(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2010-522571
(P2010-522571A)
(43)公表日 平成22年7月8日(2010.7.8)

(51)Int.Cl.		F I		テーマコード(参考)
C12N 15/09	(2006.01)	C12N 15/00	A	4B024
C12Q 1/68	(2006.01)	C12Q 1/68	A	4B063

審査請求 未請求　予備審査請求 未請求　(全 29 頁)

(21)出願番号	特願2010-501272(P2010-501272)
(86)(22)出願日	平成20年3月28日(2008.3.28)
(85)翻訳文提出日	平成21年11月24日(2009.11.24)
(86)国際出願番号	PCT/US2008/058795
(87)国際公開番号	WO2008/119084
(87)国際公開日	平成20年10月2日(2008.10.2)
(31)優先権主張番号	60/908,606
(32)優先日	平成19年3月28日(2007.3.28)
(33)優先権主張国	米国(US)
(31)優先権主張番号	60/940,321
(32)優先日	平成19年5月25日(2007.5.25)
(33)優先権主張国	米国(US)

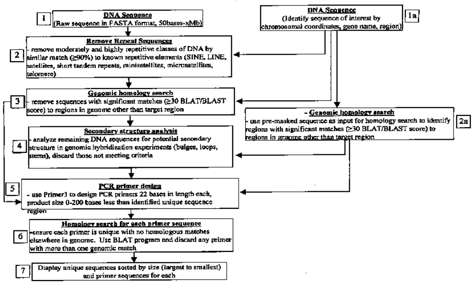
(71)出願人 502416567
ザ　チルドレンズ　マーシー　ホスピタル
アメリカ合衆国　ミズーリ州　６４１０８
カンザス　シティ　ギルハム　ロード
２４０１
(74)代理人 100080159
弁理士　渡辺　望稔
(74)代理人 100090217
弁理士　三和　晴子
(72)発明者 ニューカーク　ヘザー
アメリカ合衆国　６４１０８　ミズーリ州
カンザス　シティ　ギルハム　ロード
２４０１

最終頁に続く

(54)【発明の名称】少コピー数の核酸セグメントを同定および選択する方法

(57)【要約】
本発明は、既知の核酸配列内から少コピー数の核酸セグメントを同定し、同定した少コピー数セグメントの中からハイブリダイゼーション実験に熱力学的に好適に用いられるセグメントを選択する方法に係わる。

FIG. 6



【特許請求の範囲】
【請求項１】
少コピー数の核酸セグメントを同定する方法であって、
（ａ）対象ゲノム領域から高頻度および中頻度繰り返し配列を除去し、非繰り返しゲノムセグメントを表示するステップ、
（ｂ）非繰り返しゲノムセグメントを、対象領域以外のゲノム領域とのホモロジーに関して検索し、対象としないゲノム領域と相同なセグメントを総て廃棄するステップ、
（ｃ）非繰り返しゲノムセグメントにおいて予測される二次構造モチーフを同定するステップ、および
（ｄ）ステップａ、ｂおよびｃのうちの少なくとも一つによって同定した非繰り返しセグメントからプローブを設計し、このプローブを対象ゲノム領域および対象としないゲノム領域と比較してユニークさに関し解析するステップ
の中から二つ以上のステップを含む方法。
【請求項２】
ステップａ～ｄの中から少なくとも三つのステップを含む、請求項１に記載の方法。
【請求項３】
前記ステップａの非繰り返しゲノムセグメントが１ｋｂより大きいサイズを有する、請求項１に記載の方法。
【請求項４】
ステップｃが熱力学的解析によって行なわれる、請求項１に記載の方法。
【請求項５】
実施した方法によって得られるゲノムセグメントに対応するＰＣＲプライマーを設計するステップをさらに含む、請求項１に記載の方法。
【請求項６】
前記ＰＣＲプライマーがユニーク配列しか含まないことを確実にするステップをさらに含む、請求項５に記載の方法。
【請求項７】
ハイブリダイゼーション実験に用いられるプローブを選択する方法であって、
（ａ）対象配列から繰り返し配列を除去して配列セグメントを用意するステップ、
（ｂ）前記配列セグメントを、対象配列を含む領域以外のゲノム領域とそれぞれ比較して、前記ゲノム領域のいずれかの箇所とマッチする前記セグメントを総て廃棄し、残ったユニーク配列を保存するステップ、
（ｃ）前記ユニーク配列を予測される二次構造モチーフに関して評価するステップ、および
（ｄ）予測される二次構造モチーフを有しない前記ユニーク配列に基づきプローブを選択するステップ
を含む方法。
【請求項８】
前記プローブのためのＰＣＲプライマーを設計するステップをさらに含む、請求項７に記載の方法。
【請求項９】
前記ＰＣＲプライマーがゲノムの他の箇所とマッチしないことを確実にするステップをさらに含む、請求項８に記載の方法。
【請求項１０】
ステップ（ｃ）が熱力学的解析によって行なわれる、請求項７に記載の方法。
【請求項１１】
前記熱力学的解析がギブズ自由エネルギーの式に基づき、そのときギブズ自由エネルギーが０と５０との間である、請求項１０に記載の方法。
【請求項１２】
ギブズ自由エネルギーの式においてΔＨ＜－１０００、ΔＳ＜－３５００、およびＴｍ

≧３７℃である、請求項１１に記載の方法。

【請求項１３】

Ｔｍが４２℃以上である請求項１２に記載の方法。

【請求項１４】

Ｔｍが６０℃以上である請求項１２に記載の方法。

【請求項１５】

配列番号１～５７からなる群から選択された核酸配列。

【発明の詳細な説明】

【技術分野】

【０００１】

［関連出願］

本出願は、２００７年3月２８日付出願の米国仮特許出願第６０／９０８，６０６号および２００７年5月２５日付出願の米国仮特許出願第６０／９４０，３２１号に関連し、これらの優先権を主張する。これらの仮出願はいずれもその全体が本明細書に参考として含まれる。

上記出願は総て所有者が共通する。

【０００２】

［配列表］

１　本出願は、米国特許規則§１．８２１～１．８２５に則り、かつＥＦＳ－Ｗｅｂ要件に合致して電子形式で提出された配列表を含む。この配列表はその全体が本明細書に参考として含まれる。

【０００３】

本発明は、既知の核酸配列内からハイブリダイゼーション実験に好適に用いられる少コピー数の核酸セグメントを同定する方法に係わる。さらに本発明は、同定した少コピー数の核酸セグメントの中からハイブリダイゼーション実験に熱力学的に好適に用いられるセグメントを優先的に選択する方法に係わる。

【背景技術】

【０００４】

少コピー数プローブを核酸配列上の標的相同セグメントに対して用いることは当該技術分野で公知である。幾つかの従来方法は、標的配列セグメントを繰り返し配列のデータベースに照らしてスキャンすることに依拠し、プローブ配列は隣接する２個の繰り返し配列間に位置するものとして同定された。しかし、それらの方法の信頼性は繰り返し配列のデータベースの品質によって左右された。その上、それらの方法で同定されたプローブ配列の中には、例えば二次構造コンホメーション（ヘアピンループ、ステム、バルジなど）に起因してハイブリダイゼーションに好適でないものも有った。プローブとして用いられる少コピー数核酸配列を同定するその他の方法は、通例多くのステップにおいて知識の豊富な研究者による膨大な検討および解析を必要とする面倒なプロセスを伴った。

【０００５】

ユニーク配列領域の同定に通常用いられるコンピュータ利用の方法は、ＲｅｐｅａｔＭａｓｋｅｒ（ワールドワイドウェブ上の、該当部分に「ｒｅｐｅａｔｍａｓｋｅｒ．ｏｒｇ」と記載されたウェブサイトにて公開）およびＢＬＡＴ（ワールドワイドウェブ上の、該当部分に「ｇｅｎｏｍｅ．ｕｃｓｃ．ｅｄｕ」と記載されたウェブサイトにて公開）などのウェブベースプログラムを含む。これらのプログラムはいずれも、ゲノム配列をゲノム領域の熱力学特性に関して評価するものではない。従って、これらのプログラムから抽出されるプローブはユニーク配列を含み得るが、そのような配列はハイブリダイゼーションに好適でない恐れが有る。現在、そのような配列がハイブリダイゼーションに好適であるか判断するには、当該配列を物理的にプローブまたはプライマーとしなければならず、これには通常時間と経費が掛かる。

【０００６】

潜在的プローブ配列の熱力学特性の評価に用いられるコンピュータ利用の方法は、始め

に配列を同定することができない。例えば、ゲノム配列の熱力学的評価に通常用いられるプログラムであるＭｆｏｌｄ（ワールドワイドウェブ上の、該当部分に「ｂｉｏｉｎｆｏ．ｒｐｉ．ｅｄｕ」と記載されたウェブサイトにて公開）はゲノム配列を、ユニーク配列としての性質に関しては評価しない。従って、同定された、熱力学的に安定な配列がユニークなものかは試験してみるまでユーザーには判然としない。プローブの試験には時間も費用も掛かるので、遺伝子セグメント内で熱力学的に安定なユニーク配列を同定する、より確実な方法を見いだすことが望ましい。

【０００７】

以下の文献の教示および内容は総て特定的に本明細書に参考として含まれる。

【先行技術文献】

【特許文献】

【０００８】

【特許文献１】米国特許第７，０１４，９９７号、「シングルコピープローブを用いての染色体構造異常の位置特定」、ＲｏｇａｎおよびＫｎｏｌｌ、２００６年。

【特許文献２】米国特許第７，０１３，２２１号、「反復型プローブ設計、およびフレキシブルｉｎ　ｓｉｔｕ合成アレイを用いての詳細発現プロファイリング」、Ｆｒｉｅｎｄら、２００６年。

【特許文献３】米国特許第７，１１５，７０９号、「高ｃｏｍｐｌｅｘｉｔｙ核酸プローブを用いる標的染色体ＤＮＡ染色方法」、Ｇｒａｙら、２００６年。

【特許文献４】米国特許第６，８２８，０９７号、「シングルコピーゲノムハイブリダイゼーションプローブとその生成方法」、ＲｏｇａｎおよびＫｎｏｌｌ、２００４年。

【特許文献５】米国特許第６，２４２，１８４号、「シングルコピーおよびマルチコピー核酸配列のｉｎ　ｓｉｔｕハイブリダイゼーション」、Ｓｉｎｇｅｒら、２００１年。

【非特許文献】

【０００９】

【非特許文献１】Ａｎｄｒｅｓｓｏｎ，Ｒ．、Ｒｅｐｐｏ，Ｅ．、Ｋａｐｌｉｎｋｓｋｉ，Ｌ．、Ｒｅｍｍ，Ｍ．「ユニークゲノムＰＣＲプライマー設計用ＧＥＮＯＥＭＡＳＫＥＲパッケージ」、ＢＭＣ　Ｂｉｏｉｎｆｏｒｍａｔｉｃｓ、２００６年、２７（７）：１７２。

【非特許文献２】Ｋｎｏｌｌ，Ｊ．Ｈ．Ｍ．およびＲｏｇａｎ，Ｐ．Ｋ．「高分解能で行なう配列ベースの染色体異常ｉｎ　ｓｉｔｕ検出」、Ａｍｅｒｉｃａｎ　Ｊｏｕｒｎａｌ　ｏｆ　Ｍｅｄｉｃａｌ　Ｇｅｎｅｔｉｃｓ、２００３年、１２１Ａ：２４５－２５７。

【非特許文献３】Ｍｉｕｒａ，Ｆ．、Ｕｅｍａｔｓｕ，Ｃ．、Ｓａｋａｋｉ，Ｙ．、Ｉｔｏ，Ｔ．「３’末端サブ配列の安定性およびユニークさに基づく新規な高特異的ＰＣＲプライマー設計方法」、Ｂｉｏｉｎｆｏｒｍａｔｉｃｓ、２００５年、２１（２４）：４３６３－７０。

【非特許文献４】Ｎｅｗｋｉｒｋ，Ｈ．、Ｋｎｏｌｌ，Ｊ．Ｈ．Ｍ．、Ｒｏｇａｎ，Ｐ．（２００５年）「Ｃｏｔ－１　ＤＮＡによる定量的ゲノムおよび発現ハイブリダイゼーションの歪み：　該作用の緩和」、Ｎｕｃｌｅｉｃ　Ａｃｉｄｓ　Ｒｅｓｅａｒｃｈ、３３：ｅ１９１。

【非特許文献５】Ｎｅｗｋｉｒｋ，Ｈ．、Ｍｉｒａｌｌｅｓ，Ｍ．、Ｒｏｇａｎ，Ｐ．、Ｋｎｏｌｌ，Ｊ．Ｈ．Ｍ．（２００６年）「定量的マイクロスフェアハイブリダイゼーションによるゲノムコピー数測定」、Ｈｕｍａｎ　Ｍｕｔａｔｉｏｎ、２７：３７６－３８６。

【非特許文献６】Ｒｏｇａｎ，Ｐ．Ｋ．、Ｃａｚｃａｒｒｏ，Ｐ．Ｍ．、Ｋｎｏｌｌ，Ｊ．Ｈ．「蛍光ｉｎ　ｓｉｔｕハイブリダイゼーション用シングルコピーゲノムＤＮＡプローブの配列ベース設計」、Ｇｅｎｏｍｅ　Ｒｅｓｅａｒｃｈ、２００１年、１１（６）：１０８６－９４。

【非特許文献７】Ｒｏｚｅｎ，Ｓ．、Ｓｋａｌｅｔｓｋｙ，Ｈ．Ｊ．：　一般ユーザーおよび生物学者プログラマーのための、ＷＷＷ上のＰｒｉｍｅｒ３。　Ｋｒａｗｅｔｚ，Ｓ

．、Ｍｉｓｅｎｅｒ，Ｓ．（編集）「バイオインフォマティクスの方法とプロトコル：分子生物学における手法」、Ｈｕｍａｎａ　Ｐｒｅｓｓ、Ｔｏｔｏｗａ、ＮＪ、３６５−３８６（２０００年）所載。

【非特許文献８】Ｔａｔｕｓｏｖａ，Ｔ．Ａ．およびＭａｄｄｅｎ，Ｔ．Ｌ．「Ｂｌａｓｔ　２　ｓｅｑｕｅｎｃｅｓ−タンパク質およびヌクレオチド配列を比較する新しいツール」、ＦＥＭＳ　Ｍｉｃｒｏｂｉｏｌ．　Ｌｅｔｔ．、１９９９年、１７４：２４７−２５０。

【非特許文献９】Ｚｕｋｅｒ，Ｍ．「核酸フォールディングおよびハイブリダイゼーション予測のためのｍｆｏｌｄウェブサーバー」、Ｎｕｃｌｅｉｃ　Ａｃｉｄｓ　Ｒｅｓ．、３１：３４０６−３４１５（２００３年）。

【非特許文献１０】ＲｅｐｅａｔＭａｓｋｅｒ：　Ｓｍｉｔ，Ａ．Ｆ．Ａ．、Ｈｕｂｌｅｙ，Ｒ．、Ｇｒｅｅｎ，Ｐ．、未出版。現行バージョン：　ｏｐｅｎ−３．１．６

【非特許文献１１】ＢＬＡＴ：　該当部分に「ｇｅｎｏｍｅ．ｕｃｓｃ．ｅｄｕ」と記載されたアドレスを有する、ワールドワイドウェブ上のＵＣＳＣゲノムブラウザウェブサイト。

【発明の概要】

【発明が解決しようとする課題】

【００１０】

従って、当該技術分野では、既知の核酸配列からハイブリダイゼーションに好適な少コピー数の核酸セグメントを迅速かつ確実に同定する方法が必要とされている。また、長い既知核酸配列から、熱力学的にハイブリダイゼーションに好適な少コピー数の核酸セグメントを迅速に同定する方法も必要とされている。

【課題を解決するための手段】

【００１１】

本発明は、既知の核酸配列内から少コピー数の核酸セグメントを迅速かつ確実に同定する、また同定した少コピー数セグメントの中からハイブリダイゼーション実験に熱力学的に好適に用いられるセグメントを選択する方法およびコンピュータ化プロセスを提供することによって、従来技術の免れ得なかった問題点を克服し、当該技術分野に明確な進歩をもたらす。

【００１２】

本発明は、ハイブリダイゼーションの感度およびスループットを有利に改善する。本発明の方法によれば、ユーザーは他のゲノミクスプログラムを用いた時よりも長い配列を一度に解析することができ、その際任意の長さの配列を解析することがなお可能である。長い配列とは、その長さが１００キロベース（ｋｂ）、１５０ｋｂ、２００ｋｂ、２５０ｋｂ、３００ｋｂ、５００ｋｂを越え、さらには１０００ｋｂ以上にも達する配列であり得る。加えて、本発明の方法が用いるパラメーターはウェブベースプログラムで通常用いられるものより厳密である。ギブズ自由エネルギーの式に基づくΔＧ（ギブズ自由エネルギー）、ΔＨ（エンタルピー）、ΔＳ（エントロピー）およびＴｍ（融点）を含めたそれら厳密な基準によって、ゲノミクス実験用のユニーク配列プローブのみを高効率で選択することが可能となる。ギブズ自由エネルギーの式は方程式であり、変数ΔＨ、ΔＳおよびＴｍは、好ましい形態において＜５０である所望のΔＧが得られるように操作され得ると理解される。上記変数のうちの一つ以上の操作が好ましい範囲を逸脱し、それでもなおΔＧ＜５０となる場合、そのような基準もしくはパラメーターも本発明の範囲内である。好ましい形態では、基準もしくはパラメーターはΔＧ＜５０、ΔＨ＜−１０００、ΔＳ＜−３５００およびＴｍ≧６０℃であることが求められる。ＱＭＨの場合はこれらが最も好ましい基準もしくはパラメーターであるが、ＦＩＳＨの場合は最も好ましいＴｍが≧４２℃となり、アレイベースの技術では最も好ましいＴｍは≧３７℃となる。

【００１３】

本発明の方法は現行技術と比較してより包括的である。なぜなら、該方法は配列解析を熱力学的解析と組み合わせ、それによって少コピー数配列である（すなわち繰り返し配列

でなく、好ましくはゲノム中に１回しか現れないシングルコピーである）と同時にハイブリダイゼーションの際に熱力学的に安定である核酸セグメントを同定するからである。また、本発明の方法はユニーク配列を同定し、ゲノムを検索して他の非繰り返しゲノム領域が対象領域と相同でないことを確実にする。さらに、従来技術と異なり、本発明の方法では少コピー数核酸セグメントの解析がダブルチェック式に行なわれ、当該セグメントをポリメラーゼ連鎖反応（ＰＣＲ）法用のプライマーとして、あるいはまた可変温度に依拠する他の方法において好適に用いることができるかが判断される。

【発明の効果】

【００１４】

本発明は、その諸特性がユーザー定義可能である、異なる長さの様々な少コピー数核酸プローブの設計に用いることができる点で汎用性が高い。例えば、本発明はユーザーがユニーク配列プローブの長さを出力用に選択することを可能にする。

【図面の簡単な説明】

【００１５】

以下の図面は本明細書の一部を成し、本発明の態様をより明確に示すものである。これらの図面の一つ以上を本明細書中に述べた具体例の詳細な説明と共に参照することにより、本発明をより良く理解することができよう。本出願は少なくとも一つのカラー図面を含む。この特許出願公開のカラー図面付きの写しは、必要な料金を添えて申請すれば当該官庁より提供される。

【００１６】

【図１】ウェブベースのユニークゲノム配列ハンター（ＵＧＳＨ）プログラムの入力画面を示すスクリーンキャプチャーの説明図である。

【図２Ａ】ユニーク配列ゲノムプローブとその位置を表示するＵＧＳＨからの出力の一例を示すスクリーンキャプチャーの説明図である。

【図２Ｂ】ＵＧＳＨからの「プライマー選択出力」画面の一例を示すスクリーンキャプチャーの説明図である。

【図２Ｃ】ＦＡＳＴＡフォーマットで表示されたＵＧＳＨからのプライマー配列ファイルの一例を示すスクリーンキャプチャーの説明図である。

【図３】第７番染色体上のＢＡＣ　ＲＰ１１－６７７Ｆ１４由来のユニーク配列プローブを用いた蛍光ｉｎ　ｓｉｔｕハイブリダイゼーション（ＦＩＳＨ）実験で撮影した写真である。

【図４】五つの異なるユニーク配列プローブを含むユニーク配列プローブカクテルを用いたＦＩＳＨ実験で撮影した写真である。

【図５】ＵＧＳＨ法を用いて設計していないプローブを用いたＦＩＳＨ実験の結果を示す写真である。染色体上の多数の位置にハイブリダイズしたプローブ（矢印で示した薄い灰色の箇所）が、この配列が二つ以上の染色体領域と相同であり、従って純然たるユニーク配列ではないことを示している。

【図６】既知の核酸配列内から少コピー数の核酸セグメントを同定し、同定した少コピー数セグメントの中からハイブリダイゼーション実験に熱力学的に好適に用いられるセグメントを選択するためのコンピュータ化された方法の一具体例を示すフローチャートである。

【図７】既知の核酸配列内から少コピー数の核酸セグメントを同定し、同定した少コピー数セグメントの中からハイブリダイゼーション実験に熱力学的に好適に用いられるセグメントを選択するためのコンピュータ化された方法の別の具体例を示すフローチャートである。

【図８】被検者または患者から一例として取得した配列内で既知の繰り返し配列を同定するためのコンピュータ化された方法の一具体例を示すフローチャートである。

【図９】被検者または患者から取得した配列から既知の繰り返し配列を抽出し、残った配列部分をユーザーが指定したサイズパラメーターに従って選択するためのコンピュータ化された方法の一具体例を示すフローチャートである。

【発明を実施するための形態】
【００１７】
本発明は、ゲノムＤＮＡにおいてユニーク配列領域を同定するための新規なコンピュータ化プロセスを含み、ユニーク配列ゲノムセグメントを設計する方法を提供する。同定したセグメントはゲノム、もしくはゲノムの一部から、ゲノムライブラリーから、または他のゲノムＤＮＡ源から合成または増幅することができ、マイクロアレイ、アレイＣＧＨ（マイクロアレイと共に「アレイベースの」と形容される）、定量的マイクロスフェアハイブリダイゼーション（ＱＭＨ）、および蛍光ｉｎ　ｓｉｔｕハイブリダイゼーション（ＦＩＳＨ）を非限定的に含めたハイブリダイゼーション実験に用い得る。本発明のコンピュータ化プロセスおよび関連する方法はユーザー基準に適う配列のみを報告し（例えば、コンピュータプログラムウィンドウ内に表示し、データファイルに格納し、印刷その他の手段で出力し）、基準に合致しない配列は廃棄される。
【００１８】
上記方法では、二つの異なる方策を組み合わせることにより、ゲノム配列またはセグメントがユニークもしくは非繰り返し配列組成について評価され、かつ同定されたユニーク配列領域の熱力学特性が解析されて、同定された少コピー数核酸セグメントがハイブリダイゼーションアッセイにおいて最適性能を発揮することが保証されるので、該方法は従来方法に優る。
【００１９】
本発明の方法は、本明細書中でユニークゲノム配列ハンター（ＵＧＳＨ）と呼称する単一のコンピュータプログラムを用いて配列をそのゲノム上での表出すなわち分布と、熱力学特性との両方に関して解析することにより、現行技術に進歩をもたらす。本発明の方法の好ましい一形態は五つの主要ステップ、すなわち：　１）対象配列から高頻度および中頻度繰り返し配列を除去し、該当するゲノムセグメント（すなわち、繰り返し配列除去後に残ったセグメント）を表示するステップ。得られるゲノムセグメントはどのようなサイズであってもよいが、ＦＩＳ用としては、好ましくは５００ｂｐを越え、さらに好ましくは７５０ｂｐを越え、最も好ましくは１ｋｂを越えるサイズを有する；　２）各セグメントを、対象領域以外のゲノム領域とのホモロジーに関して検索し、ゲノムの他の箇所とマッチするセグメントを総て廃棄するステップ；　３）熱力学的解析により、ユニーク配列セグメントを予測される二次構造モチーフ（ヘアピンループ、ステム、バルジなど）に関して評価するステップ；　４）上記三つのステップを経たゲノムセグメントのためのＰＣＲプライマーを設計するステップ；　および５）各ＰＣＲプライマーを評価し、当該プライマーがユニーク配列しか含まず、ゲノムの他の箇所とマッチしないことを確実にするステップを含む。幾つかの好ましい形態において、このプロセスはステップ３までで中止され、他の好ましい形態ではステップ４が済んだところで中止される。しかし、用時、五つのステップを総て実行することが好ましい。上記一連のステップは、実験室内でのゲノミクス実験に用いられるユニーク配列プローブを設計するためのより頑強かつ正確なツールを提供する。これらのステップは必ずしも上記の順序で実行されなくともよい。基本方法の変形例では、上記ステップの一つ以上が除かれる。一具体例において、本発明方法の複数のステップはコンピュータプログラムを介して自動化される。好ましくは、そのコンピュータプログラムはＰｅｒｌなどの、ウェブベースのアプリケーションの創作に適したコンピュータ言語で書かれる。
【００２０】
ＵＧＳＨの開発
ＵＧＳＨ法は、ゲノムプローブを繰り返し設計し、試験することにより開発された。始めに従来技術（米国特許第６，８２８，０９７号（’０９７特許）および同第７，０１４，９９７号（’９９７特許））の方法を用いて、定量的マイクロスフェアハイブリダイゼーション（ＱＭＨ）実験用の「シングルコピー」プローブが生成された（Ｎｅｗｋｉｒｋら、２００６年、「定量的マイクロスフェアハイブリダイゼーションによるゲノムコピー数測定」、Ｈｕｍａｎ　Ｍｕｔａｔｉｏｎ、２７：３７６－３８６）。ＱＭＨアッセイは

、ビオチニル化した患者ゲノムＤＮＡにスペクトルの明らかなマイクロスフェアと結合させたユニーク配列プローブを直接ハイブリダイズさせ、その後フローサイトメトリー解析を行なうことにより高スループットのゲノムコピー数測定を可能にする（Ｎｅｗｋｉｒｋら、２００６年、米国仮特許出願第６０／７０８，７３４号）。フローサイトメトリーでは、多重反応において試験プローブと、一つの二倍体ゲノムに２コピー存在することが知られている基準プローブとに関し平均蛍光強度（ＭＦＩ）が測定される。続いてＭＦＩ比（試験プローブ対基準プローブ）を計算して、試験プローブが２コピー存在するか（ＭＦＩ比＝１）、１コピー存在するか（ＭＦＩ比＝０．５）、それとも３コピー以上存在するか（ＭＦＩ比＞１）が確認される。ＵＧＳＨ法のこのようなステップ１は、上記特許出願に記載された方法と類似してはいるが異なっている。上記特許出願の方法は、対象配列の反復配列マスキング（すなわち、対象配列をゲノム中のあらゆる既知繰り返し配列と比較して、９０％以上の配列類似性を有する配列を排除もしくは「マスキング」すること（これにより、ギャップおよびウィンドウが導入されて二つの配列がより良くマッチし得る））を行ない、ユニークプローブもしくは「シングルコピープローブ」を生成することを含む。例えば、’０９７特許の方法を用いてＡＢＬ１（ｃｈｒ９）に特異的な配列を解析後、ＱＭＨ用にプローブ（ＡＢＬＡｌｕＭｅｒ１と呼称）が設計された（Ｎｅｗｋｉｒｋら、２００５年）。基準配列としては公知のシングルコピーＨＯＸＢ１配列（Ｎｅｗｋｉｒｋら、２００６年）が用いられた。どちらのプローブ（～１００塩基）もスペクトルの明らかなマイクロスフェアと結合させてから、ビオチニル化した正常対照ゲノムＤＮＡにハイブリダイズされた。確認に用いられたのが正常対照ＤＮＡであったので、ＨＯＸＢ１およびＡＢＬＡｌｕＭｅｒ１プローブのＭＦＩ比は１になるはずであるが、ＭＦＩ比は４．５５であった。このことは、ＡＢＬＡｌｕＭｅｒ１配列がゲノム中の他の相同領域にハイブリダイズされたことを示唆している（Ｎｅｗｋｉｒｋら、２００５年、「Ｃｏｔ－１　ＤＮＡによる定量的ゲノムおよび発現ハイブリダイゼーションの歪み：　該作用の緩和」、Ｎｕｃｌｅｉｃ　Ａｃｉｄｓ　Ｒｅｓｅａｒｃｈ、３３：ｅ１９１）。

【００２１】

次に、反復配列マスキング（ステップ１）に続いてゲノムホモロジー検索（ステップ２）を行なうことを含む別の方策が用いられ、ＡＢＬに特異的なプローブ１６－１ｄが設計された（Ｎｅｗｋｉｒｋら、２００６年）。このプローブを、ＨＯＸＢ１を用いたＱＭＨ反応において二つの異なる正常ヒトゲノムＤＮＡにハイブリダイズすることにより、１．３６および１．１８のＭＦＩ比がそれぞれ得られた。これらの比は、１に近付いてはいるが、なお最適ではない。１６－１ｄプローブのその後の解析から、プローブの３’末端近傍に安定なヘアピンループ構造が存在することが判明しており（Ｎｅｗｋｉｒｋら、２００６年）、このことが、該プローブのＭＦＩ比が最適とならないことの説明となるかもしれない。方法をさらに改良するべく、二次構造解析ステップ（ステップ３）を組み込むことでＵＧＳＨ法の改善が図られた。

【００２２】

対象のＡＢＬ配列領域から反復配列を除去し、かつゲノムホモロジー検索および二次構造解析を行なって、別のプローブ１６－１ｂが開発された（１００塩基：　Ｎｅｗｋｉｒｋら、２００６年）。１６－１ｂはＨＯＸＢ１と共にＱＭＨ実験に用いられ、その際ＭＦＩ比は１．０１±０．０１（１６正常試料を試験）であった。このことは、このプローブがゲノム中の単一位置にハイブリダイズされたことを示唆している。このように、ステップ１、２および３を組み合わせることにより従来よりも優れた結果が得られた。二次構造解析のための正確なパラメーター（ΔＧ＜５０、ΔＨ＜－１０００、ΔＳ＜－３５００、およびＴｍ≧６５℃：　先に挙げた基準が満たされない場合）は、二次構造の程度が様々である複数のユニーク配列プローブを用いた実験によって確認された。従来技術で開発されたプローブの一つである１６－１ａは強い二次構造性（ΔＧ＝－１２２、ΔＨ＝－１５８４、ΔＳ＝－４７１４、Ｔｍ＝６３℃）を示した（Ｎｅｗｋｉｒｋら、２００６年）。プローブ１６－１ａはＱＭＨ反応においてＨＯＸＢ１と共にハイブリダイズされ、その際ＭＦＩ比は正常ゲノム対照試料に関し０．７３から０．９３（ｎ＝４）で、プローブの不

安定さを示した。反復配列マスキングとそれに続くゲノムホモロジー検索（上記ステップ１および２）を用いて設計された、従来技術による別のプローブ１６－２Ａも、どちらかと言えば強い二次構造性（ΔＧ＝－９１、ΔＨ＝－１２９６、ΔＳ＝－３８８６、Ｔｍ＝６０℃）を示した（Ｎｅｗｋｉｒｋら、２００６年）。

【００２３】

ＨＯＸＢ１を用いたＱＭＨ実験では、正常ゲノムＤＮＡとのＱＭＨ反応におけるＭＦＩ比が０．８４から０．９２（ｎ＝４）であり、ＭＦＩ比が１に近いことでプローブ構造が僅かながらより安定であることが示唆された。プローブ１６－１ｂ（Ｎｅｗｋｉｒｋら、２００６年）は別様の二次構造性（ΔＧ＝－９．６６、ΔＨ＝－１３８．８、ΔＳ＝－４１６．４、Ｔｍ＝６０．２℃）を示し、ＨＯＸＢ１を用いた、正常ゲノム対照ＤＮＡ試料に対する多重ハイブリダイゼーションでのＭＦＩ比は０．９６と１．０９との間（ｎ＝１１）であった（Ｎｅｗｋｉｒｋら、２００６年）。

【００２４】

図６を参照すると、ゲノムハイブリダイゼーションプローブ選択のためのユニークゲノム配列ハンター（ＵＧＳＨ）法はＤＮＡ配列を要求し（ステップ１）、該配列はＦＡＳＴＡまたはＧｅｎｂａｎｋフォーマットでＵＧＳＨプログラムに入力され得る。この配列は染色体座標、遺伝子名、または対象領域によって定義されてもよい（ステップ１ａ）。その場合（ステップ１ａ）、ＵＧＳＨがデータベースに照会するが、照会（すなわち、Ｃｈｒ１５：２１２６３４２１－２１２６３８２１、ＳＮＲＰＮ、ＰＷＳなど）に応じた適当な配列を取り出す上で特に好ましいデータベースはＵＣＳＣデータベース（ｇｅｎｏｍｅ．ｕｃｓｃ．ｅｄｕ）である。プロセスの次のステップ（ステップ２）で、入力された配列から繰り返し配列が除去される。ＵＧＳＨはこれを、ＤＮＡの高頻度繰り返しクラスの配列（ＳＩＮＥ、ＬＩＮＥ、サテライト、ショートタンデムリピート、ミニサテライト、マイクロサテライト、テロメアなど）と対象配列とのアラインメントを行なうことにより実現する。具体的には、ＵＧＳＨはＲｅｐｅａｔＭａｓｋｅｒプログラムを実行して繰り返し配列を除去するが、ＲｅｐｅａｔＭａｓｋｅｒのために厳密に定義された出力パラメーターを用いて、既知の反復配列に対し９０％以上のホモロジーマッチを有する配列は総て排除する。この手続きには、同様の反復配列マスキングプログラムであればいずれを用いてもよい。反復配列に関して既にマスキングを施した照会配列を入力することにより反復配列マスキングステップを迂回することも可能である（ステップ２Ａ）。ＵＣＳＣゲノムブラウザおよびＧｅｎｂａｎｋは、マスキングされた配列を表示するオプションを提供することで反復配列マスキングステップを不要にする。

【００２５】

方法のこの段階で、ＵＧＳＨプログラムは反復配列をマスキングしたＤＮＡ配列を生成した。プロセスの次のステップ（ステップ３）ではこの配列をゲノム中の相同配列に関して、ＵＣＳＣゲノムブラウザからのＢＬＡＴプログラムを用いてスキャンする。配列のセグメントでＢＬＡＴスコアが３０以上となるものは総てプローブ選択から除外する。ＮＣＢＩ提供のＢＬＡＳＴなど、ゲノムワイドなホモロジー検索プログラムであればいずれをＢＬＡＴに替えて用いてもよく、その際パラメーターは同一のものを用い得る（許容可能なスコアは≦３０または１から３０の間で、好ましくは２５未満（または１から２５の間）、さらに好ましくは２０未満（または１から２０の間）、さらに好ましくは１５未満（または１から１５の間）、さらに好ましくは１０未満（または１から１０の間）、さらに好ましくは８未満（または１から８の間）、さらに好ましくは６未満（または１から６の間）、さらに好ましくは５未満（または１から５の間）、さらに好ましくは４未満（または１から４の間）、さらに好ましくは３未満（または１から３の間）、さらに好ましくは２未満（または１から２の間）、最も好ましくは１）。

【００２６】

次に、反復配列を含まず、ゲノムの他の箇所とほとんど乃至全く相同でない残余配列を、プローブをゲノムハイブリダイゼーション実験に準最適としかねない潜在的二次構造（すなわち、バルジ、ループ、またはステム）について調べる（ステップ４）。好ましいＵ

ＧＳＨ法はＭｆｏｌｄプログラムを採用し、かつプローブ選択のために厳密に定義されたパラメーター（ΔＧ＜５０、ΔＨ＜－１０００、ΔＳ＜－３５００、Ｔｍ≧６０℃、またはＱＭＨやアレイベースの用途のために別に示されるもの）を用いる。それらのパラメーターが満たされない場合、配列はプローブ設計から除外される。

【００２７】

ＰＣＲプローブが求められる場合は、二次構造解析実施後に残った配列をＰＣＲプライマー設計に用いる（ステップ５）。ＵＧＳＨ法はＰｒｉｍｅｒ３プログラム（Ｒｏｚｅｎら、２０００年）を用いて、長さ１５塩基以上のプライマーを設計する。プライマーの長さは、ＦＩＳＨ用には１５～１００塩基とし得、アレイベースの用途やＱＭＨ用には１５～７０、好ましくは２５～７０塩基とし得る。ＦＩＳＨ用として特に好ましい長さは２２塩基などである。また、いずれの用途においても、生成物のサイズは入力された配列のサイズと等しくなるか、またはそれより僅かに小さくなる。好ましくは、生成物サイズが入力配列サイズと等しくなるか、またはそれより僅かに小さくなるということは前者が後者を０～２００塩基下回るということであるが、通常のプライマー選択プログラムのいずれを代用することも可能であり、入力配列が比較的長いと生成物のサイズが入力配列のサイズを２００塩基より大きく下回ることも有る。設計したプライマーはＵＣＳＣ　ＢＬＡＴプログラムを用いてＢＬＡＴ検索し（ステップ６）、ゲノムの他の箇所と相同な配列が存在しないことを確実にする。ゲノムの２箇所以上とマッチするプライマーは総て廃棄する。ＰＣＲプローブの替わりにハイブリダイゼーションオリゴヌクレオチドが求められる場合にはＰＣＲプライマー設計ステップおよびＰＣＲプライマーホモロジー検索ステップを省略し、ステップ４から得られた、マッチする相同な配列がゲノム上にない、反復配列を含まない配列をハイブリダイゼーションプローブとして用いることができる。全プロセス完了後、ＵＧＳＨはサイズでソートしたユニーク配列を、求めが有ればプライマー配列と共に表示する（ステップ７）。これがＵＧＳＨ法で実行されるプロセスの概要であるが、ステップ２から７は通常ＵＧＳＨプログラムによって自動的に実施され、ユーザーには分からない。

【００２８】

ＵＧＳＨは好ましくはインターネットもしくはウェブベースアプリケーションとして実行され、その際グラフィカルユーザーインタフェース（ＧＵＩ）は一つ以上のインターネットブラウザウィンドウから提供される。図１は、ウェブベースインタフェースを介して提供されるＵＧＳＨ入力ページのスクリーンキャプチャーである。ユーザーはジョブ名、プローブ選択のための最小サイズ、および１行当たりの表示塩基数を入力する。次にユーザーは対象配列を、ＦＡＳＴＡフォーマットで配列ボックスに入力するか、またはブラウズボタンを用いてＮＣＢＩからＧｅｎｂａｎｋファイルフォーマットでアップロードする。報告されるべきプライマーの数は、デフォルトパラメーターとしては通常２５とされるが、ユーザーが変更してもよい。プローブについて定められる最小ＰＣＲ生成物サイズもユーザーによって変更可能である。パラメーターを総て入力したユーザーは実行ボタンをクリックし、ユニーク配列プローブ選択のためのＵＧＳＨプログラムを実行する。

【００２９】

図２Ａは、ユニーク配列領域を入力配列内での位置によって表示するＵＧＳＨ出力ページのスクリーンショットである。ＵＧＳＨプログラムにＧｅｎｂａｎｋ配列ファイルがアップロードされた場合はソース欄に、いずれもＧｅｎｂａｎｋによって決定されたファイルの定義、配列の登録番号、配列のバージョン（該当する場合）および配列のＧＩ番号が記載される。ユーザーによって指定されたジョブ名が、ユーザーによって入力された配列の全長と共に表示される。入力画面で指定された、ユニーク配列プローブの選択で考慮されるべき最小サイズが示される。ユニーク配列領域の位置が（例えば「＞３１６５－４２６２」と）表示された後、その座標に含まれる実際の配列が表示される。配列情報のあとにプライマーが表示される（図２Ｂ）。

【００３０】

図２Ｂは、各ユニーク配列領域について配列数を表示する、ＵＧＳＨプログラムからの

「プライマー選択出力」画面の一例を示すスクリーンキャプチャーである。図示例において、配列はｓｅｑ１．ｐｒｉｍｅｒ、ｓｅｑ２．ｐｒｉｍｅｒなどと名付けられ、プライマー設計に用いられる各ユニーク配列領域のサイズが括弧に入れて示される。テキストファイルを開くと２５個の、または入力画面でユーザーが指定した数の実際のプライマー配列が記録されたファイルが表示される（図２Ｃ）。

【００３１】

図２Ｃは、ＦＡＳＴＡフォーマットで表示されたＵＧＳＨからのプライマー配列ファイルの一例を示すスクリーンキャプチャーである。ユーザーがプライマー配列ファイル上でクリックすると、該ファイルが表示される。「ＰＬ」はユニーク配列領域の左のプライマーを、「ＰＲ」は右のプライマーを意味する。プローブ全体を示す「ＰＦ」では括弧内に、左プライマーの開始位置、左プライマーの長さ、右プライマーの開始位置、および右プライマーの長さが入力配列との関係において表示される。その下方に、プライマーを含む規定された領域が示される。後続するプライマーはそれぞれ０からｎまでの数字を付して示され、その際ｎはＵＧＳＨ入力画面でユーザーが指定した、表示されるべきプライマーの数である。グラフィカルインタフェース（図１）を用いて配列が入力される（ステップ１または１ａ）。「実行」ボタンがクリックされた後、ユニーク配列プローブおよびプライマーが表示され（図２Ａ、２Ｂおよび２Ｃ）、これがプロセスの最終ステップとなる（ステップ７）。介在する他のステップは総て、ＵＧＳＨユーザーには分からないように（見えないところで、もしくはユーザーとの対話を要せずに）実施される。

【００３２】

図７に、以下の手続きを概略的に示す。一つまたは複数の患者配列が与えられた（入力された）時、その配列がすでにアノテーション付きである（すなわち、反復配列の位置が既知である）場合は候補のユニーク配列を直接生成し（図９参照）、そうでない場合は、プログラムは反復配列位置を決定してから次のステップへと戻る。生成した候補配列はＦＡＳＴＡファイルフォーマットで格納し、それらに対してＢＬＡＳＴまたはＢＬＡＴ（デフォルト設定）を実行することにより、ユーザー、第三者またはデフォルトの基準を満たさないセグメントを総て選び出す。残った配列をＭｆｏｌｄプログラムに通し、そこから出力される配列をＰｒｉｍｅｒ３プログラムへ送り、該プログラムに処理させる。Ｐｒｉｍｅｒ３プログラムがプローブを生成する。ＢＬＡＴまたはＢＬＡＳＴプログラムを再度実行することにより、プローブを検証する。各ステップは、本出願中で詳述されるフィルタリングしきい値を有する。

【００３３】

患者配列はしばしばＮＣＢＩデータベースから取得され、従ってその特徴（すなわち、反復配列の位置など）がアノテーションとして付加されている。図８を参照されたい。アノテーションが付加されていない場合は、ＲｅｐｅａｔＭａｓｋｅｒやＤｕｓｔなどといった無償の反復配列探索プログラムを用いて既知の繰り返し配列を患者配列内で特定する。上記プログラムがもたらす出力は、典型的にはＦＡＳＴＡフォーマットで記録されたあらゆる反復配列とその位置のリストを含む。

【００３４】

図９に示すように、反復配列を総て除去し、かつ残った配列のうちで対象サイズを有するものを総て抽出することにより候補配列を生成する。出力配列は、次のプログラムと整合するフォーマット済みファイルに（すなわち、ＦＡＳＴＡフォーマットで）格納する。

【００３５】

ＵＧＳＨプログラムの一具体例を擬似コードで示す。ここに示したプログラムは複数のモジュールに編成されており、それらはプログラムが用いられる際に互いに、またインターネット上で入手可能な他のプログラムおよびデータと対話する。本明細書に述べた方法は好ましくはコンピュータ内のプロセッサまたはプログラムによって実施されると理解される。

【００３６】

主制御機能

ウェブユーザーインタフェースを作成する｛
パラメーター
好ましい具体例に含まれるパラメーター：
（１）　ジョブ名（テキスト）
（２）　最小ユニーク配列サイズ（整数、１０００ｂｐｓ）
（３）　１行当たりの塩基対数（整数、デフォルト＝６０ｂｐｓ）
（４）　配列（アップロードファイルまたはテキスト）
（５）　報告されるプライマーの数（整数、デフォルト＝２５ｂｐｓ）
（６）　最小生成物サイズ（整数、デフォルト＝１００ｂｐｓ）

任意パラメーター：
（７）　Ｍｆｏｌｄ用パラメーター（下記リストおよび／またはＭｆｏｌｄウェブサイト参照）
（８）　ＢＬＡＴ／ＢＬＡＳＴ用パラメーター（下記リストおよび／またはＢＬＡＴ／ＢＬＡＳＴウェブサイト参照）

オプション
好ましい具体例に含まれるオプション
（１）　患者配列の処理
（２）　プライマー生成

択一的具体例に含まれるオプション
（３）　Ｍｆｏｌｄインタフェース（あとで追加）
（４）　ＢＬＡＴ／ＢＬＡＳＴインタフェース（あとで追加）
（５）　ＲｅｐｅａｔＭａｓｋｅｒインタフェース（あとで追加）

動作ボタン
（１）　アップロード
（２）　サブミット
（３）　リセット
（４）　結果を電子メールで送信（将来追加
｝
Ｉｆ　アップロードが真である｛
<u>ＵＧＳＨプロセス</u>
アップロードされたファイル内の配列に対して実施する
Ｅｌｓｅ　ｉｆ　サブミットが真である｛
<u>ＵＧＳＨプロセス</u>
ＵＧＳＨ配列テキストボックスに入力された配列に対して実施する
Ｅｌｓｅ　ｉｆ　リセットが真である｛
全パラメーターをデフォルト値にリセットする
｝
｝
Ｅｌｓｅ｛
信号（すなわちボタンのクリック）を待つ
｝

【００３７】
ＵＧＳＨプロセス
｛
入力：患者配列

出力：プローブ

配列を読み取る（ＦＡＳＴＡフォーマットが必要）
Ｉｆ　配列にアノテーションが付加されている｛
　反復配列の特徴（例えば位置）を抽出する
　非繰り返し配列を含む新しいファイルを生成する
｝
Ｅｌｓｅ｛
　反復配列探索プログラム（例えばＲｅｐｅａｔＭａｓｋｅｒ）を実行する
　反復配列の特徴を抽出する
　非繰り返し配列を含む新しいファイルを生成する
｝

【００３８】
／／以下の手続きは、典型的には連続して実行される複数のモジュールのパイ／／プラインである（各モジュールは一組のフィルタリングパラメーターを有／／する異なるプログラムを実行）。

上記生成した配列でＢＬＡＴまたはＢＬＡＳＴを実行する

ＢＬＡＴまたはＢＬＡＳＴからの出力をフィルタリングする

上記フィルタリングした配列でＭｆｏｌｄを実行する

Ｍｆｏｌｄ検査に通った配列を収集する

上記収集した配列でＰｒｉｍｅｒ３を実行する

Ｐｒｉｍｅｒ３からの出力を収集する

Ｐｒｉｍｅｒ３からの出力配列でＢＬＡＴまたはＢＬＡＳＴを実行する

検証した配列をプローブとして出力する
｝

【００３９】
<u>反復配列探索</u>
｛
入力：ファイルに入れた標的配列
出力：ファイルに入れた非繰り返し配列

デフォルトパラメーターでＲｅｐｅａｔＭａｓｋｅｒを実行する
特徴を抽出する
非繰り返し配列をファイルにセーブする
｝

【００４０】
<u>配列を読み取る</u>
｛
配列ファイルをアップロードする

```
  各行を解析する｛
      Ｉｆ　当該行が配列の名前である｛
        当該行を名前アレイに格納する
      ｝
      Ｉｆ　当該行がＤＮＡ配列である｛
        当該行を配列アレイに格納する
      ｝
      Ｉｆ　ファイルが違法な配列を含む｛
        処理を中止して警告する
        プログラムを終了する
      ｝
  ｝
｝
```

【００４１】

<u>反復配列の特徴を抽出する</u>

```
｛
  入力：アノテーション付きの標的配列
  出力：ファイルに入れた非繰り返し配列

  反復配列アノテーション中の各反復配列に関して｛
      位置および反復配列長を読み取る
      当該配列を除去して次の反復配列を出現させる
      間の非繰り返しセグメントを保持する
      ｉｆ　セグメントサイズ≧所定のしきい値である
      ｛
        セグメントに名前を付け、ファイルに格納する

           命名規則：各非繰り返し配列について、標的配列名のあとに当該配列の位置
                     範囲を付したものをその配列の名前とする
           格納フォーマット：デフォルトによりＦＡＳＴＡ配列フォーマット
      ｝
      Ｅｌｓｅ
      ｛
        セグメントをスキップする
      ｝

  ｝
｝
```

【００４２】

<u>ＢＬＡＴまたはＢＬＡＳＴを実行する</u>

```
｛
  入力：ファイルに入れた非繰り返し配列
  出力：ヒトゲノム配列に対するユニーク配列

  デフォルトパラメーターでＢＬＡＴまたはＢＬＡＳＴを実行する
  ＢＬＡＴ／ＢＬＡＳＴ出力をスキャンする｛
      Ｉｆ　該出力がユニークな相同配列である｛
          候補配列としてデータファイルに格納する
```

```
    }
    Ｅｌｓｅ｛
        配列を保存しない
    }
}
```

【００４３】

<u>Ｍｆｏｌｄを実行する</u>

```
｛
  入力：ＢＬＡＴ／ＢＬＡＳＴからのユニーク候補配列
  出力：ファイルに入れた、熱力学的に安定な配列
    オプション：配列の熱力学特性／フォールディング構造に関してＭｆｏｌｄが計算した一つ以上の変数をＵＧＳＨに渡し、それによってＵＧＳＨ　ＧＵＩウィンドウ内でユーザーに示し、および／またはデータファイルにローカルに格納する
```

【００４４】

一組の所定パラメーターでＭｆｏｌｄを実行する

<u>ＵＧＳＨによってＭｆｏｌｄに用意されるパラメーター（ほとんどのパラメーターに、Ｍｆｏｌｄプログラムにおいて確立されるデフォルト設定が用いられ得る）</u>

```
    配列名
    配列
    フォールディングの制約
        特定の塩基対またはヘリックスの形成を強いる
        特定の塩基対またはヘリックスの形成を禁じる
        一続きの塩基の対形成を強いる
        一続きの塩基の対形成を禁じる
        一続きの塩基が別の一続きの塩基と対になることを禁じる
    線状配列か環状配列かを特定する
    フォールディング温度
    イオン条件（すなわち、Ｎａ＋およびＭｇ＋＋のモル濃度）
    準際適度
    ウィンドウパラメーター
    対になった塩基同士の最大間隔
```

【００４５】

```
  Ｍｆｏｌｄ出力をスキャンする｛
    Ｉｆ　出力が、配列が熱力学的に安定であることを示唆する（基準は所定）
    ｛
        候補配列としてデータファイルに格納
    }
    Ｅｌｓｅ
    ｛
        配列を保存しない
    }
  }
}
```

【００４６】

<u>Ｐｒｉｍｅｒ３を実行する</u>

```
｛
```

入力：安定なユニーク配列
出力：ゲノムプローブ配列

【００４７】
一組の所定パラメーターでＰｒｉｍｅｒ３を実行する

<u>ＵＧＳＨによってＰｒｉｍｅｒ３に用意されるパラメーター：</u>
ＰＲＩＭＥＲ＿ＭＡＸ＿ＥＮＤ＿ＳＴＡＢＩＬＩＴＹ＝９．０
ＰＲＩＭＥＲ＿ＭＡＸ＿ＭＩＳＰＲＩＭＩＮＧ＝１２．００
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＭＡＸ＿ＭＩＳＰＲＩＭＩＮＧ＝２４．００
ＰＲＩＭＥＲ＿ＭＩＮ＿ＳＩＺＥ＝１８
ＰＲＩＭＥＲ＿ＯＰＴ＿ＳＩＺＥ＝２４
ＰＲＩＭＥＲ＿ＭＡＸ＿ＳＩＺＥ＝２７
ＰＲＩＭＥＲ＿ＭＩＮ＿ＴＭ＝５７．０
ＰＲＩＭＥＲ＿ＯＰＴ＿ＴＭ＝６０．０
ＰＲＩＭＥＲ＿ＭＡＸ＿ＴＭ＝６３．０
ＰＲＩＭＥＲ＿ＭＡＸ＿ＤＩＦＦ＿ＴＭ＝１００．０
ＰＲＩＭＥＲ＿ＭＩＮ＿ＧＣ＝２０．０
ＰＲＩＭＥＲ＿ＭＡＸ＿ＧＣ＝８０．０
ＰＲＩＭＥＲ＿ＳＥＬＦ＿ＡＮＹ＝８．００
ＰＲＩＭＥＲ＿ＳＥＬＦ＿ＥＮＤ＝３．００
ＰＲＩＭＥＲ＿ＮＵＭ＿ＮＳ＿ＡＣＣＥＰＴＥＤ＝０
ＰＲＩＭＥＲ＿ＭＡＸ＿ＰＯＬＹ＿Ｘ＝５
ＰＲＩＭＥＲ＿ＯＵＴＳＩＤＥ＿ＰＥＮＡＬＴＹ＝０
ＰＲＩＭＥＲ＿ＦＩＲＳＴ＿ＢＡＳＥ＿ＩＮＤＥＸ＝１
ＰＲＩＭＥＲ＿ＧＣ＿ＣＬＡＭＰ＝０
ＰＲＩＭＥＲ＿ＳＡＬＴ＿ＣＯＮＣ＝５０．０
ＰＲＩＭＥＲ＿ＤＮＡ＿ＣＯＮＣ＝５０．０
ＰＲＩＭＥＲ＿ＭＩＮ＿ＱＵＡＬＩＴＹ＝０
ＰＲＩＭＥＲ＿ＭＩＮ＿ＥＮＤ＿ＱＵＡＬＩＴＹ＝０
ＰＲＩＭＥＲ＿ＱＵＡＬＩＴＹ＿ＲＡＮＧＥ＿ＭＩＮ＝０
ＰＲＩＭＥＲ＿ＱＵＡＬＩＴＹ＿ＲＡＮＧＥ＿ＭＡＸ＝１００
ＰＲＩＭＥＲ＿ＷＴ＿ＴＭ＿ＬＴ＝１．０
ＰＲＩＭＥＲ＿ＷＴ＿ＴＭ＿ＧＴ＝１．０
ＰＲＩＭＥＲ＿ＷＴ＿ＳＩＺＥ＿ＬＴ＝１．０
ＰＲＩＭＥＲ＿ＷＴ＿ＳＩＺＥ＿ＧＴ＝１．０
ＰＲＩＭＥＲ＿ＷＴ＿ＧＣ＿ＰＥＲＣＥＮＴ＿ＬＴ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＧＣ＿ＰＥＲＣＥＮＴ＿ＧＴ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＣＯＭＰＬ＿ＡＮＹ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＣＯＭＰＬ＿ＥＮＤ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＮＵＭ＿ＮＳ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＲＥＰ＿ＳＩＭ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＳＥＱ＿ＱＵＡＬ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＥＮＤ＿ＱＵＡＬ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＰＯＳ＿ＰＥＮＡＬＴＹ＝０．０
ＰＲＩＭＥＲ＿ＷＴ＿ＥＮＤ＿ＳＴＡＢＩＬＩＴＹ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＰＲＯＤＵＣＴ＿ＳＩＺＥ＿ＬＴ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＰＲＯＤＵＣＴ＿ＳＩＺＥ＿ＧＴ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＰＲＯＤＵＣＴ＿ＴＭ＿ＬＴ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＰＲＯＤＵＣＴ＿ＴＭ＿ＧＴ＝０．０

ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＤＩＦＦ＿ＴＭ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＣＯＭＰＬ＿ＡＮＹ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＣＯＭＰＬ＿ＥＮＤ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＲＥＰ＿ＳＩＭ＝０．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＰＲ＿ＰＥＮＡＬＴＹ＝１．０
ＰＲＩＭＥＲ＿ＰＡＩＲ＿ＷＴ＿ＩＯ＿ＰＥＮＡＬＴＹ＝０．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＩＮ＿ＳＩＺＥ＝１８
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＯＰＴ＿ＳＩＺＥ＝２０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＡＸ＿ＳＩＺＥ＝２７
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＩＮ＿ＴＭ＝５７．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＯＰＴ＿ＴＭ＝６０．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＡＸ＿ＴＭ＝６３．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＩＮ＿ＧＣ＝２０．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＡＸ＿ＧＣ＝８０．０
ＰＲＩＭＥＲ＿ＩＮＴＥＲＮＡＬ＿ＯＬＩＧＯ＿ＭＡＸ＿ＰＯＬＹ＿Ｘ＝５
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＴＭ＿ＬＴ＝１．０
ＰＲＨＶＩＥＲ＿ＩＯ＿ＷＴ＿ＴＭ＿ＧＴ＝１．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＳＩＺＥ＿ＬＴ＝１．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＳＩＺＥ＿ＧＴ＝１．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＧＣ＿ＰＥＲＣＥＮＴ＿ＬＴ＝０．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＧＣ＿ＰＥＲＣＥＮＴ＿ＧＴ＝０．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＣＯＭＰＬ＿ＡＮＹ＝０．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＮＵＭ＿ＮＳ＝０．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＲＥＰ＿ＳＩＭ＝０．０
ＰＲＩＭＥＲ＿ＩＯ＿ＷＴ＿ＳＥＱ＿ＱＵＡＬ＝０．０

【００４８】
　Ｐｒｉｍｅｒ３からの出力を収集する

　Ｐｒｉｍｅｒ３からの出力配列でＢＬＡＴまたはＢＬＡＳＴを実行する

　検証した配列をプローブとして出力する
}

【００４９】
注記：　ＵＧＳＨとユーティリティプログラム（Ｍｆｏｌｄ、ＢＬＡＴ／ＢＬＡＳＴ、Ｐｒｉｍｅｒ３など）との間でのデータの受け渡しはテキストファイルでか、または一方のプログラムが用意する任意パラメーターを介して行なわれる。それらのパラメーターは、ウェブインタフェースを介して受け取ったり、事前に定義してファイルに格納したり、あるいはまた定数として取り扱う場合はＵＧＳＨプログラム（すなわちＰｅｒｌ）スクリプトに含めたりすることができる。
【００５０】
［定義］
　特に断らない限り、本明細書中に用いた技術用語および科学用語は総て、本発明の属する技術分野の熟練者が本出願提出時に通常理解していたのと同じ意味を有する。下記の定義が本出願のいずれか他の箇所に示された「定義」と異なるか、またはそれよりも広義である場合は下記定義の方が優先する。
【００５１】
　本明細書で「核酸」および「複数の核酸」と言う時の核酸は通常、リン酸基、糖基、ならびにプリンおよびピリミジン塩基を有する鎖状高分子のことである。一般にリボ核酸（

ＲＮＡ）とデオキシリボ核酸（ＤＮＡ）との２型が有る。この術語は、ＤＮＡとＲＮＡとのハイブリッド（ＤＮＡ／ＲＮＡ）およびリボソームＤＮＡ（ｒＤＮＡ）を包含する。天然に含まれる塩基はアデニン、グアニン、シトシン、およびチミン（ＲＮＡではウラシル）である。合成の塩基、たとえばイノシンなども存在し、それらを代用して核酸プローブを創出してもよい。合成塩基とその利用は、当業者には良く知られている。

【００５２】

「少コピー数の核酸セグメント」と「少コピー数セグメント」とは同義であり、これらの語は「ユニーク」すなわち非反復性であるか、ほとんど唯一であるか、または当業者によって反復性として分類されないほど正常な染色体もしくはゲノム中に滅多に現われない、様々な長さの核酸配列を意味する。

【００５３】

「繰り返しＤＮＡ」、「反復配列」、およびこれらに類する語は、ゲノム中で繰り返されるＤＮＡ配列を意味する。高頻度繰り返しＤＮＡと名付けられたクラスは、伸長単鎖中で何千回も繰り返される５～１００ヌクレオチドの短鎖配列から成り、サテライトＤＮＡを含む。中頻度繰り返しＤＮＡと名付けられた別のクラスは、ゲノム全体に均一に分散した約１５０～３００ヌクレオチドの、より長い配列から成り、Ａｌｕ配列およびトランスポゾンと呼ばれる配列を含む。

「配列」と「セグメント」とは交換可能な語で、様々な長さの核酸断片を意味する。

【００５４】

本明細書中に用いた「ハイブリダイゼーション」という語は通常、２本の相補的なＲＮＡおよび／またはＤＮＡ１本鎖の対（緊密な物理的結合）を形成して２本鎖分子を得ることを意味する。ハイブリダイゼーション法には、標的核酸をその生物学的構造体から分離するマイクロアレイ、サザンブロット解析および定量的マイクロスフェアハイブリダイゼーションなどの固形支持体法と、標的核酸をその生物学的構造体、例えば細胞、組織、細胞核、染色体、または他の形態学的に認識される構造体から分離しない、細胞または染色体ベースの方法との両方が含まれる。

「ＰＣＲ」とはポリメラーゼ連鎖反応のことである。

【実施例】

【００５５】

以下の実施例によって、本発明の好ましい具体例を明示する。下記実施例に開示された方法は、本発明の発明者らが本発明の実施において十分に機能することを発見した方法を代表するものであり、従って本発明実施の好ましい態様を成すと見なされ得ることは、当業者には理解されるはずである。しかし、本出願の開示に照らせば、開示された具体例に多くの変更を加え、しかもなお同様の成果を上げることが本発明の思想および範囲を逸脱することなく可能であることも、当業者には理解されるはずである。

【実施例１】

【００５６】

本発明を、定量的マイクロスフェアハイブリダイゼーション（ＱＭＨ）および蛍光ｉｎ ｓｉｔｕハイブリダイゼーション（ＦＩＳＨ）を用いて試験した。

ＱＭＨ解析

ＨＯＸＢ１（ｃｈｒ１７：　４３９６４２６１－４３９６４３６０）（本出願では、座標参照番号は総て２００６年３月版ＵＣＳＣゲノムビルドによる）およびディジョージ症候群（ＤＧ）臨界領域（ｃｈｒ２２：　１９０７９５５７－１９０７９６５６）に特異的なユニーク配列プローブ（１００ｂｐ）をＵＧＳＨ法を用いて設計し、正常対照ゲノムＤＮＡからＰＣＲで合成した（プロメガ（Ｐｒｏｍｅｇａ））。各プローブ用の順方向プライマーは炭素数６の５’リンカー、続いてアミン基を用いて合成し（インビトロジェン（Ｉｎｖｉｔｒｏｇｅｎ））、得られたプローブをスペクトルの明らかなポリスチレンカルボキシル化マイクロスフェア（ルミネックス（Ｌｕｍｉｎｅｘ））に修飾カルボジイミド結合反応を介して結合させた（Ｎｅｗｋｉｒｋら、２００６年）。二つの異なるディジョージ症候群患者ゲノムＤＮＡ試料および一つの正常対照試料に関し、全ゲノム増幅を用い

たビオチン－１６－ｄＵＴＰの導入によりハイブリダイゼーション用の標的ＤＮＡを調製した。ビオチニル化したゲノムＤＮＡを切断して平均サイズ１ｋｂとし、ディジョージ症候群プローブおよびＨＯＸＢ１プローブを多重反応においてハイブリダイズさせた。試料をデュアルレーザーフローサイトメトリーによって解析し（ルミネックス）、各プローブについて平均蛍光強度（ＭＦＩ）比を得た。ディジョージ症候群患者試料（ＤＧ－１、ＤＧ－２）および正常対照試料のデータを下記に示す。

【００５７】

【表１】

表１

試料／プローブ	HOXB1 MFI	MFI 比	DG MFI	DG MFI 比
ＤＧ－１	123	1	65	0.53
ＤＧ－２	109	1	57	0.52
正常	173	1	171	0.99

【００５８】

ＨＯＸＢ１プローブの場合ＭＦＩ値は１２３であり、ディジョージ症候群プローブではＭＦＩ値は６５であった。ＭＦＩ比は０．５以下で、ＨＯＸＢ１プローブが２コピー存在するのに対しディジョージ症候群プローブは１コピーしか存在しないことを示しており、このことはディジョージ症候群患者ＤＮＡの実際の遺伝子型を反映するものである。この例はＵＧＳＨがユニーク配列領域の同定に成功したことを示し、なぜならＭＦＩ比が０．５を上回れば、ディジョージ症候群プローブが他のゲノム領域にハイブリダイズし、従ってユニーク配列のみからなるものではないことが示唆されることになったからである。ユニーク配列領域に特異的であるようには事実上設計されていない（すなわち、従来技術による方法で設計された）ＱＭＨプローブの中に、ゲノムに欠失領域の有る患者においてＭＦＩ比が０．５以下にならないものが有り、それらはＮｅｗｋｉｒｋら、２００６年に開示された（ヒトの突然変異）。

【００５９】

ＦＩＳＨ解析

本発明を用いて、ＦＩＳＨ解析用のユニーク配列プローブも設計した。ＢＡＣ　ＲＰ１１－６７７Ｆ１４（２０３ｋｂ；　７ｑ３１）に特異的なゲノム配列をＵＧＳＨにアップロードして（図１）プログラムを実行し、ユニーク配列プローブを表示させた（図２）。一つのプローブ（ｃｈｒ７：　１１５３６７６０２－１１５３７１２０１）と、対応するプライマー配列とをＵＧＳＨ出力から選択し、プライマーを合成した（インビトロジェン）。特異的ゲノム領域をＰＣＲで増幅した（プロメガ）。標準的な直接プローブ標識法（ミラス（Ｍｉｒｕｓ）社）を用い、プローブを正常ヒト対照染色体（中期および間期）に、ＦＩＳＨを用いてハイブリダイズさせた。単一のユニーク配列プローブが非常に明るく際だったハイブリダイゼーションシグナルを生成したが（図３）、このことは他のゲノム領域へのクロスハイブリダイゼーションが無かったことを示し、従って該プローブのユニーク配列設計を検証するものである。

【００６０】

図３は、ＵＧＳＨ法を用いて設計した、第７染色体上のＢＡＣ　ＲＰ１１－６７７Ｆ１４に由来するユニーク配列プローブを用いて行なったＦＩＳＨ実験で撮影した写真である。対照プローブとして、第７染色体のセントロメアに特異的であるＣｅｎ７プローブ（緑色；ビシス（Ｖｙｓｉｓ））を正常ヒト中期染色体伸展標本にハイブリダイズさせた。同時にＢＡＣ　ＲＰ１１－６７７Ｆ１４プローブ（赤色）もハイブリダイズさせた。この実験は、ＢＡＣ　ＲＰ１１－６７７Ｆ１４プローブが他の染色体領域に非特異的に結合することは一切なかったことを示しており、従って該プローブがユニークＤＮＡ配列のみからなることを実証し、ＵＧＳＨ法の有効性を保証するものである。

【００６１】

本発明の方法を、１回のＦＩＳＨ実験に併用される５種以上のユニーク配列プローブのみからなるユニーク配列プローブカクテルの創出にも適用してみた。ＵＧＳＨ法を用いて

設計した、第３染色体に特異的な５種のユニーク配列プローブの使用から得られた結果を図４に示す。各プローブをＰＣＲ増幅して直接標識し（赤色；ミラス社）、その後混合して対照プローブ（Ｃｅｎ７、緑色；ビシス）と共に正常ヒト中期染色体にコハイブリダイズさせた。このＦＩＳＨ実験では、ユニーク配列プローブカクテルに関して得られたハイブリダイゼーションシグナル強度が単一のユニーク配列プローブの場合（図３）と比較して格段に大きく、またバックグラウンド蛍光がきわめて弱かったために、位置の特定がより素早くかつ容易に可能であった。

【００６２】

上記のようなプローブカクテルは市販のＦＩＳＨプローブにとって、はるかに大きいサイズの現行ＦＩＳＨプローブ（～３００ｋｂ）にシグナルにおいて匹敵する点で理想的であろうが、ユニーク配列プローブカクテルはそのサイズが有意に小さい（合計で１０ｋｂ以下）ことで、染色体異常のより正確な診断を可能にする。上記実験によって、本発明の新規な方法がユニーク配列ＦＩＳＨプローブの設計に有用であることが実証されている。

【００６３】

ＵＧＳＨで設計したユニーク配列プローブを、シングルコピープローブ生成に有効な従来方法（例えば’０９７および’９９７特許）と比較した。ＦＩＳＨ実験において、ＵＧＳＨ法ではなく、’０９７および’９９７特許に示された方法を用いて設計したプローブを用いた。第９染色体に特異的なＤＮＡ配列内の反復配列を、周知の反復配列ファミリーおよびクラスを用いるホモロジー検索によってマスキングし（’０９７および’９９７特許）、その結果得られた、「シングルコピー」であるとされる一領域（ＡＢＬ１プローブ１６－１：　ＫｎｏｌｌおよびＲｏｇａｎ、２００３年）に対するプライマーを設計した。

【００６４】

ＦＩＳＨ実験の結果はプローブ（赤色）が染色体上の多数の位置にハイブリダイズしたことを示し、該プローブの配列は二つ以上の染色体領域と相同であり、従ってユニーク配列のみからなるものではないことを示唆している。このＦＩＳＨ実験では、第９染色体のセントロメアに特異的な対照プローブ（ＣＥＰ９；ビシス）をコハイブリダイズさせた。ＡＢＬ１プローブ配列自体をさらに解析したところ、プローブ配列の６１．９８％がＡｌｕ、ＬＩＮＥ１およびＬＩＮＥ２を含めた繰り返し要素からなることが判明した。それらの要素はそれぞれの原型繰り返し配列と僅かながら異なるため、その配列同定に反復配列マスキングでは十分でなかった。

【００６５】

この配列をＢＬＡＴで解析すると、ゲノム全体で１５０を越える数の一致が同定され、ＢＬＡＴスコアは大半が２１５から１００であった。これに対し、ＵＧＳＨ法の場合の好ましいカットオフＢＬＡＴスコアは２５で、このスコアであればユニーク配列プローブをきわめて厳密に選択することができる。ユニーク配列プローブ選択のためのカットオフ値が上記のように限局されることの結果がどのようなものかは、図３および４を図５と比較すれば明らかである。

【００６６】

図５は、ＵＧＳＨ法ではなく、’０９７および’９９７特許に示された方法を用いて設計したプローブを用いたＦＩＳＨ実験で撮影した写真である。第９染色体に特異的なＤＮＡ配列内の反復配列を、周知の反復配列ファミリーおよびクラスを用いるホモロジー検索によってマスキングし（’０９７および’９９７特許）、その結果得られた「シングルコピー」の一領域に対するプライマーを設計した。ＦＩＳＨ実験の結果はプローブ（赤色）が染色体上の多数の位置にハイブリダイズしたことを示し、該プローブの配列は二つ以上の染色体領域と相同であり、従ってユニーク配列のみからなるものではないことを示唆している。このＦＩＳＨ実験では、第９染色体のセントロメアに特異的な対照プローブ（ＣＥＰ９；ビシス）をコハイブリダイズさせた。

【００６７】

研究者の行なう特定の実験において、配列同定のための厳密なパラメーターや熱力学的

範囲の限局がさほど求められない場合、ユーザーはこれらの変数を変更してもよい。その結果、より多数の配列が同定されようが、それらの配列はゲノムハイブリダイゼーション実験において低い性能しか示さない恐れが有る。

【００６８】

ＵＧＳＨ法の用途には、何らかのゲノムハイブリダイゼーション実験に用いられるプローブの生成も含まれる。ＵＧＳＨはマイクロアレイおよびアレイＣＧＨ実験用のユニーク配列プローブ（６０〜７０塩基）を同定し得る。プローブが短いため、このような用途でプライマー配列が必要とされることはないであろうが、ＵＧＳＨは必要なユニーク配列領域を表示する。ＵＧＳＨ法の用途にはさらに、サザンおよびノーザンブロット解析、ｉｎ ｓｉｔｕハイブリダイゼーション、多重ライゲーション依存性プローブ増幅（ＭＬＰＡ）、ならびに多重増幅プローブハイブリダイゼーション（ＭＡＰＨ）が非限定的に含まれる。

【実施例２】

【００６９】

この実施例において、本発明の方法を用いて開発した多くのプローブを提供する。各プローブは子宮頸癌の危険性の評価に単独で、または少なくとも１種の他のプローブと組み合わせて用いることができる。プローブが標的核酸配列とハイブリダイズする場合は、対象配列が存在すると分かり、子宮頸癌発症の危険性は低い。しかし、ハイブリダイゼーションが起こらない場合は対象配列が欠失しているか、または突然変異してハイブリダイゼーションを妨げるようになっている。このような状況は、当該個体において子宮頸癌発症の危険度が増していることを示唆するものである。本発明のこの態様の幾つかの形態では、配列番号１〜３１からなる群から選択された単一のプローブがハイブリダイゼーションアッセイに用いられる。ハイブリダイゼーションが起こらないということからはまた、当該個体が子宮頸癌を発症する危険性は対象配列を含むゲノムを有する個体との比較のみならず一般の人々との比較においても高いことが推測される。他の好ましい形態では複数のプローブが併用される。さらに好ましくは、配列番号１〜２５、または配列番号２６〜３０からなる群から選択された少なくとも２種のプローブが一つの方法で用いられる。配列番号１〜２５から選択されるプローブは第３染色体（３ｑ２６）に由来し、配列番号２６〜３０から選択されるプローブは第７染色体に由来する。好ましい形態の幾つかでは、複数のプローブを含むプローブカクテルが用いられる。各プローブについて配列およびハイブリダイズする位置が既知であるので、いずれのプローブのハイブリダイゼーション（またはその欠如）からも無損傷性、もしくは変異していない配列と比べた時の変異状況に関して豊富な情報が得られ、そのような情報は総て子宮頸癌の検出や個体における危険度評価の一助となり得る。

【００７０】

配列番号３２〜４３も子宮頸癌の遺伝子マーカーに関わる。配列番号３２、３５、３８および４１のうちのいずれか一つ以上のハイブリダイゼーションが起こらないことは子宮頸癌発症の危険性の増大に関連付けられ、一方これらのプローブのいずれか一つのハイブリダイゼーションは正常な遺伝子配列が存在すること、および子宮頸癌発症の危険性が高まってはいないことを示唆する。配列番号３３および３４はそれぞれ配列番号３２の順方向および逆方向プライマーであり、配列番号３６および３７はそれぞれ配列番号３５の順方向および逆方向プライマーであり、配列番号３９および４０はそれぞれ配列番号３８の順方向および逆方向プライマーであり、配列番号４２および４３はそれぞれ配列番号４１の順方向および逆方向プライマーである。配列番号１〜３１の場合同様、配列番号３２、３５、３８および４１のプローブは個別に用いても互いに組み合わせて用いてもよく、さらには配列番号１〜３１のうちのいずれかと併用することも可能である。表２に、上記プローブそれぞれの座標を示す（２００６年３月版ＵＣＳＣゲノムビルドに拠る）。

【００７１】

【表２】

表２

プローブ名	始点座標*	終点座標	プローブサイズ	配列番号
染色体3q26プローブカクテル：全プローブを1反応で一つにプール				
RP11-641D5-8	170468591	170470501	1910	1
RP11-641D5-7	170472622	170474906	2284	2
RP11-641D5-6	170491470	170494165	2695	3
RP11-641D5-5	170495466	170498705	3239	4
RP11-641D5-4	170504182	170507036	2854	5
RP11-641D5-3	170513776	170515778	2002	6
RP11-641D5-2	170551404	170553206	1802	7
RP11-641D5-1	170564835	170568441	3606	8
RP11-3K16-5	170571082	170573293	2211	9
RP11-3K16-4	170616435	170618896	2461	10
RP11-3K16-3	170633935	170636538	2603	11
RP11-3K16-1	170702962	170704398	1436	12
RP11-816J6-1	170782158	170783927	1769	13
RP11-816J6-2	170811261	170813516	2255	14
RP11-362K14-3	170821049	170822942	1893	15
RP11-362K14-2	170824210	170827979	3769	16
RP11-362K14-1	170860403	170861821	1418	17
RP11-379K17-5	171017787	171020006	2219	18
RP11-379K17-4	171031245	171034304	3059	19
RP11-379K17-3	171131084	171135002	3918	20
RP11-379K17-2	171135323	171138745	3422	21
RP11-379K17-1	171138881	171142114	3233	22
RP13-8108-1	171140257	171142304	2047	23
RP13-8108-2	171166207	171168262	2055	24
RP13-8108-3	171209493	171210861	1368	25
染色体7プローブカクテル：全プローブを1反応で一つにプール				
BAC667F14-1	115561346	115564397	3051	26
BAC667F14-2	115597264	115601247	3984	27
BAC667F14-3	115667956	115669681	1950	28
BAC667F14-4	115676311	115678653	2343	29
BAC667F14-5	115685858	115688020	2162	30
BAC667F14-6	115698372	115700626	2254	31

*：2006年3月版ＵＣＳＣゲノムビルド

【００７２】

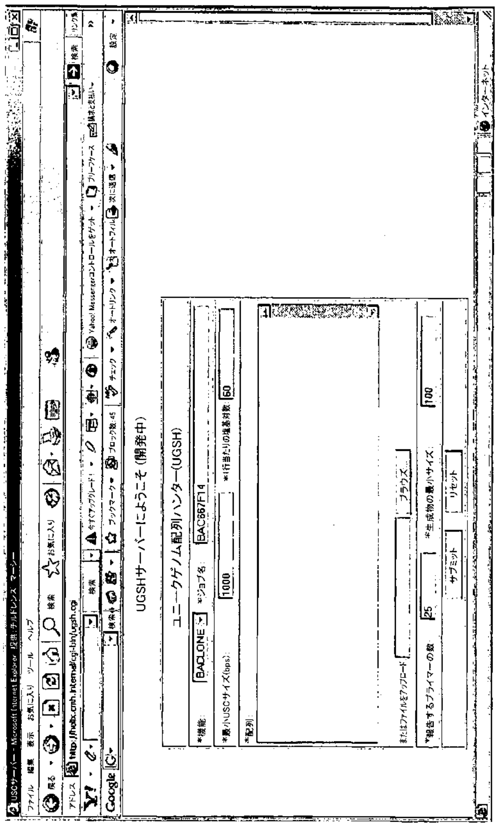
最後に、本発明に従って開発されたプローブは量子マイクロスフェアハイブリダイゼーションアッセイに特に好適に用いられる。好ましいプローブには、本発明が配列番号４４～５７として提供するものが含まれる。これらのプローブは各々、自身の由来する病原体の存在を検出するのに単独で用いられる。配列番号４４はマイコプラズマ属（Ｍｙｃｏｐｌａｓｍａ）ＦＲＸ　Ａ遺伝子に由来する（属特異的）。具体的には、配列番号４５のハイブリダイゼーションはＭ．　ｆｅｒｍｅｎｔａｎｓの存在を示し、配列番号４６のハイブリダイゼーションはＭ．　ｍｏｌｌｉｃｕｔｅｓの存在を示し、配列番号４７のハイブリダイゼーションはＭ．　ｈｏｍｉｎｉｓの存在を示し、配列番号４８のハイブリダイゼーションはＭ．　ｈｙｏｒｈｉｎｉｓの存在を示し、配列番号４９のハイブリダイゼーションはＭ．　ａｒｇｉｎｉｎｉの存在を示し、配列番号５０のハイブリダイゼーションはＭ．　ｏｒａｌｅの存在を示し、配列番号５１のハイブリダイゼーションはＡｃｈｅｏｐｌａｓｍａ　ｌａｉｄｌａｗｉｉの存在を示し、配列番号５２のハイブリダイゼーションはＭ．　ｓａｌｉｖａｒｉｕｍの存在を示し、配列番号５３のハイブリダイゼーションはＭ．　ｐｕｌｍｏｎｉｓの存在を示し、配列番号５４のハイブリダイゼーションはＭ．　ｐｎｅｕｍｏｎｉａｅの存在を示し、配列番号５５のハイブリダイゼーションはＭ．　ｐｉｒｕｍの存在を示し、配列番号５６のハイブリダイゼーションはＭ．　ｃａｐｒｉｃｏｌｕｍの存在を示し、配列番号５７のハイブリダイゼーションはＨｅｌｉｃｏｂａｃｔｅ

ｒ　ｐｙｌｏｒｉの存在を示す。

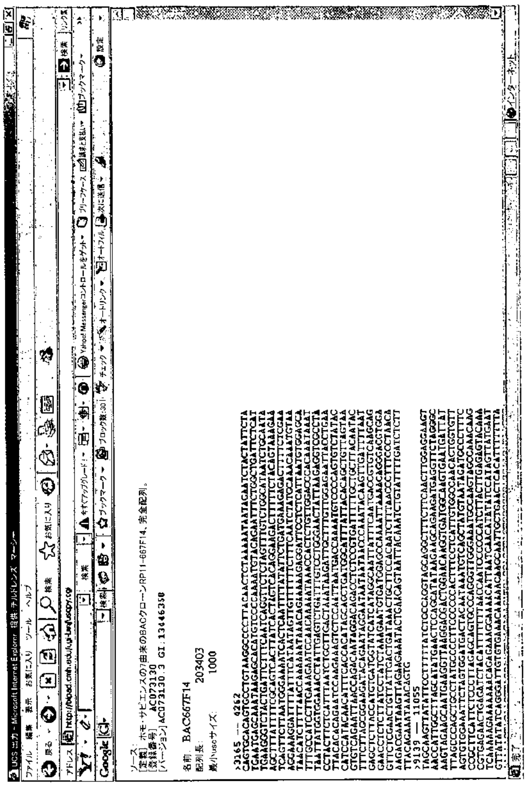
【００７３】

本明細書に開示され、また特許請求された組成物および方法はいずれも、本発明の開示に照らせば甚だしい実験を行なわなくとも製造および実施することができる。本発明の組成物および方法を好ましい具体例の観点から説明したが、該組成物および方法に、また本明細書に記載した方法の個々のステップや一連のステップに変更を加えることが本発明の思想、趣旨および範囲を逸脱することなく可能であることは、当業者には明白であろう。具体的には、本明細書に述べた薬剤に替えて化学的にも生理学的にも関連する所与の薬剤を用いてもよく、その際同等または類似の結果が得られるであろうことは明らかである。これと同様の、当業者にとって明らかな代替および変形は総て、以下に記す特許請求の範囲に規定された本発明の趣旨、範囲および思想を逸脱しないと考えられる。

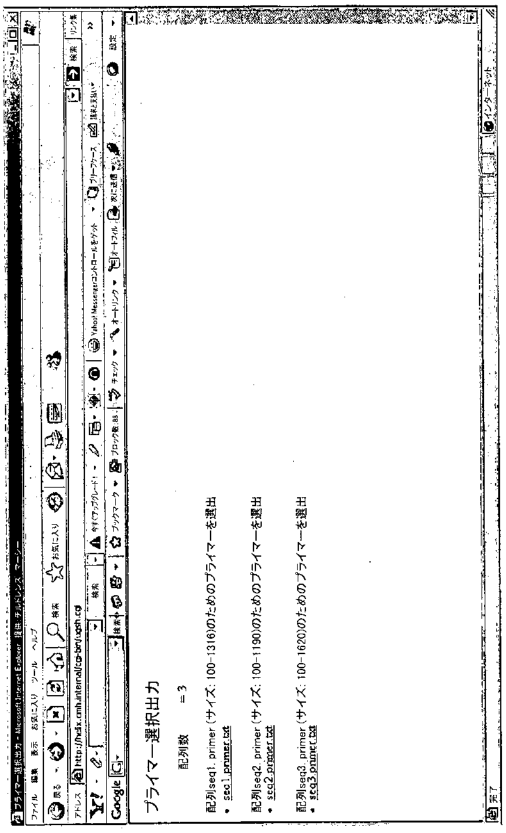
【図１】



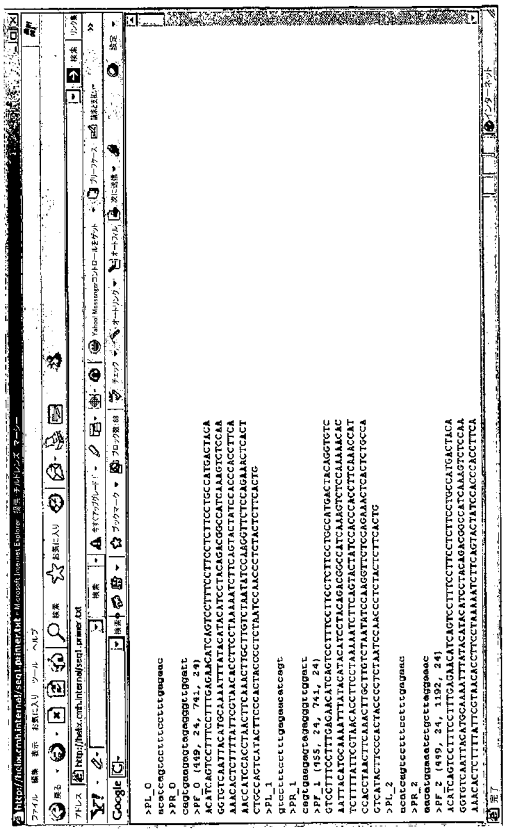
【図２Ａ】



【図２Ｂ】

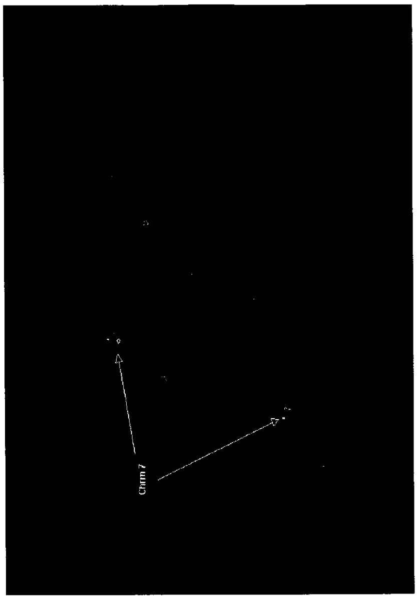


【図２Ｃ】



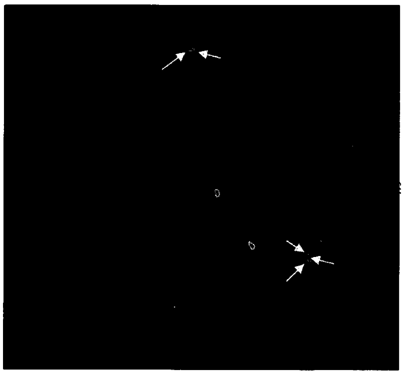
【図３】

FIG. 3



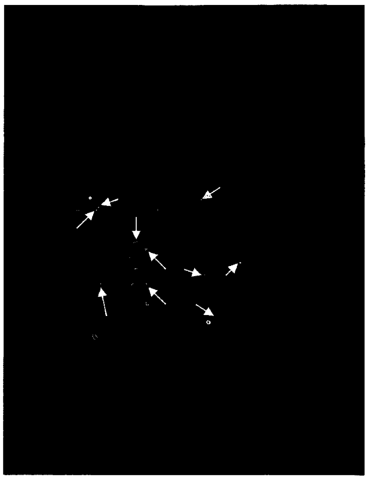
【図４】

FIG.4

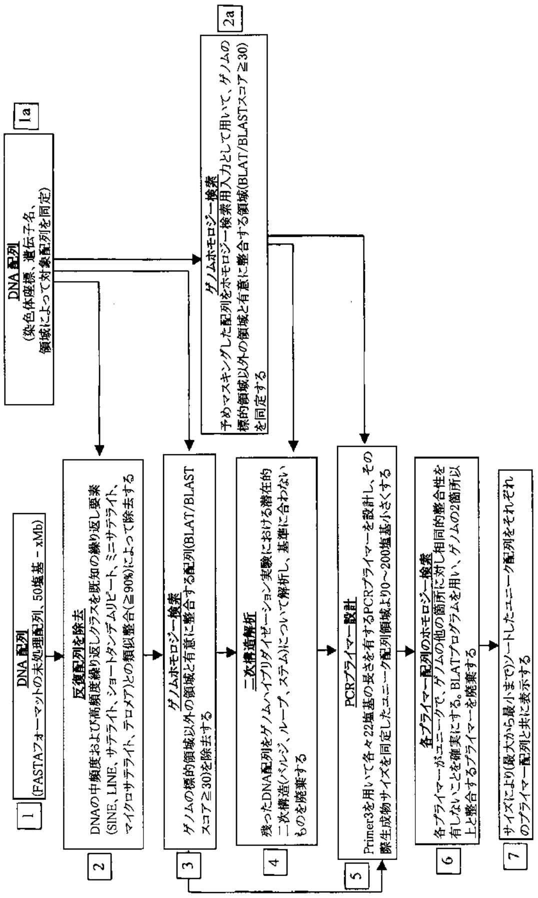


【図５】

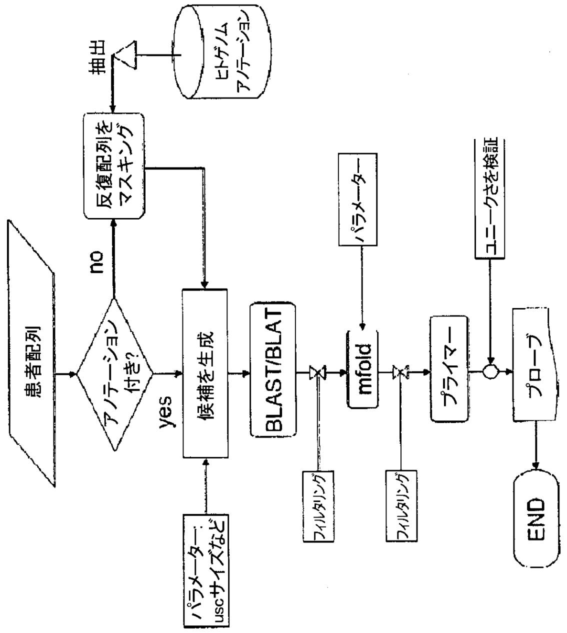
FIG. 5



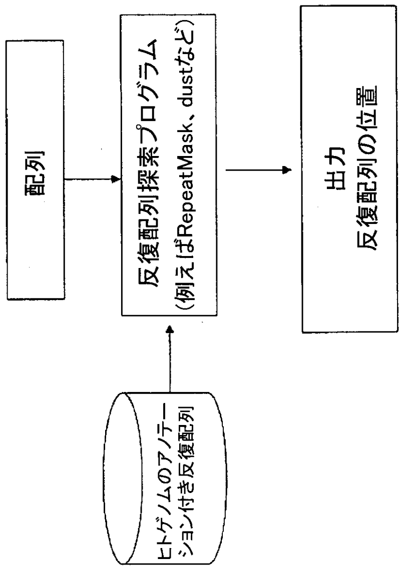
【図６】



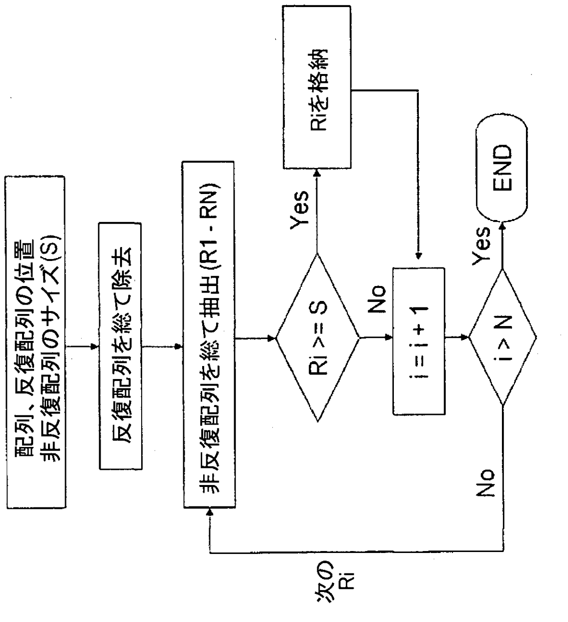
【図７】



【図８】



【図９】



【配列表】
2010522571000001.xml

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/58795

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12Q 1/68; C12P 19/34; C07H 21/04 (2008.04)
USPC - 435/6, 435/91.2, 536/24.3
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
435/6, 435/91.2, 536/24.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest: US Pre-Grant Publication Full-Text, US Patents Full-Text, EPO Abstracts, JPO Abstracts. ("low copy" or unique or particular or single) near3 sequence; (mfold or blat or blast or primer3 or repeatmasker); @pd<20070328; oligonucleotide near3 (selection or design)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/0059743 (Maass Sepulveda et al.); Paras. [0020], [0060], [0061], [0065], [0090], [0092], [0098], [0102]-[0111], Figure 1, Table 3	1-14
A	US 2006/0141501 A1 (Friend et al.) [0071] - [0168]	1-14
A	US 2006/0110744 A1 (Sampas et al.) [0022] - [0039]	1-14
A	US 2003/0108919 A1 (Kautzer et al.) [0037]-[0116]	1-14
A	US 2003/0104470 A1 (Fors et. al) [0172] - [0202]	1-14

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 August 2008 (14.08.08)	22 AUG 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/58795

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.
Group 1: Claims 1-14 are directed toward a method for identifying a low copy nucleic acid segment and a method of selecting probes used for hybridization experiments.
Groups 2-58: Claim 15 is directed toward a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1-57.
The inventions listed as Groups 1 - 58 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:
There is no common technical feature shared between the listed groups. The method of Group I does not share any special technical feature with the nucleic acid sequence of any of Groups 2-58, and vice versa. The different sequences represented by the nucleotide content of Groups 2-58 (SEQ ID: NOs 1-57 respectively) are different structures that are not common to one another but are different because they are composed of unique nucleic acid sequences.

Therefore, the inventions listed as Groups 1 - 58 lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-14

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MT,NL,NO,PL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BR,BW,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PG,PH,PL,PT,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,SV,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(72)発明者 ビ チェンペン
アメリカ合衆国 ６４１０８ ミズーリ州 カンザス シティ ギルハム ロード ２４０１

Ｆターム(参考) 4B024 AA12 AA13 AA20 CA02 CA09 HA14
4B063 QA13 QQ42 QQ52 QR32 QR35 QR55 QS25 QS34 QX01