

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6285929号
(P6285929)

(45) 発行日 平成30年2月28日 (2018. 2. 28)

(24) 登録日 平成30年2月9日 (2018. 2. 9)

(51) Int. Cl.	F I
C 1 2 N 15/09 (2006. 01)	C 1 2 N 15/00 Z N A A
C 1 2 M 1/00 (2006. 01)	C 1 2 M 1/00 A
C 1 2 Q 1/68 (2018. 01)	C 1 2 Q 1/68 A
G O 1 N 37/00 (2006. 01)	G O 1 N 37/00 1 O 2

請求項の数 9 (全 100 頁)

(21) 出願番号	特願2015-523238 (P2015-523238)	(73) 特許権者	515015023
(86) (22) 出願日	平成25年7月17日 (2013. 7. 17)		カウンスル、インコーポレーテッド
(65) 公表番号	特表2015-531588 (P2015-531588A)		アメリカ合衆国カリフォルニア州9408
(43) 公表日	平成27年11月5日 (2015. 11. 5)		O, サウス・サンフランシスコ, キンボー
(86) 国際出願番号	PCT/US2013/050965		ル・ウェイ 180
(87) 国際公開番号	W02014/015084	(74) 代理人	100140109
(87) 国際公開日	平成26年1月23日 (2014. 1. 23)		弁理士 小野 新次郎
審査請求日	平成28年4月18日 (2016. 4. 18)	(74) 代理人	100075270
(31) 優先権主張番号	13/551, 584		弁理士 小林 泰
(32) 優先日	平成24年7月17日 (2012. 7. 17)	(74) 代理人	100101373
(33) 優先権主張国	米国 (US)		弁理士 竹内 茂雄
(31) 優先権主張番号	13/551, 590	(74) 代理人	100118902
(32) 優先日	平成24年7月17日 (2012. 7. 17)		弁理士 山本 修
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 遺伝的変異を検出するためのシステムおよび方法

(57) 【特許請求の範囲】

【請求項 1】

複数の標的ポリヌクレオチドを配列決定するための装置を生成する方法であって、
 (a) 反応表面を有する固体支持体を提供することと、
 (b) 前記固体支持体に複数のオリゴヌクレオチドを結合することと、を含み、前記複数のオリゴヌクレオチドが、

(i) 複数の異なる第1のオリゴヌクレオチドであって、配列 A および配列 B を含み、配列 A が、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列 B が、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端に存在し、原因となる遺伝的変異体を含む配列または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、複数の異なる第1のオリゴヌクレオチドと、

(i i) 配列 A をそれぞれの3'末端に含む、複数の第2のオリゴヌクレオチドと、

(i i i) 配列 C をそれぞれの3'末端に含む、配列 C が、複数の異なる標的ポリヌクレオチドにより共有される配列と同じである、複数の第3のオリゴヌクレオチドと、を含み、

配列 A、B、および C が、異なる配列であり、それぞれ5個以上のヌクレオチドを含み、

ここで、1つ以上の前記複数のオリゴヌクレオチドが、1つ以上の保護基を含むものであり、

10

20

前記原因となる遺伝的変異体が、疾患または形質と関連する統計的、生物学的、および/または機能的証明が存在する遺伝的変異体である、前記方法。

【請求項 2】

配列 A、B、および C が、互いに 90% 未満の配列同一性を有する、請求項 1 に記載の方法。

【請求項 3】

前記複数のオリゴヌクレオチドが、反応部分を含み、その結果、前記反応表面と前記反応部分との間の反応が、前記複数のオリゴヌクレオチドを前記固体支持体に結合するようになる、請求項 1 に記載の方法。

【請求項 4】

前記複数の第 1 のオリゴヌクレオチドが、それぞれが異なる配列 B を含む、少なくとも約 100 個の異なる第 1 のオリゴヌクレオチドを含む、請求項 1 に記載の方法。

【請求項 5】

前記固体支持体が、フローセルのチャンネルである、請求項 1 に記載の方法。

【請求項 6】

前記複数の第 2 のオリゴヌクレオチドの量が、前記複数の第 1 のオリゴヌクレオチドの量より少なくとも 1,000 倍高く、前記複数の第 2 のオリゴヌクレオチドの量および前記複数の第 3 のオリゴヌクレオチドの量が、1 対 1 の比である、請求項 1 に記載の方法。

【請求項 7】

前記複数の第 2 のオリゴヌクレオチドの量が、前記複数の第 1 のオリゴヌクレオチドの量より少なくとも 10,000 倍 高い、請求項 6 に記載の方法。

【請求項 8】

前記複数の異なる第 1 のオリゴヌクレオチドが、配列 A および配列 B を含む追加の第 1 のオリゴヌクレオチドをさらに含み、配列 B が、それぞれの異なる追加の第 1 のオリゴヌクレオチドに対して異なり、それぞれの追加の第 1 のオリゴヌクレオチドの 3' 末端にあり、非対象配列を含む配列、または非対象配列の 200 ヌクレオチド以内にある配列に相補性である、請求項 1 に記載の方法。

【請求項 9】

複数の標的ポリヌクレオチドを配列決定する方法であって、請求項 1 の方法に従い生成された装置を標的ポリヌクレオチドおよび非標的ポリヌクレオチドを含む試料に曝露することを含み、配列決定データが、配列決定非標的ゲノム配列と比べて標的ゲノム配列に対して配列決定強化される、方法。

【発明の詳細な説明】

【背景技術】

【0001】

次世代配列決定 (NGS) は、日単位で測定されるターンアラウンドタイムでの小規模の安価なゲノム配列決定を許容する。しかしながら、NGS が一般に行われ、理解されている通り、ゲノムの全ての領域は、大まかに等しい確率で配列決定され、つまり、可能性のある突然変異を解釈するために機能が十分によく理解されている、比較的低い割合の領域から配列情報を収集するように、大量のゲノム配列が収集および破棄される。一般に、完全ゲノム試料から興味のある領域のみを精製することは、配列決定とは別のステップとして行われる。それは、通常、当該技術分野の現状において何日もかかる低効率プロセスである。

【発明の概要】

【発明が解決しようとする課題】

【0002】

直接標的配列決定 (DTS) は、Illumina, Inc. により用いられる標準配列決定プロトコルに対する修正であり、配列決定基質 (すなわち、フローセル) が同様にゲノム配列捕捉基質になるのを許容する。典型的な次世代配列決定プロトコルの通常フローに別の器具を追加することなく、DTS プロトコルは、特別に調製されたライブラリー

10

20

30

40

50

から gDNA を捕捉するように配列決定表面を修飾する。次に、捕捉されたライブラリーは、通常の gDNA ライブラリーと同様に配列決定される。しかしながら、以前の提案に従う配列決定基質の修飾および付随するライブラリー調製は、非効率性をもたらし、信頼性および再現性を低減して、価値のある試料を浪費する。したがって、DTS プロセスに対する改善が望ましい。

【課題を解決するための手段】

【0003】

発明の概要

一態様において、本発明は、複数の標的ポリヌクレオチドを配列決定するための装置、および装置を生成する方法を提供する。一実施形態において、この方法は、(a) 反応表面を有する固体支持体を提供することと、(b) その固体支持体に複数のオリゴヌクレオチドを結合することと、を含む。いくつかの実施形態において、複数のオリゴヌクレオチドは、(i) 複数の異なる第1のオリゴヌクレオチドであって、配列Aおよび配列Bを含み、配列Aは、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bは、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にあり、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、複数の異なる第1のオリゴヌクレオチドと、(ii) 配列Aをそれぞれの3'末端に含む複数の第2のオリゴヌクレオチドと、(iii) 配列Cをそれぞれの3'末端に含む複数の第3のオリゴヌクレオチドと、を含み、配列Cは、複数の異なる標的ポリヌクレオチドにより共有される配列と同じである。いくつかの実施形態において、A、B、およびCは、異なる配列であり、それぞれ5個以上のヌクレオチドを含む。

【0004】

いくつかの実施形態において、配列A、B、およびCは、互いに90%未満の配列同一性を有する。いくつかの実施形態において、複数のオリゴヌクレオチドは、反応部分を含み、その結果、反応表面と反応部分との間の反応が、複数のオリゴヌクレオチドを固体支持体に結合するようになる。いくつかの実施形態において、複数の第1のオリゴヌクレオチドは、それぞれが異なる配列Bを含む、少なくとも約100個の異なる第1のオリゴヌクレオチドを含む。いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列Bは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む。いくつかの実施形態において、固体支持体は、フローセルのチャネルである。いくつかの実施形態において、反応表面は、アクリルアミド、N-(5-ブロモアセトアミジルペンチル)アクリルアミド、テトラメチルエチレンジアミン、および過硫酸カリウムを含む重合混合物から生成され得る官能化ポリアクリルアミドを含む。いくつかの実施形態において、複数の第2のオリゴヌクレオチドの量は、複数の第1のオリゴヌクレオチドの量より少なくとも約1000倍または10000倍高く、複数の第2のオリゴヌクレオチドの量および複数の第3のオリゴヌクレオチドの量は、約1対1の比である。いくつかの実施形態において、第1のオリゴヌクレオチドのそれぞれは、固体支持体に約50pMの濃度で付加される。いくつかの実施形態において、複数の第2のオリゴヌクレオチドおよび複数の第3のオリゴヌクレオチドの濃度は、約500nMである。いくつかの実施形態において、本発明は、複数の標的ポリヌクレオチドを配列決定する方法を提供し、この方法は、本発明の方法に従い生成された装置を、標的ポリヌクレオチドおよび非標的ポリヌクレオチドを含む試料に曝露することを含み、非標的ゲノム配列と比較して、標的ゲノム配列に対して配列決定データが強化される。いくつかの実施形態において、複数の異なる第1のオリゴヌクレオチドは、配列Aおよび配列Bを含む追加の第1のオリゴヌクレオチドをさらに含み、配列Bは、それぞれの異なる追加の第1のオリゴヌクレオチドに対して異なり、それぞれの追加の第1のオリゴヌクレオチドの3'末端にあり、非対象配列を含む配列、または非対象配列の200ヌクレオチド以内にある配列に相補性である。

【0005】

一態様において、本発明は、試料中の複数の標的ポリヌクレオチドを配列決定するための方法を提供する。一実施形態において、この方法は、(a)断片化ポリヌクレオチドを生成するように、標的ポリヌクレオチドを断片化することと、(b)適合したポリヌクレオチドの両端で相補性配列Dにハイブリダイズされた配列Dを含む、適合したポリヌクレオチドを生成するように、それぞれが配列Dを含むアダプターオリゴヌクレオチドを、断片化ポリヌクレオチドに連結することと、任意に配列Dが、標的ポリヌクレオチド3'末端の伸長により生成される、連結することと、(c)適合したポリヌクレオチドを、配列C、配列D、および試料と関連付けられたバーコードを含む増幅プライマーを使用して増幅することと、配列Dが、増幅プライマーの3'末端に位置付けられる、増幅することと、(d)増幅した標的ポリヌクレオチドを、固体表面に結合した複数の異なる第1のオリゴヌクレオチドにハイブリダイズすることと、(e)固体表面上で架橋増幅を行うことと、(f)ステップ(e)からの複数のポリヌクレオチドを配列決定することと、を含む。固体表面は、本明細書に記載される複数のオリゴヌクレオチドを含んでよく、本明細書に記載され、任意に本明細書に記載される方法に従い生成される装置を含む。いくつかの実施形態において、固体表面は、(i)複数の異なる第1のオリゴヌクレオチドであって、配列Aおよび配列Bを含み、配列Aが、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bが、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にあり、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、複数の異なる第1のオリゴヌクレオチドと、(ii)配列Aをそれぞれの3'末端に含む複数の第2のオリゴヌクレオチドと、(iii)配列Cをそれぞれの3'末端に含む複数の第3のオリゴヌクレオチドと、を含む。いくつかの実施形態において、配列A、B、およびCは、異なる配列であり、それぞれ5個以上のヌクレオチドを含む。

【0006】

いくつかの実施形態において、この方法は、ステップ(d)の前に、第2の増幅ステップをさらに含み、増幅したポリヌクレオチドは、ステップ(c)において標的ポリヌクレオチドに付加された1つ以上の配列の少なくとも一部分に相補性である配列を含む3'末端を有する第2の増幅プライマーを使用して増幅される。いくつかの実施形態において、配列A、B、およびCは、互いに90%未満の配列同一性を有する。いくつかの実施形態において、複数の第1のオリゴヌクレオチドは、それぞれが異なる配列Bを含む、少なくとも約100個の異なる第1のオリゴヌクレオチドを含む。いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列Bは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む。いくつかの実施形態において、それぞれのバーコードは、少なくとも3つのヌクレオチド位置において、2つ以上の試料のプール中のバーコードと1つおきに異なる。いくつかの実施形態において、試料は、4つのヌクレオチド塩基A、G、C、およびTの全てが、プール中のそれぞれのバーコードに沿ってあらゆる位置でほぼ均一に表されるようにプールされる。いくつかの実施形態において、1つ以上のバーコードは、AGGTC A、CAGCAG、ACTGCT、TAACGG、GGATTA、AACCTG、GCCGTT、CGTTGA、GTAAC C、CTTAAC、TGCTAA、GATCCG、CCAGGT、TTCAGC、ATGATC、およびTCGGATからなる群から選択される。いくつかの実施形態において、バーコードは、配列Cと配列Dとの間に位置する。いくつかの実施形態において、この方法は、標的ポリヌクレオチドが誘導される試料を、バーコード配列に基づいて特定するステップをさらに含む。いくつかの実施形態において、断片化ポリヌクレオチドは、約200~約1000塩基対の長さの中央値を有する。いくつかの実施形態において、ステップ(f)は、(i)バーコードから3'にある位置にハイブリダイズする第1の配列決定プライマーの伸長により配列決定することと、次いで(ii)バーコードから5'にある位置にハイブリダイズする第2の配列決定プライマーの伸長により配列決定することと、を含む。いくつかの実施形態において、固体支持体は、フローセルのチャンネルである。いく

10

20

30

40

50

つかの実施形態において、ステップ (b) および (c) は、液体ハンドラー (例えば、BioMek FXP) のような自動システムにより行われる。いくつかの実施形態において、ステップ (d) は、例えば、cBotマシンを備えるシステムのような自動システムにより行われる。いくつかの実施形態において、ステップ (d) を行う自動システムは、ステップ (e) も行う。いくつかの実施形態において、配列決定データは、少なくとも約 100 個の異なる標的ポリヌクレオチドに対して生成される。いくつかの実施形態において、ステップ (d) は、単一フローセルにおいて少なくとも約 10 μ g の DNA を利用する。いくつかの実施形態において、この方法は、複数の試料上で並行して行われる。いくつかの実施形態において、ステップ (c) は、複数の試料のそれぞれに対して四重に行われる。いくつかの実施形態において、DNA の量は、ステップ (a)、(b)、および (c) のうちの 1 つ以上の完了時に測定される。いくつかの実施形態において、ステップ (a)、(b)、および (c) のうちの 1 つ以上は、次のステップで使用されるそのステップの最後に残る DNA の量の最小閾値、例えば、それぞれ 1 μ g、0.8 μ g、13 μ g を有する。いくつかの実施形態において、配列決定データは、単一反応において少なくとも約 10⁸ 個の標的配列に対して生成される。いくつかの実施形態において、配列決定データは、単一反応において 1 試料当たり約 10⁷ 個未満の標的配列に対して生成される。いくつかの実施形態において、1 つ以上の原因となる遺伝的変異体の存在または非存在は、少なくとも約 90% の精度で決定される。いくつかの実施形態において、複数の異なる第 1 のオリゴヌクレオチドは、配列 A および配列 B を含む追加の第 1 のオリゴヌクレオチドをさらに含み、配列 B は、それぞれの異なる追加の第 1 のオリゴヌクレオチドに対して異なり、それぞれの追加の第 1 のオリゴヌクレオチドの 3' 末端にあり、非対象配列を含む配列、または非対象配列の 200 ヌクレオチド以内にある配列に相補性である。

【0007】

一態様において、本発明は、試料中の複数の異なる標的ポリヌクレオチドを濃縮する方法を提供する。いくつかの実施形態において、この方法は、(a) 配列 Y を含むアダプターオリゴヌクレオチドを、標的ポリヌクレオチドのそれぞれに連結することと、(b) 複数の異なるオリゴヌクレオチドプライマーを、適合した標的ポリヌクレオチドにハイブリダイズすることとであって、それぞれのオリゴヌクレオチドプライマーが、配列 Z および配列 W を含み、配列 Z が、全てのオリゴヌクレオチドプライマーの中で共通であり、さらに配列 W が、それぞれの異なるオリゴヌクレオチドプライマーに対して異なり、それぞれのオリゴヌクレオチドプライマーの 3' 末端に位置付けられ、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の 200 ヌクレオチド以内にある配列に相補性である、ハイブリダイズすることと、(c) 伸長反応において、配列 Z および配列 Y を含む伸長したプライマーを生成するように、オリゴヌクレオチドプライマーを適合した標的ポリヌクレオチドに沿って伸長することとであって、配列 Y が配列 Y に相補性である、伸長することと、(d) 伸長したプライマーを、(i) 配列 V および配列 Z を含む第 1 の増幅プライマーであって、配列 Z が第 1 の増幅プライマーの 3' 末端に位置付けられる、第 1 の増幅プライマー、ならびに (ii) 配列 X および配列 Y を含む第 2 の増幅プライマーであって、配列 Y が第 2 の増幅プライマーの 3' 末端に位置付けられる、第 2 の増幅プライマーの対を使用して、指数関数的に増幅することと、を含む。いくつかの実施形態において、配列 W、Y、および Z は、異なる配列であり、それぞれ 5 個以上のヌクレオチドを含む。それぞれのオリゴヌクレオチドプライマーは、第 1 の結合パートナーを含んでも含まなくてもよい。いくつかの実施形態において、この方法は、ステップ (d) の前に、伸長したプライマーを、第 1 の結合パートナーに結合する第 2 の結合パートナーを含む固体表面に曝露することと、それにより伸長したプライマーを、伸長反応の 1 つ以上の成分から取り出して精製することと、をさらに含む。いくつかの実施形態において、この方法は、精製ステップを含まない。

【0008】

いくつかの実施形態において、複数のオリゴヌクレオチドプライマーは、それぞれが異なる配列 W を含む、少なくとも約 100 個の異なるオリゴヌクレオチドプライマーを含む

10

20

30

40

50

。いくつかの実施形態において、複数のオリゴヌクレオチドプライマーのうちの1つ以上の配列Wは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む。いくつかの実施形態において、標的ポリヌクレオチドは、断片化ポリヌクレオチドを含む。いくつかの実施形態において、断片化ポリヌクレオチドは、約200~約1000塩基対の長さの中央値を有する。いくつかの実施形態において、断片化ポリヌクレオチドは、ステップ(a)の前に平滑末端を生成するか、または明確なオーバーハング、例えば、アデニンからなるオーバーハングを有するように処理される。いくつかの実施形態において、第1の結合パートナーおよび第2の結合パートナーは、ストレプトアビジンおよびビオチン等の結合対のメンバーである。いくつかの実施形態において、固体表面は、磁場に反応するビーズ等のビーズである。いくつかの実施形態において、精製ステップは、ビーズを精製するための磁場の印加を含む。いくつかの実施形態において、伸長したプライマーは、標的ポリヌクレオチドから取り出して精製される。いくつかの実施形態において、この方法は、ステップ(d)の生成物を配列決定することをさらに含む。いくつかの実施形態において、配列決定は、二本鎖架橋ポリヌクレオチドを生成するように、ステップ(d)の生成物を、固体支持体に結合した結合オリゴヌクレオチドとの架橋増幅により増幅することと、架橋ポリヌクレオチドの一本鎖を、結合オリゴヌクレオチド中の切断部位で切断することと、固体支持体に結合した標的配列を含む遊離一本鎖ポリヌクレオチドを生成するように、切断された架橋ポリヌクレオチドを変性させることと、ステップ(a)、(c)、または(d)のうちの1つ以上の間に付加された1つ以上の配列の少なくとも一部分にハイブリダイズされた配列決定プライマーを伸長することにより、標的配列を配列決定することと、を含む。いくつかの実施形態において、配列決定は、結合テンプレートを生成するように、ステップ(d)の生成物を、固体支持体上の結合プライマーの伸長により増幅することと、配列決定プライマーを結合テンプレートにハイブリダイズすることと、配列決定プライマーを伸長することと、配列決定プライマーの伸長により付加されたヌクレオチドを特定することと、を含む。いくつかの実施形態において、複数の異なるオリゴヌクレオチドプライマーは、配列Zおよび配列Wを含む追加のオリゴヌクレオチドプライマーをさらに含み、配列Wは、それぞれの異なる追加のオリゴヌクレオチドプライマーに対して異なり、それぞれの追加のオリゴヌクレオチドプライマーの3'末端にあり、非対象配列を含む配列、または非対象配列の200ヌクレオチド以内にある配列に相補性である。

【0009】

一態様において、本発明は、試料中の複数の異なる標的ポリヌクレオチドを濃縮する方法を提供する。いくつかの実施形態において、この方法は、(a)複数の異なるオリゴヌクレオチドプライマーを標的ポリヌクレオチドにハイブリダイズすることであって、それぞれのオリゴヌクレオチドプライマーが、配列Zおよび配列Wを含み、配列Zが、全てのオリゴヌクレオチドプライマーの中で共通であり、さらに配列Wが、それぞれの異なるオリゴヌクレオチドプライマーに対して異なり、それぞれのオリゴヌクレオチドプライマーの3'末端に位置付けられ、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、ハイブリダイズすることと、(b)伸長反応において、伸長したプライマーを生成するように、オリゴヌクレオチドプライマーを標的ポリヌクレオチドに沿って伸長することと、(c)アダプターオリゴヌクレオチドをそれぞれの伸長したプライマーに連結することであって、アダプターオリゴヌクレオチドが、配列Yを含み、さらに配列Yが、配列Yの相補体である、連結することと、(d)伸長したプライマーを、(i)配列Vおよび配列Zを含む第1の増幅プライマーであって、配列Zが第1の増幅プライマーの3'末端に位置付けられる、第1の増幅プライマー、ならびに(ii)配列Xおよび配列Yを含む第2の増幅プライマーであって、配列Yが第2の増幅プライマーの3'末端に位置付けられる、第2の増幅プライマーの対を使用して、指数関数的に増幅することと、を含む。いくつかの実施形態において、配列W、Y、およびZは、異なる配列であり、それぞれ5個以上のヌクレオチドを含む。それぞれのオリゴヌクレオチドプライマーは、第1の結合パートナーを含んでも含まな

10

20

30

40

50

くてもよい。いくつかの実施形態において、この方法は、ステップ(d)の前に、伸長したプライマーを、第1の結合パートナーに結合する第2の結合パートナーを含む固体表面に曝露して、それにより伸長したプライマーを、伸長反応の1つ以上の成分から取り出して精製することをさらに含む。いくつかの実施形態において、この方法は、精製ステップを含まない。

【0010】

いくつかの実施形態において、複数のオリゴヌクレオチドプライマーは、それぞれが異なる配列Wを含む、少なくとも約100個の異なるオリゴヌクレオチドプライマーを含む。いくつかの実施形態において、複数のオリゴヌクレオチドプライマーのうちの1つ以上の配列Wは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む。いくつかの実施形態において、標的ポリヌクレオチドは、断片化ポリヌクレオチドを含む。いくつかの実施形態において、断片化ポリヌクレオチドは、約200~約1000塩基対の長さの中央値を有する。いくつかの実施形態において、ステップ(b)は、伸長したプライマー、およびそれらがハイブリダイズされる標的ポリヌクレオチドを、平滑末端を生成するか、またはステップ(c)の前に明確なオーバーハング、例えば、アデニンからなるオーバーハングを有するように処理することをさらに含む。いくつかの実施形態において、第1の結合パートナーおよび第2の結合パートナーは、ストレプトアビジンおよびビオチン等の結合対のメンバーである。いくつかの実施形態において、固体表面は、磁場に反応するビーズ等のビーズである。いくつかの実施形態において、精製ステップは、ビーズを精製するように磁場の印加を含む。いくつかの実施形態において、伸長したプライマーは、標的ポリヌクレオチドから取り出して精製される。いくつかの実施形態において、この方法は、ステップ(d)の生成物を配列決定することをさらに含む。いくつかの実施形態において、配列決定は、二本鎖架橋ポリヌクレオチドを生成するように、ステップ(d)の生成物を、固体支持体に結合した結合オリゴヌクレオチドとの架橋増幅により増幅することと、架橋ポリヌクレオチドの一本鎖を、結合オリゴヌクレオチド中の切断部位で切断することと、固体支持体に結合した標的配列を含む遊離一本鎖ポリヌクレオチドを生成するように、切断された架橋ポリヌクレオチドを変性させることと、ステップ(b)、(c)、または(d)のうちの1つ以上の間に付加された1つ以上の配列の少なくとも一部分にハイブリダイズされた配列決定プライマーを伸長することにより、標的配列を配列決定することと、を含む。いくつかの実施形態において、配列決定は、結合テンプレートを生成するように、ステップ(d)の生成物を、固体支持体上の結合プライマーの伸長により増幅することと、配列決定プライマーを結合テンプレートにハイブリダイズすることと、配列決定プライマーを伸長することと、配列決定プライマーの伸長により付加されたヌクレオチドを特定することと、を含む。いくつかの実施形態において、複数の異なるオリゴヌクレオチドプライマーは、配列Zおよび配列Wを含む追加のオリゴヌクレオチドプライマーをさらに含み、配列Wは、それぞれの異なる追加のオリゴヌクレオチドプライマーに対して異なり、それぞれの追加のオリゴヌクレオチドプライマーの3末端にあり、非対象配列を含む配列、または非対象配列の200ヌクレオチド以内にある配列に相補性である。

【0011】

一態様において、本発明は、対象のゲノム中の遺伝的変異を検出する方法を提供する。いくつかの実施形態において、この方法は、(a)ポリヌクレオチドの複数のクラスタを提供することであって、(i)それぞれのクラスタが、支持体に結合した核酸二本鎖の複数コピーを含み、(ii)クラスタ中のそれぞれの二本鎖が、配列A-B-G-D-Cを5から3に含む第1の分子と、配列C-D-G-B-Aを5から3に含む第2の分子と、を含み、(iii)配列Aが、配列Aに相補性であり、配列Bが、配列Bに相補性であり、配列Cが、配列Cに相補性であり、配列Dが、配列Dに相補性であり、配列Gが、配列Gに相補性であり、(iv)配列Gが、対象からの標的ポリヌクレオチド配列の一部であり、複数のクラスタのそれぞれに対して異なり、(v)配列Bが、対応する標的ポリヌクレオチド配列中の配列Gに関して5に位置する、提

供することと、(b)それぞれのクラスタに対してR1配列を生成するように、配列Gを、配列Dを含む第1のプライマーの伸長により配列決定することと、(c)それぞれのクラスタに対してR2配列を生成するように、配列Bを、配列Aを含む第2のプライマーの伸長により配列決定することと、(d)全てのR1配列を第1の参照配列に整列させるように、第1のアルゴリズムを使用して第1のアラインメントを行うことと、(e)第1の参照配列に関して挿入または欠失を含む可能性が高いとして第1のアラインメントにおいて特定されたR1配列を局所的に整列させ、それぞれの挿入または欠失の単一コンセンサスアラインメントを生成するように、第2のアルゴリズムを使用して第2のアラインメントを行うことと、(f)全てのR2配列を第2の参照配列に整列させることにより、R2アラインメントを行うことと、(g)ステップ(d)~(f)により特定された配列変異を特定するレポートを受信者に伝送することと、を含む。

10

【0012】

いくつかの実施形態において、第1の参照配列は、参照ゲノムを含む。いくつかの実施形態において、第2の参照配列は、あらゆる異なる標的ポリヌクレオチドのあらゆる配列Bからなる。いくつかの実施形態において、R2配列は、R1配列から独立して整列される。いくつかの実施形態において、この方法は、同じクラスタのR2配列が整列する第1の参照配列中の第2の位置から10,000塩基対を超えて離れた、第1の参照配列中の第1の位置に整列するR1配列を破棄することをさらに含む。いくつかの実施形態において、欠失されるR1配列の部分が、あるクラスタの配列Bの少なくとも一部分と同一であり、配列Gが、そのクラスタのR1配列より短いとき、そのクラスタのR1配列の一部を欠失させることをさらに含む。いくつかの実施形態において、この方法は、欠失されるR1配列の一部が、任意の配列Bの少なくとも一部分と同一であり、その部分が、R1の5もしくは3ヌクレオチドのいずれかを含み、(i)いかなるR2配列も、そのクラスタに対して生成されなかったか、または(ii)生成されたR2配列が、任意の配列Bと同一でないかのいずれかであるとき、クラスタのR1配列の一部を欠失させることをさらに含む。いくつかの実施形態において、第1のアルゴリズムを使用するシステムを用いて第1のアラインメントを行うことは、そのシステムが第2のアルゴリズムを使用して第1のアラインメントを行う場合にかかる、および/または使用されるより少ない時間で、および/または少ないシステムメモリを使用して全てのR1読み取り値を整列させる。いくつかの実施形態において、第1のアルゴリズムは、パローズ-ホイラー変換に基づく。いくつかの実施形態において、第2のアルゴリズムは、スミス-ウォーターマンアルゴリズムまたはハッシュ関数に基づく。いくつかの実施形態において、R1およびR2配列は、少なくとも100個の異なる標的ポリヌクレオチドに対して生成される。いくつかの実施形態において、配列A、B、C、およびDは、少なくとも5個のヌクレオチド長である。いくつかの実施形態において、あらゆるクラスタの配列Gは、1~1000ヌクレオチド長である。いくつかの実施形態において、複数のクラスタのそれぞれのプローブ配列Bは、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である。いくつかの実施形態において、クラスタのうちの1つ以上の配列Bは、配列番号22~121からなる群から選択される配列を含む。いくつかの実施形態において、R1配列は、単一反応において少なくとも約 10^8 個のクラスタに対して生成される。いくつかの実施形態において、1つ以上の原因となる遺伝的変異体の存在、非存在、または対立遺伝子比は、少なくとも約90%の精度で決定される。いくつかの実施形態において、コンセンサス配列は、標的ポリヌクレオチド中の挿入、欠失、または挿入および欠失を、少なくとも約90%の精度で特定する。いくつかの実施形態において、複数のクラスタのそれぞれのプローブ配列Bは、非対象配列、または非対象配列の200ヌクレオチド以内にある配列に相補性である。いくつかの実施形態において、1つ以上の非対象配列の存在または非存在は、少なくとも約90%の精度で決定される。いくつかの実施形態において、この方法は、対象のR1配列に基づいて複数の確率を計算することと、それらの確率をレポートに含めることと、をさらに含み、それぞれの確率は、対象または対象の子孫が疾患または形質を有するか、または発症する確率で

20

30

40

50

ある。

【 0 0 1 3 】

いくつかの実施形態において、それぞれの第 1 の分子は、バーコード配列を含む。いくつかの実施形態において、それぞれのバーコードは、並行して分析される複数の異なるバーコード中のバーコードと 1 つおきに異なる。いくつかの実施形態において、バーコード配列は、単一反応において配列決定された試料のプール中の単一試料と関連付けられる。いくつかの実施形態において、複数のバーコード配列のそれぞれは、単一反応において配列決定された試料のプール中の単一試料と一意に関連付けられる。いくつかの実施形態において、バーコード配列は、配列 D から 5 に位置する。いくつかの実施形態において、この方法は、第 3 のプライマーを配列 C にハイブリダイズすることと、それぞれのクラスタに対してバーコード配列を生成するように、バーコード配列を、第 3 のプライマーの伸長により配列決定することと、をさらに含む。いくつかの実施形態において、この方法は、バーコード配列に基づいて、クラスタから配列を分類することをさらに含む。いくつかの実施形態において、この方法は、バーコード配列分類内で同じ配列およびアラインメントを有する複数の R 1 配列を、そのうちの 1 つを除いて全てを破棄することをさらに含む。

【 0 0 1 4 】

一態様において、本発明は、対象のゲノム中の遺伝的変異を検出する方法を提供する。いくつかの実施形態において、この方法は、(a) ポリヌクレオチドの複数のクラスタに配列決定データを提供することとあって、(i) それぞれのクラスタが、支持体に結合した核酸二本鎖の複数コピーを含み、(i i) クラスタ中のそれぞれの二本鎖が、配列 A - B - G - D - C を 5 から 3 に含む第 1 の分子と、配列 C - D - G - B - A を 5 から 3 に含む第 2 の分子と、を含み、(i i i) 配列 A が、配列 A に相補性であり、配列 B が、配列 B に相補性であり、配列 C が、配列 C に相補性であり、配列 D が、配列 D に相補性であり、配列 G が、配列 G に相補性であり、(i v) 配列 G が、対象からの標的ポリヌクレオチド配列の一部であり、複数のクラスタのそれぞれに対して異なり、(v) 配列 B が、対応する標的ポリヌクレオチド配列中の配列 G に関して 5 に位置し、(v i i i) 配列決定データが、配列 D を含む第 1 のプライマーの伸長により生成された R 1 配列を含み、(v i) 配列決定データが、配列 A を含む第 2 のプライマーの伸長により生成された R 2 配列を含む、提供することと、(b) 全ての R 1 配列を第 1 の参照配列に整列させるように、第 1 のアルゴリズムを使用して第 1 のアラインメントを行うことと、(c) 第 1 の参照配列に関して挿入または欠失を含む可能性が高いとして、該第 1 のアラインメントにおいて特定された R 1 配列を局所的に整列させ、それぞれの挿入または欠失の単一コンセンサスアラインメントを生成するように、第 2 のアルゴリズムを使用して第 2 のアラインメントを行うことと、(d) 全ての R 2 配列を第 2 の参照配列に整列させることにより、R 2 アラインメントを行うことと、(e) ステップ (b) ~ (d) により特定された配列変異を特定するレポートを受信者に伝送することと、を含む。

【 0 0 1 5 】

いくつかの実施形態において、第 1 の参照配列は、参照ゲノムを含む。いくつかの実施形態において、第 2 の参照配列は、あらゆる異なる標的ポリヌクレオチドのあらゆる配列 B からなる。いくつかの実施形態において、R 2 配列は、R 1 配列から独立して整列される。いくつかの実施形態において、この方法は、1 0 , 0 0 0 より多くの塩基対である第 1 の参照配列中の第 1 の位置に整列する R 1 配列を、同じクラスタの R 2 配列が整列する第 1 の参照配列中の第 2 の位置から破棄することをさらに含む。いくつかの実施形態において、この方法は、欠失される R 1 配列の一部が、あるクラスタの配列 B の少なくとも一部分と同一であり、配列 G が、そのクラスタの R 1 配列より短いとき、そのクラスタの R 1 配列の一部を欠失させることをさらに含む。いくつかの実施形態において、この方法は、欠失される R 1 配列の一部が、任意の配列 B の少なくとも一部分と同一であり、その部分が、R 1 の 5 もしくは 3 ヌクレオチドのいずれかを含み、(i) いかな

る R 2 配列も、クラスタに対して生成されなかったか、または (i i) 生成された R 2 配列が、任意の配列 B と同一でないかのいずれかであるとき、クラスタの R 1 配列の一部を欠失させることをさらに含む。いくつかの実施形態において、第 1 のアルゴリズムを使用するシステムを用いて第 1 のアラインメントを行うことは、そのシステムが第 2 のアルゴリズムを使用して第 1 のアラインメントを行う場合にかかる、および / または使用されるより少ない時間で、および / または少ないシステムメモリを使用して全ての R 1 読み取り値を整列させる。いくつかの実施形態において、第 1 のアルゴリズムは、パローズ - ホイラー変換に基づく。いくつかの実施形態において、第 2 のアルゴリズムは、スミス - ウォーターマンアルゴリズムまたはハッシュ関数に基づく。いくつかの実施形態において、配列決定データは、少なくとも 100 個の異なる標的ポリヌクレオチドの R 1 および R 2 配列を含む。いくつかの実施形態において、配列 A、B、C、および D は、少なくとも 5 個のヌクレオチド長である。いくつかの実施形態において、あらゆるクラスタの配列 G は、1 ~ 1000 ヌクレオチド長である。いくつかの実施形態において、複数のクラスタのそれぞれのプローブ配列 B は、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の 200 ヌクレオチド以内にある配列に相補性である。いくつかの実施形態において、クラスタのうちの 1 つ以上の配列 B は、配列番号 22 ~ 121 からなる群から選択される配列を含む。いくつかの実施形態において、配列決定データは、単一反応からの少なくとも約 10⁸ 個の R 1 配列を含む。いくつかの実施形態において、1 つ以上の原因となる遺伝的変異体の存在、非存在、または対立遺伝子比は、少なくとも約 90 % の精度で決定される。いくつかの実施形態において、コンセンサス配列は、標的ポリヌクレオチド中の挿入、欠失、または挿入および欠失を、少なくとも約 90 % の精度で特定する。いくつかの実施形態において、複数のクラスタのそれぞれのプローブ配列 B は、非対象配列を含む配列、または非対象配列の 200 ヌクレオチド以内にある配列に相補性である。いくつかの実施形態において、1 つ以上の非対象配列の存在または非存在は、少なくとも約 90 % の精度で決定される。いくつかの実施形態において、この方法は、R 1 配列に基づいて複数の確率を計算することと、それらの確率をレポートに含めることと、をさらに含み、それぞれの確率は、対象または対象の子孫が疾患または形質を有するか、または発症する確率である。

【0016】

いくつかの実施形態において、それぞれの第 1 の分子は、バーコード配列を含む。いくつかの実施形態において、それぞれのバーコードは、並行して分析された複数の異なるバーコード中のバーコードと 1 つおきに異なる。いくつかの実施形態において、バーコード配列は、単一反応において配列決定された試料のプール中の単一試料と関連付けられ、配列決定データに表される。いくつかの実施形態において、複数のバーコード配列のそれぞれは、単一反応において配列決定された試料のプール中の単一試料と一意に関連付けられる。いくつかの実施形態において、バーコード配列は、配列 D から 5 に位置する。いくつかの実施形態において、配列決定データは、配列 C を含む第 3 のプライマーの伸長により生成されたそれぞれのクラスタのバーコード配列をさらに含む。いくつかの実施形態において、この方法は、バーコード配列に基づいて、クラスタからの配列を分類することをさらに含む。いくつかの実施形態において、この方法は、バーコード配列分類内で同じ配列およびアラインメントを有する複数の R 1 配列を、そのうちの 1 つを除く全てを破棄することをさらに含む。

【0017】

参照による組み込み

本明細書において言及する全ての刊行物、特許、および特許出願は、それぞれ個別の刊行物、特許、または特許出願が、参照により組み込まれることが具体的かつ個別に示されたのと同じ程度で参照により本明細書に組み込まれる。

【0018】

本発明の新規特徴が、添付の特許請求の範囲において詳細に説明される。本発明の特徴および利点に関するより良い理解は、本発明の原理が利用される例証の実施形態を説明す

10

20

30

40

50

る以下の詳細な説明、および添付の図面を参照することにより得られる。

【図面の簡単な説明】

【0019】

【図1】結合したオリゴヌクレオチドを含む例示の固体支持体の一部分、および標的ポリヌクレオチドを増幅する例示の架橋増幅プロセスの第1のステップを示す。

【図2】本発明の実施形態に従う例示の捕捉および増幅プロセスを示す。

【図3-1】例示の原因となる遺伝的変異体の表を提供する。

【図3-2】例示の原因となる遺伝的変異体の表を提供する。

【図3-3】例示の原因となる遺伝的変異体の表を提供する。

【図3-4】例示の原因となる遺伝的変異体の表を提供する。

10

【図3-5】例示の原因となる遺伝的変異体の表を提供する。

【図3-6】例示の原因となる遺伝的変異体の表を提供する。

【図3-7】例示の原因となる遺伝的変異体の表を提供する。

【図3-8】例示の原因となる遺伝的変異体の表を提供する。

【図3-9】例示の原因となる遺伝的変異体の表を提供する。

【図3-10】例示の原因となる遺伝的変異体の表を提供する。

【図3-11】例示の原因となる遺伝的変異体の表を提供する。

【図3-12】例示の原因となる遺伝的変異体の表を提供する。

【図3-13】例示の原因となる遺伝的変異体の表を提供する。

【図3-14】例示の原因となる遺伝的変異体の表を提供する。

20

【図3-15】例示の原因となる遺伝的変異体の表を提供する。

【図3-16】例示の原因となる遺伝的変異体の表を提供する。

【図3-17】例示の原因となる遺伝的変異体の表を提供する。

【図3-18】例示の原因となる遺伝的変異体の表を提供する。

【図3-19】例示の原因となる遺伝的変異体の表を提供する。

【図3-20】例示の原因となる遺伝的変異体の表を提供する。

【図3-21】例示の原因となる遺伝的変異体の表を提供する。

【図3-22】例示の原因となる遺伝的変異体の表を提供する。

【図3-23】例示の原因となる遺伝的変異体の表を提供する。

【図3-24】例示の原因となる遺伝的変異体の表を提供する。

30

【図3-25】例示の原因となる遺伝的変異体の表を提供する。

【図3-26】例示の原因となる遺伝的変異体の表を提供する。

【図3-27】例示の原因となる遺伝的変異体の表を提供する。

【図3-28】例示の原因となる遺伝的変異体の表を提供する。

【図3-29】例示の原因となる遺伝的変異体の表を提供する。

【図3-30】例示の原因となる遺伝的変異体の表を提供する。

【図3-31】例示の原因となる遺伝的変異体の表を提供する。

【図3-32】例示の原因となる遺伝的変異体の表を提供する。

【図3-33】例示の原因となる遺伝的変異体の表を提供する。

【図3-34】例示の原因となる遺伝的変異体の表を提供する。

40

【図3-35】例示の原因となる遺伝的変異体の表を提供する。

【図3-36】例示の原因となる遺伝的変異体の表を提供する。

【図3-37】例示の原因となる遺伝的変異体の表を提供する。

【図3-38】例示の原因となる遺伝的変異体の表を提供する。

【図4-1】例示の特定標的配列に相補性である例示の配列の表を提供する。

【図4-2】例示の特定標的配列に相補性である例示の配列の表を提供する。

【図4-3】例示の特定標的配列に相補性である例示の配列の表を提供する。

【図5】本発明の実施形態に従う例示の増幅プロセスを示す。

【図6】標的増幅、架橋増幅、および配列決定の例示のプロセスを示す。

【図7】本発明の実施形態に従う例示の増幅プロセスを示す。

50

【図 8】本発明の方法において有用なコンピュータシステムの非限定的な例を示す。

【図 9 - 1】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 2】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 3】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 4】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 5】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 6】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図 9 - 7】異なる集団を区別する多数の A I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用される I D 番号である。

【図9-8】異なる集団を区別する多数のA I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用されるI D番号である。

【図9-9】異なる集団を区別する多数のA I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用されるI D番号である。

【図9-10】異なる集団を区別する多数のA I Mを提供する。これらのエントリーは、d b S N Pデータベース、すなわち米国政府により維持される遺伝的変異体のデータベース：www.ncbi.nlm.nih.gov/projects/SNP/における項目を指す。d b S N P中の精選された記録は、遺伝的変異体の配列および位置を説明する情報を含み、また、得ることができる場合には、異なる集団におけるそのような変異体の対立遺伝子の頻度を含む。r s 番号（例えば、r s 3 3 2、r s 2 5等）は、d b S N Pデータベースの一部分にインデックスを付けるために使用されるI D番号である。

【図10】配列決定データを整列させるための例示のデータ処理プロセスを示す。

【図11】配列決定データを使用してアラインメントを生成するための例示のプロセスを示す。

【図12】図12 Aおよび12 Bは、例示のアラインメントプロセスにおけるf i x _ a l i g nステップの前および後のアラインメントを示す。

【図13】図13 Aおよび13 Bは、例示の局所アラインメントステップの前および後のアラインメントを示す。

【図14】ユーザーが、希な遺伝的疾患の保因者である確率を配信する例示のプロセスを説明する。

【図15】ユーザーが、希な遺伝的疾患の保因者である確率を配信する例示のプロセスを説明する。

【図16】ユーザーが、希な遺伝的疾患の保因者である確率を配信する例示のプロセスを説明する。

【図17】ユーザーが、希な遺伝的疾患の保因者である確率を配信する例示のプロセスを説明する。

【図18】2人の仮定の親：母親H e n (J a n e D o e)および父親H e n (J o h n D o e)に対するレポート作成のための例示の入出力ステップを示す。

【図19】試料プローブ設計プロセスのステップにおいて考慮される配列領域の位置的関係を示す。

【発明を実施するための形態】

【0020】

「ポリヌクレオチド」、「ヌクレオチド」、「ヌクレオチド配列」、「核酸」、および「オリゴヌクレオチド」という用語は、同義に使用される。それらは、任意の長さのヌクレオチドのポリマー形態、デオキシリボヌクレオチドもしくはリボヌクレオチドのいずれか、またはそれらの類似体を指す。ポリヌクレオチドは、任意の3次元構造を有することができる、既知または未知の任意の機能を行うことができる。以下は、ポリヌクレオチドの非限定的な例である：遺伝子または遺伝子断片のコードまたは非コード領域、遺伝子間D N A、連鎖分析から定義される遺伝子座（複数可）、エクソン、イントロン、メッセンジャーR N A (m R N A)、転移R N A、リボソームR N A、短干渉R N A (s i R N A)

10

20

30

40

50

、短ヘアピンRNA (shRNA)、マイクロRNA (miRNA)、小核小体RNA、リボザイム、cDNA、組み換えポリヌクレオチド、分岐ポリヌクレオチド、プラスミド、ベクター、任意の配列の単離DNA、任意の配列の単離RNA、核酸プローブ、アダプター、およびプライマー。ポリヌクレオチドは、修飾ヌクレオチド、例えば、メチル化ヌクレオチドおよびヌクレオチド類似体を含み得る。存在する場合、ヌクレオチド構造への修飾は、ポリマーの組み立ての前または後に付与され得る。ヌクレオチドの配列は、非ヌクレオチド成分により中断され得る。ポリヌクレオチドは、重合後、例えば、標識成分、タグ、反応部分、または結合パートナーとの共役によりさらに修飾され得る。ポリヌクレオチド配列は、提供されるとき、特に明記しない限り5'から3'の方向で列挙される。

10

【0021】

本明細書において使用される場合、「標的ポリヌクレオチド」という用語は、本発明の1つ以上のオリゴヌクレオチドがハイブリダイズするように設計される標的配列を有する核酸分子の集団中の核酸分子またはポリヌクレオチドを指す。いくつかの実施形態において、標的配列は、試料から誘導された配列、例えば、特定のゲノム、ミトコンドリア、細菌、ウイルス、またはRNA (例えば、mRNA、miRNA、一次miRNA、またはpre-miRNA) 配列を一意に特定する。いくつかの実施形態において、標的配列は、例えば、複数の異なる標的ポリヌクレオチドにより共有される共通の配列、例えば、異なる標的ポリヌクレオチドに連結された共通のアダプター配列である。「標的ポリヌクレオチド」は、一方もしくは両方の鎖上に標的配列を含む二本鎖核酸分子、または標的配列を含む一本鎖核酸分子を指すように使用され得、核酸分子の任意の供給源、またはそれを単離もしくは生成するためのプロセスから誘導され得る。標的ポリヌクレオチドは、1つ以上 (例えば、1、2、3、4、5、6、7、8、9、10、またはそれより多く) の標的配列を含んでよく、それらは同じであっても異なってもよい。一般に、異なる標的ポリヌクレオチドは、異なる配列、例えば、1つ以上の異なるヌクレオチドまたは1つ以上の異なる標的配列を含む。

20

【0022】

「ハイブリダイゼーション」および「アニーリング」は、1つ以上のポリヌクレオチドが反応して、ヌクレオチド残基の塩基間の水素結合を介して安定化される複合体を形成する反応を指す。水素結合は、ワトソンクリック塩基対合、フーグスティーン結合によるか、または任意の他の配列特異的方法において起こり得る。この複合体は、二本鎖構造を形成する2つの鎖、多鎖複合体を形成する3つ以上の鎖、単一自己ハイブリダイジング鎖、またはこれらの任意の組み合わせを含んでよい。ハイブリダイゼーション反応は、より広範囲のプロセスにおけるステップ、例えば、PCRの開始、またはリボザイムによるポリヌクレオチドの酵素切断を構成することができる。第2の配列のヌクレオチド残基の塩基との水素結合を介して安定化され得る1の配列は、第2の配列に「ハイブリダイズ可能」とあると言われる。そのような場合、第2の配列が、第1の配列にハイブリダイズ可能であるということもできる。

30

【0023】

一般に、所与の配列の「相補体」は、所与の配列に完全に相補性であり、ハイブリダイズ可能な配列である。一般に、第2の配列または第2の配列の群にハイブリダイズ可能な第1の配列は、第2の配列または第2の配列の群に特異的または選択的にハイブリダイズ可能であり、第2の配列または第2の配列の群へのハイブリダイゼーションが、ハイブリダイゼーション反応中、非標的配列とのハイブリダイゼーションよりも好ましいようにする (例えば、当該技術分野において一般に使用される厳しい条件のような所与の条件群の下で熱力学的により安定である)。典型的に、ハイブリダイズ可能な配列は、それらそれぞれの長さの全部または一部分にわたってある程度の配列相補性、例えば、25%~100%の相補性 (少なくとも約25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、および100%の配列相補性を含

40

50

む)を共有する。

【0024】

ポリヌクレオチドに適用される「ハイブリダイズされた」という用語は、ヌクレオチド残基の塩基間の水素結合を介して安定化された複合中のポリヌクレオチドを指す。水素結合は、ワトソンクリック塩基対合、フーグスティーン結合、または任意の他の配列特異的方法により起こり得る。複合体は、二本鎖構造を形成する2つの鎖、多鎖複合体を形成する3つ以上の鎖、単一自己ハイブリダイジング鎖、またはこれらの任意の組み合わせを含んでよい。ハイブリダイゼーション反応は、より広範囲のプロセスにおけるステップ、例えば、PCRの開始、ライゲーション反応、配列決定反応、または切断反応を構成することができる。

10

【0025】

本発明の実施は、別段の指示が無い限り、当該技術分野の技術の範囲内である、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノムおよび組み換えDNAの従来技術を用いる。例えば、Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989)、CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1987))、シリーズMETHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995))、Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R. I. Freshney, ed. (1987))を参照されたい。

20

【0026】

一態様において、本発明は、複数の標的ポリヌクレオチドを配列決定するための装置を生成する方法を提供する。一実施形態において、この方法は、(a)反応表面を有する固体支持体を提供することと、(b)複数のオリゴヌクレオチドを固体支持体に結合することと、を含む。いくつかの実施形態において、複数のオリゴヌクレオチドは、(i)複数の異なる第1のオリゴヌクレオチドであって、配列Aおよび配列Bを含み、配列Aが、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bが、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にあり、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、第1のオリゴヌクレオチドと、(ii)配列Aをそれぞれの3'末端に含む複数の第2のオリゴヌクレオチドと、(iii)複数の異なる標的ポリヌクレオチドにより共有される配列と同じである、配列Cを3'末端に含む複数の第3のオリゴヌクレオチドと、を含む。いくつかの実施形態において、配列A、B、およびCのうちの1つ以上は、異なる配列である。いくつかの実施形態において、配列A、B、およびCのうちの1つ以上は、配列A、B、およびCのその他のうちの1つ以上とは約5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより多いか、それ未満、またはそれ以上異なる(例えば、約10%、20%、30%、40%、50%、60%、70%、80%、90%、またはそれより高い配列同一性未満の配列同一性を有する)。いくつかの実施形態において、配列A、B、およびCのうちの1つ以上は、それぞれ約1、2、3、4、5、6、7、8、9、10、15、20、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドを含む。

30

40

【0027】

様々な好適な固体支持体材料は、当該技術分野において既知である。固体支持体材料の非限定的な例としては、ガラス、溶融シリカ、および他のシリカ含有材料のようなシリカ系基質；ポリエチレン、ポリスチレン、ポリ(塩化ビニル)、ポリプロピレン、ナイロン

50

、ポリエステル、ポリカーボネート、ポリ（メチルメタクリレート）、および環状オレフィンポリマー基質のようなシリコン水素化物またはプラスチック材料；および金、二酸化チタン、またはケイ素支持体のような他の固体支持体が挙げられる。固体支持体材料は、任意の好適な形態で提供され得、ビーズ、ナノ粒子、ナノ結晶、ファイバー、マイクロファイバー、ナノファイバー、ナノワイヤー、ナノチューブ、マット、平板、平板ウエハーまたはスライド、多ウェルプレート、光学スライド、フローセル、およびチャンネルが挙げられるが、これらに限定されない。固体支持体は、1つ以上の追加構造、例えば、チャンネル、マイクロ流体チャンネル、毛細管、およびウェルをさらに含むことができる。いくつかの実施形態において、固体支持体は、フローセルのチャンネルである。

【0028】

10

分子（例えば、核酸）の固体支持体への固定または結合について言及するとき、「固定した」および「結合した」という用語は、本明細書において同義に使用され、両方の用語は、別段の指示が無い限り、直接または間接、共有または非共有結合を包含することが意図される。本発明のいくつかの実施形態において、共有結合が好まれ得るが、一般に、唯一必要なことは、分子（例えば、核酸）が、支持体を使用することが意図される条件の下、例えば、核酸増幅および/または配列決定の適用において、支持体に固定または結合したままであることである。

【0029】

いくつかの実施形態において、固体支持体材料は、指定条件の下、分子（例えば、オリゴヌクレオチドまたは修飾オリゴヌクレオチド）が、固体支持体の表面に直接結合され得るように、反応性の材料を含む。いくつかの実施形態において、固体支持体材料は、例えば、ポリヌクレオチド等の生分子への結合（例えば、共有結合）を許容する、反応基を含む中間材料の層またはコーティングの適用により、「官能化」された不活性基質またはマトリックス（例えば、ガラススライド、ポリマービーズ、または他の固体支持体材料）を含む。そのような支持体の例としては、ガラスのような不活性基質上に支持されるポリアクリルアミドハイドロゲルが挙げられるが、これに限定されない。そのような実施形態において、生体分子（例えば、オリゴヌクレオチド）は、中間材料（例えば、ヒドロゲル）に直接共有結合され得るが、この中間材料は、それ自体が基質またはマトリックス（例えば、ガラス基質）に非共有結合されてもよい。

20

【0030】

30

反応表面の非限定的な例としては、タンパク質の表面上への物理吸着によるビオチン基の安定した結合を形成するためのビオチン化アルブミン（BSA）の使用を含む。共有結合修飾は、分子を固体支持体に、通常はガラススライドに結合するために使用されている、シランを使用して行うことができる。例として、テトラエトキシシランおよびトリエトキシ-プロモアセトアミドプロピル-シランの混合物（例えば、1：100の比）を使用して、チオリン酸塩またはホスホロチオエート官能性を含む核酸の結合を許容する官能化ガラススライドを調製することができる。ビオチン分子は、アミノ表面と反応するビオチン-PEG-スクシンイミジルエステル等の適切な反応種を使用して表面に結合され得る。

【0031】

40

いくつかの実施形態において、固体支持体に結合されるオリゴヌクレオチドは、反応部分を含む。一般に、反応部分は、反応表面と反応することにより、固体支持体への結合を促進する任意の部分を含む。いくつかの実施形態において、官能化ポリアクリルアミドヒドロゲルは、反応部分を含む複数のオリゴヌクレオチドに結合するために使用され、この反応部分は、硫黄含有求核基である。適切な硫黄求核基含有ポリヌクレオチドの例は、Zhaoら（Nucleic Acids Research, 2001, 29(4), 955~959）およびPirrungaら（Langmuir, 2000, 16, 2185~2191）に開示され、例えば、単純チオール、チオリン酸塩、およびチオホスホルアミデートが挙げられる。好ましいヒドロゲルは、（i）アクリルアミド、メタクリルアミド、ヒドロキシエチルメタクリレート、またはN-ビニルピロリジノンである、第1のコ

50

モノマーと、(i i) N - (5 - プロモアセトアミジルペンチル) アクリルアミド、テトラメチルエチレンジアミンのような官能化アクリルアミドまたはアクリレートである、第2のコモノマーとの混合物から形成されるものである。いくつかの実施形態において、官能化ポリアクリルアミドを含む反応表面は、アクリルアミド、N - (5 - プロモアセトアミジルペンチル) アクリルアミド、テトラメチルエチレンジアミン、および過硫酸カリウムを含む重合混合物から生成される。支持体材料および反応表面のさらなる非限定的な例は、米国公開第20120053074号および国際公開第2005065814号により提供され、参照によりそれら全体が本明細書に組み込まれる。

【0032】

固体支持体が、結合のために曝露されるオリゴヌクレオチドは、任意の好適な長さであり得、1つ以上の配列要素を含み得る。配列要素の例としては、1つ以上の増幅プライマーアニーリング配列またはその相補体、1つ以上の配列決定プライマーアニーリング配列またはその相補体、複数の異なるオリゴヌクレオチドまたは異なるオリゴヌクレオチドのサブセットの中で共有される1つ以上の共通配列、1つ以上の制限酵素認識部位、1つ以上の標的ポリヌクレオチド配列に相補性の1つ以上の標的認識配列、1つ以上のランダムまたはニアランダム配列（例えば、1つ以上の位置にある2つ以上の異なるヌクレオチドの群からランダムに選択される1つ以上のヌクレオチド、異なるヌクレオチドのそれぞれは、そのランダム配列を含むオリゴヌクレオチドのプール中に表される1つ以上の位置で選択される）、1つ以上のスペーサー、およびこれらの組み合わせが挙げられるが、これらに限定されない。2つ以上の配列要素は、互いに非隣接であり得るか（例えば、1つ以上のヌクレオチドにより分離される）、互いに隣接し得るか、部分的に重なり得るか、または完全に重なり得る。例えば、増幅プライマーアニーリング配列は、配列決定プライマーアニーリング配列としても機能し得る。配列要素は、オリゴヌクレオチドの3'末端もしくはその付近、5'末端もしくはその付近、またはその内部に位置し得る。一般に、本明細書において使用される場合、「3'末端に」位置する配列要素は、オリゴヌクレオチドの最も3'側のヌクレオチドを含み、「5'末端に」位置する配列要素は、オリゴヌクレオチドの最も5'側のヌクレオチドを含む。いくつかの実施形態において、配列要素は、約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長である。いくつかの実施形態において、オリゴヌクレオチドは、約5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長である。

【0033】

スペーサーは、反復した単一ヌクレオチド（例えば、一列に1、2、3、4、5、6、7、8、9、10、もしくはそれより多くの同じヌクレオチド）、または1、2、3、4、5、6、7、8、9、10、もしくはそれより多くの回数反復される2、3、4、5、6、7、8、9、10、もしくはそれより多くのヌクレオチドの配列からなり得る。スペーサーは、試料中の任意の標的配列にハイブリダイズしない配列のような特定配列を含むか、またはそれからなり得る。スペーサーは、ランダムに選択されたヌクレオチドの配列を含むか、またはそれからなり得る。

【0034】

いくつかの実施形態において、それぞれが全ての第1のオリゴヌクレオチドの中で共通である配列Aと、それぞれの異なる第1のオリゴヌクレオチドに対して異なる配列Bとを含む、複数の異なる第1のオリゴヌクレオチドは、固体支持体に結合される。いくつかの実施形態において、それぞれの第1のオリゴヌクレオチドの配列Bは、異なる標的配列に相補性である。いくつかの実施形態において、複数の第1のオリゴヌクレオチドは、それぞれが異なる配列Bを含む、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、またはそれより多いか、それ未満、また

はそれ以上の異なる第1のオリゴヌクレオチドを含む。いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列Bは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む(例えば、それぞれが図4-1~3からの異なる配列を持つ1、5、10、25、50、75、または100の異なるオリゴヌクレオチド)。いくつかの実施形態において、配列Bまたはそれが特異的にハイブリダイズする標的配列は、原因となる遺伝的変異体を含む。いくつかの実施形態において、配列Bまたはそれが特異的にハイブリダイズする標的配列は、原因となる遺伝的変異体の約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。原因となる遺伝的変異体は、典型的に、第1のオリゴヌクレオチドの下流に位置し、原因となる遺伝的変異体の少なくとも一部分が、第1のオリゴヌクレオチドの伸長のためのテンプレートとなるようにする。一般に、原因となる遺伝的変異体は、疾患または形質との関連の統計的、生物学的、および/または機能的証明が存在する遺伝的変異体である。単一原因となる遺伝的変異体を、複数の疾患または形質と関連付けることができる。いくつかの実施形態において、原因となる遺伝的変異体を、メンデル形質、非メンデル形質、または両方と関連付けることができる。原因となる遺伝的変異体は、1、2、3、4、5、6、7、8、9、10、20、50、またはそれより多くの配列差といったポリヌクレオチド中(例えば、原因となる遺伝的変異体を含むポリヌクレオチドと、同じ相対ゲノム位置において原因となる遺伝的変異体を欠失するポリヌクレオチドとの間)の変化として呈し得る。原因となる遺伝的変異体の種類の非限定的な例としては、単一ヌクレオチド多型(SNP)、欠失/挿入多型(DIP)、コピー数変異体(CNV)、短縦列反復(STR)、制限断片長多型(RFLP)、単純配列反復(SSR)、可変数の縦列反復(VNTR)、ランダム増幅多型DNA(RAPD)、増幅断片長多型(AFLP)、レトロトランスポゾン間増幅多型(IRAP)、長短散在要素(LINE/SINE)、長縦列反復(LTR)、可動要素、レトロトランスポゾンマイクロサテライト増幅多型、レトロトランスポゾン系挿入多型、配列特異的増幅多型、および遺伝性後成的修飾(例えば、DNAメチル化)が挙げられる。原因となる遺伝的変異体は、密接に関連する原因となる遺伝的変異体の群であってもよい。いくつかの原因となる遺伝的変異体は、RNAポリヌクレオチドの配列変異として影響を及ぼし得る。このレベルで、いくつかの原因となる遺伝的変異体は、ある種のRNAポリヌクレオチドの存在または非存在によっても示される。またいくつかの原因となる遺伝的変異体は、タンパク質ポリペプチドの配列変異をもたらす。多数の原因となる遺伝的変異体が、当該技術分野において既知である。SNPである原因となる遺伝的変異体の一例は、鎌状赤血球貧血を引き起こすヘモグロビンのHb S変異体である。DIPである原因となる遺伝的変異体の一例は、嚢胞性線維症を引き起こすCFTR遺伝子の508突然変異である。CNVである原因となる遺伝的変異体の一例は、ダウン症候群を引き起こすトリソミー21である。STRである原因となる遺伝的変異体の一例は、ハンチントン病を引き起こす縦列反復である。図3-1~38は、原因となる遺伝的変異体および関連疾患の非限定的な例の表を提供する。原因となる遺伝的変異体の非限定的な例は、米国公開第20100022406号にも説明され、参照によりその全体が本明細書に組み込まれる。

【0035】

原因となる遺伝的変異体は本来、個体、族、および集団の遺伝子型および表現型の統計的および分子遺伝学的分析により発見され得る。メンデル形質の原因となる遺伝的変異体は、典型的に2段階プロセスで特定される。第1段階において、複数の個体とその形質を有する族が、遺伝子型および表現型について審査される。これらの族から得た遺伝子型および表現型データを使用して、メンデル形質の存在と多数の遺伝子マーカーの存在との間の統計的関連を確立する。この関連は、原因となる遺伝的変異体がマップする可能性が高い候補領域を確立する。第2段階において、原因となる遺伝的変異体自体が特定される。第2のステップは、典型的に、候補領域を配列決定することを含む。より洗練された1段

10

20

30

40

50

階プロセスが、原因となる遺伝的変異体の直接特定、またはより小さな候補領域の特定を許容する、より進んだ技術を用いて可能である。ある形質に対して1つの原因となる遺伝的変異体が発見された後、同じ形質の追加の変異体が、単純な方法により発見され得る。例えば、形質と関連付けられる遺伝子は、その形質またはそれらの相対物を有する個体において配列決定され得る。原因となる遺伝的変異体を発見するための新たな方法の発明は、活発な研究分野である。既存の方法の適用および新たな方法の組み込みは、本明細書に記載のデバイス、システム、および方法により使用または試験され得る追加の原因となる遺伝的変異体の発見をもたらし続けることが期待される。多くの原因となる遺伝的変異体は、オンライン版人間のメンデルの遺伝 (Online Mendelian Inheritance in Man (OMIM)) およびヒト遺伝子変異データベース (Human Gene Mutation Database (HGMD)) を含むデータベースにおいて目録化される。原因となる遺伝的変異体は、学術文献、カンファレンス、および学者間の私的通信においても報告されている。

【0036】

原因となる遺伝的変異体は、特定の集団内に任意の頻度で存在し得る。いくつかの実施形態において、原因となる遺伝的変異体のうちの少なくとも1つは、参照集団において1%を超えない発生率を有する形質を引き起こす。別の実施形態において、原因となる遺伝的変異体のうちの少なくとも1つは、参照集団において10,000分の1を超えない発生率を有する形質を引き起こす。いくつかの実施形態において、原因となる遺伝的変異体は、疾患または形質と関連付けられる。いくつかの実施形態において、原因となる遺伝的変異体は、その存在が、疾患または形質を有するか、または発症する危険性を約1%、5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、200%、300%、400%、500%、もしくはそれより多いか、それ未満、またはそれ以上だけ増加させる、遺伝的変異体である。いくつかの実施形態において、原因となる遺伝的変異体は、その存在が、疾患または形質を有するか、または発症する危険性を約1倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、25倍、50倍、100倍、500倍、1000倍、10000倍、もしくはそれより多いか、それ未満、またはそれ以上だけ増加させる、遺伝的変異体である。いくつかの実施形態において、原因となる遺伝的変異体は、その存在が、疾患または形質を有するか、または発症する危険性を、任意の統計的に有意な量 (例えば、約0.1、0.05、 10^{-3} 、 10^{-4} 、 10^{-5} 、 10^{-6} 、 10^{-7} 、 10^{-8} 、 10^{-9} 、 10^{-10} 、 10^{-11} 、 10^{-12} 、 10^{-13} 、 10^{-14} 、 10^{-15} 、もしくはそれより小さいか、またはそれ未満のp値を有する増加) だけ増加させる、遺伝的変異体である。

【0037】

いくつかの実施形態において、原因となる遺伝的変異体は、2つ以上の異なる個体集団間、例えば、2つ以上のヒト集団間で疾患または形質との異なる程度の関連を有する。いくつかの実施形態において、原因となる遺伝的変異体は、1つ以上の集団、例えば、1つ以上のヒト集団内でのみ疾患または形質との統計的に有意な関連を有する。ヒト集団は、共通の遺伝的継承を共有するヒトのグループ、例えば民族グループ (例えば、白色人種) であり得る。ヒト集団は、ハプロタイプ集団またはハプロタイプ集団のグループ (例えば、ハプロタイプH1、M52) であり得る。ヒト集団は、国民グループ (例えば、アメリカ人、イギリス人、アイルランド人) であり得る。ヒト集団は、年齢、性別、および社会経済学的因子により最も良く描出されるもの等の人口統計的集団であり得る。ヒト集団は、歴史的集団であり得る。集団は、分布の極端にある個体が決して互いに出会うことがないような大きな地理的領域にわたって分布する個体からなり得る。集団の個体は、不連続領域に地理的に分散し得る。集団は、生物地理的祖先についての情報を提供することができる。集団は、祖先により定義することもできる。遺伝学的研究が、集団を定義することができる。いくつかの実施形態において、集団は、祖先および遺伝学に基づくことがあり、主要なヒト集団は、西ユーラシア人、サハラ以南アフリカ人、東アジア人、および先住アメリカ人を含む、大陸規模の分類に対応する。ほとんどのヒトは、祖先に基づいてこれ

らの集団のうちの少なくとも1つに割り当てられ得る。多数のより小さな集団も、オーストラリア先住民、オセアニア人、ブッシュマンを含む大陸分類として区別される。

【0038】

非常に頻繁に、集団を亜集団にさらに分解することができる。集団と亜集団との間の関係は、階層的であり得る。例えば、オセアニア人集団は、ポリネシア人、メラネシア人、およびミクロネシア人を含む亜集団にさらに細分することができる。西ユーラシア人集団は、ヨーロッパ人、西ノ中央アジア人、南アジア人、および北アフリカ人を含む亜集団にさらに細分することができる。ヨーロッパ人集団は、北西ヨーロッパ人、南ヨーロッパ人、およびアシュケナージ系ユダヤ人集団を含む亜集団にさらに細分することができる。北西ヨーロッパ人集団は、イギリス人、アイルランド人、ドイツ人、フィンランド人等を含む国民集団にさらに細分することができる。東アジア集団は、中国人、日本人、および韓国人亜集団にさらに細分することができる。南アジア集団は、インド人およびパキスタン人集団にさらに細分することができる。インド人集団は、ドラヴィダ族、ブラフミー族、カンナダ族、マラヤーラム族、タミル族、テルグ族、ツルバ族、およびゴンド族にさらに細分することができる。亜集団は、原因となる遺伝的変異体を特定する目的で集団として機能し得る。

【0039】

いくつかの実施形態において、原因となる遺伝的変異体は、希な遺伝的疾患等の疾患と関連付けられる。原因となる遺伝的変異体が関連付けられ得る疾患の例として、21-ヒドロキシラーゼ欠乏症、ABCC8-関連インスリン過剰症、ARSA CS、軟骨形成不全、色覚異常、アデノシンリン酸デアミナーゼ1、神経細胞障害を伴う脳梁欠損症、アルカプトン尿症、1-抗トリプトファン欠乏症、マンノシド症、サルコグリカン異常症、サラセミア、アルツハイマー、アンジオテンシンII受容体、I型、アポリポタンパク質E遺伝子型決定、アルギニノコハク酸尿症、アスパルチルグリコサミン尿症、ビタミンE欠乏症を伴う運動失調、運動失調-毛細管拡張症、多腺性自己免疫症候群1型、BRCA1遺伝性乳癌/卵巣癌、BRCA2遺伝性乳癌/卵巣癌、1つ以上の他の種類の癌、バルデ-ピードル症候群、ベスト卵黄様黄斑変性症、サルコグリカン異常症、サラセミア、ピオチニダーゼ欠乏症、ブラウ症候群、ブルーム症候群、CFTR関連障害、CLN3-関連神経セロイド-リポフスチン症、CLN5-関連神経セロイド-リポフスチン症、CLN8-関連神経セロイド-リポフスチン症、カナバン病、カルニチンパルミトイルトランスフェラーゼIA欠乏症、カルニチンパルミトイルトランスフェラーゼII欠乏症、軟骨毛髪形成不全症、脳海綿状血管奇形、全脈絡膜萎縮、コーエン症候群、先天性白内障、顔面異形症および神経障害、先天性グリコシル化異常症Ia、先天性グリコシル化異常症Ib、フィンランド型先天性ネフローゼ、クローン病、シスチン症、DFNA9(COCH)、糖尿病および難聴、早発性原発性ジストニア(DYT1)、接合型表皮水疱症ヘルリツ-ピアソン型、FANCC-関連ファンconi貧血、FGFR1-関連頭蓋骨癒合症、FGFR2-関連頭蓋骨癒合症、FGFR3-関連頭蓋骨癒合症、第V因子ライデン栓友病、第V因子R2突然変異栓友病、第XI因子欠乏症、第XIII因子欠乏症、家族性腺腫性ポリポージス、家族性自律神経失調症、家族性高コレステロール血症B型、家族性地中海熱、遊離シアル酸蓄積症、パーキンソニズム-17を伴う前頭側頭型認知症、フマラーゼ欠乏症、GJB2-関連DFNA3非症候性難聴および聴覚消失、GJB2-関連DFNB1非症候性難聴および聴覚消失、GNE-関連筋障害、ガラクトース血症、ゴーシェ病、グルコース-6-リン酸デヒドロゲナーゼ欠乏症、グルタル酸血症I型、糖原貯蔵症Ia型、糖原貯蔵症Ib型、糖原貯蔵症II型、糖原貯蔵症III型、糖原貯蔵症V型、薄束(Gracile)症候群、HFE-関連遺伝性ヘモクロマトーシス、ハイデルAIM、ヘモグロビンS-サラセミア、遺伝性フルクトース不耐症、遺伝性膀胱炎、遺伝性チミン-ウラシル尿症、ヘキソサミニダーゼA欠乏症、発汗性外胚葉形成不全症2、シスタチオニン-シンターゼ欠乏により引き起こされるホモシスチン尿症、高カリウム性周期性四肢麻痺1型、高オルニチン血症-高アンモニア血症-ホモシトルリン尿症症候群、原発性1型高シュウ酸尿症、原発性2型高シュウ酸尿症、軟骨低形

10

20

30

40

50

成症、低カリウム血性周期性麻痺 1 型、低カリウム血性周期性麻痺 2 型、低ホスファターゼ症、小児筋障害および乳酸アシドーシス（致命のおよび非致命的形態）、イソ吉草酸血症、クラッペ病、L G M D 2 I、レーバー遺伝性視神経萎縮症、リー症候群フランス系カナダ型、長鎖 3 - ヒドロキシアシル - C o A デヒドロゲナーゼ欠乏症、M E L A S、M E R R F、M T H F R 欠乏症、M T H F R 熱不安定性異型、M T R N R 1 - 関連難聴および聴覚消失、M T T S 1 - 関連難聴および聴覚消失、M Y H - 関連ポリポーシス、メーブルシロップ尿症 1 A 型、メーブルシロップ尿症 1 B 型、マクーン - オルブライト症候群、中鎖アシル - コエンザイム A デヒドロゲナーゼ欠乏症、皮質下嚢胞を伴う巨脳性白質脳症、異染色性白質ジストロフィー、ミトコンドリア心筋症、ミトコンドリア D N A - 関連リー症候群および N A R P、ムコリピド症 I V、ムコ多糖症 I 型、ムコ多糖症 I I I A 型、ムコ多糖症 V I I 型、多内分泌腺腫瘍 2 型、筋 - 眼 - 脳病、ネマリ筋障害、神経学的表現型、スフィンゴミエリナーゼ欠乏に起因するニーマン - ピック病、ニーマン - ピック病 C 1 型、ナイミーヘン染色体不安定症候群、P P T 1 - 関連ニューロンセロイド - リポフスチン症、P R O P 1 - 関連下垂体ホルモン欠乏症、パリスト - ホール症候群、先天性筋緊張症、ペンドレッド症候群、ペルオキシソーム二機能酵素欠乏症、広汎性発達障害、フェニルアラニンヒドロキシラーゼ欠乏症、プラスミノゲン活性化因子抑制剤 I、多発性嚢胞腎常染色体劣性、プロトロンビン G 2 0 2 1 0 A 栓友病、ブソイドビタミン D 欠乏くる病、濃化異骨症、網膜色素変性、常染色体劣性ボスニア型、レット症候群、肢根型点状軟骨異形成症 1 型、短鎖アシル - C o A デヒドロゲナーゼ欠乏症、シュバックマン - ダイアモンド症候群、シェーグレン - ラルソン症候群、スミス - レムリ - オピッツ症候群、痙性対麻痺 1 3、硫酸輸送体 - 関連骨軟骨異形成、T F R 2 - 関連遺伝性ヘモクロマトーシス、T P P 1 - 関連神経セロイド - リポフスチン症、致死性異形成症、トランスチレチンアミロイド症、三機能タンパク質欠乏症、チロシンヒドロキシラーゼ - 欠乏 D R D、チロシン血症 I 型、ウィルソン病、X - 結合若年網膜隔離症、およびゼルウィガー症候群スペクトルが挙げられるが、これらに限定されない。

【 0 0 4 0 】

いくつかの実施形態において、複数の第 1 のオリゴヌクレオチドのうちの 1 つ以上の配列 B またはそれが特異的にハイブリダイズする標的配列は、非対象配列を含む。いくつかの実施形態において、配列 B またはそれが特異的にハイブリダイズする標的配列は、非対象配列の約 1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。一般に、非対象配列は、試験される個体以外の生物から誘導されるポリヌクレオチド、例えば、細菌、古細菌、ウイルス、原生生物、真菌、または他の生物からの D N A または R N A に対応する。非対象配列は、ある生物または生物の類の同一性を示すことがあり、さらに感染のような疾患状態を示し得る。生物を特定する際に有用な非対象配列の例としては、限定されないが、r R N A 配列、例えば 16 s r R N A 配列が挙げられる（例えば、国際公開第 2 0 1 0 1 5 1 8 4 2 号を参照）。いくつかの実施形態において、非対象配列は、原因となる遺伝的変異体の代わりに、またはそれとは別に分析される。いくつかの実施形態において、原因となる遺伝的変異体および非対象配列は、並行して、例えば同じ試料中（例えば、第 1 のオリゴヌクレオチド、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つもの、および非対象配列を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つものの混合物を使用する）および / または同じレポートにおいて分析される。

【 0 0 4 1 】

いくつかの実施形態において、複数の第 1 のオリゴヌクレオチドのうちの 1 つ以上の配列 B、またはそれが特異的にハイブリダイズする標的配列は、祖先情報提供マーカー（A I M）を含む。いくつかの実施形態において、配列 B またはそれが特異的にハイブリダイズする標的配列は、A I M の約 1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、5

10

20

30

40

50

00、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。一般に、AIMは、2つ以上のヒト集団等の2つ以上の個体集団間で頻度が異なる遺伝的変異体であり、単独または1つ以上の他のAIMとの組み合わせのいずれかで対象の祖先を推理するために使用され得る。AIMは、原因となる遺伝的変異体のうちの1つの危険性が高い集団等の1つ以上の集団に属するか、または属しない人を分類するために使用され得る。例えば、AIMは、ある形質の罹患率が高い集団についての診断手段となり得る。場合によっては、AIMは、より細かい精度で集団、例えば、亜大陸群または関連する民族グループを区別する。いくつかの実施形態において、AIMは、原因となる遺伝的変異体および/または非対象配列の代わりに、またはそれとは別に分析される。いくつかの実施形態において、AIM、原因となる遺伝的変異体、および/または非対象配列は、並行して、例えば、同じ試料中（例えば、第1のオリゴヌクレオチド、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つもの、およびAIMを含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つものの混合物を使用する）および/または同じレポート中で分析される。AIMの種類の特異的な例としては、単一ヌクレオチド多型（SNP）、欠失/挿入多型（DIP）、コピー数変異体（CNV）、短縦列反復（STR）、制限断片長多型（RFLP）、単純配列反復（SSR）、可変数の縦列反復（VNTR）、ランダム増幅多型DNA（RAPD）、増幅断片長多型（AFLP）、レトロトランスポゾン間増幅多型（IRAP）、長および短散在要素（LINE/SINE）、長縦列反復（LTR）、可動要素、レトロトランスポゾンマイクロサテライト増幅多型、レトロトランスポゾン系挿入多型、配列特異的増幅多型、および遺伝性後成的修飾（例えば、DNAメチル化）が挙げられる。AIMは、RNAポリヌクレオチド中の配列変異でもあり得る。いくつかのAIMは、ある種のRNAポリヌクレオチドの存在または濃度により示すこともできる。いくつかのAIMは、タンパク質ポリペプチド中の配列変異でもあり得る。いくつかのAIMは、ある種のタンパク質ポリペプチドの存在または非存在により示すこともできる。多数の祖先情報提供マーカーが、図9-1~10に特定される。他のAIMは、米国公開第2007/0037182号に説明されている。AIMは、原因となる遺伝的変異体であってもなくてもよい。例えば、ダフィーヌル（FY*0）遺伝的変異体は、血液抗原の非存在を引き起こす。この異型は、サハラ以南アフリカ人集団においてほぼ100%の頻度で存在し、サハラ以南アフリカ外の集団においてほぼ0%の頻度で存在する。色素沈着と関連付けられる多くの原因となる遺伝的変異体もAIMである。原因となる遺伝的変異体でないAIMは、他のAIMにより引き起こされた形質と間接的に関連付けることができる。

【0042】

AIMは、複数の集団における遺伝的変異体の頻度を決定することにより発見することができる。これは、様々な集団からの個体における既知の異型の頻度を決定することにより達成され得る。異型発見のプロセス中に本質的に達成されることもある。両方のタスクは、SNP多型を目録化したInternational HapMapプロジェクトにより行われた。祖先情報提供マーカーは、それらの予知力を判断する多様な測定値によりランク付することができる。1つの測定値は、FstまたはFSTと呼ばれるライトのF統計である。この変数は、固定指数を含む他の名前でも知られている。AIMをランク付するための別の基準は、情報提供性である。AIMをランク付する別の方法は、PaschouらのPCA関連SNP法である（Paschou et al. PCA-correlated SNPs for structure identification in worldwide human populations. PLoS Genet (2007) 第3巻(9) pp. 1672~86）。

【0043】

祖先情報提供マーカーに基づいて、祖先推理において前選択した程度の信頼性を達成するため（例えば、少なくとも約80%、85%、90%、95%、97.5%、99%、またはそれ以上）、および複数の集団の祖先推理を達成するために、複数の祖先情報提供マーカーを審査することが必要であり得る。ランダムに選択された遺伝的変異体の十分に

大きな一団を使用して、祖先を推理することができる（例えば、約 5、10、15、25、50、100、250、500、1000、2500、5000、もしくはそれより多いか、またはそれ以上の AIM）。特に適切な AIM の標的群を構築することができる。多くの研究者は、示唆される祖先情報提供マーカーの一覧表を公開している（例えば、Seldin et al. Application of ancestry informative markers to association studies in European Americans. PLoS Genet (2008) 第 4 巻 (1) pp. e5、Halder et al. A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. Hum Mutat (2008) 第 29 巻 (5) pp. 648~58、Tian et al. Analysis and application of European genetic substructure using 300K SNP information. PLoS Genet (2008) 第 4 巻 (1) pp. e4、Price et al. Discerning the ancestry of European Americans in genetic association studies. PLoS Genet (2008) 第 4 巻 (1) pp. e236、Paschou et al. PCA-correlated SNPs for structure identification in world wide human populations. PLoS Genet (2007) 第 3 巻 (9) pp. 1672~86、および Bauchet et al. Measuring European population stratification with microarray genotype data. Am J Hum Genet (2007) 第 80 巻 (5) pp. 948~56)。これらのおよび同様の一覧表を使用し、本明細書に記載されるデバイスまたは方法が試験するように構成され得る AIM の一団を構築することができる。

【0044】

いくつかの実施形態において、複数の第 2 のヌクレオチドおよび複数の第 3 のヌクレオチドは、複数の第 1 のヌクレオチドに加えて、固体支持体に結合される。いくつかの実施形態において、第 2 のヌクレオチドは全て、配列 A を 3' 末端に含み、複数の第 2 のオリゴヌクレオチド中の配列 A は、第 1 のオリゴヌクレオチドの全ての配列 A と同じである。いくつかの実施形態において、第 3 のオリゴヌクレオチドは、配列 C を 3' 末端に含み、配列 C は、複数の異なる標的ポリヌクレオチドにより共有される配列に相補性である。いくつかの実施形態において、テンプレートとなる標的ポリヌクレオチドに沿った第 1 のオリゴヌクレオチドの伸長は、配列 C を含む伸長生成物を生成し、これは配列 C に相補性であり、特異的にハイブリダイズ可能である。いくつかの実施形態において、固体支持体に曝露される複数の第 2 のオリゴヌクレオチドの量は、例えば、固体支持体に結合された複数のオリゴヌクレオチドに対する反応において、固体支持体に曝露される複数の第 1 のオリゴヌクレオチドの量より約 10 倍、50 倍、100 倍、1000 倍、5000 倍、7500 倍、10000 倍、12500 倍、15000 倍、20000 倍、50000 倍、100000 倍、もしくはそれより多いか、それ未満、またはそれ以上高い。いくつかの実施形態において、複数の第 2 のオリゴヌクレオチドの量対固体支持体に曝露された第 3 のオリゴヌクレオチドの量の比（または反比）は、約 1 : 1、2 : 1、3 : 1、4 : 1、5 : 1、6 : 1、7 : 1、8 : 1、9 : 1、10 : 1、もしくはそれより多いか、それ未満、またはそれ以上である。いくつかの実施形態において、複数の第 1 のオリゴヌクレオチドは、固体支持体に約 0.5 pM、1 pM、5 pM、10 pM、25 pM、50 pM、75 pM、100 pM、200 pM、500 pM、1 nM、10 nM、100 nM、500 nM、もしくはそれより高いか、それ未満、またはそれ以上の濃度で付加される。いくつかの実施形態において、複数の第 2 のオリゴヌクレオチドおよび / または第 3 のオリゴヌ

10

20

30

40

50

クレオチドの濃度は、約 0.5 nM、1 nM、5 nM、10 nM、25 nM、50 nM、75 nM、100 nM、200 nM、500 nM、1 μM、5 μM、10 μM、25 μM、50 μM、100 μM、500 μM、もしくはそれより高いか、それ未満、またはそれ以上である。

【0045】

いくつかの実施形態において、1つ以上の複数のオリゴヌクレオチドは、1つ以上の保護基を含む。一般に、保護基は、例えば、ポリメラーゼ、リガーゼ、および/または他の酵素によりオリゴヌクレオチドの3'末端の伸長を防ぐ任意の修飾である。保護基は、オリゴヌクレオチドが固体支持体に結合される前または後に付加されてよい。いくつかの実施形態において、保護基は、増幅または配列決定プロセスの間に付加される。保護基の例としては、アルキル基、非ヌクレオチドリinker、ホスホロチオエート、アルカン-ジオール残基、ペプチド核酸、および例えば、コルジセピンを含む3'-OHを欠くヌクレオチド誘導体が挙げられるが、これらに限定されない。

10

【0046】

いくつかの実施形態において、基質に結合されたオリゴヌクレオチドのうちの1つ以上は、切断部位を含み、その部位での切断が、切断されたポリヌクレオチドの全てまたは一部分を、固体支持体への結合から解放するようにする。いくつかの実施形態において、切断は、ポリヌクレオチドテンプレートに沿って伸長され得る3'末端を生成する。いくつかの実施形態において、複数の第1、第2、および/または第3のオリゴヌクレオチドの一部分のみが、切断部位を含む（例えば、約10%、20%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより多いか、それ未満、またはそれ以上）。切断部位は、化学的、酵素的、および光化学的切断を含むが、これらに限定されない任意の好適な手段により切断可能であり得る。切断基は、第1のヌクレオチドと固体支持体との間、またはオリゴヌクレオチド中の任意の数のヌクレオチド、例えば、固体支持体への結合点から約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド、あるいはその後位置付けられ得る。

20

【0047】

化学的、酵素的、および光化学切断のプロセス、ならびにそのようなプロセスにより切断される切断部位は、当該技術分野において既知である。切断手段の例としては、切断部位が、二本鎖テンプレートの片方または両方の鎖の切断を配向する酵素の適切な制限部位である、制限酵素消化；切断部位が1つ以上のリボヌクレオチドを含み得る、デオキシリボヌクレオチドとリボヌクレオチドとの間の結合のRNase消化または化学切断；切断部位が、適切なジスルフィド結合を含む必要がある、還元剤（例えば、TCEP）を用いたジスルフィド結合の化学還元；切断部位が、ジオール結合を含む必要がある、過ヨウ素酸塩を用いたジオール結合の化学切断；脱塩基部位の生成および後次の加水分解が挙げられるが、これらに限定されない。切断に続いて、例えば、ポリメラーゼ、リガーゼ、および/または他の酵素により伸長することができない3'末端の生成を遮断することができる。遮断薬の例としては、切断剤の付加前、付加中、または付加後に付加され得るアミン（例えば、エタノールアミン）が挙げられるが、これに限定されない。切断プロセスおよび切断部位の追加の非限定的な例は、米国公開第20120053074号に説明され、参照によりその全体が組み込まれる。

30

40

【0048】

いくつかの実施形態において、複数の標的ポリヌクレオチドが、複数の標的ポリヌクレオチドを含む試料を本発明の装置に曝露することを含む方法に従い増幅される。いくつかの実施形態において、増幅プロセスは、架橋増幅を含む。標準架橋増幅を行うための一般的な方法は、当該技術分野において既知である。例として、国際公開第1998/044151号および同第2000/018957号はともに、複数の同一の固定されたポリヌクレオチド鎖および複数の同一の固定された相補鎖から形成されたクラスタまたは「コロニー」からなるアレイを形成するために、増幅生成物が固体支持体上に固定されるのを許

50

容する、核酸増幅の方法を説明する。いくつかの実施形態において、複数のポリヌクレオチドは、複数の標的ポリヌクレオチドを含む試料を本発明の装置に曝露することを含む方法に従い配列決定される。固体支持体に結合された複数のオリゴヌクレオチドを使用して配列決定を行うための一般的な方法、例えば、参照によりそれら全体が本明細書に組み込まれる、米国公開第20120053074号および同第20110223601号に開示される方法は、当該技術分野において既知である。本発明の方法および装置に従い標的ポリヌクレオチドを増幅および/または配列決定するための非限定的な例示的方法が、本明細書で提供される。一般に、特定の標的ポリヌクレオチドの増幅は、非標的ゲノム配列と比較して、標的ゲノム配列のような標的ポリヌクレオチドに対して強化される配列決定データの生成を許容する。いくつかの実施形態において、非標的ポリヌクレオチドと比較して、標的ポリヌクレオチドに対する配列決定データ（特に原因となる遺伝的変異体の配列決定データ）の強化は、約または少なくとも約10倍、100倍、500倍、1000倍、5000倍、10000倍、50000倍、100000倍、1000000倍、またはそれ以上である。

【0049】

オリゴヌクレオチドを含む基質、それらの生成のための方法、ならびにそれらの操作のためのシステムおよび方法の非限定的な例は、国際公開第2008/002502号に提供され、参照によりその全体が本明細書に組み込まれる。

【0050】

一態様において、本発明は、試料中の複数の標的ポリヌクレオチドを配列決定するための方法を提供する。一実施形態において、この方法は、(a)断片化ポリヌクレオチドを生成するように、標的ポリヌクレオチドを断片化することと、(b)適合したポリヌクレオチドの両端で相補性配列Dにハイブリダイズされた配列Dを含む、適合したポリヌクレオチドを生成するように、それぞれが配列Dを含むアダプターオリゴヌクレオチドを、断片化ポリヌクレオチドに連結することと、任意に配列Dが、標的ポリヌクレオチド3'末端の伸長により生成される、連結することと、(c)適合したポリヌクレオチドを、配列C、配列D、および試料と関連付けられたバーコードを含む増幅プライマーを使用して増幅することと、配列Dが、増幅プライマーの3'末端に位置付けられる、増幅することと、(d)増幅された標的ポリヌクレオチドを、固体表面に結合した複数の異なる第1のオリゴヌクレオチドにハイブリダイズすることと、(e)固体表面上で架橋増幅を行うことと、(f)ステップ(e)からの複数のポリヌクレオチドを配列決定することと、を含む。固体表面は、本明細書に記載され、任意に本明細書に記載の方法に従い生成される装置を含む、本明細書に記載される複数のオリゴヌクレオチドを含み得る。いくつかの実施形態において、固体表面は、(i)複数の異なる第1のオリゴヌクレオチドであって、配列Aおよび配列Bを含み、配列Aが、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bが、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にあり、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、複数の異なる第1のオリゴヌクレオチドと、(ii)配列Aをそれぞれの3'末端に含む、複数の第2のオリゴヌクレオチドと、(iii)配列Cをそれぞれの3'末端に含む、複数の第3のオリゴヌクレオチドと、を含む。いくつかの実施形態において、配列A、B、C、およびDのうちの1つ以上は、異なる配列である。いくつかの実施形態において、配列A、B、C、およびDのうちの1つ以上は、配列A、B、C、およびDのうちの他の1つ以上とは約5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより多いか、それ未満、またはそれ以上異なる（例えば、約10%、20%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより多い配列同一性未満の配列同一性を有する）。いくつかの実施形態において、配列A、B、C、およびDのうちの1つ以上は、それぞれ約1、2、3、4、5、6、7、8、9、10、15、20、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドを含む。

10

20

30

40

50

【0051】

標的ポリヌクレオチドが誘導される試料は、同じ個体からの複数の試料、異なる個体からの試料、またはこれらの組み合わせを含むことができる。いくつかの実施形態において、試料は、単一個体からの複数のポリヌクレオチドを含む。いくつかの実施形態において、試料は、2つ以上の個体からの複数のポリヌクレオチドを含む。個体は、標的ポリヌクレオチドが誘導され得る任意の生物またはその一部分であり、その非限定的な例としては、植物、動物、真菌、原生生物、モネラ界生物、ウイルス、ミトコンドリア、および葉緑体が挙げられる。試料ポリヌクレオチドは、細胞試料、組織試料、流体試料、またはそこから誘導される器官試料（またはそれらのうちのいずれかから誘導される細胞培養物）等の対象から単離することができ、例えば、培養細胞株、生検、血液試料、口腔粘膜検体、または細胞を含有する流体試料（例えば、唾液）が挙げられる。対象は、動物であり得、限定されないが、ウシ、ブタ、マウス、ラット、ニワトリ、ネコ、イヌ等が挙げられ、通常はヒトのような哺乳類である。試料は、例えば、化学合成により人工的に誘導することもできる。いくつかの実施形態において、試料は、DNAを含む。いくつかの実施形態において、試料は、ゲノムDNAを含む。いくつかの実施形態において、試料は、ミトコンドリアDNA、葉緑体DNA、プラスミドDNA、細菌人工染色体、酵母人工染色体、オリゴヌクレオチドタグ、試料が採取された対象以外の生物（例えば、細菌、ウイルス、または真菌）からのポリヌクレオチド、またはこれらの組み合わせを含む。いくつかの実施形態において、試料は、増幅により、例えば、プライマーおよびDNAポリメラーゼの任意の好適な組み合わせを使用するプライマー伸長反応（ポリメラーゼ連鎖反応（PCR）、逆転写、およびこれらの組み合わせを含むが、これらに限定されない）により生成されたDNAを含む。プライマー伸長反応のテンプレートがRNAである場合、逆転写の生成物は、相補DNA（cDNA）と称される。プライマー伸長反応において有用なプライマーは、1つ以上の標的に特異的な配列、ランダム配列、部分的ランダム配列、およびこれらの組み合わせを含むことができる。プライマー伸長反応に好適な反応条件は、当該技術分野において既知である。一般に、試料ポリヌクレオチドは、試料中に存在する任意のポリヌクレオチドを含む、標的ポリヌクレオチドを含んでも含まなくてもよい。いくつかの実施形態において、単一個体からの試料は、複数の別個の試料に分割され（例えば、2、3、4、5、6、7、8、9、10、またはそれより多くの別個の試料）、それらは独立して本発明の方法、例えば、二重、三重、四重、またはそれ以上の分析に供される。

【0052】

核酸の抽出および精製のための方法は、当該技術分野において周知である。例えば、核酸は、フェノール、フェノール/クロロホルム/イソアミルアルコール、またはTRIzolおよびTri Reagentを含む同様の製剤を用いる有機抽出により精製することができる。抽出技法の他の非限定的な例としては、（1）例えば、フェノール/クロロホルム有機試薬（Ausubel et al., 1993）を使用するエタノール沈殿が続く有機抽出（自動核酸抽出器、例えば、Applied Biosystems（Foster City, Calif.）から入手できるモデル341 DNA抽出器を使用するか、または使用しない）、（2）固定相吸着法（米国特許第5,234,809号、Walsh et al., 1991）、および（3）食塩誘発性核酸沈殿法（Miller et al., (1988)、そのような沈殿法は、典型的に「塩析」法と称される）が挙げられる。核酸単離および/または精製の別の例としては、核酸が特異的に、または非特異的に結合することができる磁性粒子の使用に続く、磁石を使用し、核酸をビーズから洗浄および溶出するビーズの単離が挙げられる（例えば、米国特許第5,705,628号を参照）。いくつかの実施形態において、上記単離法は、試料から望ましくないタンパク質を排除する、例えば、プロテイナーゼKまたは他の同様のプロテイナーゼを用いる消化を助けるように、酵素消化ステップが先行し得る。例えば、米国特許第7,001,724号を参照されたい。所望される場合、RNAase阻害剤は、溶解緩衝液に付加されてよい。ある特定の細胞または試料の種類の場合、タンパク質変性/消化ステップをプロトコルに追加することが望ましい場合がある。精製方法は、DNA、RNA、または

両方を単離するように配向され得る。DNAおよびRNAの両方が、抽出手順の間またはその後一緒に単離されるとき、さらなるステップを用いて、片方または両方を他方から別個に精製することができる。抽出された核酸の亜分画は、例えば、サイズ、配列、または他の物理的もしくは化学的特徴ごとに精製することにより生成することもできる。最初の核酸単離ステップに加えて、例えば、過剰な、または望ましくない試薬、反応物質、または生成物を取り除くために、本発明の方法における任意のステップ後に核酸の精製を行うことができる。試料中の核酸の量および/または純度を決定するための方法は、当該技術分野において既知であり、吸光度（例えば、260nm、280nmでの光の吸光度、およびそれらの比）および標識の検出（例えば、SYBRグリーン、SYBRブルー、DAPI、ヨウ化プロピジウム、ヘキスト染色、SYBRゴールド、臭化エチジウム等の蛍光染色および挿入剤）が挙げられる。

10

【0053】

いくつかの実施形態において、標的ポリヌクレオチドは、1つ以上の特定のサイズ範囲（複数可）の断片化ポリヌクレオチドの集団に断片化される。いくつかの実施形態において、断片化に供される試料ポリヌクレオチドの量は、約50ng、100ng、200ng、300ng、400ng、500ng、600ng、700ng、800ng、900ng、1000ng、1500ng、2000ng、2500ng、5000ng、10μg、もしくはそれより多いか、それ未満、またはそれ以上である。いくつかの実施形態において、断片は、約1、10、100、1000、10000、100000、300000、500000、もしくはそれより多いか、それ未満、またはそれ以上の開始DNAのゲノム相当物から生成される。断片化は、化学的、酵素的、および機械的断片化を含む、当該技術分野において既知の方法により達成され得る。いくつかの実施形態において、断片は、約10~約10,000ヌクレオチドの長さの平均または中央値を有する。いくつかの実施形態において、断片は、約50~約2,000ヌクレオチドの長さの平均または中央値を有する。いくつかの実施形態において、断片は、約100~2500、200~1000、10~800、10~500、50~500、50~250、または50~150ヌクレオチド、それ未満、それを超える、またはその間の長さの平均または中央値を有する。いくつかの実施形態において、断片は、約200、300、500、600、800、1000、1500ヌクレオチド、もしくはそれより多いか、それ未満、またはそれ以上の長さの平均または中央値を有する。いくつかの実施形態において、断片化は、機械的に達成され、試料ポリヌクレオチドを超音波処理に供することを含む。いくつかの実施形態において、断片化は、1つ以上の酵素が二本鎖核酸切断を生成するのに好適な条件下で、試料ポリヌクレオチドを1つ以上の酵素で処理することを含む。ポリヌクレオチド断片の生成に有用な酵素の例としては、配列特異的および非配列特異的ヌクレアーゼが挙げられる。ヌクレアーゼの非限定的な例としては、DNase I、断片化酵素、制限エンドヌクレアーゼ、それらの異型、およびそれらの組み合わせが挙げられる。例えば、DNase Iを用いる消化は、Mg++の非存在下、およびMn++の存在下でDNA中のランダム二本鎖切断を誘発することができる。いくつかの実施形態において、断片化は、試料ポリヌクレオチドを1つ以上の制限エンドヌクレアーゼで処理することを含む。断片化は、5'オーバーハング、3'オーバーハング、平滑末端、またはこれらの組み合わせを有する断片を生成することができる。いくつかの実施形態において、例えば、断片化が1つ以上の制限エンドヌクレアーゼの使用を含むとき、試料ポリヌクレオチドの切断は、予測可能な配列を有するオーバーハングを残す。いくつかの実施形態において、この方法は、アガロースゲルからのカラム精製または単離等の標準方法を介して断片をサイズ選択するステップを含む。いくつかの実施形態において、この方法は、断片化後に断片長の平均および/または中央値を決定することを含む。いくつかの実施形態において、所望の閾値を上回る断片長の平均および/または中央値を有する試料は、再度断片化に供される。いくつかの実施形態において、所望の閾値を下回る断片長の平均および/または中央値を有する試料は、破棄される。

20

30

40

【0054】

50

いくつかの実施形態において、断片化ポリヌクレオチドの5'および/または3'末端ヌクレオチド配列は、1つ以上のアダプターオリゴヌクレオチド(「アダプター」とも称される)とのライゲーションの前に修飾されない。例えば、制限エンドヌクレアーゼによる断片化を使用して、予測可能なオーバーハングを残すことができ、ポリヌクレオチド断片上の予測可能なオーバーハングに相補性であるオーバーハングを含む1つ以上のアダプターオリゴヌクレオチドとのライゲーションが続く。別の例において、予測可能な平滑末端を残す酵素による切断の後に、平滑末端を含むアダプターオリゴヌクレオチドに対する平滑末端化ポリヌクレオチド断片のライゲーションが続く。いくつかの実施形態において、断片化ポリヌクレオチドは、アダプターに連結される前に、平滑末端を有するポリヌクレオチド断片を生成するように平滑末端研磨される(または「末端修復される」)。平滑末端研磨ステップは、3'~5'エキソヌクレアーゼ活性および5'~3'ポリメラーゼ活性の両方を有するDNAポリメラーゼ、例えば、T4ポリメラーゼといった好適な酵素でのインキュベーションにより達成され得る。いくつかの実施形態において、末端修復は、オーバーハングを生成するために、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、もしくはそれより多くのヌクレオチド、例えば、1つ以上のアデニン(「Aテーリング」)、1つ以上のチミン、1つ以上のグアニン、または1つ以上のシトシンの付加が後に続くか、またはそれで終了する。オーバーハングを有するポリヌクレオチド断片を、例えば、ライゲーション反応において、相補性オーバーハングを有する1つ以上のアダプターオリゴヌクレオチドに連結することができる。例えば、単一のアデニンを、テンプレート独立ポリメラーゼを使用して、末端修復されたポリヌクレオチド断片の3'末端に付加することができ、後にそれぞれが3'末端に突出するチミンを有する1つ以上のアダプターへのライゲーションが続く。いくつかの実施形態において、アダプターオリゴヌクレオチドは、1つ以上のヌクレオチドを持つ3'末端の伸長に続いて、5'リン酸化により修飾された平滑末端二本鎖DNA断片分子に連結することができる。場合によっては、3'末端の伸長は、例えば、クレノウポリメラーゼもしくは当該技術分野において既知の任意の他の好適なポリメラーゼ等のポリメラーゼを用いるか、またはターミナルデオキシヌクレオチドトランスフェラーゼの使用により、マグネシウムを含有する好適な緩衝液中の1つ以上のdNTPの存在下で行われ得る。いくつかの実施形態において、平滑末端を有する標的ポリヌクレオチドは、平滑末端を含む1つ以上のアダプターに連結される。断片化ポリヌクレオチドの5'末端のリン酸化は、例えば、ATPおよびマグネシウムを含有する好適な緩衝液中のT4ポリヌクレオチドを用いて行われ得る。断片化ポリヌクレオチドは、任意に、例えば、ホスファターゼ等の当該技術分野において既知の酵素を使用することにより、5'末端または3'末端を脱リン酸化するように処理されてよい。

【0055】

いくつかの実施形態において、断片化の後に、断片化ポリヌクレオチドに対するアダプターオリゴヌクレオチドのライゲーションが続く。アダプターオリゴヌクレオチドは、配列を有する任意のオリゴヌクレオチドを含み、その少なくとも一部分は、標的ポリヌクレオチドに連結され得ることが知られている。アダプターオリゴヌクレオチドは、DNA、RNA、ヌクレオチド類似体、非標準ヌクレオチド、標識ヌクレオチド、修飾ヌクレオチド、またはこれらの組み合わせを含むことができる。アダプターオリゴヌクレオチドは、一本鎖、二本鎖、または部分的二本鎖であり得る。一般に、部分的二本鎖アダプターは、1つ以上の一本鎖領域および1つ以上の二本鎖領域を含む。二本鎖アダプターは、互いにハイブリダイズされた2つの別個のオリゴヌクレオチドを含むことができ(「オリゴヌクレオチド二本鎖」とも称される)、ハイブリダイゼーションは、1つ以上の平滑末端、1つ以上の3'オーバーハング、1つ以上の5'オーバーハング、ミスマッチおよび/または非対合ヌクレオチドから生じる1つ以上の隆起、またはこれらの任意の組み合わせを残し得る。いくつかの実施形態において、一本鎖アダプターは、互いにハイブリダイズすることができる2つ以上の配列を含む。2つのそのようなハイブリダイズ可能な配列が、一本鎖アダプターに含有されるとき、ハイブリダイゼーションは、ヘアピン構造を(ヘアピ

10

20

30

40

50

ンアダプター)を生じる。アダプターの2つのハイブリダイズされた領域が、非ハイブリダイズ領域により互いから分離されるとき、「バブル」構造が生じる。バブル構造を含むアダプターは、内部ハイブリダイゼーションを含む単一アダプターオリゴヌクレオチドからなり得るか、または互いにハイブリダイズされた2つ以上のアダプターオリゴヌクレオチドを含み得る。例えば、アダプター中の2つのハイブリダイズ可能な配列間の部配列ハイブリダイゼーションは、一本鎖アダプターオリゴヌクレオチドにおいて二本鎖構造を生成することができる。異なる種類のアダプターは、ヘアピンアダプターおよび二本鎖アダプター、または異なる配列のアダプター等の組み合わせで使用することができる。異なるアダプターは、連続反応において、または同時に標的ポリヌクレオチドに連結され得る。いくつかの実施形態において、同一のアダプターが、標的ポリヌクレオチドの両端に付加される。例えば、第1および第2のアダプターを、同じ反応に付加することができる。アダプターを、標的ポリヌクレオチドと合わせる前に操作することができる。例えば、末端リン酸塩を付加または除去することができる。

【0056】

いくつかの実施形態において、アダプターは、2本の鎖がアニールされるとき、少なくとも1つの二本鎖領域および少なくとも1つの不適合領域を提供するように、2つの部分的に相補性のポリヌクレオチド鎖をアニールすることにより形成されたミスマッチアダプターである。アダプターの「二本鎖領域」は、短い二本鎖領域であり、典型的に、2つの部分的に相補性のポリヌクレオチド鎖のアニールリングにより形成された5個以上の連続塩基対を含む。この用語は、単に二本鎖がアニールされる核酸の二本鎖領域を指し、いかなる特定の構造配置も暗示するものではない。いくつかの実施形態において、二本鎖領域は、約5、10、15、20、25、30、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長である。一般に、ミスマッチアダプターの二本鎖領域が、機能を喪失することなく可能な限り短いことが有益である。この文脈において、「機能」とは、二本鎖領域が、酵素触媒された核酸ライゲーション反応のための標準反応条件下で安定した二本鎖を形成することを意味し、これらの条件は、当業者に既知であり(例えば、酵素に適切なライゲーション緩衝液中4 ~ 25 の範囲の温度でのインキュベーション)、アダプターを形成する2つの鎖が、標的分子へのアダプターのライゲーションの間、部分的にアニールされたままであるようにする。二本鎖領域が、典型的にプライマー伸長またはPCR反応のアニールリングステップにおいて使用される条件下で安定することは、必ずしも必要とは限らない。典型的に、二本鎖領域は、アダプターの「ライゲーション可能な」末端、すなわち、ライゲーション反応において標的ポリヌクレオチドに連結される末端に隣接する。アダプターのライゲーション可能な末端は、平滑であるか、または他の実施形態において短いことがある。1つ以上のヌクレオチドの5' または3' オーバーハングは、ライゲーションを平易化/促進するように存在し得る。アダプターのライゲーション可能な末端における5' 末端ヌクレオチドは、典型的に、試料ポリヌクレオチド上の3' ヒドロキシル基へのホスホジエステル結合を可能にするようにリン酸化される。「不適合領域」という用語は、アダプターを形成する2つのポリヌクレオチド鎖の配列が、非相補性の程度を呈し、2つの鎖がプライマー伸長またはPCR反応のための標準アニール条件下で互いにアニールできないようになるアダプターの領域を指す。不適合領域における2本の鎖は、酵素触媒されたライゲーション反応の標準反応条件下で、ある程度のアニールを呈し得るが、但し、2つの鎖は、アニール条件下で一本鎖形態に戻る。

【0057】

アダプターオリゴヌクレオチドは、多様な配列要素のうちの1つ以上を含有することができ、1つ以上の増幅プライマーアニールリング配列またはその相補体、1つ以上の配列決定プライマーアニールリング配列またはその相補体、1つ以上のバーコード配列、複数の異なるアダプターまたは異なるアダプターのサブセットの中で共有される1つ以上の共通配列、1つ以上の制限酵素認識部位、1つ以上の標的ポリヌクレオチドオーバーハングに相補性である1つ以上のオーバーハング、1つ以上のプローブ結合部位(例えば、配列決定

10

20

30

40

50

プラットフォーム、例えば、本明細書に記載される装置等の超並列配列決定のためのフローセル、またはIllumina, Inc.により開発されたフローセルへの結合のため)、1つ以上のランダムまたはニアランダム配列(例えば、1つ以上の位置にある2つ以上の異なるヌクレオチドの群からランダムに選択される1つ以上のヌクレオチド、異なるヌクレオチドのそれぞれは、ランダム配列を含むアダプターのプールに表される1つ以上の位置で選択される)、およびこれらの組み合わせが挙げられるが、これらに限定されない。2つ以上の配列要素は、互いに非隣接であり得るか(例えば、1つ以上のヌクレオチドにより分離される)、互いに隣接し得るか、部分的に重なり得るか、または完全に重なり得る。例えば、増幅プライマーアニーリング配列は、配列決定プライマーアニーリング配列としても機能し得る。配列要素は、3'末端もしくはその付近、5'末端もしくはその付近、またはアダプターオリゴヌクレオチドの内側に位置し得る。アダプターオリゴヌクレオチドが、ヘアピン等の二次構造を形成することができるとき、配列要素は、部分的もしくは完全に二次構造の外側に位置し得るか、部分的もしくは完全に二次構造の内側に位置し得るか、または二次構造に関与する配列間に位置し得る。配列要素は、任意の好適な長さ、例えば、約3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長であってよい。アダプターオリゴヌクレオチドは、それらが含まれる1つ以上の配列要素を収容するのに少なくとも十分な任意の好適な長さを有することができる。いくつかの実施形態において、アダプターは、約10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、200、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長である。

【0058】

いくつかの実施形態において、1つの試料からの断片化ポリヌクレオチドに連結されたアダプターオリゴヌクレオチドは、全てのアダプターオリゴヌクレオチドに共通の1つ以上の配列と、その特定の試料のポリヌクレオチドに連結されたアダプターに一意のバーコードとを含み、バーコード配列を使用して、1つの試料またはアダプター連結反応から生じるポリヌクレオチドを、別の試料またはアダプター連結反応から生じるポリヌクレオチドと区別できるようにする。いくつかの実施形態において、アダプターオリゴヌクレオチドは、1つ以上の標的ポリヌクレオチドオーバーハングに相補性である5'オーバーハング、3'オーバーハング、または両方を含む。相補的オーバーハングは、1つ以上のヌクレオチド長であり得、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、またはそれより多くのヌクレオチド長が挙げられるが、これらに限定されない。相補的オーバーハングは、固定配列を含み得る。アダプターオリゴヌクレオチドの相補的オーバーハングは、1つ以上のヌクレオチドのランダム配列を含み得、1つ以上のヌクレオチドが、1つ以上の位置にある2つ以上の異なるヌクレオチドの群からランダムに選択され、異なるヌクレオチドのそれぞれが、ランダム配列を含む相補的オーバーハングを持つアダプターのプールに表される1つ以上の位置で選択されるようになる。いくつかの実施形態において、アダプターオーバーハングは、制限エンドヌクレアーゼ消化により生成された標的ポリヌクレオチドオーバーハングに相補性である。いくつかの実施形態において、アダプターオーバーハングは、アデニンまたはチミンからなる。

【0059】

いくつかの実施形態において、アダプターオリゴヌクレオチドは、配列要素配列Dを含む1本の鎖を含む。いくつかの実施形態において、アダプターオリゴヌクレオチドは、相補配列D'にハイブリダイズされた配列Dを含み、配列D'は、配列Dと同じまたは異なる鎖上にある。いくつかの実施形態において、標的ポリヌクレオチドの3'末端は、アダプターオリゴヌクレオチドに沿って伸長され、相補配列D'を生成する。好ましい実施形態において、断片化ポリヌクレオチドおよびアダプターオリゴヌクレオチドは、両末端でアダプターオリゴヌクレオチド配列に連結された断片化ポリヌクレオチド配列を含む二本鎖の適合したポリヌクレオチドを生成するように組み合わせられ、(例えば、ライゲーションにより、および任意に断片伸長により)処理され、適合したポリヌクレオチドの両末端

は、配列 D にハイブリダイズされた配列 D を含む。いくつかの実施形態において、アダプター連結に供される断片化ポリヌクレオチドの量は、約 50 ng、100 ng、200 ng、300 ng、400 ng、500 ng、600 ng、700 ng、800 ng、900 ng、1000 ng、1500 ng、2000 ng、2500 ng、5000 ng、10 µg、もしくはそれより多いか（例えば、閾値量）、それ未満、またはそれ以上である。いくつかの実施形態において、断片化ポリヌクレオチドの量は、アダプター連結を進める前に決定され、アダプター連結は、その量が閾値量を下回る場合は行われない。

【0060】

「連結」および「ライゲーション」という用語は、アダプターオリゴヌクレオチドおよび試料ポリヌクレオチドのような2つのポリヌクレオチドに関して本明細書において使用される場合、連続する骨格を持つ単一のより大きなポリヌクレオチドを生成するための2つの別個のポリヌクレオチドの共有結合を指す。2つのポリヌクレオチドを連結するための方法は、当該技術分野において既知であり、酵素的および非酵素的（例えば、化学的）方法が挙げられるが、これらに限定されない。非酵素的であるライゲーション反応の例としては、参照により本明細書に組み込まれる、米国特許第5,780,613号および同第5,476,930号に記載される非酵素的ライゲーション技法が挙げられる。いくつかの実施形態において、アダプターオリゴヌクレオチドは、リガーゼ、例えば、DNAリガーゼまたはRNAリガーゼにより断片化ポリヌクレオチドに連結される。それぞれが特徴付けられた反応条件を有する複数のリガーゼは、当該技術分野において既知であり、限定されないが、tRNAリガーゼ、Taq DNAリガーゼ、好熱菌（*Thermus filiformis*）DNAリガーゼ、大腸菌（*Escherichia coli*）DNAリガーゼ、Tth DNAリガーゼ、*Thermus scotoductus* DNAリガーゼ（IおよびII）、熱安定リガーゼ、Ampligase 熱安定DNAリガーゼ、Vanc型リガーゼ、9°N DNAリガーゼ、Tsp DNAリガーゼ、および生物資源調査により発見された新規リガーゼを含むNAD⁺依存性リガーゼ；T4 RNAリガーゼ、T4 DNAリガーゼ、T3 DNAリガーゼ、T7 DNAリガーゼ、Pfu DNAリガーゼ、DNAリガーゼ1、DNAリガーゼIII、DNAリガーゼIV、および生物資源調査により発見された新規リガーゼを含む、ATP依存性リガーゼ；ならびにそれらの野生型、変異体アイソフォーム、および遺伝子操作された異型が挙げられる。ライゲーションは、相補性オーバーハングのようなハイブリダイズ可能な配列を有するポリヌクレオチドの間にあり得る。ライゲーションは、2つの平滑末端の間にもあり得る。一般に、5'リン酸塩は、ライゲーション反応において利用される。5'リン酸塩は、断片化ポリヌクレオチド、アダプターオリゴヌクレオチド、または両方により提供され得る。5'リン酸塩は、必要に応じて、連結されるポリヌクレオチドに付加され得るか、またはそれから除去され得る。5'リン酸塩の付加または除去のための方法は、当該技術分野において既知であり、限定されないが、酵素的および化学的プロセスが挙げられる。5'リン酸塩の付加および/または除去に有用な酵素としては、キナーゼ、ホスファターゼ、およびポリメラーゼが挙げられる。いくつかの実施形態において、ライゲーション反応において連結される2つの末端の両方（例えば、アダプター末端および断片化ポリヌクレオチド末端）は、2つの共有結合が2つの末端を連結する際に、断片化ポリヌクレオチドの片方または両端で行われるように、5'リン酸塩を提供する。いくつかの実施形態において、3'リン酸塩は、ライゲーションの前に除去される。いくつかの実施形態において、アダプターオリゴヌクレオチドは、断片化ポリヌクレオチドの両端に付加され、各末端の片方または両方の鎖は、1つ以上のアダプターオリゴヌクレオチドに連結される。いくつかの実施形態において、別個のライゲーション反応は、各試料に少なくとも1つの異なるバーコード配列を含む異なるアダプターオリゴヌクレオチドを使用する異なる試料に対して実行され、どのバーコード配列も並行して分析される複数の試料の標的ポリヌクレオチドに連結されないようになる。

【0061】

アダプターオリゴヌクレオチドの非限定的な例としては、C A C T C A G C A G C A C

G A C G A T C A C A G A T G T G T A T A A G A G A C A G T (配列番号 17) を G R G A G T C G T C G T G C T G C T A G T G T C T A C A C A T A T T C T C T G T C (配列番号 18) にハイブリダイズすることにより形成された二本鎖アダプターが挙げられる。アダプターオリゴヌクレオチドの追加の非限定的な例は、参照により本明細書に組み込まれる、米国公開第 20110319290 号および同第 20070128624 号に記載される。

【0062】

いくつかの実施形態において、適合したポリヌクレオチドは、試料中の標的ポリヌクレオチドを増幅する増幅反応に供される。いくつかの実施形態において、増幅は、配列 C、配列 D、および試料と関連付けられるバーコードを含むプライマーを使用し、配列 D は、増幅プライマーの 3' 末端に位置付けられる。増幅プライマーは、約 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、もしくはそれより多くのヌクレオチド、それ未満、またはそれ以上の任意の好適な長さであり得、それらの任意の部分または全ては、プライマーがハイブリダイズする対応する標的配列に相補性であり得る（例えば、約 5、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド）。「増幅」は、標的配列のコピー数を増加させる任意のプロセスを指す。標的ポリヌクレオチドのプライマー配向増幅のための方法は、当該技術分野において既知であり、限定されないが、ポリメラーゼ連鎖反応（PCR）に基づく方法が挙げられる。PCR による標的配列の増幅に好ましい条件は、当該技術分野において既知であり、プロセス中の多様なステップにおいて最適化されることができ、例えば、標的型、標的濃度、増幅される配列長、標的および / または 1 つ以上のプライマーの配列、プライマー長、プライマー濃度、使用されるポリメラーゼ、反応量、1 つ以上の要素と 1 つ以上の他の要素との比、およびその他の反応における要素の特徴に依存し、それらのうちのいくつか、または全ては変更することができる。一般に、PCR は、増幅される標的の変性（二本鎖の場合）、標的に対する 1 つ以上のプライマーのハイブリダイゼーション、および DNA ポリメラーゼによるプライマーの伸長というステップを必要とし、これらのステップは、標的配列を増幅するために反復（または「サイクル化」）する。このプロセスにおけるステップは、収率を増強する、疑似生成物の形成を減少させる、および / またはプライマーアニーリングの特異性を増減させる等の様々な結果に対して最適化することができる。最適化の方法は、当該技術分野において周知であり、増幅反応における要素の種類もしくは量、および / またはプロセスにおける所与のステップの条件（例えば、特定のステップでの温度、特定のステップの期間、および / またはサイクル数）に対する調整を含む。いくつかの実施形態において、増幅反応は、少なくとも 5、10、15、20、25、30、35、50、またはそれより多くのサイクルを含む。いくつかの実施形態において、増幅反応は、5、10、15、20、25、35、50、またはそれより多くを超えないサイクルを含む。サイクルは、任意の数のステップ、例えば、1、2、3、4、5、6、7、8、9、10、もしくはそれより多くのステップを含むことができる。ステップは、鎖変性、プライマーアニーリング、およびプライマー伸長を含むが、これらに限定されない所与のステップの目的を達成するのに好適な任意の温度または温度の勾配を含むことができる。ステップは、手動で中断されるまで無期限に、約 1 秒、5 秒、10 秒、15 秒、20 秒、25 秒、30 秒、35 秒、40 秒、45 秒、50 秒、55 秒、60 秒、70 秒、80 秒、90 秒、100 秒、120 秒、180 秒、240 秒、300 秒、360 秒、420 秒、480 秒、540 秒、600 秒、もしくはそれより多いか、それ未満、またはそれ以上であるが、これらに限定されない任意の期間であり得る。異なるステップを含む任意の数の周期は、任意の順序で組み合わせることができる。

【0063】

いくつかの実施形態において、増幅は、増幅プライマーの 3' 末端にある配列 D と、適合したポリヌクレオチドの配列 D との間のハイブリダイゼーション、増幅プライマーから誘導された配列 D と、プライマー伸長中に生成された配列 D とを含むプライマー伸長

10

20

30

40

50

生成物を生成するための適合したポリヌクレオチドに沿った増幅プライマーの伸長を含む。いくつかの実施形態において、増幅プロセスは、プライマー伸長生成物をテンプレートポリヌクレオチドから変性させること、およびそのプライマー伸長生成物をさらなるプライマー伸長反応のテンプレートとして使用するプロセスを反復することにより、1回以上反復される。いくつかの実施形態において、プライマー伸長の第1のサイクルは、第1のプライマー伸長反応において使用されるプライマーと同じプライマーを使用して、約5、10、15、20、25、30、35、50、またはそれより多いサイクルの間に反復される。いくつかの実施形態において、増幅プライマーによる1つ以上のプライマー伸長の後に、第1の増幅プライマーを持つ増幅により適合したポリヌクレオチドに付加された配列に対して配列相補性（例えば、配列Cの相補体またはその一部分に相補性）を含む3末端を有する第2の増幅プライマーを使用する1つ以上の増幅サイクルが続く。いくつかの実施形態において、第2の増幅プライマーは、配列C、またはその一部分を3末端に含む。第2の増幅プライマーの非限定的な例は、CGAGATCTACACGCCCTCCCTCGCGCCATCAG（配列番号19）を含む。いくつかの実施形態において、第2の増幅プライマーによる増幅は、約5、10、15、20、25、30、35、50、もしくはそれより多いか、それ未満、またはそれ以上のサイクルを含む。いくつかの実施形態において、増幅に供される適合したポリヌクレオチドの量は、約50 ng、100 ng、200 ng、300 ng、400 ng、500 ng、600 ng、700 ng、800 ng、900 ng、1000 ng、1500 ng、2000 ng、2500 ng、5000 ng、10 µg、もしくはそれより多いか（例えば、閾値量）、それ未満、またはそれ以上である。いくつかの実施形態において、適合したポリヌクレオチドの量は、増幅を進める前に決定され、その量が閾値量を下回る場合、増幅は行われない。

【0064】

いくつかの実施形態において、増幅プライマーは、バーコードを含む。本明細書において使用される場合、「バーコード」という用語は、バーコードが関連付けられるポリヌクレオチドのいくつかの特徴が特定されるのを許す既知の核酸配列を指す。いくつかの実施形態において、特定されるポリヌクレオチドの特徴は、ポリヌクレオチドが誘導される試料である。いくつかの実施形態において、バーコードは、約または少なくとも約3、4、5、6、7、8、9、10、11、12、13、14、15、もしくはそれより多いヌクレオチド長である。いくつかの実施形態において、バーコードは、10、9、8、7、6、5、または4ヌクレオチド長より短い。いくつかの実施形態において、いくつかのポリヌクレオチドと関連付けられるバーコードは、他のポリヌクレオチドと関連付けられるバーコードとは異なる長さである。一般に、バーコードは十分な長さであり、それらが関連付けられるバーコードに基づく試料の特定を許すように十分に異なる配列を含む。いくつかの実施形態において、バーコード、およびそれが関連付けられる試料の供給源は、バーコード配列中の1つ以上のヌクレオチドの突然変異、挿入、または欠失後、例えば、1、2、3、4、5、6、7、8、9、10、もしくはそれより多いヌクレオチドの突然変異、挿入、または欠失後に正確に特定することができる。いくつかの実施形態において、複数のバーコード中の各バーコードは、少なくとも3つのヌクレオチド位置、例えば、少なくとも3、4、5、6、7、8、9、10、もしくはそれより多いヌクレオチド位置において複数のバーコードと1つおきに異なる。複数のバーコードは、試料のプールに表されることがあり、それぞれの試料は、そのプール中の他の試料から誘導されるポリヌクレオチドに含有されるバーコードとは異なる1つ以上のバーコードを含むポリヌクレオチドを含む。1つ以上のバーコードを含むポリヌクレオチドの試料は、それらが連結されるバーコード配列に基づいてプールされ得、ヌクレオチド塩基A、G、C、およびTの4つ全てが、プール中の各バーコードに沿って1つ以上の位置にほぼ均一に表される（例えば、1、2、3、4、5、6、7、8、もしくはそれより多い位置、またはバーコードの全ての位置）。いくつかの実施形態において、本発明の方法は、標的ポリヌクレオチドが、標的ポリヌクレオチドが連結されるバーコード配列に基づいて誘導される試料を特定することをさらに含み、一般に、バーコードは、標的ヌクレオチドに連結されるとき、標的ポリヌ

10

20

30

40

50

クレオチドが誘導された試料の識別子として機能する核酸配列を含む。

【 0 0 6 5 】

いくつかの実施形態において、別個の増幅反応は、それぞれの試料に対して少なくとも1つの異なるバーコード配列を含む増幅プライマーを使用して、別個の試料に対して実行され、いずれのバーコード配列も2つ以上の試料のプール中の複数の試料の標的ポリヌクレオチドに連結されるようになる。いくつかの実施形態において、異なる試料から誘導され、異なるバーコードを含む増幅ポリヌクレオチドが、ポリヌクレオチドの後次操作を進める前にプールされる（例えば、固体支持体上の増幅および/または配列決定前）。プールは、全体反応量を含む総構成増幅反応の任意の分画を含むことができる。試料は、均一または不均一にプールされ得る。いくつかの実施形態において、標的ポリヌクレオチドは、それらが連結されるバーコードに基づいてプールされる。プールは、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、20、25、30、40、50、75、100、もしくはそれより多いか、それ未満、またはそれ以上の異なる試料から誘導されたポリヌクレオチドを含み得る。試料は、ヌクレオチド塩基A、G、C、およびTをバーコードに沿って均一に1つ以上の位置で表すために、例えば、4、8、12、16、20、24、28、32、36、40、44、48、52、56、60、64、96、128、192、256、384等の4の倍数でプールすることができる。バーコードの非限定的な例としては、AGGTC A、CAGCAG、ACTGCT、TAA C G G、G G A T T A、A A C C T G、G C C G T T、C G T T G A、G T A A C C、C T T A A C、T G C T A A、G A T C C G、C C A G G T、T T C A G C、A T G A T C、およびT C G G A Tが挙げられる。いくつかの実施形態において、バーコードは、増幅プライマーの配列Dと配列Cの間、または配列Cおよび配列Dの後に5 から3 配向（「下流」）に位置付けられる。いくつかの実施形態において、増幅プライマーは、配列C G A G A T C T A C A C G C C T C C C T C G C G C C A T C A G X X X X X X C A C T C A G C A G C A C G A C G A T C A C（配列番号21）を含むか、またはそれからなり、それぞれの「X」は、バーコード配列のゼロ、1つ、またはそれ以上のヌクレオチドを表す。

【 0 0 6 6 】

増幅プライマーの非限定的な例を、表1に提供する。

【 0 0 6 7 】

【表1】

配列番号 1	CGAGATCTACACGCCTCCCTCGCGCCATCAGAGGTCACACTCAGCAGCACGACGATC
配列番号 2	CGAGATCTACACGCCTCCCTCGCGCCATCAGCAGCAGCACTCAGCAGCACGACGATC
配列番号 3	CGAGATCTACACGCCTCCCTCGCGCCATCAGACTGCTCACTCAGCAGCACGACGATC
配列番号 4	CGAGATCTACACGCCTCCCTCGCGCCATCAGTAACGGCACTCAGCAGCACGACGATC
配列番号 5	CGAGATCTACACGCCTCCCTCGCGCCATCAGGGATTACACTCAGCAGCACGACGATC
配列番号 6	CGAGATCTACACGCCTCCCTCGCGCCATCAGAACCTGCACTCAGCAGCACGACGATC
配列番号 7	CGAGATCTACACGCCTCCCTCGCGCCATCAGGCCGTTCACTCAGCAGCACGACGATC
配列番号 8	CGAGATCTACACGCCTCCCTCGCGCCATCAGCGTTGCACTCAGCAGCACGACGATC
配列番号 9	CGAGATCTACACGCCTCCCTCGCGCCATCAGGTAACCCACTCAGCAGCACGACGATC
配列番号 10	CGAGATCTACACGCCTCCCTCGCGCCATCAGCTTAACCACTCAGCAGCACGACGATC
配列番号 11	CGAGATCTACACGCCTCCCTCGCGCCATCAGTGCTAACACTCAGCAGCACGACGATC
配列番号 12	CGAGATCTACACGCCTCCCTCGCGCCATCAGGATCCGCACTCAGCAGCACGACGATC
配列番号 13	CGAGATCTACACGCCTCCCTCGCGCCATCAGCCAGGTCCTCAGCAGCACGACGATC
配列番号 14	CGAGATCTACACGCCTCCCTCGCGCCATCAGTTCAGCCACTCAGCAGCACGACGATC
配列番号 15	CGAGATCTACACGCCTCCCTCGCGCCATCAGATGATCCACTCAGCAGCACGACGATC
配列番号 16	CGAGATCTACACGCCTCCCTCGCGCCATCAGTCGGATCACTCAGCAGCACGACGATC

10

20

30

40

50

【 0 0 6 8 】

いくつかの実施形態において、標的ポリヌクレオチドは、本明細書に記載される任意の装置のような固体支持体に結合された複数のオリゴヌクレオチドにハイブリダイズされる。ハイブリダイゼーションは、アダプター連結および増幅のような1つ以上の試料処理ステップの前または後であり得る。好ましい実施形態において、標的ポリヌクレオチドは、アダプター連結および1つ以上の増幅反応の両方の後に固体支持体上のオリゴヌクレオチドにハイブリダイズされる。固体支持体上のオリゴヌクレオチドは、ランダムポリヌクレオチド配列、複数の異なる標的ポリヌクレオチドに共通する特定配列（例えば、配列D、D₁、もしくはその一部分等のアダプターオリゴヌクレオチドから誘導される1つ以上の配列）、増幅プライマーから誘導される1つ以上の配列（例えば、配列C、C₁、もしくはその一部分、またはこれらの組み合わせ）、異なる標的ポリヌクレオチドに特定配列（例えば、本明細書に記載される配列Bにより表される）、またはこれらの組み合わせにハイブリダイズすることができる。いくつかの実施形態において、固体支持体は、配列Aおよび配列Bを含む複数の異なる第1のオリゴヌクレオチドを含み、配列Aは、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bは、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にある。いくつかの実施形態において、複数の第1のオリゴヌクレオチドは、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の異なるオリゴヌクレオチドを含み、それぞれが、異なる配列Bを含む。いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列Bは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む（例えば、1、5、10、25、50、75、または100個の異なるオリゴヌクレオチドは、それぞれが図4-1~3からの異なる配列を持つ）。いくつかの実施形態において、配列B、またはそれが特異的にハイブリダイズする標的配列は、本明細書に記載されるように、原因となる遺伝的変異体を含む。いくつかの実施形態において、配列Bまたはそれが特異的にハイブリダイズする標的配列は、本明細書に記載されるように、原因となる遺伝的変異体の約1、2、3、4、5、6、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、またはそれより多いヌクレオチド以内にある。原因となる遺伝的変異体は、典型的に、第1のオリゴヌクレオチドの下流に位置し、原因となる遺伝的変異体の少なくとも一部分が、第1のオリゴヌクレオチドの伸長のためのテンプレートとして機能するようになる。固体支持体は、本明細書に記載されるように、それぞれの第2のオリゴヌクレオチドの3'末端に配列Aを含む複数の第2のオリゴヌクレオチドと、それぞれの第3のオリゴヌクレオチドの3'末端に配列Cを含む複数の第3のオリゴヌクレオチドと、をさらに含むことができる。

【 0 0 6 9 】

いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列B、またはそれが特異的にハイブリダイズする標的配列は、非対象配列を含む。いくつかの実施形態において、配列B、またはそれが特異的にハイブリダイズする標的配列は、非対象配列の約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。一般に、非対象配列は、試験される個体、または試料が採取された個体以外の生物から誘導されたポリヌクレオチド、例えば、細菌、古細菌、ウイルス、原生生物、真菌、または他の生物からのDNAまたはRNAに対応する。非対象配列は、胎児からの核酸、例えば、胎児からの無細胞核酸（細胞外核酸とも称される）を含むこともできる。非対象配列は、生物の同一性を示し得るか、またはさらに感染等の疾患状態を示し得る。生物を特定する際に有用な非対象配列の例としては、限定されないが、rRNA配列、例えば、16S rRNA配列が挙げられる（例えば、国際公開第2010151842号を参照）。いくつかの実施形

10

20

30

40

50

態において、非対象配列は、原因となる遺伝的変異体の代わりに、またはそれとは別に分析される。いくつかの実施形態において、原因となる遺伝的変異体および非対象配列は、例えば、同じ試料中（例えば、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つものと、非対象配列を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つものとの第 1 のオリゴヌクレオチドの混合物を使用する）および / または同じレポートにおいて並行して分析される。

【 0 0 7 0 】

いくつかの実施形態において、複数の第 1 のオリゴヌクレオチドのうちの 1 つ以上の配列 B、またはそれが特異的にハイブリダイズする標的配列は、祖先情報提供マーカー（A I M）を含む。いくつかの実施形態において、配列 B またはそれが特異的にハイブリダイズする標的配列は、A I M の約 1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にあり、A I M は、原因となる遺伝的変異体のうちの 1 つに対して危険性が高い集団のような 1 つ以上の集団に属するか、または属しない人を分類するために使用され得る。例えば、A I M は、ある形質の保有率が高い集団に対して特徴的であり得る。場合によっては、A I M は、より細かい精度で集団、例えば、亜大陸グループまたは関連する民族グループを区別する。いくつかの実施形態において、A I M は、原因となる遺伝的変異体および / または非対象配列の代わりに、またはそれとは別に分析される。いくつかの実施形態において、A I M、原因となる遺伝的変異体、および / または非対象配列は、同じ試料中（例えば、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つものと、A I M を含むか、またはその付近にある配列に特異的にハイブリダイズする配列 B を持つものとの第 1 のオリゴヌクレオチドの混合物を使用する）および / または同じレポートにおいて並行して分析される。

【 0 0 7 1 】

いくつかの実施形態において、この方法は、固体支持体上で架橋増幅を行うことをさらに含む。一般に、架橋増幅は、テンプレートへのプライマーのアニールリング、プライマー伸長、およびテンプレートからの伸長プライマーの分離の反復ステップを使用する。これらのステップは、一般に、P C R（または逆転写酵素 + P C R）技法において当業者に既知の試薬および条件を使用して行うことができる。したがって、核酸ポリメラーゼは、好適なテンプレートの存在下で、プライマーを伸長するようにヌクレオシド三リン酸塩分子（または D N A / R N A に存在するヌクレオチドの前駆体として機能する他の分子、例えば、修飾ヌクレオシド三リン酸塩）の供給と一緒に使用することができる。望ましくは、過剰なデオキシリボヌクレオシド三リン酸塩が望ましくは提供される。好ましいデオキシリボヌクレオシド三リン酸塩は、d T T P（デオキシチミジンヌクレオシド三リン酸塩）、d A T P（デオキシアデノシンヌクレオシド三リン酸塩）、d C T P（デオキシシトシンヌクレオシド三リン酸塩）、および d G T P（デオキシグアノシンヌクレオシド三リン酸塩）と省略される。好ましいリボヌクレオシド三リン酸塩は、U T P、A T P、C T P、および G T P である。しかしながら、代替が可能である。これらは、天然に存在し得るか、または非天然に存在し得る。一般に P C R 反応に使用される種類の緩衝液が提供されてもよい。プライマー伸長中にヌクレオチドを組み込むために使用される核酸ポリメラーゼは、好ましくは、それが数回使用され得るために利用される反応条件下で安定している。したがって、加熱を使用して新たに合成された核酸鎖をそのテンプレートから分離する場合、核酸ポリメラーゼは、好ましくは使用される温度で熱安定性である。そのような熱安定性ポリメラーゼは、当業者に既知である。それらは、好熱性微生物から得られ、T a q ポリメラーゼとして知られる D N A 依存性 D N A ポリメラーゼ、およびその熱安定性誘導体も含む。

【 0 0 7 2 】

典型的に、プライマーのそのテンプレートへのアニールリングは、25 ~ 90 の温度

10

20

30

40

50

で行う。この範囲の温度は、典型的に、プライマー伸長中にも使用され、アニーリングおよび/または変性中に使用される温度と同じであり得るか、または異なり得る。アニーリングを許し、また所望の程度のプライマー伸長が起こるのを許すのに十分な時間が経過した後、所望される場合、鎖分離を許すように温度を高めることができる。この段階で、温度は、典型的に、60 ~ 100 の温度に高められる。アニーリングの前に非特異的プライミング問題を低減するため、および/または増幅開始のタイミングを制御するため、例えば、多数の試料の増幅開始を同期させるために、高い温度を使用することもできる。代替として、低塩および高pH(12超)の溶液での処理によるか、またはカオトロピック塩(例えば、グアニジン塩酸塩)を使用することによるか、または有機溶媒(例えば、ホルムアミド)により、鎖が分離され得る。

10

【0073】

鎖分離(例えば、加熱による)に続いて、洗浄ステップが行われ得る。この洗浄ステップは、例えば、同じテンプレートを固定プライマーの近接に維持することが所望される場合に、アニーリング、プライマー伸長、および鎖分離の初期段階の間で省略され得る。これは、コロニー形成を開始するためにテンプレートが数回使用されるのを許す。固体支持体上の増幅により生成されるコロニーのサイズは、例えば、アニーリング、プライマー伸長、および鎖分離が起こるサイクル数を制御することにより制御することができる。コロニーのサイズに作用する因子を制御することもできる。これらは、固定プライマーの表面上の数および配置、プライマーが固定される支持体の構成、テンプレートおよび/またはプライマー分子の長さおよび剛性、温度、ならびに上述のサイクルが行われ得る流体のイオン強度および粘度を含む。

20

【0074】

本発明の方法に従う増幅プロセスの非限定的な例が、図1に示され、以下に記載される。第1に、固体支持体に結合され、配列Bをその3'末端に含む第1のオリゴヌクレオチドは、相補性標的配列B'、例えば、複数の異なる標的ポリヌクレオチド(例えば、特定のゲノムDNA配列)中の特定の標的ポリヌクレオチドに固有の配列にハイブリダイズする。このようにして、配列Bはプローブとして機能する。図1の標的ポリヌクレオチドは、アダプターオリゴヌクレオチド(例えば、配列DおよびD')および増幅プライマー(例えば、CおよびC')から誘導される配列を含む。第1のオリゴヌクレオチドの伸長は、固体支持体に結合された第1の伸長を生成し、第1の伸長生成物は、5'から3'に配列A、B、C'、およびD'を含み、配列C'は、配列Cに相補性であり、配列D'は、配列Dに相補性である。次に、第1の伸長生成物は、標的ポリヌクレオチドテンプレートから(例えば、熱または化学変性により)分離される。次に、第1の伸長生成物の配列C'は、固体支持体に結合された複数の第3のオリゴヌクレオチドのうちの1つにハイブリダイズし、第3のオリゴヌクレオチドは、配列Cをその3'末端に含む。第3のオリゴヌクレオチドの伸長は、固体支持体に結合された第2の伸長生成物を生成し、第2の伸長生成物は、5'から3'に配列C、D、B'、およびA'を含み、配列B'は、配列Bに相補性であり、配列A'は、配列Aに相補性である。2つの伸長生成物は、二本鎖ポリヌクレオチド「架橋」を形成し、一本の鎖が両端で固体支持体に結合される。次に、第1および第2の伸長生成物は変性され、伸長生成物と他のオリゴヌクレオチドとの間の部分配列ハイブリダイゼーションに続く伸長は、第1および第2の伸長生成物を複製する。例えば、それぞれの第1の伸長生成物は、第2の伸長生成物の追加のコピーを生成するように、さらなる第3のオリゴヌクレオチドにハイブリダイズすることができる。さらに、第2の伸長生成物は、固体支持体に結合された複数の第2のオリゴヌクレオチドのうちの1つにハイブリダイズすることができ、第2のオリゴヌクレオチドは、配列Aをその3'末端に含む。第2のオリゴヌクレオチドの伸長は、第1の伸長生成物の配列を含む伸長生成物を生成する。伸長生成物に沿った一連の伸長は、最初の第1の伸長生成物から放射状に外に広がり、第1の伸長生成物のクラスタまたは「コロニー」および単一の標的ポリヌクレオチドから誘導されるそれらの相補性第2の伸長生成物を生成する。このプロセスは、異なる配列または配列配置を含むオリゴヌクレオチド、異なる標的ポリヌクレオチドまたは標

30

40

50

的ポリヌクレオチドの組み合わせ、固体支持体の種類、および特定の架橋増幅反応に依存する他の考慮事項に対応するように修飾され得る。一般に、このプロセスは、標的ポリヌクレオチドおよび非標的ポリヌクレオチドを含む試料ポリヌクレオチドからの特定の標的ポリヌクレオチドの固体支持体上の増幅を提供する。一般に、標的ポリヌクレオチドは、選択的に増幅するが、試料中の非標的ポリヌクレオチドは、増幅しないか、またははるかに低い程度、例えば、1つ以上の標的ポリヌクレオチドより約10倍、100倍、500倍、1000倍、2500倍、5000倍、10000倍、25000倍、50000倍、100000倍、1000000倍、もしくはそれより多いか、それ未満だけ低く増幅する。

【0075】

いくつかの実施形態において、架橋増幅に供された以前の増幅ステップから増幅したポリヌクレオチドの量は、約50 ng、100 ng、500 ng、1 μg、2 μg、3 μg、4 μg、5 μg、6 μg、7 μg、8 μg、9 μg、10 μg、11 μg、12 μg、13 μg、14 μg、15 μg、20 μg、25 μg、26 μg、27 μg、28 μg、29 μg、30 μg、40 μg、50 μg、もしくはそれより多いか、それ未満、またはそれ以上である（例えば、閾値量）。いくつかの実施形態において、以前の増幅ステップから増幅したポリヌクレオチドの量は、架橋増幅を進める前に決定され、その量が閾値量を下回る場合、架橋増幅は行われない。

【0076】

いくつかの実施形態において、架橋増幅に続いて、固体支持体に結合された複数のオリゴヌクレオチドを配列決定する。試薬および反応条件を含む、固体支持体に結合されたポリヌクレオチドを配列決定するための一般的な方法は、当該技術分野において既知である。いくつかの実施形態において、配列決定は、単一末端配列決定を含むか、またはそれからなる。いくつかの実施形態において、配列決定は、対合末端配列決定を含むか、またはそれからなる。配列決定は、任意の好適な配列決定技法を使用して実行することができ、ヌクレオチドは、連続して遊離3'ヒドロキシル基に付加され、5'から3'配向のポリヌクレオチド鎖の合成をもたらす。付加されるヌクレオチドの同一性は、好ましくは、それぞれのヌクレオチド付加後に決定される。全ての隣接する塩基が配列決定されるとは限らないライゲーションによる配列決定を使用する配列決定技法、および塩基が表面上の鎖に付加されるのではなく、そこから除去される超並列シグネチャー配列決定（MPSS）もまた、ピロリン酸放出（ピロ配列決定）の検出を使用する技術であるため、本発明の範囲内である。そのようなピロ配列決定に基づく技法は、特に、ビーズがライブラリー分子からの単一テンプレートが、それぞれのビーズ上で増幅されるように乳剤中で増幅されている、ビーズの配列決定アレイに適用可能である。

【0077】

本発明の方法で使用することができる1つの特定配列決定方法は、可逆的連鎖停止剤として作用することができる修飾ヌクレオチドの使用に依存する。そのような可逆的連鎖停止剤は、例えば、国際公開第04018497号および米国特許第7057026号に記載されるように、除去可能な3'保護基を含む。そのような修飾ヌクレオチドが、配列決定されるテンプレートの領域に相補性の成長ポリヌクレオチド鎖に組み込まれると、さらなる配列伸長を配向するために使用できる遊離3'-OH基は存在しないため、ポリメラーゼは、さらなるヌクレオチドを付加することができない。成長鎖に組み込まれる塩基の同一性が決定されると、3'保護基は、次の連続するヌクレオチドの付加を許すように除去され得る。これらの修飾ヌクレオチドを使用して誘導される生成物を順序付けることにより、DNAテンプレートのDNA配列を推測することが可能である。そのような反応は、修飾ヌクレオチドのそれぞれが、特定の塩基に対応することが知られている異なる標識に結合した場合、修飾ヌクレオチドのそれぞれが単一実験において行い、それぞれの統合ステップにおいて付加された塩基間の区別を容易にすることができる。好適な標識の非限定的な例は、国際公開第2007/135368号に記載され、その内容は、参照によりそれら全体が本明細書に組み込まれる。代替として、個別に付加された修飾ヌクレオチド

10

20

30

40

50

のそれぞれを含有する別個の反応を実行することができる。

【0078】

修飾ヌクレオチドは、それらの検出を容易にするための標識を担持し得る。特定の実施形態において、標識は蛍光標識である。それぞれのヌクレオチド型は、異なる蛍光標識を担持し得る。しかしながら、検出可能な標識は、蛍光標識である必要はない。DNA配列へのヌクレオチドの組み込みの検出を許す任意の標識を使用することができる。蛍光標識されたヌクレオチドを検出するための方法は、標識ヌクレオチドに特異的な波長のレーザー光線を使用すること、または他の好適な光源の使用を含む。組み込まれたヌクレオチド上の標識からの蛍光は、CCDカメラまたは他の好適な検出手段により検出され得る。好適な検出手段は、国際公開第2007/123744号に記載され、その内容は、参照によりそれら全体が本明細書に組み込まれる。

10

【0079】

いくつかの実施形態において、第1の配列決定反応は、固体支持体に結合したオリゴヌクレオチドに含有される切断部位での切断により形成された3'末端から進行し、そのオリゴヌクレオチドは、架橋増幅中に伸長された。いくつかの実施形態において、切断鎖は、結合したオリゴヌクレオチドの伸長により配列決定する前に、その相補鎖から分離される。いくつかの実施形態において、切断により形成された新たに遊離された3'末端を有する結合したオリゴヌクレオチドは、鎖置換活性を有するポリメラーゼを使用して伸長され、切断した鎖は、新たな鎖が伸長されると置換される。いくつかの実施形態において、結合したオリゴヌクレオチドの伸長は、増幅反応からのテンプレート伸長生成物の全長に沿って進み、いくつかの実施形態において、最後に特定されたヌクレオチドを越える伸長を含む。いくつかの実施形態において、次にテンプレート伸長生成物は、固体支持体に結合したオリゴヌクレオチドに含有される切断部位で切断され、配列決定反応中に伸長したオリゴヌクレオチドは、遊離された第1の配列決定伸長生成物を生成するために線形化される。次に、第1の配列決定生成物の5'末端は、配列決定プライマー（例えば、本明細書に記載される配列決定プライマー）の伸長によるか、または切断部位での切断により形成される3'末端からの伸長により進めることができる、第2の配列決定反応のテンプレートとして機能し得る。いくつかの実施形態において、配列決定されるテンプレートポリヌクレオチドに沿って特定されたヌクレオチド数の平均または中央値は、約10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、300、400、500、もしくはそれより多いか、それ未満、またはそれ以上である。

20

30

【0080】

いくつかの実施形態において、配列決定は、少なくとも部分的に一本鎖であるテンプレートを生成するために、実質的に全てを除去するか、または「架橋」構造における固定鎖のうちの1つの少なくとも一部分を除去もしくは置換するように、架橋増幅生成物を処理することを含む。したがって、一本鎖であるテンプレートの一部分は、配列決定プライマーを用いるハイブリダイゼーションに使用可能である。架橋二本鎖核酸構造における1つの固定鎖の全てまたは一部分を除去するプロセスは、本明細書において「線形化」と称されることがあり、国際公開第07010251号にさらに詳細に記載され、その内容は、参照によりそれら全体が本明細書に組み込まれる。

40

【0081】

架橋テンプレート構造は、制限エンドヌクレアーゼを用いた片方もしくは両方の鎖の切断によるか、または切断エンドヌクレアーゼを用いた1つの鎖の切断により線形化され得る。切断の他の方法は、制限酵素または切断酵素の代替として使用することができ、化学切断（例えば、過ヨウ素を用いるジオール結合の切断）、エンドヌクレアーゼを用いた切断による脱塩基部位の切断（例えば、NEBにより供給される「USER」、パーツ番号M5505S）、熱またはアルカリへの曝露による、それ以外ではデオキシリボヌクレオチドからなる増幅生成物に組み込まれるリボヌクレオチドの切断、光化学切断またはペプチドリナーの切断を含むが、これらに限定されない。いくつかの実施形態において、線

50

形化ステップは、例えば、共有結合で固定化した1つのみの増幅オリゴヌクレオチドを用いて、別の増幅オリゴヌクレオチドを溶液中に含まないで固相増幅反応が行われるときに回避され得る。切断ステップに続いて、切断に使用される方法に関係なく、切断反応の生成物は、固体支持体に結合されない切断鎖（複数可）の一部分（複数可）を除去するために変性条件に供され得る。好適な変性条件は、例えば、水酸化ナトリウム溶液、ホルムアミド溶液、または熱は、例えば、標準分子生物学プロトコルに記載されるように、当該技術分野において既知である（Sambrook et al., 2001, Molecular Cloning, A Laboratory Manual, 3rd Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, NY; Current Protocols, eds Ausubel et al.）。変性は、部分的または実質的に一本鎖である配列決定テンプレートの生成をもたらす。次に、配列決定反応は、テンプレートの一本鎖部分への配列決定プライマーのハイブリダイゼーションにより開始されてよい。したがって、本発明は、核酸シーケンシング反応が、配列決定プライマーを線形化増幅生成物の一本鎖領域にハイブリダイズすることと、1つ以上のヌクレオチドを配列決定される増幅テンプレート鎖の領域に相補性のポリヌクレオチド鎖に連続的に組み込むことと、組み込まれたヌクレオチド（複数可）のうちの1つ以上に存在する塩基を特定することと、それによりテンプレート鎖の領域の配列を決定することと、を含む方法を包含する。

【0082】

いくつかの実施形態において、配列決定プライマーは、アダプターオリゴヌクレオチドから誘導された1つ以上の配列に相補性の配列、増幅プライマー、固体支持体に結合されたオリゴヌクレオチド、またはこれらの組み合わせを含む。いくつかの実施形態において、配列決定プライマーは、配列Dまたはその一部分を含む。いくつかの実施形態において、配列決定プライマーは、配列Cまたはその一部分を含む。配列決定プライマーは、任意の好適な長さ、例えば、約5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドであり得、その任意の部分または全ては、プライマーがハイブリダイズする対応する標的配列に相補性であり得る（例えば、約5、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド）。いくつかの実施形態において、配列決定プライマーは、配列C A C T C A G C A G C A C G A C G A T C A C A G A T G T G T A T A A G A G A C A G（配列番号20）を含む。

【0083】

一般に、配列決定プライマーの伸長は、配列決定伸長生成物を生成する。配列決定プロセス中で特定される配列決定伸長生成物に付加されるヌクレオチドの数は、テンプレート配列、反応条件、使用される試薬、および他の因子を含む多数の因子に依存し得る。いくつかの実施形態において、成長配列決定プライマーに沿って特定されたヌクレオチド数の平均または中央値は、約10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、300、400、500、もしくはそれより多いか、それ未満、またはそれ以上である。いくつかの実施形態において、配列決定プライマーは、増幅反応からのテンプレートプライマー伸長生成物の全長に知って伸長され、いくつかの実施形態において、最後に特定されたヌクレオチドを超える伸長を含む。

【0084】

いくつかの実施形態において、配列決定伸長生成物は、テンプレートを部分的または完全に一本鎖にし、第2の配列決定プライマーとのハイブリダイゼーションに使用可能にするために、配列決定伸長生成物を、それがハイブリダイズされる結合したテンプレート鎖から除去するように変性条件に供される。第2の配列決定プライマーは、第1の配列決定プライマーと同じであり得るか、または異なり得る。いくつかの実施形態において、第2の配列決定プライマーは、第1の配列決定プライマーがハイブリダイズする配列よりも標

的核酸の5'末端の近くに位置する配列にハイブリダイズする。いくつかの実施形態において、第2の配列決定プライマーは、第1の配列決定プライマーがハイブリダイズする配列よりも標的核酸の3'末端の近くに位置する配列にハイブリダイズする。いくつかの実施形態において、第1および第2の配列決定プライマーのうちの1つのみが、バーコード配列に沿って伸長し、それによりそのバーコード配列中のヌクレオチドを特定する。いくつかの実施形態において、1つの配列決定プライマー（例えば、第1の配列決定プライマー）は、バーコードから5'に位置する配列にハイブリダイズし（この配列決定プライマーの伸長が、バーコードに相補性の配列を生成しなくなる）、別の配列決定プライマー（例えば、第2の配列決定プライマー）は、バーコードから3'に位置する配列にハイブリダイズする（この配列決定プライマーの伸長が、バーコードに相補性の配列を生成するようになる）。いくつかの実施形態において、第2の配列決定プライマーは、配列番号19を含む。

【0085】

本発明は、本質的にポリヌクレオチド鎖へのヌクレオチドの連続組み込みに依存する任意の配列決定手法を使用することができるため、上記のような配列決定方法の使用に限定されることを意図しない。好適な技法としては、例えば、米国特許第6306597号、米国公開第20090233802号、同第20120053074号、および同第20110223601号に記載されるものが挙げられ、それらは参照によりそれら全体が組み込まれる。鎖再合成が用いられる場合において、両方の鎖は、固定化された鎖の一部分の後次解放を許すように、表面に固定化されなければならない。これは、国際公開第07010251号に記載される多数の機構を通じて達成することができ、その内容が参照によりその全体が本明細書に組み込まれる。例えば、1つのプライマーは、ウラシルヌクレオチドを含有することができ、これは、ヌクレオチド塩基を除去する酵素ウラシルDNAグリコシラーゼ（UDG）、および脱塩基性ヌクレオチドを切除するエンドヌクレアーゼVIIIIを使用して、ウラシル塩基において鎖を切断できることを意味する。この酵素の組み合わせは、New England Biolabs（NEBパーツ番号M5505）からUSER（商標）として入手可能である。第2のプライマーは、8-オキソグアニンヌクレオチドを含み得、次に、酵素FPG（NEBパーツ番号M0240）により切断可能である。プライマーのこの設計は、どのプライマーがプロセスのどの時点で切断されるか、およびクラスタのどこで切断が起こるかの完全な制御を提供する。プライマーは、例えば、特定の位置での化学的切断を許すジスルフィドまたはジオール修飾を用いて化学的に修飾され得る。

【0086】

いくつかの実施形態において、配列決定データは、単一反応容器（例えば、フローセル中のチャンネル）内の試料から約5、10、25、50、100、150、200、250、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の異なる標的ポリヌクレオチドに対して生成される。いくつかの実施形態において、配列決定データは、複数の試料、例えば、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して並行して生成される。いくつかの実施形態において、配列決定データは、単一反応容器（例えば、フローセル内のチャンネル）内の複数の試料、例えば、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して生成され、配列決定データは、配列決定されたポリヌクレオチドが起源とする試料に従い後次に分類される。単一反応において、配列決定データは、架橋増幅反応から約または少なくとも約 10^6 、 10^7 、 10^8 、 2×10^8 、 3×10^8 、 4×10^8 、 5×10^8 、 10^9 、 10^{10} 、もしくはそれより多い標的ポリヌクレオチドまたはクラスタに対して生成され得、その反応において、それぞれの試

料に対して約 10^4 、 10^5 、 10^6 、 2×10^6 、 3×10^6 、 4×10^6 、 5×10^6 、 10^7 、 10^8 、もしくはそれより多いか、それ未満、またはそれ以上の標的ポリヌクレオチドまたはクラスターの配列決定データを含み得る。いくつかの実施形態において、約 5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の原因となる遺伝的変異体の存在、非存在、または遺伝子型は、配列決定データに基づいて試料に対して決定される。1つ以上の原因となる遺伝的変異体の存在、非存在、または遺伝子型は、約 80%、85%、90%、95%、97.5%、99%、99.5%、99.9%、もしくはそれより高いか、またはそれ以上の精度で決定され得る。

10

【0087】

いくつかの実施形態において、本発明の方法におけるステップのうちの1つ以上、または全ては、例えば、1つ以上の自動デバイスの使用により自動化される。一般に、自動デバイスは、ヒトの指示なしに作動することができるデバイスであり、自動システムは、例えば、自動デバイスがさらなるヒトの操作なしに1つ以上のステップを行った後にコンピューターに命令を入力することにより、ヒトが機能を促進するように任意の動作を行った後の期間中に機能を行うことができる。本発明の実施形態を実現するコードを含むソフトウェアおよびプログラムは、CD-ROM、DVD-ROM、テープ、フラッシュドライブ、もしくはディスク、または他の適切なコンピューター可読媒体等のいくつかの種類のデータ記憶媒体上に記憶されてもよい。本発明の様々な実施形態は、ハードウェアのみで、またはソフトウェアおよびハードウェアの組み合わせで実現することもできる。例えば、一実施形態において、従来のパーソナルコンピューターではなく、プログラマブル論理制御装置(PLC)が使用される。当業者に既知の通り、PLCは、汎用コンピューターにかかる費用が不要である多様なプロセス制御アプリケーションにおいて頻繁に使用される。PLCは、1つまたは多様な制御プログラムを実行する既知の方法で構成され得、ユーザーもしくは別のデバイスからの入力を受け取ることができ、および/またはパーソナルコンピューターのそれに似た方法で、ユーザーもしくは別のデバイスに出力を提供することができる。したがって、本発明の実施形態は、汎用コンピューターに関して説明されているが、他の構成が使用され得るため、汎用コンピューターの使用は単なる例示であることを理解されたい。

20

30

【0088】

いくつかの実施形態において、自動化は、1つ以上の液体ハンドラーおよび関連ソフトウェアの使用を含み得る。いくつかの市販の液体ハンドリングシステムを利用して、これらのプロセスの自動化を行うことができる(例えば、例として、Perkin-Elmer、Beckman Coulter、Caliper Life Sciences、Tecan、Eppendorf、Apricot Design、Velocity 11からの液体ハンドラーを参照)。いくつかの実施形態において、自動ステップは、断片化、末端修復、A-テリング(アデニンオーバーハングの付加)、アダプター連結、PCR増幅、試料定量化(例えば、DNAの量および/または純度)、および配列決定のうちの1つ以上を含む。いくつかの実施形態において、固体表面に結合した増幅ポリヌクレオチドのハイブリダイゼーション、テンプレートとしての増幅ポリヌクレオチドに沿った伸長、および/または架橋増幅が、自動化される(例えば、Illumina cBotの使用により)。架橋増幅を行うためのデバイスの非限定的な例は、国際公開第2008002502号に記載される。いくつかの実施形態において、配列決定は自動化される。多様な自動配列決定機器が市販されており、Life Technologies(SOLiDプラットフォーム、およびpHに基づく検出)、Roche(454プラットフォーム)、Illumina(例えば、フローセルに基づくシステム、例えば、ゲノムアナライザー、HiSeq、またはMiSeqシステム)により製造されるシーケンサーを含む。2、3、4、5、またはそれ以上の自動デバイス間(例えば、液体ハンドラー、架橋増幅デバイス、および配列決定デバイスのうちの1つ以上の間)での伝達は、手動また

40

50

は自動であり得る。いくつかの実施形態において、本発明の方法における1つ以上のステップ（例えば、全てのステップまたは全ての自動化ステップ）は、約72、48、24、20、18、16、14、12、10、9、8、7、6、5、4、3、2、1、もしくはそれ以下、またはそれ未満の時間内に完了される。いくつかの実施形態において、試料受領、DNA抽出、断片化、アダプター連結、増幅、または架橋増幅からの配列決定データの生成までの時間は、約72、48、24、20、18、16、14、12、10、9、8、7、6、5、4、3、2、1、もしくはそれ以下、またはそれ未満の時間である。

【0089】

一態様において、本発明は、試料中の複数の異なる標的ポリヌクレオチドを富化する方法を提供する。いくつかの実施形態において、この方法は、（a）配列Yを含むアダプターオリゴヌクレオチドを標的ポリヌクレオチドのそれぞれに連結することと、（b）複数の異なるオリゴヌクレオチドプライマーを、適合した標的ポリヌクレオチドにハイブリダイズすることであって、それぞれのオリゴヌクレオチドプライマーが、配列Zおよび配列Wを含み、配列Zが、全てのオリゴヌクレオチドプライマーの中で共通であり、さらに配列Wが、それぞれの異なるオリゴヌクレオチドプライマーに対して異なり、それぞれのオリゴヌクレオチドプライマーの3'末端に位置付けられ、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の200ヌクレオチド以内にある配列に相補性である、ハイブリダイズすることと、（c）伸長反応において、配列Zおよび配列Yを含む伸長したプライマーを生成するように、オリゴヌクレオチドプライマーを適合した標的ポリヌクレオチドに沿って伸長することであって、配列Yが、配列Yに相補性である、伸長することと、（d）（i）第1の増幅プライマーであって、配列Vおよび配列Zを含み、配列Zが第1の増幅プライマーの3'末端に位置付けられる、第1の増幅プライマー、ならびに（ii）第2の増幅プライマーであって、配列Xおよび配列Yを含み、配列Yが第2の増幅プライマーの3'末端に位置付けられる、第2の増幅プライマーの対を使用して、指数関数的に増幅することと、を含む。いくつかの実施形態において、それぞれのオリゴヌクレオチドプライマーは、第1の結合パートナーを含む。いくつかの実施形態において、この方法は、ステップ（d）の前に、伸長したプライマーを、第1の結合パートナーに結合する第2の結合パートナーを含む固体表面に曝露することと、それにより伸長したプライマーを、伸長反応の1つ以上の成分から取り出して精製することをさらに含む。いくつかの実施形態において、配列V、W、X、Y、およびZのうちの1つ以上は、異なる配列である。いくつかの実施形態において、配列Vおよび配列Xは同じである。いくつかの実施形態において、配列Vおよび/または配列Xは、それらそれぞれのプライマーに含まれない。いくつかの実施形態において、配列V、W、X、Y、およびZのうちの1つ以上は、配列V、W、X、Y、およびZのその他の1つ以上と、約5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより多いか、それ未満、またはそれ以上異なる（例えば、約10%、20%、30%、40%、50%、60%、70%、80%、90%、もしくはそれより高い配列同一性、またはそれより低い配列同一性を有する）。いくつかの実施形態において、配列V、W、X、Y、およびZのうちの1つ以上は、それぞれ約1、2、3、4、5、6、7、8、9、10、15、20、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドを含む。いくつかの実施形態において、本発明の他の態様に関して記載されるように、配列Vまたは配列Zは、配列Aに相当し、配列Wは、配列Bに相当し、配列Xは、配列Cに相当し、および/または配列Yは、配列Dに相当する。

【0090】

一態様において、本発明は、試料中の複数の異なる標的ポリヌクレオチドを富化する方法を提供する。いくつかの実施形態において、この方法は、（a）複数の異なるオリゴヌクレオチドプライマーを、標的ポリヌクレオチドにハイブリダイズすることであって、それぞれのオリゴヌクレオチドプライマーが、配列Zおよび配列Wを含み、配列Zが、全てのオリゴヌクレオチドプライマーの中で共通であり、さらに配列Wが、それぞれの異なるオリゴヌクレオチドプライマーに対して異なり、それぞれのオリゴヌクレオチドプライマ

一の 3 末端に位置付けられ、原因となる遺伝的変異体を含む配列、または原因となる遺伝的変異体の 200ヌクレオチド以内にある配列に相補性である、ハイブリダイズすることと、(b)伸長反応において、伸長したプライマーを生成するように、オリゴヌクレオチドプライマーを標的ポリヌクレオチドに沿って伸長することと、(c)アダプターオリゴヌクレオチドをそれぞれの伸長したプライマーに連結することと、アダプターオリゴヌクレオチドが、配列 Y を含み、さらに配列 Y が、配列 Y の相補体である、連結することと、(d)(i)第 1 の増幅プライマーであって、配列 V および配列 Z を含み、配列 Z が第 1 の増幅プライマーの 3 末端に位置付けられる、第 1 の増幅プライマー、ならびに(ii)第 2 の増幅プライマーであって、配列 X および配列 Y を含み、配列 Y が第 2 の増幅プライマーの 3 末端に位置付けられる、第 2 の増幅プライマーの対を使用して、指数関数的に増幅することと、を含む。いくつかの実施形態において、それぞれのオリゴヌクレオチドプライマーは、第 1 の結合パートナーを含む。いくつかの実施形態において、この方法は、ステップ(c)の前に、伸長したプライマーを、第 1 の結合パートナーに結合する第 2 の結合パートナーを含む固体表面に曝露することと、それにより伸長したプライマーを、伸長反応の 1 つ以上の成分から取り出して精製することをさらに含む。いくつかの実施形態において、配列 V、W、X、Y、および Z のうちの 1 つ以上は、異なる配列である。いくつかの実施形態において、配列 V および配列 X は同じである。いくつかの実施形態において、配列 V および / または配列 X は、それらそれぞれのプライマーに含まれない。いくつかの実施形態において、配列 V、W、X、Y、および Z のうちの 1 つ以上は、配列 V、W、X、Y、および Z のその他の 1 つ以上と、約 5 %、10 %、15 %、20 %、25 %、30 %、40 %、50 %、60 %、70 %、80 %、90 %、もしくはそれより多いか、それ未満、またはそれ以上異なる(例えば、約 10 %、20 %、30 %、40 %、50 %、60 %、70 %、80 %、90 %、もしくはそれより高い配列同一性、またはそれより低い配列同一性を有する)。いくつかの実施形態において、配列 V、W、X、Y、および Z のうちの 1 つ以上は、それぞれ約 1、2、3、4、5、6、7、8、9、10、15、20、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドを含む。いくつかの実施形態において、本発明の他の態様に関して記載されるように、配列 V または配列 Z は、配列 A に相当し、配列 W は、配列 B に相当し、配列 X は、配列 C に相当し、および / または配列 Y は、配列 D に相当する。

【0091】

標的ポリヌクレオチドが誘導される試料は、同じ個体からの複数の試料、異なる個体からの試料、またはそれらの組み合わせを含むことができる。いくつかの実施形態において、試料は、単一個体からの複数のポリヌクレオチドを含む。いくつかの実施形態において、試料は、2 つ以上の個体からの複数のポリヌクレオチドを含む。試料ポリヌクレオチドの供給源の例およびそれらを精製するための方法は、例えば、本発明の他の態様に関して本明細書に記載される。

【0092】

いくつかの実施形態において、標的ポリヌクレオチドは、1 つ以上の特定のサイズ範囲(複数可)の断片化ポリヌクレオチドの集団に断片化される。いくつかの実施形態において、断片化に供される試料ポリヌクレオチドの量は、約 50 ng、100 ng、200 ng、300 ng、400 ng、500 ng、600 ng、700 ng、800 ng、900 ng、1000 ng、1500 ng、2000 ng、2500 ng、5000 ng、10 μg、もしくはそれより多いか、それ未満、またはそれ以上である。いくつかの実施形態において、断片は、約 1、10、100、1000、10000、100000、300000、500000、もしくはそれより多いか、それ未満、またはそれ以上の開始 DNA のゲノム相当物から生成される。断片化は、化学的、酵素的、および機械的断片化を含む当該技術分野において既知の方法により達成され得る。いくつかの実施形態において、断片は、約 10 ~ 約 10,000ヌクレオチド長の平均または中央値を有する。いくつかの実施形態において、断片は、約 50 ~ 約 2,000ヌクレオチドの長さの平均または中央値を有する。いくつかの実施形態において、断片は、約 100 ~ 2500、200 ~

1000、10～800、10～500、50～500、50～250、または50～150ヌクレオチド長であるか、それ未満、それ以上、またはその間の平均または中央値を有する。いくつかの実施形態において、断片は、約200、300、500、600、800、1000、1500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長の平均または中央値を有する。断片化および任意の末端修復（任意のA-テリングを含む）の例示の方法は、例えば、本発明の他の態様に関して本明細書に記載される。末端修復は、アダプターオリゴヌクレオチドの連結前、例えば、オリゴヌクレオチドプライマーの伸長前または伸長後の任意のステップにおいて行われ得る。

【0093】

いくつかの実施形態において、オリゴヌクレオチドプライマー伸長の断片化の後に、断片化または伸長ポリヌクレオチドへのアダプターオリゴヌクレオチドのライゲーションが続く（例えば、図5および7を参照）。アダプターオリゴヌクレオチドの例、およびそれらの操作および標的ポリヌクレオチドへの連結のための方法は、例えば、本発明の他の態様に関して本明細書に記載される。いくつかの実施形態において、アダプターオリゴヌクレオチドは、配列要素配列Yを含む1本の鎖を含む。いくつかの実施形態において、アダプターオリゴヌクレオチドは、配列Yの相補体である、配列要素配列Yを含む1本の鎖を含む。いくつかの実施形態において、アダプターオリゴヌクレオチドは、相補性配列Yにハイブリダイズされた配列Yを含み、配列Yは、配列Yと同じまたは異なる鎖上にある。いくつかの実施形態において、標的ポリヌクレオチドの3'末端または伸長したプライマーは、アダプターオリゴヌクレオチドに沿って伸長し、配列Yまたは配列Yを生成する。いくつかの実施形態において、断片化ポリヌクレオチドおよびアダプターオリゴヌクレオチドは、両端でアダプターオリゴヌクレオチド配列に連結した断片化ポリヌクレオチド配列を含む二本鎖の適合したポリヌクレオチドを生成するように（例えば、ライゲーションにより、および任意に断片伸長により）合わせて処理され、適合したポリヌクレオチドの両端は、配列Yにハイブリダイズされた配列Yを含む。いくつかの実施形態において、標的ポリヌクレオチドにハイブリダイズされた伸長したプライマーは、一端で配列Yにハイブリダイズされた配列Yを含む二本鎖の適合したポリヌクレオチドを生成するように（例えば、ライゲーションにより、および任意に3'末端伸長により）合わせて処理される。いくつかの実施形態において、さらなる操作（例えば、アダプター連結またはオリゴヌクレオチドプライマー伸長）に供される断片化ポリヌクレオチドの量は、約50ng、100ng、200ng、300ng、400ng、500ng、600ng、700ng、800ng、900ng、1000ng、1500ng、2000ng、2500ng、5000ng、10μg、もしくはそれより多いか（例えば、閾値量）、それ未満、またはそれ以上である。いくつかの実施形態において、断片化ポリヌクレオチドの量は、さらなる操作を進める前に決定され、その量が閾値量を下回る場合、さらなる操作は行われない。

【0094】

いくつかの実施形態において、標的ポリヌクレオチド配列に相補性の配列を含むプライマー伸長生成物は、伸長反応において生成される。一般に、伸長反応は、標的ポリヌクレオチドにハイブリダイズされたオリゴヌクレオチドプライマーの伸長を含む。オリゴヌクレオチドプライマーは、任意の好適な長さ、例えば、約5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドであり得、その任意の部分または全ては、プライマーがハイブリダイズする対応する標的配列に相補性であり得る（例えば、約5、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド）。プライマー伸長は、PCR反応の1つ以上のサイクル、例えば、変性、プライマーアニーリング、およびプライマー伸長を含み得、これらは、逆プライマーを用いるか、または用いずに任意の回数で反復され得る。例えば、逆プライマーの非存在下で複数のサイクルを使用して、さらなる増幅のためのテンプレートとして伸長プライマーを使用せずに、対応する標的に沿ったプライ

10

20

30

40

50

マーの反復伸長により1つ以上の標的ポリヌクレオチドを直線的に増幅することができる。プライマーとして有用なオリゴヌクレオチドおよびプライマー伸長反応（例えば、増幅）におけるそれらの使用のための方法の例は、例えば、本発明の他の態様に関して本明細書に提供される。増幅方法の非限定的な例の例示が、図2に提供される。

【0095】

いくつかの実施形態において、オリゴヌクレオチドプライマーは、反応において複数の異なるオリゴヌクレオチドプライマーのそれぞれに共通する配列Zと、それぞれの異なるオリゴヌクレオチドプライマーに対して異なり、それぞれのオリゴヌクレオチドプライマーの3'末端に位置付けられる配列Wと、を含む。いくつかの実施形態において、複数のオリゴヌクレオチドプライマーは、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の異なるオリゴヌクレオチドを含み、それぞれが異なる配列Wを含む。いくつかの実施形態において、複数のオリゴヌクレオチドプライマーのうちの1つ以上の配列Wは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む（例えば、それぞれが図4-1~3からの異なる配列を持つ1、5、10、25、50、75、または100個の異なるオリゴヌクレオチド）。いくつかの実施形態において、配列W、またはそれが特異的にハイブリダイズする標的配列は、本明細書に記載される原因となる遺伝的変異体を含む。いくつかの実施形態において、配列Wまたはそれが特異的にハイブリダイズする標的配列は、本明細書に記載されるように、原因となる遺伝的変異体の約1、2、3、4、5、6、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。原因となる遺伝的変異体は、典型的に、オリゴヌクレオチドプライマーの下流に位置し、その原因となる遺伝的変異体の少なくとも一部分が、オリゴヌクレオチドプライマーの伸長のためのテンプレートとして機能する。典型的に、アダプターオリゴヌクレオチドから誘導された配列Yを含む標的ポリヌクレオチドに沿ったオリゴヌクレオチドプライマーの伸長は、プライマーから誘導された配列を5'末端に含み、アダプターから誘導された配列に相補性の配列（例えば、Yの相補体である配列Y'）を3'末端付近に含むプライマー伸長生成物を生成する。

【0096】

いくつかの実施形態において、複数のオリゴヌクレオチドプライマーのうちの1つ以上の配列W、またはそれが特異的にハイブリダイズする標的配列は、非対象配列を含む。いくつかの実施形態において、配列Wまたはそれが特異的にハイブリダイズする標的配列は、非対象配列の約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。一般に、非対象配列は、試験される個体以外の生物から誘導されるポリヌクレオチド、例えば、細菌、古細菌、ウイルス、原生生物、真菌、または他の生物からのDNAまたはRNAに対応する。非対象配列は、ある生物または生物の類の同一性を示すことがあり、さらに感染等の疾患状態を示し得る。生物を特定する際に有用な非対象配列の例としては、rRNA配列、例えば16S rRNA配列が挙げられるが、これに限定されない（例えば、国際公開第2010151842号を参照）。いくつかの実施形態において、非対象配列は、原因となる遺伝的変異体の代わりに、またはそれとは別に分析される。いくつかの実施形態において、原因となる遺伝的変異体および非対象配列は、並行して、例えば同じ試料中（例えば、第1のオリゴヌクレオチド、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Wを持つもの、および非対象配列を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Wを持つものの混合物を使用する）および/または同じレポート中で分析される。

【0097】

いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列W、またはそれが特異的にハイブリダイズする標的配列は、祖先情報提供マーカー(AIM)を含む。いくつかの実施形態において、配列Wまたはそれが特異的にハイブリダイズする標的配列は、AIMの約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。AIMは、原因となる遺伝的変異体のうちの1つの危険性が高い集団のような1つ以上の集団に属するか、または属しない人を分類するために使用され得る。例えば、AIMは、ある形質の保有率が高い集団についての診断手段となり得る。場合によっては、AIMは、より細かい精度で集団、例えば、亜大陸グループまたは関連する民族グループを区別する。いくつかの実施形態において、AIMは、原因となる遺伝的変異体および/または非対象配列の代わりに、またはそれとは別に分析される。いくつかの実施形態において、AIM、原因となる遺伝的変異体、および/または非対象配列は、並行して、例えば、同じ試料中(例えば、第1のオリゴヌクレオチド、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つもの、およびAIMを含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つものの混合物を使用する)および/または同じレポート中で分析される。

【0098】

いくつかの実施形態において、オリゴヌクレオチドプライマーは、結合対のメンバーのような第1の結合パートナーを含む。一般に、「結合パートナー」は、第1の部分および第2の部分のうちの1つを指し、第1および第2の部分は、互いに対して特定の結合親和性を有する。本発明における使用のための好適な結合対としては、限定されないが、抗原/抗体(例えば、ジゴキシゲニン/抗ジゴキシゲニン、ジニトロフェニル(DNP)/抗DNA、ダンシル-X-抗ダンシル、フルオレセイン/抗フルオレセイン、ルシファー黄色/抗ルシファー黄色、およびローダミン/抗ローダミン); ビオチン/アビジン(またはビオチン/ストレプトアビジン); カルモデュリン結合タンパク質(CBP/カルモデュリン); ホルモン/ホルモン受容体; レクチン/炭水化物; ペプチド/細胞膜受容体; タンパク質A/抗体; ハプテン/抗ハプテン; 酵素/共因子; および酵素/基質が挙げられるが、これらに限定されない。他の好適な結合対としては、FLAG-ペプチド(Hopp et al., *BioTechnology*, 6:1204~1210(1988)); KT3エピトープペプチド(Martin et al., *Science*, 255:192~194(1992)); チュープリンエピトープペプチド(Skinner et al., *J. Biol. Chem.*, 266:15163~15166(1991)); およびT7遺伝子10タンパク質ペプチドタグ(Lutz-Freyermuth et al., *Proc. Natl. Acad. Sci. USA*, 87:6393~6397(1990)) およびそれに対するそれぞれの抗体等のポリペプチドが挙げられる。結合パートナーのさらなる非限定的な例としては、細胞膜受容体、毒素、および毒のアゴニストおよびアンタゴニスト、ウイルスエピトープ、ステロイド、ホルモン受容体、ペプチド、酵素、および他の触媒ポリペプチド等のホルモン、酵素基質、共因子、小有機分子薬物を含む薬物、アヘン剤、アヘン剤受容体、エクチン、糖、ポリサッカリドを含むサッカリド、タンパク質、およびモノクローナル抗体と合成抗体断片とを含む抗体、細胞、細胞膜、および細胞膜受容体を含むその中の部分、ならびに細胞小器官が挙げられる。いくつかの実施形態において、第1の結合パートナーは、反応部分であり、第2の結合パートナーは、本発明の他の態様に関して本明細書に記載されるような反応部分と反応する反応表面である。いくつかの実施形態において、オリゴヌクレオチドプライマーは、伸長反応を開始する前に固体表面に結合される。結合パートナーのオリゴヌクレオチドへの付加のための方法は、当該技術分野において既知であり、合成中(例えば、結合パートナーを含む修飾ヌクレオチドを使用することにより)、または合成後に付加を含む。

【0099】

いくつかの実施形態において、オリゴヌクレオチドプライマーの伸長の後に、固体表面

上の伸長プライマーの精製が続く。いくつかの実施形態において、アダプター連結の後に、固体表面上の伸長プライマーの精製が続く。典型的に、固体表面は、結合対の第2のメンバーであり、第1の結合パートナーに結合する第2の結合パートナーを含む。いくつかの実施形態において、固体表面は、膜、スライド、プレート、微細機械加工されたチップ、微粒子、ビーズ等を含む広範な形態を有し得る。固体表面は、ガラス、プラスチック、ケイ素、アルカンチオール酸誘導体化金、セルロース、低架橋および高架橋ポリスチレン、シリカゲル、ポリアミド等を含むが、これらに限定されない広範囲の材料を含み得、様々な形態および特徴（例えば、ウェル、インデント、チャンネル等）を有することができる。この表面は、親水性であり得るか、または親水性にすることができ、シリカ、硫酸マグネシウム、およびアルミニウム等の無機粉末；天然ポリマー材料、特にセルロース材料およびセルロースから誘導された材料（繊維含有紙、例えば、濾紙、クロマトグラフ紙等）；合成または修飾された天然に存在するポリマー、例えば、ニトロセルロース、酢酸セルロース、ポリ（塩化ビニル）、ポリアクリルアミド、架橋デキストラン、アガロース、ポリアクリレート、ポリエチレン、ポリプロピレン、ポリ（4-メチルブチル）等；それら自体が使用されるか、または他の材料と併せて使用されるかのいずれか；バイオガラス、セラミック、金属として使用可能なガラス等を含み得る。リポソーム、リン脂質小胞、および細胞等の天然または合成アセンブリを用いることもできる。この表面は、薄片、棒、ビーズを含む粒子等の多数の形状のうちのいずれか1つを有することができる。

【0100】

いくつかの実施形態において、固体表面は、1つのビーズまたは複数のビーズを含む。これらのビーズは、任意の便宜的なサイズであり得、任意の数の既知の材料から製造され得る。そのような材料の例としては、無機物、天然ポリマー、および合成ポリマーが挙げられる。これらの材料の特定例としては、セルロース、セルロース誘導体、アクリル樹脂、ガラス、シリカゲル、ポリスチレン、ゼラチン、ポリビニルピロリドン、ビニルおよびアクリルアミドのコポリマー、ジビニルベンゼン等と架橋されたポリスチレン（例えば、Merrifield, Biochemistry 1964, 3, 1385~1390に記載される）、ポリアクリルアミド、ラテックスゲル、ポリスチレン、デキストラン、ゴム、シリコン、プラスチック、ニトロセルロース、天然スポンジ、シリカゲル、制御細孔ガラス、金属、架橋デキストラン（例えば、Sephadex）アガロースゲル（Sepharose）、および当業者に既知の他の固体支持体が挙げられる。ビーズは、一般に、直径約2~約100μm、または直径約5~約80μm、場合によっては直径約10~約40μmである。いくつかの実施形態において、ビーズは、磁性、常磁性、または他の方法で磁場に応答性であり得る。ビーズを磁場に対して応答性にすることは、例えば、磁場の印加およびビーズの単離により（例えば、溶液からのビーズの除去またはビーズからの溶液の除去により）そこに結合されたポリヌクレオチドを有するビーズの単離および精製に有用であり得る。磁場に対して応答性のビーズの非限定的な例としては、Life Technologies (Carlsbad, CA) により製造されたダイナビーズ (Dynabeads) が挙げられる。ビーズを分離する他の方法を使用することもできる。例えば、捕捉ビーズは、核酸-ビーズ複合体を蛍光にする蛍光部分で標識され得る。標的捕捉ビーズ複合体は、例えば、フローサイトメトリーまたは蛍光細胞選別装置により分離され得る。ビーズは、遠心分離により分離されることもできる。ビーズへの結合によるポリヌクレオチドの単離は、例えば、好適な洗浄緩衝液中でビーズを洗浄するステップをさらに含み得る。一般に、プライマー伸長生成物の精製は、プライマー伸長反応の1つ以上の成分から取り出して精製することを含み、伸長生成物が精製される1つ以上の成分が、例えば、10倍、5倍、100倍、500倍、1000倍、10000倍、100000倍、もしくはそれより多いか、または検出レベルを下回る量だけ低減されるようにする。いくつかの実施形態において、精製は、プライマー伸長生成物が、それらがハイブリダイズされた標的ポリヌクレオチドテンプレートから取り出して精製されるように、変性ステップを含む。

【0101】

伸長プライマーは、線形増幅または指数関数的増幅のような増幅に供され得る。増幅のための方法は、当該技術分野において既知であり、その例は、例えば、本発明の他の態様に関して本明細書に記載される。指数関数的増幅としては、PCR増幅、およびプライマー伸長生成物が、プライマー伸長のさらなる試行のためのテンプレートとして機能する任意の他の増幅方法が挙げられる。増幅は、典型的に、1つ以上の増幅プライマーを利用し、その例は、例えば、本発明の他の態様に関して本明細書に記載される。増幅プライマーは、任意の好適な長さ、例えば、約5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドであり得、その任意の部分または全ては、プライマーがハイブリダイズする対応する標的配列に相補性であり得る（例えば、約5、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド）。一般に、PCRには、増幅される標的の変性（二本鎖である場合）、1つ以上のプライマーの標的へのハイブリダイゼーション、およびDNAポリメラーゼによるプライマーの伸長のステップが関与し、ステップは、標的配列を増幅するために反復（または「サイクル化」）される。このプロセスにおけるステップは、収率を強化する、疑似生成物の形成を減少させる、および/またはプライマーアニーリングの特異性を増減させる等の様々な結果に対して最適化することができる。最適化の方法は、当該技術分野において周知であり、増幅反応における要素の種類もしくは量、および/またはプロセスにおける所与のステップの条件（例えば、特定ステップでの温度、特定ステップの期間、および/またはサイクル数）に対する調整を含む。いくつかの実施形態において、増幅反応は、少なくとも5、10、15、20、25、30、35、50、またはそれより多いサイクルを含む。いくつかの実施形態において、増幅反応は、わずか5、10、15、20、25、35、50、またはそれより多いサイクルを含む。サイクルは、任意の数のステップ、例えば、1、2、3、4、5、6、7、8、9、10、もしくはそれより多いステップを含むことができる。ステップは、鎖変性、プライマーアニーリング、およびプライマー伸長を含むが、これらに限定されない所与のステップの目的を達成するのに好適な任意の温度または温度の勾配を含むことができる。ステップは、手動で中断されるまで無期限に、約1秒、5秒、10秒、15秒、20秒、25秒、30秒、35秒、40秒、45秒、50秒、55秒、60秒、70秒、80秒、90秒、100秒、120秒、180秒、240秒、300秒、360秒、420秒、480秒、540秒、600秒、もしくはそれより多いか、それ未満、またはそれ以上であるが、これらに限定されない任意の期間であり得る。異なるステップを含む任意の数のサイクルは、任意の順序で組み合わせることができる。

【0102】

いくつかの実施形態において、増幅は、増幅プライマーの対を使用してプライマー伸長生成物を生成することを含む。増幅プライマーは、アダプターオリゴヌクレオチド配列から誘導された配列の全体または1つ以上の部分に相補性の配列、オリゴヌクレオチドプライマー配列から誘導された配列、テンプレートポリヌクレオチドに相補性でない配列（例えば、5'非相補性配列）、1つ以上の他の配列要素（例えば、本明細書に記載されるような配列要素）、またはそれらの組み合わせを含み得る。いくつかの実施形態において、第2の増幅プライマーは、配列Xおよび配列Yを含み、配列Yは、第2の増幅プライマーの3'末端に位置付けられる。

【0103】

図2は、増幅プロセスの非限定的な例を示す。例示の指数関数的増幅反応の第1のステップにおいて、第2の増幅プライマーの配列Yは、以前のオリゴヌクレオチドプライマー伸長反応からの伸長プライマーの相補性配列Y'にハイブリダイズする。第2の増幅プライマーの伸長（例えば、ポリメラーゼによる）は、配列X、Y、W'、およびZ'を5'から3'の方向で含む第2の増幅プライマー伸長生成物を生成し、配列W'は、配列Wの相補体であり、配列Z'は、配列Zの相補体である。次に、プライマー伸長生成物は変性され、さらなる第2の増幅プライマーとのハイブリダイゼーションおよびその伸長のため

のテンプレートとして機能するように、テンプレート標的ポリヌクレオチド、および第1の増幅プライマーとのハイブリダイゼーションおよびその伸長のための伸長生成物を遊離する。いくつかの実施形態において、第1の増幅プライマーは、配列Vおよび配列Zを含み、配列Zは、第1の増幅プライマーの3'末端に位置付けられる。この例示の増幅反応において、配列Zは、第2の増幅プライマー伸長生成物の配列Z'にハイブリダイズする。第1のプライマーの伸長（例えば、ポリメラーゼによる）は、配列V、Z、W、Y、およびX'を5'から3'配向に含む第1の増幅プライマー伸長生成物を生成し、配列X'は、配列Xに相補性であり、それ自体が第2の増幅プライマーの伸長のためのテンプレートとして機能することができる。したがって、変性、ハイブリダイゼーション、および伸長の反復サイクルは、配列X、Y、W、Z、およびV'（5'から3'）を含む第2の鎖にハイブリダイズされた配列V、Z、W、Y、およびX'（5'から3'）を含む1つの鎖を含むプライマー伸長生成物の二本鎖を生成する。この例示の増幅反応に従い、標的ポリヌクレオチド配列は、一般に、片方の鎖上の配列ZとY'との間、および他方の鎖上の配列Z'とYとの間に位置付けられる。

【0104】

いくつかの実施形態において、オリゴヌクレオチドプライマーおよび/または1つ以上の増幅プライマーは、バーコードを含む。バーコードの例は、例えば、本発明の他の態様に関して本明細書に記載される。いくつかの実施形態において、別個の増幅反応は、それぞれの試料に対して少なくとも1つの異なるバーコード配列を含む増幅プライマーを使用して、別個の試料に対して実行され、いかなるバーコード配列も、並行して分析される複数の試料の標的ポリヌクレオチドに連結されないようになる。いくつかの実施形態において、異なる試料から誘導され、異なるバーコードを含む増幅ポリヌクレオチドは、ポリヌクレオチドの後次操作を進める前（例えば、配列決定前）にプールされる。プールは、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、40、50、75、100、もしくはそれより多いか、それ未満、またはそれ以上の異なる試料から誘導されたポリヌクレオチドを含み得る。プールは、後次に配列決定に供され得、配列決定された標的ポリヌクレオチドの供給源試料は、それらの関連バーコードに基づいて特定され得る。

【0105】

いくつかの実施形態において、指数関数的に増幅された標的ポリヌクレオチドが配列決定される。配列決定は、例えば、本発明の他の態様を参照して本明細書に記載される配列決定プロセスを含む、当該技術分野において既知の任意の配列決定方法に従い行われ得る。テンプレートに依存する合成を使用する配列分析は、いくつかの異なるプロセスを含むことができる。例えば、広範に実施される4色Sanger配列決定方法において、テンプレート分子の集団を使用して、相補性断片配列の集団を形成する。プライマー伸長は、4つの天然に存在するヌクレオチドの存在下、色素標識された終止因子ヌクレオチド、例えばジデオキシリボヌクレオチドの亜集団を用いて実行され、それぞれの種類の終止因子（ddATP、ddGTP、ddTTP、ddCTP）は、異なる検出可能な標識を含む。結果として、断片の入れ子式集合が形成され、断片は、プライマーを超える配列中のそれぞれのヌクレオチドで終止し、終止ヌクレオチドの特定を許容する方法で標識される。次に、入れ子式断片集団は、例えば、キャピラリー電気泳動を使用して、サイズに基づく分離に供され、それぞれの異なるサイズの断片と関連付けられる標識は、終止ヌクレオチドを特定するために特定される。結果として、分離システムにおいて検出器を越えて移動する標識の配列は、合成された断片の配列情報の直接読み出し、および相補性により基礎的テンプレートを提供する（例えば、米国特許第5171534号を参照）。

【0106】

テンプレートに依存する配列決定方法の他の例としては、合成プロセスによる配列が挙げられ、個別のヌクレオチドは、それらが成長プライマー伸長生成物に付加されると反復して特定される。

【0107】

10

20

30

40

50

ピロ配列決定は、配列決定反応の副産物、つまりピロリン酸塩の存在について、得られる合成混合物をアッセイすることにより、ヌクレオチドの組み込みを特定する合成プロセスによる配列の例である。具体的に、プライマー/テンプレート/ポリメラーゼ複合体が、単一種のヌクレオチドと接触する。そのヌクレオチドが組み込まれる場合、重合反応は、ピロリン酸塩を放出する三リン酸塩鎖のリン酸塩とリン酸塩との間のヌクレオシド三リン酸塩を切断する。次に、放出されたピロリン酸塩の存在は、ピロリン酸塩をAMPでATPに変換する化学発光酵素レポーターシステムを使用して特定された後、ルシフェラーゼ酵素を使用してATPを測定し、測定可能な光シグナルを生成する。光が検出される場合、塩基が組み込まれており、光が検出されない場合は塩基が組み込まれていない。適切な洗浄ステップに続いて、様々な塩基は、複合体と周期的に接触し、テンプレート配列における後次塩基を連続的に特定する。例えば、米国特許第6210891号を参照されたい。

10

【0108】

関連するプロセスにおいて、プライマー/テンプレート/ポリメラーゼ複合体は、基質上で固定化され、複合体は、標識ヌクレオチドと接触する。複合体の固定化は、プライマー配列、テンプレート配列、および/またはポリメラーゼ酵素を通じて行われることがあり、共有結合または非共有結合であり得る。例えば、複合体の固定化は、ポリメラーゼまたはプライマーと基質表面との間の結合を介することができる。代替構成において、ヌクレオチドは、除去可能な終止基とともに、および除去可能な終止基なしに提供される。組み込む際に、標識は複合体と連結され、したがって検出可能である。終止因子を担持するヌクレオチドの場合、個別に特定可能な標識を担持する4つの異なるヌクレオチドの全ては、複合体と接触する。標識ヌクレオチドの組み込みは、終止因子の存在を理由として伸長を停止し、標識を複合体に付加して、組み込まれたヌクレオチドの特定を可能にする。次に、この標識および終止因子は、組み込まれたヌクレオチドから除去され、適切な洗浄ステップに続いて、プロセスは反復される。非終止ヌクレオチドの場合、単一種の標識ヌクレオチドが複合体に付加され、ピロ配列決定と同様に、それが組み込まれるかどうかを決定する。ヌクレオチド上の標識基の除去および適切な洗浄ステップに続いて、様々な異なるヌクレオチドが、同じプロセスにおいて反応混合物を通じてサイクル化される。例えば、あらゆる目的で参照によりその全体が本明細書に組み込まれる、米国特許第6833246号を参照されたい。例えば、Illuminaゲノム分析器システムは、参照により本明細書に組み込まれる国際公開第98/44151号に記載される技術に基づき、DNA分子は、アンカープロープ結合部位(それ以外では、フローセル結合部位と称される)を介して、配列決定プラットフォーム(フローセル)に結合され、ガラススライド上で原位置増幅する。DNA分子が増幅する固体表面は、典型的に、複数の第1および第2の結合オリゴヌクレオチドを含み、第1の結合オリゴヌクレオチドは、標的ポリヌクレオチドの付近または片方の末端にある配列に相補性であり、第2の結合オリゴヌクレオチドは、標的ポリヌクレオチドの付近または他方の末端にある配列に相補性である。この配置は、本明細書に記載されるような架橋増幅を許容する。次に、DNA分子は、配列決定プライマーにアニールされ、可逆的終止因子アプローチを使用して塩基ごとに並行して配列決定される。配列決定プライマーのハイブリダイゼーションは、架橋を係留する結合オリゴヌクレオチドのうちの1つにおいて、切断部位での二本鎖架橋ポリヌクレオチドの1本の鎖の切断により進行され得、したがって、変性により除去され得る固体基質に結合されない1つの単一鎖、および配列決定プライマーに結合され、それに対するハイブリダーゼーションに使用可能な他の鎖を残す。典型的に、Illuminaゲノム分析器システムは、8つのチャンネルを持つフローセルを利用して、18~36塩基長の配列決定読み出し値を生成し、1試験当たり1.3Gbpを超える高品質データを生成する(www.illumina.comを参照)。

20

30

40

【0109】

合成プロセスによるなおもさらなる配列において、異なる標識ヌクレオチドの組み込みは、テンプレートに依存する合成が実行されると実時間で観察される。具体的に、個別の

50

固定化プライマー／テンプレート／ポリメラーゼ複合体は、蛍光標識ヌクレオチドが組み込まれると観察され、それが付加されるとそれぞれの付加された塩基の実時間特定を許容する。このプロセスにおいて、標識基は、組み込み中に切断されたヌクレオチドの一部分に結合される。例えば、標識基の組み込み中に除去されたリン酸鎖の一部分、すなわち、 α 、 β 、 γ 、またはヌクレオシドポリリン酸塩上の他の末端リン酸基に結合することにより、標識は、新生鎖に組み込まれず、代わりに天然のDNAが生成される。個別の分子の観察は、典型的に、非常に小さな照明量内での複合体の光学的閉じ込めを必要とする。複合体を光学的に制限することにより、ランダムに拡散するヌクレオチドが非常に短期間の間存在する一方、組み込まれたヌクレオチドが、それらが組み込まれるとより長い間観察量内で保持される、監視領域を形成する。これは、組み込み事象と関連付けられる特徴的なシグナルをもたらし、これも付加される塩基に特徴的であるシグナルプロファイルによって特徴付けられる。関連態様において、蛍光共鳴エネルギー転移(FRET)色素対のような相互作用する標識成分は、ポリメラーゼまたは複合体の他の部分、および組み込みヌクレオチド上に提供され、組み込み事象が標識成分を相互に作用する近接に置くと、特徴的なシグナルが生じ、これもまた組み込まれる塩基に特徴的である(例えば、米国特許第6,056,661号、同第6,917,726号、同第7,033,764号、同第7,052,847号、同第7,056,676号、同第7,170,050号、同第7,361,466号、および同第7,416,844号、ならびに米国公開第20070134128号を参照)。

10

【0110】

20

いくつかの実施形態において、試料中の核酸は、ライゲーションにより配列決定される。この方法は、DNAリガーゼ酵素を使用して、例えば、ポロニー方法およびSOLiD技術(Applied Biosystems、現Invitrogen)において使用されるように、標的配列を特定する。一般に、固定長の全ての可能なオリゴヌクレオチドのプールが提供され、配列決定された位置に従い標識される。オリゴヌクレオチドは、アニーリングおよびライゲーションされ、マッチング配列のためのDNAリガーゼによる選好的ライゲーションは、その位置での相補性配列に対応するシグナルを生じる。

【0111】

いくつかの実施形態において、配列決定データは、複数の試料、例えば、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して並行して生成される。いくつかの実施形態において、配列決定データは、単一反応容器(例えば、フローセル中のチャンネル)内の複数の試料、例えば、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して生成され、配列決定データは、配列決定されたポリヌクレオチドが起源とする試料に従い(例えば、バーコード配列に基づいて)後次に分類される。

30

【0112】

いくつかの実施形態において、配列決定データは、単一反応容器(例えば、フローセル中のチャンネル)内の試料からの約5、10、25、50、100、150、200、250、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多い異なる標的ポリヌクレオチドに対して生成される。いくつかの実施形態において、配列決定データは、複数の試料、例えば、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して並行して生成される。いくつかの実施形態において、配列決定データは、単一反応容器(例えば、フローセル中のチャンネル)内の複数の試料、例えば、約2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、

40

50

384、768、1000、もしくはそれより多いか、それ未満、またはそれ以上の試料に対して生成され、配列決定データは、配列決定されたポリヌクレオチドが起源とする試料に従い後次に分類され、単一反応において、配列決定データは、架橋増幅反応から約、または少なくとも約 10^6 、 10^7 、 10^8 、 2×10^8 、 3×10^8 、 4×10^8 、 5×10^8 、 10^9 、 10^{10} 、もしくはそれより多い標的ポリヌクレオチドまたはクラスタに対して生成され得、反応におけるそれぞれの試料に対して約 10^4 、 10^5 、 10^6 、 2×10^6 、 3×10^6 、 4×10^6 、 5×10^6 、 10^7 、 10^8 、それ未満、またはそれ以上の標的ポリヌクレオチドまたはクラスタの配列決定データを含み得る。いくつかの実施形態において、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の原因となる遺伝的変異体の存在または非存在は、配列決定データに基づいて、試料に対して決定される。1つ以上の原因となる遺伝的変異体の存在または非存在は、約80%、85%、90%、95%、97.5%、99%、99.5%、99.9%、もしくはそれより高いか、またはそれ以上の精度で決定され得る。

【0113】

いくつかの実施形態において、本発明の方法におけるステップの1つ以上、または全ては、例えば、1つ以上の自動化デバイスの使用により自動化される。一般に、自動化デバイスは、ヒトの指示なしに作動することができるデバイスである。自動デバイスは、例えば、自動化デバイスがさらなるヒトの操作なしに1つ以上のステップを行った後にコンピューターに命令を入力することにより、ヒトが機能を促進するように任意の動作を行った後の期間中に機能を行うことができる。本発明の実施形態を実現するコードを含むソフトウェアおよびプログラムは、CD-ROM、DVD-ROM、テープ、フラッシュドライブ、もしくはディスク、または他の適切なコンピューター可読媒体等のいくつかの種類のデータ記憶媒体上に記憶されてもよい。本発明の様々な実施形態は、ハードウェアのみで、またはソフトウェアおよびハードウェアの組み合わせにて実現することもできる。例えば、一実施形態において、従来のパーソナルコンピューターではなく、プログラマブル論理制御装置(PLC)が使用される。当業者に既知の通り、PLCは、汎用コンピューターにかかる費用が不要である多様なプロセス制御アプリケーションにおいて頻繁に使用される。PLCは、1つまたは多様な制御プログラムを実行する既知の方法で構成され得、ユーザーもしくは別のデバイスからの入力を受信することができ、および/またはパーソナルコンピューターのそれに似た方法でユーザーもしくは別のデバイスに出力を提供することができる。したがって、本発明の実施形態は、汎用コンピューターに関して説明されているが、他の構成が使用される得るため、汎用コンピューターの使用は単なる例示であることを理解されたい。

【0114】

いくつかの実施形態において、自動化は、1つ以上の液体ハンドラーおよび関連ソフトウェアの使用を含み得る。いくつかの市販の液体ハンドリングシステムを利用して、これらのプロセスの自動化を試験することができる(例えば、例として、Perkin-Elmer、Beckman Coulter、Caliper Life Sciences、Tecan、Eppendorf、Apricot Design、Velocity 11からの液体ハンドラーを参照)。いくつかの実施形態において、自動化ステップは、断片化、末端修復、A-テリング(アデニンオーバーハングの付加)、アダプター連結、PCR増幅、試料定量化(例えば、DNAの量および/または純度)、および配列決定のうちの1つ以上を含む。いくつかの実施形態において、架橋増幅は、自動化される(例えば、Illumina cBotの使用により)。いくつかの実施形態において、配列決定は自動化される。多様な自動化配列決定機器は、市販されており、Life Technologies(SOLIDプラットフォーム、およびpHに基づく検出)、Roche(454プラットフォーム)、Illumina(例えば、フローセルに基づくシステム、例えば、ゲノムアナライザー、HiSeq、またはMiSeqシステム)によ

り製造されるシーケンサーを含む。2、3、4、5、またはそれ以上の自動化デバイス間（例えば、液体ハンドラー、架橋増幅デバイス、および配列決定デバイスのうちの1つ以上の間）での転移は、手動または自動であってよい。いくつかの実施形態において、本発明の方法における1つ以上のステップ（例えば、全てのステップまたは全ての自動化ステップ）は、約72、48、24、20、18、16、14、12、10、9、8、7、6、5、4、3、2、1、もしくはそれより少ないか、またはそれ以下の時間内に完了される。いくつかの実施形態において、試料受領、DNA抽出、断片化、アダプター連結、増幅、または架橋増幅からの配列決定データの生成までの時間は、約72、48、24、20、18、16、14、12、10、9、8、7、6、5、4、3、2、1、もしくはそれより少ないか、またはそれ以下の時間である。

10

【0115】

一態様において、本発明は、対象のゲノム中の遺伝的変異を検出する方法を提供する。いくつかの実施形態において、この方法は、配列決定データを生成すること、および分析することを含む。一実施形態において、この方法は、(a)ポリヌクレオチドの複数のクラスタを提供することであって、(i)それぞれのクラスタが、支持体に結合した核酸二本鎖の複数コピーを含み、(ii)クラスタ中のそれぞれの二本鎖が、配列A-B-G-D-Cを5から3に含む第1の分子と、配列C-D-G-B-Aを5から3に含む第2の分子と、を含み、(iii)配列Aが、配列Aに相補性であり、配列Bが、配列Bに相補性であり、配列Cが、配列Cに相補性であり、配列Dが、配列Dに相補性であり、配列Gが、配列Gに相補性であり、(iv)配列Gが、対象からの標的ポリヌクレオチド配列の一部であり、複数のクラスタのそれぞれに対して異なり、(v)配列Bが、対応する標的ポリヌクレオチド配列中の配列Gに関して5に位置する、提供することと、(b)配列Gを、配列Dを含む第1のプライマーの伸長により配列決定し、それぞれのクラスタのR1配列を生成することと、(c)それぞれのクラスタのR2配列を生成するように、配列Aを含む第2のプライマーの伸長により配列Bを配列決定することと、(d)全てのR1配列を第1の参照配列に整列させるように、第1のアルゴリズムを使用して第1のアラインメントを行うことと、(e)第1の参照配列に関して挿入または欠失を含有する可能性が高いとして第1のアラインメントにおいて特定されたR1配列を局部的に整列させ、それぞれの挿入または欠失の単一コンセンサスアラインメントを生成するように、第2のアルゴリズムを使用して第2のアラインメントを行うことと、(f)全てのR2配列を第2の参照配列に整列させることにより、R2アラインメントを行うことと、(g)ステップ(d)～(f)により特定された配列変異を特定するレポートを受信者に送信することと、を含む。いくつかの実施形態において、配列A、B、C、およびDは、本発明の他の態様に関して記載されるように、それぞれ配列A、B、C、およびDに対応する。

20

30

【0116】

いくつかの実施形態において、この方法は、(a)ポリヌクレオチドの複数のクラスタの配列決定データを提供することであって、(i)それぞれのクラスタが、支持体に結合した核酸二本鎖の複数コピーを含み、(ii)クラスタ中のそれぞれの二本鎖が、配列A-B-G-D-Cを5から3に含む第1の分子と、配列C-D-G-B-Aを5から3に含む第2の分子と、を含み、(iii)配列Aが、配列Aに相補性であり、配列Bが、配列Bに相補性であり、配列Cが、配列Cに相補性であり、配列Dが、配列Dに相補性であり、配列Gが、配列Gに相補性であり、(iv)配列Gが、対象からの標的ポリヌクレオチド配列の一部であり、複数のクラスタのそれぞれに対して異なり、(v)配列Bが、対応する標的ポリヌクレオチド配列中の配列Gに関して5に位置し、(vi)配列決定データが、配列Dを含む第1のプライマーの伸長により生成されたR1配列を含み、(vii)配列決定データが、配列Aを含む第2のプライマーの伸長により生成されたR2配列を含む、提供することと、(b)全てのR1配列を第1の参照配列に整列させるように、第1のアルゴリズムを使用して第1のアラインメントを行うことと、(c)第1の参照配列に関して挿入または欠失を含有する可能性が高い

40

50

として、該第1のアラインメントにおいて特定されたR1配列を局所的に整列させ、それぞれの挿入または欠失の単一コンセンサスアラインメントを生成するように、第2のアルゴリズムを使用して第2のアラインメントを行うことと、(d)全てのR2配列を第2の参照配列に整列させることにより、R2アラインメントを行うことと、(e)ステップ(b)~(d)により特定された配列変異を特定するレポートを受信者に送信することと、を含む。いくつかの実施形態において、配列A、B、C、およびDは、本発明の他の態様に関して記載されるように、それぞれ配列A、B、C、およびDに対応する。

【0117】

一般に、ポリヌクレオチドのクラスタは、支持体上の位置に共局在化する核酸二本鎖の複数のコピーを含む。多様な好適な固体支持体および支持体、材料は、当該技術分野において既知であり、その非限定的な例は、例えば、本発明の他の態様に関して本明細書に提供される。ポリヌクレオチドのクラスタは、架橋増幅により生成され得る。架橋増幅を行うための好適な方法および装置は、例えば、本発明の他の態様に関して本明細書に提供される。いくつかの実施形態において、固体支持体は、複数のクラスタを含み、それぞれのクラスタは、異なる標的ポリヌクレオチド配列の増幅により複数形成される。増幅される標的ポリヌクレオチド配列、例えば、配列Gの部分は、支持体上に固定化された第1のオリゴヌクレオチドの伸長を含むプロセスにおいて支持体に結合され得る。いくつかの実施形態において、固体支持体は、配列Aおよび配列Bを含む複数の異なる第1のオリゴヌクレオチドを含み、配列Aは、全ての第1のオリゴヌクレオチドの中で共通であり、さらに配列Bは、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3'末端にある。いくつかの実施形態において、複数の第1のオリゴヌクレオチドは、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の異なるオリゴヌクレオチドを含み、それぞれが異なる配列Bを含む。いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列Bは、図4-1~3に示される配列番号22~121からなる群から選択される配列を含む(例えば、それぞれが図4-1~3からの異なる配列を持つ1、5、10、25、50、75、または100個の異なるオリゴヌクレオチド)。いくつかの実施形態において、配列B、またはそれが特異的にハイブリダイズする標的配列は、本明細書に記載される原因となる遺伝的変異体を含む。いくつかの実施形態において、配列B、またはそれが特異的にハイブリダイズする標的配列は、本明細書に記載される原因的変異体の約1、2、3、4、5、6、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。原因となる遺伝的変異体は、典型的に、第1のオリゴヌクレオチドの下流に位置し、原因となる遺伝的変異体の少なくとも一部分が、第1のヌクレオチドの伸長のためのテンプレートとして機能するようになる。固体支持体は、本明細書に記載されるように、それぞれの第2のオリゴヌクレオチドの3'末端に配列Aを含む複数の第2のオリゴヌクレオチドと、それぞれの第3のオリゴヌクレオチドの3'末端に配列Cを含む複数の第3のオリゴヌクレオチドと、をさらに含み得る。二本鎖のクラスタを生成するために結合された第1、第2、および第3のオリゴヌクレオチドを使用する標的ポリヌクレオチド配列の一部分の架橋増幅の例は、図1に示され、配列Gは、配列BとDとの間に黒線で表され、配列Gは、配列BとDとの間に黒線で表される。

【0118】

いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列B、またはそれが特異的にハイブリダイズする標的配列は、非対象配列を含む。いくつかの実施形態において、配列Bまたはそれが特異的にハイブリダイズする標的配列は、非対象配列の約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。一般に、非対象配

10

20

30

40

50

列は、試験される個体以外の生物から誘導されるポリヌクレオチド、例えば、細菌、古細菌、ウイルス、原生生物、真菌、または他の生物からのDNAまたはRNAに対応する。非対象配列は、胎児からの核酸、例えば、胎児からの無細胞核酸（細胞外核酸とも称される）を含むこともできる。非対象配列は、生物の同一性を示し得、さらに感染等の疾患状態を示し得る。生物を特定する際に有用な非対象配列の例としては、限定されないが、rRNA配列、例えば、16S rRNA配列が挙げられる（例えば、国際公開第2010151842号を参照）。いくつかの実施形態において、非対象配列は、原因となる遺伝的変異体の代わりに、またはそれとは別に分析される。いくつかの実施形態において、原因となる遺伝的変異体および非対象配列は、例えば、同じ試料中（例えば、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つものと、非対象配列を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つものとの第1のオリゴヌクレオチドの混合物を使用する）および/または同じレポートにおいて並行して分析される。

【0119】

いくつかの実施形態において、複数の第1のオリゴヌクレオチドのうちの1つ以上の配列B、またはそれが特異的にハイブリダイズする標的配列は、祖先情報提供マーカー（AIM）を含む。いくつかの実施形態において、配列Bまたはそれが特異的にハイブリダイズする標的配列は、AIMの約1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、60、70、80、90、100、200、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド以内にある。AIMは、原因となる遺伝的変異体のうちの1つの危険性が高い集団のような1つ以上の集団に属するか、または属しない人を分類するために使用され得る。例えば、AIMは、ある形質の保有率が高い集団についての診断手段となり得る。場合によっては、AIMは、より細かい精度で集団、例えば、亜大陸グループまたは関連する民族グループを区別する。いくつかの実施形態において、AIMは、原因となる遺伝的変異体および/または非対象配列の代わりに、またはそれとは別に分析される。いくつかの実施形態において、AIM、原因となる遺伝的変異体、および/または非対象配列は、並行して、例えば、同じ試料中（例えば、第1のオリゴヌクレオチド、原因となる遺伝的変異体を含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つもの、およびAIMを含むか、またはその付近にある配列に特異的にハイブリダイズする配列Bを持つものの混合物を使用する）および/または同じレポート中で分析される。

【0120】

いくつかの実施形態において、複数のクラスターのうちの1つ以上の配列が配列決定される。配列決定の例示の方法は、例えば、本発明の他の態様に関して本明細書に記載される。配列決定データは、それぞれのクラスターの1つ以上の配列決定プライマーの伸長により生成され得る。配列決定プライマーは、任意の好適な長さ、例えば、約5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチドであり、その任意の部分または全ては、プライマーがハイブリダイズする対応する標的配列に相補性であり得る（例えば、約5、10、15、20、25、30、35、40、45、50、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド）。配列決定プライマー、プライマーDは、それが相補性配列Dに特異的にハイブリダイズするように、配列Dを含むか、またはそれからなり得る。いくつかの実施形態において、配列Dの下流にある第1のヌクレオチドは、配列Gの第1のヌクレオチドであり、プライマーDの伸長に付加された第1のヌクレオチドが配列Gに対応するようになる。配列決定プライマー、プライマーCは、それが相補性配列Cに特異的にハイブリダイズするように、配列Cを含むか、またはそれからなり得る。いくつかの実施形態において、配列Cの下流にある第1のヌクレオチドは、バーコード配列の第1のヌクレオチドであり、プライマーCの伸長に付加された第1のヌクレオチドがバーコード配列に対応するようになる。配列決定プライマー、プライマーAは、それが相補性配列Aに特異的にハイブリダイズする

ように、配列 A を含むか、またはそれからなり得る。いくつかの実施形態において、配列 A の下流にある第 1 のヌクレオチドは、配列 B の第 1 のヌクレオチドであり、プライマー A の伸長に付加された第 1 のヌクレオチドが配列 B に対応するようになる。いくつかの実施形態において、配列決定プライマーは、配列 C A C T C A G C A G C A C G A C G A T C A C A G A T G T G T A T A A G A G A C A G (配列番号 20) を含む。

【0121】

2 つ以上の異なる配列決定プライマーが、それぞれのクラスタに対して複数の配列決定読み出し値を生成するように連続配列決定反応において使用され得る。例えば、連続配列決定反応は、プライマー A、C、および D のそれぞれに対して任意の順序で行われ得る（例えば、プライマー D、次にプライマー C、次にプライマー A）。配列決定反応は、鎖切断、鎖変性、または前の反応の 1 つ以上の成分（例えば、配列決定プライマー）を除去するための洗浄ステップのうちの 1 つ以上により進められ得る。配列決定反応は、個別のヌクレオチドプライマー伸長の複数のサイクルを含み得、それぞれの付加後に、付加された塩基の同一性を決定する特定ステップが続く。個別のヌクレオチド伸長のサイクル数は、例えば、連続配列決定反応において使用される複数の配列決定プライマーのそれぞれに対して、または連続配列決定反応において使用される全ての配列決定プライマーに対して集合的に、約 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、150、200、250、500、もしくはそれより多いか、またはそれ以上である。いくつかの実施形態において、個別のヌクレオチド伸長のサイクル数は、バーコードまたはプローブ配列等の特定される配列の長さに基づいて選択され、約 30、25、20、15、10、9、8、7、6、5 サイクル、もしくはそれより少ないか、またはそれ以下であり得る。連続配列決定反応において使用される複数の配列決定プライマーのそれぞれに対するサイクル数は、異なり得る。例えば、59 サイクルのプライマー D の伸長後に、6 サイクルのプライマー C の伸長が続いてよく、次に 15 サイクルのプライマー A の伸長が続いて、合計 80 サイクルの伸長となり得る。

【0122】

第 1 の配列決定プライマー、第 2 の配列決定プライマー、およびインデックスプライマーである第 3 のプライマーの伸長は、それぞれのクラスタに対して、それぞれ R1、R2、およびバーコード配列を生成することができる。一般に、複数の配列は、クラスタのアレイ上の位置等の連続伸長反応の物理的共有化に基づいて、単一クラスタから生じると特定される。いくつかの実施形態において、配列決定データ（例えば、R1 および / または R2 配列）は、単一反応容器（例えば、フローセル中のチャンネル）内の試料からの約 5、10、25、50、100、150、200、250、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の異なる標的ポリヌクレオチドに対して、例えば、1 つ以上の配列決定プライマーの伸長により生成される。いくつかの実施形態において、配列決定データは、複数の試料、例えば、約 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多い試料に対して並行して生成される。いくつかの実施形態において、配列決定データは、単一反応容器（例えば、フローセル中のチャンネル）内の試料からの約 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、24、48、96、192、384、768、1000、もしくはそれより多い試料に対して生成され、配列決定データは、配列決定されたポリヌクレオチドが生じる試料に従って（例えば、バーコード配列に基づいて）、後次に分類される。バーコード配列に基づく配列決定データの分類は、例えば、本明細書に記載される 1 つ以上のアラインメントを行う前または後、および任意に 1 つ以上の配列を分析から除去する前に行われ得る。一般に、配列決定読み出し値がバーコードに基づいて分類されると、読み出し値のそれぞれの分類は、他の分類から独立してさらに処理される。いくつかの実施形態において、それぞれのバーコードは、並

10

20

30

40

50

行して分析された複数の異なるバーコード中のバーコードと1つおきに異なる。典型的に、バーコード配列は、単一反応において配列決定された試料のプール中の単一試料と関連付けられる。いくつかの実施形態において、複数のバーコード配列のそれぞれは、同時に配列決定された試料のプール中の単一試料と固有に関連付けられる。いくつかの実施形態において、バーコード配列は、配列 D から 5 に位置する。

【0123】

単一反応において、配列決定データ（例えば、R1および/またはR2配列）は、架橋増幅反応からの約または少なくとも約 10^6 、 10^7 、 10^8 、 2×10^8 、 3×10^8 、 4×10^8 、 5×10^8 、 10^9 、 10^{10} 、もしくはそれより多い標的ポリヌクレオチドまたはクラスタに対して生成され、反応におけるそれぞれの試料に対して約 10^4 、 10^5 、 10^6 、 2×10^6 、 3×10^6 、 4×10^6 、 5×10^6 、 10^7 、 10^8 、それ未満、またはそれ以上の標的ポリヌクレオチドまたはクラスタに対する配列決定データを含み得る。配列決定システムは、様々な出力データファイル型またはフォーマットのいずれかで配列決定データを出力することができ、*.fasta、*.csfasta、*.seq.txt、*.qseq.txt、*.fastq、*.sff、*.prb.txt、*.sms、*.srsおよび/または*.qvが挙げられるが、これらに限定されない。いくつかの実施形態において、約5、10、25、50、75、100、125、150、175、200、300、400、500、750、1000、2500、5000、7500、10000、20000、50000、もしくはそれより多いか、それ未満、またはそれ以上の原因となる遺伝的変異体の存在または非存在は、配列決定データに基づいて、試料に対して決定される。1つ以上の原因となる遺伝的変異体の存在、非存在、または対立遺伝子比は、約80%、85%、90%、95%、97.5%、99%、99.5%、99.9%、もしくはそれより高いか、またはそれ以上の精度で決定される。いくつかの実施形態において、1つ以上の非対象配列および/または1つ以上のAIMの存在、非存在、または量は、約80%、85%、90%、95%、97.5%、99%、99.5%、99.9%、もしくはそれより高いか、またはそれ以上の精度で決定される。

【0124】

いくつかの実施形態において、複数のクラスタに対する1つ以上の配列決定反応において特定された配列は、参照配列に対して整列される。一般に、アラインメントは、1つの配列を別の配列に沿って置くことと、ギャップをそれぞれの配列に沿って反復して導入することと、2つの配列がどれ程良好にマッチするかをスコア付することと、好ましくは、参照に沿って様々な位置に対して繰り返すことと、を伴う。最良のスコアマッチは、アラインメントであると見なされ、配列間の関係の程度についての推測を示す。いくつかの実施形態において、配列決定読み出し値が比較される参照配列は、参照ゲノム、例えば、対象と同じ種のメンバーのゲノムである。参照ゲノムは、完全であり得るか、または不完全であり得る。いくつかの実施形態において、参照ゲノムは、標的ポリヌクレオチドを含む領域のみからなる。いくつかの実施形態において、参照配列は、ヒトゲノムを含むか、またはそれからなる。いくつかの実施形態において、参照配列は、試験されるか、または試料が採取される個体以外の1つ以上の生物のポリヌクレオチドの配列、例えば、1つ以上の細菌、古細菌、ウイルス、原生生物、真菌、または他の生物からの配列を含むか、またはそれからなる。いくつかの実施形態において、参照配列は、複数の既知の配列、例えば、標的ポリヌクレオチド配列を増幅するために使用される全てのプローブ配列を含むか、またはそれらからなる（例えば、全ての異なる標的ポリヌクレオチドに対する全ての配列Bおよび/または配列B）。1つのプライマーの伸長から生成された配列決定データ（例えば、プライマーDからのR1配列）は、別のプライマーの伸長から生成された配列決定データ（例えば、プライマーAからのR2配列）と同じか、または異なる参照配列に整列され得る。1つのプライマーの伸長から生成された配列決定データは、参照配列に2回以上整列され得、それぞれのアラインメントは、異なるアラインメントアルゴリズムを使用する。R1配列は、R2配列から独立して整列され得、R1およびR2配列の第1

のアラインメントは、同じアラインメントアルゴリズムを使用し得る。

【0125】

アラインメントにおいて、参照における非マッチング塩基に沿った配列決定読み出し値中の塩基は、置換突然変異がその点で発生したことを示す。同様に、1つの配列が他の配列中の塩基に沿ってギャップを含む場合、挿入または欠失突然変異(「インデル」)が発生したことが推測される。1つの配列が互いに整列されることを特定することが所望される場合、アラインメントは、時として対合アラインメントと呼ばれる。複数の配列アラインメントは、一般に、例えば、一連の対合アラインメントを含む、2つ以上の配列のアラインメントを指す。いくつかの実施形態において、アラインメントをスコア付することは、置換およびインデルの確率に対して値を設定することを伴う。個別の塩基が整列される
10
とき、マッチまたはミスマッチは、置換可能性によるアラインメントスコアに寄与し、これは例えば、マッチに対して1、およびミスマッチに対して0.33であり得る。インデルは、ギャップペナルティによりアラインメントスコアから差し引き、例えば、ギャップペナルティは-1であり得る。ギャップペナルティおよび置換可能性は、配列がどのように突然変異するかについての経験的知識または演繹的仮定に基づき得る。それらの値は、得られるアラインメントに影響する。アラインメントを行うためのアルゴリズムの例としては、限定されないが、スミス-ウォーターマン(SW)アルゴリズム、ニードルマン-
20
ウンシュ(NW)アルゴリズム、パローズ-ホイラー変換に基づくアルゴリズム(BWT)、およびNovoalign(Novocraft Technologies; www.novocraft.comで入手可能)、ELAND(Illumina, San Diego, CA)、SOAP(soap.genomics.org.cnで入手可能)、およびMaq(maq.sourceforge.netで入手可能)等のハッシュ関数ライナーが挙げられる。

【0126】

いくつかの実施形態において、本発明に従うアラインメントは、コンピュータープログラムに使用して行われる。BWTアプローチを実装する1つの例示のアラインメントプログラムは、Geeknet(Fairfax, Va.)により維持されるSourceForgeウェブサイトから入手可能なパローズ-ホイラーライナー(BWA)である。BWTは、典型的に、1ヌクレオチド当たり2ビットのメモリを占有し、典型的なデスクトップまたはラップトップコンピュータを用いて、長さが4 G塩基対のヌクレオチド
30
配列をインデックス付することを可能にする。前処理は、BWTの構築(すなわち、参照をインデックス付する)および支持する補助データ構造を含む。BWAは、2つの異なるアルゴリズムを含み、ともにBWTに基づく。BWAによるアラインメントは、誤差率の低い(3%未満)最大約200 bpの短いクエリに対して設計されたアルゴリズムbwa-shortを使用して進めることができる(Li H. and Durbin R. Bioinformatics, 25:1754~60(2009))。第2のアルゴリズムであるBWA-SWは、より多くの誤差を含む長い読み出し値に対して設計されている(Li H. and Durbin R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler Transform. Bioinformatics, Epub.)。当業者
40
であれば、bwa-swが、時として「bwa-long」、「bwa longアルゴリズム」または同様に称されることを認識するであろう。

【0127】

スミス-ウォーターマンアルゴリズムのバージョンを実現するアラインメントプログラムは、Geeknet(Fairfax, Va.)により維持されるSourceForgeウェブサイトから入手可能なMUMmerである。MUMmerは、完全形態であってもドラフト形態であっても、全体ゲノムを速やかに整列させるためのシステムである(Kurtz, S., et al., Genome Biology, 5:R12(2004)、Delcher, A.L., et al., Nucl. Acids Res., 27:11(1999))。例えば、MUMmer 3.0は、2.4 GHz Linu
50

x デスクトップコンピュータ上で78MBのメモリを使用し、13.7秒以内に5 - メガベースゲノムの対の間の全20 - 塩基対以上の正確なマッチを見出すことができる。MUMmerは、不完全なゲノムを整列することもでき、ショットガン配列決定プロジェクトからの何百または何千のコンティグを容易に扱うことができ、システムとともに含まれるNUCmerプログラムを使用して、それらを別のコンティグの群またはゲノムに整列させる。

【0128】

アラインメントプログラムの他の非限定的な例としては、Kent Informatics (Santa Cruz, Calif.) からのBLAT (Kent, W. J., Genome Research 4:656~664 (2002))、Beijing Genomics institute (Beijing, CN) またはBGI Americas Corporation (Cambridge, Mass.) からのSOAP2、Bowtie (Langmead, et al., Genome Biology, 10:R25 (2009))、配列および異型のコンセンサスアセスメント (CASAVA) ソフトウェアのヌクレオチドデータベースの効率的な大規模アラインメント (ELAND) またはELANDv2 成分 (Illimuna, San Diego, Calif.)、Real Time Genomics, Inc. (San Francisco, Calif.) からのRTG Investigator、Novocraft (Selangor, Malaysia) からのNovoalign、European Bioinformatics Institute (Hinxton, UK) からのExonerate (Slater, G., and Birney, E., BMC Bioinformatics 6:31 (2005))、University College Dublin (Dublin, Ireland) からのClustal Omega (Sievers F., et al., Mol Syst Biol 7, article 539 (2011))、University College Dublin (Dublin, Ireland) からのClustalW またはClustalX (Larkin M. A., et al., Bioinformatics, 23, 2947~2948 (2007))、およびEuropean Bioinformatics Institute (Hinxton, UK) からのFASTA (Pearson W. R., et al., PNAS 85 (8):2444~8 (1988)、Lipman, D. J., Science 227 (4693):1435~41 (1985)) が挙げられる。

【0129】

いくつかの実施形態において、本発明のステップのうちのいずれか、または全ては自動化される。例えば、Perl スクリプトまたはシェルスクリプトを、上述の様々なプログラムのいずれかを起動するように書くことができる (例えば、Tisdall, Mastering Perl for Bioinformatics, O'Reilly & Associates, Inc., Sebastopol, Calif. 2003、Michael, R., Mastering Unix Shell Scripting, Wiley Publishing, Inc., Indianapolis, Ind. 2003 を参照)。代替として、本発明の方法は、1つ以上の専用プログラムにおいて全体的または部分的に具現化され得、例えば、それぞれが任意にC++等のコンパイル言語で書かれた後、バイナリとして編集および配布される。本発明の方法は、既存の配列分析プラットフォーム内で、またはその中で機能性を起動することにより、モジュールとして全体的または部分的に実現され得る。ある実施形態において、本発明の方法は、全て単一の開始キュー (例えば、ヒトの活動、別のコンピュータプログラム、またはマシン) からもたらされる引き金となるイベントの1つまたは組み合わせに反応して自動的に起動される、多くのステップを含む。したがって、本発明は、それらのステップのうちのいずれか、またはそれらのステップの任意の組み合わせがキューに反応して自動的に起こり得る方法を提供する。出力は、コンピュータファイルのフォーマットで提供され得る。ある

10

20

30

40

50

実施形態において、出力は、参照ゲノムの配列に整列された核酸の配列等の配列データを含むFASTAファイル、VCFファイル、テキストファイル、またはXMLファイルである。他の実施形態において、出力は、参照ゲノムに対して対象の核酸における1つ以上の突然変異を説明する座標または文字列を含む。当該技術分野において既知のアラインメント文字列としては、Simple UnGapped Alignment Report (SUGAR)、Verbose Useful Labeled Gapped Alignment Report (VULGAR)、およびCompact Idiosyncratic Gapped Alignment Report (CIGAR) (Ning, Z., et al., Genome Research 11(10): 1725~9 (2001))が挙げられる。いくつかの実施形態において、出力は、配列アラインメント、例えば、CIGAR文字列を含む配列アラインメントマップ(SAM)またはバイナリアラインメントマップ(BAM)ファイルである(SAMフォーマットは、例えば、Li, et al., The Sequence Alignment/Map format and SAM tools, Bioinformatics, 2009, 25(16): 2078~9に記載されている)。いくつかの実施形態において、CIGARは、1ライン当たり1つのギャップトアラインメントを表示するか、または含む。CIGARは、圧縮された対合アラインメントフォーマットでレポートされたCIGAR文字列である。

【0130】

いくつかの実施形態において、クラスタからのR1配列は、複数の異なる標的ポリヌクレオチドからの配列Gを含み、クラスタからのR2配列は、配列Bを含み、配列Bは、増幅した二本鎖のクラスタを生成するために使用されるプローブ配列である。それぞれの配列Bが、特定の標的ポリヌクレオチドを標的するように選択されるとき、参照配列(例えば、参照ゲノム)内のその配列および位置は一般に知られており、同じクラスタからのR1配列は、予想されるヌクレオチド距離内にあると予期され得る。予想されるヌクレオチド距離は、断片化試料ポリヌクレオチドを含む試料の断片長の平均もしくは中央値、またはそのような中央値もしくは平均断片長に基づいてあり得ない断片長を表す上閾値距離に基づき得る。したがって、いくつかの実施形態において、同じクラスタからのR2配列から閾値距離よりさらに離れた位置に整列するR1配列は、誤りであるため破棄される。いくつかの実施形態において、それを上回るクラスタの配列読み出し値が破棄される、同じクラスタからの整列されたR1配列とR2配列との間の参照配列に沿った上限閾値距離は、約1000、2500、5000、7500、10000、12500、15000、20000塩基対、もしくはそれより多いか、またはそれ以上である。いくつかの実施形態において、参照配列(例えば、参照ゲノム)の非固有領域に対するR1配列のアラインメントは破棄され、その配列は、参照配列内の固有の配列のより小さなサブセットに再整列される。

【0131】

典型的に、塩基品質スコアは、配列決定結果において、それぞれのヌクレオチドに対して決定され、特定の塩基呼び出しが誤りである確率に関する。塩基品質スコアの例は、Phred品質スコアQ、 $Q = -10 \log_{10} P$ であり、式中、Pは、対応する塩基呼び出しが正しくない確率を表す。いくつかの実施形態において、塩基品質スコアは、例えば、複数のアラインメントのそれぞれに対してマッピング品質スコアを決定することにより、参照配列に対する配列決定読み出し値のアラインメントを評価するために使用される。マッピング品質スコアを計算するための方法は、当該技術分野において既知である。例えば、閾値を下回る品質スコアを有するアラインメントは、破棄されるか、再整列されるか、またはより高いスコアを有する代替アラインメントと置き換えられ得る。いくつかの実施形態において、閾値を下回るマッピング品質スコアを持ち、複数の最適アラインメントを有するアラインメントは、参照配列内の配列のサブセット、例えば、標的ポリヌクレオチドを含む参照ゲノムの領域のみに再整列される。いくつかの実施形態において、閾値マッピング品質スコアは、約100、75、50、25、20、10、5、4、3、

2、1、もしくは0、またはそれ未満である。

【0132】

いくつかの実施形態において、重複する可能性が高い配列決定読み出し値は、最初のアラインメントに続いて除去される。配列決定読み出し値がマップされるとき、重複読み出し値は、アラインメントアルゴリズムにより重複としてマークされ得る。例えば、アラインメントアルゴリズム内のマーク重複サブルーチンは、整列された配列のファイル（例えば、*.BAMファイル）内のレコードの全てを審査し、どの読み出し値が他の読み出し値の重複であるかを決定する。一般的に言えば、2種類の重複が存在する：典型的に一次解析ソフトウェアにおける欠陥により引き起こされる光学的重複、および重複PCR反応により引き起こされるPCR重複。しかしながら、計算の観点から、光学的重複とPCR重複は区別できない。2つの読み出し値が重複であるか否かを決定する1つの方法は、塩基配列を比較することであり、2つの重複読み出し値は、重複塩基配列を有するはずである。しかしながら、配列決定エラーに起因して、それは2つの重複読み出し値が配列決定される場合であり得、1つの読み出し値の配列決定エラーが、その塩基配列を他の読み出し値とは著しく異なるようにする。したがって、塩基配列を比較して2つの読み出し値が重複するかどうかを決定するのではなく、代わりに、それらのアラインメント比較することができる。2つの読み出し値が重複である場合、次に、両方の読み出し値のアラインメントの群全体は、全体的に同じである。いくつかの実施形態において、重複は、アラインメントアルゴリズムとは別個の1つ以上のアルゴリズムを使用して除去および/または破棄されるようにマークされる。一般に、バーコード配列が使用されるとき、配列決定読み出し値は、同じバーコード配列分類内で発生するときに単に削除される。

【0133】

いくつかの実施形態において、第2のアルゴリズムを使用する第2のアラインメントは、第1のアルゴリズムを使用する第1のアラインメントの後に行われる。第2のアラインメントは、第1のアラインメントと同じ参照配列に関して、第1のアラインメントにおいて使用されるか、または参照配列を使用しないものとは異なる参照配列であり得る（例えば、全ての配列決定が重なるとき、特定の領域が互いに整列される）。例えば、第1の参照配列に関して挿入および/または欠失（インデル）を含有する可能性が高いと第1のアラインメントにおいて特定された配列は、標的ポリヌクレオチドに含有される挿入および/または欠失に対して単一のコンセンサス配列を生成するように局所的に整列され得る。第1のアラインメントは、個別の配列を参照配列に独立して整列し得る。場合によっては、真のインデルを持つ配列決定読み出し値は、複数のミスマッチを持つアラインメントモデルが、インデル含有アラインメントより高いスコアであるとき、インデルではなく複数のミスマッチと整列され得る。典型的に、複数の配列が、単一のヌクレオチド位置と重なるように（例えば、タイル状に）整列される。予測される量を超える配列変異を含有する重なる領域（例えば、ヒト対象のゲノム中の固有の遺伝子座に対する2つより多くの対立遺伝子）は、高い可能性のインデルの存在を示し得る。特定の参照配列に対するいくつかのインデルの位置は既知であり得、既知のインデルの位置と重なる配列は、その配列がインデルを含有する可能性が高いと特定するようになる。インデルを含む可能性は、1つ以上のそのような因子に基づいて、例えば、少なくとも約60%、70%、80%、90%、95%、99%、もしくはそれより高い可能性のように、数値で表され得る。いくつかの実施形態において、原因となる遺伝的変異体のような関心領域に重なり、また任意にインデルを含むか、または含む可能性がある全ての配列は、関心領域に対して単一のコンセンサス配列を生成するために、第2のアルゴリズムを使用して局所的に整列される。関心領域は、任意の好適なサイズ、例えば、約5、10、15、20、25、50、100、250、500、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長であり得る。第2のアラインメントは、1つ以上のヌクレオチド位置に重なる全ての配列決定読み出し値の局所的複数配列アラインメントであり得る。いくつかの実施形態において、第2のアラインメントは、ある位置での全ての配列決定のアラインメントを最適化することにより、単一のコンセンサス配列を特定する。いくつかの実施形態において、第2

10

20

30

40

50

のアラインメントにより生成されたコンセンサス配列は、参照配列に関して、コンセンサス配列を生成するために再整列された配列の1つ以上より少なくとも1、2、3、4、5、6、7、8、9、10、15、20、または25より少ないヌクレオチドミスマッチを含有する。いくつかの実施形態において、第2のアラインメントを行うために使用されるアルゴリズムは、参照配列と比較して、約1、2、3、4、5、6、7、8、9、10、15、20、25、もしくはそれより多いか、またはそれ以上のヌクレオチドの挿入および/または欠失を、約80%、85%、90%、95%、97%、99%、もしくはそれより高いか、またはそれ以上の精度で特定することができる。

【0134】

典型的に、第2のアルゴリズムは、第1のアルゴリズムとは異なり、第2のアルゴリズムは、同じ数のアラインメントを行うために、アルゴリズムを実行するシステム（例えば、コンピューターシステム）のより多くの資源を必要とし得る。例えば、第1のアルゴリズムを使用するシステムを用いて第1のアラインメントを行うことは、そのシステムが第2のアルゴリズムを使用して、全てのR1読み出し値の第1のアラインメントを行う場合にかかる時間よりも短い時間で全てのR1読み出し値を整列し得る。いくつかの実施形態において、第1のアルゴリズムを用いて第1のアラインメントを行うことは、第2のアルゴリズムを使用して同じシステムにより第1のアラインメントを行うためにかかる時間より約90%、80%、70%、60%、50%、40%、30%、25%、20%、15%、10%、5%、もしくはそれより短い、またはそれ以下の時間がかかる。さらなる例として、第1のアルゴリズムを使用するシステムを用いて第1のアラインメントを行うことは、このシステムが第2のアルゴリズムを使用して全てのR1読み出し値の第1のアラインメントを行う場合に使用されるよりも少ないシステムメモリを使用して、全てのR1読み出し値を整列することができる。いくつかの実施形態において、第1のアルゴリズムを用いて第1のアラインメントを行うことは、第2のアルゴリズムを使用して第1のアラインメントを行うために同じシステムにより使用されるよりも約90%、80%、70%、60%、50%、40%、30%、25%、20%、15%、10%、5%、もしくはそれより少ないか、またはそれ以下のメモリを使用する。第1のアルゴリズムは、経験則を使用して第1のアラインメントを行うことができる。いくつかの実施形態において、第1のアルゴリズムは、パローズ - ホイラー変換、例えば、パローズ - ホイラーライナーに基づく。

【0135】

いくつかの実施形態において、試料からの1つ以上の標的ポリヌクレオチドから誘導される増幅配列（例えば、全てのクラスタの少なくとも75%、80%、85%、90%、95%、または100%からの配列G）は、1ヌクレオチド長から約10、25、50、100、250、500、1000、2000、5000、もしくはそれより多いか、それ未満、またはそれ以上のヌクレオチド長である。一般に、クラスタの標的ポリヌクレオチドから誘導された増幅配列（例えば、配列G）が、増幅配列（「読み出し値長」とも称される）を配列決定する際に行われるヌクレオチド伸長のサイクル数より短い場合、その配列決定読み出し値に戻された配列データは、最初に増幅配列（例えば、BまたはB₁）を捕捉するために使用される第1のオリゴヌクレオチドの配列を含有する可能性が高い。配列BまたはB₁が参照配列（例えば、参照ゲノム）に存在するとき、配列は正しく整列し得るが、標的ポリヌクレオチド中の任意の真の突然変異は、マスクされ得るか、またはより低い信頼度で推測され得る。R1配列に含有される第1のオリゴヌクレオチド配列の負の効果を避けるために、クラスタの配列BまたはB₁に対応する可能性のある塩基呼び出しは、欠失され得る。クラスタの配列BまたはB₁は、例えば、R2配列を生成する別個の配列決定反応において特定され得る。次に、R1配列は、同じクラスタのR2配列と比較され、R1の1つ以上のヌクレオチドが配列BまたはB₁に対応するか否かを決定し得る。クラスタについてR2配列（または任意の配列Bを含むR2配列）が得られない場合、第1のオリゴヌクレオチド配列を欠失することは、欠失されるR1配列の一部が任意の配列B₁の少なくとも一部分と同一であり（例えば、任意のクラスタにおいて見出さ

れるか、または標的ポリヌクレオチドを増幅するために使用される任意の配列 B に対応する任意の B)、この部分が、R 1 の 5 または 3 ヌクレオチドのいずれかを含み、(i) いかなる R 2 配列もクラスタに対して生成されなかったか、または (i i) 生成された R 2 配列が任意の配列 B と同一でないかのいずれかであるとき、クラスタの R 1 配列の一部分を欠失することを含み得る。一般に、R 1 配列の内部に見出される配列 B または B を含むヌクレオチド配列 (つまり、配列決定読み出し値の 5 または 3 末端を含まない) は、増幅した配列が、さらに離れた配列 B を使用して捕捉されたことを示す。

【 0 1 3 6 】

いくつかの実施形態において、本発明の方法により検出された遺伝的変異を使用して、複数の確率を計算する。それぞれの確率は、疾患または形質を有するか、または発症する対象または対象の現在もしくは将来の子孫の確率であり得る。いくつかの実施形態において、それぞれの確率は、対象の R 1 配列に基づき、1 つ以上のそのような確率は、分析結果のレポートに含まれ得る。一般に、試験された対象が疾患または形質を有するか、または発症する確率の計算は、1 つ以上の試験された原因となる遺伝的変異体、非対象配列、および / または A I M と関連付けられた危険性のレベルに基づく。例えば、2 つの原因となる遺伝的変異体が、付加的に疾患を発症する危険性に寄与する場合、対象における両方の原因となる遺伝的変異体の存在は、対象におけるその疾患の危険性が、それぞれと関連付けられる危険性を付加することから得られる値だけ増加することを示す。一般に、対象の子孫が疾患または形質を有する確率の計算は、1 つ以上の試験された原因となる遺伝的変異体および / または A I M、および子孫がその原因となる遺伝的変異体および / または A I M を受け継ぐ可能性と関連付けられた危険性のレベルに基づく。危険性計算は、1 つ以上のデータベースに維持される危険性相関に基づき得、そのデータベースは、試験された対象の遺伝子型決定結果および関連表現型に関する外部レポートおよび / または記録に基づいて更新され得る。いくつかの実施形態において、計算は、コンピューター可読媒体に含まれる命令に従い、コンピューターにより行われる。いくつかの実施形態において、対象または対象の子孫が疾患または形質を有するか、または発症する確率の統計的信頼度は、少なくとも約 70 %、80 %、85 %、90 %、95 %、97.5 %、99 %、もしくはそれより高い。信頼度は、配列決定精度における信頼度、試験される関連遺伝的変異体の数、およびそれぞれの遺伝的変異体との危険性関連における信頼度等の多数の因子に基づき得る。確率を計算するための例示の方法は、米国公開第 20100022406 号に記載される。

【 0 1 3 7 】

本発明の任意の態様のいくつかの実施形態において、コンピューターシステムを使用して、記載される方法の 1 つ以上のステップを実行する。図 8 は、本発明の方法において有用なコンピューターシステムの非限定的な例を示す。いくつかの実施形態において、コンピューターシステムは、液体ハンドラー、架橋増幅システム (例えば、Illumina cBot)、および / または配列決定システム (例えば、Illumina ゲノム分析器、HiSeq、または MiSeq システム) 等の分析システムに統合され、その一部である。いくつかの実施形態において、コンピューターシステムは、分析システムに接続されるか、または移植される。いくつかの実施形態において、コンピューターシステムは、ネットワーク接続により分析システムに接続される。コンピューターシステム (またはデジタルデバイス) は、結果を受信および記憶し、その結果を分析し、および / または結果および分析のレポートを生成するために使用され得る。コンピューターシステムは、媒体 (例えば、ソフトウェア) および / またはネットワークポート (例えば、インターネットから) からの命令を読み出すことができる論理装置として理解され得、任意に固定媒体を有するサーバーに接続され得る。コンピューターシステムは、CPU、ディスクドライブ、キーボードおよび / またはマウス等の入力デバイス、およびディスプレイ (例えば、モニター) のうちの 1 つ以上を含み得る。データ通信、例えば、命令またはレポートの伝達は、ローカルまたはリモート位置にあるサーバーに対して通信媒体を通じて達成され得る。通信媒体は、データを伝達および / または受信する任意の手段を含み得る。例えば、通

信媒体は、ネットワーク接続、ワイヤレス接続、またはインターネット接続であり得る。そのような接続は、ワールドワイドウェブ上の通信を提供することができる。本発明に関するデータは、受信のため、および/または受信者によるレビューのために、そのようなネットワークまたは接続（あるいは情報を伝達するための任意の他の好適な手段、プリントアウト等の物理的レポートを郵送することを含むが、これに限定されない）上で伝達され得ることが想定される）。受信者は、限定されないが、個人、ヘルスケア提供者、ヘルスケア管理者、または電子システム（例えば、1つ以上のコンピューター、および/または1つ以上のサーバー）であり得る。いくつかの実施形態において、コンピューター可読媒体は、生物学的試料の分析の結果の伝達に好適な媒体を含む。この媒体は、個人の遺伝子プロファイルの分析に関する結果を含むことができ、そのような結果は、本明細書に記載される方法を使用して誘導される。データおよび/または結果は、モニター等のディスプレイ上にいつでも表示され得、遺伝的レポートの形態で記憶または印刷され得る。

【0138】

表現型と関連付けられる原因となる遺伝的変異体は、化学的文献から得られ、対象からの試料の配列結果との比較のためにコンピューターシステムに送信され得る。原因となる遺伝的変異体の遺伝子型および生物学的試料からの結果は、コンピューターシステム（または他のデジタルデバイス）に送信され、記憶され、分析され得、ゲノムデータの結果および分析のレポートを生成する。これらの結果および分析は、ヘルスケア提供者のような受信者により、オンラインポータルまたはウェブサイトを介してオンラインでアクセスされ得る。これらの結果および分析は、オンラインで見られるか、受信者のコンピューターに保存されるか、印刷されるか、または受信者に郵送され得る。結果は、例えば、医師または他の医療従事者の指示において、個人向け健康管理に使用され得る。例えば、対象は、遺伝カウンセリングを受けるように遺伝カウンセラーに照会または接触され得る。

【0139】

データベースは、例えば、本発明の方法により生成された配列決定結果についてより多くの情報を提供する、様々な任意の構成要素のうちの1つ以上を有し得る。いくつかの実施形態において、コンピューターが特定された原因となる遺伝的変異体と関連付けられた機能を実行するための命令を含む、コンピューターで実行可能なソフトウェアでコードされたコンピューター可読媒体が提供される。そのようなコンピューターシステムは、遂行されることが所望される評価の種類に応じて、そのようなコードまたはコンピューターで実行可能なソフトウェアの任意の組み合わせを含み得る。コンピューターシステムは、配列（例えば、原因となる遺伝的変異体の遺伝子型）のそれぞれを少なくとも1つの表現型、例えば、医学的状態（その表現型を有するか、または発症する危険性が挙げられるが、これに限定されない）等の状態に関連付けるためのコードも有し得る。それぞれの医学的状態を、次いで医療専門家による少なくとも1つの推奨およびその推奨を含むレポートを生成するためのコードに関連付けることができる。このシステムは、レポートを生成するためのコードも有し得る。異なる種類のレポート、例えば、受信者が希望するか、または支払った詳細レベルに基づくレポートが生成され得る。例えば、受信者は、状態のような単一表現型に対する分析を注文した可能性があり、したがってレポートは、条件等のその単一表現型の結果を含み得る。別の受信者は、パネルまたは器官系の遺伝子プロファイルを要求した可能性、または別の個人は、全ての臨床的に関連する原因となる遺伝的変異体の分析を含む包括的な遺伝子プロファイルを要求した可能性がある。レポートは、対象の情報（例えば、氏名、生年月日、民族グループ、試料の種類、試料採取日、および/または試料受領日）；分析方法（複数可）の説明；試験された全ての原因となる遺伝的変異体の結果；試験された全ての疾患または形質の結果；陽性スコアを有する疾患または形質の結果（例えば、閾値レベルを上回る危険性、例えば、約1/50000、1/25000、1/10000、1/5000、1/2500、1/1000、1/500、1/100、1/50、1/10、もしくはそれより高いか、またはそれ以上）；陽性スコアを有する疾患または形質と関連付けられる原因となる遺伝的変異体の結果；2人以上の個人の結果（例えば、親であるか、または子供を持つことを計画している個人）；疾患または形

質を有するか、または発症する危険性；現在または将来の子が、疾患または形質を有するか、または発症する危険性；胎児が疾患または形質を有するか、または発症する危険性；危険性計算の方法；さらなる措置についての推奨のうちの1つ以上を含み得る。

【0140】

生成されたレポートは、遺伝カウンセラーおよび/または他の医療専門家、例えば、管理医師もしくは有資格の医師、または他の第3者によりレビューされ、さらに分析され得る。遺伝カウンセラーもしくは医療専門家、または両者、あるいは他の第3者は個人と面会して、その結果、分析、および遺伝レポートについて話し合うことができる。話し合いは、原因となる遺伝的変異体（複数可）、例えば、試験された原因となる遺伝的変異体（複数可）（存在、非存在、および/または遺伝子型）、その原因となる遺伝的変異体（複数可）が、どのように遺伝または伝達され得るか（例えば、質問票から生成された家系図を使用する）、原因となる遺伝的変異体（複数可）の保有率；関連表現型の保有率または発生率；および関連表現型についての情報（例えば、医学的または臨床的に関連する状態等の特定の状態または形質）、例えば、その表現型がどのように個人に影響し得るか、および講じられ得る予防策についての情報を含む。遺伝カウンセラーまたは医療専門家は、他の遺伝情報またはその分析における質問票および個人との話し合いからの情報等の他の情報を組み込むことができる。条件または形質等の表現型についての情報は、個人の医師または他のヘルスケア提供者のためのさらなる遺伝カウンセリング等のフォローアップ提案等の推奨、予測医療の推奨、または予防医療の推奨を含み得る。スクリーニング情報、例えば、乳癌スクリーニングの方法は、例えば、個人が、より高い乳癌の危険性にあることが見出された場合に話し合われ得る。話し合われ得る他のトピックとしては、ライフスタイルの修正および薬物治療が挙げられる。例えば、ライフスタイルの修正が提案され得、例えば、食事変化および特定の食事計画が推奨され得るか、運動計画が提案され得、特定の運動施設またはトレーナーが個人に照会され得る。一般的な誤解が含まれることもあるが、個人が、それが機能するか、または有用であると思っているかもしれないが、公開された文献において有益でないか、または実際に有害であることが示された予防措置または他の介入に気付くのを可能にする。代替療法、例えば、栄養補助食品等の代替医療、または鍼療法もしくはヨガ等の代替療法が含まれ得る。家族計画のオプションとともに、表現型の存在または表現型の進行を検出するか、または監視を助けることができるスクリーニング審査または研究室試験等のモニタリングオプションが含まれ得る。表現型、例えば、個人が素因を持っている疾患を予防することができ、その発症を制限するか、またはその進行を遅延し得る医薬、または高い効能と低い副作用を持つ医薬、あるいは有害反応（複数可）の可能性に起因して個人が避ける必要がある医薬または医薬群が個人に提案され得る。例えば、医療従事者は、代謝、効能、および/または安全性を含む、可能性の高い個人の薬物反応の評価を行うことができる。医療従事者は、個人の遺伝プロフィールからの情報のみに基づいて、または個人の環境要因（例えば、ライフスタイル、習慣、診断された病状、現在の薬物治療、およびその他）についての情報と併せて、処置の潜在的な必要性または処置の効果について、予防的処置およびモニタリング（例えば、診察および検査、放射線検査、自己検査、または研究室試験）等の治療処置を話し合うこともできる。追加の資源が列挙され得、例えば、個人または個人の医師もしくは他のヘルスケア専門家が、表現型、原因となる遺伝的変異体（複数可）、または両方についての追加の情報、例えば、表現型に関する情報を含むウェブサイトへのリンク、例えば、その表現型の全国組織等の遺伝レポートまたは外部ウェブサイトを生成する企業からの内部ウェブサイトを得ることを含む。追加の資源は、個人が表現型、原因となる遺伝的変異体（複数可）、または両方についてより多くの情報を得るために求め得る電話番号、書籍、または人々への照会も含み得る。

【0141】

一態様において、本発明は、第1のサービスと、任意に第2のサービスを提供することを含む方法を提供し、a) 第1のサービスは、夫婦の子が、原因となる遺伝的変異体により引き起こされた複数の形質のそれぞれを有する可能性を予測することを含み、この予測

10

20

30

40

50

は、その夫婦の2人の個人のそれぞれの遺伝子型に基づき、b)第2のサービスは、複数の形質に対してその夫婦の子の可能性のある表現型を予測することを含み、この可能性は、それぞれの表現型および/またはその夫婦の個人の家族歴に基づいて決定される。一実施形態において、少なくとも1つの予測は、さらに個人のそれぞれの遺伝的に推測される祖先に基づく。別の実施形態において、第1のサービスは、有料サービスとして提供され、第2のサービスは、無料サービスとして提供される。

【0142】

一態様において、本発明は、a)夫婦のそれぞれのメンバーから得た家族歴情報を記憶するように構成されたコンピューター可読媒体と、b)その夫婦のそれぞれのメンバーについての遺伝情報を含むデータを記憶するように構成されたコンピューター可読媒体と、c)実行されると、i)遺伝情報において特定された対立遺伝子により引き起こされた形質に関して、それぞれの個人の保因状態を予測するか、またはii)家族歴および/または遺伝情報により決定され可能な夫婦の子の可能性のある形質を予測する、コンピューターコードを含むコンピューター可読媒体と、d)i)その夫婦の少なくとも1人のメンバーの保因状態、またはii)その子の可能性のある形質を表示するディスプレイと、を備えるシステムを提供する。いくつかの実施形態において、システムは、e)DNA試験キットを購入する申し出を受け入れるように構成されたウェブページをさらに備える。いくつかの実施形態において、ディスプレイは、電子的であり、例えば、ウェブページである。いくつかの実施形態において、システムは、e)遺伝情報に基づいて遺伝カウンセラーおよび/または他の医療従事者(例えば、遺伝医学者または産科医/産婦人科医)への照会を表示するディスプレイをさらに備える。

【0143】

インターネットおよびワールドワイドウェブは、情報へのアクセスおよび配布を提供する。いくつかの実施形態において、ウェブサイトは、顧客が遺伝子検査を購入し、遺伝子検査の結果を受け取るのを許容するために様々な機能性を効率良く提供するのに特に適切であり得る。システムは、典型的に、ウェブサイトが存在するサーバーを含む。ユーザーは、コンピューターモニターまたは電話スクリーン等のサーバーに接続されたインターフェースを使用して、情報をポップアップ表示するか、またはユーザーを別のウェブページに導くリンクをクリックまたはロールオーバーすることにより、ウェブサイトとインタラクトする。ウェブサイトは、典型的に双方向性であり、ユーザーがインターフェース上に情報またはクエリを入力し、応答を得るのを許容する。

【0144】

システムおよびビジネス方法のいくつかの実施形態において、ウェブサイトは、顧客が遺伝子検査の結果を購入、管理、および見ることとともに、将来の子孫が疾患または形質を発症する確率についてより広く学ぶのを許容することができる。例えば、顧客は、彼らの子孫がメンデル病を発症する危険性があるかどうかを学ぼうとする、これから親になる夫婦であり得る。顧客には、(i)顧客の保因状態、(ii)顧客が1つ以上の疾患または形質を発症する可能性、および(iii)顧客の子孫が1つ以上の疾患または形質を発症する確率のうちの1つ以上を、顧客のDNAにおいて特定された原因となる遺伝的変異体に基づいて決定するための遺伝子検査を購入する申し出が提示され得る。

【0145】

顧客が遺伝子検査を購入することを選ぶ場合、次に顧客は、遺伝子検査、その会社の製品に関する遺伝カウンセラーとの直接電話相談、および/または遺伝カウンセラーおよび/もしくは他の関連医療従事者への照会と引き換えに、例えば、オンラインクレジットカード取引を通じて料金を支払い得る。遺伝子検査および照会は、購入時に料金を支払うことができるか、または初回ユーザー登録料に含めることができる。いくつかの実施形態において、サービスは無料であり、収益は、他の製品を特定の製品と併せて宣伝することによりその会社により生成される。例えば、顧客がオンライン注文をした後、その注文は処理するためにサーバーに送られる。支払いが確認されると、注文処理サーバーは、電子通知を出荷ベンダーに送信し、DNA採取キットを顧客に郵送することができる。一実施形

態において、DNA採取キットは、遺伝子検査サービスとは別個であるか、あるいはユーザーもしくは顧客が既にDNA採取キットを有しているか、または別の供給源から入手している。注文確認および注文に関する更新および出荷状態を含む通知を、定期的に顧客に電子的送信することもできる。本発明のビジネス方法のいくつかの実施形態において、顧客は、試料を採取キットに投入することができる。当業者に明らかな任意の試料を、採取キット中または上に投入することができる。試料は、分析される核酸を含有する任意の材料であり得、唾液または血液のような体液のように当業者に明らかである。次に、採取キットは、遺伝子型決定研究室に送るために会社に戻され得るか、または処理するために遺伝子型決定研究室に直接戻され得る。会社内ある、会社と協働するように契約している、またはその会社外にある遺伝子型決定研究室は、顧客のDNAを提供された試料から単離することができる。DNAが試料から単離された後、遺伝子型決定デバイス（例えば、本明細書に記載される装置）を使用して、（i）祖先情報提供マーカー、（ii）原因となる遺伝的変異体、および（iii）非対象配列（それらのうちの1つ以上は、本明細書において未加工の遺伝子型情報とも称される）のうちの1つ以上の存在についてDNAを検査することができる。いくつかの実施形態において、DNAは、未加工の遺伝子型情報の存在についてDNAを検査するために、試料から単離される必要はない。

10

【0146】

未加工の遺伝子型情報は、記憶および処理のためにサーバーに電子的送信され得る。サーバー上のコンピューターコードが、未加工の遺伝子型情報に対して実行され、顧客の祖先を推測する、および/または存在する場合は原因となる遺伝的変異体および/または非対象配列の存在を確認することができる。次に、処理された遺伝子型情報は、サーバーに電子的に送信されることができ、サーバー上のコンピューターコードが、処理された遺伝子型情報に対して実行され、顧客の子孫が、顧客の処理された遺伝子型情報に存在することが見出された原因となる遺伝的変異体により引き起こされた複数の形質のうちのそれぞれを有する可能性を予測することができる。次に、結果は、記憶するためにサーバーに電子的に伝送され得る。

20

【0147】

一例において、通知が、結果の入手可能性を顧客に警告するために顧客に送信され得る。この通知は、電子的であり得、その非限定的な例としては、テキストメッセージ、Eメール、または他のデータパケットが挙げられ、または通知は非電子的であってもよく、その非限定的な例としては、遺伝カウンセラーからの電話または郵送されるレポート等の印刷通信物が挙げられる。顧客に提供される結果は、1つ以上の疾患または形質について、顧客の保因状態、および/または顧客もしくは顧客の将来の子孫が1つ以上の疾患もしくは形質を発症する可能性を顧客に知らせることができる。顧客が結果および照会を受け取った後、顧客の注文は満たされたと見なすことができ、結果および照会は、オンラインウェブサイトアカウントを通じて顧客にアクセス可能のままであり得る。次に、顧客がさらにオフラインで照会を購入することを望むが、ウェブサイトの権限外にある場合、顧客はそうするように選択することができる。

30

【0148】

一態様において、本発明は、上記の方法において使用され得る組成物を提供する。本発明の組成物は、本明細書に記載される要素のうちのいずれか1つ以上を含むことができる。例えば、組成物は、結合されるオリゴヌクレオチドを含む1つ以上の固体支持体、固体支持体への結合のための1つ以上のオリゴヌクレオチド、1つ以上のアダプターオリゴヌクレオチド、1つ以上の増幅プライマー、第1の結合パートナーを含む1つ以上のオリゴヌクレオチドプライマー、第2の結合パートナーを含む1つ以上の固体表面（例えば、ビーズ）、1つ以上の配列決定プライマー、これらのいずれのいずれかを利用するための試薬、これらのうちのいずれかを含む反応混合物、およびこれらのうちのいずれかを使用するための指示のうちの1つ以上を含み得る。

40

【0149】

一態様において、本発明は、上記の方法および組成物において開示される要素のうちの

50

いずれか1つ以上を含有するキットを提供する。いくつかの実施形態において、キットは、1つ以上の容器に本発明の組成物を含む。例えば、キットは、それに結合されるオリゴヌクレオチドを含む1つ以上の固体支持体、固体支持体への結合のための1つ以上のオリゴヌクレオチド、1つ以上のアダプターオリゴヌクレオチド、1つ以上の増幅プライマー、第1の結合パートナーを含む1つ以上のオリゴヌクレオチドプライマー、第2の結合パートナーを含む1つ以上の固体表面（例えば、ビーズ）、1つ以上の配列決定プライマー、これらのいずれのいずれかを利用するための試薬、これらのうちのいずれかを含む反応混合物、およびこれらのうちのいずれかを使用するための使用説明書のうちの1つ以上を含み得る。いくつかの実施形態において、このキットは、（a）DNAリガーゼ、（b）DNA依存性DNAポリメラーゼ、（c）RNA依存性DNAポリメラーゼ、（d）ランダムプライマー、（e）少なくとも4つのチミジンを3'末端に持つプライマー、（f）DNAエンドヌクレアーゼ、（g）3'～5'エキソヌクレアーゼ活性を有するDNA依存性DNAポリメラーゼ、（h）それぞれのプライマーが複数の選択された配列のうちの1つを有する複数のプライマー、（i）DNAキナーゼ、（j）DNAエキソヌクレアーゼ、（k）磁性ビーズ、および（l）キットに含まれる要素のうちの1つ以上に好適な1つ以上の緩衝液のうちの1つ以上をさらに含む。アダプター、プライマー、他のオリゴヌクレオチド、および試薬は、限定されないが、本明細書に記載されるもののいずれかであり得る。キットの要素はさらに、限定されないが、任意の量および/または組み合わせで（例えば、同じキットまたは同じ容器で）提供され得る。キットは、本発明の方法に従って使用するための追加の薬剤をさらに含み得る。キット要素は、任意の好適な容器中に提供され得、容器には、試験管、バイアル瓶、フラスコ、ボトル、アンプル、シリンジ等が挙げられるが、これらに限定されない。薬剤は、本発明の方法において直接使用され得る形態、または使用前に調製を必要とする形態、例えば、凍結乾燥した薬剤の再構成で提供され得る。薬剤は、単回使用のためのアリコート、または複数使用、例えば多数の反応において得られ得るストックとして提供され得る。

【実施例】

【0150】

以下の例は、本発明の様々な実施形態を説明する目的で示され、本発明をいかなる方法においても制限するものではない。本実施例は、本明細書に記載される方法とともに、現在の好ましい実施形態を代表し、例示的であって、本発明に対する制限として意図されない。特許請求の範囲により定義される本発明の趣旨に包含されるそこでの変化および他の使用は、当業者であれば思い付くであろう。

【0151】

実施例1：試料調製および配列決定プロセス

ゲノムDNA（gDNA）を、96-ウェルフォーマットに抽出し、ウェルA1、G12、およびH12は空のまま残す（後に、それぞれ無テンプレート対照、試験された全ての原因となる遺伝的変異体を欠失するCoriell試料NA12878ゲノムDNAを含有する汎用陰性標準、および複数の既知の原因となる遺伝的変異体のうちの1つを含む試料を含有する）。それぞれのウェルから50μLを、吸光度プレートの対応するウェルに移す。260nmでの吸光度を、DNA量を計算するために、Tecan M200プレートリーダーを使用して測定する。50μLのgDNAを、吸光度プレートからEppendorf twin.tecプレートに移す。対照試料を、このtwin.tecプレート上のそれらそれぞれの位置に付加する。gDNAおよび対照を、以下のプロトコルに従い、10'でSonicMan（Matrical, Spokane WA）超音波破砕機内で断片化する：前冷却180秒、サイクル100、超音波破砕3.0秒、パワー35%、蓋冷却1.0秒、プレート冷却0、後冷却0。2μLの試料を、Fragment Analyzer（Advanced Analytical Technologies, Ames IA）を使用して、断片化サイズ分布について分析する。少なくとも200塩基対および1000bpを超えない断片サイズの中央値を有する試料が、さらな

る処理に供される。200bpを下回る断片サイズの中央値を持つ試料は破棄され、抽出されたgDNAから再処理される。1000bpを上回る断片サイズの中央値を持つ試料は、所望のサイズ範囲に達するようにさらなる超音波破碎に供されるか、または破棄され、抽出されたgDNAから再処理されるかのいずれかである。

【0152】

超音波破碎したgDNAを、Beckman Biomek FXPと併用するために、丸底試料プレートに移す。Biomekは、末端修復の処理、アデニンオーバーハングの付加、およびアダプターライゲーションを自動化する。Biomekシステムは、Agencourt SPRI Plate超磁石プレート、マルチチャンネル分注器およびSpan-8分注器を備えるBiomek FXP双腕システム（ポンプ制御モジュール、コンピューターおよびモニター、ペルチェコントローラー、2つの廃棄容器、および2つの水容器を備える）、およびBiomek FXP制御ソフトウェアを含む。このプロセスは、末端修復緩衝液および酵素、a-テリング緩衝液および酵素、ライゲーション緩衝液および酵素、ならびにAgencourt AMPure XPビーズを含有するSPRIworks HT断片化ライブラリーキットを利用する。それぞれの反応後、処理されたgDNAを、磁性ビーズ分離を使用して清浄する。アダプターライゲーションに続いて、Tecan M200により測定される260nmでの吸光度を使用して、処理された試料中のDMAを定量化する。900ng未満の試料は、さらに処理されないが、代わりに元の抽出された試料から再処理される。吸光度の読み取り後、試料プレートを、PCR増幅のためにBiomek FXPに戻す。第1のステップは、384-ウェルプレート上でそれぞれの試料を4つの別個の試料に分割することであり、それぞれの試料供給源に対する増幅が4重複で行われるようにする。増幅プライマーは、配列の試料供給源の特定を可能にするバーコード配列を含む。PCRは、二重384-ウェルブロック、1.5mL管ラック、24-チャンネル200µLマルチチャンネル分注器、および96-ウェルアルミニウムプレートホルダーを備えるABI GeneAmp PCRシステム9700の使用を含む。試料を、以下のプロトコルに従い自動的に熱サイクル処理する：95で5分間、98で20秒間の27サイクル、65Cで15秒間、72Cで1分間。増幅が完了した後、それぞれの試料供給源からの4つの副試料を、96-ウェルプレートの単一ウェルに再度合わせる。

【0153】

増幅ポリヌクレオチドを、磁性ビーズ分離により精製する。1.8試料量の磁性ビーズを、それぞれの試料に付加し、室温で約5分間静置する。スラリーが完全に清浄され、全てのビーズがそれぞれのウェルの側方に回収されるまで、このプレートを、磁性分離器上に約2分間置く。次に、緩衝溶液を吸引し、200µLの70%エタノールを付加する。エタノールを、吸引される前に室温で約30秒間静置する。次に、このプレートを磁石から除去し、DNAを約40µLの溶出緩衝液中に溶出する（EB；10mM トリス-HCl、pH8.5）。このプレートを磁石に戻し、ビーズがウェルの側方に回収されるまで、室温で約2分間静置する。次に、それぞれのウェルから40µLの試料を、新たな吸光度定量プレートの対応するウェルに移す。それぞれのウェル中のDNA量を、上記のように260nmでの吸光度を測定することによりチェックする。少なくとも500ng/µLの濃度を有する試料を、配列決定のためにさらに処理する。より低い濃度を持つウェルは失敗とし、対応する試料を再増幅する。

【0154】

増幅試料を96-ウェルプレートの列にわたってプールして、12試料のプールを生成し、それぞれの試料の増幅ポリヌクレオチドは、プール中の12試料の中でその試料に一意のバーコードを含む。プールに付加されるそれぞれの試料の量は、配列決定のために提出される試料中のDNAの総量が約11.25µgであるように計算される。それぞれのプールを、上記のように、38.5µL EB中の溶出液を用いて磁性ビーズ上のクリーンアップにより濃縮する。1µLのそれぞれのプールを使用して、NanoDropマシン（Thermo Scientific, Wilmington DE）上で総DNA

を定量化する。10 μ gを下回る試料は失敗とし、プーリングおよびクリーンアップを反復する。少なくとも10 μ gを有する試料を、配列決定のためにさらに処理する。

【0155】

それぞれのプール中のポリヌクレオチドが結合され、架橋増幅され、配列決定される前に、cBot試薬プレートを調製する。試薬プレートを、市販のHF緩衝液を含むPhusion High-Fidelity PCRマスターミックス(New England Biolabs)、洗浄剤フリーPhusion HF緩衝液パック(New England Biolabs)、0.1N NaOH、HT1緩衝液(5 \times SSC+0.05% Tween 20)、およびHT2緩衝液(0.3 \times SSC+0.05% Tween 20)を使用して、10枚一度に調製する。5つのNova Biostorage 8-管ストリップを、10個の別個のNova Biostorage Roboockの位置1、2、3、7、および10に配置する。1.25mLのPhusionマスターミックスを15mL管に付加し、続いて1.25mLのRNase-およびDNaseを含まない水を付加し、10秒間攪拌して、1 \times Phusionマスターミックスを生成する。440 μ Lの5 \times Phusion HF緩衝液を、「HF」と標識された別の15mL管に付加し、続いて1760 μ LのRNase-およびDNaseを含まない水を付加し、混合して1 \times HF緩衝液を生成する。試薬を以下のように試薬プレートの列に分注する：列1-720 μ L HT1緩衝液；列2-230 μ L Phusionマスターミックス；列3-200 μ L 1 \times HF緩衝液；列7-300 μ L HT2緩衝液；および列10-215 μ L 0.1N NaOH。次に、それぞれの管ストリップをNova Biostorage管キャップでカバーし、必要になるまで全てのプレートを冷凍する。

【0156】

次に、それぞれの試料プールを、配列決定のためにフローセルへの結合により調製する。結合および架橋増幅のためのシステムは、cBotシステム、NanoDrop吸光度分光計、Applied Biosystems Veriti 96-ウェル熱サイクル装置(0.2mL)、Veriti熱サイクル装置プログラム、ならびにcBot結合および架橋増幅プログラム、を備える。試料を、95℃に5分間加熱する。12.5 μ Lの4 \times ハイブリダイゼーション緩衝液(10 \times SSC+0.2% Tween-20)をそれぞれの試料に付加し、Illumina cBotマシン上に載せるまで氷上に置く。次に、シッパーコーム、フローセル、試薬プレート、および試料管をcBot上に載せる。それぞれの試料プールに対し、チャンネルの表面に結合されたオリゴヌクレオチドの伸長により、ポリヌクレオチドをフローセルのチャンネルに結合する(図1の「標的捕捉」ステップ)。結合したオリゴヌクレオチドは、選択された原因となる遺伝的変異体の上流に位置する約5000の異なる取り調べ位置の集合のメンバーに特異的にハイブリダイズする異なるオリゴヌクレオチドの集合を含む。次に、架橋増幅配列のクラスタが、標準手順を使用してcBot上で生成される。

【0157】

クラスタを、ゲノム分析器IIX(GAIIx; Illumina, San Diego, CA)を使用して配列決定する。配列決定システムは、ゲノム分析器IIX、対合末端モジュール、配列決定制御ソフトウェア、GAIIxプログラム(配列決定、前洗浄、プライム、後洗浄)、500mL容量プラスチックビーカー、大きな正方形の氷バケツ、および0.1g公差のスケールを備える。配列決定を2回行う。第1回目に、配列決定データを、バーコードの下流(伸長した鎖に沿って3')および標的ゲノム配列に隣接してハイブリダイズする第1のプライマーから生成し、それにより、原因となる遺伝的変異体を含む標的gDNA領域に対する配列決定データを生成する。第2回目に、配列決定データを、バーコード配列の上流(伸長した鎖に沿って5')でハイブリダイズする第2のプライマーから生成し、バーコード配列データがそれぞれのクラスタに対して生成されるようにする。これらの配列決定反応の順序は逆にすることができる。次に、それぞれのクラスタのバーコードを、それらの対応するgDNA配列にマッチさせて、それぞれのgDNA

A配列の試料供給源が特定できるようにする。G A I I xからの未加工データを、標準 Illumina ソフトウェアを使用して、それぞれが品質スコアを持つ個別の読み出し値に組み入れる。読み出し値を、パローズ - ホイラーアライナーを使用して参照ゲノムに整列させ、ゲノム分析ツールキット G A T K を使用して、このアラインメントから変異体を見出す。G A T K 一覧表からの出力ファイルは全て、配列決定読み出し値間の矛盾を見出し、次に、参照アセンブリを使用して、遺伝子型レポートを生成し、試料を提供した患者と相談するために、要求した医師に安全に送信される。

【 0 1 5 8 】

実施例 2：増幅および配列決定プロセス

複数の異なる標的ポリヌクレオチドの増幅のための例示のプロセスが、図 2 および 5 に示され、それらは主に図 2 の固相精製ステップの包含において異なる。図 7 も例示の増幅プロセスを示し、アダプター連結後の代わりに、主にオリゴヌクレオチドプライマー伸長がアダプター連結前に行われるという点で図 2 に示されるプロセスとは異なる。増幅は、固相精製ステップを含んでも含まなくてもよい。図 6 は、図 5 と同様に増幅プロセスを示すとともに、例示の架橋増幅および配列決定プロセスも示す。図 6 に示される増幅プロセスは、任意の架橋増幅方法および関連する配列決定方法と併用され得る。

【 0 1 5 9 】

第 1 に、部分的一本鎖アダプターが、断片化ポリヌクレオチドにライゲーションされる。部分的一本鎖アダプターは、一端に二本鎖領域（相補性配列 U にハイブリダイズされた配列 U）および使用されるハイブリダーゼーションおよび伸長条件下で標的ポリヌクレオチドにハイブリダイズしない一本鎖配列 Y を有する。ライゲーションは、配列 Y を標的ポリヌクレオチドの両方の 5' 末端に付加する。次に、それぞれが 3' 末端に異なる標的的特異的配列 W を有する複数の異なるオリゴヌクレオチドプライマーが、それらそれぞれの標的ポリヌクレオチドにハイブリダイズされ、伸長されて 3' 末端に配列 Y （Y の相補体）を持つ伸長されたオリゴヌクレオチドを生成する。伸長は、例えば、図 7 に示されるように、アダプターライゲーション前に行われ得る。オリゴヌクレオチドプライマーは、図 5 におけるように第 1 の結合パートナーを欠き得るか、または図 2 および図 7 の小さな突出する円のように第 1 の結合パートナーを含み得る。伸長されたオリゴヌクレオチドが結合パートナーを含む場合、それらは、図 2 のビーズ（より大きな円）のように、第 1 の結合パートナーに結合する第 2 の結合パートナーを含む固体表面に選択的に結合することにより精製され得る。結合および伸長されたオリゴヌクレオチドは、例えば、反応溶液を除去し、ビーズを洗浄し、新たな反応溶液を付加する間、磁場の存在下で磁気応答性ビーズ上に固定することにより精製され得る（例えば、さらなる増幅反応の成分）。次に、精製されるか、または精製されない伸長されたオリゴヌクレオチドは、増幅プライマーの対を用いて増幅する。1 つの増幅プライマーは、配列 X および配列 Y を含み、配列 Y へのハイブリダイゼーションのために 3' 末端に配列 Y を持つ。X - Y プライマーは、伸長されたオリゴヌクレオチドに沿って伸長され、配列 X、Y、W、および Z を含む複数の伸長された X - Y オリゴヌクレオチドを生成する（5' から 3' ；W は W の相補体であり、Z は Z の相補体である）。別の増幅プライマーは、配列 V および Z を含み、伸長された X - Y プライマーの配列 Z へのハイブリダイゼーションのために 3' 末端に Z を持つ。V - Z プライマーは、伸長された X - Y プライマーに沿って伸長され、V、Z、Y、および X を含む複数の配列を生成し（5' から 3' ；X は X の相補体である）、次に、増幅プロセスにおけるそれぞれの連続するプライマー伸長反応に対して、それらはさらなる X - Y プライマーの伸長のためのテンプレートとして機能し得、次に、それらはさらなる V - Z プライマーの伸長のためのテンプレートとして機能し得る。主要な増幅配列は、複数の異なる標的ポリヌクレオチドを含み、それぞれが、配列 V、Z、W、Y、および X （5' から 3'）を含む 1 つの鎖と、配列 X、Y、W、Z、および V （5' から 3'）を含む別の鎖と、を含むポリヌクレオチド中に、Z / Y の間および Z / Y の間に位置する標的ポリヌクレオチド配列とともに含有される。次に、これらの増幅ポリヌクレオチドは、配列決定に供され得る。

10

20

30

40

50

【0160】

配列決定は、図6の下半分に示されるプロセスに従い得る。第1の結合オリゴヌクレオチドは、典型的に、指数関数的増幅ステップ中に付加された配列に対する相補性により、増幅ポリヌクレオチドの付近または3 末端にある配列にハイブリダイズされる（それにより、指数関数的に増幅した生成物を、特異的に増幅する、および最終的に配列決定する）。それぞれの第1の結合オリゴヌクレオチドの伸長は、同じ配列を持つ二本鎖架橋ポリヌクレオチドのクラスタを生成するように、架橋増幅のための核形成点を提供する。第1の結合オリゴヌクレオチドの伸長生成物は、ハイブリダイズされたテンプレートを除去するように変性される。次に、伸長された第1の結合オリゴヌクレオチドは、典型的に、3 10
末端またはその付近にある配列に相補性であり、指数関数的増幅ステップ中に付加された配列から誘導される第2の結合オリゴヌクレオチドにハイブリダイズする。次に、伸長された第2の結合オリゴヌクレオチドは、さらなる第1のオリゴヌクレオチドの伸長のためのテンプレートとして機能し得、次に、それらはさらなる第2のオリゴヌクレオチドの伸長のためのテンプレートとして機能し得る。ここで、いくつかまたは全ての第1のオリゴヌクレオチドは、架橋増幅プロセスを完了した後に切断される切断部位を含む。次に、結合されたポリヌクレオチドは、加熱（例えば、約95 ）のような変性条件に供されるか、または化学的に変性され、複数の結合された架橋ポリヌクレオチドの1つの鎖を除去する。次に、残りの結合された鎖は、図6において上に「第1の読み出し値」と示される、配列決定プライマーを用いるハイブリダイゼーションのために使用される。次に、配列決定データが、ヌクレオチド伸長および検出の連続ステップにより生成され、配列決定プライマーを伸長する。次に、伸長された第1の配列決定プライマーは、第1の配列決定プライマーとは異なる第2の配列決定プライマーから配列決定プロセスを反復するために、変性され、テンプレートから除去され得る。バーコード配列を特定するのに十分な配列決定データを生成するためだけに1つの配列決定プライマーが使用される場合、その配列決定反応は、他の配列決定反応より著しく短くあり得る（例えば、約3、4、5、6、7、8、9、10、11、12、13、14、15、またはそれより多いヌクレオチド付加サイクルより短い）。図6は、単一標的ポリヌクレオチドの架橋増幅および配列決定を示すに過ぎないが、架橋増幅および配列決定は、典型的に、以前の増幅ステップにおいて増幅した複数の異なる標的ポリヌクレオチドが関与し、それらの全ては、並行して架橋増幅および配列決定される。 20 30

【0161】

実施例3：非対象配列の特定

ポリヌクレオチド（例えば、DNAおよび/またはRNA）が、当該技術分野において既知の標準方法を使用して、ウイルスおよび/または細菌ポリヌクレオチドを含有することが疑われる対象からの試料から抽出される。試料ポリヌクレオチドを、実施例1のように、断片化、末端修復、およびA-テリングする。次に、配列Dを含むアダプターオリゴヌクレオチドを、試料ポリヌクレオチドに連結し、次に、配列C、配列D、およびバーコードを含む増幅プライマーを使用して増幅する。増幅標的ポリヌクレオチドを、固体表面に結合された複数の異なる第1のオリゴヌクレオチドにハイブリダイズする。それぞれの第1のオリゴヌクレオチドは、配列Aおよび配列Bを含み、配列Bは、それぞれの異なる第1のオリゴヌクレオチドに対して異なり、それぞれの第1のオリゴヌクレオチドの3 40
末端にあり、非対象配列を含む配列、または非対象配列の200ヌクレオチド以内にある配列に相補性である。具体的に、第1のオリゴヌクレオチドは、対象のゲノムの外側に高い深さを有する配列、例えば、特定の群、順序、族、属、種、または他の分類群のウイルスまたは細菌に固有のウイルスまたは細菌配列を増幅するように選択される。増幅した配列は、16s rRNA配列を含み得る。健常な対照からのポリヌクレオチドを同時に処理する。次に、標的ポリヌクレオチドを、本発明の方法に従い架橋増幅して、配列決定する。非対象配列から生成された配列決定データを、感染因子を特定するために使用することができる。非対象配列に対して生成された配列決定データは、細菌の異なる分類群の相対レベル（例えば、1つ以上の分類群と1つ以上の他の分類群との比）、またはこれら 50

の推移を検出するために使用され得る。次に、細菌または感染因子の識別子または相対レベルは、医学的推奨を作成するか、または医療行為を行うための基礎として使用される。

【0162】

実施例4：遺伝的変異を検出するための核酸配列のアラインメント

この例示の配列操作およびアラインメント手順（「パイプライン」）は、ゲノム分析器 I I x (G A I I x) または H i S e q シーケンサー (I l l u m i n a ; S a n D i e g o , C A) からの未加工データで開始し、遺伝子型を推測し、患者試料からのメトリックスを計算する。配列決定データは、本発明の方法に従い、1フローセルレーン当たり 12 × 多重化構成のバーコード化試料の実行から生成される。シーケンサーの未加工データは、塩基呼び出し (B C L ファイル) および様々な品質制御および較正メトリックスを含む。未加工塩基呼び出しおよびメトリックスは、最初に Q S E Q ファイルに編集され、次に試料特異的 F A S T Q ファイルにフィルタリング、マージ、および逆多重化される (バーコード配列に基づいて)。F A S T Q 読み出し値は、H G 19 ゲノムに整列され、初期 B A M ファイルを作成する。この B A M ファイルは、いくつかの変換を経てアラインメントをフィルタリング、クリップ、および精練し、品質メトリックスを再較正する。最終 B A M ファイルを、既知の変異体の遺伝子型を推測し、新規の遺伝子型を発見するために使用し、呼び出しセットを生成する。次に、呼び出しセット (V C F ファイル) を、様々な呼び出しメトリックスを使用してフィルタリングし、1試料当たり高信頼度 (例えば、約 80 %、85 %、90 %、95 %、99 %、もしくはそれ以上、またはそれを超える信頼度) の変異体呼び出しの最終群を作成する。最後に、様々なメトリックスを、1試料、1レーン、および1パッチ当たりで計算し、メトリックスを、可視化、レビュー、および最終レポート生成のために実験室情報管理システムにロードする。パイプラインは、ローカルに (全体または一部が) および/または A m a z o n クラウドのようなクラウドコンピューティングを使用して実行することができる。ユーザーは、任意の好適な通信機構を使用して、パイプラインと対話することができる。例えば、対話は、D j a n g o 管理コマンド (D j a n g o S o f t w a r e F o u n d a t i o n , L a w r e n c e , K S)、パイプラインのそれぞれのステップを実行するためのシェルスクリプト、または好適なプログラミング言語 (例えば、P H P、R u b y o n R a i l s、D j a n g o、または A m a z o n E C 2 等のインターフェース) で書き込まれたアプリケーションプログラミングインターフェースを介し得る。この例示のパイプラインの操作に関する概要が、図10および11に示される。

【0163】

配列決定は、8レーンを持つフローセル上で起こる。それぞれのレーンは、12 (またはそれ以上の H i S e q を持つ) 試料を有し、それぞれが固有の 6 ~ 7 ヌクレオチドバーコード配列を持つ。それぞれのレーンは、いくつかのタイルに細分される (G A I I x の場合 120、H i S e q の場合 48)。シーケンサーは、1フローセルクラスタ当たり3つの読み出し値を出力する。読み出し値1 (R 1) は、第1のプライマーの伸長により生成された g D N A 断片の一端の配列 (59 b p) である。読み出し値2は、第3のプライマーの伸長により生成されたバーコード配列 (6 b p) である。読み出し値3 (R 2) は、第2のプライマーの伸長により生成されたプローブ配列 (15 b p) の一部である。

【0164】

パイプラインの第1のステップ (F A S T Q ファイルを作成する) において処理される未加工配列決定データは、典型的に、大きい (例えば、約 100 G B、150 G B、200 G B、250 G B、300 G B、400 G B、500 G B、1000 G B、もしくはそれ以上、またはそれを超える)。したがって、分析ステップのいくつかまたは全てにクラウドコンピューティングを利用することが有利であり得る。この例において、第1のステップは、ローカルに実行され、得られる F A S T Q ファイルは、A m a z o n S 3 (A m a z o n (S e a t t l e , W A) により提供されるオンライン記憶ウェブサービス) にアップロードされ、A m a z o n E C 2 インスタンス (A m a z o n (S e a t t l e , W A) により提供されるクラウドコンピューティングウェブサービス) を使用して処

理される。Amazonのシンプルキューサービス (SQS) を使用してタスクを割り当てる。次に、最終呼び出しおよびメトリックスがダウンロードされ、ローカルデータベースにロードされる。EC2インスタンスは、タスクおよびFASTQファイルをそれぞれSQSおよびS3からプルし、それら进行处理して、結果をS3にアップロードする。インスタンスは、手動で開始および/または終了され得るか、または部分的もしくは完全に自動化され得る。

【0165】

図10は、利用される一時およびアーカイブ記憶空間、ならびにIllumina GAIIx配列決定データ上のパイプラインに対する処理ランタイムの例を示す。BCL2FASTQステップは、全体バッチ上で実行され、後次ステップは1試料当たりで実行される。処理時間を低減するために、BCL2QSEQはローカルに実行され、その後、残りのステップは、1試料当たり1つの96 Amazon EC2インスタンス上で実行される。クラウドを使用するパイプラインを実行することは、バッチ収率に依存して7~10時間かかる。ステップ进行处理するデータのうちの1つ以上に対するクラウドコンピューティングの使用は、約10%、25%、50%、75%、90%、もしくはそれより高い割合で、試料の最終アラインメントを生成するために必要な総時間を低減し得る。

【0166】

図11は、例示の配列決定データ操作プロセスを示す。BCL2FASTQプロセスは、未加工塩基呼び出しをフィルタリング、マージ、および逆多重化された読み出し値に変換し、bcl2qseqおよびprocess_laneステップを含む。bcl2qseqに対する入力、配列決定バッチの未加工データディレクトリであり、Illuminaツールを使用してQSEQファイル(1タイルおよび読み出し番号あたり1つ)に変換され、これは全体バッチ上でローカルに実行される。QSEQファイルは、process_laneにおいて処理され、不良な読み出し値をフィルタリングし(Illumina「清純(chastity)フィルター」を使用する)、異なるタイルからの読み出し値をマージし、読み出し値2中のバーコードを使用して、読み出し値1および読み出し値3を試料特異的FASTQファイルに逆多重化する。それぞれのレーンは、並行してmnであり得る。

【0167】

FASTQ2BAMプロセスは、読み出し値をゲノムに対して整列し、そのアラインメントをさらに処理する。必要に応じて、フォーマット変更、ソート、およびインデックス付を行う。全てのステップは、個別の試料に対してファイル上で実行され、バッチに表される全ての試料は、異なるマシン上で並行して実行され得る。FASTQ2BAM中のステップは、align_bwa、fix_align、mark_duplicates、realign_bam、recalibrate_bam、およびclip_alignmentを含む。align_bwaステップにおいて、FASTQファイル中の読み出し値は、BWAアライナーを使用して参照ゲノムに対して整列される。このステップは、2回呼び出しがなされ、1回は読み出し値1を参照ゲノムに整列するため、1回は読み出し値3を、標的ポリヌクレオチドを増幅するために使用されるプローブ配列の集合に整列するためである。出力配列アラインメント/マップ(SAM)ファイルを、インデルの検出を改善するために、バイナリアラインメント/マップ(BAM)ファイルに変換した後、ソートおよびインデックス化し、デフォルトBWAパラメータを以下のように修飾する：シード長を16に減少する；アラインメント中の許容されるギャップの数を3に増加する；ギャップオープンおよび伸長ペナルティをそれぞれ6および2に減少する；許容されるギャップ伸長の数を20に増加する(より大きな変異体の場合、カスタムコンティグは、それらの存在を推測するためにゲノムに付加され得る)。次に、fix_alignmentステップは、アラインメントのいくつかを修正して、それらの精度を改善し、誤りである可能性が高いアラインメントを除去する。読み出し値3は、このシナリオにおいて読み出し値1から可変距離にプローブから誘導された配列を含むため、アライナーにより作製された統計的仮定(例えば、読み出し値間の予想距離)のいくつかに適合せず、

10

20

30

40

50

慣習的な対合末端マッピングは効率が悪い。アラインメント精度を改善するために、読み出し値 1 および読み出し値 3 は、独立して整列され（一般に、対合末端マッピングより速い）、次いで固定アラインメントステップは、結果を処理して、読み出し値 1 および読み出し値 3 が異なる鎖上にあるか、または互いから 1 0 0 0 0 塩基対より多く離れて位置付けられる（読み出し値 1 は、読み出し値 3 がマップされない場合はフィルタリングされない）任意の読み出し値を破棄し、複数の最良スコアリング位置を持つ読み出し値を関心領域のみからなるゲノムのサブセットにリマップする（ROI；例えば、原因となる遺伝的変異体、非対象配列、または AIM を含有する領域、典型的にプローブ配列付近）。図 1 2 A および 1 2 B は、示される配列パイルアップ中の人工ギャップを閉じる、それぞれ `fix-align` 前および後の非固有領域を持つ C F T R エクソン中の読み出し値の例示のアラインメントを示す。

10

【0168】

`mark_duplicates` ステップは、Picard ツールセットからのツールを使用して、PGT および光学的重複を特定およびタグ付する（BAM フォーマットで配列決定データを処理するための java ベースのコマンド・ラインユーティリティ；Geeknet (Fairfax, Va.) により維持される SourceForge ウェブサイトから入手可能）。重複を破棄しないと、参照と代替対立遺伝子との間の非均一 PCR 効率性は、対立遺伝子バイアスにつながる可能性があり、変異体の参照および代替対立遺伝子の数は、PCR によりバイアスされ得る。追加の情報なしに、同じ位置にマップする試料からの 2 つの読み出し値は重複であり、したがって 1 つを除いて全てが重複として

20

【0169】

`realign_bam` ステップは、複数の配列のスミス・ウォーターマンアラインメントをインデルの周りで行い、典型的に、真のインデルをより良く特定するか、または偽陽性 SNP の数を低減もしくは排除する効果を有する。最初のアライナー（BWA または任意の同様のツール）は、それぞれの読み出し値を独立して推測的に整列する。真の INDEL を持つ読み出し値は、そのアラインメントモデルが、使用されるパラメータの群および経験則を考慮すると、INDEL を持つものより良好にスコアするため、SNP のクラスタを持つ読み出し値として整列し得る。同じ読み出し値の複数の配列アラインメントにおいて、アライナーは、全ての読み出し値のアラインメントモデルのスコアを最適化しようとし（参照に対して、および互いに対して）、したがって、SNP の同じクラスタが全ての読み出し値を整列することができない限りは、真のアラインメントは、典型的により高いスコアとなる。リアラインメントステップは、ROI に見出される任意の INDEL の周りで複数の配列正確なリアラインメントを行う。図 1 3 A および 1 3 B は、ローカルリアラインメントの前および後の同じ読み出し値を示す。リアラインメントは、ROI 中の任意のインデルの周りで行われ得る。代替または追加として、リアラインメントは、既知のインデルの周り、例えば、インデルの 1 つ以上の参照群（例えば、Mills et al., Genome Res. (2011) June; 21(6): 830~839、Durbin et al., Nature (2010) October 28; 467(7319): 1061~1073、および Bhargale et al., Nature Genetics (2006) 38, 1457~1462 において報告される群）中のインデルの周りで行われ得る。

30

40

【0170】

`recalibrate_bam` ステップは、経験的バッチデータを使用して塩基品質を再校正し、Illumina ソフトウェアは、配列決定化学反応の単純モデルを使用して様々な品質制御メトリックスを使用し、それぞれの塩基の品質値を推定する。特定のエラーモードは考慮に入れない。このステップにおいて、高スコアリングアラインメントを

50

使用するGATKツールを使用して、経験的塩基品質を決定し、配列決定された塩基の多くの特徴（報告された品質、周囲の塩基、読み出し位置等）の間の経験的品質の共変動を分析し、共変動モデルを使用して全ての品質を再校正する。このステップは、より正確な塩基品質を提供し、より正確な呼び出し統計につながる。

【0171】

`clip` アラインメントステップは、対応するクラスタからのプローブ配列を含む整列した読み出し値から塩基を除去する。読み出し値の長さより短い対象からの増幅配列の場合、読み出し値1は、典型的に、対応するプローブからの配列を含有する。プローブ配列が参照ゲノムから誘導されるとき、これらの読み出し値は、ゲノムに対して整列するが、任意の真のSNPをマスクするため、参照配列に向かって対立遺伝子バイアスを導入する。このステップは、読み出し値がそれ自体のプローブと重なるときに特定され、その重なる塩基を読み出し値から選択的に除去する。読み出し値1および読み出し値3の両方がマップし、読み出し値1がプローブと重なる全ての読み出し値の場合、読み出し値1の重なる塩基は、アラインメントから除去される（「クリップされる」）。読み出し値の読み出し値3はマップせず、次に読み出し値1は、それが任意のプローブと重なる場合にクリップされるが、その重なりが読み出し値のいずれかの末端で起こる場合に限られる（末端塩基を含まないプローブ配列は、読み出し値がそのプローブから生成されなかったことを示す）。いずれかの場合、クリッピングは、クリップされた塩基に「S」操作を含めるように、CIGARアラインメント文字列を修飾することにより行われ、開始位置を更新し、クリップされた塩基の塩基品質を0に設定する。したがって、配列は依然として存在するが、アラインメントは、クリップされた塩基を除外するように修飾される。

【0172】

BAM2VCFプロセスは、最終アラインメント（FASTQ2BAMプロセスからのBAMファイル）を使用して、配列決定データおよび以前の知識を考慮して変異体の確率を計算するベジアン方法を使用し、遺伝子型を決定する。全てのステップは、個別の試料のファイル上で実行され、並行して異なるマシン上で実行され得る。変異体特定のための戦略（「呼び出し」）は、感度を最大化し、次いで呼び出しメトリックスおよび他の基準に基づいてそれをフィルタリングするように非常に寛大な閾値を使用して、特定された変異体の初期セット（「呼び出しセット」）を作成することである。BAM2VCFプロセスにおけるステップは、`genome_whitelist`、`genotype_novel`、および`hard_filter_vcf`ステップを含む。`genome_whitelist`ステップは、既知の変異体位置およびそれらの位置での対立遺伝子の所与の参照一覧表との比較に基づいて、入力BAMファイルの遺伝子型を推測する。このステップにおいて、変異体を特定するようにプログラム化されたコンピューターアルゴリズム（「呼び出し元」）は、全ての変異体を出力し、任意の信頼度に基づくフィルタリングをスキップするように構成される。このステップの出力は、変異体呼び出しフォーマット（VCF）ファイルであり、追加のステップにおいてさらに処理される。

【0173】

遺伝子型新規ステップは、参照ゲノムとは異なるROI内の変異体部位を特定し、それらの部位での遺伝子型を推測する。このステップにおいて、呼び出し元は、以前のステップにおいて使用される既知の変異体の参照一覧表に含まれない遺伝子型のみを出力し、任意の信頼度に基づくフィルタリングをスキップするように構成される。`genotype_novel`ステップからの呼び出しは、多くの偽陽性を含み得る。`hard_filter_vcf`ステップは、いくつかの呼び出しメトリックスを使用して、遺伝子型決定をフィルタリングする。これらのメトリックスは、2つの広いカテゴリーに分類される：（1）塩基呼び出し、マッピング、変異体、または遺伝子型決定の信頼度を定量化するもの、および（2）鎖バイアス、位置バイアス、またはIllumina化学誤差を引き起こすことが知られているホモポリマー実行等の配列特徴の存在等の共通の誤差原因の可能性を定量化するもの。修飾された閾値は、ゲノム分析ツールキット（GATK）による推奨に基づき得る。代替または追加として、所望の感度および特異性のための閾値を特定する

ように、マシン学習アプローチが使用されてもよい。

【0174】

この実施例におけるプロセスの評価を支援するために、それぞれの試料の記録は、対応するバーコードを持つ読み出し値、ゲノムにマップされる読み出し値、`fix_align` ステップ後の読み出し値、PCR / 光学的重複を除外した後の読み出し値、読み出し値1 および読み出し値3が10000bp超離れてマップする読み出し値、`fix_align` ステップがリマップを試行する非固有領域における読み出し値、成功裏にリマップされた読み出し値、クリップされ、対応する読み出し値3を有する読み出し値、クリップされ、対応する読み出し値3を有しない読み出し値、XおよびY染色体における読み出し値(性別を推測するために使用され得る)、および参照ゲノム中のSNP同一性にマッチするSNP呼び出しの数からなる。これらのメトリックスのいずれか1つ、または組み合わせに関する閾値は、その閾値を下回る任意の試料の結果が破棄されるように設定され得る。これらのメトリックスのうちのいずれか1つ、または組み合わせは、1レーン当たりの読み出し値の総数および初期フィルターを通過する読み出し値の数に加えて、全体配列決定レーンに対して集約され得る。配列決定レーンメトリックスのいずれか1つ、または組み合わせに関する閾値が、閾値(複数可)を通過するのに失敗したレーンから生じるデータの除外のために設定され得る。任意の2つの呼び出しセット間の一致および不一致は、検証研究のため、または今後の試料分析のための閾値を設定するために分析され得る。

10

【0175】

特定された任意のゲノム変異体に対して、任意の変異体呼び出しの基礎となる全ての読み出し値のアラインメントを示す、パイルアップ画像が生成され得る。パイルアップ画像は、`Integrative Genomics Viewer (IGV; Broad Institute, Cambridge, MA)`により提供される)等のゲノミクスデータの可視化装置を使用して生成され得る。これを行うために、(1)ゲノムおよびBAMファイルをロードし、(2)それぞれの変異体位置を通じて繰り返し、パイルアップのスナップショットPNGを出力する、IGVスクリプトが生成される。IGVは、仮想フレームバッファ(例えば、`xvfb`)の下で実行され、得られるPNGファイルは、IGVクロームを除去するように(コマンドライン`Imagick`ツールを使用して)クロップされる。

20

【0176】

実施例5：プローブ配列を選択する

30

増幅および配列決定のための標的配列の初期捕捉のための最適なプローブ配列を選択するプロセス(「プローブ設計」とも称されるプロセス)において、アルゴリズムが用いられる。次に、プローブ配列は、オリゴヌクレオチドプライマーまたは固体支持体に結合された第1のオリゴヌクレオチドの集合の生成において使用され得る。プローブ設計プロセスは、変異体および配列決定される対応する標的配列の一覧表に追加を組み込むように反復され得る。したがって、アルゴリズムは、以前に設計されたROIにより既にカバーされている領域が再設計されないように、以前に設計された関心領域およびプローブの付加を可能にする。

【0177】

プローブ設計の最初の単位は、関心領域(ROI)であり、遺伝子のエクソン、単一ゲノム塩基、非コーディングである領域もしくは点、または重なる可能性があり得るこれらの組み合わせの一覧表であり得る。このプロセスにおける第1のステップは、プローブが設計される異なる領域の全てをロードした後に調整することである。「設計エンジン」クラスは、考慮される全てのROI、および後にそれぞれのROIに対して設計されたプローブの全てを追跡する。点突然変異として最初に入力された変異体等の小さなROIは、さらに処理される前に100bpの長さにパッドされる。次に、全ての重なるROIは、重複プローブが設計されないように、単一ROIに組み入れられる。

40

【0178】

2つのROIは、それらが合わされたROIをカバーするROIタイルの数を低減する

50

場合にのみマージされる。2つの別個の領域をカバーするタイルの数は、仮定的に連結されたROIをカバーするタイルの数と一緒に計算される。最も少ないタイルを必要とする場合は、プローブ設計の後次ステップに使用される。アルゴリズムは、所与のゲノム領域をカバーするタイルの数を決定するために使用される。

【0179】

ROIがパッドされ、マージされると、全てのROIは少なくとも100bp長であり、どれも重ならない。得られるROIは長い（例えば、特定されたタイル長より長い）または短い（例えば、特定されたタイル長より短い、またはそれに等しい）。長いROIは、ROIタイルに細分され、それらはプローブが設計される配列の単位長である。タイル長より短い、またはそれに等しい短いROIは細分されない。それぞれの可能性のあるROIタイルは、それから設計されたプローブがどれ程良好に動作するかについて評価される。タイルの最大数は、この計算の上限としても計算される。可能性のあるタイルの最小値と最大値との間の全てのタイル数は、タイルの最小数から最大数の順に考慮される。これらのタイル数は、どの数が計算されるかに応じて、分割されるROIにおける塩基の数の上限をROIの最小または最大長さで割ったものに等しい。これらの数は、 $ROI_TILE(250 \text{ 塩基})$ および MAX_ROI_TILE であり、 $MAX_ROI_TILE = (TILE_SIZE) - (READ_LENGTH) - (RECESS)$ である。 $TILE_SIZE$ は、300~440塩基長である。 $READ_LENGTH$ は、40~60塩基長であり、 $RECESS$ は、10塩基長に設定される。プローブ設計が全ての有効なプローブを生じる多数のタイルが見出されると、繰り返しは終了する。これは事実上、ある領域をカバーするために必要とされるタイルの数を最小化すると同時に、最良のプローブが以下の基準に従い選択されることを保証する。

【0180】

プローブ設計アルゴリズムは、他のROIタイルからの単離において、所与のROIタイル上で作用するため、ROIタイルは、このプローブ設計プロセスの基本単位として考慮することができる。それぞれのROIタイルは、そのために設計された前方向および逆方向タイルを有し、全ての塩基が配列決定時にいずれかの鎖から証明され得るようにする。プローブ設計アルゴリズムは、それぞれのROIタイルに対して計算される「プローブ設計ウィンドウ」内の全ての $READ_LENGTH$ タイルに対する前方向および逆方向プライマーを考慮することにより機能する。次に、このウィンドウ内のそれぞれのプローブは、以下に記載される基準に基づいてスコア付され、最も重要なスコアから最も重要でないスコアの群を作成し、全てのスコアに対して低いスコアほど良い。したがって、最良のプローブは、単にプローブスコア群の複数フィールド昇順ソートにおいて最初に出現するものである。それぞれのROIタイル分配は、プローブ設計アルゴリズムを可能性のあるROIタイルのそれぞれに対して実行させる。繰り返しは、最も少ないROIタイルの条件で開始し、そのような分配が有効なプローブを生じない場合（その条件は以下に記載される）、ROIタイルの数が増加し、分配が再度行われる。

【0181】

プローブウィンドウは、以下のように定義される：（1）プローブウィンドウの長さは、 $(TILE_SIZE) - (\text{現在のROIタイルの長さ}) - (RECESS)$ として定義される；（2）次に、開始座標は、ROI自体の末端から $RECESS$ bp離れたものとして定義され、停止座標は、上記のプローブウィンドウの長さを $RECESS$ 座標に加算することにより計算される；および（3）次に、この範囲内の全ての40merは、評価のためのプローブとして考慮される。図19は、このステップにおいて考慮される配列領域の位置関係の図を提供する。

【0182】

それぞれのプローブを評価するために使用される基準は、考慮される順に、「ニア24-mer」の固有性、ニア24-mer中の任意の共通SNPとの重なり、全体40merのマッピング可能性、NtBspQIの重なり、反復マスキング、全体40mer中の任意の共通SNPとの重なり、疾患変異体とのニア24merの重なり、疾患変異体との40

merの重なり、GC%、およびROIまでの距離を含む。理想的なプローブは、ゲノム全体で固有であるが、時として任意の部位の付近にそのような位置を見出すことは不可能である。相殺するために、「ニア24-mer」（プローブ配列またはその相補体を含むオリゴヌクレオチドの最も3側の24塩基として定義される）は、可能な限り固有であるように選択される。捕捉されたゲノム種の伸長は、5から3で起こるため、二本鎖と一本鎖の連結の最も近くにある塩基対結合部位の品質は、捕捉の効率性に大きな影響を及ぼし、より強い結合は、捕捉された配列が伸長される可能性を高める。結合品質を測定するために、カリフォルニア大学サンタクルーズ24merマップ可能性追跡（UCSCゲノムブラウザを通じて入手可能）が使用され、これがゲノム中のそれぞれの塩基に対して、その塩基で開始する24merのマッピングスコアを付与する。このスコアは、1/Nとして付与され、Nは、ゲノム中の24merに対するマッチの数である。この試験から2つの結果のみが考慮される：スコアが1に等しいかどうか（すなわち、固有である）、または1未満であるかどうか（すなわち、複数の結合部位を有する）。最初の場合が好ましい。

【0183】

ニア24-mer中の共通のSNPとの重なりは望ましくない。捕捉プローブ結合部位におけるいかなるミスマッチも、結合効率を低減する。近くのSNPは、多くの場合、連鎖不均衡であるため、この結合効率の差は、大量の対立遺伝子バイアスを導入する。したがって、見出される任意の対立遺伝子が、その遺伝子中の他の対立遺伝子と同じプローブ結合部位を有することができる最大の機会を有することが望ましい。UCSC共通SNP追跡が、この計算を行うために使用される。重なりがないもの、および1つ以上の重なりがあるものという2つのカテゴリがある。前者が非常に好ましい。

【0184】

全体40merのマッピング可能性は、24merのマッピング可能性と同じカテゴリを用いるが、24mer追跡の代わりにUCSC40merマッピング可能性追跡を使用して、同じスコアを決定する。この新たな追跡は、40merマッピング可能性のみが24merのマッピング可能性の代わりに考慮される。

【0185】

酵素NtBspQIは、オリゴヌクレオチド合成または操作において有用であり得る。したがって、プローブと酵素の認識部位との間の重なりがスコア付される。

【0186】

反復マスキングを評価する際に、UCSC反復マスク追跡（注釈反復）を使用して、それぞれ40merを構成する塩基に対する値の加重平均を計算する。反復マスク追跡は、それがマスクされるか否かに依存して、値0または1をそれぞれの塩基に割り当てる。したがって、スコアが高いほど、それはより多くマスクされる。マスクされた塩基に対するプローブを設計することは望ましくないため、より低い、理想的には0スコアがより良い。これらのスコアは、四分位に分割されるため、マスクされる最大25%はある群を含み（0としてスコア付される）、最大50%は別の群を含み、75%および100%も同様である。

【0187】

プローブのニア24merの任意の共通SNPとの重なりが評価されるように、全体40merプローブ配列の任意の共通SNPとの重なりも評価およびスコア付される。

【0188】

ニア24merが疾患変異体と重なるか否かも評価される。この試験は、ニア24merが配列決定される任意の原因となる遺伝的変異体との重なりについて評価されることを除いて、共通SNP重なり試験と同様である。スコア0は、重なりがない場合に付与され、スコア1は、重なりが存在する場合に付与される。次に、同様の分析は、全体40merに対して行われる。

【0189】

プローブ配列のGC%は、20~80%のGC%には0、この範囲外には1の2つのス

10

20

30

40

50

コアのうちの1つを受ける。最後に、ROIまでの距離が評価される。他の全てが等しい場合、ROIに最も近いプローブが好ましい。ROI距離に対するスコアは、プローブの末端とそれが標的とするROIの開始との間の塩基対の数に等しい。

【0190】

有効なプローブは、以下の基準の全てを満たし、ROIのプローブ設計のさらなる繰り返しが必要とされないプローブである：(1) 1/3.5以上のニア24merのマップ可能性、(2) 1/3.5以上の全体40merのマップ可能性、(3) 0.25以下の反復分率、(4) ニア24mer中に共通SNPとの重なりがない、プローブ配列中のNtBspQI認識部位がない。

【0191】

実施例6：試料収集および分析

ユーザーが希な遺伝的疾患の保因者である確率を配信する例示のプロセスが、図14~17に示される。図14~15は、それぞれウェブおよび医療顧客の注文履行のためのパイプラインを示す。注文は、医師または顧客により発注され得る。注文は、単一検査または夫婦もしくは家族のためになされ得る。この注文は、ウェブサイトを通じて受け入れられ得る。注文システムは、連絡先情報、人口統計学的詳細、および請求情報を受け入れることができる。連絡先情報は、限定されないが、氏名、住所、電話番号、およびEメールアドレスが挙げられ得る。人口統計学的情報としては、限定されないが、性別、生年月日、および自己報告された民族性が挙げられ得る。注文確認通知は、提供された連絡先情報を使用して送信され得る。受け入れ可能な注文は、データベースに追加され、これらの注文の状態は、状態マシンにより後次に維持され得る。

【0192】

次に、試料採取キットがユーザーに送信される。任意のヒト組織または液体である試料が採取される。試料は、ヒトから単離されたDNAでもあり得る。この実施例に有用な試料の例としては、唾液、血液、尿、口腔細胞、羊水、細胞剥離物、および細胞培養物が挙げられるが、これらに限定されない。次に、試料は、本明細書に記載されるデバイスを使用して遺伝子型決定する。表現型勧誘、例えば、ユーザーの表現型形質の自己特定を引き出すことは、試料処理と並行して行われ得る。

【0193】

試料採取は、自宅、診療所、または特別な採集場所で行われ得る。試料採取および返却は、注文追跡状態マシンの状態を進行させることにより追跡され得る。受け入れ施設により受領された試料は、状態マシンにおけるそれらの状態を進行させることによりデータベースシステムに登録され得る。受け入れ施設での受け入れ後、試料は遺伝子型決定施設に配送され得る。遺伝子型決定施設は、保護ファイル転送プロトコルにより保護データストレージに未加工ゲノムデータを戻すことができる。ファイルアップロードは、状態マシンの進行をトリガすることができる。この進行は、データストレージサーバーから未加工ゲノムデータとともに、遺伝子型呼び出しを行うように構成されたサーバーをトリガして、その注文と関連付けられた任意の表現型データを検索することができる。遺伝子型決定アルゴリズムは、完全に確率的な遺伝子型呼び出しを生成することができる。

【0194】

図16~17は、それぞれ高レベルの試料処理パイプラインおよび詳細な計算パイプラインを示す。試料のバッチは、品質制御の目的で受領され、測定される(バッチはQCを通過する)。家族歴、性別、または自己報告された祖先等の情報は、それぞれの試料に対する呼び出しの独立した検査として機能するように使用される(バッチ試料に対して検索された表現型データ)。このプロセスと並行して、この予測を含むレポートが定期的に更新される。第1の検査前危険性計算が、表現型に基づいて配送される(例えば、家族歴およびオンライン質問票に対する他の回答)。遺伝子型試料が受領および処理されると、検査後計算が付与される。次に、レポートが生成され、図16に示されるように、研究室スタッフおよび医師の承認を得るためにパイプラインの最終段階に送信される。

【0195】

品質制御メトリックスが、呼び出しプロセスから生成され得る。例示の品質制御メトリックは、少なくとも1つの遺伝子型が閾値より高い事後確率を有する確率的遺伝子型呼び出しのパーセンテージである。試料のバッチが一緒に処理される。バッチとして処理されると、個別の確率的遺伝子型呼び出しを使用して、バッチレベル品質制御統計を生成することができる。確率的遺伝子型呼び出しが、データベースに記憶され得る。良好な遺伝子型呼び出しは、注文状態の進行をトリガすることができる。夫婦または家族からの注文の場合、状態マシンは、注文全体の完了のために保持することができ、他の単一注文は次の状態に移行され得る。表現型データが、危険性計算に必要とされる場合、状態マシンは、全ての表現型データが収集されるまで処理を遅延させることができる。状態マシンは、表現型が必要とされるという患者への通知をトリガすることもできる。全ての遺伝子型および表現型データの用意ができた場合、状態マシンは進行し、危険性計算サーバーをトリガして危険性計算を行うことができる。危険性計算の結果は、連載され、結果報告システムに転送され得る。これは、結果のマシン可読フォーマットである。状態マシンは、転送が完了したときに注文を進めることができる。結果報告サーバーは、確率的危険性計算と適切なテキストおよびフォーマット設定を組み合わせ、ヒト可読レポートを生成することができる。このヒト可読レポートは、ウェブサイト上で表示するためにさらにフォーマット設定され得る。このヒト可読レポートは、印刷するためにPDFファイル等の他の媒体に対してフォーマット設定され得る。最終結果報告は、自己検証システムを使用して自動的に発表され得る。ヒトは、発表するために報告をレビューすることができる。確認者は、臨床検査科学者および医師であり得る。結果は、結果のビューおよび品質制御メトリックスの要約にリンクする、ウェブポータルを介してアクセスされる。臨床検査科学者によりレポートが受け入れられると、結果を医師に発表する。医師は、同様のポータルで結果をレビューし、結果の最終発表を承認することができる。

10

20

【0196】

図18は、2人の仮想の親：母親Henおよび父親Henに対するレポート生成のための例示の入出力ステップを示す。父母の遺伝子型、父母の表現型、ならびに相対遺伝子型および表現型を組み込む子の予測が生成される。これらの変数のいずれか、または全ては、欠損値であり得、人口統計学的に類似する個体から（およびこれが既知でない場合は世界人口から）初期化されたデフォルトを含む。得られる子の予測は、疾患または形質の危険性だけでなく、身長および体重等の他の変数も含み得る。子の予測における異なる変数は、異なる重量の遺伝子型および表現型を使用する。

30

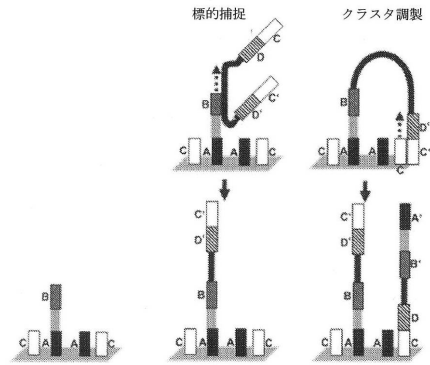
【0197】

本発明の好ましい実施形態が本明細書に図示および説明されたが、そのような実施形態が単なる例として提供されることは、当業者に明らかとなるであろう。ここで当業者は、多くの変動、変化、および置換は、本発明から逸脱することなく思い付くであろう。本明細書に記載される本発明の実施形態に対する様々な代替が、本発明を実施する際に用いられ得ることを理解されたい。以下の特許請求の範囲は、本発明の範囲を定義すること、それらの特許請求の範囲内の方法および構造、ならびにそれらの相当物がそれにより網羅されることが意図される。

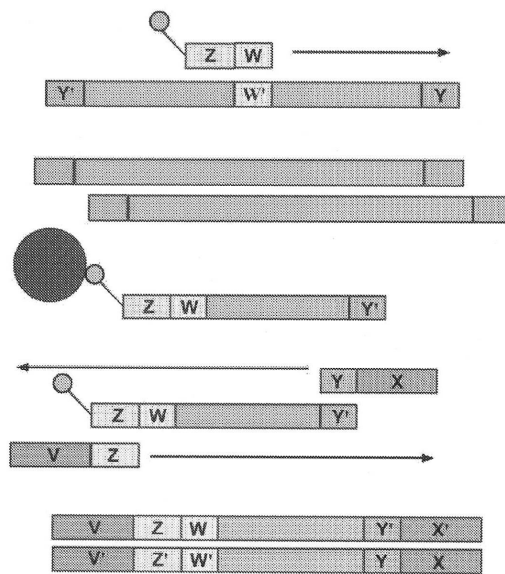
【図 1】

ハイブリダイゼーション、伸長、および変性

固定化 DNA



【図 2】



【図 3 - 1】

疾患	遺伝子	変異体名
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	F306+1
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	F306+1
21-ヒドロキシラ ーゼ欠乏症	CYP21A3	g.655A/C>G
21-ヒドロキシラ ーゼ欠乏症	CYP21A4	g.655A/C>G
21-ヒドロキシラ ーゼ欠乏症	CYP21A6	G110del8nt
21-ヒドロキシラ ーゼ欠乏症	CYP21A5	G110del8nt
21-ヒドロキシラ ーゼ欠乏症	CYP21A7	1172N, rs34607927
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	I236N
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	M239K, rs6476
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	P30L
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	P453S
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	Q318X
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	R356W
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	V237E, rs12530380
21-ヒドロキシラ ーゼ欠乏症	CYP21A2	V281L, rs6471
ABCC8-関連イン スリン過剰症	ABCC8	3992-9G>A
ABCC8-関連イン スリン過剰症	ABCC8	delF1388
ABCC8-関連イン スリン過剰症	ABCC8	delF1388
軟骨形成不全	ABCC8	V187D
軟骨形成不全	FGFR3	G375C
色覚異常	FGFR3	G380R, rs28931614
色覚異常	CNGB3	c.1148delC
色覚異常	CNGB3	c.819-826del8
色覚異常	CNGB3	c.819-826del8
色覚異常	CNGB3	c.886-896del11insT
色覚異常	CNGB3	c.886-896del11insT
色覚異常	CNGB3	c.991-3T>G
色覚異常	CNGB3	p.Arg403Gln
色覚異常	CNGB3	p.Glu336X

【図 3 - 2】

疾患	遺伝子	変異体名
アデノシンーリン 酸デアミナーゼ 1	AMPD1	P48L
アデノシンーリン 酸デアミナーゼ 1	AMPD1	Q12X, rs17602729
神経細胞傷害を伴う 脳梁欠損症	SLC12A6	c.2436delG
神経細胞傷害を伴う 脳梁欠損症	SLC12A6	c.2436delG
アルカプトン	HGD	c.174delA
アルカプトン	HGD	c.174delA
アルカプトン	HGD	c.457_458insG
アルカプトン	HGD	c.457_458insG
アルカプトン	HGD	G161R
アルカプトン	HGD	G270R
アルカプトン	HGD	IVS1-1G>A
疾患	遺伝子	変異体名
アルカプトン	HGD	IVS5+1G>A
アルカプトン	HGD	Met368Val
アルカプトン	HGD	P230S
アルカプトン	HGD	S47L
アルカプトン	HGD	V300G
α-1-アンチトリプシ ン欠乏症	SERPINA1	Arg101His, rs709932
A-1-アンチトリプシ ン欠乏症	SERPINA1	Glu264Val
A-1-アンチトリプシ ン欠乏症	SERPINA1	Glu342Lys, rs28929474
A-1-アンチトリプシ ン欠乏症	SERPINA1	Glu376Asp, rs1303
α-マンノシドーシス	MAN2B1	IVS14+1G>C
α-マンノシドーシス	MAN2B1	p.L809P
α-マンノシドーシス	MAN2B1	p.R750W
α-サルコグリカン異 常症	SGCA	R77C, rs28933693
α-サラセミア	HBA2	H19D
α-サラセミア	HBA1	HbQ
α-サラセミア	HBA1, HBA2	3.7 kb (I 型) 欠失 α-2
α-サラセミア	HBA1, HBA2	3.7 kb (I 型) 欠失 α-2
α-サラセミア	HBA1, HBA2	HbVar データベース id # 1086, -- (SEA); α-グロビ ン遺伝子 α-Thal-1 の両方を含 む約 20 kb の欠失

【 図 3 - 4 】

疾患	遺伝子	変異体名
		HbVar データベース id# 1086,- -(SEA); α -グロビン遺伝子 α - Thal-1 の両方を含む約 20 kb の 欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1087,- -(MED-I); α -グロビン遺伝子 α - Thal-1 の両方を含む約 17.5 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1087,- -(MED-I); α -グロビン遺伝子 α - Thal-1 の両方を含む約 17.5 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1088,- -(α)20.5; この 20.5 kb 欠失は、 α 2 および α 1 の 5'末端; α - Thal-1 を含む
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1088,- -(α)20.5; この 20.5 kb 欠失は、 α 2 および α 1 の 5'末端; α - Thal-1 を含む
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1094,- -(FIL); α 1, α 2, および 遺伝子 α -Thal-1 を含む 30-34 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1094,- -(FIL); α 1, α 2, および 遺伝子 α -Thal-1 を含む 30-34 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1095,- -(THAI); α 1, α 2, および 遺伝子 α -Thal-1 を含む 30-34 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1095,- -(THAI); α 1, α 2, および 遺伝子 α -Thal-1 を含む 30-34 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1097,- -(MED-II); 2つの α および 遺伝子 α -Thal-1 を含む 26.5 kb の欠失
α -サラセミア	HBA1,HBA2	HbVar データベース id# 1097,- -(MED-II); 2つの α および 遺伝子 α -Thal-1 を含む 26.5 kb の欠失
α -サラセミア α -サラセミア	HBA1,HBA2 HBA2	HbVar データベース id# 187

疾患	遺伝子	変異体名
		HbVar データベース id # 2598, IVS 1-5 (G>A)
α-サラセミア	HBA2	HbVar データベース id # 703
α-サラセミア	HBA1,HBA2	HbVar データベース id # 704, HbVar データベース id # 705,
α-サラセミア	HBA1,HBA2	Koya Dura HbVar データベース id # 707, rs41412046
α-サラセミア	HBA1,HBA2	HbVar データベース id # 87, HbVar データベース id # 969, Poly A . (A>G); AATAAA->AATGAA β +
α-サラセミア	HBA1,HBA2	HbVar データベース id # 971, Poly A . (A>G); AATAAAA->AATAAG β +
α-サラセミア	HBA1,HBA2	MIT
α-サラセミア	HBA2	W14X
アンジオテンシン II 受容体 I 型	AGTR1	A1166C
アポリポタンパク質 E 遺伝子型決定	APOE	p.C112R, rs429358
アポリポタンパク質 E 遺伝子型決定	APOE	p.R158C, rs7412
アルギニノコハク酸尿症	ASL	R385C
ARSACS	SACS	5254C>T
ARSACS	SACS	6594delT
ARSACS	SACS	6594delT
アスパルチルグリコシミン尿症	AGA	c.199_200delGA
アスパルチルグリコシミン尿症	AGA	c.199_200delGA
アスパルチルグリコシミン尿症	AGA	C163S
ヒタミン E 欠乏症を伴う運動失調	TTPA	744delA
ヒタミン E 欠乏症を伴う運動失調	TTPA	744delA
血管拡張性心臓病	ATM	R35X
多臓器自己免疫症候群 I 型	AIRE	c.1163_1164insA
多臓器自己免疫症候群 I 型	AIRE	c.1163_1164insA
多臓器自己免疫症候群 I 型	AIRE	c.769C>T

【 図 3 - 5 】

疾患 多属性自己免疫症候群 1型	遺伝子	変異体名
多属性自己免疫症候群 1型	AIRE	c.967_979del
多属性自己免疫症候群 1型	AIRE	c.967_979del
パーデット・ビー ドル症候群	BBS1	M390R
パーデット・ビー ドル症候群	BBS10	p.C91LfsX4
パーデット・ビー ドル症候群	BBS10	p.C91LfsX4
ペスト卵黄様黄 斑変作症	BEST1	c.G383C
β-サルコグリ カン異常症	SGCB	S114F
R-サラセミア	HB	-28 (A->G) β +
β-サラセミア	HB	-29 (A->G) β +
R-サラセミア	HB	-29A>G
β-サラセミア	HB	-30 (T->A) β +
R-サラセミア	HB	-37 (C->G) β +
β-サラセミア	HB	-48C>T
β-サラセミア	HB	CAPl+1 (A->C) β + コドン 15 (G->A); TGG(Trp)- >TAG(停止コドン) β 0 , rs347116011
β-サラセミア	HB	コドン 15 (G->A); TGG(Trp)- >TAG(停止コドン) β 0 , rs347116011
β-サラセミア	HB	コドン 16 (-C); GGC(Gly)- >GG-β 0 コドン 16 (-C); GGC(Gly)- >GG-β 0 コドン 17 (A->T); AAG(Lys)->TAG(停止 コドン) β 0 コドン 24 (T->A); GGT(Gly)- >GAG(Gly) β + コドン 39 (C->T); CAG(Gln)- >TAG(停止コドン) β 0 コドン 5 (-CT); CCT(Pro)- >C-β 0 コドン 5 (-CT); CCT(Pro)- >C-β 0 コドン 6 (-A); GAG(Glu)->G- G β 0
β-サラセミア	HB	コドン 5 (-CT); CCT(Pro)- >C-β 0 コドン 6 (-A); GAG(Glu)->G- G β 0
β-サラセミア	HB	コドン 6 (-A); GAG(Glu)->G- G β 0

【 図 3 - 6 】

病態	遺伝子	変異体名
β -サラセミア	HBB	コドン 6 (-A); GAG(Glu)->G. G β 0
β -サラセミア	HBB	コドン 8 (-AA); AAG(Lys)-> -G β 0
β -サラセミア	HBB	コドン 8 (-AA); AAG(Lys)-> -G β 0 コドン 41/42 (-TTCT); TCTT(TT(Phe-Pho))->----TT β 0
β -サラセミア	HBB	コドン 41/42 (-TTCT); TCTT(TT(Phe-Pho))->----TT β 0 コドン 71/72 (+A); TTT AGT(Phe Ser)->TTT A AGT; β 0
β -サラセミア	HBB	コドン 71/72 (+A); TTT AGT(Phe Ser)->TTT A AGT; β 0 コドン 8/9 (+G); AAG TCT(Lys,Ser)->AAG G TCT β 0
β -サラセミア	HBB	コドン 8/9 (+G); AAG TCT(Lys,Ser)->AAG G TCT β 0 HbVar データベース id # IVS-II-654 (C->T); AAGCAATA- -AAG*GTAAATA 0 (複製)
β -サラセミア	HBB	HbVar データベース id # IVS-II-705 (T->G); GATGTAAAGA- -AAG*GTAAAGA β + HbVar データベース id # IVS-II-745 (C->G); CAGTACCAT- -CAG*GTACCAT β + HbVar データベース id # 979, bp 欠失 0
β -サラセミア	HBB	619 bp 欠失 β 0 IVS-I-1 (G->A); AG*GTTGGT- -AGATTGGT β 0
β -サラセミア	HBB	IVS-I-1 (G->T); AG*GTTGGT->AGTTTGGT β 0 IVS-I-110 (G->A) β + 複製変異 は、変容体スプライス部位 に対して 5' の 21 ヌクレオ チド GC
β -サラセミア	HBB	

【図 3 - 7】

疾患	遺伝子	変異体名
β -サラセミア	HBB	IVS-J-5 (G>C) β^+ (重症)
β -サラセミア	HBB	IVS-II-1 (G>A); β^0
β -サラセミア	HBB	IVS-II-844 (C>G); β^+
β -サラセミア	HBB	IVS11+6T>C
β -サラセミア	HBB	IVS11-849A>C
β -サラセミア	HBB	IVS11-849A>G
ピオチニダーゼ欠乏症	BTB	A171T, rs13073139
ピオチニダーゼ欠乏症	BTB	D252G, rs28934601
ピオチニダーゼ欠乏症	BTB	D444H, rs13078881
ピオチニダーゼ欠乏症	BTB	P403V
ピオチニダーゼ欠乏症	BTB	G98:d7i3
ピオチニダーゼ欠乏症	BTB	G98:d7i3
ピオチニダーゼ欠乏症	BTB	Q456H
ピオチニダーゼ欠乏症	BTB	R157H
ピオチニダーゼ欠乏症	BTB	R538C
ピオチニダーゼ欠乏症	BTB	E383K
ブラウ症候群	NOD2	L469F
ブラウ症候群	NOD2	R334Q
ブラウ症候群	NOD2	R334W
ブルーム症候群	BLM	2407insT
ブルーム症候群	BLM	2407insT
ブルーム症候群	BLM	736delATCTGAinsTAGATT
ブルーム症候群	BLM	C (2281)del6/ins7)
ブルーム症候群	BLM	736delATCTGAinsTAGATT
ブルーム症候群	BLM	C (2281)del6/ins7)
BRCA1 遺伝性乳癌/卵巣癌	BRCA1	185delAG
BRCA1 遺伝性乳癌/卵巣癌	BRCA1	185delAG
BRCA1 遺伝性乳癌/卵巣癌	BRCA1	5382insC
BRCA1 遺伝性乳癌/卵巣癌	BRCA1	5382insC
BRCA1 遺伝性乳癌/卵巣癌	BRCA1	Tyr978X
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	6174delT
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	6174delT
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	8765delAG
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	8765delAG
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	A305E (914C>A), rs28940574
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	E285A (R54A>C), rs28940279
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	IVS2-2A>G (433-2A>G)
BRCA2 遺伝性乳癌/卵巣癌	BRCA2	Y231X (693C>A)
カナバン病	ASPA	
カナバン病	ASPA	
カナバン病	ASPA	
カナバン病	ASPA	

【図 3 - 8】

疾患	遺伝子	変異体名
カルニチンパルミトイルトランスフェラーゼ IA 欠乏症	CPT1A	G710E
カルニチンパルミトイルトランスフェラーゼ IA 欠乏症	CPT1A	P479L
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	G549D
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	L178F 534 ins/25 bp del
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	L178F 534 ins/25 bp del
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	P227L
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	P50H
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	P604S
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	Q413fs
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	Q413fs
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	Q550R
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	R124X
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	R503C
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	R631C
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	S113L
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	s38fs

【図 3 - 9】

疾患	遺伝子	変異体名
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	s38fs
カルニチンパルミトイルトランスフェラーゼ II 欠乏症	CPT2	Y628S, rs28936673
軟骨毛髪形成不全症	RMRP	g.262G>T
軟骨毛髪形成不全症	RMPR	g.70A>G
CFTR-関連障害	CFTR	1811+1.6kbA->G
CFTR-関連障害	CFTR	2183AA>G
CFTR-関連障害	CFTR	2183AA>G
CFTR-関連障害	CFTR	3849+10kbC>T
CFTR-関連障害	CFTR	A455E
CFTR-関連障害	CFTR	A559T
CFTR-関連障害	CFTR	C524X
CFTR-関連障害	CFTR	574delA, 574delA
CFTR-関連障害	CFTR	574delA, 574delA
CFTR-関連障害	CFTR	2108delA, 2108delA
CFTR-関連障害	CFTR	2108delA, 2108delA
CFTR-関連障害	CFTR	3171delC, 3171delC
CFTR-関連障害	CFTR	3171delC, 3171delC
CFTR-関連障害	CFTR	621+1G->T
CFTR-関連障害	CFTR	2105-2117del13insAGAAA
CFTR-関連障害	CFTR	2105-2117del13insAGAAA
CFTR-関連障害	CFTR	711+1G->T
CFTR-関連障害	CFTR	711+5G->A
CFTR-関連障害	CFTR	712-1G->T
CFTR-関連障害	CFTR	1288insTA, 1288insTA

【図 3 - 10】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	1288insTA, 1288insTA
CFTR-関連障害	CFTR	936delTA
CFTR-関連障害	CFTR	936delTA
CFTR-関連障害	CFTR	[δ] JF311
CFTR-関連障害	CFTR	[δ] JF311
CFTR-関連障害	CFTR	1078delT, 1078delT
CFTR-関連障害	CFTR	1078delT, 1078delT
CFTR-関連障害	CFTR	1161delC, 1161delC
CFTR-関連障害	CFTR	1161delC, 1161delC
CFTR-関連障害	CFTR	1609delCA, 1609delCA
CFTR-関連障害	CFTR	1609delCA, 1609delCA
CFTR-関連障害	CFTR	[δ] JF507
CFTR-関連障害	CFTR	[δ] JF507
CFTR-関連障害	CFTR	rs332, [δ] JF508
CFTR-関連障害	CFTR	rs332, [δ] JF508
CFTR-関連障害	CFTR	1677delTA, 1677delTA
CFTR-関連障害	CFTR	1677delTA, 1677delTA
CFTR-関連障害	CFTR	1717-1G->A
CFTR-関連障害	CFTR	1812-1G->A
CFTR-関連障害	CFTR	1898+1G->A
CFTR-関連障害	CFTR	1898+1G->T
CFTR-関連障害	CFTR	1898+5G->T
CFTR-関連障害	CFTR	1949del84, 1949del84
CFTR-関連障害	CFTR	1949del84, 1949del84
CFTR-関連障害	CFTR	2043delG, 2043delG

【図 3 - 1 1】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	2043delG, 2043delG
CFTR-関連障害	CFTR	2055del9->A
CFTR-関連障害	CFTR	2055del9->A
CFTR-関連障害	CFTR	2143delT, 2143delT
CFTR-関連障害	CFTR	2143delT, 2143delT
CFTR-関連障害	CFTR	2184delA, 2184delA
CFTR-関連障害	CFTR	2184delA, 2184delA
CFTR-関連障害	CFTR	2184insA, 2184insA
CFTR-関連障害	CFTR	2184insA, 2184insA
CFTR-関連障害	CFTR	2307insA, 2307insA
CFTR-関連障害	CFTR	2307insA, 2307insA
CFTR-関連障害	CFTR	296+12T->C
CFTR-関連障害	CFTR	2789+5G->A
CFTR-関連障害	CFTR	2869insG, 2869insG
CFTR-関連障害	CFTR	2869insG, 2869insG
CFTR-関連障害	CFTR	3120G->A
CFTR-関連障害	CFTR	3120+1G->A
CFTR-関連障害	CFTR	3272-26A->G
CFTR-関連障害	CFTR	3659delC, 3659delC
CFTR-関連障害	CFTR	3659delC, 3659delC
CFTR-関連障害	CFTR	3667del4, 3667del4
CFTR-関連障害	CFTR	3667del4, 3667del4
CFTR-関連障害	CFTR	3791delC, 3791delC
CFTR-関連障害	CFTR	3791delC, 3791delC
CFTR-関連障害	CFTR	3821delT, 3821delT

【図 3 - 1 2】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	3821delT, 3821delT
CFTR-関連障害	CFTR	3905insT, 3905insT
CFTR-関連障害	CFTR	3905insT, 3905insT
CFTR-関連障害	CFTR	4016insT, 4016insT
CFTR-関連障害	CFTR	4016insT, 4016insT
CFTR-関連障害	CFTR	394delTT, 394delTT
CFTR-関連障害	CFTR	394delTT, 394delTT
CFTR-関連障害	CFTR	405+1G->A
CFTR-関連障害	CFTR	405+3A->C
CFTR-関連障害	CFTR	444delA
CFTR-関連障害	CFTR	444delA
CFTR-関連障害	CFTR	3876delA, 3876delA
CFTR-関連障害	CFTR	3876delA, 3876delA
CFTR-関連障害	CFTR	457TAT->G
CFTR-関連障害	CFTR	457TAT->G
CFTR-関連障害	CFTR	3199del6, 3199del6
CFTR-関連障害	CFTR	3199del6, 3199del6
CFTR-関連障害	CFTR	406-1G->A
CFTR-関連障害	CFTR	663delT, 663delT
CFTR-関連障害	CFTR	663delT, 663delT
CFTR-関連障害	CFTR	935delA, 935delA
CFTR-関連障害	CFTR	935delA, 935delA
CFTR-関連障害	CFTR	CFTR dele2,3 (21kb)
CFTR-関連障害	CFTR	CFTR dele2,3 (21kb)
CFTR-関連障害	CFTR	D1152H

【図 3 - 1 3】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	E60X
CFTR-関連障害	CFTR	E92X
CFTR-関連障害	CFTR	F508C, rs1800093
CFTR-関連障害	CFTR	G178R
CFTR-関連障害	CFTR	G330X
CFTR-関連障害	CFTR	G480C
CFTR-関連障害	CFTR	G542X
CFTR-関連障害	CFTR	G551D
CFTR-関連障害	CFTR	G622D
CFTR-関連障害	CFTR	G85E
CFTR-関連障害	CFTR	G91R
CFTR-関連障害	CFTR	H48T, rs35516286
CFTR-関連障害	CFTR	I506V
CFTR-関連障害	CFTR	IVS8-5T
CFTR-関連障害	CFTR	IVS8-7T
CFTR-関連障害	CFTR	IVS8-9T
CFTR-関連障害	CFTR	K710X
CFTR-関連障害	CFTR	L206W
CFTR-関連障害	CFTR	M1101K, rs36210737
CFTR-関連障害	CFTR	N1303K
CFTR-関連障害	CFTR	P574H
CFTR-関連障害	CFTR	Q1238X
CFTR-関連障害	CFTR	Q359K/T360K_wt
CFTR-関連障害	CFTR	Q493X
CFTR-関連障害	CFTR	Q552X

【図 3 - 1 4】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	Q890X
CFTR-関連障害	CFTR	R1066C
CFTR-関連障害	CFTR	R1070Q
CFTR-関連障害	CFTR	R1158X
CFTR-関連障害	CFTR	R1162X
CFTR-関連障害	CFTR	R117C
CFTR-関連障害	CFTR	R117H
CFTR-関連障害	CFTR	R1283M
CFTR-関連障害	CFTR	R334W
CFTR-関連障害	CFTR	R347H
CFTR-関連障害	CFTR	R347P
CFTR-関連障害	CFTR	R352Q
CFTR-関連障害	CFTR	R553X
CFTR-関連障害	CFTR	R560T
CFTR-関連障害	CFTR	R709X
CFTR-関連障害	CFTR	R75X
CFTR-関連障害	CFTR	R764X
CFTR-関連障害	CFTR	S1196X
CFTR-関連障害	CFTR	S1235R, rs34911792
CFTR-関連障害	CFTR	S1251N
CFTR-関連障害	CFTR	S1255X
CFTR-関連障害	CFTR	S364P
CFTR-関連障害	CFTR	S549I
CFTR-関連障害	CFTR	S549N
CFTR-関連障害	CFTR	S549R

【図 3 - 1 5】

疾患	遺伝子	変異体名
CFTR-関連障害	CFTR	S549R
CFTR-関連障害	CFTR	T338I
CFTR-関連障害	CFTR	V520F
CFTR-関連障害	CFTR	W1089X
CFTR-関連障害	CFTR	W1204X
CFTR-関連障害	CFTR	W1204X
CFTR-関連障害	CFTR	W1282X
CFTR-関連障害	CFTR	Y1092X
CFTR-関連障害	CFTR	Y122X
全胚総腸萎縮	CHM	c.1609+2dupT
全胚総腸萎縮	CHM	c.1609+2dupT
CLN3-関連神経セロイドリポフスチン症	CLN3	c.461_677del
CLN3-関連神経セロイドリポフスチン症	CLN3	c.461_677del
CLN3-関連神経セロイドリポフスチン症	CLN3	c.791_1056del
CLN3-関連神経セロイドリポフスチン症	CLN3	c.791_1056del
CLN5-関連神経セロイドリポフスチン症	CLN5	c.1175_1176delAT
CLN5-関連神経セロイドリポフスチン症	CLN5	c.1175_1176delAT
CLN5-関連神経セロイドリポフスチン症	CLN5	c.225G>A
CLN8-関連神経セロイドリポフスチン症	CLN8	c.70C>G
コーエン症候群	VPS13B	c.3348_3349delCT
コーエン症候群	VPS13B	c.3348_3349delCT
先天性白内障		
顔面異形症および神経障害	CTDP1	IVS6+389C>T

【図 3 - 1 6】

疾患	遺伝子	変異体名
先天性グリコシル化異常症	PMM2	p.F119L
先天性グリコシル化異常症	PMM2	p.R141H
先天性グリコシル化異常症	MP1	R295H, rs28928906
フィンランド型先天性ネフローゼ	NPHS1	c.121_122del
フィンランド型先天性ネフローゼ	NPHS1	c.121_122del
フィンランド型先天性ネフローゼ	NPHS1	c.3325C>T
クローン病	NOD2	3020 ins C
クローン病	NOD2	3020 ins C
クローン病	NOD2	G908R, rs2066845
クローン病	NOD2	R702W, rs2066844
シスチン症	CTNS	1035insC
シスチン症	CTNS	1035insC
シスチン症	CTNS	537del21
シスチン症	CTNS	537del21
シスチン症	CTNS	57kb 欠失
シスチン症	CTNS	57kb 欠失
シスチン症	CTNS	D205N
シスチン症	CTNS	L158P
シスチン症	CTNS	W138X
DFNA 9 (COCH)	COCH	P51S
糖尿病および難聴	mtDNA	3234A>G
糖尿病および難聴	mtDNA	3271T>C
糖尿病および難聴	mtDNA	G8363A
糖尿病および難聴	mtDNA	T14709C
早発性原発性ジストニア (DYT1)	TOR1A	904_906delGAG
早発性原発性ジストニア (DYT1)	TOR1A	904_906delGAG
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	3024delT
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	3024delT
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	p.Q243X
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	R144X

【図 3 - 1 7】

疾患	遺伝子	変異体名
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	R42X
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMB3	R635X
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMA3	R650X
接合型表皮水疱症		
ヘルリッツ-ピアソ	LAMC2	R95X
第 V 因子ライデ	F5	H1299R
第 V 因子ライデ	F5	R506Q, rs6025
第 V 因子 R2 突然変異	F5	rs6027
検査病	F11	E117X (576G>T)
第 XI 因子欠乏症	F11	F283L (1074T>C)
第 XI 因子欠乏症	F11	IVS14+1G>A
第 XI 因子欠乏症	F11	IVS14del14
第 XI 因子欠乏症	F11	IVS14del14
第 XIII 欠乏症	F13A1	V34L, rs5985
家族性腺腫性ポリポージス	APC	11307K, rs1801155
家族性自律神経失調症	IKBKAP	2507+6T>C
家族性自律神経失調症	IKBKAP	P914L
家族性自律神経失調症	IKBKAP	R696P
家族性高コレステロール血症 B 型	APOB	R3500Q, rs5742904
家族性高コレステロール血症 B 型	APOB	R3500W
家族性高コレステロール血症 B 型	APOB	R3531C, rs12713559
家族性地中海熱	MEFV	A744S (2230G>T)
家族性地中海熱	MEFV	delI692 (del2076_2078)
家族性地中海熱	MEFV	delI692 (del2076_2078)
家族性地中海熱	MEFV	E148Q (442 G>C), rs3743930

【図 3 - 1 8】

疾患	遺伝子	変異体名
家族性地中海熱	MEFV	E167D (501 G>C)
家族性地中海熱	MEFV	F479L (1437 C>G)
家族性地中海熱	MEFV	K695R (2084A>G)
家族性地中海熱	MEFV	M680I (2040G>C)
家族性地中海熱	MEFV	M694I (2082G>A), rs28940578
家族性地中海熱	MEFV	M694V (2080A>G)
家族性地中海熱	MEFV	P369S (1105 C>T), rs11466023
家族性地中海熱	MEFV	R408Q (1223G>A), rs11466024
家族性地中海熱	MEFV	R653H (1958G>A)
家族性地中海熱	MEFV	R761H (2282G>A)
家族性地中海熱	MEFV	T267I (800 C>T)
家族性地中海熱	MEFV	V726A (2177T>C), rs28940579
FANCC-関連ファンconi貧血	FANCC	322delG
FANCC-関連ファンconi貧血	FANCC	322delG
FANCC-関連ファンconi貧血	FANCC	IVS4+4A>T (711 +4A>T)
FANCC-関連ファンconi貧血	FANCC	Q13X (37C>T)
FANCC-関連ファンconi貧血	FANCC	R547X
FGFR1-関連頭蓋骨癒合症	FGFR1	P252R
FGFR2-関連頭蓋骨癒合症	FGFR2	P253R
FGFR2-関連頭蓋骨癒合症	FGFR2	S252W
FGFR3-関連頭蓋骨癒合症	FGFR3	A391E, rs28931615
FGFR3-関連頭蓋骨癒合症	FGFR3	P250R, rs4647924
近離シアル酸蓄積症	SLC17A5	c.1007_1008delTA
近離シアル酸蓄積症	SLC17A5	c.1007_1008delTA
近離シアル酸蓄積症	SLC17A5	c.115C>T

【図 3 - 19】

疾患	遺伝子	変異体名
パーキンソニズム-17 を伴う前頭側頭型認知症	MAPT	IVS10+16
パーキンソニズム-17 を伴う前頭側頭型認知症	MAPT	P301L
パーキンソニズム-17 を伴う前頭側頭型認知症	MAPT	P301S
パーキンソニズム-17 を伴う前頭側頭型認知症	MAPT	R406W
フマラーゼ欠乏症	FH	c.1431_1433dupAAA
フマラーゼ欠乏症	FH	c.1431_1433dupAAA
ガラクトース血症	GALT	5.0Kb 遺伝子欠失
ガラクトース血症	GALT	5.0Kb 遺伝子欠失
ガラクトース血症	GALT	5'UTR-119del
ガラクトース血症	GALT	5'UTR-119del
ガラクトース血症	GALT	IVS2-2 A>G
ガラクトース血症	GALT	K285N
ガラクトース血症	GALT	L195P T>C
ガラクトース血症	GALT	L218L
ガラクトース血症	GALT	N314D, rs2070074
ガラクトース血症	GALT	Phe171Ser
ガラクトース血症	GALT	Q169K
ガラクトース血症	GALT	Q188R
ガラクトース血症	GALT	S135L
ガラクトース血症	GALT	T138M C>T
ガラクトース血症	GALT	X380R
ガラクトース血症	GALT	Y209C A>G
ゴーシェ病	GBA	1035insG
ゴーシェ病	GBA	1035insG
ゴーシェ病	GBA	84insG
ゴーシェ病	GBA	84insG
ゴーシェ病	GBA	D409H, rs1064651
ゴーシェ病	GBA	D409V
ゴーシェ病	GBA	IVS2(+1)G>A
ゴーシェ病	GBA	L444P (1448T>C), rs35095275
ゴーシェ病	GBA	N370S
ゴーシェ病	GBA	R463C
ゴーシェ病	GBA	R463H
ゴーシェ病	GBA	R496H (1604G>A)
ゴーシェ病	GBA	V394L
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	167delT

【図 3 - 20】

疾患	遺伝子	変異体名
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	167delT
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	235delC
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	235delC
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	35delG
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	35delG
GJB2-関連 DFNA3 非症候性難聴および聴覚消失	GJB2	IVS1+1G>A
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	101delAG
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	313del14
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	313del14
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	delE120
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	delE120
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	M34T, rs35887622

【図 3 - 21】

疾患	遺伝子	変異体名
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	Q124X
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	R184P
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	V37I
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	W24X
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	W77R
GJB2-関連 DFNA1 非症候性難聴および聴覚消失	GJB2	W77X
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	A335V
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	R459L
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	R459P
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	rs1050828, rs1050828
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	rs1050829, rs1050829
グルコース-6-リン酸デヒドロゲナーゼ欠乏症	G6PD	rs5030868, rs5030868
グルタル酸血症 I 型	GCDH	A421V
グルタル酸血症 I 型	GCDH	R402W
糖原貯蔵症 Ia 型	G6PC	459insTA
糖原貯蔵症 Ia 型	G6PC	459insTA

【図 3 - 22】

疾患	遺伝子	変異体名
糖原貯蔵症 Ia 型	G6PC	727G>T
糖原貯蔵症 Ia 型	G6PC	del F327
糖原貯蔵症 Ia 型	G6PC	del F327
糖原貯蔵症 Ia 型	G6PC	G188R
糖原貯蔵症 Ia 型	G6PC	G270V
糖原貯蔵症 Ia 型	G6PC	Q242X
糖原貯蔵症 Ia 型	G6PC	Q27fsdelC
糖原貯蔵症 Ia 型	G6PC	Q27fsdelC
糖原貯蔵症 Ia 型	G6PC	Q347X
糖原貯蔵症 Ia 型	G6PC	R83C
糖原貯蔵症 Ia 型	G6PC	R83H
糖原貯蔵症 Ib 型	G6PT1	I211delCT
糖原貯蔵症 Ib 型	G6PT1	A367T
糖原貯蔵症 Ib 型	G6PT1	G339C
糖原貯蔵症 Ib 型	G6PT1	G339D
糖原貯蔵症 Ib 型	G6PT1	W118R
糖原貯蔵症 II 型	GAA	Arg854X
糖原貯蔵症 II 型	GAA	Asp645Glu, rs28940868
糖原貯蔵症 II 型	GAA	IVS1(-13c>g)
糖原貯蔵症 III 型	AGL	I484delT
糖原貯蔵症 III 型	AGL	I484delT
糖原貯蔵症 III 型	AGL	I7delAG
糖原貯蔵症 III 型	AGL	I7delAG
糖原貯蔵症 V 型	PYGM	Q6X
糖原貯蔵症 V 型	PYGM	G204S

【図 3 - 2 3】

疾患	遺伝子	変異体名
糖原貯蔵症 V 型	PYGM	K542T
糖原貯蔵症 V 型	PYGM	K542X
糖原貯蔵症 V 型	PYGM	R49X
GNE-関連ミオパシー	GNE	M712T, rs28937594
薄皮症候群	BCSL1	c.232A>G, rs28937590
ヘモグロビン S β-サ ラセミア	HBG	c.19G>A
ヘモグロビン S β-サ ラセミア	HBG	c.20A>T
ヘモグロビン S β-サ ラセミア	HBG	c.79G>A
ヘモグロビン S β-サ ラセミア	HBG	Hb CS
ヘモグロビン S β-サ ラセミア	HBG	Hb D
ヘモグロビン S β-サ ラセミア	HBG	Hb O
遺伝性フルクトース不 耐症	ALDOB	A149P, rs1800546
遺伝性フルクトース不 耐症	ALDOB	A174D
遺伝性フルクトース不 耐症	ALDOB	Delta4E4
遺伝性フルクトース不 耐症	ALDOB	Delta4E4
遺伝性フルクトース不 耐症	ALDOB	N334K
遺伝性フルクトース不 耐症	ALDOB	Y203X
遺伝性肺炎	PRSS1	A16V
遺伝性肺炎	SPINK1	MIT
遺伝性肺炎	PRSS1	N29I
遺伝性肺炎	SPINK1	N34S, rs17107315
遺伝性肺炎	PRSS1	R122C
遺伝性肺炎	PRSS1	R122H
遺伝性チミン-ウラシル 尿症	DPYD	rs3918290
ヘキシサミニダーゼ A 欠乏症	HEXA	I278insTATC
ヘキシサミニダーゼ A 欠乏症	HEXA	I278insTATC
ヘキシサミニダーゼ A 欠乏症	HEXA	G269S (805G>A)
ヘキシサミニダーゼ A 欠乏症	HEXA	IVS12 +1G>C

【図 3 - 2 4】

疾患	遺伝子	変異体名
ヘキシサミニダーゼ A 欠乏症	HEXA	IVS7 +1G>A
ヘキシサミニダーゼ A 欠乏症	HEXA	IVS9 +1G>A
ヘキシサミニダーゼ A 欠乏症	HEXA	R178C
ヘキシサミニダーゼ A 欠乏症	HEXA	R178H
ヘキシサミニダーゼ A 欠乏症	HEXA	R247W (739C>T)
ヘキシサミニダーゼ A 欠乏症	HEXA	R249W (745C>T)
HFE-関連遺伝性ヘ モクロマトーシス	HFE	E168Q
HFE-関連遺伝性ヘ モクロマトーシス	HFE	E168X
HFE-関連遺伝性ヘ モクロマトーシス	HFE	HM971246, H63H
HFE-関連遺伝性ヘ モクロマトーシス	HFE	P160delC
HFE-関連遺伝性ヘ モクロマトーシス	HFE	P160delC
HFE-関連遺伝性ヘ モクロマトーシス	HFE	Q127H, rs28934595
HFE-関連遺伝性ヘ モクロマトーシス	HFE	Q283P
HFE-関連遺伝性ヘ モクロマトーシス	HFE	rs1799945, rs1799945
HFE-関連遺伝性ヘ モクロマトーシス	HFE	rs1800562, rs1800562
HFE-関連遺伝性ヘ モクロマトーシス	HFE	rs1800730, rs1800730
HFE-関連遺伝性ヘ モクロマトーシス	HFE	V53M
HFE-関連遺伝性ヘ モクロマトーシス	HFE	V59M

【図 3 - 2 5】

疾患	遺伝子	変異体名
HFE-関連遺伝性ヘ モクロマトーシス	HFE	W169X
発汗性外胚葉形成不 全症 2	GJB6	A88V, rs28937872
発汗性外胚葉形成不 全症 2	GJB6	G11R
発汗性外胚葉形成不 全症 2	GJB6	V37E
シスタチオニン β-シ ンターゼ欠乏により 引き起こされるホモ システチン尿症	CBS	G307S 919G>A
シスタチオニン β-シ ンターゼ欠乏により 引き起こされるホモ システチン尿症	CBS	I278T 833T>C, rs5742905
高カリウム性周期性四 肢麻痺 1 型	SCN4A	I693T
高カリウム性周期性四 肢麻痺 1 型	SCN4A	L689I
高カリウム性周期性四 肢麻痺 1 型	SCN4A	L689V
高カリウム性周期性四 肢麻痺 1 型	SCN4A	M1360V
高カリウム性周期性四 肢麻痺 1 型	SCN4A	M1592V
高カリウム性周期性四 肢麻痺 1 型	SCN4A	p.A1156T
高カリウム性周期性四 肢麻痺 1 型	SCN4A	p.M1370V
高カリウム性周期性四 肢麻痺 1 型	SCN4A	p.R1448C
高カリウム性周期性四 肢麻痺 1 型	SCN4A	p.T1313M
高カリウム性周期性四 肢麻痺 1 型	SCN4A	R675G
高カリウム性周期性四 肢麻痺 1 型	SCN4A	R675Q
高カリウム性周期性四 肢麻痺 1 型	SCN4A	R675W
高カリウム性周期性四 肢麻痺 1 型	SCN4A	T704M
高カリウム性周期性四 肢麻痺 1 型	SCN4A	V781I
高オルニチン血症-高ア ンモニア血症-ホモシ トルリン尿症候群	SLC25A15	F188del

【図 3 - 2 6】

疾患	遺伝子	変異体名
高オルニチン血症-高ア ンモニア血症-ホモシ トルリン尿症候群	SLC25A15	F188del
原発性 1 型高シュ ウ酸尿症	AGXT	33insC
原発性 1 型高シュ ウ酸尿症	AGXT	33insC
原発性 1 型高シュ ウ酸尿症	AGXT	F152I
原発性 1 型高シュ ウ酸尿症	AGXT	G170R
原発性 1 型高シュ ウ酸尿症	AGXT	I244T
原発性 2 型高シュ ウ酸尿症	GRHPR	I03delG
原発性 2 型高シュ ウ酸尿症	GRHPR	I03delG
軟骨低形成症	FGFR3	Asn328Ile
軟骨低形成症	FGFR3	I538V
軟骨低形成症	FGFR3	K650M
軟骨低形成症	FGFR3	K650N 1950G>T
軟骨低形成症	FGFR3	K650Q
軟骨低形成症	FGFR3	N540K 1620C>A
軟骨低形成症	FGFR3	N540S
軟骨低形成症	FGFR3	N540T
低カリウム血性周期性 麻痺 1 型	CACNA1S	R528G
低カリウム血性周期性 麻痺 1 型	CACNA1S	R528H
低カリウム血性周期性 麻痺 1 型	CACNA1S	rs28930068, rs28930068
低カリウム血性周期性 麻痺 1 型	CACNA1S	rs28930069, rs28930069
低カリウム血性周期性 麻痺 2 型	SCN4A	R669H
低カリウム血性周期性 麻痺 2 型	SCN4A	R672C
低カリウム血性周期性 麻痺 2 型	SCN4A	R672G
低カリウム血性周期性 麻痺 2 型	SCN4A	R672H
低カリウム血性周期性 麻痺 2 型	SCN4A	R672S
低ホスファターゼ症	ALPL	Asp361Val
低ホスファターゼ症	ALPL	c.1559delT
低ホスファターゼ症	ALPL	c.1559delT
低ホスファターゼ症	ALPL	E174K
低ホスファターゼ症	ALPL	G317D
低ホスファターゼ症	ALPL	Phc310Lcu

【図 3 - 27】

疾患	遺伝子	変異体名
イソ吉草酸血症	IVD	A282V
イソ吉草酸血症	IVD	rs28940889
クラッペ病	GALC	EX11-17DEL
クラッペ病	GALC	EX11-17DEL
クラッペ病	GALC	G270D
クラッペ病	GALC	rs1805078, rs1805078
クラッペ病	GALC	rs398607
レーバー遺伝性 視神経萎縮症	mtDNA	14484T>C
レーバー遺伝性 視神経萎縮症	mtDNA	15257G>A
レーバー遺伝性 視神経萎縮症	mtDNA	G14459A
レーバー遺伝性 視神経萎縮症	mtDNA	G3460A
レーバー遺伝性 視神経萎縮症	mtDNA	m.11778G>A
レーバー遺伝性 視神経萎縮症	mtDNA	m.13708G>A
レーバー遺伝性 視神経萎縮症	mtDNA	m.15812G>A
レーバー遺伝性 視神経萎縮症	mtDNA	m.3394T>C
レーバー遺伝性 視神経萎縮症	mtDNA	m.4216T>C
レーバー遺伝性 視神経萎縮症	mtDNA	m.4917A>G
リー症候群フラン ス系カタグ型 LGMID21	LRPPRC FKRP	A354V L276I, rs28937900
長鎖 3-ヒドロキシ アシル-CoA デヒド ロゲナーゼ欠乏症	HADHA	E474Q c.1528G>C
長鎖 3-ヒドロキシ アシル-CoA デヒド ロゲナーゼ欠乏症	HADHA	Q342X 1132C>T
メーブルシロップ尿 症 1A 型	BCKDHA	Y438N
メーブルシロップ尿 症 1B 型	BCKDHB	E372X
メーブルシロップ尿 症 1B 型	BCKDHB	G278S
メーブルシロップ尿 症 1B 型	BCKDHB	R183P
マクカーソーオル ブライト症候群	GNAS	R201C
マクカーソーオル ブライト症候群	GNAS	R201G

【図 3 - 28】

疾患	遺伝子	変異体名
マクカーソーオル ブライト症候群	GNAS	R201H
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	244insT
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	244insT
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	250C>T
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	583G>A
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	616C>T
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	617G>A
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	799G>A
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	K304E
中鎖アシル-CoA デヒド ロゲナーゼ欠乏症	ACADM	Y42H
皮質下囊胞を伴う 巨脳性白質脳症	MLC1	135insC
皮質下囊胞を伴う 巨脳性白質脳症	MLC1	135insC
MELAS	mtDNA	3243A>G
MELAS	mtDNA	3250T>C
MELAS	mtDNA	3252A>G
MELAS	mtDNA	A12770G
MELAS	mtDNA	C3256T
MELAS	mtDNA	G13513A

【図 3 - 29】

疾患	遺伝子	変異体名
MELAS	mtDNA	T3291C
MELAS	mtDNA	T8356C
MELAS	mtDNA	T9957C
MERRF	mtDNA	8361G>A
MERRF	mtDNA	A8296G
MERRF	mtDNA	m.8344A>G
異染性白質ジス トロフィー	ARSA	c.459+11G>A
異染性白質ジス トロフィー	ARSA	p.P426L, rs28940893
異染性白質ジス トロフィー	ARSA	p.T274M
異染性白質ジス トロフィー	ARSA	P377L
ミトコンドリア 心筋症	mtDNA	A3260T
ミトコンドリア 心筋症	mtDNA	A4300G
ミトコンドリア 心筋症	mtDNA	C3303T
ミトコンドリア 心筋症	mtDNA	T9997C
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	5537insT
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	5537insT
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	8993T>C
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	8993T>G
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	C11777A
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	T10158C
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	T10191C
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	T8851C
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	T9176C

【図 3 - 30】

疾患	遺伝子	変異体名
ミトコンドリア DNA- 関連リー症候群および NARP	mtDNA	T9176G
MTHFR 欠乏症	MTHFR	1298A>C
MTHFR 欠乏症	MTHFR	rs1801133, rs1801133
MTRNR1-関連難聴お よび聴覚消失	mtDNA	1095T>C
MTRNR1-関連難聴お よび聴覚消失	mtDNA	1494C>T
MTRNR1-関連難聴お よび聴覚消失	mtDNA	1555A>G
MTRNR1-関連難聴お よび聴覚消失	mtDNA	961T>G
MTRNR1-関連難聴お よび聴覚消失	mtDNA	A7445G
MTTS1-関連難聴お よび聴覚消失	mtDNA	7443A>G
MTTS1-関連難聴お よび聴覚消失	mtDNA	7444G>A
MTTS1-関連難聴お よび聴覚消失	mtDNA	7472insC
MTTS1-関連難聴お よび聴覚消失	mtDNA	7472insC
MTTS1-関連難聴お よび聴覚消失	mtDNA	7510T>C
MTTS1-関連難聴お よび聴覚消失	mtDNA	7511T>C
MTTS1-関連難聴お よび聴覚消失	mtDNA	7512T>C
ムコリビド症 IV	MCOLN1	delEx1 3 Ex7 (511>6944del)
ムコリビド症 IV	MCOLN1	delEx1 3 Ex7 (511>6944del)
ムコリビド症 IV	MCOLN1	IVS-2A>G
ムコ多糖症 I 型	IDUA	c.46_57del
ムコ多糖症 I 型	IDUA	c.46_57del
ムコ多糖症 I 型	IDUA	p.A327P

【図 3 - 3 1】

疾患	遺伝子	変異体名
ムコ多糖症 I 型	IDUA	p.P533R
ムコ多糖症 I 型	IDUA	Q70X
ムコ多糖症 I 型	IDUA	W402X
ムコ多糖症 IIIA 型	SGSH	p.R245H
ムコ多糖症 IIIA 型	SGSH	p.R74C
ムコ多糖症 IIIA 型	SGSH	p.S66W
ムコ多糖症 VII 型	GUSB	p.D152N
多内分泌腺腫瘍 2 型	RET	2047T>A
多内分泌腺腫瘍 2 型	RET	2047T>A
多内分泌腺腫瘍 2 型	RET	2047T>C
多内分泌腺腫瘍 2 型	RET	2047T>G
多内分泌腺腫瘍 2 型	RET	2048G>A
多内分泌腺腫瘍 2 型	RET	A883F 2647 G>T
多内分泌腺腫瘍 2 型	RET	Glu768Asp G>C
多内分泌腺腫瘍 2 型	RET	M918T
筋-眼-脳病	POMGNT1	c.1539+1G>A
MYH-関連ポリポージス	MUTYH	c.1376C>A
MYH-関連ポリポージス	MUTYH	c.494A>G, rs34612342
MYH-関連ポリポージス	GENE_SYMBOL_TBD	rs36053993, rs36053993
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	c.990delC
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	c.990delC
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	fsP330

【図 3 - 3 2】

疾患	遺伝子	変異体名
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	fsP330
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	L302P
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	R496L
スフィンゴミエリナーゼ欠乏に起因するニーマン-ピック病	SMPD1	R608del
ニーマン-ピック病 C1 型	NPC1	I1061T
ナイミールヘン染色体不安定症候群	NBN	657del5
ナイミールヘン染色体不安定症候群	NBN	657del5
バリスター-ホル症候群	GLI3	2012delG
バリスター-ホル症候群	GLI3	2012delG
バリスター-ホル症候群	GLI3	2023delG
バリスター-ホル症候群	GLI3	2023delG
ベンドレッド症候群	SLC26A4	I197delT
ベンドレッド症候群	SLC26A4	I197delT
ベンドレッド症候群	SLC26A4	E384G
ベンドレッド症候群	SLC26A4	IV58+1(G->A)
ベンドレッド症候群	SLC26A4	L236P
ベンドレッド症候群	SLC26A4	T416P
ベルオキシソーム二機能酵素欠乏症	HSD17B4	c.302+1G>C
ベルオキシソーム二機能酵素欠乏症	HSD17B4	c.303-1G>A
広汎性発達障害	NLGN4X	D396X
広汎性発達障害	NLGN4X	D396X

【図 3 - 3 3】

疾患	遺伝子	変異体名
広汎性発達障害	NLGN4X	NLGN4X:1253delAG
広汎性発達障害	NLGN4X	NLGN4X:1253delAG
広汎性発達障害	NLGN3	R451C
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	G272X
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	I65T
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	IVS12+1G>T
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	L48S, rs5030841
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	R158Q, rs5030843
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	R252W, rs5030847
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	R261Q, rs5030849
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	R408Q, rs5030859
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	R408W, rs5030858
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	rs5030855
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	rs5030861
フェニルアラニンヒドロキシラーゼ欠乏症	PAH	Y414C, rs5030860
プラスミノーゲン活性化因子抑制剤 I	SERPINE1	-844 G>A
プラスミノーゲン活性化因子抑制剤 I	SERPINE1	4G/5G

【図 3 - 3 4】

疾患	遺伝子	変異体名
多発性嚢胞腎、常染色体劣性	PKHD1	c.10412T>G
多発性嚢胞腎、常染色体劣性	PKHD1	c.107C>T
多発性嚢胞腎、常染色体劣性	PKHD1	c.1486C>T
多発性嚢胞腎、常染色体劣性	PKHD1	c.5895dupA
多発性嚢胞腎、常染色体劣性	PKHD1	c.5895dupA
多発性嚢胞腎、常染色体劣性	PKHD1	c.9689delA
多発性嚢胞腎、常染色体劣性	PKHD1	c.9689delA
PPT1-関連ニューロセロイドーリボフスチン症	PPT1	c.364A>T
PPT1-関連ニューロセロイドーリボフスチン症	PPT1	p.L10X
PPT1-関連ニューロセロイドーリボフスチン症	PPT1	p.R151X
PPT1-関連ニューロセロイドーリボフスチン症	PPT1	T75P
PROPI-関連下垂体ホルモン欠乏症	PROPI	301-302delAG
PROPI-関連下垂体ホルモン欠乏症	PROPI	301-302delAG
プロトロンビン検査病	F2	rs1799963
プロトロンビン検査病	F2	rs6025, rs6025
ブソイドビタミン D 欠乏くる病	CYP27B1	エクソン 8 中 7bp 重複
ブソイドビタミン D 欠乏くる病	CYP27B1	エクソン 8 中 7bp 重複
ブソイドビタミン D 欠乏くる病	CYP27B1	958delG

【 図 3 - 3 6 】

疾患	遺伝	変異体名
スミス・レムリ・オピッ ン症候群	DHCR7	R352Q
スミス・レムリ・オピッ ン症候群	DHCR7	R352W
スミス・レムリ・オピッ ン症候群	DHCR7	R404C
スミス・レムリ・オピッ ン症候群	DHCR7	R446Q
スミス・レムリ・オピッ ン症候群	DHCR7	T93M
スミス・レムリ・オピッ ン症候群	DHCR7	V326L
スミス・レムリ・オピッ ン症候群	DHCR7	W151X
スミス・レムリ・オピッ ン症候群	DHCR7	W151X
癌性対麻痺 ¹³ 硫酸輸送体-関連骨軟骨 異形成	HSPD1	V72I
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	340delV
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	340delV
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	c.837C>T
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	C653S
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	IVS1+2T>C
硫酸輸送体-関連骨軟骨 異形成	SLC26A2	R178X
TFR2-関連遺伝性ヘ モクロマトーシス	TFR2	AVAQ594-597del
TFR2-関連遺伝性ヘ モクロマトーシス	TFR2	AVAQ594-597del
TFR2-関連遺伝性ヘ モクロマトーシス	TFR2	F60X

【 図 3 - 3 8 】

疾患	遺伝	変異体名
チロシンヒドロキシターゼ欠乏症 DRD	TH	R202H
チロシン血症Ⅰ型	FAH	E357X
チロシン血症Ⅰ型	FAH	IVS12+5 G>A
チロシン血症Ⅰ型	FAH	IVS7-6 T>G
チロシン血症Ⅰ型	FAH	IVS8-1 G>C
チロシン血症Ⅰ型	FAH	p.W262X
チロシン血症Ⅰ型	FAH	P261L
チロシン血症Ⅰ型	FAH	Q64H
ウイルソン病	ATP7B	130delG4
ウイルソン病	ATP7B	3402delC
ウイルソン病	ATP7B	3402delC
ウイルソン病	ATP7B	H106Q9
ウイルソン病	ATP7B	R778G
ウイルソン病	ATP7B	W779X
ウイルソン病	ATP7B	W779X
X-結合若年膜隔離症	RS1	E72K
X-結合若年膜隔離症	RS1	G109R
X-結合若年膜隔離症	RS1	G74V
ゼルウィガー症候群スペクトル	PEX1	c.2097_2098in
ゼルウィガー症候群スペクトル	PEX1	c.2097_2098in
ゼルウィガー症候群スペクトル	PEX1	c.2916delA
ゼルウィガー症候群スペクトル	PEX1	c.2916delA
ゼルウィガー症候群スペクトル	PEX1	n.G843D

【図 4 - 1】

配列番号22	GAGGATTTCAGGAGTCAATGACTGAGGATGGGACTCTCTGA
配列番号23	TGAGGGCTTGGACCAAAATCTTCAAGCAAAAACAGAAAAACA
配列番号24	ACTGGACCGCCCTCCACCGCCCTCCACCGCGGGCCCT
配列番号25	TCCTTTTCCGAGACAACTTCATCTCGAAAGGCTGTCA
配列番号26	TTTGTGAGCTTACAGTGGAAATGCTATTAATTTCTTTC
配列番号27	ACTTAGAAAGTTAAAGTAAGAAATATTAATAICTCTAT
配列番号28	TCAGAAAGGGCAAAAGCTTCTCTCTCGCATCCCTCATG
配列番号29	TTATTTTGTCTCTGCTGTTCATGGCATAGTTTGGTGGCG
配列番号30	AATGGCTTGCACCTGAGAAATCTATTGTATATGGCAAGAC
配列番号31	AGAGCAAGAGGACCTGGGAGGTGCTGCACCCCATACCA
配列番号32	CAAAATAGAAATGCTCTTATAGATGAGTATCAAAAATAAT
配列番号33	TCCTCCGCTCTCTCGCGGGGTGCTGAAACAGCCCGGG
配列番号34	ACCCGGGCTGAGCGCTGCTGGGCCGCTGCGCTTCCCGG
配列番号35	TTCTACCTGTGAGCAGGAATCTAGGACACAGTCCTGTAC
配列番号36	TGCCCTGTGAGACCTACACGCCCCCACCATTGCCCCAC
配列番号37	GGCGCTCTGTGCTGCGGCTGGTGGGCTACTACATCTTC
配列番号38	GAGCGTGATTAGGTACTGGACACCTGCCAAGTGTGGGCT
配列番号39	CTGGGATTGAGGGTTTTCATACACCTTCTGCTAGGATAA
配列番号40	TAAAAATTAATAAATACAGTTAAAAATCATGGTCATATA
配列番号41	CCGCTGCACTGACTTCATTTCTTAGACAAGACACAGTGT
配列番号42	ACTGCAACATTTTCAAGCAAAAGAAATCCCGTTGCTGTCG
配列番号43	CTTAGCTACGCTCCAGGCTGTGACGAGAAATACAGGGAC
配列番号44	ATTTTAGATTCAAAATTTGATGCCGATTACATTTTCTCAA
配列番号45	AGGCAAGCTGTCTCCAGGCTTTTATCAGACAGTGCCCGC
配列番号46	TTTAAAGTTTCTGTGACCTTTGTTAGAAAGTTTATAAATG
配列番号47	AATAGTAGGCTGTGTGACATTTCTCAACTTACTTATAAA
配列番号48	AACAGCTTCTGCTGCTGTAAGTGGCTGTCTGCCACA
配列番号49	CTTAGAATCTCTCTGAGGCTGAGCTGAGGGCCCTTGGG
配列番号50	CTATTTCTCTCTCTTATTTTCAGAAATAGAAAAGCAATTC
配列番号51	CACGGACATACGATACCGGCCAGTGACAGTCAGGGCAA
配列番号52	ATGCCATTACACAGTAATAACCAATGACATAGTACA
配列番号53	TCCTTTGGAAACAGTGTGGACCCAGTGTGCTCCGACA
配列番号54	TGTACAGGATGTACTGTACTGAGTGTGACAGCAACTAT
配列番号55	CACCTGCTGATGAGAGGTGGGCTGGGGCACTAAACCCG
配列番号56	CTGCGAGTGGGATTTCTCTGAAAGAGACACAGTGAACAG
配列番号57	AACCTATTATTTTAACTACCTGCTTCTATTGTTGAAAGAAA
配列番号58	AGCCACTGTGCGCGGCTCAGATATTTCTTCAAGTAAATGA
配列番号59	CATTGCTGTGAGTGGCCCTCAGTTTACATAGTGTATCTT
配列番号60	TAATTTTATTCGCCATTAGGATGAAATCCATATTCACAAA
配列番号61	TCAAGCCAGCTGGGAAGGAGATGGAAGCTGCTGCGC
配列番号62	TGCTGTTAAGATGTTACTTCTTTAAAAAAGATGGGTTAT
配列番号63	AAAAAATATGCTTATTAGAATCAAAATATGATAGCAAAAC
配列番号64	GTTAATATTTTATGCTAATGACAGCAATATATATTAAGT
配列番号65	GCTGTACAGAGCTATATATCATATAATTATTTATACATATG
配列番号66	CAGGAGGATCAGTCTCTGTAGAGGACGGGAGGAGCTGGGG
配列番号67	TATTTTCAGGTACTTCTGAAATGATAGCAATTTGTG
配列番号68	TGTTAGTTTTTCAGTTTCTCTGAAAGTCTATCTCTAGA
配列番号69	TGTAGCCCTTTGAGCATGAGGTATGCAATAGAACATAATG
配列番号70	CTTGCATCAAGTAAGGTGAAATATTCATATCTGTTCT

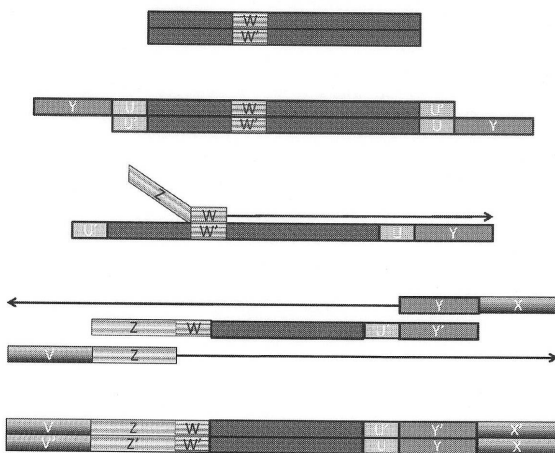
【図 4 - 2】

配列番号71	CGCGGGCCGCTAGGCTGATTGGCTGCTGCAGCCACCC
配列番号72	TTTCGGGAGGAGGAGAGGGTGGGGTGGCGGGTGCAGACT
配列番号73	CCTCAGCCACAACCATTAGCTGCAACGGTCCAGGCTCGTG
配列番号74	ATCTCTCGCATTTCTGCTGAGGCTGTCTTTTCTTCTT
配列番号75	CCCGGGCAGTCTGGGCTTGAACGTGTGTGTCAGCCGCGC
配列番号76	GGAAACAAGGGGTCTCCGAGCAGCCCGAGCCCTCCCTTC
配列番号77	GCACCTTCCCGCAGCGGCTGGGTGAGCCCTGGGAGCTGA
配列番号78	AAGGTTTTCAGAAAGTTAAATGGAATAGAAACATTTTGG
配列番号79	AAATCAATTTCTGTTTCTTAAAGTAAATTTCTCATGAGCAT
配列番号80	TAGGCACTTCCAGCTGGTGTCAATCCCTGATCACTGGGAG
配列番号81	CAGAACTTATAAAATATTGATAGGACGCTCTTGGGAG
配列番号82	CCAGCTTGGGACTATGCCCATGAGTGCCTGGCCATGCCCG
配列番号83	CGCCCTCAGAGTCCAGGCAAGGCAAGGAGGAGGAGG
配列番号84	GTTCTGGAAGAAATCTGCTCTGGGAGGAGGAGTGGGCGAGG
配列番号85	TGTCATCCCGAGCTCATCTCTCATCTGCTCAGTTTTC
配列番号86	AGGACTGTCTGTGGCATTTCCCTCTGGGATCTGAATGATGG
配列番号87	CCGGAGGAAAAATCTCTCATCTTTTGAAGCTATTGAG
配列番号88	ATATGTTGAGTATTTTGAATATCTCATACAATATAGCCTA
配列番号89	GGGCTGTGTTGTACCCCTGAGGATCTGCTGCAATGCCA
配列番号90	TACTAGTGTTTTCATTGGTATTAAGCTTGTGTAATATTT
配列番号91	AGCCACACGCTGGCCAGACTCAGAGAAATGAATACAAT
配列番号92	GTGGCGAGAAAGCATGAGGAATGGAGATGGAAGGAGGACAG
配列番号93	TCTCAGTTTGGCTGAGAAAGCAGGAGTGGGGCTTGAACCA
配列番号94	CAACATAGCAAGACCCGCTCACTATAAAATGAAAGAAC
配列番号95	CCAGTGGTGGGAGCCCGGGGAGGAGGAGGAGGAGGAGG
配列番号96	TTATTTTATGGATGTAACAGGCTCTTTGATGTTATA
配列番号97	TCCTGAACAAAGCAATTAAGAGGGAATTAACATAAATAA
配列番号98	TTATGTTATGAACATGATTTGTATCAATATGTGAGTTT
配列番号99	CTTTTCTTTTATAAAGGAGGACTCTTTGGCTGATATCTG
配列番号100	CCTGAAGTCTCAGTTTCCATTACATTAACCTCACTACC
配列番号101	CCACCCCTGGGCTCCTGGGGGCTGGGATCCAGATGGT
配列番号102	ACGGGATGAGGAGGGCGTGGTGTCTATGTGGCGGGGA
配列番号103	CCTCGGATTGAAGAAAGTCTGGTACTCATCTGTTGGCGTA
配列番号104	GTTTAAAAATTTGCTCTTATTTGTTCCAAATGTCTGCTTC
配列番号105	TAATGTGTAATGATAGGTCTTGTCAAAATGTTTAAATAGT
配列番号106	GAGTCCGATGCGGCTGACTGTCACTGCCACCATTCATCC
配列番号107	CTGAATGTTGCAAAATCTTAAATAAACATGTTCCAGAGGAGA
配列番号108	GCCTTATTTCCGTTTCCACTCTCTCTCTCTAGTTCATCC
配列番号109	TCAGGAAATCTTACAGTCCACACTCAGTCAGCCCAAGGA
配列番号110	CCTTCTCGGATCTCAAGAGCAGAGGGTTAAACATCATGA
配列番号111	GGGGCGCGGCCCTCAAGTCCGAGGACCTCCCTCTGGGG
配列番号112	AGTTCCTCAGGGCGCCCTGTGGCGGCGCCGCTGCACT
配列番号113	ACGACCTTATTACTCTATAACGATGAGTCTAGCAAGTAC
配列番号114	ACAAAAAAGGTAACTATGTAAAGACATGTTAATTAGC
配列番号115	CTTCGAGAAATCTGAAAAAAGTCAAAAGGTTTGAATGTTG
配列番号116	CTATTGAAGATTGTTCATCAATATTTGATGCATGATAGG
配列番号117	TTCCAGGCAAGCAGTACCTAAGGTTTACAGCTGATGA
配列番号118	CCCATCCAAGGAAATTTAGAAAAAGGAGGGATGTGTA

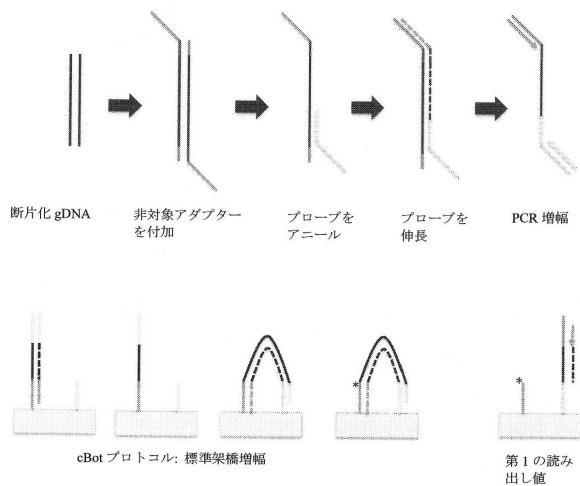
【図 4 - 3】

配列番号119	GAGTGGGAGGGGTAAAAGGGCTATAAAAAAATCTAAA
配列番号120	CCAATCATTTGCAACAAACAGAAACAGCTCTGACAGAGAAAGG
配列番号121	AATTGGAGGACACCAGTGGCATCAGGCTCTCTGTGTTGC

【図 5】



【図 6】



【図 9 - 3】

rs1129038, rs3769005, rs2596501, rs1364394, rs969539, rs974020, rs1922286, rs2171209, rs1560569, rs1744366, rs2596834, rs1890131, rs2418844, rs1157492, rs1922086, rs6432110, rs8041327, rs1416467, rs33706, rs379973, rs10496610, rs275974, rs822739, rs1517407, rs2187684, rs7997100, rs959763, rs1003306, rs1854226, rs2367191, rs202546, rs2905347, rs2236876, rs10508372, rs959260, rs920599, rs986642, rs2219248, rs2923031, rs495347, rs103294, rs2003092, rs1373557, rs16891982, rs1073321, rs2014303, rs1873195, rs1408794, rs9290675, rs7108371, rs1045873, rs253776, rs7163907, rs10853962, rs1777689, rs1032143, rs10512122, rs7908825, rs4859259, rs10516982, rs7965049, rs153595, rs7277342, rs10486207, rs3809125, rs2847502, rs10509384, rs2097884, rs4639533, rs10255965, rs1976033, rs1107820, rs10483853, rs4686497, rs3822616, rs4832640, rs2086085, rs2685159, rs12502036, rs9861816, rs2251432, rs8040452, rs7552348, rs3007711, rs1879558, rs1476162, rs17864053, rs477627, rs2047058, rs1065674, rs9860730, rs1448314, rs554788, rs1582398, rs1660964, rs10509594, rs17457687, rs4307284, rs11586379, rs12124147, rs7548659, rs760607, rs1981135, rs749663, rs1334804, rs10889750, rs1288367, rs213496, rs11207865, rs4915691, rs1566246, rs3101336, rs544858, rs10493649, rs12758138, rs6660484, rs2968485, rs712886, rs12061503, rs12046602, rs731756, rs10776798, rs6587597, rs912572, rs1419074, rs488150, rs1322780, rs859362, rs6880701, rs2146060, rs2784101, rs2494302, rs10494878, rs1930903, rs2275302, rs4916113, rs587913, rs6723966, rs10929646, rs4260216, rs4304747, rs1489688, rs4665797, rs1438131, rs13032262, rs10205008, rs605832, rs7594173, rs6544718, rs6713506, rs7577894, rs2075375, rs6720799, rs41420, rs13032535, rs11164050, rs1123861, rs6739285, rs1955394, rs16829231, rs6730157, rs2117742, rs1037266, rs6752189, rs1227131, rs6749847, rs7558428, rs9288172, rs3769823, rs10189499, rs1955117, rs10189760, rs6436553, rs6436883, rs10510268, rs3804989, rs17043611, rs342042, rs2128162, rs3916092, rs11914832, rs744751, rs285327, rs11544593, rs6773085, rs1130841, rs6795735, rs2054956, rs9310279, rs6551458, rs7630043, rs9289584, rs535234, rs7634702, rs4678297, rs9839394, rs10513729, rs7636818, rs12635682, rs4859259, rs6775595, rs6449375, rs13132286, rs6831024, rs6448770, rs9884706, rs10008492, rs12186184, rs1389037, rs6832891, rs1109501, rs4261956, rs3923243, rs1455313, rs6820697, rs1343921, rs13349989, rs7669241, rs2850971, rs2194860, rs1488817, rs1557815, rs4574434, rs1541745, rs128688, rs692157, rs1390009, rs10519410, rs1507500, rs10517660, rs1492468, rs1430975, rs11730565, rs1435442, rs4241802, rs10032784, rs2736122, rs7718757, rs1533019, rs29460, rs464923, rs1001935, rs2279095, rs1373967, rs173686, rs6452788, rs174015, rs10478046, rs6863510, rs6866231, rs346650, rs1528961, rs171718, rs4868204, rs2135053, rs7727897, rs248327, rs1933652, rs2320106, rs4960257, rs2876167, rs1992387, rs2237148, rs7763768, rs6918101, rs12660883, rs382259, rs2076173, rs8048447, rs12524885, rs4610536, rs638473, rs2451688, rs9389124, rs7573036, rs4896663, rs17551120, rs11756366, rs10428622, rs10950641, rs2041362, rs2189947, rs10267453, rs697518, rs2529015, rs10951140, rs1003549, rs933360, rs17135491, rs4020771, rs7791143, rs4730287, rs7781715, rs12705973, rs12706793, rs1593306, rs11761774, rs4571660, rs7795021, rs7880648, rs1548353, rs4565458, rs12548107, rs13273386, rs1127379, rs10109984, rs7830163, rs6993747, rs7827918, rs2380646, rs4634634, rs11708821, rs11784678, rs1026804, rs1841316, rs4871195, rs4736413, rs396861, rs10758823, rs568620, rs10962589, rs2840790, rs10511715, rs10812520, rs867469, rs2130118, rs10746763, rs10909807, rs884886, rs10373927, rs7866625, rs1009473, rs7090242, rs7921493, rs11255712, rs1324322, rs4948508, rs3858126, rs4746826, rs7072160, rs4691, rs4999437, rs817459, rs7097946, rs11186543, rs701873, rs1023331, rs7071247, rs10509826, rs1325172, rs540609,

【図 9 - 4】

rs11146457, rs2045272, rs7931276, rs1477569, rs4755844, rs2218868, rs2237997, rs12804561, rs2282504, rs7342241, rs2015747, rs488753, rs7131355, rs4936969, rs740851, rs7976721, rs10770437, rs699039, rs12370505, rs2088170, rs2860493, rs2723891, rs11107018, rs7307510, rs12231308, rs7312155, rs991817, rs11066320, rs2955105, rs2650170, rs1405050, rs3815210, rs11059335, rs2398513, rs377318, rs7323018, rs1979558, rs9539642, rs13379032, rs4885167, rs2175075, rs1678386, rs9556553, rs4984240, rs7152286, rs2038210, rs10498472, rs7492698, rs2180611, rs1531631, rs8003492, rs71126387, rs9671457, rs4906226, rs1947745, rs11634609, rs2611605, rs1153860, rs1906433, rs10519005, rs2292745, rs1320205, rs387727, rs8025028, rs3922394, rs1510058, rs8043261, rs2072986, rs1891325, rs896401, rs949429, rs30237, rs8047148, rs7205880, rs12446160, rs1872678, rs9937047, rs4238802, rs3751834, rs6539986, rs2253820, rs7215135, rs7220080, rs685098, rs2165846, rs11870879, rs196948, rs12939848, rs8071270, rs736632, rs400839, rs1786153, rs11564361, rs16948113, rs4941246, rs17079195, rs11151863, rs4799268, rs4536588, rs6510672, rs1552046, rs715159, rs8111998, rs2082455, rs17496703, rs7976, rs7252868, rs8103620, rs4814697, rs6076623, rs6118234, rs1041200, rs8116153, rs1884783, rs6126462, rs6129532, rs6032343, rs477627, rs4811158, rs2041317, rs6092326, rs3916504, rs6071491, rs977712, rs1735899, rs2832643, rs2836824, rs11702531, rs2253538, rs487828, rs1147509, rs4131364

北ヨーロッパ人とアシェクナージ系ユダヤ人を区別する AIM:

rs3769005, rs7551844, rs2596501, rs3007711, rs1129038, rs6996185, rs847851, rs17443616, rs1922086, rs740200, rs7108371, rs1065674, rs554788, rs1364394, rs1890131, rs1922286, rs10516982, rs477627, rs6448770, rs13084044, rs1355170, rs6432110, rs9860730, rs12713956, rs6028505, rs9576338, rs986642, rs4756052, rs7983897, rs10506555, rs822759, rs6751522, rs1991718, rs10508372, rs10493430, rs33706, rs379773, rs202546, rs920590, rs4307284, rs7163907, rs2251432, rs10762340, rs1660964, rs10853962, rs764255, rs4859259, rs923031, rs959260, rs8029021, rs2596834, rs2418844, rs998401, rs10496610, rs10509384, rs2003092, rs1560569, rs8016025, rs6477998, rs1517407, rs224378, rs17079195, rs10460810, rs6687300, rs2252815, rs1032143, rs1476162, rs17457687, rs4642918, rs4832640, rs7637803, rs7977100, rs6432398, rs6490700, rs4771165, rs359291, rs1373557, rs2171209, rs1073321, rs4885167, rs920559, rs90213, rs10497233, rs3822616, rs7819806, rs12430668, rs12441220, rs10486207, rs2377689, rs4336881, rs1598452, rs605832, rs2219248, rs8078782, rs1045873, rs552976, rs10431948, rs2793471, rs2905347, rs7316723, rs1118962, rs1032355, rs1582398, rs6841252, rs1743789, rs1408794, rs710807, rs914505, rs2154328, rs1870599, rs1270679, rs1854226, rs2014303, rs881433, rs10868791, rs10504275, rs6502048, rs1904406, rs17026435, rs8041327, rs1400174, rs11744977, rs9290675, rs1462368, rs518457, rs1435260, rs2063092, rs1011340, rs718387, rs9285474, rs2236876, rs2847502, rs1777689, rs706397, rs9993173, rs2363127, rs217627, rs10512122, rs725974, rs7699900, rs7038314, rs7301622, rs10495413, rs381795, rs2047058, rs969539, rs10494870, rs489381, rs28044756, rs7975512, rs9327745, rs7149147, rs11713873, rs1994952, rs523776, rs10520475, rs349187, rs618897, rs6021183, rs4815707, rs131478, rs3844253, rs10509954, rs12502036, rs1531746, rs7561808, rs163018, rs2176246, rs11248060, rs11245544, rs6864793, rs9861816, rs7130125, rs1157492, rs103294, rs7501724, rs764796, rs3809125, rs284509, rs11586379, rs153595, rs2419063, rs730566, rs2187684, rs838958, rs10825992, rs4782202, rs12500040, rs7965049, rs10484547, rs7920193, rs4686497, rs2241332, rs7689609, rs12251462, rs10458810, rs2097884, rs7598757, rs2287101

【図 9 - 5】

南ヨーロッパ人とアシェクナージ系ユダヤ人を区別する AIM:

rs7551844, rs6996185, rs1129038, rs6477998, rs1710807, rs1743789, rs6841252, rs6448770, rs11586379, rs1462368, rs8029021, rs13084044, rs47851, rs764255, rs1079195, rs10431948, rs12706799, rs8016025, rs3007711, rs12713956, rs7983897, rs793471, rs8060233, rs12441220, rs1355170, rs4756052, rs554788, rs881433, rs6687300, rs1065674, rs10493430, rs1011340, rs10825992, rs8078782, rs40459, rs920559, rs605832, rs9285474, rs4815707, rs818958, rs10460810, rs1976033, rs10506555, rs9576338, rs4832398, rs914505, rs969539, rs477627, rs1531746, rs6028505, rs224378, rs787195, rs6502048, rs4642918, rs6751522, rs1991718, rs4797909, rs11744977, rs1870590, rs2685159, rs2154328, rs7598757, rs9860730, rs163018, rs6678209, rs7108371, rs1994952, rs8040452, rs10516982, rs262216, rs618897, rs284509, rs6490700, rs17026435, rs11713873, rs6907950, rs11259206, rs10762340, rs9993173, rs7149147, rs4693933, rs1381795, rs7316723, rs359291, rs2446653, rs273477, rs82063092, rs10255965, rs4976006, rs7155872, rs730566, rs4307284, rs2377689, rs7689609, rs2411128, rs10517807, rs1435260, rs3927745, rs10458810, rs17310125, rs1660964, rs7301622, rs1163928, rs1416467, rs998401, rs313478, rs11854557, rs4131048, rs13008, rs2241332, rs4243408, rs951412, rs7920193, rs3844253, rs1173187, rs777710, rs4771165, rs132549, rs1389560, rs3856744, rs6460989, rs2252815, rs2251432, rs391871, rs2363127, rs10504275, rs7372209, rs495347, rs4770401, rs1410418, rs1118962, rs7975512, rs7038314, rs1904406, rs489381, rs7277342, rs7699900, rs10497233, rs10494870, rs9323490, rs1157492, rs1032355, rs552976, rs10495413, rs12430668, rs4040431, rs6063178, rs12645879, rs1598452, rs2285369, rs7679675, rs7561808, rs7900607, rs1024487, rs6378803, rs217627, rs4787923, rs17457687, rs1400174, rs17439723, rs1435090, rs7819806, rs90213, rs2547116, rs3764796, rs9435, rs986642, rs2589654, rs16922018, rs1352411, rs1984282, rs2287101, rs4859259, rs1107820, rs1476162, rs959763, rs6511703, rs12500040, rs6021183, rs10868791, rs7501724, rs7163907, rs10853962, rs4336881, rs959567, rs6864793, rs10520475, rs11245545, rs17799799, rs518457, rs10509384, rs2366687, rs1003306, rs2367191, rs10504924, rs10507632, rs2728945, rs1905135, rs10516096,

アイルランド人とイギリス人を区別する AIM (および他の北ヨーロッパ人集団を互いに区別する):

rs354690, rs1200826, rs3133219, rs4937688, rs17595617, rs592229, rs7632151, rs10158939, rs6673923, rs11761305, rs2824976, rs1314572, rs11735755, rs1491610, rs1697143, rs12677633, rs2331295, rs12028395, rs1035218, rs11812285, rs17487804, rs2831507, rs1119274, rs2866904, rs3211663, rs2180046, rs1791648, rs1874777, rs11169282, rs10735745, rs11629910, rs7637171, rs1975944, rs17751406, rs3133719, rs7802273, rs4682664, rs1424341, rs17495345, rs2101869, rs622917, rs11234095, rs11579554, rs2574824, rs985246, rs7172022, rs17043135, rs3934253, rs6796183, rs9400660, rs6542847, rs9537662, rs4391081, rs921656, rs12621435, rs6808082, rs969517, rs2279578, rs4741859, rs697389, rs9939430, rs300386, rs201930, rs6569343, rs6978997, rs2740259, rs2616982, rs7834280, rs17057633, rs4521229, rs3132069, rs13322103, rs4568821, rs2241466, rs1546963, rs1039621, rs11152343, rs7615026, rs11246602, rs7998134, rs1168553, rs8027997, rs7585767, rs7015718, rs4854502, rs7916977, rs11759031, rs12282752, rs924247, rs115769, rs2489772, rs2066284, rs12589358, rs2284759, rs692612, rs1420109, rs17166207, rs6855114, rs2292702, rs150881, rs213759, rs7937428, rs150139, rs3029121, rs3805331, rs9366778, rs2389197, rs2442426, rs2154254, rs1345941,

【図 9 - 6】

rs2120991, rs384267, rs10982246, rs283310, rs2832146, rs3936367, rs906671, rs11163055, rs725926, rs9353982, rs2251530, rs13032261, rs9981318, rs4561129, rs2193402, rs4854647, rs10045343, rs3937773, rs10016411, rs12618502, rs4857192, rs1706497, rs1361168, rs603089, rs7521746, rs3096700, rs1859218, rs6727787, rs3798315, rs1812458, rs12327626, rs10496354, rs6593652, rs12202888, rs12427378, rs126280, rs1496292, rs13281284, rs10197026, rs3764622, rs632853, rs1334811, rs11258614, rs1432770, rs762324, rs10506479, rs17583414, rs4889126, rs11635677, rs7538393, rs4607068, rs2148379, rs1544908, rs1604528, rs170148024, rs5761163, rs1495085, rs6727258, rs8718, rs1409590, rs13071295, rs12132550, rs4693421, rs1861532, rs10896958, rs6738275, rs9348266, rs1925391, rs42985, rs116925, rs11162802, rs4771561, rs12706465, rs13215804, rs4818059, rs201739, rs13059876, rs17745881, rs3823518, rs4903820, rs2214020, rs6992907, rs9556820, rs10249706, rs3798236, rs7701328, rs12229047, rs10846585, rs936110, rs11851852, rs9300943, rs1926150, rs782752, rs4868253, rs6586395, rs4832680, rs1979689, rs17325754, rs2803353, rs2109302, rs2840131, rs12645288, rs4241627, rs12065099, rs10511771, rs17264096, rs3026886, rs569908, rs749268, rs7591709, rs1450878, rs8015618, rs17068823, rs4669356, rs6926578, rs10095717, rs11841049, rs17735123, rs1918098, rs17158154, rs1928027, rs7685881, rs4130719, rs6742202, rs1294875, rs12319392, rs9788522, rs6061663, rs12407446, rs6705555, rs9551221, rs4688636, rs949791, rs11748141, rs10885355, rs4395073, rs6824429, rs1204020, rs2644640, rs6455468, rs10163400, rs2005104, rs7907942, rs1007042, rs6452431, rs2901100, rs6799064, rs16880382, rs1805313, rs12199314, rs1243974, rs10171287, rs352418, rs10953428, rs4974389, rs7254735, rs7257916, rs639813, rs6431565, rs1435887, rs10761660, rs10118378, rs12533837, rs2503848, rs12363937, rs7652254, rs1296896, rs17537169, rs2268933, rs6939639, rs3775479, rs1691018, rs17463958, rs6741326, rs7019271, rs4509106, rs4743564, rs13395560, rs4751890, rs8058014, rs7946208, rs2023952, rs6066892, rs153645, rs4443343, rs9297395, rs10895115, rs10419661, rs1833208, rs129074, rs1889085, rs12127377, rs32441, rs10886671, rs1363926, rs13221445, rs10501445, rs11811998, rs13225753, rs8034124, rs7906816, rs11189831, rs11147671, rs2809823, rs6538980, rs500906, rs13182055, rs7143468, rs4636972, rs10519510, rs7521399, rs278297, rs1468329, rs10515441, rs566821, rs1882396, rs11906231, rs7909464, rs12030971, rs10047997, rs17866606, rs170644, rs4129526, rs1869092, rs2149973, rs4660438, rs12404676, rs16928653, rs11138141, rs11756439, rs3925053, rs10843894, rs4259484, rs11772445, rs1598589, rs6476962, rs6676680, rs2836604, rs4843426, rs12626495, rs4621050, rs1794275, rs1801212, rs10889272, rs3847580, rs808868, rs10741808, rs2622892, rs17497293, rs7171366, rs7685182, rs12280220, rs4944000, rs6030839, rs437683, rs11000000, rs1122821, rs6767356, rs10795942, rs10917268, rs605383, rs9057881, rs10780596, rs7720751, rs1928091, rs3760440, rs12129709, rs1259866, rs17600119, rs17141252, rs10954006, rs137296, rs9623117, rs10517796, rs1758333, rs17680913, rs4321843, rs10209427, rs7123425, rs17590228, rs16957064, rs1941603, rs4403789, rs7697628, rs7682694, rs2335629, rs1405134, rs2651458, rs4678230, rs7152796, rs531539, rs1874494, rs4843507, rs17273267, rs7320534, rs10868262, rs2113574, rs4814983, rs7587228, rs17624713, rs17508376, rs1997458, rs2176313, rs251867, rs413710, rs7257183, rs1554808, rs1572228, rs7006668, rs338926, rs7135506, rs6427832

【図 9 - 7】

rs1523902, rs9721224, rs10510254, rs2718324, rs4740358, rs12440104, rs6582999, rs4257073, rs2127247, rs17006704, rs1165281, rs11642715, rs4968938, rs10235162, rs2218935, rs9353512, rs13096852, rs8037001, rs992499, rs7035812, rs12643434, rs7094594, rs9330294, rs7798936, rs1008645, rs2227286, rs11757369, rs4886026, rs243099, rs1131620, rs10279772, rs7619189, rs10792322, rs11539202, rs2145449, rs1119145, rs12570141, rs38591384, rs527507, rs11590681, rs3804639, rs7584576, rs201328, rs201430, rs10097270, rs133074, rs1964196, rs2135091, rs116885, rs4394088, rs16943989, rs11055697, rs9646474, rs16854813, rs9290227, rs3928425, rs6949364, rs7586333, rs12091966, rs6831436, rs10083961, rs10932040, rs6798997, rs861079, rs12867481, rs6881628, rs753281, rs110496342, rs3117582, rs1463242, rs2408239, rs7744432, rs564968, rs7870926, rs609305, rs4887443, rs1931902, rs11692435, rs2048161, rs4684146, rs4628026, rs1930146, rs11735972, rs1866997, rs3745509, rs12412656, rs295117, rs11686077, rs10282703, rs10869031, rs2047171, rs2616612, rs16874127, rs12249377, rs845787, rs4072374, rs2552353, rs1377554, rs400688, rs1507599, rs4728490, rs10484867, rs8545561, rs4334271, rs385773, rs4480424, rs2304572, rs11117791, rs10964808, rs9813516, rs2034781, rs11607746, rs626364, rs6536415, rs6090443, rs2328492, rs1379411, rs2583692, rs9938659, rs9315681, rs11622517, rs13359372, rs10143058, rs7590571, rs12064002, rs10835941, rs17134141, rs1343227, rs10077241, rs10132, rs2761291, rs6432594, rs10499748, rs6746287, rs7709399, rs12669805, rs11793897, rs7762729, rs12596290, rs11153162, rs6940057, rs1316883, rs186783, rs6982890, rs937816, rs7725438, rs10504180, rs8094458, rs241529, rs980274, rs855395, rs10948327, rs12149783, rs12275853, rs2155854, rs4816612, rs2669331, rs497321, rs669336, rs1333592, rs2270834, rs6076248, rs9692165, rs4776567, rs4652192, rs6942609, rs6656261, rs1339287, rs631661, rs12345503, rs10491367, rs1348338, rs2876507, rs17069574, rs6416581, rs1370031, rs10174187, rs10783028, rs7904364, rs16952027, rs12993717, rs12568609, rs550010, rs4896634, rs6574859, rs1558160, rs1772579, rs1878424, rs288057, rs12647859, rs130539, rs7934426, rs7278046, rs1886690, rs1950221, rs17747379, rs13034680, rs7790116, rs6031215, rs10016081, rs966149, rs793096, rs8044334, rs16987794, rs12156286, rs803183, rs7871735, rs187763, rs858550, rs2301570, rs1010656, rs1815619, rs2184925, rs10521057, rs11184219, rs2726675, rs10485678, rs7782814, rs489466, rs1093309, rs12439063, rs2396081, rs4887111, rs6979324, rs4233628, rs2189864, rs1267472, rs7006436, rs2839475, rs6874663, rs952579, rs10794696, rs11634215, rs1256992, rs6748280, rs8024685, rs1477798, rs8109759, rs6076954, rs17034929, rs4915886, rs2560301, rs6700849, rs1368406, rs2724031, rs2292719, rs1553664, rs3935801, rs4148298, rs4968844, rs4459724, rs1423089, rs12021671, rs16839940, rs8021670, rs4948235, rs13200910, rs308793, rs10924440, rs7329258, rs4780267, rs282152, rs10824914, rs16907003, rs7153598, rs3010888, rs6957465, rs12441839, rs10192315, rs727637, rs11786174, rs1488214, rs11635749, rs2729910, rs4869578, rs9459128, rs212434, rs1863471, rs12372169, rs724665, rs2908201, rs6010652, rs807704, rs6695712, rs1282557, rs10496954, rs2295682, rs9820762, rs12198061, rs11698898, rs2139747, rs700278, rs2192883, rs6312, rs1486731, rs2684796, rs12795982, rs2839589, rs2184221, rs9640302, rs4746826, rs632547, rs7764278, rs6651164, rs16750, rs13020203, rs9905820, rs1951797, rs10969030, rs10002315, rs4920104, rs7257017, rs9514044, rs2872338, rs1750111, rs1512155, rs4785173, rs2151842, rs2218724, rs1822917, rs1688015, rs11221731, rs887106, rs7809872, rs4532570, rs12554199, rs2773207, rs12314080, rs11881242, rs7225060, rs2518201, rs1010779, rs8048863, rs9299261, rs4586385, rs12642133, rs10777342, rs16901689, rs1859569, rs9300717, rs1154278, rs1174549, rs7140955, rs10108974, rs7852159, rs9419387, rs10831905,

【図 9 - 9】

rs17279118, rs373531, rs12372697, rs10357, rs2822548, rs1387910, rs4859864, rs17319408, rs1265203, rs12904370, rs10518627, rs10050980, rs3782814, rs17820124, rs475357, rs1795065, rs4414922, rs7487814, rs13194998, rs7807708, rs17660328, rs4471501, rs8077267, rs4304313, rs2387673, rs3846677, rs17012485, rs4234474, rs11731647, rs6712681, rs2058131, rs2270409, rs6089784, rs17698283, rs6494039, rs2937785, rs260445, rs9641312, rs4911442, rs6881004, rs905570, rs99521, rs12997609, rs504341, rs1748168, rs11196967, rs1791468, rs17455729, rs1393861, rs4623797, rs4819591, rs1009728, rs17019119, rs4943731, rs1036295, rs2793471, rs9283119, rs4248914, rs733044,

スペイン人を白色人種と区別する AIM:

rs118003, rs166673, rs1699088, rs261056, rs280181, rs286804, rs347413, rs374672, rs382515, rs382558, rs455249, rs787363, rs600558, rs640330, rs688858, rs707077, rs717251, rs717471, rs717614, rs718996, rs719628, rs719750, rs720030, rs720229, rs721128, rs721246, rs721288, rs721361, rs721362, rs721363, rs724130, rs724658, rs726512, rs726610, rs727293, rs728506, rs741933, rs754291, rs831451, rs869538, rs894198, rs950132, rs951572, rs951784, rs952020, rs952562, rs953111, rs955770, rs955771, rs956193, rs956377, rs957862, rs958985, rs959763, rs961463, rs965892, rs965917, rs967294, rs967906, rs979050, rs979322, rs988462, rs1016120, rs1023534, rs1032143, rs1033042, rs1035701, rs1073354, rs1074149, rs1074716, rs1074805, rs1075263, rs1075906, rs1108929, rs1113082, rs1113144, rs1114000, rs1116138, rs1116853, rs1158619, rs1202645, rs1210110, rs1331819, rs1343809, rs1359247, rs1363234, rs1363235, rs1366233, rs1368387, rs1368697, rs1373904, rs735575, rs1375989, rs1376392, rs1377470, rs1378702, rs1380415, rs1382859, rs1384118, rs1388175, rs1390476, rs1391961, rs1393736, rs1395162, rs1395604, rs1399664, rs1400347, rs1402458, rs1404759, rs1405633, rs1410592, rs1411445, rs1412171, rs1414149, rs1417503, rs1418706, rs1419702, rs1420117, rs1449050, rs1451371, rs1454027, rs1478810, rs1479285, rs1485254, rs1498458, rs1501626, rs1537371, rs1537523, rs1546561, rs1548299, rs1549944, rs1572902, rs1579029, rs1585217, rs1585946, rs1587120, rs1599710, rs1603681, rs1605727, rs1820072, rs1824683, rs1830876, rs1854673, rs1900672, rs1917152, rs1917199, rs1927014, rs1929646, rs913621, rs1951095, rs1951207, rs1951507, rs1951519, rs1961491, rs1986508, rs1990542, rs1992539, rs1999333, rs2008163, rs2008927, rs2010392, rs2011812, rs2014303, rs2014533, rs2015963, rs2017892, rs2017903, rs2028021, rs2040847, rs2040848, rs2057127, rs2067084, rs2078426, rs2134897, rs2139246, rs2194127, rs2196302, rs2210254, rs2216629, rs2226672, rs2351002, rs2355205, rs2358295, rs2366101, rs2366977, rs2369860, rs2373584, rs2374545, rs2389774, rs2409253, rs2418400, rs2419925, rs2865116, rs2868932, rs2886937, rs2887280, rs2898975

南、中央、および東アジア人を区別する AIM:

rs32826, rs63319, rs170872, rs187145, rs187861, rs191644, rs205474, rs219562, rs235946, rs266175, rs266177, rs303450, rs305851, rs353599, rs355191, rs366220, rs409914, rs411301, rs428294, rs498046, rs632159, rs664574, rs714857, rs717205, rs717757, rs718137, rs719006, rs719866, rs719910, rs720376, rs720578, rs720844, rs721002, rs721407, rs721684, rs722559, rs722653, rs722711, rs723195, rs723212, rs723686, rs723698, rs723723, rs723794, rs724246, rs724247, rs724258, rs724259, rs725389, rs726108, rs726217, rs726472, rs726493, rs726692, rs727424, rs727708, rs728497, rs728618, rs728647, rs730991, rs791856, rs864736, rs896019, rs910552, rs923976, rs950198, rs950303, rs950400, rs951062, rs951199, rs951998, rs952503,

【図 9 - 8】

rs10498377, rs269850, rs9454096, rs6967514, rs13132569, rs12327342, rs11799558, rs1325156, rs2844651, rs2556097, rs2891409, rs12538916, rs8005808, rs2456220, rs6859290, rs924884, rs12501016, rs11644921, rs907238, rs2331175, rs2800257, rs717966, rs10825066, rs7112734, rs1524600, rs10491030, rs1387329, rs7996510, rs17461905, rs11839068, rs9312336, rs7204900, rs1672376, rs12677340, rs9575619, rs2765584, rs205649, rs2839349, rs1454149, rs4907404, rs2166488, rs1230313, rs908822, rs2402800, rs913984, rs17651643, rs10979066, rs3024536, rs17173637, rs567384, rs9562077, rs10942957, rs4658702, rs7785682, rs2392147, rs613479, rs12710777, rs2791163, rs12094260, rs483241, rs12083651, rs12279152, rs7971883, rs17527048, rs2231895, rs512458, rs7321658, rs9923061, rs2298608, rs865151, rs11754258, rs207274, rs17567580, rs8021112, rs6506515, rs1553255, rs2492501, rs3781578, rs10780514, rs11695049, rs2022725, rs6959225, rs12796185, rs1369256, rs10063082, rs11966566, rs4427879, rs2395653, rs12018676, rs10154899, rs135551, rs9538094, rs10518452, rs17482565, rs2761384, rs10496557, rs7707491, rs943068, rs4713949, rs4233167, rs767249, rs12191488, rs6855792, rs1541335, rs2868146, rs6564940, rs1497305, rs1465010, rs11603501, rs10230087, rs899069, rs7864459, rs10517677, rs4708459, rs1545255, rs11634465, rs2284944, rs894392, rs10021708, rs7748977, rs9385824, rs7116354, rs11642445, rs3806300, rs4705295, rs2303802, rs791656, rs4387618, rs16972092, rs4891825, rs11712075, rs1978368, rs16858789, rs1446394, rs2060530, rs157237, rs1482363, rs9502971, rs3757791, rs11768706, rs10888101, rs16936082, rs6585474, rs3857176, rs9510565, rs7787246, rs10954638, rs10968042, rs2700230, rs7652486, rs11602256, rs12492695, rs1197701, rs4918844, rs12646895, rs12377016, rs11995144, rs227906, rs4753339, rs7673403, rs7855536, rs760560, rs913259, rs9442235, rs11940490, rs3742076, rs9407554, rs7321548, rs1523250, rs5751691, rs2691344, rs6089457, rs849794, rs897111, rs9613630, rs1432053, rs2616667, rs10159191, rs12619236, rs7660552, rs12731666, rs2841455, rs6454198, rs1466835, rs637001, rs7594872, rs470490, rs1071104, rs10403583, rs7526907, rs10484854, rs1355829, rs12273605, rs691127, rs1930718, rs1543603, rs4936578, rs722113, rs3846456, rs3934711, rs11101312, rs10068902, rs11988550, rs4331742, rs12923978, rs382962, rs12247354, rs10923142, rs9912143, rs2468638, rs11543651, rs16997896, rs2827845, rs10112398, rs1412471, rs1012769, rs153785, rs4368537, rs7820789, rs3813135, rs1554690, rs4890891, rs2114088, rs2327506, rs445275, rs6842443, rs9423942, rs11593576, rs244050, rs1007018, rs2192754, rs945249, rs17048973, rs17088268, rs6105221, rs2798950, rs9855729, rs4808416, rs10883479, rs6729229, rs875569, rs2475229, rs2047153, rs10493770, rs10128640, rs7154324, rs17780310, rs1944362, rs3793504, rs12452093, rs4283980, rs20707713, rs1118045, rs12585880, rs7598637, rs10887047, rs915538, rs17293152, rs904269, rs2987380, rs717751, rs6765166, rs10486461, rs7858808, rs6599127, rs11136596, rs16943315, rs7336525, rs9366950, rs6073394, rs1908984, rs2280699, rs2238508, rs7460771, rs7502772, rs7038373, rs1020042, rs1791419, rs11780780, rs11111689, rs743082, rs7242960, rs10500956, rs10431150, rs1698803, rs12743521, rs11778372, rs17271148, rs1360646, rs8066468, rs11831940, rs2166487, rs7018906, rs11180273, rs544603, rs4786780, rs7658820, rs13068538, rs7963499, rs4902562, rs10515998, rs4331170, rs204538, rs666649, rs12691565, rs16876243, rs10504268, rs4397025, rs6793927, rs2735202, rs11172147, rs11632449, rs10733358, rs2048816, rs10942889, rs1475048, rs7698580, rs12504106, rs7416953, rs12111723, rs7075768, rs11854173, rs10520656, rs616582, rs13295839, rs824453, rs4596664, rs9289539, rs10275488, rs4981674, rs4955713, rs2387698, rs17277113, rs250413, rs17590785, rs10462465, rs830310, rs7814885, rs10733333, rs6772196, rs1342744, rs10168477, rs10273, rs7221778, rs7187133, rs17827858,

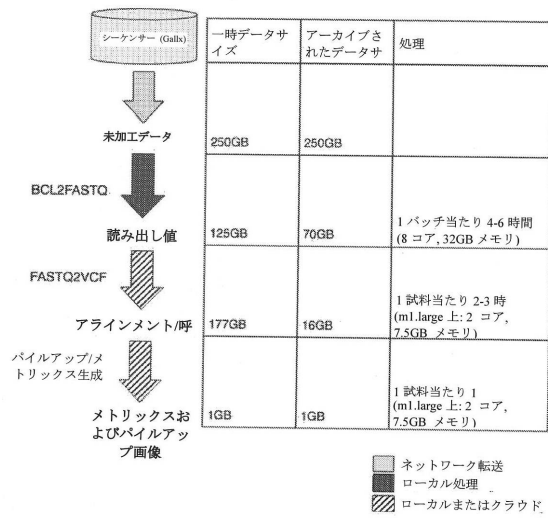
【図 9 - 10】

rs952785, rs952976, rs952993, rs953624, rs953680, rs956070, rs956408, rs958478, rs958976, rs963096, rs963170, rs963171, rs966134, rs966285, rs967990, rs978486, rs978488, rs985933, rs985934, rs988913, rs1000313, rs1004246, rs1011873, rs1013063, rs1018616, rs1029554, rs103076, rs1073109, rs1073119, rs1073985, rs1114000, rs1178328, rs1318710, rs1336672, rs1340846, rs1343981, rs1352514, rs1356763, rs1359932, rs1362987, rs1366627, rs1371063, rs1374499, rs1381331, rs1381972, rs1383965, rs1383970, rs1391221, rs1391961, rs1392096, rs1404501, rs1405567, rs1406012, rs1406218, rs1407043, rs1408830, rs1413724, rs1416615, rs1417503, rs1418102, rs1422004, rs1442232, rs1422895, rs1427356, rs1428702, rs1434193, rs1446966, rs1448229, rs1470248, rs1480951, rs1509421, rs1532517, rs1549690, rs1563353, rs1565138, rs1565534, rs1570699, rs1578243, rs1583355, rs1586149, rs1586597, rs1812624, rs1817186, rs1822811, rs1822812, rs1822813, rs1823100, rs1826873, rs1876545, rs1892114, rs1903189, rs1903746, rs1924257, rs1928045, rs1934008, rs1940006, rs1944086, rs1945146, rs1950284, rs1954119, rs1985835, rs2007397, rs2008011, rs2008686, rs2013708, rs2035799, rs2035801, rs2050137, rs20552291, rs2065443, rs2077811, rs2010551, rs2134897, rs2138212, rs2145231, rs2192874, rs2327790, rs2369898, rs2370672, rs2382193, rs2392002, rs2406519, rs2421172, rs2877455

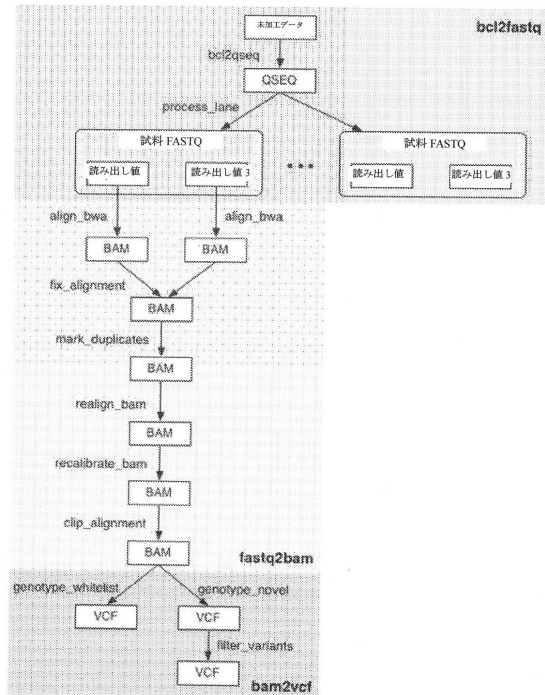
中国人集団と日本人集団を区別する AIM:

rs466256, rs1024052, rs6462942, rs6685064, rs244757, rs3924094, rs11087655, rs944664, rs11087655, rs2402820, rs2280100, rs11118295, rs721953, rs2377527, rs296364, rs10894532, rs2842186, rs6462942, rs4820389, rs4632283, rs4838113, rs16966142, rs1500127, rs4925258, rs1118295, rs2280100, rs766144, rs6462942, rs766144, rs2241880, rs1159114, rs4962061, rs5924101, rs7900633, rs11087655, rs715846, rs6817562, rs2185369, rs331332, rs4937787, rs2241880, rs12484697, rs4669614, rs11087655, rs2377527, rs5924101, rs4838113, rs2838665, rs12493799, rs12505641, rs3019711, rs5924101, rs1286041, rs4669614, rs614502, rs2280100, rs4362557, rs5924766, rs6937974, rs782758, rs10488533, rs614502, rs2248000, rs7003901, rs4362557, rs1858234, rs2724863, rs1748944, rs1077834, rs1259807, rs12576800, rs2970869, rs9308712, rs6855567, rs6457300, rs331332, rs2293072, rs4829981, rs1077834, rs4947803, rs7016174, rs7991928, rs2271738, rs9397264, rs4849159, rs1554656, rs1383887, rs12660417, rs1077834, rs2358521, rs1383887, rs12025272, rs2763002, rs9397264, rs4669614, rs7660417, rs2114776, rs4749159, rs1383887, rs452172, rs3805466, rs7016174, rs7991928, rs9397264, rs16993741, rs7016174, rs1259807

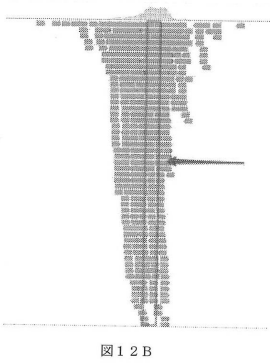
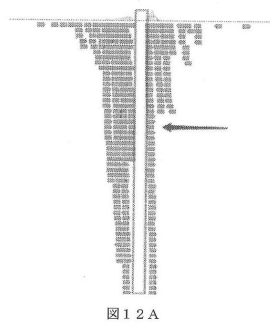
【図 10】



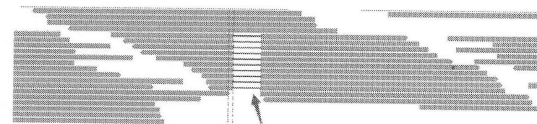
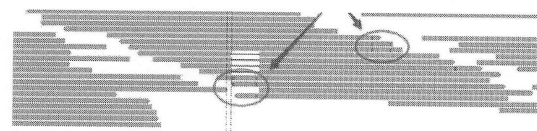
【図 11】



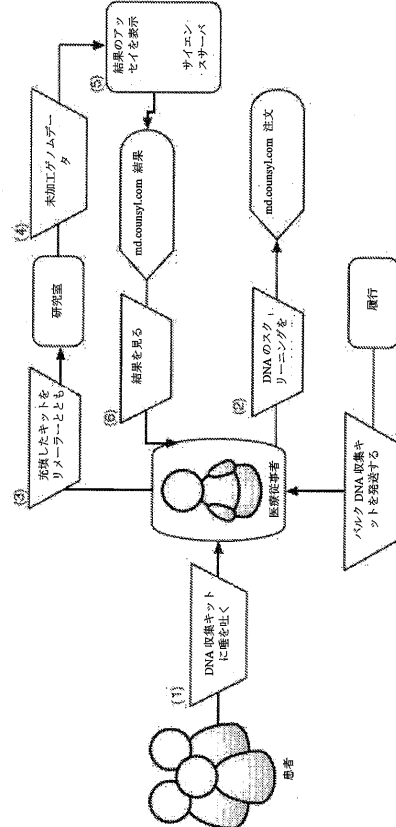
【図 12】



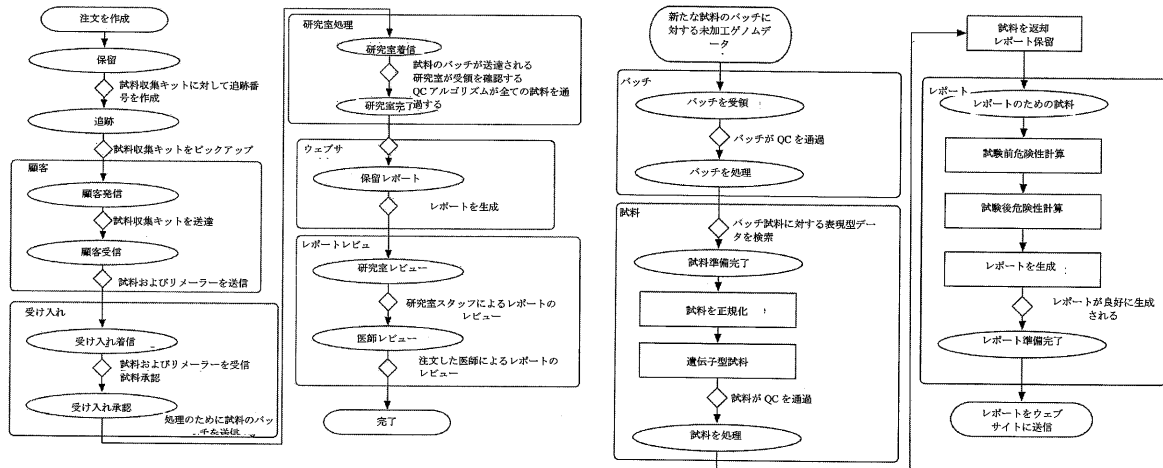
【図 13】



【 図 1 5 】



【 図 1 7 】



【 図 1 8 】

表現型質問票および家族歴

母親 Hem のバックグラウンド	性別	年齢	身長	体重	家族歴
イギリス人	男	1 (4)	170	65	心臓病
生年月日	男	1 (4)	170	65	心臓病
身長	男	1 (4)	170	65	心臓病
体重	男	1 (4)	170	65	心臓病
家族歴	男	1 (4)	170	65	心臓病
イギリス人	男	1 (4)	170	65	心臓病
生年月日	男	1 (4)	170	65	心臓病
身長	男	1 (4)	170	65	心臓病
体重	男	1 (4)	170	65	心臓病
家族歴	男	1 (4)	170	65	心臓病

表現型質問票および家族歴

【 図 1 9 】

```
-----[PROBE_WINDOW][RECESS][ROI_TILE][RECESS][PROBE_WINDOW]-----
-----[...PROBE...]-----[...PROBE...]-----
```

フロントページの続き

(31)優先権主張番号 13/551,587

(32)優先日 平成24年7月17日(2012.7.17)

(33)優先権主張国 米国(US)

(31)優先権主張番号 13/665,671

(32)優先日 平成24年10月31日(2012.10.31)

(33)優先権主張国 米国(US)

(72)発明者 リチャーズ, ハンター

アメリカ合衆国カリフォルニア州94131, サンフランシスコ, セザー・チャベス・ストリート
4140, アパートメント 10

(72)発明者 エヴァンス, エリック

アメリカ合衆国カリフォルニア州94066, サン・ブルーノ, ナショナル・アベニュー 105
1, アパートメント 151

(72)発明者 スリニヴァサン, パラジ

アメリカ合衆国カリフォルニア州94080, サウス・サンフランシスコ, キンボール・ウェイ
180

(72)発明者 スリニヴァサン, スブラマニウム

アメリカ合衆国ニューヨーク州11803, プレインビュー, リッチフィールド・ストリート 6
4

(72)発明者 シャー, アビク

アメリカ合衆国カリフォルニア州94117, サンフランシスコ, ウェブスター・ストリート 4
41

(72)発明者 パターソン, エイ・スコット

アメリカ合衆国カリフォルニア州94131, サンフランシスコ, レイドリー・ストリート 54
6

(72)発明者 チュウ, クレメント

アメリカ合衆国カリフォルニア州94107, サンフランシスコ, テネシー・ストリート 112
8, アパートメント 1

審査官 中村 勇介

(56)参考文献 国際公開第2012/040387(WO, A1)

特表2003-521252(JP, A)

国際公開第2012/027572(WO, A1)

(58)調査した分野(Int.Cl., DB名)

C12Q1/00-3/00

JSTPlus/JMEDPlus/JST7580(JDreamIII)