個のアレイ上の DNA 標的の塩基呼び出しセットを比較することによる変異の発見を示す図である。図は、HIV 遺伝子の 435 位における A G 置換を図示する。図 6 a は、参照（左パネル）および標的（右パネル）に対するアレイ像の小さい部分を示す。各々のアレイ上の円で囲んだプローブ（尋問位置 435）は、第 3 の位置にある単一塩基置換によって異なる。図 6 b は、参照および標的の両方についての 435 位を尋問する 8 残基コールセットの標準化された強度を示す。これらの塩基呼び出しセット間の強度の違い（参照に対する標的）がまた示される。図 6 c は、参照および標的あるいはそれらの複合の差（I 未知 - I 参照）に対する複合の塩基呼び出しセット（組み合わせた両方の鎖）を示す。

20

【図 7】変異の容易な同定を可能にする配列トレースの直接比較を示す図である。図 7 a および 7 b は、2 つの 1.2 kb HIV 遺伝子ポリヌクレオチド（参照および標的）から生じる配列トレースを示す。図 7 c は、配列トレースにおける差異（変異スキャン）が、どのように変異部位（A 435 G）において 12 ~ 14 塩基の幅のフットプリントを示すかを図示する。

30

【図 8】HIV pol-1 遺伝子中における 1041 塩基の変異スキャンを示す図である。異なるトレースの正および負の包絡線は、1.2 kb HIV 配列の 1041 塩基についてプロットされる。図は、この方法によって正しく同定された 11 の単一の塩基置換を示す。星印は、ピークが変異に特徴的な広範なプロフィールを表示し損なうために負とスコアされる、潜在的な変異を示す。

【図 9】p53 遺伝子における挿入の検出を示す図である。p53 遺伝子の 11 のエキソンが同時にスキャンされた（プライマーを含み約 1700 塩基）。図 9 a は、エキソン 4 の置換スキャン（正の包絡）が、どのように正しく G C 塩基変化を同定した変異フットプリントを明らかにしたかを図示する。図 9 b は、エキソン 5 の変異スキャンが置換スキャン（正の包絡）および挿入スキャンの両方においてフットプリントを示すことを示す。挿入スキャンは、変異を標的サンプルにおける C 挿入（相補的な G 挿入プローブの増加）として同定し、一方置換スキャンは不明瞭であった。

40

【図 10】同型接合性欠失および異型接合性欠失についての完全一致（PM）差プロットディスプレイを示す図である。図 10 a は、野生型 C F T R ポリヌクレオチド配列を、3-bp D F 508 (T T T) 同型接合性欠失を含む標的と比較する、170 bp 領域の変異スキャンを示す。変異スキャンはまた、置換（正の包絡のみ）、挿入、および欠失プローブについて示される。3-bp 欠失（分析ソフトウェアによって「T T T」として同

50

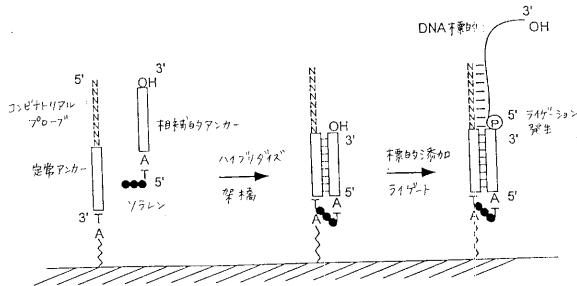
定される)は、欠失スキャンによって容易に検出された。図10bは、CFTR遺伝子の1079bp領域において野生型CFTR標的を3-bp F508「人工」異型接合性欠失変異体と比較する、変異スキャンを示す。野生型参照配列を有するDNAが両方のサンプルにおいて存在するため、PM差スキャンにおいてフットプリントは検出されなかった。しかし、欠失スキャンは、異種接合性サンプルにおいて3-bp「TTT」欠失を正しく同定した。

【図11】2.5kbpミトコンドリアアンプリコンの変異スキャンを示す図である(同型接合性対異型接合性)。図11aは、同型サンプルの300bp変異スキャンを示す。このスキャンは3つの単一塩基置換を示した。8つの異なるサンプルの間で総計176の配列の差異が試験された。同型ポリヌクレオチドの変異スキャンは、これらの配列の差異を90%以上も正しく同定した。さらに、偽陽性の割合は非常に低かった(スクリーニングされた3900bpあたり1未満)。図11bは、サンプルが野生型および変異した異型ポリヌクレオチドの50:50混合物である以外は図11aと同様である300bpスキャンを示す。このスキャンは3つの単一の塩基置換を示したが、減少したシグナル/ノイズ比を有した。

10

【図1】

二重鎖プローブ量体アレイの設計

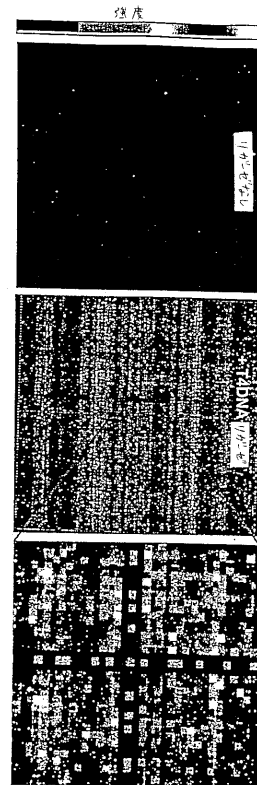


【図2】

2a

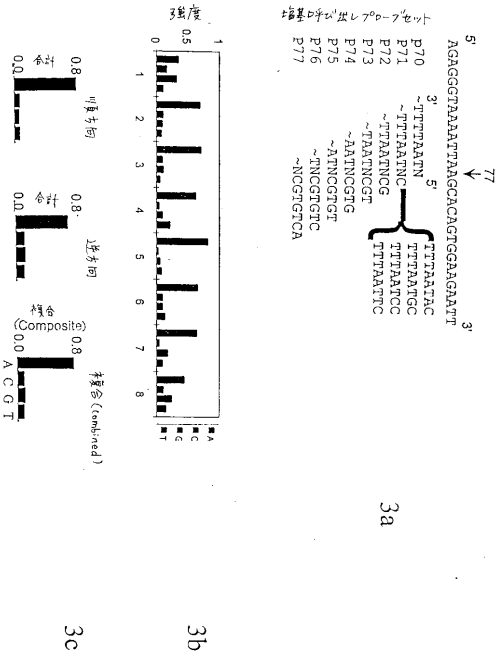
2b

2c



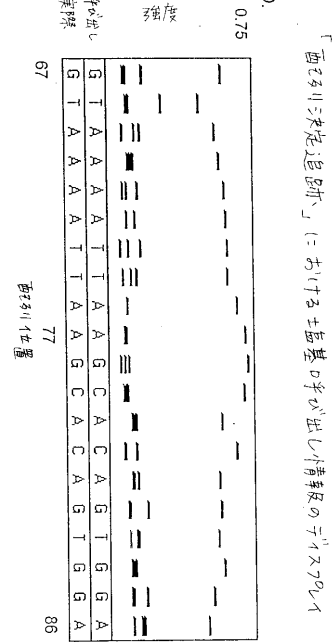
2kb HIV-1 T29011-1の2重鎖量体アレイへのライゲーション

【図3】

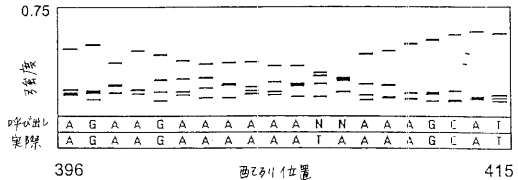


参照面2列1を1列とした塩基の呼び出しセットの構築

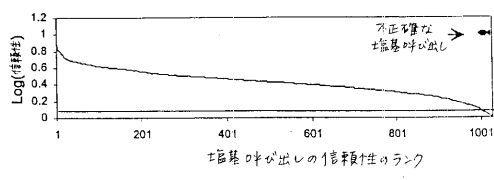
【図4】



【図5】

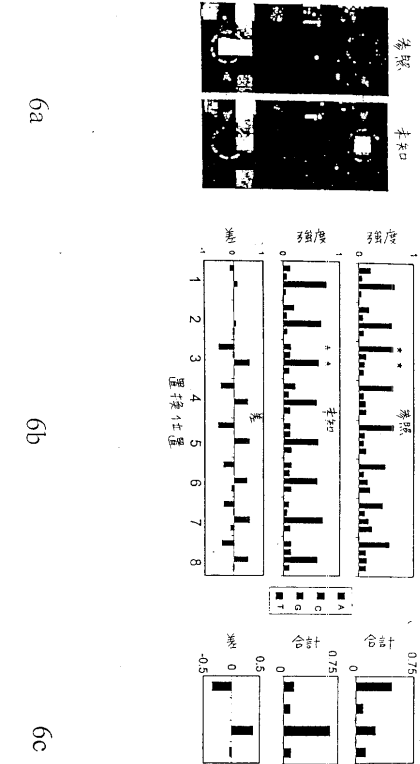


5a



5b

【図6】



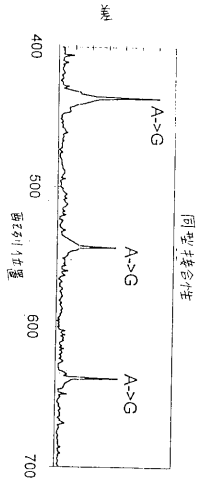
6a

6b

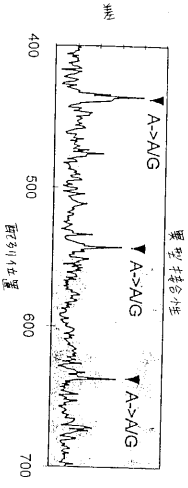
6c



2,5kbP ミトコンドリア DNA の変異不十分  
(同型挿合性対異型挿合性)



11a



11b

【 1 1 1 】

---

フロントページの続き

(72)発明者 ケビン エル． ガンダーソン  
アメリカ合衆国 カリフォルニア 92024, エンシニタス, ジュニパー ヒル ドライブ  
543

(72)発明者 マーク エス． チー  
アメリカ合衆国 カリフォルニア 92014, デル マー, 15ティーエイチ ストリート  
ナンバー24 155

(72)発明者 デイビッド ジェイ． ロックハルト  
アメリカ合衆国 カリフォルニア 92014, デル マー, トリー ポイント ロード 5  
10

Fターム(参考) 4B024 AA11 CA02 HA12  
4B063 QA12 QQ43 QR32 QR55 QS34 QX01

专利名称(译)	使用完整的N-mer阵列进行核酸分析		
公开(公告)号	<a href="#">JP2006204302A</a>	公开(公告)日	2006-08-10
申请号	JP2006044556	申请日	2006-02-21
[标]申请(专利权)人(译)	阿菲梅特里斯公司		
申请(专利权)人(译)	Affymetrix公司公司		
[标]发明人	ケビンエルガンダーソン マークエステー デイビッドジェイロックハルト		
发明人	ケビン エル. ガンダーソン マーク エス. チー デイビッド ジェイ. ロックハルト		
IPC分类号	C12Q1/68 C12N15/09 C12M1/00 C12Q1/6809 C12Q1/6827 C12Q1/6837 G01N33/53 G01N37/00		
CPC分类号	C12Q1/6837 C12Q1/6809 C12Q1/6827 C12Q2600/156 C12Q2565/501 C12Q2535/131		
FI分类号	C12Q1/68.ZNA.A C12N15/00.A C12Q1/68.AZN.A C12Q1/6809.C C12Q1/6809.Z C12Q1/6827.C C12Q1/6827.Z C12Q1/6837.C C12Q1/6837.Z		
F-TERM分类号	4B024/AA11 4B024/CA02 4B024/HA12 4B063/QA12 4B063/QQ43 4B063/QR32 4B063/QR55 4B063/QS34 4B063/QX01		
代理人(译)	夏木森下		
优先权	60/100393 1998-09-15 US		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

要解决的问题：提供表征多核苷酸突变的方法。具有单链突出端的相同或基本相同的多核苷酸探针的阵列，其构成参考多核苷酸的片段和靶多核苷酸的片段的n-聚物的完整或基本完整的集合。分别连接并比较其杂交模式。该方法也可以用于检测多态性的存在。还可以确定未知序列的两个或更多个多核苷酸是否相同。另外，可以基于末端序列列出和区分多核苷酸的片段。这些方法可用于医学，药物遗传学，生物化学和法医科学。[选择图]无

標的の複雑性、プローブ長、および連結方法の関数としての塩基呼び出し正確度

		540bp CFTR	1.2kb HIV	2.5kbp Mito	5.4kbp φX174
酵素連結	8マー(%)	99.0(0.2)	98.2(0.1)	88.9(1.3)	73.9(4.1)
	9マー(%)	99.8(0.0)	99.9(0.0)	94.5(0.3)	85.2(0.5)
化学連結	8マー(%)	99.0(0.4)	96.1(0.5)	87.8(1.5)	69.0(2.8)
	9マー(%)	100.0(0.0)	99.5(0.0)	93.4(0.)	89.0(0.4)