

(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2005-58238

(P2005-58238A)

(43) 公開日 平成17年3月10日(2005.3.10)

| (51) Int. Cl. ⁷ | F I | テーマコード (参考) |
|----------------------------|-------------------------------------|-------------|
| C 1 2 Q 1/68 | C 1 2 Q 1/68 Z N A A | 2 G O 5 4 |
| C 1 2 M 1/00 | C 1 2 M 1/00 A | 4 B O 2 4 |
| C 1 2 N 15/09 | G O 1 N 21/78 C | 4 B O 2 9 |
| G O 1 N 21/78 | G O 1 N 33/53 M | 4 B O 6 3 |
| G O 1 N 33/53 | G O 1 N 37/00 1 O 2 | |
| | 審査請求 有 請求項の数 25 O L (全 39 頁) 最終頁に続く | |

| | | | |
|--------------|-------------------------------------|----------|---|
| (21) 出願番号 | 特願2004-293140 (P2004-293140) | (71) 出願人 | 399125757 |
| (22) 出願日 | 平成16年10月5日 (2004.10.5) | | アフィメトリックス インコーポレイテッド |
| (62) 分割の表示 | 特願2003-362608 (P2003-362608) の分割 | | アメリカ合衆国 カリフォルニア州 95 051 サンタ クララ セントラル エ クスプレスウェイ 3380 |
| 原出願日 | 平成6年6月24日 (1994.6.24) | (74) 代理人 | 100078282 |
| (31) 優先権主張番号 | 08/082, 937 | | 弁理士 山本 秀策 |
| (32) 優先日 | 平成5年6月25日 (1993.6.25) | (74) 代理人 | 100062409 |
| (33) 優先権主張国 | 米国 (US) | | 弁理士 安村 高明 |
| | | (74) 代理人 | 100113413 |
| | | | 弁理士 森下 夏樹 |
| | | | 最終頁に続く |

(54) 【発明の名称】 核酸配列のハイブリダイゼーションと配列決定

(57) 【要約】

【課題】 部分的にミスマッチの配列のハイブリダイゼーションのために標的核酸にハイブリダイズしたプローブのアレイを分析することにおける実質的な困難性の解決。

【解決手段】 標的よりも少数の塩基をもつ多数の核酸プローブにより標的核酸を配列決定する方法であって：そのプローブをその標的と接触させ；その標的に特異的にハイブリダイズする第1プローブを同定し；その第1プローブのA, C, T, U、及びG伸長物の少なくとも2を含んで成る伸長プローブの第1セットを選び；そしてその伸長プローブの第1セットの他のものよりも強くその標的に特異的にハイブリダイズするその伸長プローブの第1セットの中の1を同定し、それによりその伸長プローブの中の1がその標的核酸の中の1の塩基を同定する、段階を含んで成る方法

【選択図】 なし

【特許請求の範囲】

【請求項 1】

標的よりも少数の塩基をもつ多数の核酸プローブにより標的核酸を配列決定する方法であって、以下の段階：

その標的とそのプローブを接触させ；

その標的への高いアフィニティーをもつプローブを同定し；

その高いアフィニティープローブから多数の候補標的配列を同定し；そして

その候補標的配列のそれぞれについてのミスマッチ・プローブの標的アフィニティーを分析し、ここで、そのミスマッチ・プローブが少なくとも 1 塩基ミスマッチを除き候補標的配列の一部と相補的である、を含んで成る方法。

10

【請求項 2】

候補標的配列が、その標的に等しい塩基数をもつ核酸配列の全てを含み、そしてその高アフィニティー・プローブの少なくとも 1 に相補的な部分をもつ、請求項 1 に記載の方法。

【請求項 3】

予想されたアフィニティー・パターンと、その候補標的配列の各々についてのアフィニティー・パターンを比較する段階をさらに含んで成る、請求項 1 に記載の方法。

【請求項 4】

予想されたアフィニティー・パターンが、塩基ミスマッチがミスマッチ・プローブの内部にあるときよりもそのミスマッチ・プローブの末端にある場合により高いアフィニティーを示す、請求項 3 に記載の方法。

20

【請求項 5】

予想されたアフィニティー・パターンに最も近いアフィニティー・パターンをもつ候補標的配列を選択する段階をさらに含んで成る、請求項 3 に記載の方法。

【請求項 6】

標的よりも少数の塩基をもつ多数の核酸プローブにより標的核酸配列を配列決定する方法であって、以下の段階：

その標的とそのプローブを接触させ；

その標的アフィニティーに基づきそのプローブの各々について確率を決定し、各々の確率が、対応のプローブがその標的の一部に相補的である見込みを特定し；そして

その確率に基づいてその標的を同定するプローブのセットを選択する、

を含んで成る方法。

30

【請求項 7】

2 以上の核酸分子コレクションの、核酸配列を比較する方法であって、

(a) 標的核酸を含み、固体表面に結合された複数の標的エレメントを提供する工程と、

(b) その標的エレメントを下記と接触させる工程と、

(i) 標識され標的核酸配列と実質的に相補的な配列を含む核酸の第 1 のコレクション、および

(i i) 標的核酸配列と相補的な配列を有する少なくとも第 2 の標識された核酸

ここで、第 1 と第 2 の標識は互いに識別可能である

(c) 第 1 と第 2 の標識された相補的核酸の標的核酸に対する結合を検出する工程と

を含むことを特徴とする方法。

40

【請求項 8】

第 1 の細胞タイプにおける複数の遺伝子の各々の差次的な発現を、第 2 の細胞タイプにおける同一の遺伝子の発現に対して検出する方法であって、

これらの 2 つの細胞タイプからの標識された核酸の混合物を、これらの 2 つの細胞タイプ由来の複数の既知の遺伝子を表すポリヌクレオチドのアレイに、このアレイ中の相補的配列のポリヌクレオチドにハイブリダイゼーションさせる条件下で、加える工程と、

一方の細胞タイプ由来の核酸と主としてハイブリダイズするこのアレイ中のポリヌクレオチドが、明確な蛍光エミッション色を与え、他の細胞タイプ由来の核酸とハイブリダイズするこのアレイ中のポリヌクレオチドが、異なる蛍光エミッション色を与えるような蛍

50

光励起条件下で、このアレイを蛍光で検査する工程とを含むことを特徴とする方法。

【請求項 9】

第 1 の細胞タイプにおける複数の遺伝子の各々の差次的な発現を、第 2 の細胞タイプにおける同一の遺伝子の発現に対して検出する方法であって、

この 2 つの細胞タイプから単離された、標識された mRNA あるいは mRNA 生成物をつくる工程と、

少なくともこの 2 つの細胞タイプ由来の複数の既知の遺伝子を表すポリヌクレオチドの高密度アレイのこの 2 つの細胞タイプからの標識された mRNA あるいは mRNA 生成物の混合物を、このアレイ中の相補的配列のポリヌクレオチドにハイブリダイゼーションさせる条件下で、加える工程と、

一方の細胞タイプ由来の核酸と主としてハイブリダイズするこのアレイ中のポリヌクレオチドが、明確な蛍光エミッション色を与え、他の細胞タイプ由来の核酸とハイブリダイズするこのアレイ中のポリヌクレオチドが、異なる蛍光エミッション色を与えるような蛍光励起条件下で、このアレイを蛍光で検査する工程とを含むことを特徴とする方法。

10

【請求項 10】

2 以上の核酸分子コレクション中の核酸配列を検出する方法であって、

(a) 同定可能な核酸配列を含み固体表面に結合されたポリヌクレオチドのアレイを提供する工程と、

(b) このポリヌクレオチドアレイを下記と接触させる工程と、

(i) このアレイの核酸と実質的に相補的な配列を有する標識された核酸の第 1 のコレクション、および

(i i) このアレイの核酸と実質的に相補的な配列を有する標識された核酸の少なくとも第 2 のコレクション

ここで、第 1 と第 2 の標識は互いに識別可能である

(c) 第 1 と第 2 の標識された相補的核酸のアレイの核酸に対する結合を検出する工程とを含むことを特徴とする方法。

20

【請求項 11】

標識された核酸の 2 以上のコレクション中の核酸配列のコピー数を比較する方法であって、

核酸を含み固体表面に結合された複数のプローブを提供する工程と、

このプローブを下記と接触させる工程と、

第 1 の標識を有する標識された核酸の第 1 のコレクション、および第 2 の標識を有する標識された核酸の少なくとも第 2 のコレクションここで、第 1 と第 2 の標識は互いに識別可能であるプローブに相補的な第 1 の標識をされた核酸および第 2 の標識をされた核酸のそれぞれの結合量を検出し、これによりこれらの核酸のコピー数を比較する工程と

を含むことを特徴とする方法。

30

【請求項 12】

2 つの細胞タイプからの標識された核酸が第 1 及び第 2 の蛍光レポーターで標識されていることを特徴とする請求項 11 に記載の方法。

40

【請求項 13】

核酸の試験サンプルがある個体からのものであるかどうかを決定する方法であって、

(a) その個体からの核酸の参照サンプルを提供する工程と、

(b) 既知の部位に複数の異なる核酸を有する支持体を提供する工程と、

(c) 参照サンプルを支持体に適用して参照ハイブリダイゼーションパターンを得る工程と、

(d) 試験サンプルを支持体に適用し試験ハイブリダイゼーションパターンを得る工程と、

(e) 参照ハイブリダイゼーションパターンと試験ハイブリダイゼーションパターンとを比較し、この試験サンプルがその個体からのものであるかどうかを決定する工程とを含む

50

ことを特徴とする方法。

【請求項 14】

異なる個人間の遺伝的バリエーションを検出する方法であって、

- (a) 各個人のサンプルであってその個人の核酸を含むサンプルを提供する工程と、
- (b) 既知の部位に複数のプローブを有する支持体を提供する工程と、
- (c) 各サンプルを支持体に適用してハイブリダイゼーションパターンを得る工程と、
- (d) 得られたハイブリダイゼーションパターンを比較して遺伝的バリエーションを検出する工程と

を有することを特徴とする方法。

【請求項 15】

異なる組織を診断する方法であって、

- (a) 異なる組織のサンプルであってその異なる組織の転写産物を含むサンプルを提供する工程と、
- (b) 既知の部位に複数の異なるプローブを有する支持体を提供する工程と、
- (c) サンプルを支持体に適用し発現パターンを得る工程と、
- (d) この発現パターンを、複数の既知の組織の異なる発現パターンを有する参照データベースと比較し、これにより異なる組織を診断する工程と

を含むことを特徴とする方法。

【請求項 16】

組織の細胞構成物を決定する方法であって、

- (a) 組織のサンプルであってその組織の核酸を含むサンプルを提供する工程と、
- (b) 既知の部位に複数のプローブを有する支持体を提供する工程と、
- (c) サンプルをこの支持体に適用し発現パターンを得る工程と、
- (d) この発現パターンを、複数の既知の細胞の発現パターンを有する参照データベースと比較しこれによりその組織の細胞構成物を決定する工程とを含むことを特徴とする方法。

【請求項 17】

サンプルを分析する方法であって下記の工程を含むことを特徴とする方法

1 cm² あたり 100 より多いオリゴヌクレオチドを含む、1 以上のオリゴヌクレオチド高密度アレイを、1 以上の固体サポート上に載せる或いは形成する工程であって、このオリゴヌクレオチドはこの固体サポート上に形成されるか或いは前もって形成されてからこのサポートの既知の場所に載せられる、

メッセンジャー RNA サンプルを少なくとも 2 つの細胞集団から抽出し、このメッセンジャー RNA あるいはこの少なくとも 2 つの細胞集団からのメッセンジャー RNA サンプルの生成物を標識する工程と、

この抽出及び標識工程の生成物をこの 1 以上の高密度アレイにさらす工程と、

この生成物がこの高密度アレイのどこでハイブリダイズしたかを検出する工程と、

この検出工程に基づいて、この少なくとも 2 つの細胞集団におけるこのメッセンジャー RNA の発現レベルを決定する工程。

【請求項 18】

第 1 の細胞タイプにおける複数の遺伝子の各々の差次的な発現を、第 2 の細胞タイプにおける同一の遺伝子の発現に対して検出する方法であって、

これらの 2 つのタイプの細胞からの標識された核酸の混合物を、これらの 2 つのタイプの細胞由来の複数の既知の遺伝子を表すポリヌクレオチドのアレイに、アレイの相補的配列のポリヌクレオチドにハイブリダイゼーションさせる条件下で、加える工程と、

一方の細胞タイプ由来の標識された核酸とハイブリダイズするアレイのポリヌクレオチドが、明確な蛍光エミッション色を与え、他の細胞タイプ由来の標識された核酸とハイブリダイズするアレイのポリヌクレオチドが、異なる蛍光エミッション色を与えるような蛍光励起条件下で、このアレイを蛍光で検査する工程とを含むことを特徴とする方法。

【請求項 19】

10

20

30

40

50

複数のサンプルをマーキングする方法であって、

(a) それぞれ固有の配列でコードされる異なるポリマーからなる

複数のマーカを提供する工程と、

(b) 異なるマーカを各サンプルに包含させる工程と

を有することを特徴とする方法。

【請求項 20】

試薬基質 (reagent matrix) を調製する方法であって、

(a) 複数の第 1 の標的エレメントを、少なくとも 1 cm^2 あたり 103 個の異なる第 1 の標的エレメントが支持体上に存在するように、支持体の特定の部位に提供する工程、

(b) 少なくとも 1 つの第 1 の標的エレメントと特異的に結合可能であり、試薬にリンクされている第 2 の標的エレメントを、提供する工程と、

(c) 特異的結合可能な条件下で、この第 2 の標的エレメントを支持体に接触させ、これにより試薬基質を調製する工程とを含むことを特徴とする方法。

【請求項 21】

固体支持体に付けられた複数の位置的識別可能な配列特異的試薬を含む組成物であって、

この試薬が、少なくとも 3 つのサブユニットを有する前もって選ばれたマルチサブユニット

長さの所定のサブユニット配列と特異的に結合でき、またこの試薬が、この前もって選ばれた長さの可能な配列のすべてを実質的に表すことを特徴とする組成物。

【請求項 22】

個人における突然変異を検出する方法であって、

(a) 個人の核酸のサンプルを提供する工程と、

(b) 少なくともいくつかのオリゴヌクレオチドは突然変異配列を有し少なくとも他のいくつかのオリゴヌクレオチドは野生型配列を有する複数のオリゴヌクレオチドを既知の部位に有する支持体を提供する工程と、

(c) 各サンプルを支持体に適用してハイブリダイゼーションパターンを得る工程と、

(d) このハイブリダイゼーションパターンを分析してその個人における突然変異を検出する工程とを含むことを特徴とする方法。

【請求項 23】

複数の配列をお互いに対してマッピングする方法であって

(a) 複数の配列特異的プローブが位置的に付着された支持体を調製する工程と、

(b) それぞれの配列を支持体にさらし、これによりこの配列特異的プローブと配列との相互作用のパターンを決定する工程と、

(c) この配列特異的プローブの相互作用のこの配列における相対的位置を決定し、この配列のオーバーラップと順番を決定する工程と

を含むことを特徴とする方法。

【請求項 24】

複数のポリヌクレオチド配列をお互いに対してマッピングする方法であって

(a) 既知の位置に複数の配列特異的プローブが、少なくとも 1 cm^2 あたり約 100 プローブの割合で、付着された支持体を調製する工程と、

(b) それぞれのポリヌクレオチド配列を、ハイブリダイゼーション条件下で、支持体に付着されたプローブに接触させる工程と、

(c) この配列におけるハイブリダイゼーションの相対的位置を決定し、この複数のポリヌクレオチド配列のオーバーラップと配列の順番を決定する工程と

を含むことを特徴とする方法。

【請求項 25】

表面に、 1 cm^2 の表面積あたり少なくとも 10^3 の別個のポリヌクレオチド或いはポリペプチド・バイオポリマーのマイクロアレイをもつ支持体であって、各別個のバイオポリマーサンプルは (i) このアレイの別々の限定された位置に配置され、(ii) 少なくとも

50 サブユニットの長さを有し、(iii) 標識された標的サンプルとハイブリダイズされたときに検出可能な量であることを特徴とする支持体。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、核酸分析、検出、及び配列決定の分野に関する。より特に、1の態様においては、本発明は、核酸のアレイ (arrays) の合成、核酸のハイブリダイジング、一本鎖プローブと標的核酸から成る二本鎖核酸内のミスマッチの検出、及びDNA又はRNAあるいは他のポリマーの配列の決定のための改良技術を提供する。

【背景技術】

【0002】

(政府の権利)

本明細書中に記載する発明は、エネルギー省とAffymaxとの間の契約DE-FG03-92ER 81275号(認可第21012-92-II)の中で又はその下で；並びにNIH 契約第1R01HG 00813-01号の中で又はその下で生じた。

【0003】

(発明の背景)

多くの分野において核酸の配列を決定することは重要である。なぜなら、例えば核酸は酵素、構造タンパク質、及び生物学的機能の他のエフェクターをコードしているからである。ポリペプチドをコードする核酸のセグメントに加えて、遺伝子発現の制御及び調節に関係する多くの核酸配列が存在する。

【0004】

ヒト：ゲノム計画は、核酸配列決定技術を使用した計画の1例である。この計画は、ヒト生物のゲノムの完全な配列を決定することを目的としている。このような配列は、各々の特定の個人の配列とは必ずしも一致しないであろうけれども、その生物一般に関する有意義な情報と特定の個人からのゲノム・セグメント内に含まれる特定配列を提供するであろう。このヒト・ゲノム計画は、さらに詳細な研究に有用なマッピング情報をも提供するであろう。

【0005】

配列決定計画、例えばヒト・ゲノム計画の要求におけるよりも明らかな、高迅速で、正確で、そして安価が配列決定技術についての必要性は、どこにもない。ヒト・ゲノムの配列決定を完成するためには、約 3×10^9 、又は30億の、塩基対の決定が必要であろう。

【0006】

配列決定のために今日典型的に使用される手順は、Sanger et al., Proc. Natl. Acad. Sci. USA (1977) 74: 5463-5467、とMaxam et al., Methods in Enzymology (1980) 65: 499-559中に記載された方法を含む。このサンガー法は、鎖延長停止ジデオキシ・ヌクレオチドによる酵素的伸長を使用する。マキサム・ギルバート法は、塩基特異的解裂反応を示す化学反応を使用する。両方法は、多数の複雑な操作、例えば同種DNA断片の単離、サンプルの骨の折れる、かつ、のろい調製、分離用ゲルの調製、そのゲルへのサンプルの適用、そのゲル上でのサンプルの電気泳動、最終ゲルの調製、及びその手順の結果の分析を必要とする。

【0007】

他の技術も核酸の配列決定のために提案されてきた。全目的をもって引用により本明細書中に取り込むPCT特許公開第92/10588号は、標識された、標的核酸の配列が支持体上の核酸プローブのアレイへのハイブリダイゼーションにより決定されるような1の改良技術を記載している。各プローブは、その支持体上に位置的に区別できる場所に位置する。この標識された標的がその支持体に晒されるとき、それは、相補的ヌクレオチド配列を含む位置において結合する。その結合位置におけるプローブの配列の知識を通じて、ある者は、標的核酸のヌクレオチド配列を決定することができる。この技術は、ひじょうに大きなアレイの核酸プローブが使用されるとき、特に効率的である。このようなアレ

10

20

30

40

50

イは、Pirrung et al. に付与された米国特許第 5,143,854 号中に記載された技術に従って形成されることができる。また、米国出願逐次番号第 07/805,727 号をも参照のこと。両方を、全目的をもって引用により本明細書中に取り込む。

【0008】

これらの核酸プローブが標的よりも短い長さをもつとき、ある者は、より短いプローブからのアフィニティー・データに基づき、より大きな標的の配列を決定するための再構築技術を使用することができる。Drmanac et al. への米国特許第 5,202,231 号、と Southern への PCT 特許公開第 89/10977 号を参照のこと。この困難性を克服するための 1 つの技術は、ハイブリダイゼーション又は SBH による配列決定といわれてきた。例えば、12-mer の標的 DNA 5'-AGCCTAGCTGAA が全部で 8 ヌクレオチドのプローブのアレイと混合されると仮定する。この標的が正確に相補的なヌクレオチド配列をもつようなプローブだけと結合する場合、65,536 のオクタマーの中のたった 5 つのプローブ (3'-TCGGATCG, CGGATCGA, GGATCGAC, GATCGACT、と ATCGACT) がこの標的にハイブリダイズするであろう。これらのハイブリダイジング・プローブからの重複配列の整列は、その元の 12-mer の標的の相補物を再構築する：

```

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

```

【発明の開示】

【発明が解決しようとする課題】

【0009】

多くの樂觀に会ったにもかかわらず、従来技術は、特定の限定とも会った。例えば、従事者は、他の困難性の中で、部分的にミスマッチの配列のハイブリダイゼーションのために標的核酸にハイブリダイズしたプローブのアレイを分析することにおける実質的な困難性に遭遇した。本発明は、このようなアレイによる配列決定における有意な進歩を提供する。

【課題を解決するための手段】

【0010】

項目 1 . 標的よりも少数の塩基をもつ多数の核酸プローブにより標的核酸を配列決定する方法であって：

そのプローブをその標的と接触させ；

その標的に特異的にハイブリダイズする第 1 プローブを同定し；

その第 1 プローブの A, C, T, U、及び G 伸長物の少なくとも 2 を含んで成る伸長プローブの第 1 セットを選び；そしてその伸長プローブの第 1 セットの他のものよりも強くその標的に特異的にハイブリダイズするその伸長プローブの第 1 セットの中の 1 を同定し、それによりその伸長プローブの中の 1 がその標的核酸の中の 1 の塩基を同定する、段階を含んで成る方法。

【0011】

項目 2 . 核酸プローブの実質的に全てが n ヌクレオチドを含んで成り、そして伸長プローブが第 1 プローブの n - 1 ヌクレオチドを含んで成る、項目 1 に記載の方法。

【0012】

項目 3 . 伸長プローブの第 1 セットから反対の方向において伸長する A, C, T, U、及び G 伸長プローブの第 2 セットを選び；そして

その伸長プローブの第 2 セットの他のものよりも強く標的に特異的にハイブリダイズす

るその伸長プローブの第2セットの中の1を同定し、それにより、その伸長プローブの第2セットの中の1が標的核酸内の第2の塩基を同定する、
段階をさらに含んで成る、項目1に記載の方法。

【0013】

項目4．伸長プローブのセットを選ぶ段階を繰り返し、そして5回以上伸長プローブを同定する、

段階をさらに含んで成る、項目1に記載の方法。

【0014】

項目5．同定の段階が、1塩基ミスマッチ・プローブであって、伸長プローブの第1セットのA、C、T、U、及びGモノ置換の中の少なくとも2を含んで成るものを同定し；
その1塩基ミスマッチ・プローブのハイブリダイゼーション・アフィニティー・データを記録し；そして

そのハイブリダイゼーション・アフィニティー・データが、その1塩基ミスマッチ・プローブの予想ハイブリダイゼーションアフィニティー・データを確信させるときその第1プローブの正しい伸長として伸長プローブの第1セットの中の1を選ぶ、
段階をさらに含んで成る、項目1に記載の方法。

項目6．予想されるハイブリダイゼーション・データが：伸長プローブの末端においてミスマッチをもつプローブ/標的複合体についてのより高い結合アフィニティー；及び
その複合体の内部においてミスマッチをもつプローブ/標的複合体についてのより低い結合アフィニティー、
を含んで成る、項目5に記載の方法。

【0015】

項目7．ハイブリダイゼーション・データが伸長プローブの中の1についてのハイブリダイゼーション値に対して正規化され；そして

同定の段階が、その伸長プローブの中の1つの正規化された値よりも高い正規化されたハイブリダイゼーション値をもたない末端1塩基ミスマッチ・プローブをもつ伸長プローブの中の1を選ぶことを含んで成る、
項目6に記載の方法。

【0016】

項目8．同定の段階が、他の伸長プローブよりも高い、標的に対する結合アフィニティーを示す伸長プローブのセットの中の1を選ぶ段階を含んで成る、項目1に記載の方法。

【0017】

項目9．同定の段階が、適当にプログラムされたコンピューター内で行われる、項目1に記載の方法。

【0018】

項目10．標的核酸のヌクレオチド配列が第1核酸の配列と同一であるかどうかを決定する方法であって：

その標的核酸を多数の核酸プローブに接触させ；

その配列に1塩基ミスマッチを除き同一であるプローブへのその標的のアフィニティーを測定し；そして

その標的の核酸配列がその第1核酸と同一であるかを、1塩基ミスマッチを除き同一であるプローブへのその標的のアフィニティーが所定のパターンに従う場合に、決定する、
ことを含んで成る方法。

【0019】

項目11．所定のパターンが、配列の完全相補物のアフィニティーに正規化された1塩基ミスマッチのプローブのアフィニティーを含んで成る、項目10に記載の方法。

【0020】

項目12．1塩基ミスマッチ・プローブのアフィニティーが、アフィニティー対ミスマッチ位置としてプロットされ、そして配列の完全相補物のアフィニティーに正規化される、
項目11に記載の方法。

10

20

30

40

50

【0021】

項目13．標的の核酸配列が、1塩基ミスマッチに相補的なプローブへのその標的のアフィニティーが所定のパターンに従わない場合に、第1核酸と同一ではないことを決定する段階をさらに含んで成る、項目10に記載の方法。

【0022】

項目14．核酸のプローブ・アレイであって、標的核酸への正確な相補物、並びにその正確な相補物の1塩基ミスマッチを含んで成るように可能性のあるプローブの全てから選ばれているプローブ・アレイ。

【0023】

項目15．核酸プローブが約8と15塩基の間の長さをもつ、項目14に記載のライブラリー。 10

【0024】

項目16．ライブラリーが、単一支持体上にある、項目14に記載のライブラリー。

【0025】

項目17．ライブラリーがn-塩基以下のプローブを含んで成り、そしてそのライブラリーがn-塩基の可能性のあるプローブの全ての50%未満を含んで成る、項目14に記載のライブラリー。

【0026】

項目18．ライブラリーがn-塩基以下のプローブを含んで成り、そしてそのライブラリーがn-塩基の可能性のあるプローブの全ての10%未満を含んで成る、項目14に記載のライブラリー。 20

【0027】

項目19．コア核酸プローブであって、核酸標的に正確に相補的なもの、並びにそのコア・プローブの選択されたA、C、T、U、及びGの1塩基置換物、を含んで成る核酸プローブ・キット。

【0028】

項目20．コア・プローブ並びにそのコア・プローブのA、C、T、及びGのT塩基置換物から本質的に成る、項目19に記載の核酸プローブ・キット。

【0029】

項目21．標的サンプルが標的と同一であるか又はこれと異なるかどうかを決定するための指示書をさらに含んで成る、項目19に記載の核酸プローブ・キット。 30

【0030】

項目22．コア・プローブが8と15の間の塩基を含んで成る、項目19に記載の核酸プローブ・キット。

【0031】

項目23．プローブが、鎌状赤血球貧血、P-53突然変異、嚢胞性線維症突然変異、HLAクラス1遺伝子、及びHLAクラス2遺伝子から成る群から選ばれた遺伝子特徴について標的サンプルを評価するように選ばれる、項目19に記載の核酸プローブ・キット。

【0032】

項目24．プローブが、鎌状赤血球貧血について標的サンプルを評価するように選ばれる、項目19に記載の核酸プローブ・キット。 40

【0033】

項目25．標的核酸の配列決定方法であって、以下の段階：

プローブのアレイとその標的を接触させ；

少なくとも同定されたハイブリダイゼーション・アフィニティー・レベルをもつアレイ内で選択された高アフィニティー・プローブを同定し；

その高アフィニティー・プローブから誘導された多数の候補標的配列を同定し；

ミスマッチ・プローブであって、少なくとも1の塩基ミスマッチを除きその高アフィニティー・プローブと同一の塩基配列を含むものを同定し； 50

そのミスマッチ・プローブについてのアフィニティー・パターンと、その候補標的配列についての予想ミスマッチ・アフィニティー・パターンとを比較し；そして

その候補標的配列から標的配列を、選択された標的配列が最良のミスマッチ・パターンをもつように、選択する、
を含んで成る方法。

【0034】

項目26．予想されたミスマッチ・アフィニティー・パターンが；

伸長プローブの末端においてミスマッチをもつプローブ／標的複合体についてのより高い結合アフィニティー；及び

その複合体の内部においてミスマッチをもつプローブ／標的複合体についてのより低い結合アフィニティー、
を含んで成る、項目25に記載の方法。 10

項目27．アフィニティー・パターンを比較する段階が、以下の段階；

予想されたミスマッチ・アフィニティー・パターン内のそれぞれのミスマッチ位置について予想されるアフィニティー値を同定し；そして

その予想されたアフィニティー値からの同定された変動以下のアフィニティー値をもつその多数のミスマッチ位置を選ぶ、
を含んで成る、項目25に記載の方法。

【0035】

項目28．以下の段階； 20

より短いプローブのアレイと標的を接触させ；

そのプローブの各々についてその標的の部分との完全マッチの確率を決定し；

そのプローブの全て各々組合せについて、その完全マッチの確率の全てを決定し；そして

それについて完全マッチの確率の全てが最大化される、そのプローブの組合せからそのプローブの同定された組合せを選択し、それによりそのプローブの同定された組合せがその標的核酸の配列を同定する、
を含んで成る、標的核酸の配列決定方法。

【0036】

項目29．完全マッチの確率の決定段階が、以下の段階； 30

多数の既知の断片についてのハイブリダイゼーション・データを集め；

そのハイブリダイゼーション・データから完全マッチの既知の確率を予測するアルゴリズムを同定し；そして

そのアルゴリズムを使用してそのプローブについての完全マッチの確率を同定する、
を含んで成る、項目28に記載の方法。

【0037】

(発明の要約)

核酸(オリゴヌクレオチド)を合成し、ハイブリダイズし、分析し、そして配列決定するための改良された技術を本発明により提供する。

【0038】 40

本発明の1の態様に従って、標的オリゴヌクレオチドが多数の、より短い長さの固定化されたプローブに晒される。これらのプローブを、集合的に、“アレイ(array)”という。本法においては、ある者は、標的に高いアフィニティーをもつコア・プローブを最初に固定し、そして次にこのコア・プローブと比較したとき1塩基のミスマッチをもつ全てのプローブの結合特性を評価することにより、標的核酸がそのアレイ内のプローブに相補的であるかどうかを同定する。この1塩基ミスマッチ・プローブが特徴的な結合又はアフィニティー・パターンを示す場合、このときそのコア・プローブは、その標的核酸の少なくとも一部に正確に相補的である。

【0039】

本法は、コア・プローブの“左”と“右”伸長といわれることができるプローブの結合 50

アフィニティーを評価することによりそのアレイ内のいずれのプロープよりも大きな標的核酸を配列決定するために拡大されることができ、コアの正しい左と右伸長は、1塩基ミスマッチ・プロープの特定のハイブリダイゼーション・パターン及び/又は最も強い結合アフィニティーを示すようなものである。1塩基ミスマッチ・プロープの結合アフィニティー特定は、その3'又は5'末端上にミスマッチをもつプロープ/標的複合体が内部ミスマッチをもつプロープ/標的複合体よりも安定である特性パターンに従う。この方法は、次に、核酸標的の配列を提供するために、コア・プロープの追加の左と右伸長を決定するために繰り返される。

【0040】

いくつかの態様、例えば診断においては、標的は、特定の配列をもつと予想される、標的が予想配列をもつかどうかを決定するために、相補的プロープ並びに1塩基ミスマッチ・プロープの全ての全サブセット又はその中のいくつかのサブセットを含むプロープのアレイが合成される。このようなプロープへのこの標的のハイブリダイゼーション・パターンの分析を通じて、その標的が予想配列をもつかどうか決定されることができ、そしてそうでない場合その標的の配列が場合により決定されることができ、

10

【0041】

核酸標的の分析のためのキットも、本発明により提供される。1の態様に従えば、キットは、核酸プロープのアレイを含む。このプロープは、標的核酸に対する完全な相補物を含むことができる。これらのプロープは、その完全に相補的なプロープの1塩基置換物であるプロープをも含む。このキットは、完全相補物の1以上のA, C, T, G、及び/又はU置換を含むことができる。このようなキットは、特定の遺伝子配列についての標的の分析、を含むさまざまな用途、例えば遺伝子疾患のための分析におけるものをもつであろう。

20

【発明の効果】

【0042】

本発明によって、部分的にミスマッチの配列のハイブリダイゼーションのために標的核酸にハイブリダイズしたプロープのアレイを分析することにおける実質的な困難性を解決するための、アレイによる配列決定における有意な進歩を提供した。

【発明を実施するための最良の形態】

【0043】

本明細書中における本発明の性質と利点のさらなる理解は、本明細書の残りの部分と添付図面を参照することにより理解されるであろう。

30

【0044】

(好ましい態様の説明)

目次

- A. 合成
- B. ハイブリダイゼーション
- C. ミスマッチ分析
- D. 適用
- E. 結論

40

(定義)

プロープ - 典型的には固体表面上で形成された公知の組成又はモノマー配列の分子であって、標的分子に晒され又は晒されてもよく、そしてそのプロープがその標的にハイブリダイズされたかどうかを決定するために検査され又は検査されてもよいもの。“コア (core)”プロープは、標的に強いアフィニティーを示すプロープである。“伸長 (extension)”プロープは、コア・プロープ配列の全部又は一部に加えそのコア・プロープ配列の1以上の可能性のある伸長を含むプロープである。本明細書は、プロープの3'-末端における伸長として“左”伸長と、そしてプロープの5'-末端における伸長として“右”伸長という。但し、反対の表記法も明らかに採用されることができであろう。

50

【0045】

標的 - それについて、その組成又はモノマー配列を研究することが望ましい、典型的には、未知の組成又はモノマー配列をもつ分子。標的は、より大きな分子の一部、例えば、より長い核酸内の数塩基であることができる。

【0046】

n - 塩基ミスマッチ - n が 1 以上である、コアプローブ内の対応モノマーとは異なる、n モノマーをその中にもつプローブ。

【0047】

A, T, C, G, U - は、それぞれ、以下のヌクレオチド、アデニン、チミン、シトシン、グアニン、とウリジンについての略号である。

【0048】

ライブラリー - 1 以上の支持体内にしばしば形成される、所定ヌクレオチド配列の核酸プローブのコレクションであって、標的核酸のハイブリダイゼーション試験において使用されるもの。

A. 合成

光指定オリゴヌクレオチド合成のための方法を図 1 中に示す。このような戦略は、本発明者の譲受人に譲渡され、そして全目的をもって引用により本明細書中に取り込む米国特許第 5, 143, 854 号中により詳細に記載されている。

【0049】

図 1 において説明する光指定合成法においては、光不安定性保護基又は基 (X) により誘導体化された表面 (2) を、フォトリトグラフ・マスク (photolithographic mask) M_1 を通して照射して、反応性ヒドロキシル (OH) 基を露出させる。(光不安定性の保護基により 5' - ヒドロキシルにおいて保護された) ホスホルアミジット活性化ヌクレオシドのシリーズの第 1 の (T - X) を次に表面全体に晒す。先行する照射の間に先に晒された部位においてのみカップリングが生じる。

【0050】

上記カップリング反応が完了した後、支持体を濯ぎ、そしてその表面を再び、新たな又は平行移動されたマスク M_2 を通して照射して、カップリングのための異なる基を露出させる。(光不安定性の保護基により 5' - ヒドロキシルにおいて再び保護された) 新たなホスホルアミジット活性化ヌクレオシド C - X を添加し、そして上記露出部位にカップリングさせる。上記工程を、光脱保護とカップリングのサイクルを通じて繰り返して、支持体上にオリゴヌクレオチド・プローブの所望のセットを作り出す。フォトリトグラフィが使用されるので、この工程は微細にされることができる。さらに、反応が光により場所的にアドレスされた部位においてのみ生じるので、各部位におけるプローブのヌクレオチド配列は正確に知られ、そして標的分子 (標的核酸又は、他の態様においては、タンパク質、例えばレセプタのいずれか) との、各部位におけるオリゴヌクレオチド・プローブの相互作用を評価することができる。5' - O - (- メチル - 6 - ニトロピペロニルオキシカルボニル) - N - アシル - 2' - デオキシヌクレオシド、又は MeNPOc - N - アシル - デオキシヌクレオシド、MeNPOc - dT, MeNPOc - dC^{i b u}, MeNPOc - dG^{P A C}、及び MeNPOc - dA^{P A C} を含む光保護デオキシヌクレオシドを本工程のために開発した。保護基化学は、PCT 特許公開第 92 / 10092 号及び 1990 年 12 月 6 日に出願された米国出願逐次番号第 07 / 624, 120 号、及び 1992 年 12 月 2 日に出願された第 07 / 971, 181 号であって、共に本発明の譲受人に譲渡され、そして全目的をもって引用により本明細書中に取り込まれたものの中かなり詳細に開示されている。

実施例

1. 保護基

塩基は 280 nm 領域内に強い * 遷移をもつので、光除去性保護基の脱保護波長は、不所望のヌクレオシド光化学を回避するために 280 nm よりも長い波長でなければならない。さらに、4 つのデオキシヌクレオシドの光脱保護速度は、光が、全照射合成部

10

20

30

40

50

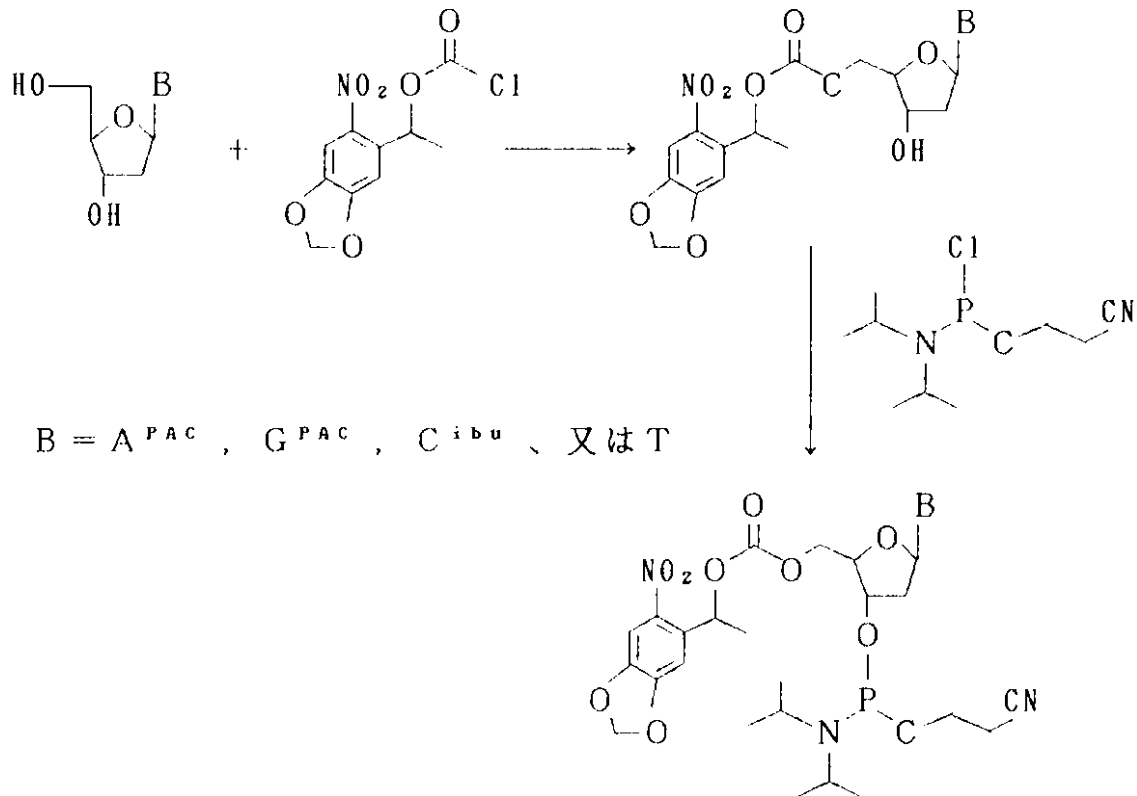
位内で、ヒドロキシル（又は他の官能基、例えばスルフヒドリル又はアミノ基）を均しく脱保護するであろうように、近似していなければならない。

【0051】

これらの基準に適合するために、1セットの、5'-O-(2-メチル-6-ニトロピペロニルオキシカルボニル)-N-アシル-2'-デオキシヌクレオシド (MeNPOc-N-アシル-デオキシヌクレオシド) が光指定合成のために開発され、そして保護されたヌクレオシドの光速度論的挙動が測定された。5'-O-(2-メチル-6-ニトロピペロニルオキシカルボニル)-N-アシル-2'-デオキシヌクレオシド・ホスホルアミジットを合成するための合成系路をスキーム I に示す。

【0052】

【数1】



スキーム I

第1段階において、N-アシル-2'-デオキシヌクレオシドを1-(2-ニトロ-4,5-メチレンジオキシフェニル)-エタン-1-クロロホルメートと反応させて、5'-MeNPOc-N-アシル-2'-デオキシヌクレオシドを作る。第2段階において、その3'-ヒドロキシルを、標準的な手順を用いて2-シアノエチル-N,N'-ジイソプロピルクロロホスホルアミジットと反応させて、5'-MeNPOc-N-アシル-2'-デオキシヌクレオシド-3'-O-ジイソプロピルクロロホスホルアミジットを作る。これらの試薬は、4においてアルゴン下で乾燥保存されるとき長時間安定であった。

【0053】

4つのデオキシヌクレオシド、MeNPOc-dT, MeNPOc-dC^{ibu}, MeNPOc-dG^{PAC}, とMeNPOc-dA^{PAC}のそれぞれの0.1 mM 溶液をジオキサン中で調製した。アリコート(200 mL)を、さまざまな時間にわたり狭い経路(2 mm)の石英キュベット内で14.5 mW/cm²の365 nm 光で照射した。4~5の時間点を、各塩基について採取し、そしてそれらの溶液を、0.1(v/v)% TFAを含む水中60(v/v)%の移動相により溶出して、280 nmにおけるHPLC装置とヌクレオジル5-C₈HPLCカラムにより出発材料の損失について分析した(MeNPOc-dTは水中70(v/v)%メタノールの移動相を必要とした。)。残った

10

20

30

40

50

MeNPoc - N - アシル - デオキシヌクレオシドのピーク面積を計算して、MeNPoc - dT, MeNPoc - dC^{i b u}, MeNPoc - dG^{P A C}、とMeNPoc - dA^{P A C}について、それぞれ28秒、31秒、27秒、と18秒の光分解半減期を得た。その後のリトグラフ実験において、4, 5分間の照射時間 ($9 * t_{1/2}^{MeNPoc - dC}$) は、99%以上のMeNPoc保護基の除去を導いた。

【0054】

光指定合成においては、全体の合成収率は、光脱保護の収率、光脱保護のコントラスト、と化学的カップリングの効率に依存する。光速度論的条件は、好ましくは、光脱保護の収率が99%を超えることを確保するように選ばれる。支持体の通暗い領域内の不所望の光分解は、その合成忠実度 (fidelity) に悪影響を及ぼすことができるが、高吸光度 (50DU) によるリトグラフ・マスクを使用することにより、そして光学表面の注意深いインデックス・マッチング (index matching) により最小化されることができる。ガラス支持体へのDMT - N - アシル - デオキシヌクレオシド・ホスホルアミジットの縮合効率は、95% ~ 99%のレンジ内で測定された。MeNPoc - N - アシル - デオキシヌクレオシド・ホスホルアミジットの縮合効率も、90%を上廻って測定された。但し、それらの効率は、合成間で変動することができ、そして監視されなければならない。

10

【0055】

2. カップリング効率の測定

光保護されたヌクレオシドのカップリング効率を調べるために、4つのMeNPoc - アミジットのそれぞれを最初に、(DMT化学物質を介して) 支持体にカップリングした。その支持体の領域を照射し、そしてMeNPoc - ホスホルアミジットを保護基なしで添加した。その支持体の新たな領域を次に照射し; 蛍光デオキシヌクレオシド・ホスホルアミジット (FAM - ホスホルアミジット Applied Biosystems) をカップリングして; そしてその支持体をシグナルについて走査した。蛍光標識されたホスホルアミジットが、新たに露出されたヒドロキシル基と先に未反応とヒドロキシル基の両方において反応する場合、その時、2つの部位の間の蛍光強度の比は、そのカップリング効率の尺度を提供する。この計測は、表面光分解の収率が不変に近いということを受け入れる。この又は近似の検定を使用した化学的カップリング収率は、変動性であるが、高く、80 ~ 95%の間のレンジにある。

20

30

【0056】

別個の検定において、化学的カップリング効率は、ヘキサエチレングリコール誘導体化支持体上で測定される。まず、グリコール・リンカーを脱トリチル化し、MeNPoc - デオキシヌクレオシド - O - シアノエチルホスホルアミジットをキャッピングなしでその樹脂にカップリングした。次に、DMT - デオキシヌクレオシド - シアノエチルホスホルアミジット (リポーター - アミジット) をその樹脂にカップリングした。このリポーター - アミジットは、上記第1段階からのいずれかの未反応のヒドロキシル基にカップリングする。このトリチル溶出液を集め、そして吸収分光光度計により定量した。溶出液を、そのデリバリー・ライン内に残った残存トリチルを測定するためにMeNPoc - ホスホルアミジット・カップリング直後にそのラインからも集めた。この検定において、カップリング効率は、100%の、リポーター - アミジットのカップリング効率を呈すると測定される。ヘキサエチレングリコール・リンカーへのMeNPoc - デオキシリボヌクレオシド - O - シアノエチルホスホルアミジットのカップリング効率と16のジヌクレオチドの効率が測定され、そしてDMT - デオキシヌクレオシド・ホスホルアミジットとは区別できなかった。

40

【0057】

3. オリゴヌクレオチド・プローブの場所に指定された合成

オリゴヌクレオチド・プローブの合成を始めるために、支持体を調製し、そしてMeNPoc - dC^{i b u} - 3' - O - ホスホルアミジットを合成リンカーを通じて合成サポートに付着させた。このサポートに付着させた。このサポートの領域を、リトグラフ・マス

50

クの $800 \times 1280 \mu\text{m}$ の孔を通しての照射により合成のために活性化した。7つの追加のホスホルアミジット合成サイクルを（対応のDMT保護されたデオキシヌクレオシドをもって）行い、S-3'-CGCATCCGを作り出した。室温において4時間濃NH₄OHにより、ホスフェートと環外アミン保護基の除去の後、支持体を、水ジャケット付熱制榔されたハイブリダイゼーション・チャンバー内に載せた、この支持体を、以下に述べるミスマッチ実験において使用した。

【0058】

B. ハイブリダイゼーション

オリゴヌクレオチド・アレイは、ハイブリダイゼーション研究を含む、多種多様の用途において使用されることができる。ハイブリダイゼーション研究においては、そのアレイは、図1中に示すように、着目のレセプタ（R）に晒されることができる。このレセプタは、適当な標識（*）、例えばフルオレセインにより標識されることが出来る。このレセプタが結合した支持体上の位置が測定され、そして、その位置におけるオリゴヌクレオチド・プローブの配列の知識を通じて、ある者は次に、そのレセプタがオリゴヌクレオチドである場合、そのレセプタ配列を決定することができる。

10

【0059】

ハイブリダイゼーションにより配列決定（SBH）は、各部位におけるプローブの同一性が知られているアレイを形成するために、多くのプローブを表面に付着させることにより最も効率良く行われる。標識された標DNA又はRNAは次にそのアレイにハイブリダイズされ、そしてそのハイブリダイゼーション・パターンが、そのアレイ内の全相補的プローブの同一性を決定するために調べられる。ミスマッチ・プローブ/標的複合体が重要でないことを教示する従来の技術とは反対に、本発明は、ミスマッチ・プローブ/標的複合体のハイブリダイゼーション・シグナルが、そのアレイ上の完全にマッチしたプローブ/標的複合体の同一性を同定し又は確認するような分析方法を提供する。

20

【0060】

オリゴヌクレオチドのアレイは、光指定合成技術を用いたハイブリダイゼーション研究のために効率的に生成される。以下に討議するように、全テトラヌクレオチドのアレイを16サイクルにおいて作り出した。これは完成するのにたった4時間しか必要としなかった。組合せ戦略が使用されるので、アレイ上の異なる化合物の数は合成の間に指数関数的に増加する。一方、化学的カップリング・サイクルの数は、直線的に増加する。例えば、4⁸（65,536）オリゴヌクレオチドの完全セットにこの合成を拡張することは、16の追加のサイクルが必要であるためにその合成にほんの4時間（又はこれ未満）を追加する。さらに、組合せ合成戦略を、いずれかの所望のプローブ組成を作り出すために行われることができる。例えば、ドデカマー（4¹²）の完全セットは、48以下の光分解とカップリング・サイクルにおいて作り出されるので（bⁿ化合物は多くともb×nサイクルを必要とする。）、（より短いオリゴヌクレオチドのいずれかのサブセットを含むドデカマーのいずれかのサブセットは、48以下の化学的カップリング段階において構築されることができる。アレイ内の化合物の数は、合成部位の密度と全体のアレイ・サイズによってのみ制限される。本発明は、片面上25ミクロンの正方形部位内で合成されたプローブをもつアレイにより行われている。この解像度において、65,536オクタヌクレオチドの完全セットは、たった0.64cm²の寸法のアレイ内に置かれることができる。1,048,576ドデカヌクレオチドのセットは、この個々のプローブ部位サイズにおいてたった2.56cm²のアレイを必要とする。

30

40

【0061】

ゲノム配列決定プロジェクトの成功は、効率的なDNA配列決定技術に依存する。最近の方法は、複雑な手順に高く頼り、そして実質的な手作業を必要とする。SBHは、最近の実施における手作業の多くを自動化するための潜在能力を提供する。光指定合成は、SBHのためだけでなく多くの他の用途のためにも微細化されたアレイの大規模生産のための効率的な手段を提供する。

【0062】

50

オリゴヌクレオチド・アレイは、一次的な配列決定用途のために使用されることができ
 るけれども、多くの診断方法は、標的核酸配列内のほんの数個のヌクレオチドの位置の分
 析を含む。1塩基の変更は、プローブ・アレイ上の標的のハイブリダイゼーション・パタ
 ーン内に多数の変化を引き起こすので、本発明に係るオリゴヌクレオチド・アレイと方法
 は、ある者が、先に明白されたDNA配列の正確さをチェックし、又は標的核酸内での特
 定の特異的配列内の変化又は突然変異について走査することを可能にする。例えば、後者
 は、遺伝子、疾患、晶質管理、及び法延分析のために重要である。オクタヌクレオチド・
 プローブセットにより、標的核酸内の1塩基変化は、8つの完全ハイブリッドの損失によ
 り、そして8つの新たな完全ハイブリッドの生成により、検出されることができ。この
 1塩基変化は、アレイ上の変更されたミスマッチ・プローブ/標的複合体の形成を通じて
 検出されることもできる。おそらく、さらに驚ろくべきことに、複雑な核酸内のこのよう
 な1塩基対の変化は、アレイ上の標的の全ハイブリダイゼーション・パターンを劇的に変
 更する。本発明に従って、この全ハイブリダイゼーション・パターンにおけるこのような
 変化は、この分析を実際に単純化するために使用される。

【0063】

光指定オリゴヌクレオチド・アレイの高い情報含有量は、遺伝子診断テストにかなり有
 益である。100~1,000の異なる突然変異の配列比較は、1時間に1回の形式にお
 けるものの代わって、同時に検定されることができ。アレイは、多種多様な病因生物の
 迅速な同定についての遺伝子マーカーを含み、そして、RNA/RNA, RNA/DNA
 、タンパク質/RNA又はタンパク質/DNA、相互作用の配列特異性を研究するために
 、構築されることもできる。ある者は、アンチセンス、トリプル・ヘリックス、又は他の
 用途のために非ワトソン・クリック・オリゴヌクレオチド及び新規の合成ヌクレオチド・
 アナログを使用することができる。好適に保護されたRNAモノマーが、RNA合成のため
 に使用されることができ、そして多種多様な合成及び非天然核酸アナログが、実施者の
 動機付けに依存して、使用されることができ。例えば、引用により本明細書に取り込む
 、PCT特許公開第91/19813号、第92/05285号、及び第92/1484
 3号を参照のこと。さらに、オリゴヌクレオチド検定は、オリゴヌクレオチド複合体の形
 成及び安定性を支配する熱力学と速度論の規則を演繹するために使用されることができ
 る。

【0064】

実施例

1. 表面オリゴヌクレオチドへの標的のハイブリダイゼーション

先に討議した支持体結合オクタヌクレオチド・プローブを、15 における15分間の
 インキュベーションによりハイブリダイゼーション・チャンバー内で5' GCGTAGG
 C-フルオレセインの標的にハイブリダイズさせた。アレイ表面を次に蛍光外 (epifluorescence) 顕微鏡 (488nm アルゴン・イオン励起) に応答信号を送
 った (interrogated)。この走査の蛍光画像を図2に示す。この蛍光強度パ
 ターンは、プローブの合成を指定するために使用される800x1280µmストライプ
 にマッチする。さらに、シグナル強度は、(ガラス支持体の背景を4倍上廻って) 高く、
 そのプローブへの標的の特異的結合を立証する。

【0065】

標的・プローブ複合体の挙動を、そのハイブリダイゼーション溶液の温度を上昇させる
 ことにより調べた。各温度において10分間の平衡化の後、支持体をシグナルについて走
 査した。2本鎖は、研究下配列について予想される温度レンジ内でメルトした ($T_m = [2 \cdot (A + T) + 4 \cdot (G + C)]$ の規則から得られた $T_m = 28$)。アレイ内のプロ
 ーブは、標的DNAの再ハイブリダイゼーションにより立証されるように、標的・プロ
 ーブ複合体の温度変性に対して安定であった。

【0066】

2. 標的ハイブリダイゼーションの配列特異性

標的ハイブリダイゼーションの配列特異性を立証するために、2つの異なるプローブを

800 × 1280 μm のストライプ内で合成した。図 3 A は、2 つのプロープの位置を同定する。プロープ S - 3' - CGCATCCG をストライプ 1, 3 と 5 内で合成した。プロープ S - 3' - CGCTTCCG をストライプ 2, 4 と 6 内で合成した。図 3 B は、15 における支持体への 5' - GCGTAGGC - フルオレセイン標的のハイブリダイゼーションの結果を示している。これらのプロープはたった 1 の内部塩基が異なるけれども、上記標的は、(ストライプ 1, 3 と 5 内の背景を ~ 500 カウント上廻って) その相補的配列に特異的にハイブリダイズし、位置 2, 4 と 6 (~ 10 カウント) 内には検出可能なシグナルはほとんど又は全くない。図 3 C は、両配列への標的のハイブリダイゼーションの結果を示す。図 3 中のすべての位置内のシグナルは、図 3 B におけるシグナルの非存在が 1 塩基ミスマッチの不安定性にのみ依ることを説明している。標的には、等モル濃度において存在するけれども、図 3 B 中のストライプ 2, 4 と 6 内のシグナルの比は、領域 1, 3 と 5 内のシグナルよりも約 1.6 倍高い。この 2 本鎖は、領域 2, 4 と 6 を占める 2 本鎖よりもわずかに高い予想 T_m をもつ。これらの 2 本鎖は、15 分間で 45 までその温度を上昇させることにより解離され、そしてそのハイブリダイゼーションは、上記と逆の順番で繰り返され (図 3 D と 3 E)、これは、逆方向におけるハイブリダイゼーションの特異性を立証している。

10

【0067】

3. プロープ・マトリックスの組合せ合成とそれへの核酸標的のハイブリダイゼーション

光指定合成においては、製品の位置と組成は、照射と化学的カップリング試薬の順番に依存する (完全な説明については、Fodor et al., Science (1991) 251: 767-773 を参照のこと。) 図 4 において説明するように、256 テトラヌクレオチドの合成について考える。マスク 1 は、合成の第 1 ラウンドにおける 4 つのヌクレオチドの第 1 とのカップリングのために支持体表面の 1/4 を活性化する。サイクル 2 において、マスク 2 は第 2 ヌクレオチドとのカップリングのためにその支持体の別の 1/4 を活性化する。この工程がモノヌクレオチドの 4 つの領域を作り上げるために続けられる。ラウンド 2 のマスクは、ラウンド 1 のものに垂直であり、そしてラウンド 2 の各サイクルは、4 つの新たなヌクレオチドを作り出す。この工程は、図 4 において説明するように 16 のヌクレオチドを形成するためにラウンド 2 まで続けられる。ラウンド 3 のマスクは、各カップリングのサイクルが 16 のトリマーを作り出すように上記合成領域をさらに垂分割する。この支持体の垂分割は、テトラヌクレオチドを形成するためにラウンド 4 まで続けられる。このプロープ・マトリックスの合成は、 $(A + C + G + T)^4$ の多項表記法において簡単に表現されることができる。この多項式の展開は、256 のテトラヌクレオチドを作り出す。

20

30

【0068】

プロープ・マトリックスを生り出すために光指定組合せ合成により合成された 256 プロープのアレイの適用を、図 5 A に説明する。この合成のための多項式は： $3' - CG(A + G + C + T)^4 CG$ により与えられる。この合成マップを図 5 B に与える。全ての可能性のあるテトラヌクレオチドを、3' - 末端と 5' - 末端における CG に隣接するように合成した。15 におけるこのアレイへの標的 5' - GCGGCGGC - フルオレセインのハイブリダイゼーションは、最も強い位置 (2, 698 カウント) として S - 3' - CGCCGCCG 相補プロープを正確に作り出した。以下のミスマッチ：

40

【0069】

【数 2】

S - 3' - CGCAGCCG (554 カウント) 、
 S - 3' - CGCCGACG (317 カウント) 、 S - 3' - CGCCGTCG (272
 カウント) 、 S - 3' - CGACGCCG (242 カウント) 、 S - 3' - CGTC
 GCCG (203 カウント) 、 S - 3' - CGCCGCCG (180 カウント) 、 S - 3
 ' - CGCTGCCG (163 カウント) 、 S - 3' - CGCCACCG (125 カウント)
 、 と S - 3' - CGCCICCG (78 カウント)

10

についても有意な強度が観察された。

C. ミスマッチ分析

先に討議したアレイを、より短い長さ k のプローブのアレイを使用して長さ n のオリゴヌクレオチドの核酸配列を決定するための本発明に係る方法において使用されることができ。図 6 は単純な例について説明する。この標的は、配列 5' - X X Y X Y - 3' { ここで、X と Y は相補的核酸、例えば A と T 又は C と G である。 } をもつ。討議目的のために、図 6 における説明を、たった 2 つの塩基とひじょうに短い配列の使用により単純化するが、この技術は、例えば全部で 4 つの RNA 又は DNA の塩基によるより大きな核酸に容易に拡張されることができ。

20

【0070】

標的の配列は、一般に最初から (a b i n i t i o) 知られていない。ある者は、より短いプローブのアレイを用いて本発明を使用して標的の配列を決定することができる。本実施例においては、全ての可能性のある X と Y の 4 - m e r のアレイを合成し、そして次に 5 m e r 標的の配列を決定するために使用する。

【0071】

最初に、“コア (c o r e) ” プローブを同定する。このコア・プローブは、本発明に係るミスマッチ分析法を使用して標的内の配列に正確に相補的である。このコア・プローブは、以下の基準の中の 1 又は両方を使用して同定する：

30

1. コア・プローブは、他のプローブよりも強い標的への結合アフィニティー、典型的には、(先の分析サイクル内でコア・プローブとして同定されていない) アレイ内のいずれかのプローブの最高の結合アフィニティーを示す。

【0072】

2. コア・プローブ配列に比較したとき、標的とミスマッチであるプローブは、以下により詳細に討議するような特徴的なパターンを示す。ここで、そのプローブの 3' - と 5' - 末端においてミスマッチするプローブは、内部の位置においてミスマッチするプローブよりも強く結合する。この特定の実施例においては、選定基準 # 1 は、そのプローブが標的にハイブリダイズしたものとして示されている図 6 A 中に示すように、配列 3' - Y Y X Y をもつ標的に最高の結合アフィニティーをもつコア 4 - m e r プローブを同定する。(標的の 5' - X X Y X 位に一致する) プローブ 3' - Y Y X Y は、それ故、“コア” プローブとして選ばれる。

40

【0073】

選定基準 # 2 は、このコア・プローブが標的核酸に正確に相補的であることを確保するための“チェック (c h e c k) ”として使用される。この第 2 の選定基準は、そのコア・プローブと比較するとき 1 塩基ミスマッチをもつプローブのハイブリダイゼーション・データ (例えば、支持体上のプローブのアレイにハイブリダイズした標識された標的の蛍光強度、但し、他の技術も当業者によく知られている。) を評価する。この特定のケースにおいては、コア・プローブは S - 3' - Y Y X Y として選ばれた。このコア・プローブ

50

の1塩基ミスマッチ・プローブは： $S - 3' - XYXY$ 、 $S - 3' - YXXY$ 、 $S - 3' - YYY Y$ 、と $S - 3' - YYXY$ である。これらの1塩基ミスマッチの結合アフィニティー特性は、“正しい (correct)” コアが選ばれたことを確かめ、又は類似の結合アフィニティーを示すプローブのセットの中からそのコア・プローブを選ぶために使用される。

【0074】

予想される結合アフィニティー対ミスマッチ位置の例示的、仮定的なプロットを図6Bにおいて提供する。これらの結合アフィニティー値(典型的には、プローブにハイブリダイズされた標識された標的の蛍光強度、但し、アフィニティーに関する多くの他の要因を使用することもできる。)は、このグラフの左手の位置上に1の値としてプロットされた、標的への $S - 3' - YYXY$ の結合アフィニティーに対して全て正規化される。たった2つのヌクレオチドが本例において含意されるので、1位(そのプローブの3'-末端におけるヌクレオチド)においてミスマッチするプローブについてプロットされた値は、 $S - 3' - XYXY$ の正規化された結合アフィニティーである。2位におけるミスマッチについてプロットされた値は、 $S - 3' - YXXY$ の正規化されたアフィニティーである、3位におけるミスマッチについてプロットされた値は、 $S - 3' - YYY Y$ の正規化されたアフィニティーであり、そして4位におけるミスマッチについてプロットされた値は、 $S - 3' - YYXX$ の正規化されたアフィニティーである。上記のように、“アフィニティー (affinity)” は、例えば、標的上の蛍光マーカからの光子カウントの数を含む、多くの方法において測定されることができる。

【0075】

全部で3つのミスマッチのアフィニティーは、この説明においてそのコアよりも低い。その上、アフィニティー・プロットは、そのプローブの3'-末端におけるミスマッチが、この特定のケースにおいてはプローブの5'-末端におけるミスマッチよりもより少ない影響力をもつことを示している、但し、これは、にそうではない。さらに、このプローブの末端におけるミスマノチは、このプローブの中央におけるミスマッチよりも少ない妨害をもたらす。図6B中において示すようにプロットされたとき“スマイル (smile)” 形状のグラフをもたらすこれらの特徴は、“正しい” コア・プローブの選定後、又は、標的配列の他の部分に関してのコア・プローブであるミスマッチ・プローブを考慮した後の、1塩基ミスマッチのほとんどのプロットにおいて見られるであろう。この情報は、コア・プローブの最初の選定、又は正しくマッチしたコア・プローブが選ばれていることを確保するためのチェックのいずれかにおいて使用されることができるであろう。もちろん、特定の情報においては、先のセクションB中に述べたように、例えば、法延の又は遺伝子の研究、等において必要とされるコアの同定が、全てである。

【0076】

配列決定研究においては、この工程を次に、このコア・プローブの左及び/又は右伸長のために繰り返す。図6において説明する例においては、このコア・プローブの右伸長だけが可能である。このコア・プローブの可能性のある4-mer伸長プローブは $3' - YXY Y$ と $3' - YXY X$ である。再び同一の選定基準を使用する。 $3' - YXY Y$ と $3' - YXY X$ の間では、 $3' - YXY X$ が最も強い結合アフィニティーをもつであろうことが通見い出されるであろうし、そしてこのプローブは、正しいプローブ伸長として選ばれる。この選定は、上記コア・プローブと比較するとき1塩基ミスマッチをもつプローブの正規化された結合アフィニティーを再びプロットすることにより確かめられることができる。推定プロットを図6Cに説明する。再び、特徴的な“スマイル”パターンが観察され、これは、“正しい”伸長、すなわち $3' - YXY X$ が選ばれていることを示している。この情報から、ある者は、標的の配列が $5 - XXXYXY$ であると正しく結論付けるであろう。

実施例

1. $(A + T)^8$ アレイと1塩基ミスマッチの安定性

20ステップの、4-レプリカ組合せ合成を、Men Poc - dAとMen Poc 50

される。

【0081】

最高の強度をもつプローブは、その標的への完全なマッチであるはずである。対応のミスマッチ曲線を図11A~11Cに示す。ある者は、そのプローブが標的に正確に相補的であることを確認するために最高強度をもつプローブ（このケースにおいてはS-3'-TGGGT T T G）についてのミスマッチ曲線を最初にプロットする。このプローブが標的の断片に相補的であると仮定して、我々は、その標的の3'-末端上に1塩基の“伸長”を考える。この場合、2つのプローブの選択がある。2つの8-mer プローブS-3'-GGGT T T G TとS-3'-GGGT T T G Gの中の1は、標的核酸に正確に相補的であろう。これらの2つのプローブのそれぞれについての並びにプローブS-3'-TGGGT T T Gについてのミスマッチ曲線を図11Aに強度値により示す。プローブS-3'-TGGGT T T GがプローブS-3'-TGGGT T T Gのもの（典型的な“スマイル”プロット）と最も近似したミスマッチ曲線をもつということに注目のこと。それ故、ある者は、正しい伸長プローブがS-3'-GGGT T T G Gであると結論付けるであろう。

10

【0082】

図11Bは、標的配列の3'-末端を評価するための本工程の反復を示す。プローブS-3'-GGT T T G GはコアS-3'-GGGT T T G Gに最も近似したスマイル形のミスマッチ曲線をもち、そしてプローブS-3'-GGT T T G G Tはそうではないので、ある者は、プローブS-3'-GGGT T T G Gが正しい伸長プローブであると正しく結論付けるであろう。この工程は、曲線が正しい形をもつまでか、又は標的の“末端”に到達していることを示す。絶対強度が最高強度のものよりもかなり下廻るまでのいずれかまで、繰り返されることのできる。同様の方法は、5'-末端に伸長する標的の配列を提供する。図11Cは、完全にマッチしたプローブの全てのミスマッチ曲線を示しており；各曲線は、この標的について予測された無矛盾の形状をもつ。

20

【0083】

上記技術は、もちろん、図12A~12Dの各パネルにおいて説明するように、いずれかの長さの核酸に容易に延長されることができる。図12Aにおいて示すように、10-mer 標的が配列決定されるであろうし、そしてその配列は、5'-N₁N₂N₃N₄N₅N₆N₇N₈N₉N₁₀-3' {ここで、Nはいずれかのヌクレオチド又は核酸モノマーであり、そして下付き添字はそのプローブ内のヌクレオチド位置を示し、3はその3'-末端モノマーであることを示す。}により示される。当業者は、プローブが支持体に付着された5'-末端をもって合成された場合、本発明に係る方法が、適当な修飾をもって適用されることができることを理解する。

30

【0084】

より短いオリゴヌクレオチドのアレイを、本発明の1の態様に従ってより大きなヌクレオチドを配列決定するために使用することができる。図12A~12Dに示す特定の例においては、4-mer（長さ4モノマーのオリゴヌクレオチド・プローブ）を未知の10-mer 標的を配列決定するために使用する。実際には、より長いプローブと標的が典型的には使用されるが、この例示的な例は、本発明の理解を容易にする。4-mer アレイの1のメンバーは図12Aに示され、そして配列S-3'-P₃P₄P₅P₆ {ここで、各P（プローブ）ヌクレオチドは、その用途に依存してA, T, C, U, Gと他のモノマーからなる群が選ばれ、そしてその添字は、標的に対する位置を示す。}をもつ。討議目的のために、このハイブリダイゼーション・データは1のアレイから入手されることができると推定される。しかしながら、ある者は、本法を実施するために、多数のアレイ、異なる時において合成されたアレイ、又はさらに個々のプローブを使用することができる。上述のように、4の長さのプローブを討議を容易にするために選ばれ；実際には、より長いプローブが典型的には使用されるであろう。

40

【0085】

S-3'-P₃P₄P₅P₆は、標的への強い結合アフィニティーと正しいミスマッチ

50

特性をそれが示すために、そのアレイからのコア・プローブとして選ばれる。全 4 - m e r のアレイにおいて、配列 S - 3' - P₃ P₄ P₅ P₆ がそのコア配列として選ばれる。なぜなら、フルオレセイン標識された標的 (図 1 2 A 中 5' - N₁ N₂ N₃ N₄ N₅ N₆ N₇ N₈ N₉ N₁₀ - 3' とし示す) が支持体に晒されるとき、その標的が図 1 2 A に示す矢印により示されるように、そのプローブにハイブリダイズし、そして高い蛍光強度 (すなわち、多数の光子カウント) がその支持体の他の部分と比較するとき、プローブ S - 3' - P₃ P₄ P₅ P₆ を含む支持体の部分内で観察されるからである。通、最も強い結合アフィニティーを示す配列が第 1 コア配列として選ばれるであろう。

【 0 0 8 6 】

ある者は、好ましくは、最初に選ばれたコア配列が、1塩基においてコア・プローブと異なるアレイ内のプローブの蛍光強度を調べることにより標的に完全に相補的であるかどうかを確認する。図 1 2 B は、1塩基ミスマッチの相対強度対 S - 3' - P₃ P₄ P₅ P₆ コア・プローブについてのミスマッチの位置の典型的なプロットを定量的に説明する。簡単な例として S - 3' - P₃ P₄ P₅ P₆ においてヌクレオチド C が存在しないと仮定する。図 1 2 B は、配列 S - 3' - P₃ P₄ P₅ P₆ 内への C の置換によりそのコア配列プローブと異なるプローブの正規化された蛍光強度を定性的に説明しており、そして、ここで、C-含有ミスマッチ・プローブのいずれも標的内の他の配列に正確に相補的ではない。従って、図 1 2 B は、それらが標的とハイブリダイズされるとき、そのコア・プローブに正規化された、以下のプローブ・セット

S - 3' - C P₄ P₅ P₆ ,
S - 3' - P₃ C P₅ P₆ ,
S - 3' - P₃ P₄ C P₆ 、及び
S - 3' - P₃ P₄ P₅ C

の相対的蛍光強度をプロットする。他の態様においては、平均曲線が、各位置における可能性のあるヌクレオチドの全ての置換 (ミスマッチ・プローブの “ ファミリー ”) についてプロットされ、又は最も高い強度が各位置についてプロットされる。従って、図 1 2 B におけるグラフの X 軸上の 0 位は、置換がないことを表し、そしてコア・プローブ S - 3' - P₃ P₄ P₅ P₆ への標的ハイブリダイゼーションのための蛍光強度を示す。図 1 2 B 中の全ての値がこの値に関して正規化されているので、“置換無し”のケースは、1の正規化強度をもつ。C が 3, 4, 5 と 6 位において置換されるとき、その相対強度値は、通、小さい。なぜなら、これらの配列のいずれも本例においては標的に正確に相補的でないからである。

【 0 0 8 7 】

3' - 又は 5' - 末端においてミスマッチをもつプローブ / 標的複合体の相対蛍光強度が、プローブ / 標的複合体の中央においてミスマッチをもつ複合体よりも典型的には高い。なぜなら、このプローブの末端におけるミスマッチは、プローブ / 標的複合体の中央におけるミスマッチよりもより少なく脱安定化される傾向をもつからである。このプローブの 3' - 末端にミスマッチをもつプローブ / 標的複合体は、推定するに、この態様においては支持体表面へのそのプローブの 3' - 末端の近接のために、そのプローブの 5' - 末端におけるミスマッチをもつものよりもより少なく影響を及ぼす (そしてそれ故、より高い蛍光強度をもつ) ことができる。それ故、結合アフィニティーに関する正規化係数対ミスマッチ位置をプロットする曲線は、図 1 2 B 中に示すような “ ゆがんだスマイル (c r o o k e d s m i l e) ” の形をもつ傾向がある。

【 0 0 8 8 】

この方法論を使用して、ある者は、そのコア・プローブと同一の配列であるが一端において伸長しており、そして場合により他端においてより短いものをもつアレイ上のプローブを調べることによりそのコア配列を延長することができる。これらのプローブは、標的に完全にハイブリダイズする。プローブを決定するための第 2 のコア・配列として評価されることができる。この工程の繰返しにより、ある者は、その標的のヌクレオチド配列の全体を決定することができる。

【0089】

本法を説明するために、図12Cは、コア・プローブS-3'-P₃P₄P₅P₆の4つの可能性のある4-員“左伸長”を示す。図示するように、S-3'-P₃P₄P₅P₆に相補的な標的の配列に隣接するヌクレオチドはA、T、C、又はGのいずれかであり、又はその標的上に隣接するヌクレオチドは全く存在しない(すなわち、P₃は標的の5'-末端である。)。それ故、P₃P₄P₅P₆コア・プローブの可能性のある左伸長は、プローブS-3'-AP₃P₄P₅、S-3'-TP₃P₄P₅、S-3'-CP₃P₄P₅、とS-3'-GP₃P₄P₅である。この説明の目的のために、Tが実際に“正しい”ものであると推定される。なぜならAは、その標的核酸の相補位置内にあるからである。

10

【0090】

図12Dにおける上左のプロットは、S-3'-AP₃P₄P₅プローブのミスマッチ特性について予想されるハイブリダイゼーション・データを説明する。全てのデータがS-3'-AP₃P₄P₅に対して正規化されている。2-5位の各々における置換の全てについてのデータ点を示すが、各々の位置における3つの値についての平均データも使用されることができ、各の位置における1の置換を使用することができ、3つの値の最高のものを利用してよく、又はいくつかの他の組合せでもよい。S-3'-AP₃P₄P₅グラフにおいて示されるように、1の点は残りのものよりもかなり高い結合アフィニティーを示す。これは、2位におけるAについてのT置換である。AP₃P₄P₅グラフ内の残りのデータは、図12B中に示す通の“スマイル”特徴をもつ。同様のプロットが図12Bの底部分において示されたCとG置換について顕出された。それぞれのケースにおいては、全データ点は、そのグラフ内の推定“コア”プローブに対して正規化されている。

20

【0091】

図12Dの上右部分において示されるT伸長グラフは、3'-AP₃P₄P₅グラフその他のような異曲線をもたないであろう。なぜなら、3'-TP₃P₄P₅プローブの2位におけるモノ置換のいずれも標的に正確に相補的であるであろうからである。従って、2位におけるA、C、とGの置換は全て、標的に対して1塩基ミスマッチをもつプローブについて予想される特徴的な“スマイル”プロットを作り出す。さらに、T置換プローブ/標的複合体の蛍光強度は、C、G、とAプローブ/標的複合体の蛍光強度よりも通高いであろう。これらのデータは、各種組合せにおいてどの伸長が“正しい”かどうかを、そしてそれによりその標的核酸の配列を決定するために、使用されることができ。

30

【0092】

図12A~12D中に示されたデータから、ある者は、コア・プローブに相補的な配列に対してその標的の左伸長物に正確に相補的なプローブが、その標的内の2位においてAモノマーをもつと結論付ける。

【0093】

この工程は、グラフのいずれも適当な特徴をもたなくなるまで繰り返され、この時に、標的の末端に達している。同様に、右伸長が、その標的の末端(又は着目の配列の末端)に達するまで評価される。

40

【0094】

上記技術は、明らかに、ハイブリダイゼーション・データのマニュアル観察を通して行うことができる。しかしながら、好ましい態様においては、データは、1以上の適切にプログラムされたデジタル・コンピューターを使用して分析される。例示のシステムを、図13中に説明する。図中に示すように、このシステムは、CPUの制御下でオペレートされ、そしてメモリー(304)、例えばハード・ディスク、とメモリー(306)、例えばダイナミック・ランダム・アクセス・メモリー(dynamic random access memory)を含む1又は複数のコンピューターを含む。このコンピューターは、支持体(2)の部分にカップリングされた標識された標的ヌクレオチドからの蛍光強度又は他の関連情報を測定するスキャニング装置(308)を制御するために使用され

50

る。

【0095】

この支持体(2)は、その上に既知の位置において既知配列のプローブ・ヌクレオチドを含む。ユーザーは、インプット装置(313)を介してインプットを提供する。

【0096】

蛍光強度又は他の関連情報は、メモリー(304/306)内に保存される。CPU(310)は、その蛍光データを処理して、プリント装置(312)又はディスプレイ(314)の1又は両方にアウトプットを提供する。本明細書中に記載する方法に従ってデータを処理し、そして先に示したようなグラフの形態における、又は核酸モノマーの配列における、又は単純な(+)/(-)アウトプットにおけるアウトプット又はこのようなデータの分析の他の結果を得ることができる。好適なコンピューターは、例えば、IBM PC又は互換性SPARCワークステーション、又は同様の装置を含む。

10

【0097】

図14は、*n-mer*sのアレイを評価するための、そして(ミスマッチ分析のための)正確に相補的な配列又は(配列決定又は他の目的のための)より大きな*k-mer*を同定するための、典型的なコンピューター・プログラムのためのフローチャートである。図中に示すように、システムは最初に、例えば、プローブのいくつかの特定のセットの最も高い結合アフィニティーをもつプローブを選ぶことによりステップ(402)においてコア・プローブを同定する。本法は、しばしば、反復工程においてオペレーション可能であろう。ここでアレイ内の最も高いアフィニティーのプローブは、第1反復(*iteration*)後に選ばれず、そして他の場合においては、例えば、1, 2, 3、又はそれ以上の最高結合プローブを選び、そして各々について左と右の伸長を行い、そして次にこの情報を保存し、そして最終的なアウトプットを提供する前に他のデータとこの情報を比較することが価値のあることである。これらの結果は、正しい配列の確認において援助することができる。

20

【0098】

ステップ(404)において、このシステムは、コア*n-mer*の左伸長の全てを同定する。ステップ(406)において、本システムは：

- どの左伸長物がプリセットされたモノマー置換パターンと最も無矛盾である挙動を示すかを決定し、そして/又は

30

- その最高の結合アフィニティーを示す左伸長を選択する、

の1又は両方により、適切な左伸長選定する。

【0099】

上記選択基準その他は、いくつかの態様において、AND法において使用されることができる。すなわち、これらの基準の両方が適合し、又はこのシステムは、1が末端モノマーに達したか又はそのシステムが許容可能な状態で行われないうずれかを仮定する。他の態様においては、上記基準の中の1が、主な選択機構として選ばれることができ、そして他のものが、注意、すなわち、可能性のある誤った選定、又は他の選定をユーザーに提供するために使用されることができる。

【0100】

その後、本システムは、選定基準がステップ408においていくつかの最小標準に適合したかどうかを決定する。そうでない場合、本システムは、その配列の末端がステップ(410)に達したと仮定する。この選定基準が適合する場合、先のコアからの正しい伸長物として選ばれた新たな“コア”によりステップ(404)の始めに戻される。

40

【0101】

その後、本工程は、右伸長について有効に繰り返される。ステップ(412)において右伸長が同定される。ステップ(414)において、プリセットされたミスマッチ特性のプローブ及び/又は高アフィニティー右伸長物が同定される。ステップ(416)において、本システムは、この分子の末端に達しているかどうかを決定する。そうでない場合、この工程はステップ(412)に戻される。そうである場合、その分子は配列決定されて

50

いると仮定し、そしてこの工程は、プリンター又は他のアウトプット装置への適当なアウトプットをもって終了される。

【0102】

ミスマッチ分析の使用法の他の態様を、図15a中に例示する。この図は、未知の標的断片の配列を決定するためにデザインされた典型的なコンピューター・プログラムのためのフローチャートを表す。このような方法は通、適当にプログラムされたデジタル・コンピューター、例えばIBM PC又は同等なもの、Sunワークステーション、あるいは他の類似のコンピューター・システムにおいて行われるであろう。ステップ(502)に示すように、ハイブリダイゼーション・データ(例えば、上記のような蛍光強度、但し当業者によく知られた他の技術を通じて得られた他のデータも使用することができる。)をこのシステムに入れる。いくつかの態様においては、このデータは実験システムから直接的にインプットされることができ、一方、その他においては、これらのデータは、別個のシステムにおいて採取されることができ、ステップ(504)において、このシステムは、このデータ・セット内の“最良”データに対応するプローブ配列を選択する。蛍光強度データを使用する態様においては、この最良データは、最高強度レベルを示すようなプローブに典型的には対応するであろう。このデータ・セットの不履行パーセンテージが使用されることができ、又は場合により、オペレーターは、そのデータ・セットのサイズ、そのシステムの計算能力、及びそのデータ・セットのサイズのような要因に基づいて選ばれたパーセンテージにより、最強強度データの特定パーセンテージを選ぶためにそのシステムに指令する。ステップ(506)においては、このシステムは、ステップ(504)において選ばれたプローブ配列の組合せを通じて誘導されることができ、特定長さの全断片のセットを同定する。このように同定された断片を“候補”断片という。

10

20

【0103】

ステップ(508)においては、このシステムは、核酸プローブから測定されたハイブリダイゼーション・データを使用して、上記セット内の各候補断片についての完全にマッチしたプローブについて先に記載されたようなミスマッチ分析アフィニティー・プロットを構築する。ステップ(510)においては、各アフィニティー・プロットは、それが、その標的DNAの他の部分とのそのプローブの交差ハイブリダイゼーションについて考慮した後アフィニティー・プロットの予測された形状にどのくらいよく一致するかを反映するように格付けされる。

30

【0104】

この格付けステップ(510)が一旦終了すると、このプログラムは、各候補断片(ステップ512)についてのすべてのアフィニティー・プロットからの等級を合計し、そしてステップ(514)において次の候補断片を検査するために前進する。ステップ(516)において、全候補断片が検査されたとき、このシステムは、個々の候補断片についての合計されたアフィニティー・プロット等級を比較する。最後に、ステップ(518)において、このシステムは、それについて、その等級がその鋳型との最も近い全体的なマッチを示すところの候補断片(単数又は複数)を選択する。この候補断片の配列は、未知の標的断片の配列に対応する。

【0105】

図15bは、アフィニティー・プロットを格付けするための1の好ましい態様についてのフローチャートである。最初に、予想、された結合アフィニティー対ミスマッチ位置(“鋳型”)の所定プロットをシステムにエンターする(ステップ602)。典型的には、この鋳型は上記のような“スマイル”形を示すであろう。ステップ(604)においては、このシステムは、ステップ606~620において特定されるようにその鋳型と各アフィニティー・プロットを比較する。ステップ(606)と(608)においては、このシステムは、その鋳型内の対応ミスマッチ位置の値と各ミスマッチ位置についてのアフィニティー値を比較する。ステップ(610)においては、このアフィニティー値を、最初に、測定された値が鋳型の値の選ばれた間隔内にあるかどうかを決定するために検査される。いくつかの態様においては、この間隔は、不履行レベルにセットされ;その他において

40

50

は、それは、その実験誤差の予想される大きさのような要因に基づいて、オペレーターにより選ばれる。

【0106】

所定のミスマッチ位置における値がその間隔内にある場合、ステップ(612)において、このシステムは、完全候補断片とミスマッチを含む塩基配列(“ミスマッチ配列”)を比較する。ステップ(614)において、このプロセスは、そのミスマッチ配列と上記候補断片のいずれかの部分との間の完全なマッチが存在するかどうかを決定する。このようなマッチが見つかった場合、このシステムは、そのプロットのその等級付けからそのミスマッチ位置を除外し(ステップ616)、そしてステップ(620)に続く;その他の場合、その位置は、その後の計算における包含のためにマークされる(618)。ステップ(620)において、この処理は、各ミスマッチ位置が検査されるまで続けられる。鑄型値と実験値のフィットは次に、ステップ(618)における包含のためにマークされた各ミスマッチ位置についてのステップ(622)の計算において決定される。いくつかの態様においては、この計算は、その間隔内又はその間隔の外側のいずれかとして各点を等級付けすることを含むことができ、その他の態様は、それによりそれぞれの含まれた位置がその予想された値からのその逸脱を反映するように等級付けされるような方法であって、根-平均-二乗(Root-Mean-Squares)分析としてよく知られるような技術を含むものを利用する。最後に、このシステムは、ステップ(624)において各

実施例

8-mer プローブのアレイを、Pirrung et al. に付与された米国特許第5,143,854号と米国出願逐次番号第07/805,727号(両方を全目的をもって引用により本明細書中に取り込む)中に記載された方法に従って構築した。このアレイを16-6mer 標的とインキュベートし、そして公知の方法を使用して蛍光強度について走査した(例えば、引用により全目的をもって本明細書中に取り込む、米国出願逐次番号第08/195,889号を参照のこと。)。得られた蛍光強度データを、これまでの方法に従って分析した、アフィニティー・プロットを作り、そして等級付けした後、69-71のレンジにある等級をもつ4つの候補断片: 5'-AGTTGTAGTTGGATGGT, TGTGTAGTTGGATGGT, GGTTGTAGTTGGATGGT、とCGTTGTAGTTGGATGGTを同定した。この4つの類似の断片は、その断片の5'末端における塩基の同一性においてのみ異なり、それは、他の位置の受ける2本鎖分析の利益を受けない。他の候補断片についての等級からのこれらの4つの候補についての等級の分離は実質的であり;次に高い等級の候補が54の等級を受け取った。

【0107】

さらに他の本発明の態様においては、先に討議したアレイを、以下の方法により未知の標的の断片の配列を同定するために使用することができる。

【0108】

長さ“n”のプローブからのハイブリダイゼーション・データのセット(ここで、“n”は、プローブ内の塩基の数に等しく、そして未知の標的断片の長さ末端である。)を、適当にプログラムされたコンピューター・システムにエンターする。本明細書中に企図されたハイブリダイゼーション・データ・セットは、先に記載されたような蛍光強度データのセット、並びに当業者に知られた他の技術を通じて得られたデータのセットを含む。このシステムは、(Harary, F., Graph Theory, Addison-Wesley, Reading, MA (1969)及びAhuja, R.K., Magnanti, T.L., Orlin, J.B., Network Flows, Theory, Algorithms, and Applications, Prentice Hall, New York (1993)、(両方を全目的をもって本明細書中に取り込む)中に記載されたような)指定グラフを構築した、ここで、頂点(vertices)(又は“節点(nodes)”)は、上記データ・セットの全(n-1)-mersに対応する塩基配列であり、そしてその端は、そのセットの全n-mersに対応する塩基配

列である。各端は、その端が最初に含まれた $(n-1)$ -mer から終りに含まれた $(n-1)$ -mer まで接続するように構築される。標的 k -mer が A C T G T T G として知られており、そして n -mers が 3-mers である上記のような指定グラフのサンプルは、図 16 として包含される。図 16 においては、3-mer 配列に対応する端だけが含まれ；その標的塩基配列が知られていない典型的な場合においては、このシステムは、その分析において n -mer 節点到に接続する全ての可能性のある端を含むであろう。

【0109】

さらに、このシステムは、 $(n-1)$ -mer 節点のいずれとも異なる、指定グラフ内のソース節点、並びにそのソース節点からそのグラフ内の $(n-1)$ -mer 節点のそれぞれまでの端を作り出す。最終的にこのシステムは、先に作り出された節点のいずれとも異なる、グラフ内の沈み節点 (sink nodes)、並びにそのグラフの $(n-1)$ -mer 節点の各々からその沈み節点までの端を作り出す。

10

【0110】

本システムは次に上記グラフの端のそれぞれに値を“コスト (cost)”を指定する。出願節点からの端先導 (edges leading) と沈み節点までの端先導は全てゼロのコストと指定される。他の全ての端は、その端に対応する塩基配列が標的 k -mer 内のセグメントの完全マッチである確率の関数としてのコストを指定される。この確率を、“完全マッチの確率 (probability of a perfect match)”という。好ましい態様においては、各 n -mer 端に指定されたコストは、その端に対応する配列が完全にマッチしない確率により除された、その端に対応する配列がその標的断片内のセグメントの完全マッチである確率の自然対数の負に等しい。いくつかの態様においては、これらの端は、別々の、追加の値“キャパシティー (capacity)”を指定されることができ；このような態様については、出発節点から他の節の各々までの端先導、並びに沈み節点までの端先導は、全て1のキャパシティーを指定され；このグラフの他の全ての端は、無限キャパシティーを指定される。

20

【0111】

出発節点から始まり、そしてその指定端に沿って続き、多くの経路がそのグラフを通じてトレースされることができ、沈み節点に終わり、それが、標的断片の配列におそらく対応することができる。このグラフの各々端がコストを指定されているので、これらの経路の各々についての全コストを決定することができる。この方法の最終段階においては、このシステムは、出発節点に始まり、そして沈み節点において終り、グラフを通してのそれぞれの可能性のある経路についてのこの全コストを決定するためにこのグラフにアルゴリズムを適用する。この段階における使用についてもくろまれるアルゴリズムは、当業者によく知られた最小コスト/最大フロー・プログラミング・アルゴリズム、例えば、引用により本目的をもって本明細書中に取り込まれる Ahuja et al., Network Flows, Theory, Algorithms, and Applications, Prentice Hall, New York (1993)、中に記載されるものを含む。このようなアルゴリズムを使用して、このシステムは、それについて完全マッチの最高の組合せの確率が存在する合計コストをもつ経路を選び出す。例えば、最小コスト/最大フロー・アルゴリズムを使用する態様においては、完全マッチの最大確率をもつ経路は、最小コストをもつ経路に対応する。それ故、この経路に関連する塩基の配列は、その標的についての最も類似する配列に対応する。

30

40

【0112】

好ましい態様においては、個々のプローブについての完全マッチの確率は、以下のように蛍光ハイブリダイゼーション・データから決定される。選ばれたアレイの各々のプローブについて、一連の関連プローブについてのハイブリダイゼーション・データがこのシステムにエンターされ；着目のプローブに関連するプローブの強度のパターンを次に、上記のミスマッチ分析を使用して公知の標的配列とのマッチについて分析する。この段階により考慮される関連プローブのセットは、典型的には、利用可能なデータ・セットのサイズと計算キャパシティーのような要因に基づいてオペレーターにより選ばれるような、1以

50

上の塩基のミスマッチをもつようなプローブを含む。このセットに含まれるデータは、典型的には、そのプローブの強度と塩基組成を含み、そして場合により、そのプローブがその標的内のどこかと完全マッチになることを引き起こすであろう塩基置換に関するデータ（“クロストーク(crosstalk)”）を含むであろう、好ましくは、これらのデータは、標的として既知のDNA配列を使用した大シリーズの実験のために集められるであろうし；異なる実験からの投入データを次に神経ネットワーク(Neural Network)、又は(例えば、Rich, E., Knight, K. Artificial Intelligence, 2nd, McGraw-Hill, Inc., New York (1991)、これを全目的をもって本明細書中に取り込む、中に記載されているような)他のよく知られた学習に基づく分類法を通じて分析して、プローブが強度、塩基組成、及び場合によりクロストークの関数として完全マッチであるかどうかを予測するためのアルゴリズムを開発することができる。このような予測の信頼レベルは、このアルゴリズム内で推定されることができ；この信頼レベルは、コスト値を上記指定グラフの端に指定することにおいて使用される完全マッチの確率に対応する。

10

D. 適用

本明細書中に記載する技術は、特に、標的核酸が特定のヌクレオチド配列又は既知配列と異なるいくつかの他の配列をもつかどうかを決定することが望まれる場合に、広いレンジの適用をもつであろう。例えば、本明細書中の本発明の1の適用は、突然変異検出に見い出される。これらの技術は、診断、法廷、生体分析論、その他を含む多種多様な分野に適用されることができ。

20

【0113】

例えば、“野生型”の核酸が配列 $5' - N_1 N_2 N_3 N_4$ {ここで、再びNは、モノマー、例えば核酸中のヌクレオチドを表し、そして添字は位置番号を表す。}をもつと仮定する。標的核酸が、それが $5' - N_1 N_2 N_3 N_4$ と同一であるかどうか又はそれがこの配列と異なり、そしてそれが突然変異又は突然変異体配列を含むかどうかを決定するように評価されると仮定する。この標的核酸は、先に討議したように、典型的にはより短いプローブのアレイに最初に晒される。その後、1以上の“コア”配列が同定され、その各々が、その標的が突然変異配列又は突然変異を含まない場合、その標的に高い結合アフィニティーをもつと予想されるであろう。この特定の例においては、高い結合アフィニティーを示すと予想されるであろう1のプローブは、3-merアレイが使用されると仮定して、 $5' - N_1 N_2 N_3$ ($3' - P_1 P_2 P_3$)に相補的であるであろう。再び、このプローブ及び/又は標的がより長い核酸分子の一部であることができると認められるであろう。

30

【0114】

最初のスクリーニング・ツールとして、 $3' - P_1 P_2 P_3$ プローブへの標的の絶対結合アフィニティーがその標的の最初の3つの位置が予想された配列をもつかどうかを決定するために使用されるであろう。 $5' - N_1 N_2 N_3$ への相補性とその標的への強い結合を示さない場合、その標的が野生型をもたないということが適切に結論付けられることができる。

【0115】

1塩基ミスマッチ特性も、標的が突然変異体又は野生型配列を含むかどうかを決定するために本発明に従って使用されることができ。図17Aと17Bは、野生型(図17A)と突然変異体(図17B)である標的から生じる典型的な例示プロットを説明する。図示するように、野生型標的のための1塩基ミスマッチは、一般的には、典型的な、スマイル形のプロットに従う。これに反し、標的が特定の位置において突然変異をもつとき、特定コア・プローブへの標的の絶対結合アフィニティーが小さいであろうばかりでなく、その1塩基ミスマッチ特定が予想される挙動から逸脱するであろう。

40

【0116】

本発明の1の態様に従って、核酸の選択された群をもつ支持体(他には、核酸の“ライブラリー”と本明細書中にいう。)を、特定の核酸が野生型又は他の予想された核酸と同

50

一であるか又は異なるかどうかの決定において使用される。核酸のライブラリーは、通、プローブのアレイ又は“プローブ・アレイ”として提供されるであろう。このようなプローブ・アレイは、好ましくは、プローブの同一性がその支持体上のその位置によって決定されるような単一支持体上で形成される。場合により、このような支持体は、その標的のヌクレオチドが野生型と同一であるかどうかを決定するだけでなく、その標的に関する配列情報をも提供するであろう。このような支持体は、上述の分野、例えば、法延、診断、その他における用途を見い出されるであろう、単に特定の例により、本発明は、鎌状赤血球貧血に関連する診断、多数のP-53突然変異のいずれかの検出、多数の嚢胞性線維症突然変異のいずれかのために、高く多形のHLAクラス1又はクラス2遺伝子（特にクラス2DP, DQとDRベータ遺伝子）に関連するいずれかの特定変異体配列のために、並びに遺伝子疾患、遺伝子素因、及び遺伝子評価に関連する多くの他の配列のために、使用されることができる。

10

【0117】

支持体がこのような用途において使用されるとき、その支持体上に特定の長さの可能性のある核酸の全てを提供することは必要でない。実際、本発明を使用して可能性のある配列の全ての比較的小さなサブセットだけを提供することが必要であろう。例えば、標的核酸が特に着目の5-塩基配列を含んで成り、そしてある者がその5-塩基配列内の1置換を検出するために使用されることができる支持体を開発することを欲すると想定する。本発明の1の態様に従って、支持体を、5塩基配列の1塩基ミスマッチ・プローブの全て又はほとんどと一緒に、その表面上に形成された予想5塩基配列により形成されるであろう。従って、支持体上に全ての可能性のある5塩基配列を含むことは必要でないであろう。但し、より大きなアレイがしばしば好ましいであろう。典型的には、本発明に従う支持体上の核酸プローブの長さは、約5と100塩基の間、約5と50塩基の間、約8と30塩基の間、又は約8と15塩基の間であろう。

20

【0118】

特定の長さの全ての可能性のあるプローブの中での1塩基ミスマッチ・プローブの選択により、支持体上のプローブの数かなり制限されることができる。例えば、3塩基配列においては、69の可能性のあるDNA塩基配列が存在するが、予想された配列に正確に相補的なたった1のものと9つの可能性のある1塩基ミスマッチ・プローブが存在するであろう。これらのプローブだけを選ぶことにより、スクリーニングに必要な多様性が減少されるであろう。好ましくは、必ずしもそうではないが、このような1塩基ミスマッチ・プローブの全てが1の支持体上で合成される。支持体はその1塩基ミスマッチに加えて着目の他のプローブを含むようにしばしば形成されるけれども、このような支持体は通、n-塩基の可能性のあるプローブの全ての50%未満、しばしば、n-塩基の可能性のあるプローブの全ての20%未満、n-塩基の可能性のあるプローブの10%未満、そしてしばしばn-塩基の可能性のあるプローブの5%未満をもつであろう。

30

【0119】

核酸プローブは、しばしば、特定の遺伝子配列の分析のためのキットにおいて提供されるであろう。1の態様に従えば、このキットは、着目の標的核酸に相補的なプローブを含むであろう。さらに、キットは、標的の1塩基ミスマッチを含むであろう。このキットは通、このようなプローブの1以上のC, G, T, A及び/又はUの1塩基ミスマッチを含むであろう。このようなキットは、しばしば、本明細書中の教示に従って特定の核酸サンプルの配列の決定において、相補的プローブと1塩基ミスマッチの使用のための適当な指示を提供されるであろう。本発明の1の態様に従って、本キットは、1塩基ミスマッチだけと共に、その標的に対する相補的を提供する。このようなキットは、しばしば、それが特定の遺伝子の特徴を示すかどうかを決定するための遺伝子材料の適当なサンプルの評価において使用されるであろう。このようなキットは、鎌状赤血球貧血の検出、多数のP-53突然変異の中のいずれかの検出、多数の嚢胞性線維症突然変異の検出、高く多形のHLAクラス1又はクラス2遺伝子（特にクラス2DP, DQ及びDRベータ遺伝子）に関連した特定の変異体配列の検出、並びに遺伝子疾患、遺伝子素因、及び遺伝子評価に

40

50

関連した他の多くの配列の検出における、上述のようなサンプルの評価において使用されることができる。

【0120】

従って、本発明に従って選ばれるプローブをもつ支持体は、多くの突然変異の検出並びに他の働きを行うことができるであろうが、このような働きを行うために限定された数のプローブだけを必要とするであろう。

実施例

1. (G+T)⁸ アレイと示差配列決定 (Differential Sequencing)

(G+T)⁸ アレイを調製し、そして (5' - AACCCCAA CCC と比べたとき 突然変異体配列を表す) 1 n M 10

【0121】

【数3】

5' - AACCCAACCCC - フルオレセイン ;

20

と共にインキュベートし、そしてその配列が“野生型”又は“突然変異体”であるかどうかをテストするために走査した。得られた画像を図18に与える。4つの重複した、正確に相補性のオクタヌクレオチド・プローブ/標的ハイブリダイゼーションが、ある者が、その標的がプローブ; S-3'-TTGGGT TG, TGGGT TG, GGGTTGGG、とGGTTGGGGにより5'-AACCCAAACCでなければならないと仮定する場合に予想される。これらの結果は、1塩基変更の効果が、特に、アレイ上に形成する異なるミスマッチ・プローブ/標的複合体の数と同一性においてきわめて劇的であるということを示す。図18においてシグナルを作り出す標的核酸が5'-AACCCAAACC (すなわち、野生型) であるとある者が仮定する場合、相補的プローブS-3'-TTGGGT TTのためミスマッチ特性を図19Aに示す。このミスマッチ特性は、予想された形状をもたず、そしてプローブ/標的複合体は低い蛍光強度をもつ。8位におけるミスマッチに対応する強いピークは、この標的内のこの位置内の“正しい”塩基がおそらくAであるということを示している。なぜなら、AとCだけがこの実験において標的内にあるからである。6位のミスマッチも小さなピークを示す。これに反し、コアとしてS-3'-TTGGGT TGプローブ配列を用いた同様のプロットは、“スマイル”形と高い蛍光強度を作り出した。図19Bにおいて、次の8-merプローブについての同一の特性を示す。このピークは、左に1位ミスしており、再び、この配列が標的内の8位において野生型から変化していることを確かなものとする。これらは、その元の11-merの標的断片内の同一の位置に対応していた。これらのデータは、野生型に比べたとき、標的の8位において1塩基の変化が存在することを予測させる。 30 40

【0122】

仮定断片5'-AACCCAAACCに対応するミスマッチ・プローブの特性を図19Cに示す。ある者は、上記配列から下位に“動く(moving)”突然変異体の位置を観察する。最終的に、図19Dにおいて、このミスマッチ・プロットは、予想されたスマイル特徴を伴って、5'-AACCCAAACCに相補する4つのプローブに対応することが示される。

E. 結論

本発明は、ヌクレオチド配列及び核酸の他の分子との相互作用の研究のための改良方法と装置を提供する。これまでの記載は例示的であり、そして制限的なものではない。本発 50

明の多くの変更は、本開示のレビューに基づき当業者に明らかとなるであろう。単に、実施例によって、本明細書中に記載した本発明の特定のものが、他のポリマー、例えばペプチドとタンパク質への適用をもつであろうし、そして他の合成技術を使用することができる。それ故、本発明の範囲は、これまでの説明を参照して決められるべきではないが、その代わりに、それらの十分な均等の範囲と共に添付クレームを参照して決定されなければならない。

【図面の簡単な説明】

【0123】

【図1】図1は、オリゴヌクレオチドの光指定合成 (light-directed synthesis) を説明する。光保護されたヒドロキシル (OH) を担持する表面 (2) が、その光脱保護された領域内で遊離のヒドロキシル (OH) を生成するフォトリソグラフィ・マスク (photolithographic mask) (M_1) を通して照射される。これらのヒドロキシル基は次に、5' - 光保護されたデオキシヌクレオチド・ホスホoramidite (例えば、T-X) にカップリングされる。新たなマスク (M_2) を、その表面上に新たなパターンを照射するために使用し、そして第2の光保護されたホスホoramidite (例えば、C-X) を次にカップリングする。照射とカップリングのラウンドを、オリゴヌクレオチド・プローブの所望のセットが得られるまで繰り返す。場合により標識 (*) をもつ、標的 (R) を、これらのオリゴヌクレオチドに晒す。標的がアレイに結合する位置を、その標的の配列を決定するために使用する；

【図2】図2は、オリゴヌクレオチドのハイブリダイゼーションと熱解離を説明し、プローブのアレイにハイブリダイズした標的核酸 (5' - GCGTAGGC - フルオレセイン) の蛍光スキャンを示す。この支持体表面は、488 nm アルゴン・イオン・レーザー励起を使用して Zeiss AxioScop 20 顕微鏡により走査された。520 nm を超える蛍光エミッションを、光子計数モードにおいて操作された冷却光増幅装置 (Hamatsu 934-02) を使用して検出した。シグナル強度は、この画像の右に示したスケール上に示される。温度は において各パネルの右に示される；

【図3】図3は、ハイブリダイゼーションの配列特異性を説明する。(A) は、各合成部位におけるプローブ組成の係数である。3' - CGCATCCG 表面固定プローブ (本明細書中 S-3' - CGCATCCG という。) がストライプ 1, 3, と 5 内で合成され、そしてプローブ S-3' - CGCTTCCG がストライプ 2, 4, と 6 内で合成された。(B) は、標的核酸 (10 nM 5' - GCGTAGGC - フルオレセイン) とのこの支持体のハイブリダイゼーションを示す蛍光画像である。ハイブリダイゼーションを 15 分間、15 において $6 \times$ SSPE, 0.1% Triton X-100 中で行った。(C) は、(B) のハイブリダイゼーション溶液に添加された第2核酸 (10 nM 5' - GCGAAGGC) とのハイブリダイゼーションを示す蛍光画像である。(D) は、(1) (C) からのフルオレセイン化標的の高湿度解離；及び (2) 15 分間 15 における標的核酸 (10 nM 5' - GCGAAGGC) との支持体のインキュベーション、の後のハイブリダイゼーション結果を示す蛍光画像である。(E) は、(D) のハイブリダイゼーション溶液に添加された第2核酸 (10 nM 5' - GCGTAGGC) とのハイブリダイゼーションを示す蛍光画像である；

【図4】図4は、 4^4 テトラヌクレオチドの組合せ合成 (combinatorial synthesis) を説明する。ラウンド1においては、合成領域の $1/4$ が、第1 MeNPoc-ヌクレオチド (このケースにおいては T) のカップリングのためにマスク1を通しての照射により活性化される。ラウンド1のサイクル2においては、マスク2は、その合成支持体の異なる $1/4$ セクションを活性化し、そして異なるヌクレオチド (C) がカップリングされる。さらなるアレイのリソグラフィ的垂直分割及び化学的カップリングは、完全な 256 テトラヌクレオチドのセットを生り出す；

【図5A】図5Aと5Bは、256 オクタヌクレオチドのアレイへのハイブリダイゼーションを説明する。図5Aは、15 における 15 分間 $6 \times$ SSPE, 0.1% Triton

10

20

30

40

50

on X - 100 中での標的核酸 (10 nM 5' - GCGGC GGC - フルオレセイン) とのアレイのハイブリダイゼーション後の蛍光画像である。図 5 B は、S - 3' CG (A + G + C + T)⁴ CG の合成の間に作られた各々のプローブがどこに位置するかを示すマトリックス解読器である。プローブ配列 S - 3' - CGCGCCCG を含む部位を暗領域として示す。本明細書中に使用する組合せ合成の表記法 (notation) は、全目的をもって引用により本明細書中に取り込む米国出願逐次番号第 07 / 624, 120 号中に十分に記載されている；

【図 5 B】図 5 A と 5 B は、256 オクタヌクレオチドのアレイへのハイブリダイゼーションを説明する。図 5 A は、15 における 15 分間 6 x S S P E, 0.1% Triton X - 100 中での標的核酸 (10 nM 5' - GCGGC GGC - フルオレセイン) とのアレイのハイブリダイゼーション後の蛍光画像である。図 5 B は、S - 3' CG (A + G + C + T)⁴ CG の合成の間に作られた各々のプローブがどこに位置するかを示すマトリックス解読器である。プローブ配列 S - 3' - CGCGCCCG を含む部位を暗領域として示す。本明細書中に使用する組合せ合成の表記法 (notation) は、全目的をもって引用により本明細書中に取り込む米国出願逐次番号第 07 / 624, 120 号中に十分に記載されている；

【図 6 A】図 6 A ~ 6 C は、k - mer プローブを使用した n - mer 標的を配列決定するための技術を説明する。図 6 A は、支持体上のプローブにハイブリダイズした標的を説明する。図 6 B と 6 C は、通の結合アフィニティ - 対ミスマッチ位置のプロットを説明する；

【図 6 B】図 6 A ~ 6 C は、k - mer プローブを使用した n - mer 標的を配列決定するための技術を説明する。図 6 A は、支持体上のプローブにハイブリダイズした標的を説明する。図 6 B と 6 C は、通の結合アフィニティ - 対ミスマッチ位置のプロットを説明する；

【図 6 C】図 6 A ~ 6 C は、k - mer プローブを使用した n - mer 標的を配列決定するための技術を説明する。図 6 A は、支持体上のプローブにハイブリダイズした標的を説明する。図 6 B と 6 C は、通の結合アフィニティ - 対ミスマッチ位置のプロットを説明する；

【図 7】図 7 は、ハイブリダイゼーション実験の蛍光画像を説明する；

【図 8】図 8 は、1 塩基ミスマッチの関数としてハイブリダイゼーション事件をグラフにより説明する；

【図 9】図 9 は、ミスマッチの対の関数としての蛍光強度を説明する；

【図 10】図 10 は、1 塩基ミスマッチ実験の蛍光画像を説明する；

【図 11 A】図 11 A ~ 11 C は、各種 1 塩基ミスマッチ特性を説明する；

【図 11 B】図 11 A ~ 11 C は、各種 1 塩基ミスマッチ特性を説明する；

【図 11 C】図 11 A ~ 11 C は、各種 1 塩基ミスマッチ特性を説明する；

【図 12 A】図 12 A ~ 12 D は、より短い k - 員プローブからのハイブリダイゼーション結果に基づく n - 員 (そのヌクレオチジ内のモノマーの数) の標的オリゴヌクレオチドのヌクレオチド配列の決定方法を説明する。特に、図 12 A ~ 12 D は、4 - 塩基プローブにより 10 塩基標的を配列決定するための本発明の適用について説明する；

【図 12 B】図 12 A ~ 12 D は、より短い k - 員プローブからのハイブリダイゼーション結果に基づく n - 員 (そのヌクレオチジ内のモノマーの数) の標的オリゴヌクレオチドのヌクレオチド配列の決定方法を説明する。特に、図 12 A ~ 12 D は、4 - 塩基プローブにより 10 塩基標的を配列決定するための本発明の適用について説明する；

【図 12 C】図 12 A ~ 12 D は、より短い k - 員プローブからのハイブリダイゼーション結果に基づく n - 員 (そのヌクレオチジ内のモノマーの数) の標的オリゴヌクレオチドのヌクレオチド配列の決定方法を説明する。特に、図 12 A ~ 12 D は、4 - 塩基プローブにより 10 塩基標的を配列決定するための本発明の適用について説明する；

【図 12 D】図 12 A ~ 12 D は、より短い k - 員プローブからのハイブリダイゼーション結果に基づく n - 員 (そのヌクレオチジ内のモノマーの数) の標的オリゴヌクレオチド

10

20

30

40

50

のヌクレオチド配列の決定方法を説明する。特に、図12A~12Dは、4-塩基プローブにより10塩基標的を配列決定するための本発明の適用について説明する；

【図13】図13は、ヌクレオチド配列の決定のためのコンピューター・システムを説明する；

【図14】図14は、標的核酸のミスマッチ分析のための、そして、その決定のためのコンピューター・プログラムを説明する；

【図15A】図15Aと15Bは、いくつかの可能性の中の選択により標的核酸のヌクレオチド配列を決定するためのコンピューター・プログラム、並びにこのコンピューター・プログラムにおける使用のためのスコアリング・ルーチン(Scoring routine)の例を説明する；

【図15B】図15Aと15Bは、いくつかの可能性の中の選択により標的核酸のヌクレオチド配列を決定するためのコンピューター・プログラム、並びにこのコンピューター・プログラムにおける使用のためのスコアリング・ルーチン(Scoring routine)の例を説明する；

【図16】図16は、ヌクレオチド配列の決定における使用のための指定グラフを説明する；

【図17A】図17Aと17Bは、1塩基ミスマッチ特性を使用した野生型と突然変異の分析を説明する；

【図17B】図17Aと17Bは、1塩基ミスマッチ特性を使用した野生型と突然変異の分析を説明する；

【図18】図18は、1塩基ミスマッチ・テストの蛍光画像である；

【図19A】図19A~19Dは、核酸配列同定のための技術を説明する。

【図19B】図19A~19Dは、核酸配列同定のための技術を説明する。

【図19C】図19A~19Dは、核酸配列同定のための技術を説明する。

【図19D】図19A~19Dは、核酸配列同定のための技術を説明する。

【図1】

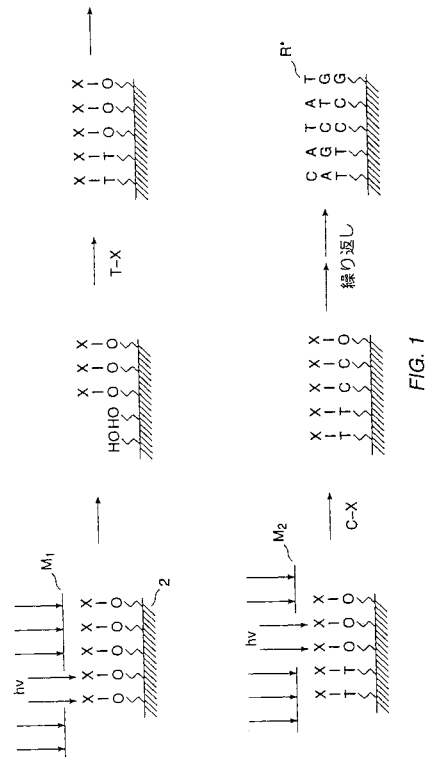


FIG. 1

【図2】

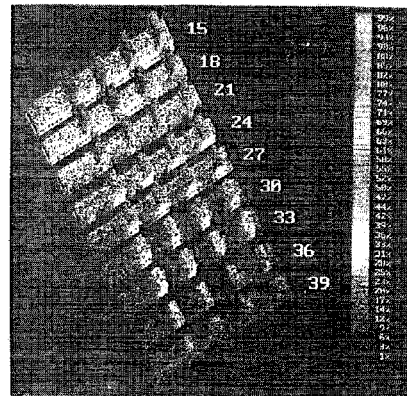


FIG. 2

10

20

【 図 3 】

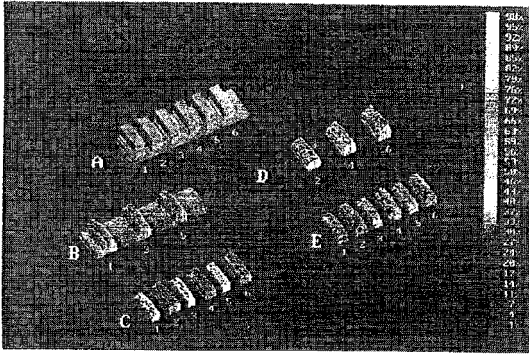


FIG. 3

【 図 4 】

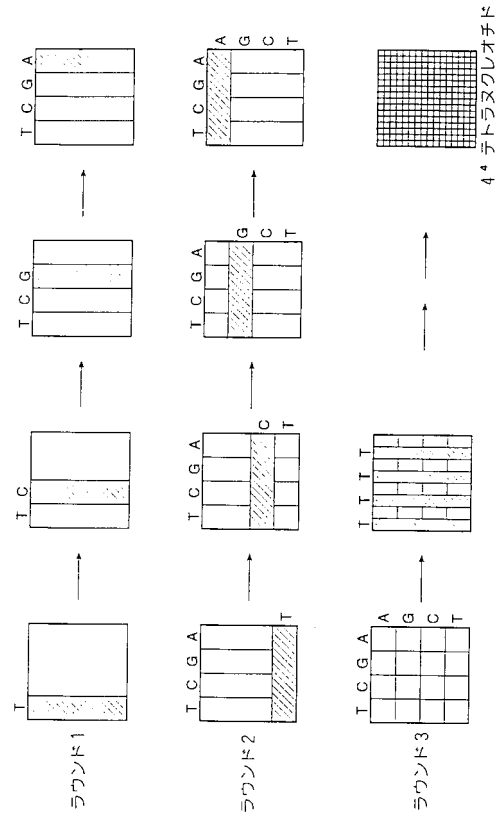


FIG. 4

【 図 5 A 】

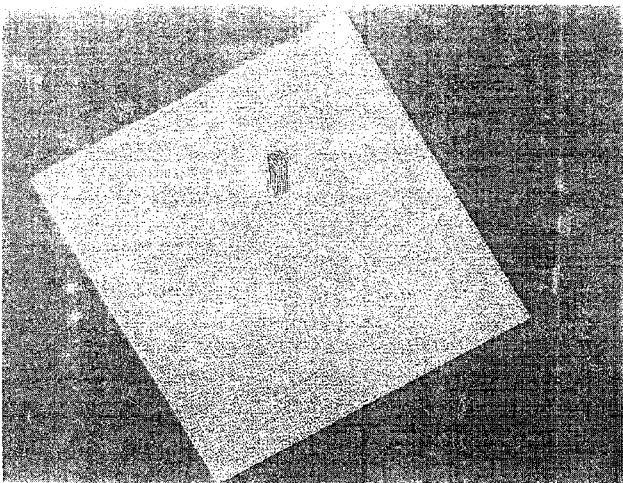


FIG. 5a

【 図 5 B 】

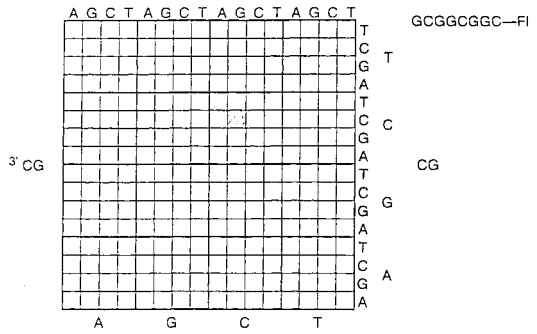


FIG. 5B

【 図 6 A 】

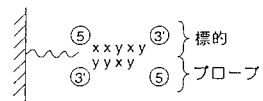


FIG. 6A

【 図 6 B 】

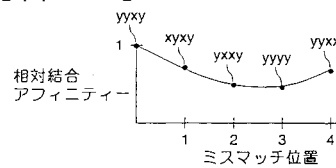
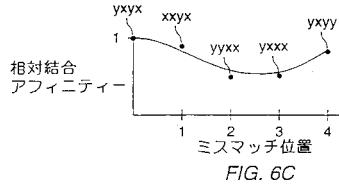
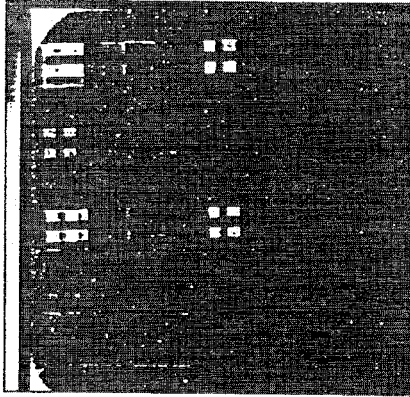


FIG. 6B

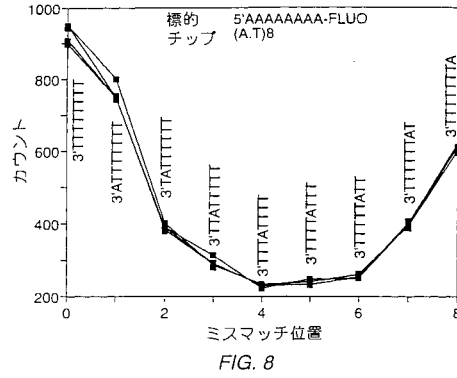
【 図 6 C 】



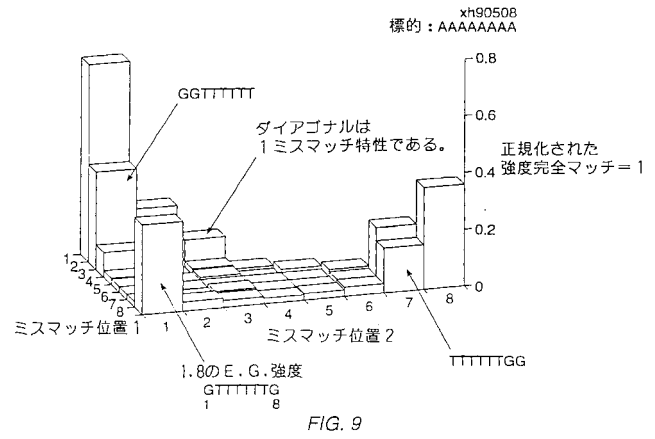
【 図 7 】



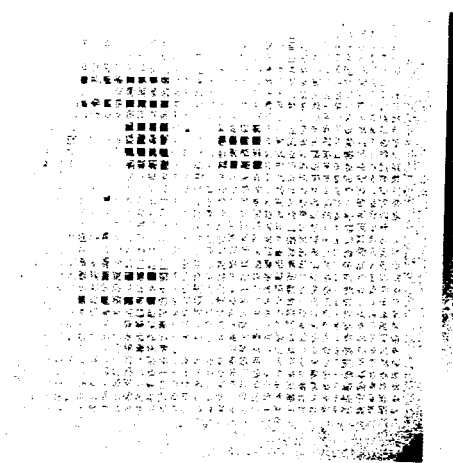
【 図 8 】



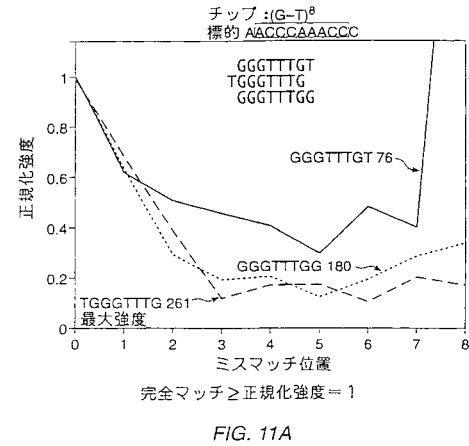
【 図 9 】



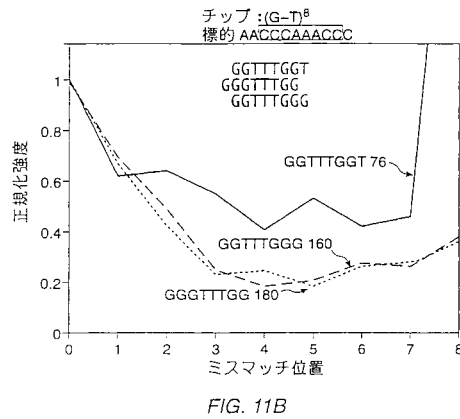
【 図 10 】



【 図 11 A 】



【 図 11 B 】



【図 1 1 C】

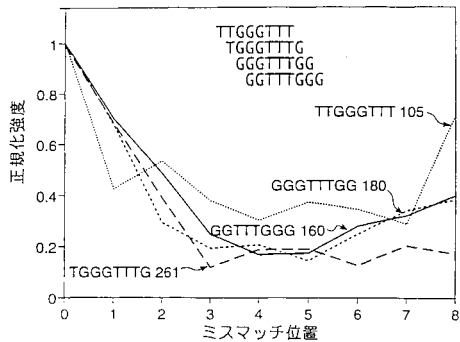


FIG. 11C

【図 1 2 A】

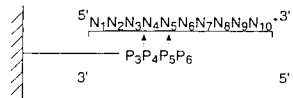


FIG. 12A

【図 1 2 B】

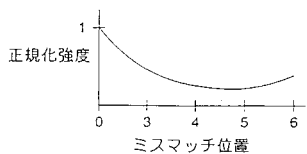


FIG. 12B

【図 1 2 C】

AP₃P₄P₅
 TP₃P₄P₅
 CP₃P₄P₅
 GP₃P₄P₅

FIG. 12C

【図 1 2 D】

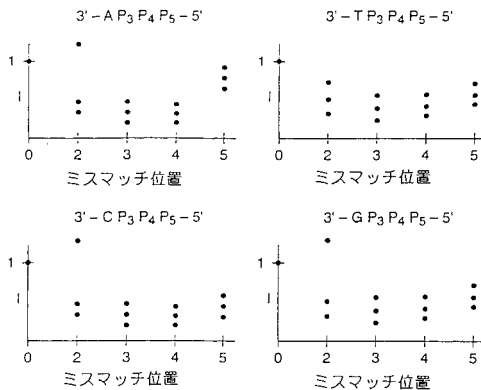


FIG. 12D

【図 1 3】

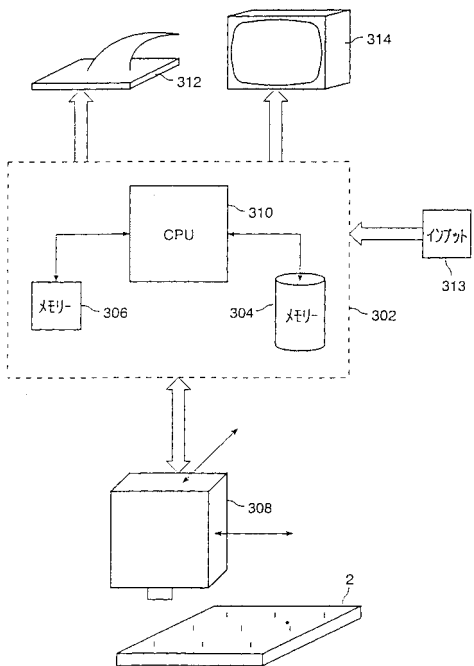


FIG. 13

【図 1 4】

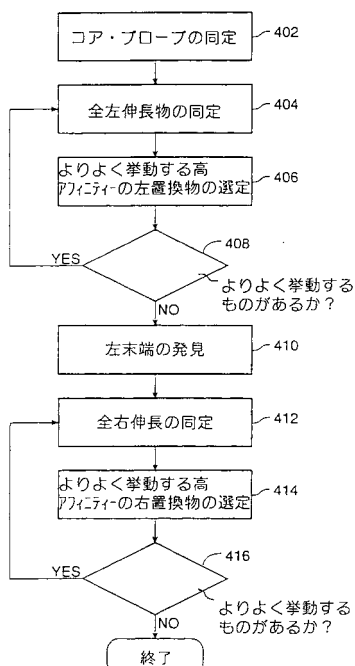


FIG. 14

【図 15 A】

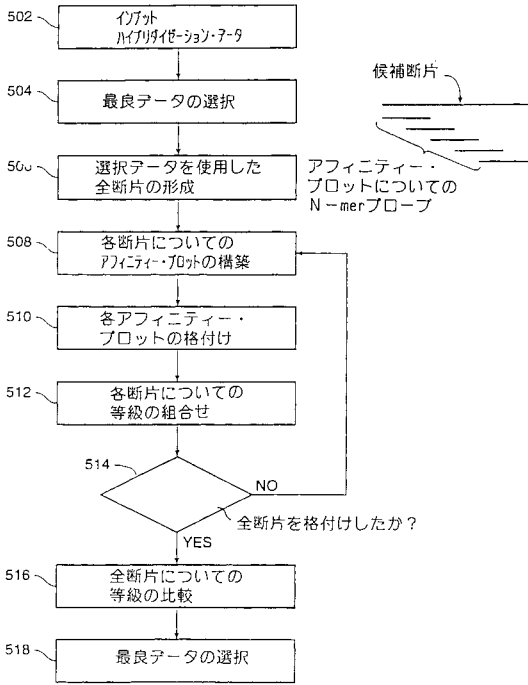


FIG. 15A

【図 15 B】

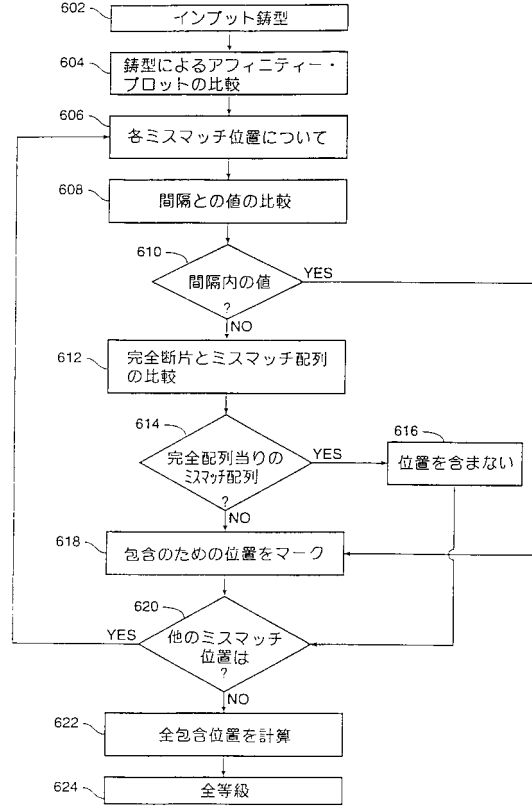


FIG. 15B

【図 16】

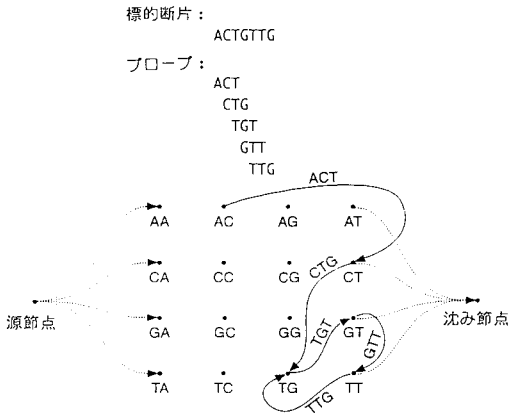


FIG. 16

【図 17 B】

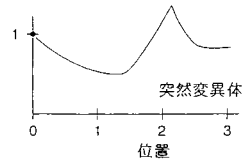


FIG. 17B

【図 17 A】

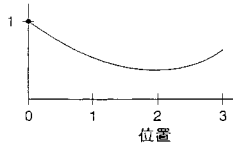


FIG. 17A

【 図 18 】

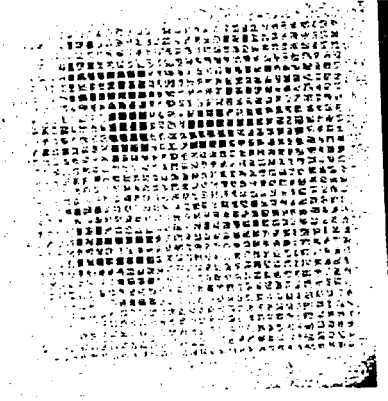
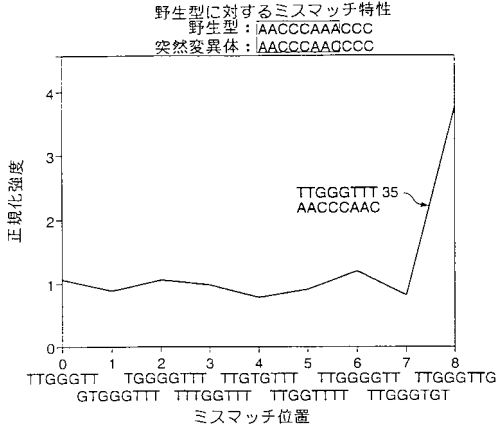
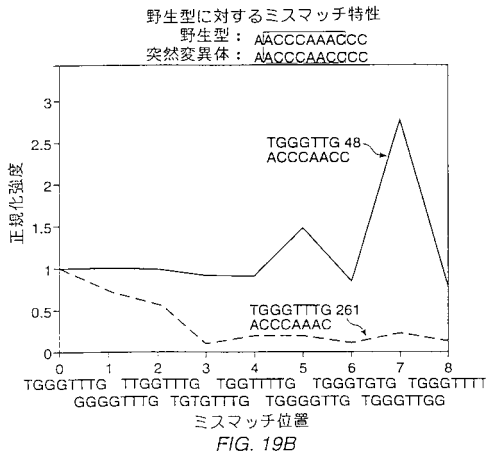


FIG. 18

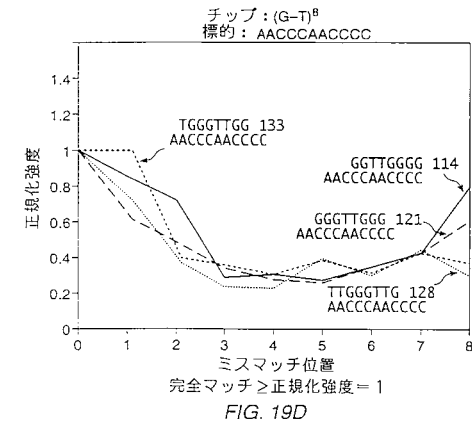
【 図 19 A 】



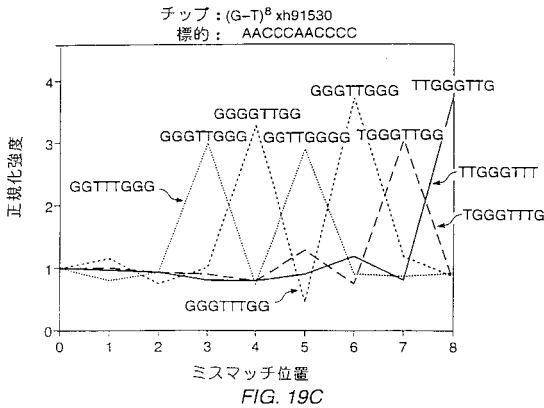
【 図 19 B 】



【 図 19 D 】



【 図 19 C 】



フロントページの続き

| (51) Int.Cl. ⁷ | F I | テーマコード(参考) |
|---------------------------|--|------------|
| G 0 1 N 37/00 | C 1 2 N 15/00 | A |
| | C 1 2 N 15/00 | F |
| | | |
| (72)発明者 | スティーブン ピー.エー. フォドア | |
| | アメリカ合衆国, カリフォルニア 9 4 3 0 3, パロ アルト, ナザン ウェイ 3 8 6 3 | |
| (72)発明者 | ロバート ジェイ. リップシュッツ | |
| | アメリカ合衆国, カリフォルニア 9 4 3 0 1, パロ アルト, パロ アルト アベニュー 9 7 0 | |
| (72)発明者 | シャオフア ファン | |
| | アメリカ合衆国, カリフォルニア 9 4 0 4 3, マウンテン ビュー. ジャクソン ストリート 9 3 7 | |
| (72)発明者 | ルイス キャロル ジェボンズ | |
| | アメリカ合衆国, カリフォルニア 9 4 0 8 7, シニーベール, ラモンネ アベニュー 7 0 1 | |
| F ターム(参考) | 2G054 AA08 BB13 CA22 CE02 GA04 GE01 JA01 | |
| | 4B024 AA11 CA04 CA09 CA11 HA14 | |
| | 4B029 AA07 AA23 BB20 CC03 | |
| | 4B063 QA08 QA12 QA17 QA18 QA19 QQ02 QQ42 QQ52 QR56 QR62 | |
| | QR82 QS34 QS39 QX02 | |