

(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2014-507164
(P2014-507164A)

(43) 公表日 平成26年3月27日(2014.3.27)

(51) Int.Cl.	F 1	テーマコード (参考)
C 12 N 15/09 (2006.01)	C 12 N 15/00	A 2 G 04 5
C 12 Q 1/68 (2006.01)	C 12 Q 1/68	A 4 B 02 4
G 01 N 37/00 (2006.01)	C 12 Q 1/68	Z 4 B 06 3
G 01 N 33/50 (2006.01)	C 12 N 15/00	F
	GO 1 N 37/00	1 O 2

審査請求 未請求 予備審査請求 未請求 (全 35 頁) 最終頁に続く

(21) 出願番号	特願2013-555620 (P2013-555620)	(71) 出願人	513212280 イルミナ, インコーポレイテッド アメリカ合衆国, カリフォルニア州 92 122, サン デイエゴ, 5200 イル ミナ ウェイ
(86) (22) 出願日	平成24年2月24日 (2012. 2. 24)	(74) 代理人	100114775 弁理士 高岡 亮一
(85) 翻訳文提出日	平成25年10月8日 (2013. 10. 8)	(74) 代理人	100121511 弁理士 小田 直
(86) 國際出願番号	PCT/US2012/026623	(74) 代理人	100191086 弁理士 高橋 香元
(87) 國際公開番号	W02012/116331	(72) 発明者	ファン, ジャンービン アメリカ合衆国, カリフォルニア州 92 122, サン デイエゴ, 5200 イル ミナ ウェイ
(87) 國際公開日	平成24年8月30日 (2012. 8. 30)		
(31) 優先権主張番号	61/446, 890		
(32) 優先日	平成23年2月25日 (2011. 2. 25)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	61/509, 960		
(32) 優先日	平成23年7月20日 (2011. 7. 20)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

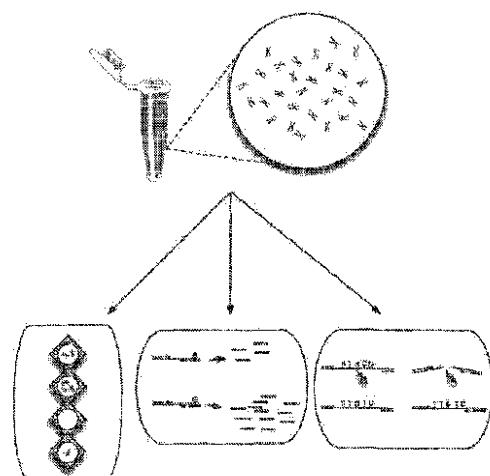
(54) 【発明の名称】ハプロタイプ決定のための方法およびシステム

(57) 【要約】

本開示の実施形態は、生体試料のハプロタイプを決定するための方法およびシステムを提供する。特定の実施形態は、ゲノムの長距離ハプロタイプ決定方法を提供する。

【選択図】図1

FIGURE 1



【特許請求の範囲】**【請求項 1】**

母系染色体と父系染色体との寄与が不均等な核酸試料の1つ以上の画分を用意する工程と、核酸試料の前記1つ以上の画分中の2つ以上の目的配列間の不均衡を検出する工程と、前記検出可能な不均衡に基づき前記核酸試料のハプロタイプを決定する工程とを含む、核酸試料のハプロタイプの決定方法。

【請求項 2】

前記核酸試料はゲノムもしくはその断片に由来する、請求項1に記載の方法。

【請求項 3】

前記ゲノムは1つ以上の細胞に由来する、請求項2に記載の方法。

10

【請求項 4】

前記1つ以上の細胞は約10～100個の細胞である、請求項3に記載の方法。

【請求項 5】

前記核酸試料は哺乳類に由来する、請求項1に記載の方法。

【請求項 6】

前記哺乳類はヒトである、請求項5に記載の方法。

【請求項 7】

前記母系および父系の染色体は、一塩基多型、コピー数多型、ゲノムの挿入およびゲノムの欠失からなる群から選択される1つ以上の変異型配列を含む、請求項1に記載の方法。

20

【請求項 8】

母系染色体と父系染色体との前記不均等な寄与は、1:1の比以外の染色体の比を含む、請求項1に記載の方法。

【請求項 9】

前記ハプロタイプを蛍光によって決定する、請求項1に記載の方法。

【請求項 10】

前記ハプロタイプを核酸配列決定法によって決定する、請求項1に記載の方法。

【請求項 11】

前記ハプロタイプをマイクロアレイ上で行われる遺伝子型決定法によって決定する、請求項1に記載の方法。

30

【請求項 12】

前記ハプロタイプを定量的ポリメラーゼ連鎖反応によって決定する、請求項1に記載の方法。

【請求項 13】

ハプロタイプ決定のための画分の調製方法であって、

a) 前記試料にとって自然な比の母系および父系の染色体成分を含む核酸試料を用意する工程と、

b) 複数の画分を作成する工程であって、1つ以上の画分が非対称な比の母系および父系の染色体成分を含み、前記非対称な比が前記個体にとって自然な比とは実質的に異なり、それによりハプロタイプ決定のための画分を調製する工程と、
を含む方法。

40

【請求項 14】

前記作成する工程は、母系および父系の染色体成分を複数の画分中の1つ以上の画分に非対称に分配させる工程を含む、請求項13に記載の方法。

【請求項 15】

前記作成する工程は、前記複数の画分のうちの1つ以上の画分中の前記母系もしくは父系の染色体成分の1種以上を示差的に分解することを含む、請求項13に記載の方法。

【請求項 16】

前記作成する工程は、前記複数の画分のうちの1つ以上の画分中の前記母系もしくは父系の染色体成分の1つを示差的に増幅させる、請求項13に記載の方法。

50

【請求項 17】

前記核酸試料は哺乳類に由来する、請求項13に記載の方法。

【請求項 18】

哺乳類はヒトである、請求項17に記載の方法。

【請求項 19】

前記核酸試料は複数の細胞に由来する、請求項13に記載の方法。

【請求項 20】

前記複数の細胞を中期で同期させる、請求項19に記載の方法。

【請求項 21】

前記複数の細胞は約5～約300個の細胞である、請求項19に記載の方法。

10

【請求項 22】

前記複数の細胞は約10～約100個の細胞である、請求項19に記載の方法。

【請求項 23】

試料中の複数の目的配列のハプロタイプを決定する方法であって、

請求項13に記載の1つ以上の画分を用意する工程と、

前記1つ以上の画分からライブラリーを作成する工程と、

前記複数の目的配列の検出可能な信号を検出する工程と、

20

前記検出可能な信号の差に基づいて前記複数の目的配列のハプロタイプを決定する工程と、

を含む方法。

【請求項 24】

前記2つ以上の目的配列は同じ染色体上にある、請求項23に記載の方法。

【請求項 25】

前記2つ以上の目的配列は、同じ染色体上の2つ以上の異なる遺伝子座に位置している、請求項23に記載の方法。

【請求項 26】

同じ染色体上の前記2つ以上の異なる遺伝子座は、少なくとも10000個のヌクレオチドだけ離れている、請求項24に記載の方法。

【請求項 27】

前記2つ以上の異なる遺伝子座は、同じ染色体上に位置しており、少なくとも10000個のヌクレオチドだけ離れている、請求項24に記載の方法。

30

【請求項 28】

前記2つ以上の異なる遺伝子座は、同じ染色体上に位置しており、少なくとも10000個のヌクレオチドだけ離れている、請求項24に記載の方法。

【請求項 29】

前記2つ以上の異なる遺伝子座は、同じ染色体上に位置しており、少なくとも20000個のヌクレオチドだけ離れている、請求項24に記載の方法。

【請求項 30】

前記1つ以上の画分は個体生物に由来する、請求項23に記載の方法。

【請求項 31】

前記1つ以上の画分は哺乳類に由来する、請求項23に記載の方法。

40

【請求項 32】

前記1つ以上の画分はヒトに由来する、請求項23に記載の方法。

【請求項 33】

工程b)の前に、母系と父系との染色体の比を測定する工程をさらに含む、請求項23に記載の方法。

【請求項 34】

前記ハプロタイプを決定する前記工程は、画分の定量的ポリメラーゼ連鎖反応分析を含む、請求項23に記載の方法。

【請求項 35】

50

前記ハプロタイプを決定する前記工程は、画分のマイクロアレイ分析を含む、請求項 23 に記載の方法。

【請求項 3 6】

前記ハプロタイプを決定する前記工程は、前記複数の目的配列のそれぞれについて配列読み取り数の差を検出する工程と、同様な配列読み取り数を有する前記目的配列を組み合わせる工程と、前記組み合わせた目的配列に基づいてハプロタイプを決定する工程とを含む、請求項 23 に記載の方法。

【請求項 3 7】

前記検出可能な信号は蛍光である、請求項 23 に記載の方法。

【請求項 3 8】

前記検出可能な信号は蛍光である、請求項 36 に記載の方法。

【請求項 3 9】

前記 2 つ以上の目的配列は、対立遺伝子、一塩基多型、コピー数多型、ゲノムの挿入およびゲノムの欠失を含む群から選択される、請求項 23 に記載の方法。

【請求項 4 0】

前記検出する工程は核酸配列決定法を含む、請求項 23 に記載の方法。

【請求項 4 1】

前記検出する工程は、マイクロアレイ上で行われる遺伝子型決定法を含む、請求項 23 に記載の方法。

【請求項 4 2】

前記検出する工程は、定量的ポリメラーゼ連鎖反応遺伝子型決定法を含む、請求項 23 に記載の方法。

【請求項 4 3】

前記配列決定法は、前記複数の目的配列における総読み取り数のうちの前記複数の目的配列の読み取り数の差を検出する、請求項 40 に記載の方法。

【請求項 4 4】

複数の読み取りを検出する工程は、前記複数の目的配列において生成された蛍光信号の数を検出することを含む、請求項 43 に記載の方法。

【請求項 4 5】

複数の遺伝子座における対立遺伝子の相を検出する方法であって、

a) 核酸分子の非対称な分配を行う工程であって、前記非対称な分配は複数の画分を含み、前記個々の画分は前記対立遺伝子の複数のコピーを含み、かつ前記個々の画分は異なる量の前記対立遺伝子を含むことを特徴とする工程と、

b) 1 つ以上の個々の画分中に存在する前記核酸分子の前記コピー中の前記対立遺伝子を識別する工程と、

c) 前記 1 つ以上の個々の画分中に存在する前記異なる量の前記対立遺伝子を評価する工程と、

d) 前記対立遺伝子を識別する工程および前記異なる量の前記対立遺伝子を評価する工程から、複数の遺伝子座における前記対立遺伝子の相を決定する工程と、
を含む方法。

【請求項 4 6】

前記評価する工程は、総読み取り数のうちの複数の遺伝子座における前記対立遺伝子の蛍光配列決定読み取り数の差を検出することを含む、請求項 45 に記載の方法。

【請求項 4 7】

前記核酸分子は個体生物に由来する、請求項 45 に記載の方法。

【請求項 4 8】

前記異なる量を評価する工程は、前記複数の遺伝子座における対立遺伝子比を測定することを含む、請求項 45 に記載の方法。

【請求項 4 9】

前記対立遺伝子を識別する工程は、前記複数の遺伝子座に存在する 1 つ以上のヌクレオ

10

20

30

40

50

チドの同一性を決定することを含む、請求項 4 5 に記載の方法。

【請求項 5 0】

前記対立遺伝子を識別する工程は核酸配列決定法を含む、請求項 4 5 に記載の方法。

【請求項 5 1】

前記対立遺伝子を識別する工程は、マイクロアレイ上で行われる遺伝子型決定法を含む、請求項 4 5 に記載の方法。

【請求項 5 2】

前記複数の遺伝子座は、同じ染色体上に位置しており、少なくとも 1 0 0 0 0 個のヌクレオチドだけ離れている、請求項 4 5 に記載の方法。

【請求項 5 3】

前記複数の遺伝子座は、同じ染色体上に位置しており、少なくとも 1 0 0 0 0 0 個のヌクレオチドだけ離れている、請求項 4 5 に記載の方法。

【請求項 5 4】

前記複数の遺伝子座は、同じ染色体上に位置しており、少なくとも 1 0 0 0 0 0 0 0 個のヌクレオチドだけ離れている、請求項 4 5 に記載の方法。

【請求項 5 5】

前記複数の遺伝子座は、同じ染色体上に位置しており、少なくとも 2 0 0 0 0 0 0 0 個のヌクレオチドだけ離れている、請求項 4 5 に記載の方法。

【請求項 5 6】

ハプロタイプを決定するための核酸画分であって、前記核酸画分は、非対称に分配された母系および父系の染色体成分を含み、前記非対称に分配された染色体成分は、母系対父系の比が、前記個体の自然な前記比とは異なる非対称な比の染色体成分である核酸画分。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本出願は、2011年2月25日に出願された米国仮特許出願第 6 1 / 4 4 6 , 8 9 0 号および2011年6月20日に出願された米国仮特許出願第 6 1 / 5 0 9 , 9 6 0 号の優先権を主張するものであり、それらの両開示内容全体が参照により本明細書に組み込まれる。

【背景技術】

【0 0 0 2】

ヒトゲノム計画の取り組みにより、ヒトの遺伝暗号を幅広く知る機会が開かれた。ヒトゲノムをさらに解読するための研究が、例えば高処理配列決定法を用いて現在進行している。H a p M a p (ハプロタイプマップ) 計画は、特定の疾患に罹患していない人々とその疾患に罹患している人々からのゲノム情報を比較して、疾患に繋がる遺伝的変異を発見することを目的とした世界的規模の科学的取り組みである。特定の遺伝子の対立遺伝子すなわちD N A配列の1つ以上の形態は、1つ以上の異なる遺伝的変異を含む可能性があり、ハプロタイプ、すなわち特定の染色体上の異なる位置または遺伝子座にある対立遺伝子の組み合わせを同定することが、H a p M a p 計画の主眼である。2つの群が異なることが同定されたハプロタイプは、疾患を引き起こす遺伝子異常の位置と相關している可能性がある。従って、H a p M a p の結果は、ヒトにおける遺伝的変異の一般的なパターンおよびそれらの変異が疾患に相關している可能性を秘めているか否かを説明するのに役立つ。

【0 0 0 3】

これらの取り組みから得られる情報は、配列が不完全であったり、ギャップや時には誤りがあつたりすることもあるが、疾患や障害の裏にある遺伝的性質の解読を支援する有益な手段を提供する。残念なことに、そのような大規模な配列決定を行うコストはまだ非常に高く、単一染色体のハプロタイプ決定すなわち対立遺伝子もしくは標的配列の相の特定などのより深い情報を提供するための技術は幻想である。ヒトゲノムからより多くの情報を解読するためのさらなる手段および技術が必要である。

10

20

30

40

50

【発明の概要】

【発明が解決しようとする課題】

【0004】

現在の遺伝子型決定法は、研究者に研究対象の遺伝子構造を提供することができる。しかし、上記遺伝子型決定法は、ある染色体上でどの配列が互いに隣接または近接しているかを、別の染色体上で隣接または近接するものと対比して決定する好都合で測定可能な手段に関して制限されている。図2は、研究対象の遺伝子型を決定することはできても、目的配列（例えば、対立遺伝子、一塩基多型（S N P）、コピー数多型（C N V）、遺伝子の挿入や欠失（i n d e l）など）が、別の目的配列と同じ染色体上に位置しているか否かを決定するのに十分な情報は得られないというジレンマを例示している。例えば、研究対象からの試料中の染色体の混合集団（図2A）について、データから例示的な遺伝子型（図2B）を決定することは可能かもしれない。しかし、ヘテロ接合型対立遺伝子が、ある染色体上でどのように組み合わされているのかを決定（ハプロタイプ決定）するのに十分な情報は得られない。例えば、親A（P_a）が対立遺伝子'および'、親B（P_b）が対立遺伝子'および'を提供しているか否か（図2C）、あるいはそれらが混合されているか否か（図2D）は不明である。目的配列が染色体上で互いに遠く離れて、遠位に、あるいは長距離に位置している場合、どの目的配列が同じ染色体上に存在しているのかを決定し、それにより染色体の長距離ハプロタイプを決定すること、すなわち対立遺伝子の相を特定することはなあさらに難しい。

10

【0005】

20

本開示の実施形態は、染色体上の互いに対するそれらの位置（例えば、近位または遠位）に関わらず、相が特定された対立遺伝子を決定するための新規な解決法を提供する。現在のハプロタイプ決定の課題を解決することを目的とした実験中に、遺伝物質の不均衡すなわち非対称な分配を行うことにより、研究対象の正確なハプロタイプ決定の問題に対する新規な解決法が得られることを発見した。不均衡な分配後に標的配列を任意に増幅することが特に有用である。本発明は、特定の機序に限定されない。実際に、本発明を実施するのに、機序の理解は必要ではない。それにも関わらず、不均衡な材料の示差的増幅に部分的に基づいて、増幅信号強度により染色体のハプロタイプを決定することを想定している。例えば、異なる対立遺伝子の信号の比により、どちらが单一染色体上に認められるかを決定し、それにより試料の相が特定されたハプロタイプを決定する。図3は、そのような一実施形態を例示している。最初の試料分配時の不均衡（図3Bおよび図3Dに示すような不均衡）を利用し、示差的増幅により、対立遺伝子がP_aの'対立遺伝子と同じ相にある、すなわちそれと組み合せられていること、および'がP_b上の'と同じ相にあること（図3E）を示している。さらに、実施形態は、一倍体試料に限定されず、それどころか、二倍体試料（例えば、対合染色体、D N A挿入断片、Y A C、B A C、コスミド、フォスミドなど）または一倍体試料（例えば、精子、卵子、全胞状奇胎などからの遺伝的相補体）のどちらか一方を利用すると効果的である。

30

【0006】

30

本明細書に記載の方法を実施することによって得られるゲノム中の対立遺伝子の相の特定から得られる情報は、一般的調査および発見努力ならびに、例えば、疾患検出、治療学、および移植拒絶反応を減少させるためのH L A適合性におけるより高い信頼において有用性が見出される。例えば、公知のハプロタイプを、ごく少ない例を挙げると、薬物代謝、創薬、病状、癌、障害、移植拒絶反応リスクおよび個別の健康管理戦略と相關させることができる。実際に、個別の健康管理に関して、研究対象の個体のハプロタイプ、次いで研究対象の特異的疾患との相関が分かると、治療の選択肢を具体的に設計して、その研究対象の必要性に応じることができる。

40

【課題を解決するための手段】

【0007】

50

本開示の一実施形態は、核酸試料中の2つ以上の目的配列間に検出可能な不均衡を含む試料の画分を得て、前記検出可能な不均衡に基づき核酸試料のハプロタイプを決定するこ

とによる、核酸試料のハプロタイプを決定する方法を含む。いくつかの実施形態では、核酸試料は、ゲノムまたはその断片からのものであり、ここで、前記ゲノムは、1つ以上の細胞、例えば約1～100個の細胞に由来している。いくつかの実施形態では、核酸試料は、哺乳類、好ましくはヒトに由来する。他の実施形態では、核酸試料は、非ヒト哺乳類、植物またはウイルスに由来する。いくつかの実施形態では、核酸試料は、目的配列に野生型配列を含み、他の実施形態では、核酸試料は、目的配列に変異型配列を含む。いくつかの実施形態では、目的配列は、ある目的配列に野生型配列を含み、別の目的配列に変異型配列を含むか、あるいはそれらの組み合わせを含む。いくつかの実施形態では、変異型配列は、一塩基多型、コピー数多型、ゲノムの挿入およびゲノムの欠失を含む群から選択される。いくつかの実施形態では、試料中の2つ以上の目的配列間の検出可能な不均衡を蛍光によって測定する。いくつかの実施形態では、試料中の2つ以上の目的配列間の検出可能な不均衡を、核酸配列決定法、例えばマイクロアレイ上で行われる遺伝子型決定法、または定量的ポリメラーゼ連鎖反応によって測定する。

10

20

30

40

【0008】

本開示の一実施形態は、染色体成分を含む核酸試料を用意し、かつ染色体成分を複数の画分中に非対称に分配し、それによりハプロタイプ決定のための画分を調製することを含む、ハプロタイプ決定のための画分の調製方法を含む。いくつかの実施形態では、染色体成分の非対称な分配は、不均等な量の染色体成分を複数の画分の異なる画分まで送達することを含む。いくつかの実施形態では、非対称に分配された染色体成分の比は、最初の細胞集団中の染色体成分の比と同じではない。いくつかの実施形態では、染色体成分の非対称な分配は、複数の画分の異なる画分中の染色体成分を示差的に分解することを含む。いくつかの実施形態では、染色体成分の非対称な分配は、複数の画分の異なる画分中の染色体成分を示差的に増幅させることを含む。いくつかの実施形態では、核酸試料は、哺乳類、好ましくはヒトに由来する。他の実施形態では、核酸試料は、非ヒト哺乳類、植物またはウイルスに由来する。いくつかの実施形態では、核酸試料は、複数の細胞、例えば約5～300個の細胞または約10～100個の細胞に由来する。いくつかの実施形態では、複数の細胞は、中期で同期されているが、他の実施形態では、複数の細胞は、中期に同期されていない。いくつかの実施形態では、染色体成分は、異なる遺伝子座に2つ以上の対立遺伝子を含み、対立遺伝子は、1つ以上の目的配列をさらに含む。

30

【0009】

本開示の一実施形態は、画分中の染色体成分が非対称に分配されている画分を用意し、画分からライブラリーを作成し、ライブラリー中の2つ以上の目的配列の検出可能な信号を検出し、かつ検出可能な信号における前記差に基づいて2つ以上の目的配列の相を決定することを含む、2つ以上の目的配列の相を決定する方法を含む。いくつかの実施形態では、検出可能な信号は、蛍光信号である。いくつかの実施形態では、2つ以上の目的配列は、同じ染色体上にあり、さらに、同じ染色体上の2つ以上の異なる遺伝子座に位置している。いくつかの実施形態では、同じ染色体上に位置する2つ以上の異なる遺伝子座は、少なくとも10,000個、少なくとも100,000個、少なくとも100,000,000個、または少なくとも200,000,000個のヌクレオチドだけ離れている。いくつかの実施形態では、画分は、個体生物に由来する。いくつかの実施形態では、画分は、哺乳類、例えばヒトに由来する。他の実施形態では、画分は、非ヒト哺乳類、植物またはウイルスに由来する。いくつかの実施形態では、相の決定のための画分を用意する前に、画分中の2つ以上の目的配列間の非対称性の程度を測定する。いくつかの実施形態では、非対称性の程度は、画分の定量的ポリメラーゼ連鎖反応分析を含む。いくつかの実施形態では、非対称性の程度の測定は、画分のマイクロアレイ分析を含む。いくつかの実施形態では、非対称性の程度の測定は、画分中の2つ以上の目的配列間の信号雑音比を測定することを含む。いくつかの実施形態では、画分中の2つ以上の目的配列間の信号雑音比は、他の画分中の信号雑音比よりも大きい。いくつかの実施形態では、信号雑音比を、蛍光検出によって測定する。

40

【0010】

50

本開示の一実施形態は、2つ以上の異なる遺伝子座に対立遺伝子を含む核酸分子の非対称な分配を行う工程であって、非対称な分配は複数の画分を含み、個々の画分は、対立遺伝子の複数のコピーを含み、かつ個々の画分は、異なる量の対立遺伝子を含むことを特徴とする工程と、1つ以上の個々の画分中に存在する核酸分子のコピー中の対立遺伝子を識別する工程と、1つ以上の個々の画分中に存在する異なる量の対立遺伝子を評価する工程と、対立遺伝子を識別する工程および異なる量の対立遺伝子を評価する工程から、2つ以上の異なる遺伝子座にある対立遺伝子の相を決定する工程とを含む、2つ以上の異なる遺伝子座にある対立遺伝子の相を決定する方法を含む。いくつかの実施形態では、評価する工程は、2つ以上の異なる遺伝子座にある対立遺伝子の総読み取り数のうちの2つ以上の異なる遺伝子座にある対立遺伝子の蛍光配列決定読み取り数の差を検出することを含む。いくつかの実施形態では、核酸分子の非対称な分配は、個体生物に由来する。いくつかの実施形態では、異なる量の対立遺伝子を評価する工程は、2つ以上の異なる遺伝子座における対立遺伝子の比を測定することを含む。いくつかの実施形態では、異なる量を評価する工程は、2つ以上の異なる遺伝子座における対立遺伝子を計数することを含む。いくつかの実施形態では、対立遺伝子を識別する工程は、核酸配列決定法を含み、他の実施形態では、対立遺伝子を識別する工程は、マイクロアレイ上で行われる遺伝子型決定法を含む。特定の場合には、核酸配列決定法およびアレイ系遺伝子型決定法を使用することができる。いくつかの実施形態では、2つ以上の異なる遺伝子座は、同じ染色体上にあり、少なくとも10,000個のヌクレオチドだけ離れている。いくつかの実施形態では、同じ染色体上に位置する2つ以上の異なる遺伝子座は、少なくとも100,000個、少なくとも100,000,000個または少なくとも200,000,000個のヌクレオチドだけ離れている。10

【図面の簡単な説明】

【0011】

【図1】母系および父系の染色体成分の不均衡な分配を含む遺伝物質のプールを生成するための実施形態を示す。20

【図2】両親からの染色体の混合集団の例と、混合集団からハプロタイプを決定する際の課題を示す。20

【図3】例示的な染色体集団と、ハプロタイプの決定におけるそれらの使用を示す。

【図4】遺伝物質の不均衡な分配を含む、本明細書に記載の方法を実施するために利用可能な例示的な遺伝子型決定情報を示す。30

【図5】所与のアッセイから有用な情報を生成する確率（すなわち、測定可能な差の確率）に対する例示的な充填率（充填される標的分子の期待数／アッセイウェルまたは位置×100）を示す。

【図6】2つの代表的な対立遺伝子、すなわち対立遺伝子Aと対立遺伝子Bとを含む遺伝物質を不均衡に分配させるための偏った增幅方法の実施形態を示す。

【図7】遺伝物質を不均衡に分配させるための鋳型の偏った分解方法の例を示す。

【図8】2つの代表的な対立遺伝子すなわち対立遺伝子Aと対立遺伝子Bとを含む遺伝物質を不均衡に分配させるための偏った分解方法の実施形態を示す。

【図9】正常な二倍体個体の蛍光生強度の例示的な散布図と、ヘテロ接合型SNPをそれの一倍体成分に分解する本明細書に記載の方法の能力を示す。40

【図10】図9の二倍体試料から取り出した12個の希釈した試料のうちの6個からの、A(Y軸上)およびB(X軸上)を任意に付した2つの遺伝子座の蛍光生強度の一連の例示的な散布図を示す。

【図11】上のパネル内の細胞HG01377(上)およびNA18507(下)に由来する不均衡な遺伝物質のプールからの配列比較されるセグメント、および下のパネル内の併合されたハプロタイプブロック(それぞれHG01377およびNA28507)を示す。

【図12】細胞NA18506由来の正常な個体の全ヒトゲノムからの不均衡な遺伝物質のプールからの配列比較されるセグメント(上のパネル)、および下のパネル内の併合さ50

れたハプロタイプブロックを示す。

【発明を実施するための形態】

【0012】

定義

本明細書に使用されている「ハプロタイプ」という用語は、典型的には1つの単位として継承され、かつ例えば転座イベント時に結びつけられた一倍体遺伝子型、すなわち染色体上の異なる位置または遺伝子座に存在する対立遺伝子すなわちDNA配列の組み合わせもしくは集まりを指す。ハプロタイプは、個体特有の遺伝型を提供することができる。ハプロタイプは、所与の組の遺伝子座間に生じる組み換えイベント数に応じて、1つの遺伝子座、いくつかの遺伝子座または染色体全体について決定することができる。対立遺伝子すなわちDNA配列は、どんな特定の型にも限定されず、例えば、正常な遺伝子配列（すなわち非変異形）または変異型遺伝子配列を含む。例えば、一塩基多型（SNP）、短いタンデム反復（STR）などは、変異型遺伝子配列とみなすことができる。「相が特定された対立遺伝子」という用語は、单一染色体上での特定の対立遺伝子の分配を指す。従って、2つの対立遺伝子の「相」とは、対立遺伝子が单一染色体または2つの別個の染色体（例えば、母系もしくは父系遺伝の染色体）上に位置しているか否かの特性評価または決定を指すことができる。特に明記しない限り、「ハプロタイプ」および「相が特定された対立遺伝子」は、同義であるものとみなす。

10

【0013】

本明細書に使用されている「単離された」、「精製された」または「精製すること」という用語は、試料から成分（例えば、汚染物質）を除去した産物またはその行為を指す。例えば、汚染性宿主細胞または他のタンパク質、塩、酵素、および核酸をその現在の環境から単離するのに使用される緩衝液などの除去により、核酸を細胞残屑または単離試薬から分離または単離する。

20

【0014】

本明細書に使用されている「試料」という用語は、生物学および化学の技術分野におけるその意味と一致するように使用されている。ある意味では、生体および環境試料などの任意の提供源から得られた検体または培養物からの核酸を含むことが意図されている。生体試料は、限定されるものではないが、ヒト、非ヒト霊長類、非ヒト動物（限定されるものではないが、齧歯類、ヒツジ類、ウシ類、反芻動物、ウサギ類、ブタ類、ヤギ類、ウマ類、イヌ類、ネコ類、鳥類などの脊椎動物）などを含む動物から得られたものであってもよい。生体試料としては、血液製剤、組織および細胞などの流体が挙げられるが、これに限定されない。さらに、生体試料は、限定されるものではないが農作植物、景観植物および苗などの植物由来、すなわち单子葉もしくは双子葉植物、落葉性もしくは常緑性植物、草本もしくは木本植物であってもよい。環境試料は、細菌、ウイルス、真菌由来などであってもよい。好ましい試料は、真核生物由来である。基本的に、相が特定された対立遺伝子を決定するのに研究者にとって興味深い任意の生物体核酸試料源が本発明に受け入れられる。試料は、合成の核酸も含むことができる。増幅されたコピーまたは化学修飾種などの核酸の誘導体または産物も含まれる。

30

【0015】

本明細書に使用されている「核酸」という用語は、例えば、ヌクレオチドのポリマーすなわちポリヌクレオチドであってもよい。この用語は、单一の分子または分子の集合体を意味するように使用することができる。核酸は、一本鎖であっても二本鎖であってもよく、コード領域および様々な調節要素の領域、非コード領域、染色体全体、部分的な染色体、断片およびそれらの変異体を含んでもよい。

40

【0016】

本明細書に使用されている「非対称な」、「不均衡な」、「不均等な」または「偏った」という用語は、同様の要素の分配に関して使用されている場合、特に明記しない限り同義であるものとみなす。この用語は、異なる量の同様の要素が2つ以上の個々の画分で生じるように、複数の画分、一定分量、サブセットなどにわたって分配される同様の要素、

50

10 例えば染色体または染色体成分の集合体を指す。複数の画分中の2つ以上の個々の画分は、同様の要素を有することができる。但し、複数の画分中の全ての画分が要素を有している必要はなく、むしろ1つ以上の画分、一定分量、サブセットなどは、要素を全く有していない。個々の画分は、存在する要素に関して同種であってもよく、あるいは、複数の同様の要素が1つ以上の異なる要素と共に存在するように、要素の異種の集合体が個々の画分に存在していてもよい。同様の要素は、実質的に同様または同一であってもよい。例えば、同様の要素は、一般的な配列を有する染色体、一般的な配列を有する染色体の断片、一般的な配列を有する染色体の少なくとも一部のコピー、または一般的な配列を有する他の核酸分子であってもよい。試料を、その成分比が最初の集団の比と同じではない画分、一定分量、サブセットなどに分けて、同様の要素の非対称もしくは不均衡な試料を調製することができる。同様の要素の非対称な分配は、例えば、画分中に不均等分配、例えば、0.5:1、1:1.5、1:2、1:3、2:3などの比の2つの親染色体の寄与が生じる2つの親染色体の寄与の分配（例えば、一方が母系由来の染色体、もう一方が父系由来の染色体）である。画分、一定分量、サブセットなどは、例えば、管、ウェル（例えば、マイクロタイタープレート内のウェル）、マイクロアレイ内のフィーチャー（feature）、表面もしくは基板上のスポット、ビーズまたは粒子などであってもよい。

【0017】

20 試料中の非対称性、不均衡性または偏りは、相対的な特徴であってもよいこと、または、相対的な方法で決定できることが理解されるであろう。例えば、試料は、試料が取り出された個体、組織または細胞中に存在する染色体または染色体成分の量とは異なる染色体または染色体成分の量によって特徴づけられる非対称性、不均衡性または偏りを染色体または染色体成分中に有することができる。従って、試料が取り出された個体、組織または細胞が少なくとも1つの染色体または染色体成分の量に天然に生じる非対称性、不均衡性または偏りを有することができ、当該試料を、少なくとも1つの染色体または染色体成分の量に天然に生じない非対称性、不均衡性または偏りを有するように非対称にすることができることが理解されるであろう。

【0018】

30 本開示の実施形態は、生体試料のハプロタイプを決定するための方法およびシステムを提供する。特定の実施形態は、ゲノムの長距離ハプロタイプの決定方法を提供する。ゲノムのハプロタイプ決定の重要性は、例えば、個別の健康管理システムへの寄与およびその駆動、ならびに臓器や組織の移植成功への寄与において、大きな意義を有する。

【0019】

40 従来の遺伝子型決定法（例えば、マイクロアレイ、配列決定、PCRなど）では、特に、目的配列が染色体上で遠く離れている場合に、単一染色体のハプロタイプの決定において困難に直面する。例えば、現在実施されているようなマイクロアレイおよびPCR分析は、通常はハプロタイプ決定情報を提供せず、配列の存在の有無のみを提供する。キャピラリー系配列分析方法などの現在実施されている第一世代の配列決定法は、システムに応じて、近位、例えば1,000 bp以下の範囲内の目的配列を検出することができる。次世代の配列決定（NGS）法のスケーラビリティは、長距離ハプロタイプを決定することに関して、比較的短い配列決定読み取り断片（例えば、システムに応じて数百個の塩基対）によって限定されているため、現在実施されている次世代の配列決定の読み取りは、その間のいずれかの範囲である。本明細書に記載されている実施形態は、ゲノム中の隣接すなわち近位および遠位すなわち長距離の対立遺伝子の相の特定を提供することにより、これらの上に述べた技術によって残されているギャップを埋める。実際に、本明細書に記載されている実施形態は、他に類を見ないほど長距離ハプロタイプを同定するのに適している。本方法は、使用されている特定の技術で検出される核酸断片の長さよりも長い範囲を有するハプロタイプを同定するのに特によく適している。例えば、本明細書に記載の方法のNGS系の実施形態を使用して、用いられるNGS技術の読み取り長さよりも長い範囲を有するハプロタイプを同定することができる。本明細書に記載の方法を実施して得られる相が特定された対立遺伝子から得られる情報は、例えば、疾患検出および個別の健康管理

理（P H C）において有用性が見出される。例えば、個体のハプロタイプを、薬物代謝、創薬、病状、癌、障害および移植拒絶反応リスクなどに相関させることができる。実際に、個別の健康管理に関して、研究対象の相が特定されたハプロタイプ、次いで研究対象特異的疾患との相関が分かると、治療の選択肢を具体的に設計して、その研究対象の必要性に応じることができる。

【0020】

本明細書に記載されている実施形態は、ハプロタイプ決定のための他の方法と比較して優れた代替となる。本開示は、例えば使用が容易で高処理用途に受け入れられる方法を提供し、試料が一倍体であるか二倍体であるかに関わらず、また試料が目的の対立遺伝子にホモ接合型であるかヘテロ接合型であるかに関わらず、長距離の対立遺伝子の相を特定する能力を有する。

10

【0021】

ハプロタイプ決定のための遺伝物質のプールを生成するための実施形態が、図1に例示されている。ゲノムまたは染色体の大部分に母系および父系の染色体成分の不均衡な分配を有する遺伝物質のプールの生成方法の一実施形態は、遺伝物質の不均等分配（左の矢印）を得るためにポアソン乱数を利用することを含む。例えば、正常なDNA試料は、母系と父系の染色体比が1:1である。本明細書に開示されている方法を実施してその試料を分画して、母系と父系の染色体の比（またはその逆）を、1:1以外、例えば、少なくとも1:0.5、少なくとも1:2、少なくとも1:3、少なくとも1:4、少なくとも2:1、少なくとも2:3などにすることにより染色体を不均衡に分配させることができる。

20

【0022】

ポアソン乱数を利用して遺伝物質の不均等分配を得ることを含む本開示の実施形態が、図2および図3に例示されている。遺伝子型決定用試料は、両親からの染色体の混合集団で構成されていてもよい（図2A）。この種の分析では、患者の遺伝子型を決定することはできるが（図2B）、ヘテロ接合型対立遺伝子が染色体上でどのように組み合わせられているのかは分からぬ。この例では、親Aが遺伝子^aおよび^bにおいてどちらにも例示的な（-）対立遺伝子を提供しているか否か、親Bが例示的な（+）対立遺伝子を提供しているか否か（図2C）、あるいはそれらが混合されているか否か（図2D）は分からぬ。ハプロタイプを決定する1つの方法は、各染色体をそれ自体の区画内に単離することを含み（図3D）、それを別個の試料として取り扱う。このように、区画中には各遺伝子の1つのコピーしか存在しないため、各試料は全ての対立遺伝子においてホモ接合型になる。しかし、この方法の欠点は、多くの空のアッセイウェルが存在すること（図3C）（但し、空のウェルは、陰性アッセイ対照として利用するのに有利になり得る）と、単一染色体を含むウェルからの信号が非常に低くなり得ることである。本明細書に記載の方法は、アッセイウェルまたは区画などの画分中に、より高濃度でこれらの画分全体に非対称に分配させた染色体試料を提供する。各親からの不均等な数の染色体（または、染色体に由来する配列を有する核酸分子）が存在する限り（図3B）、例えば、等しい数の親の染色体を示す図3Aとは対照的に、より大きな数を有する染色体からの対立遺伝子は、より高い検出信号（例えば、蛍光、発光など）を示すことができ、それにより、互いに関連づけて、異なる染色体のハプロタイプを決定することができる（図3E）。

30

【0023】

本開示の特定の方法を実施することに関する推定される改良により、既存の技術と比較して、充填密度を2~3倍増加させ、かつ所与のアッセイから使用可能な総データを5~6倍増加させることができると想定される（図4および図5）。例えば、図4Aは、染色体がアッセイにおいて単一分子レベルまで希釈されている標準的な希釈アッセイから入手可能な遺伝子型決定情報の程度を示している。1つの染色体が存在しているそれらのアッセイウェルのみで、有用なデータ、例えば $P_a = 1$ 、 $P_b = 0$ またはその逆が得られる。逆に言えば、例えば、2つの異なる対立遺伝子間の検出差が測定閾値シータ（ θ ）よりも大きい限り、体積当たりで任意の数の染色体を使用することができるため、有用な情報量

40

50

の大きな増加は、本明細書に記載の方法の実施形態を実施することにより得られる（図4B）。

【0024】

本開示の方法の実施形態を実施することにより、充填密度がより大きくなり、かつ所との数の画分に対して体積または画分当たりでより高い確率のデータを生成することができるため、ハプローム（すなわち一倍体ゲノム）のカバー率は、0～1倍希釈法などの他の方法を実施する場合と比較すると、より高い（図5）。例えば、0または1倍希釈の場合（例えば、図5Aに例示されているような例）の最大値は、使用可能なデータを生成するアッセイウェルの36%のみを用いて、24%の充填において認めることができる。あるいは、図5Bは、本明細書に開示されている非対称充填方法により、使用可能なデータを生成するアッセイウェルの76%を用いて、最大100%の充填を得ることができることを示している。検出システムの分解能すなわち感度は、使用可能なデータを得るのに必要なアッセイ画分数に影響を及ぼすと想定される。標的分子（すなわち、染色体成分）は、全染色体、染色体の断片、BAC、YAC、MAC、フォスミド、コスミドなどに存在するクローン化された染色体の挿入断片を含む。さらに、本開示の方法は、0～1倍希釈法と比較した場合、より少ない画分でハプロームの等しいカバー率を潜在的に提供することができる。

10

【0025】

一実施形態では、偏ったすなわち不均衡な增幅方法は、1組の相が特定された対立遺伝子が、遺伝物質の不均衡分配を得るように想定された增幅された集団中で識別可能であるように、異なる効率で対立遺伝子を増幅するためのプライマーおよび／または増幅条件を含む（図1、真ん中の矢印）。偏ったすなわち不均衡なポリメラーゼ連鎖反応（PCR）などの偏ったすなわち不均衡な増幅を使用して、例えば、対立遺伝子のうちの1つの増幅を（部分的に）遮断することにより、2つの対立遺伝子の不均衡分配を得ることができる。例えば、一実施形態は、Re xら（2009, J. Virol.

20

Meth. 158:24-29）およびSenescauら（2005, J. Clin. Micro.

43:3304-3308）（それらの両開示内容全体が参照により本明細書に組み込まれる）に記載されているようなプロッキングプローブの使用を含む。例えば、プロッキングプローブは、対立遺伝子のうちの1つに相補的であってもよく（図6A、上の反応：Aヌクレオチドを跨ぐように示されているプロッキングプローブ）、PCRの伸長温度に適合するTmを有し、DNAポリメラーゼによるその伸長を防止する3'プロッキング基を有する。DNAポリメラーゼ（例えば、非鎖置換型）がプローブに出会うと、鎖伸長は停止し、最終PCR産物混合物中の1つの対立遺伝子の発現は減少する。逆に言えば、他の対立遺伝子の鎖伸長は、プロッキングプローブの存在により妨害されないため、最終PCR産物混合物中にその対立遺伝子の正常な発現が生じ、それにより、PCR産物混合物中に1つの対立遺伝子の偏った発現が生じる（図6A、対立遺伝子Aよりも多い対立遺伝子B）。

30

【0026】

別の実施形態では、偏ったすなわち不均衡な増幅方法は、遺伝物質の不均衡なプールを生成するために、増幅反応において耐熱性のMutSタンパク質と、対立遺伝子特異的プローブ、例えば対立遺伝子特異的プロッキングプローブとを含む（図6B）。MutSは、Mg²⁺の存在下でヘテロ二本鎖DNAに強く結合するDNAミスマッチ結合タンパク質である（Lishanski et al., 1994, Proc. Natl. Acad. Sci. 91:2674-2678; Stanislawska-Sachadyn and Sachadyn, 2005, Acta Biochim. Pol. 52:575-583、その両開示内容全体が参照により本明細書に組み込まれる）。例えば、1つの対立遺伝子に相補体である対立遺伝子特異的プロッキングプローブを、2つの対立遺伝子鑄型を用いてホモ二本鎖DNAおよびヘテロ二本鎖DNAの両方を形成する鑄型DNA分子にアニールすることができる。MutSは、非相補的対立遺伝子と対をなしているプロッキングプローブに選択的に結合することができる（図6Bの上の反応：B対立遺伝子上に示されているヘテロ二本鎖形成と、下の反応において円で示されているMutS結合）。鎖置換型DNAポリメラーゼ（例えば、phi 29 DNAポリメラーゼ、Bst DNAポリメラーゼ大型

40

50

断片、Vent (登録商標) (エキソ-) DNAポリメラーゼ、Deep Vent (登録商標) (エキソ-) DNAポリメラーゼ、9°N_m DNAポリメラーゼなど)を用いて、Mut Sによって結合されていないプローブを除去して(例えば、抗Mut Sを用いた陰性抗体選択により)、完全にマッチする鑄型分子の鎖伸長を可能にするが、Mut S複合プローブは、適所に留まっているため、ミスマッチな鑄型分子の鎖伸長は停止し、それにより、最終產生混合物中に対立遺伝子の不均衡な発現が生じる(図6B、対立遺伝子Bよりも多い対立遺伝子A)。

【0027】

別の実施形態における偏ったすなわち不均衡な増幅方法が、図6Cによって例示されている。図6C(上の組の対立遺伝子)では、短いプローブは、遺伝子座の片側とハイブリッド形成することができる。特異的対立遺伝子にマッチするそれらのプローブでは、プローブの伸長およびライゲーションを生じさせることができる。但し、プローブおよび対立遺伝子が非相同的である場合、プローブの伸長およびライゲーションは全く生じないか最小である(上の組の2つ目の対立遺伝子)。伸長およびライゲーション後に、伸長およびライゲーションされたそれらのプローブは鑄型とハイブリッド形成したままであるが、伸長されていない短いプローブは鑄型から解放されるように、温度を上昇させることができる(3つ目の組の対立遺伝子)。ハイブリッド形成および伸長されたプローブは、鑄型と架橋し、それによりPCR増幅を遮断し、その結果、対立遺伝子の一方が他方よりも多くなる(この場合、対立遺伝子Aよりも多い対立遺伝子B)。

10

【0028】

別の実施形態における偏ったすなわち不均衡な増幅方法が、図6Dによって例示されている。図6Dは、対立遺伝子特異的PCRの使用を示しており、そこでは、プライマーの1つが、その3'末端で、多型部位(すなわち、SNPまたは他の多型の位置)の近くでアニールする。ミスマッチなプライマーは複製を開始しないが、マッチするプライマーはそれ自体が複製することができるため、一方の対立遺伝子が他方の対立遺伝子よりも多くなる(図6D、対立遺伝子Bよりも多い対立遺伝子A)(Newton, 1989, Nucleic Acid Res. 17:2503-2516、その開示内容全体が参照により本明細書に組み込まれる)。

20

【0029】

一実施形態では、遺伝物質の不均衡分配を得る工程は、対立遺伝子の偏った分解を含む(図1、右の矢印)。例えば、1つの対立遺伝子のみ(例えば、未消化の対立遺伝子)が増幅し、従って、増幅された鎖上の全ての対立遺伝子が同じ相を共有するように、プライマー間の2つの遺伝子座(例えば、ATACCおよびTTGTCを含む例示的な遺伝子座)上の対立遺伝子特異的位置で鑄型を消化することができる(図7)。試料をいくつかの別個の画分(A、BおよびC)に分割することができる。いくつかの遺伝子座は、対立遺伝子標的(AおよびG)においてヘテロ接合型であり(図7A)、ここでは、分解後に生じる集団は、単一の一倍体成分(この例では、遺伝子座TTGTCおよび対立遺伝子G)を過剰発現しているため、例えば、別個の反応物のインデックス付けおよび配列決定後に領域内の全ての対立遺伝子の相を特定することができる。いくつかの遺伝子座は、対立遺伝子標的(対立遺伝子T)においてホモ接合型であり(例えば、図7Bおよび図7C)、2つの一倍体染色体の寄与間で等しく増幅された集団が生成されるか(図7B)、あるいは増幅が僅かに生じるか全く生じない(図7C、対立遺伝子C)。

30

【0030】

図8は、偏った分解方法のいくつかの例示的な実施形態を示す。図6Bの例示的な修正形態として、図8Aは、完全にマッチする二本鎖分子を、例えば二本鎖特異的核酸分解酵素DSNで選択的に破壊することができるが、Mut Sに結合したミスマッチな二本鎖は切断から保護されていることを示す。図8Aは、耐熱性Mut Sタンパク質(円)、対立遺伝子特異的プローブ、および二本鎖特異的核酸分解酵素(はさみ)の使用を示し、ここでは、二本鎖特異的核酸分解酵素は、対立遺伝子Aよりも多い対立遺伝子Bの偏った増幅のためにホモ二本鎖DNAを切断することができる。

40

50

【 0 0 3 1 】

別の実施形態では、偏った分解方法は、単一ヌクレオチドのミスマッチのための強力な標的部位選択性を有するファージ Mu トランスポゾン (Yanagihara and Mizuuchi, 2002, Proc. Natl. Acad. Sci. 99:11317-11321、その開示内容全体が参照により本明細書に組み込まれる) と、対立遺伝子特異的プローブとを含む。Mu は、それ自体がミスマッチを有するヘテロ二本鎖 DNA 内に選択的に挿入することができるため、ライプラリー調製手順 (図 8 B、円で示されている Mu トランスポゾン) におけるその使用により、例えばミスマッチ対立遺伝子の鋳型分子を断片化するように機能することができるが、完全にマッチする対立遺伝子の鋳型分子は無傷なままであり、PCR 増幅における鋳型として機能するため、ハプロタイプ決定のための偏ったすなわち不均衡な遺伝子プール (図 8 B、対立遺伝子 B よりも多い対立遺伝子 A) が生成される。

10

【 0 0 3 2 】

別の実施形態では、偏ったすなわち不均衡な増幅方法が、図 8 B の修正形態である図 8 C によって例示されている。図 8 C には、ビオチン化された対立遺伝子特異的プローブ (B で示す) が鋳型 DNA とハイブリッド形成することが示されている。ストレプトアビジントランスポゾン融合タンパク (例えば、円で示されている Epicentre Biotechnologies 社製 Nextera DNA 試料調製キットに例示されている Mu トランスポゾン) を、ストレプトアビジン - ビオチン相互作用による二本鎖ハイブリッド形成部位に補充することにより、完全にマッチする対立遺伝子の断片化を生じさせ、かつ一方の対立遺伝子を他方の対立遺伝子よりも多くすることができる (図 8 C、対立遺伝子 A よりも多い対立遺伝子 B)。

20

【 0 0 3 3 】

別の実施形態では、偏った分解方法は、図 8 D に示すように、制限酵素を含むことができる。例えば、単位複製配列当たり約 1 つの制限部位が存在するように (例えば、公知のヘテロ接合型遺伝子座を標的とすることにより、あるいは単位複製配列長さに基づく統計により)、1 種以上の制限酵素を選択してもよい。増幅が不可能になるように、標的部位を含む単位複製配列を分解する (すなわち、円で示されている制限酵素によって制限する) ことができる。未消化の対立遺伝子を選択的に増幅して、ハプロタイプ決定のための対立遺伝子の不均等な発現を生じさせることができる (図 8 D、対立遺伝子 B よりも多い対立遺伝子 A)。

30

【 0 0 3 4 】

本開示は、ゲノムのハプロタイプの決定方法を提供する。一実施形態では、本開示の方法により、研究対象からの二倍体もしくは一倍体ゲノム試料からの遺伝物質 (すなわち染色体成分) の不均衡分配を得る。標準的な方法 (例えば、マイクロアレイ、配列決定、PCR、ゲル系など) を用いた不均衡な遺伝物質の遺伝子型決定により、長距離ハプロタイプ決定のために、広いゲノム領域にわたるハプロタイプの決定が可能となる。例えば、ハプロタイプ決定のために遺伝物質を非対称すなわち不均衡に分配するために本明細書に記載の方法を利用する場合、特定のゲノム領域内の 1 組の目的の標的配列が、別の組の対立遺伝子よりも、増幅信号強度 (例えば、マイクロアレイによる) が 3 倍高いか読み取り断片 (配列決定) が 3 倍多ければ、2 つのそれぞれの組は、2 つの異なるハプロタイプに対応すると推定される。不均衡な遺伝物質プール中の各目的の標的配列の相対的な量を測定したら、正常な二倍体ゲノムまたはプールされている正常なゲノムから測定した量と比較し、それにより試験試料中の異常を決定する。

40

【 0 0 3 5 】

本開示は、ハプロタイプ決定のために、試料の不均等な、不均衡な、偏ったまたは非対称な分配を含む方法を提供する。不均等分配は、例えば、希釀、非対称 PCR、標的化分解などによるものであってもよい。特に、本明細書に記載されている実施形態は、基板 (例えば、プレート内のウェル、スライド上の領域、複数の毛細管、フレキシブルテープの中/上のウェルなど) 上のアッセイ位置などの画分間に不均一に分配される研究対象からの遺伝物質を提供する。特定の実施形態では、試料の遺伝物質の不均一な分配は、基板上

50

の1つ以上のアッセイ位置に位置する染色体の不均等分配を表している。いくつかのアッセイ位置は遺伝物質を全く含んでおらず、これらの位置は、図3Cに例示されているようにアッセイ内で陰性対照としての有用性が見出されるものと想定される。基板としては、シリカまたは高密度プラスチックスライドおよびチップなどのマイクロアレイ基板、96、384、1536ウェルアッセイプレートなどのプレート、例えばPCRによる流れのために使用されるような毛細管、フレキシブルな高処理アッセイストリップ(例えば、Douglas Scientific社製Array Tape(商標))、ビーズ、ナノ粒子などが挙げられるが、これらに限定されない。本明細書に記載の方法は、その上または中でアッセイが行われる基板によって限定されない。

【0036】

10

本明細書に記載の方法の特定の実施形態を使用して、例えば、染色体上で互いに對して近位および遠位にある目的配列のハプロタイプを決定することができる。目的配列は、どんな特定の距離でも離れておらず、例えば、目的配列が、染色体上で互いに對して隣接するか近位にあってもよいことが想定されている。逆に、目的配列が、染色体上で互いから遠位に離れている、すなわち長距離にあることが想定されている。実際に、本明細書に記載されている実施形態を実施することは、長距離ハプロタイプを決定する際に特に有益であり得る。目的配列間の距離は、本方法を限定するものではなく、例えば目的配列は、少なくとも100個、200個、300個、400個、500個、750個または少なくとも1000個の塩基対だけ離れていてもよい。但し、実施形態は、染色体上でかなり離間されている、例えば、少なくとも10,000個、少なくとも100,000個、少なくとも1,000,000個、少なくとも10,000,000個、少なくとも150,000,000個、少なくとも200,000,000個、少なくとも247,000,000個以上の塩基対だけ離れている場合に、目的配列のハプロタイプの決定のために特に有用性が見出される。従って、本明細書に記載されている実施形態は、決定のための試料が一倍体であるか二倍体であるかに関わらず、個体のゲノムの長距離ハプロタイプ決定に特に適した方法を提供することができる。

20

【0037】

30

本開示の実施形態では、ハプロタイプ、特に染色体上で遠位に位置する目的配列の決定方法が提供される。いくつかの実施形態では、目的配列は、一塩基多型すなわちSNPである。いくつかの実施形態では、SNPは互いに隣接している、すなわち近位にあるが、他の実施形態では、SNPは互いに遠位にある、すなわち長距離にある。いくつかの実施形態では、目的配列は、ゲノム内の配列の挿入もしくは欠失すなわちinsertionである。いくつかの実施形態では、目的配列は、ゲノムのコピー数多型すなわちCNVである。他の実施形態では、目的配列は、対立遺伝子すなわち染色体上の特定の位置に位置する遺伝子または配列の他の形態である。いくつかの実施形態では、対立遺伝子は、野生型すなわち正常型の認識されている配列であるが、他の実施形態では、対立遺伝子は、SNP、CNV、insertionなどの、野生型と比較されるような1種以上の突然変異を有していてよい。

【0038】

40

そのような突然変異を同定して、癌および遺伝性疾患などの病状と直接相關させてもよい。突然変異した対立遺伝子は、研究者にとって特に興味深く、本開示の実施形態の実施により、研究者が対立遺伝子の突然変異およびそれらのハプロタイプを研究することができる有益な手段を提供することができる。ハプロタイプは、個体の二倍体ゲノムの遺伝子構造を定義するのに有益である。ハプロタイプ決定情報により、限定されるものではないが、薬物代謝、創薬、個別の健康管理戦略、移植成功のためのHLAタイピング、集団遺伝学、複合疾患の関連、遺伝人類学、疾患および癌の遺伝医学、癌や他の疾患における構造変化、対立遺伝子特異的発現や対立遺伝子特異的メチル化パターンなどの修正形態、および新規なゲノムアセンブリを含む科学的研究の多くの領域において、さらに深い理解が得られ、かつさらに広い有用性が見出され得る。偏った増幅および偏った分解を含む実施形態は、ハプロタイプ決定のための目的の対立遺伝子が小さなゲノム領域に由来する場合

50

に、特に有利である。従って、数キロベースまたは1つ以上のゲノムの領域にわたるハプロタイプ決定が望まれるHLA遺伝子型決定（例えば、HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1、HLA-DQA1など）などの臨床用途は、本明細書に開示されている方法の実施により多大な恩恵を受けるであろう。

【0039】

対立遺伝子を染色体に割り当てる能力（すなわち、ハプロタイプ決定）は、例えば、ゲノム中の組み換えイベントに関する情報を提供することにより、臨床的関連性に関する情報を提供することができるため、強力である。そのような情報は、疾患を引き起こす突然変異の場所を突き止めるのに重要であり得、連鎖不平衡すなわちゲノム中の2つの多型の存在間の関連性（疾患ゲノムと広範な疾患との関連研究の主要な特性）を決定するのに役立つことができる。例えば、1つの多型（すなわちSNP）における遺伝子型を知ることは、2つの多型間の関連性（すなわち連鎖不平衡）が高い場合に、別の多型（すなわちSNP）の遺伝子型の予想に役立つことができる。それらのハプロタイプを決定することにより、ヒトの白血球抗原（HLA）により完全に適合させることができれば、例えば移植患者の臨床的予後を大きく改善するであろう（Crawford and Nickerson, 2004, Ann. Rev. Med. 56:303-320、その開示内容全体が参照により本明細書に組み込まれる）。例えば、本明細書に開示されている方法を実施することにより、移植患者および臓器提供者候補の遺伝子型を主要組織適合遺伝子複合体に沿った複数のマーカーで同定し、そのハプロタイプを生成されたデータから決定することができる。そのような配列比較の例は、本明細書に開示されている実施例に記載されている。そのような配列比較により、移植患者と臓器提供者との非常に正確なHLA適合が得られ、その結果、そのように適合していない患者および臓器提供者よりも良好な移植結果が得られる。

10

20

30

40

50

【0040】

さらに、特定の遺伝子座における遺伝子型ではなくハプロタイプによって、疾患の重症度を予測することができる疾患もある。それは、そのような正確なハプロタイプが、特定の患者の疾患の重症度を決定する際に広範な有用性を有するだけでなく、異なる治療選択肢は異なる病状および/または重症度レベルに相関し得ることから、その診断および/または予後診断に基づいて可能な治療選択肢を決定する際にも、臨床医に情報を提供するからである。例えば、特定の鎌型赤血球貧血-グロビン遺伝子座のハプロタイプは、重症度の低い鎌型赤血球貧血と相関されており、IL10プロモーター領域のハプロタイプは、細胞移植を受けた患者の移植片対宿主病および死亡の発生率の低さと関連づけられている。従って、ゲノム試料のハプロタイプ決定を提供する本方法は、例えば、疾患の相関研究、疾患診断および予後診療、ならびに治療計画の適用に大きな影響を有し得る。但し、ハプロタイプ決定は、農業および他の園芸術、特に家畜や作物の育種においても非常に重要であり、そこでは、疾患または有利な特性を、動物や植物における特定のハプロタイプと相関させることができる。

【0041】

本明細書に提供されている実施形態は、試料中の相が特定された対立遺伝子の決定方法について記載している。典型的には、試料は核酸試料を含む。いくつかの実施形態では、核酸試料は、研究対象からの体液、例えば、血液、痰、尿、髄液などに由来している。他の実施形態では、生体試料は、研究対象からの固体、例えば、組織、生検、細胞剥離物(cell scraping)、細胞学試料または細胞試料などに由来している。一実施形態では、生体試料は、精製された単一の染色体もしくはその断片、または、例えば、コスマド、フォスマド、プラスミド、酵母人工染色体(YAC)、細菌人工染色体(BAC)、哺乳類人工染色体(MAC)、植物クローニング系(例えば、アグロバクテリウム・ツメファシエンスT-DNAクローニング系、バイナリーベクタークローニング系など)もしくはその断片などの中のDNA挿入物である。好ましい実施形態では、生体試料は、1つ以上の細胞中に存在する二倍体DNA試料である。但し、本明細書に記載の方法の実施形態は、一倍体試料(例えば、卵子、精子、胞状奇胎、および機械的に分離および/または単離された染色体、その断片、クローン化されたDNA断片などに由来する核酸)も本明細書に記載

の方法を実施するのに等しく受け入れられるため、二倍体試料に限定されない。

【0042】

一実施形態では、試料は、細胞試料または組織試料である。細胞試料または組織試料は、例えば、分離された組織からの細胞、血液または他の体液からの細胞、細胞診標本からの細胞、非ヒト動物からの細胞、植物からの細胞などの任意の提供源からのものであってもよい。好ましい実施形態では、細胞は、哺乳類由来、好ましくはヒト由来である。但し、本明細書に記載の方法は、細胞試料源に限定されない。いくつかの実施形態では、本明細書に記載の方法を実施する際に使用されるゲノム材料は、複数の細胞に由来している。いくつかの実施形態では、複数の細胞は、少なくとも2～1000個の細胞、少なくとも5～500個の細胞、少なくとも10～300個の細胞、少なくとも10～100個の細胞である。本明細書に記載の方法の実施は、特に反する明記がない限り、ウイルス学、免疫学、微生物学、分子生物学および当該技術分野の範囲内の組み換えDNA技術の従来の方法を用いることができる。そのような技術は、以下の文献に十分に説明されている。例えば、1995, Ausubel et al., Short Protocols in Molecular Biology, (3rd ed.), Wiley & Sons, 2001, Sambrook and Russell, Molecular Cloning: A Laboratory Manual (3rd Edition)、1982, Maniatis et al., Molecular Cloning: A Laboratory Manual, DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.)、1984, Oligonucleotide Synthesis (N. Gait, ed.)、1985, Nucleic Acid Hybridization (B. Hames & S. Higgins, eds.)、1986, Animal Cell Culture (R. Freshney, ed.)、1984, Perbal, A Practical Guide to Molecular Cloningを参照されたい。ゲノム材料は、当該技術分野で知られている方法で採取することができ、本明細書に記載の方法は、必ずしも、ゲノム材料の単離のためのどんな特定の方法にも限定されない。当業者であれば、そのような単離のために、無数の市販および独自で考案した代替方法が存在することを理解しているであろう。

【0043】

一実施形態では、ハプロタイプ決定のための試料は、研究対象によって提供される。研究対象は、その実体からハプロタイプを決定することを望む研究者にとって興味深い任意の生物学的実体であればよい。従って、試験のための試料は、必ずしも特定の研究対象に限定されず、研究対象は、例えば動物または植物由来であってもよい。例えば、試料を提供する研究対象は、動物、ヒトもしくは非ヒト、または植物、例えば経済的に妥当な作物などであってもよい。好ましい実施形態では、研究対象はヒトである。他の好ましい実施形態では、研究対象は経済的に妥当な動物またはその派生物である。他の実施形態では、研究対象は、経済的に妥当な植物またはその派生物である。

【0044】

本開示の方法を実施して得られる非対称分配試料は、下流用途にも容易に適用される。いくつかの実施形態では、配列決定装置またはハプロタイプ決定に関連する他の装置の前に、試料に対して下流プロセスを行うことが想定されている。いくつかの実施形態では、非対称分配試料の一定分量または画分を使用して、次世代の配列決定用のクラスター形成のために、DNAライブラリーを調製する。例えば、Nextera (商標) DNA試料調製キット (Epicentre (登録商標) Biotechnologies社、ウィスコンシン州マディソン)、GL FLXチタンライブラリー調製キット (454 Life Sciences社、コネティカット州プランフォード)、SOLID (商標) ライブラリー調製キット (Applied Biosystems (商標) Life Technologies社、カリフォルニア州カールズバッド) などに記載されている方法を行って、そのようなライブラリーを生成する。本明細書に記載されている試料は通常、例えば多置換増幅 (MDA) 法による配列決定またはマイクロア

10

20

30

30

40

50

レイアッセイのためにさらに増幅する。MDA後の配列決定のために、例えば、メイトペアライブラリー調製キット、ゲノムDNA試料調製キット、またはTruSeq(商標)試料調製もしくはエクソーム濃縮キット(Illumina(登録商標)社、カリフォルニア州サンディエゴ)に記載されているようにDNAライブラリーを作成して、増幅された試料ライブラリーを調製する。有用なクラスター増幅方法は、例えば、米国特許第5,641,658号、米国特許出願公開第2002/0055100号、米国特許第7,115,400号、米国特許出願公開第2004/0096853号、米国特許出願公開第2004/0002090号、米国特許出願公開第2007/0128624号、および米国特許出願公開第2008/0009420号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる。表面における別の有用な核酸の増幅方法は、例えば、Lizardi et al., Nat. Genet. 19:225-232 (1998)および米国特許出願公開第2007/0099208号に記載されているようなローリングサークル増幅(RCA)法であり、それらの各開示内容全体が参照により本明細書に組み込まれる。エマルジョンPCR法も有用な例示的方法であり、それについては、Dressman et al., Proc. Natl. Acad. Sci. USA 100:8817-8822 (2003)、国際公開第05/010145号、米国特許出願公開第2005/0130173号または第2005/0064460号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる。本開示の方法は、本明細書に記載されている非対称分配試料が、そのような目的のための当該技術分野で知られ、かつ/または市販されている各種方法のいずれにも受け入れられるものと想定されているため、必ずしもどんな特定のライブラリー調製または増幅方法によつても限定されない。
10

【0045】

例えば、遺伝物質の不均衡な分配を含むDNAライブラリーをフローセルなどの基板に固定し、例えば合成法による配列の配列決定前に固定されたポリヌクレオチドに対してブリッジ増幅を行うことができる。ブリッジ増幅では、固定されたポリヌクレオチド(例えば、DNAライブラリーからのポリヌクレオチド)を、固定されたオリゴヌクレオチドプライマーとハイブリッド形成させる。固定されたポリヌクレオチド分子の3'末端は、固定されたオリゴヌクレオチドプライマーから伸長する、ポリメラーゼで触媒される鋳型による伸長反応(例えば、プライマー伸長)のための鋳型を提供する。得られた二本鎖産物は2つのプライマーと「ブリッジ」を形成し、両方の鎖は、共有結合で担体に取り付けられる。次のサイクルでは、変性後、固体担体に固定された一対の一本鎖(固定された鋳型および伸長されたプライマー産物)が得られ、両方の固定された鎖は、新しいプライマー伸長のための鋳型として機能することができる。従って、第1および第2の部分を増幅させて、複数のクラスターを生成することができる。「クラスター」および「コロニー」という用語は同義で使用され、核酸配列および/または表面に取り付けられたその相補体の複数のコピーを指す。典型的には、クラスターは、核酸配列および/またはそれらの5'末端を介して表面に取り付けられたその相補体の複数のコピーを含む。例示的なブリッジ増幅およびクラスター形成方法については、例えば、国際公開第00/18957号および国際公開第98/44151号、米国特許第5,641,658号、米国特許出願公開第2002/0055100号、米国特許第7,115,400号、米国特許出願公開第2004/0096853号、米国特許出願公開第2004/0002090号、米国特許出願公開第2007/0128624号、および米国特許出願公開第2008/0009420号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる。本明細書に記載されている組成物および方法は、クラスターを含むフローセルを利用する合成法による配列に特に有用である。
20

【0046】

配列決定前に核酸を増幅するためのエマルジョンPCR法は、本明細書に記載の方法およびシステムと組み合わせて使用することもできる。エマルジョンPCRは、油中水型エマルジョン中で、アダプターを両末端に付加したショットガン法によりDNAライブラリ
30

10

20

30

40

50

ーを P C R 増幅することを含む。上記 P C R は、多鑄型 P C R であり、単一のプライマー対のみが使用される。P C R プライマーのうちの 1 つを、マイクロスケールビーズの表面に繋ぐ(5'付加)。低い鑄型濃度により、1つの鑄型分子のみが存在する最も多くのビーズを含有するエマルション微小胞が得られる。生産的エマルション微小胞(ビーズと鑄型分子の両方が存在するエマルション微小胞)では、P C R 単位複製配列を、ビーズの表面に捕捉することができる。エマルションを破壊した後、ビーズ担持増幅産物を選択的に濃縮することができる。クローンとして増幅された各ビーズは、その表面に、鑄型ライブライリーからの單一分子の増幅に対応する P C R 産物を担持している。エマルション P C R 法の様々な実施形態が、例えば、Dressman et al., Proc. Natl. Acad. Sci. USA 100:88 17-8822 (2003)、国際公開第 05/010145 号、米国特許出願公開第 2005/0130173 号、第 2005/0064460 号、および米国特許出願公開第 2005/0042648 号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる。

10

20

30

40

50

【0047】

D N A ナノボールを本明細書に記載の方法およびシステムと組み合わせて使用することもできる。ゲノムの配列決定のために D N A ナノボールを作成および利用する方法は、例えば、米国特許第 7,910,354 号および米国特許出願公開第 2009/0264299 号、第 2009/0011943 号、第 2009/0005252 号、第 2009/0155781 号、第 2009/0118488 号ならびに例えば、Drmanac et al., 2010, Science 327(5961): 78-81 に記載されており、それら全ての開示内容全体が参照により本明細書に組み込まれる。簡潔に言うと、ゲノム D N A の断片化後に、アダプターライゲーション、増幅および消化の連続するラウンドにより、(例えば、環状リガーゼ(circle ligase)によるライゲーションによって)一本鎖 D N A に環状化させ、かつローリングサークルで増幅させた(例えば、Lizardi et al., Nat. Genet. 19:225-232 (1998) および米国特許出願公開第 2007/0099208 A1 号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる)、環状ゲノム D N A 鑄型 / アダプターパターンの複数のコピーからなる頭 - 尾結合コンカテマーが得られる。コンカテマーのアダプター構造により、一本鎖 D N A の螺旋化が促進され、それにより、小型の D N A ナノボールが作成される。D N A ナノボールを基板上に捕捉して、好ましくは、各ナノボール間の距離が維持されるように順序付けまたはパターン化されたアレイを作成することができ、それにより別個の D N A ナノボールの配列決定が可能となる。

【0048】

いくつかの実施形態では、非対称分配試料をさらに処理した後に、配列決定、マイクロアレイ分析、遺伝子型決定または他の下流用途に適用する。例えば、製造業者の手順に従って、Illumina 社(HiSeq 1000、HiSeq 2000、Genome Analyzers、MiSeq、HiScan システム)、454 Life Sciences 社(FLX Genome Sequencer、GS Junior)、Applied Biosystems(商標)Life Technologies 社(ABI PRISM(登録商標)配列検出システム、SOLID(商標)システム)、Ion Torrent(登録商標)Life Technologies 社(Personal Genome Machine 配列決定装置)によって提供され、かつさらに、例えば、米国特許第 5,888,737 号、第 6,175,002 号、第 5,695,934 号、第 6,140,489 号、第 5,863,722 号および米国特許出願公開第 2007/007991 号、第 2009/0247414 号、第 2010/0111768 号ならびに国際公開第 2007/123744 号(それらの各開示内容全体が参照により本明細書に組み込まれる)に記載されているようなシステムを用いて配列決定を行うことができる。

【0049】

いくつかの実施形態では、本明細書に記載されているハプロタイプの決定方法は、配列決定、例えば合成(SBS)法による配列決定で使用する場合に特に有用性が見い出される。合成による配列決定は一般に、1種以上の標識されたヌクレオチドを、ポリメラーゼを用いて 5' から 3' 方向に伸びるポリヌクレオチド鎖に連続的に付加することを含む。

伸長されたポリヌクレオチド鎖は、基板（例えば、フローセル、チップ、スライドなど）に固定することができ、かつ標的配列を含む核酸鑄型に相補的である。SBSで使用される標識されたヌクレオチドは、種々の蛍光色素分子、質量標識、電子的に検出可能な標識または他の種類の標識のうちのいずれかを含むことができる。SBSで使用される標識されたヌクレオチドは、SBSサイクルごとに、1つのヌクレオチドのみが付加されるよう、可逆的なターミネーター基も含むことができる。組み込まれたヌクレオチドが検出された後、非ブロック化剤(deblocking agent)を添加して、その後のサイクルで、付加されたヌクレオチドを伸長可能にすることができる。SBS法は、核酸試料の異なる配列断片の平行分析に特に有用である。例えば、数百、数千、数百万以上の異なる配列断片を、公知のSBS法を用いて、単一の基板上で同時に配列決定することができる。例示的な配列決定法は、例えば、Bentley et al., *Nature* 456:53-59 (2008)、国際公開第04/018497号、米国特許第7,057,026号、国際公開第91/06678号、国際公開第07/123744号、米国特許第7,329,492号、第7,211,414号、第7,315,019号、第7,405,281号および米国特許出願公開第2008/0108082号に記載されており、それらの各開示内容全体が参照により本明細書に組み込まれる。

10

【0050】

本開示のハプロタイプ決定方法は、ライゲーションによる配列決定、ハイブリッド形成による配列決定、および他の配列決定法で使用される場合にも有用性が見い出される。ライゲーション方法による例示的な配列は、Applied Biosystems社製SOLID (商標)配列決定システム (Voelkerding et al., 2009, *Clin Chem* 55:641-658、その開示内容全体が参照により本明細書に組み込まれる)によって利用される2ベースエンコーディング (例えば、カラースペース配列決定) である。

20

【0051】

本明細書に開示されているハプロタイプ決定方法は、ハイブリッド形成法による配列に利用することができる。ハイブリッド形成による配列は、断片化および標識された標的DNA (例えば、Drmanac et al., 2002, *Adv Biochem Eng Biotechnol* 77:75-101; Lizardi et al., 2008, *Nat Biotech* 26:649-650、米国特許第7,071,324号に記載されており、それらの開示内容全体が参照により本明細書に組み込まれる) がそこに付加されるヌクレオチドプローブの短配列のアレイの使用を含む。ハイブリッド形成による配列に対するさらなる改良は、例えば、米国特許出願公開第2007/0178516号、第2010/0063264号および第2006/0287833号 (それらの開示内容全体が参照により本明細書に組み込まれる) に記載されている。ハイブリッド形成とライゲーション生化学を組み合わせた配列決定法が開発および販売されており、例えば、Complete Genomics社 (カリフォルニア州マウンテンビュー) によって実施されているゲノム配列決定法などがある。例えば、コンビナトリアルプローブ-アンカーライゲーションすなわちcPAL (商標) (Drmanac et al., 2010, *Science* 327(5961): 78-81)は、ライゲーション生化学を使用し、ハイブリッド形成による配列の利点を利用している。例えば、Pushkarevら (2009, *Nat. Biotechnol.* 27:847-52、その開示内容全体が参照により本明細書に組み込まれる) に記載され、かつHeliScope (商標) Single Molecule Sequencer (Helicos社、マサチューセッツ州ケンブリッジ) によって実施されている単一の分子配列決定法も、本開示のハプロタイプ決定方法の利点を利用することができる。

30

【0052】

本明細書に記載の方法は、任意の特定の配列決定用試料調製方法によって限定されず、変形形態は、当業者には容易に明らかであり、本開示の範囲内にあるとみなされる。但し、合成法または例えば、1種以上のポロニー (polony) 配列決定法 (Dover Systems社)、ハイブリッド形成蛍光プラットフォームによる配列決定 (Complete Genomics社)、STOP技術 (Industrial Technology Research I

40

50

n s t i t u t e 社) および合成による配列決定 (Illumina, Life Technologies 社) によって実施されているような他の関連の配列決定法によって、本明細書中の方法を、配列を実施するためのフローセルまたはアレイなどの配列決定装置に適用する場合に、特定の有用性が見い出される。

【0053】

いくつかの実施形態では、本明細書に記載されている非対称分配試料を、MDAによって処理し、マイクロアレイおよび/または他の遺伝子型分析アッセイのためにさらに処理する。例えば、いくつかの実施形態では、試料を定量的PCR (qPCR) で処理して、信号雑音比について、個々の画分または一定分量を特性評価する(例えば、Eco PCRシステム (Illumina (登録商標) 社) を利用して行う)。そのような特性評価は、下流配列決定またはマイクロアレイ分析から最高確率の解釈可能なデータを潜在的に提供する画分または一定分量を定めるのに有用である。いくつかの実施形態では、マイクロアレイ分析前に、調製のためのさらなる処理を行う。例えば、限定されるものではないが、ライプラリー試料の調製のために先に上に記載したものを含む種々の方法によって、MDAによる増幅および/またはqPCRによる特性評価後にマイクロアレイ分析用の非対称分配試料を調製する。

【0054】

有用な例示的マイクロアレイとしては、Illumina (登録商標) 社 (カリフォルニア州サンディエゴ) から入手可能なSentrix (登録商標) アレイまたはSentrix (登録商標) BeadChipアレイまたは、例えば、米国特許第6,266,459号、第6,355,431号、第6,770,441号および第6,859,570号ならびに国際公開第00/63437号(それらの各開示内容全体が参照により本明細書に組み込まれる)に記載されているようなウェルの中のビーズを含む他のアレイが挙げられるが、これらに限定されない。

【0055】

表面に粒子を有する他のアレイとしては、米国特許出願公開第2005/0227252号、第2006/0023310号、第2006/006327号、第2006/0071075号、第2006/0119913号、米国特許第6,489,606号、第7,106,513号、第7,126,755号、第7,164,533号、国際公開第05/033681号および国際公開第04/024328号(それらの各開示内容全体が参照により本明細書に組み込まれる)に記載されているものが挙げられる。また、本開示の方法を実施することにより提供されるような非対称分配試料をアッセイするのに有用なビーズアレイは、フローサイトメーターまたは同様な装置の流体流などの流体形態であってもよい。ビーズを識別するための市販されている流体形態としては、例えば、LumineX社のXMAP (商標) 法またはLynx Therapeutics社のMPPS (商標) 法に使用されているものが挙げられる。

【0056】

本開示の方法を実施することにより提供される試料を用いて使用することができる市販されているマイクロアレイのさらなる例としては、例えばAffymetrix (登録商標) GeneChip (登録商標) マイクロアレイまたは、例えば、米国特許第5,324,633号、第5,744,305号、第5,451,683号、第5,482,867号、第5,491,074号、第5,624,711号、第5,795,716号、第5,831,070号、第5,856,101号、第5,858,659号、第5,874,219号、第5,968,740号、第5,974,164号、第5,981,185号、第5,981,956号、第6,025,601号、第6,033,860号、第6,090,555号、第6,136,269号、第6,022,963号、第6,083,697号、第6,291,183号、第6,309,831号、第6,416,949号、第6,428,752号および第6,482,591号(それらの各開示内容全体が参照により本明細書に組み込まれる)に記載されているような、VLIPS (商標) (Very Large Scale Immobilized Polymer Sy-

10

20

30

40

50

nt h e s i s (非常に大規模な固定されたポリマー合成)) 法とも呼ばれることがある技術に従って合成される他のマイクロアレイが挙げられる。

【0057】

本開示の方法を実施することによって提供される試料を用いて、スポットマイクロアレイを使用することもできる。例示的なスポットマイクロアレイは、Amersham Biosciences社から入手可能なCodeLink (商標)アレイである。有用な別のマイクロアレイは、Agilent Technologies社から入手可能なSurePrint (商標)法などのインクジェット印刷法を用いて製造されるものである。使用することができる他のマイクロアレイとしては、Butte, 2002, Nature Reviews Drug Discov. 1:951-60または米国特許第5,429,807号、第5,436,327号、第5,561,071号、第5,583,211号、第5,658,734号、第5,837,858号、第5,919,523号、第6,287,768号、第6,287,776号、第6,288,220号、第6,297,006号、第6,291,193号および第6,514,751号ならびに国際公開第93/17126号および国際公開第95/35505号(それらの各開示内容全体が参照により本明細書に組み込まれる)に記載されているものが挙げられるが、これらに限定されない。

【0058】

配列決定、マイクロアレイまたは他の遺伝子型決定方法または装置からの出力は、どんな種類のものであってもよい。例えば、蛍光または発光などの光を生成する読み取り可能な出力を利用する技術もあれば、電気もしくはイオン放出を測定する技術もある。但し、本発明は、特定の目的配列の出力信号における差を測定することができるものである限り、上記種類の読み取り可能な出力に限定されない。本明細書に記載の方法を実施して得られた出力を特性評価するために使用し得る分析ソフトウェアの例としては、Pipeline、CASAVER、Genome Studio Data Analysis、Bead Studio GenotypingおよびKaryo Studioデータ分析ソフトウェア(Illumina (登録商標)社)、Signal MapおよびNimbleScanデータ分析ソフトウェア(Roche NimbleGen社)、GS Analyzer分析ソフトウェア(454 Life Sciences社)、SOLID (商標)、DNASTAR (登録商標) SeqMan (登録商標) NGen (登録商標)およびPartek (登録商標) Genomics Suite (商標)データ分析ソフトウェア(Life Technologies社)、Feature Extraction and Agilent Genomics Workbenchデータ分析ソフトウェア(Agilent Technologies社)、Genotyping Console (商標)、Chromosome Analysis Suite and Gene Chip (登録商標) Sequence Analysisデータ分析ソフトウェア(Affymetrix (登録商標)社)が挙げられるが、これらに限定されない。当業者であれば、マイクロアレイ、配列決定およびPCR生成出力のデータ分析ためのさらなる数多くの商業的および学究的に入手可能なソフトウェアの変形形態を知っているであろう。本明細書に記載されている実施形態は、どんなデータ分析方法にも限定されない。

【0059】

特定の装置に必要な特定の試料調製は、本明細書に記載されている非対称分配試料と共に使用するのに受け入れられるものと想定されるため、本開示の例示的な方法は、必ずしも特定の配列決定、マイクロアレイまたは遺伝子型決定システムによって限定されない。但し、任意の所与の検出システムの分解能すなわち感度は、解釈可能な結果を得るためにアッセイすることができる画分の数に影響を及ぼし得るものと想定される。分解能の差は、図3B()および図4B()に例示されている。

【0060】

以下の例は、非対称に生成された試料を用いた配列決定によるSNPハプロタイプの決定方法について説明するものである。この特定の例では、Nextera (商標)DNA試料調製キットなどの低いインプットDNAレベル(例えば、10~100 pg)を利用する調製方法は、このキットによって処理される試料が配列決定にすぐに使用できる状態にあり、かつ多置換增幅法などのさらなる処理を全く必要としないため、特に有用である

。それ以外の場合、MDAなどのさらなる増幅工程が必要となるかもしれない。調製された試料を、例えば、 Illumina社製Genome Analyzer、HiSeq、MiSeq、TrueSeqまたは他の配列決定プラットフォームで配列決定することができ、ここでは、分析のために各蛍光標識されたヌクレオチドに対応する蛍光計測値が生成される。例示のために、以下の配列決定結果は、非対称分配試料の調製から得られる。

【化1】

305	295	501	494	303	505	310	301
A G	C G	A G	T T	T T	C G	C G	C T
A G	T G	C T	A G	C G	C G	A A	A G
499	511	302	304	499	298	492	508

10

【0061】

この例では、個々の遺伝子座の核酸は、2本の破線によって、不連続で場合により遠くに位置している染色体領域から分離されている。1つの位置に記載されている2つのヌクレオチドは、目的配列中のヘテロ接合型配列変異体すなわち一塩基多型(SNP)を表している。ヌクレオチドの上下の数は、総読み取り数、例えばこの場合は約800回の読み取りのうちの、その特定のヌクレオチド位置での読み取り数を表している。長距離SNPの相の特定は、以下のような同様の読み取り数を有するSNP位置を組み合わせることによって決定する。

【化2】

20

305	295	501	494	303	505	310	301
A G	C G	A G	T T	T T	C G	C G	C T
A G	T G	C T	A G	C G	C G	A A	A G
499	511	302	304	499	298	492	508

【0062】

この例では、丸で囲まれた数は、異なるSNP位置における同様の読み取りを表しているため、どのSNPが同じ染色体または染色体の断片もしくはセグメント上に位置しているか、従って、相中にあるかを決定し、それにより試料のハプロタイプを決定する。従って、複数のSNPの読み取り数の計数およびそれらの読み取り計数の組み合わせを使用して、試料のハプロタイプを決定することができる。2つの染色体の親の寄与(例えば、上の配列は母系の寄与であり、下の配列は父系の寄与である)についてハプロタイプを以下であると決定する。

30

【化3】

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

40

【0063】

説明のために、核酸配列出力の代わりに、各SNPおよび介在配列のコンピュータで算出されたハイブリッド形成強度のアナログ値に対応するデジタルで得られた色計測値が得られること以外は、同様の方法でマイクロアレイによるハプロタイプ決定を例示することができる。

【0064】

いくつかの実施形態では、ライブラリーの調製または増幅などの任意のさらなる処理の前、あるいは先に例示したようなNextera(商標)キットを利用するライブラリーの調製前に、非対称分配試料を特性評価することができる。例えば、実施例1に記載されているように試料を分画または等分し、各画分を、配列決定またはマイクロアレイ分析で

50

別々に処理する。実施例 1 に記載されているように、試料を 10 個の画分に分けると、1 つの最初の試料に対して 10 回の下流プロセスを行うことができる。試料を複数の画分または一定分量に分けることで、限定されるものではないが、1 つの試料に対して複数回の分析を行えることや、コストや労力（例えば、試薬および他の消耗品、研究者の時間など）が減るといった多くの利点が得られる。コストや労力をさらに減らすために、最大の非対称性、最大の信号雑音比、および所望の標的の最大のカバー率を有するこれらの試料に対する分析前に、複数の画分を特性評価および / または定量することができる。例えば、複数の配列の画分の q P C R 系遺伝子型決定は、画分の非対称性および信号雑音比を決定するのに十分である。さらに、マイクロアレイ分析または深さの浅い配列決定方法を使用して画分の非対称性および信号雑音比も決定することもできる。最大の信号雑音比を有するこれらの画分のみのハプロタイプ決定により、例えば、解釈可能な結果を得る確率が最高となり、それにより時間、労力およびお金が節約されると想定される。

10

【0065】

当業者であれば、ハプロタイプ決定（例えば、配列決定、マイクロアレイ分析、q P C R、P C R など）のために使用される異なる検出システムの分解能が異なることを理解しているであろう。従って、任意の所与のシステムによって生成される最も有用なデータを提供する非対称性の程度、信号雑音比などを決定する際に、所与のシステムの分解能限界を考慮するべきであると想定される。

20

【0066】

以下の実施例は、本発明の特定の実施形態および態様を実証し、さらに例示するために提供されており、その範囲を限定するものとして解釈されるべきではない。

【実施例】

【0067】

実施例 1：試料の非対称な分配

ゲノムの長距離ハプロタイプの決定方法を評価する際に、基本的な染色体の寄与（例えば、母系由来の染色体の寄与および父系由来の染色体の寄与）からの信号を互いに識別するような試料の分配により、成功したゲノムの長距離ハプロタイプ決定において優れた結果が得られることを確認した。

30

【0068】

以下の方法は、長距離ハプロタイプの決定のために使用することができる試料の非対称な分配の生成方法の例示である。染色体 6 (C h r 6) は例示的な染色体であるが、任意の組織、細胞型、細胞株、不死化細胞または初期細胞などからの任意の染色体的に由来する遺伝物質を使用できることを理解されたい。

【0069】

当該試料は、それぞれ M 6 および P 6 と呼ばれる C h r 6 の母系および父系寄与の混合物を含有する。本実施例の目的のために、当該試料は、中期に同期された細胞に由来している。中期の同期は、本記載の方法を実施するために必要ではないが、本実施例は、2 つの親の寄与について相が特定された対立遺伝子の決定を実証するものであるため、これを達成するための 1 つの方法は、中期に同期された細胞で開始することである。

40

【0070】

試料細胞数は、蛍光活性化細胞選別機すなわち F A C S 、細胞数測定または他の公知の方法によって測定する。試料細胞数を測定したら、約 10 μ l の最終的な総細胞体積中に約 100 細胞 / μ l になるまで試料を希釈する。平均すると細胞は、M 6 および P 6 のそれぞれ 1 つのコピーを含むため、2 つの寄与の比は 1 : 1 である。希釈後、細胞を溶解し、当業者に知られている確立された技術によって、D N A を回収する。本実施例の目的のために、試料を 10 個の画分に分ける。すなわち、D N A 試料を 15 n l の 10 個の画分に分けることにより、1.5 個の染色体を含む最適な確率の試験試料を得て（図 5 ）、総数が約 670 個の可能な試験画分を得る。最終画分体積は、例えば D N A の回収前の細胞濃度の差、標的染色体成分（例えば、この場合、M 6 および P 6 ）における差および下流分析のために使用される測定技術（例えば、配列決定、マイクロアレイアッセイ、P C R

50

など)の感度に応じて異なる。試料を等分または分画する方法は変えることができ、例えば本実施例の目的のために、試料を下流用途用の所望の数のアッセイチャンバーに分画するのにマイクロ流体デバイスを使用するが、所望の体積がその方法の範囲内であれば、任意の手動もしくは自動の試料分画方法が適当である。

【0071】

全ての染色体からのハプロタイプデータが必要でない場合、例えば数個の染色体からのハプロタイプ決定用データが望まれる場合、図5を使用して下流処理のために必要な画分数を決定して、最高確率の解釈可能なデータを得ることができる。本実施例では、Chr 6からの1.5個の染色体を含む試料画分では、72%の画分で、P6対立遺伝子の検出信号からM6対立遺伝子を識別する(例えば、対立遺伝子の相の特定)あるいはその逆を可能にするのに十分な程の非対称性を有するデータが生成されることが想定されている。従って、僅か15n1の10個の画分が、使用可能なデータ(例えば、配列決定、マイクロアレイアッセイ、PCRなどの下流プロセスの分解能よりも大きな基本的な成分の非対称性)を提供する99.99%の確率の少なくとも1つの画分を有している。但し、グラフの特徴、すなわち解釈可能なデータを生成するアッセイのための画分数は、データ収集のために使用される方法(例えば、配列決定方法、マイクロアレイ法、PCR法など)の分解能に応じて変化するものと想定される。

10

【0072】

分画したら、研究者の必要性(例えば、配列決定、マイクロアレイ分析、エピジェネティック解析、PCRなど)に応じて試料を処理する。上述したように、非対称な分配方法の結果は、数多くの試料の一定分量または画分を提供し、そのうちの全てを研究者が直接分析することができ、後の分析のために適切にプールすることができる。研究者は、最高度の非対称性を含む試料の一定分量を分析することのみを望んでいることがあり、従って、最大の信号雑音比を得た時点で、研究者は、その特徴について画分の特性評価を行い、それに関する研究者の必要性に適合するもののみを使用する。非対称分配試料画分を分析するために使用することができる例示的な下流用途としては、上述したように、DNAライブラリー調製、増幅(例えば、PCR、qPCRおよびMDAなど)、マイクロアレイ分析、配列決定、遺伝子型決定およびハプロタイプ決定が挙げられるが、これらに限定されない。

20

【0073】

30

実施例2：非対称な試料の分配方法を用いたハプロタイプの決定

正常な個体からのヒトゲノムDNAを、3μlの水当たり0.5個の一倍体コピー(5.00E-07μg/μl)になるまで希釈した。希釈したゲノムDNA(3μl)を複数の管に等分して、各管に平均して0.5個のヒトゲノムの一倍体コピーを得た。各管に、3μlの緩衝液D2(0.25μlの1MのDTTを含む2.75μlのDLB緩衝液)を添加した(Qiagen REPLI-g(登録商標)UltraFast Mini Handbook、カタログ番号150035)後に、BioRad DNA Engineサーマルサイクラー(BioRad Part #PTC-0200G)で、4で10分間インキュベートした。3μlのREPLI-g UltraFast Stop Solutionを添加した後、33μlのMaster mixを添加した。Master mixは、30μlのREPLI-g UltraFast Reaction Buffetと、2μlのREPLI-g

40

UltraFast DNA Polymeraseと、6,000個のオリゴヌクレオチドを含む1μlの7.56mMのヒト化9merプール(1つのオリゴ当たり0.03μMの最終濃度)とを含んでいた。反応物をBioRad Tetrad2サーマルサイクラー(BioRad Part #PTC-0240G)で、30で90分間インキュベートした後に、試料を65で3分間加熱して、REPLI-g UltraFast DNA Polymeraseを熱不活性化した。

【0074】

製造業者の手順に従い、DNA Clean & Concentrator(商標)-5スピンカラム(Zymo Research社カタログ番号D4003)を用いて、多置換増幅(MDA)産物を精製した。2:1のDNA結合緩衝液:MDA産物の体積比を使用した。精

50

製したMDA産物を12μlの水で溶離した。

【0075】

Illumina社の手順11230143 Rev A中の指示に従い、Illumina(登録商標)300K HumanCytoSNP-12 BeadChipを用いるInfinium(登録商標)遺伝子型決定アッセイを、4μlの各精製された産物を用いて行った。iCS3.3.28 (Infinium II Assay Lab Setup and Procedures Guide, Illumina part # 11207963)に従い、iScanでBeadChipを走査した後、GenomeStudio(商標)Genotyping Module v1.0(Illumina Part # 11318815)を用いて、データをGenomeStudio(商標)2008.1 Framework内にインポートした。

【0076】

試料当たりの生のXおよびY強度の散布図は、出発物質中の任意にAおよびBと呼ばれる2つの遺伝子座の半接合体の存在、例えば(X, 0)および(0, Y)を示すために使用した。ここでは、A/AまたはA/0遺伝子型では、X軸に沿ったデータ点が得られ、B/BまたはB/0遺伝子型では、Y軸に沿ったデータ点が得られ、A/B遺伝子型では、X軸とY軸との間の対角線に沿ったデータ点が得られる。1つにまとめた全ゲノム(X, Y)データ点は、ヒトゲノムの2つ以上の一倍体コピーが出発物質中に存在していた場合に存在するヘテロ接合型対立遺伝子を示す。

【0077】

図9～図10は、ヘテロ接合型SNPをそれらの一倍体成分に分解するための方法の能力を示す。図9の散布図は、正常な二倍体個体の例示的な生強度を表しており、B/B遺伝子型遺伝子座強度がY軸(0, Y)に沿って濃縮され、A/A遺伝子型遺伝子座強度がX軸(X, 0)に沿って濃縮され、A/B遺伝子型遺伝子座強度がX軸とY軸との間の中ほど(X/Y)で濃縮されることを示している。図10は、図9の二倍体試料中のA&B遺伝子座に関する12個の希釈した試料のうちの6個の例示的な生強度を表している。A/A遺伝子座の強度データは、X軸に沿って濃縮されており、B/B遺伝子座の強度データは、Y軸に沿って濃縮されている。

【0078】

実施例3：非対称な試料の分配方法および次世代配列決定を用いた2種類の雄のゲノムDNA混合物中のX染色体デュシェンヌ型筋ジストロフィー遺伝子(DMD)のハプロタイプの決定

配列決定済みのゲノムNA18507(Bentley et al., 2008, Nature 456: 53-59)およびHG01377(Durbin et al., 2010, Nature 467: 1061-1073)(Coriell Cell Repositories社、ニュージャージー州カムデン)を含む2種類の正常な雄のゲノムDNA試料を等比で合わせて、公知のハプロタイプを有する二倍体X染色体を含む人工試料を得た。この試料を希釈し、一定分量当たり0.2個の一倍体コピー(6.00E-07 μg/μl)で、96個の一定分量の中に分注した。各一定分量の希釈した錆型DNAを、実施例2に記載したようにMDAで個々に増幅した。一定分量の半接合およびゲノムカバー率を評価するために、Illumina(登録商標)300K HumanCytoSNP-12 BeadChipで、Infinium(登録商標)遺伝子型決定により、4μlの増幅した材料をアッセイした。Nextera(商標)法を用い、製造業者の手順(Illumina社、カリフォルニア州サンディエゴ)に従って、100ナノグラムの精製されたMDA産物または50ngの希釈していないゲノムDNAを配列決定用ライプラリーに変換した。制限サイクルPCR時に各試料にバーコードを付した。

【0079】

計8個のプールの配列決定前に、最大12個の配列決定用ライプラリーをプールした。製造業者のガイドライン(Beckman Coulter Genomics社、マサチューセッツ州ダンバーズ)に従って、配列決定用ライプラリーを、0.6比でAMPure XPビーズで精製した。DMD遺伝子の1Mb隣接領域の標的化プルダウンのために、プローブプールを設計した。ビオチン化プローブは80ntの長さであり、190～370bpの間隔で、DMD遺伝子

10

20

30

40

50

の5'領域とハイブリッド形成するように設計した。プール後、TruSeq(商標)Custom Enrichment Kit(Illumina社、カリフォルニア州サンディエゴ)の手順に従って、配列決定用ライプラリーを1MbのDMD遺伝子領域のために濃縮した。濃縮してインデックスを付けたライプラリーを、75+35または75+75の読み取り長さのためにペアエンド配列決定を行い、Genome Analyzer IIx(Illumina社、カリフォルニア州サンディエゴ)で配列決定した。各レーンは、12個の試料のうちの1つのプールを含んでいた。各雄を別々に濃縮し、配列決定して、混合されたDNA内で真のハプロタイプ構造を確認し、そして、混合した試料を独立して濃縮し、配列決定して、領域内のヘテロ接合型SNPの全てを確認し、かつDMDオリゴ濃縮プールの性能を評価した。

【0080】

配列読み取り断片を分離し、Illumina CASAVA v1.8.1ソフトウェアパッケージを用いてヒトゲノムと配列比較し、インデックスを付けた各希釈試料のために、配列比較されるbamファイルを作成した。隣接領域を、SAMtools(Li et al., 2009, Bioinformatics 25: 2078-2079)を用いてbamファイルから抽出し、標的を切断した(Kitzman et al., 2011, Nat. Biotechnol 29: 59-63)。各隣接断片内で、「二倍体」配列決定データから、公知のSNP位置で塩基判定(base call)を行った。断片を連続的なホモ接合型セグメントに分けて、すなわち重なり合っているDNA断片を除去し、ReFHap(Duitama et al., 2011, Nucleic Acids Res. doi:10.1093/nar/gkr1042)を行って、ハプロタイプ決定する断片を併合してハプロタイプロックにする。得られたハプロタイプを、2つの個々の雄のgDNAの公知のハプロタイプと比較した。

10

20

【0081】

図11は、上のパネル内のHG01377(上)およびNA18507(下)に由来する個々の連続的なホモ接合型の配列比較されるセグメントならびにカリフォルニア大学Santa Cruz Genome Browserにロードされたカスタムトラックとしての下のパネル内の併合されたハプロタイプロックを示す。2つの併合されたハプロタイプロック間のギャップは、ヒトゲノム内の配列比較不可能な領域に起因する。ハプロタイプ決定される領域全体は989kbであり、平均的なハプロタイプロックのサイズは494kbである。

30

【0082】

実施例4：非対称試料分配方法および次世代配列決定を用いた全ヒトゲノムのハプロタイプ決定

30

正常な個体NA18506(Coriell Cell Repository社、ニュージャージー州カムデン)からのヒトゲノムDNAを、1μlの水当たり0.5または1.0個の一倍体コピー(1.50E-06μg/μlまたは3.00E-06μg/μl)になるまで希釈した。希釈したゲノムのDNA(1μl)を、1希釈物当たり24個の管に等分し、平均して各管に0.5または1.0個のヒトゲノムの一倍体コピーを得た。各管に、1μlの緩衝液D1(0.875μlの水を含む0.125μlのDLB緩衝液)を添加した(Qiagen REPLI-g(登録商標)UltraFast Mini Handbook、カタログ番号150035)後に、室温で3分間インキュベートした。1マイクロリットルの緩衝液N1(1.8μlの水を含む0.2μlのREPLI-g UltraFast Stop Solution)、次いで、17μlのMastermixを添加した。Mastermixは、15μlのREPLI-g UltraFast Reaction Bufferと、1μlのREPLI-g

40

UltraFast DNA Polymeraseと、1μlの水とを含んでいた。反応物をBioRad Tetrad2サーマルサイクラー(BioRad Part # PTC-0240G)で、30で90分間インキュベートした後、試料を65で3分間加熱して、REPLI-g UltraFast DNA Polymeraseを熱不活性化した。

【0083】

DNA Clean & Concentrator(商標)-5スピンカラム(Zymo Research社、カタログ番号D4003)を用い、製造業者の手順に従って、MDA産物を精製した。2:1=DNA結合緩衝液:MDA産物の体積比を使用した。精製したMDA産物を17μlの水で溶離した。Nextera酵素を100倍に希釈して転位反応(tra

50

nsposition reaction)への低いDNA鑄型インプット量を補い、かつdsDNA/Nextera酵素比を増加させたこと以外は製造業者の手順(Illumina社、カリфорニア州サンディエゴ)に従い、Nextera(商標)法を用いて、精製したMDA産物(15μl)を配列決定用ライブラリーに変換した。これにより、あまりに小さすぎるライブラリー挿入サイズの生成を防止した。制限サイクルPCR時に、各試料にバーコードを付した。計4つのプールの配列決定前に、最大12個の配列決定用ライブラリーをプールした。製造業者のガイドライン(Beckman Coulter Genomics社、マサチューセッツ州ダンバーズ)に従って、配列決定用ライブラリーを、0.6比でAMPure XPビーズで精製した。100+100の読み取り長さに対してペアエンド配列決定を用いて、HiSeq 2000配列決定システム(Illumina社、カリфорニア州サンディエゴ)でライブラリーを配列決定した。2つのレーンで、12個の試料の各プールを配列決定した。実施例4に記載されているように配列読み取り断片の分析を行った。親の遺伝子型から統計的計算により得られた分解されたハプロタイプと比較して、精度を確認した。

10

20

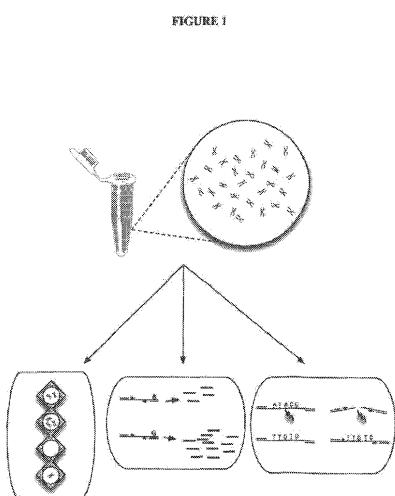
【0084】

図12は、上のパネル内の個々の連続的なホモ接合型セグメントおよび下のパネル内のカリфорニア大学Santa Cruz Genome Browserにロードしたカスタムトラックとしての併合されたハプロタイプロックの例を示す。この例は、全ゲノム二倍体試料に対してハプロタイプ決定を行うことができるることを実証している。得られた最も大きく正確なハプロタイプロックは303.5kbであり、合計1.27Gbは、分解されたハプロタイプであった。

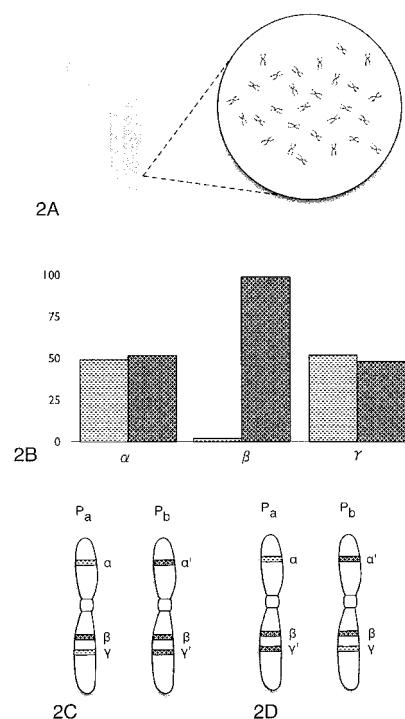
【0085】

本出願で言及されている全ての刊行物および特許は、参照により本明細書に組み込まれる。本発明の本記載の方法および組成物の各種修正形態および変形形態は、本発明の範囲および趣旨から逸脱することなく当業者には明らかであろう。本発明について、特定の好ましい実施形態と関連させて記載してきたが、特許請求される本発明は、そのような特定の実施形態に過度に限定されるべきではないことを理解されたい。実際に、関連分野において当業者に明らかである本発明を実施するための本記載の様式の各種修正形態は、以下の特許請求の範囲の範囲内に含まれることが意図されている。

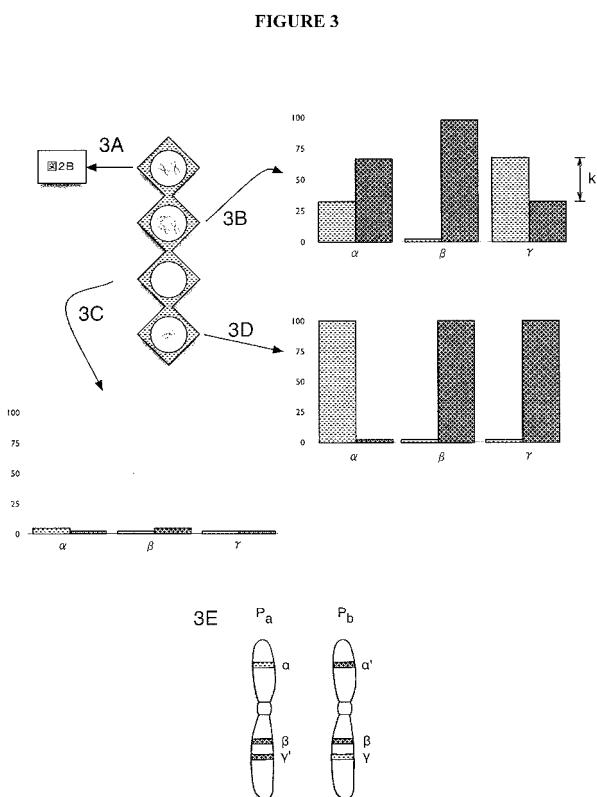
【図1】



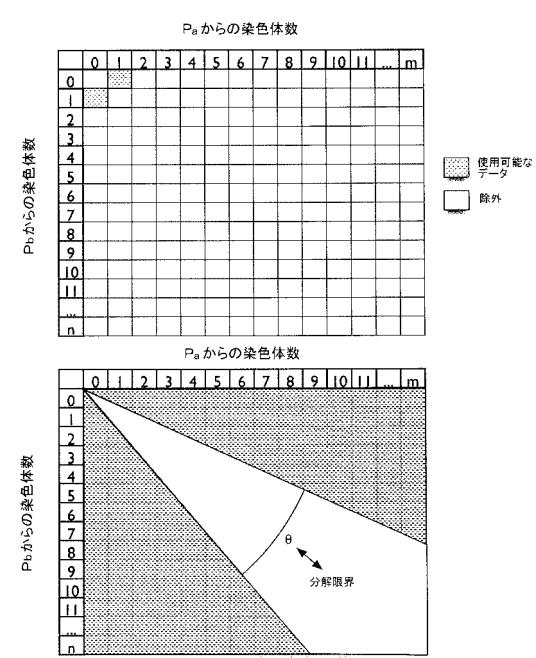
【図2】



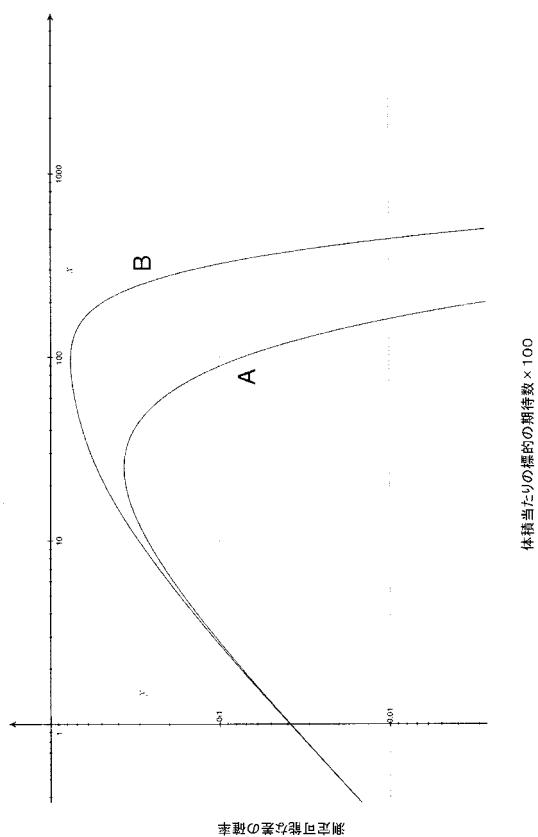
【図3】



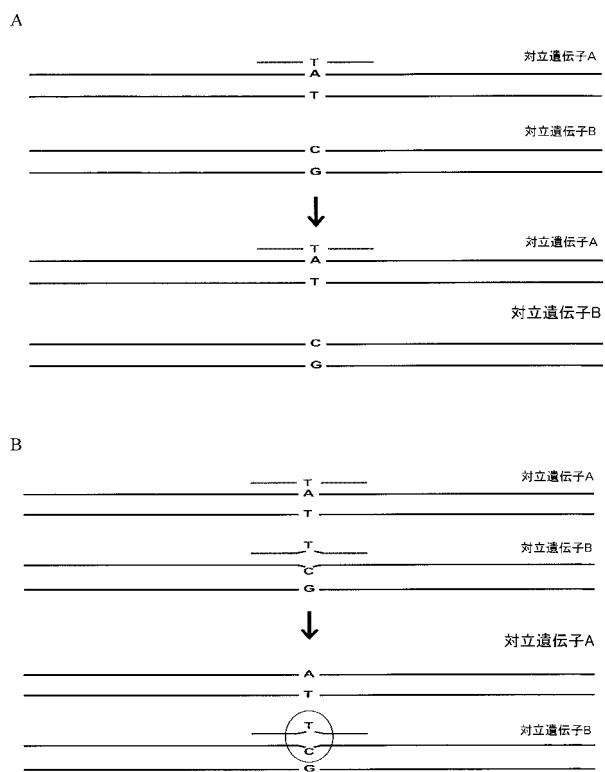
【図4】



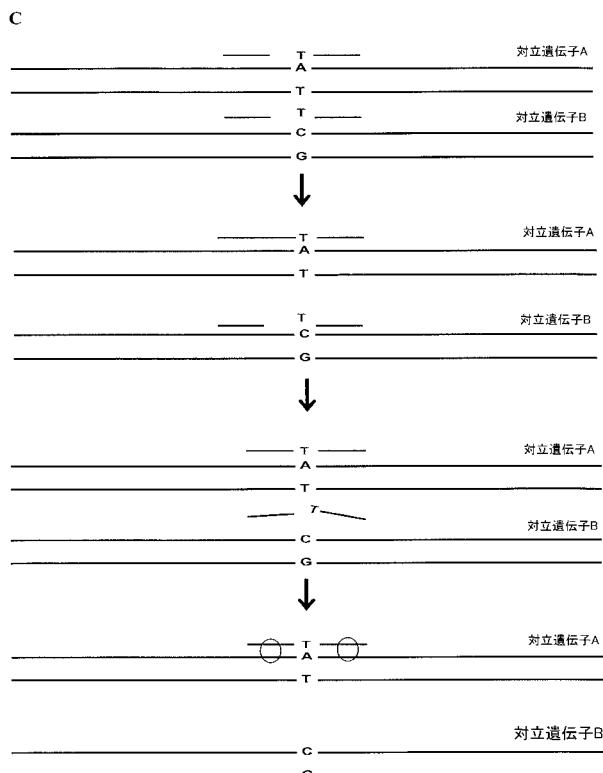
【図5】



【図6-1】

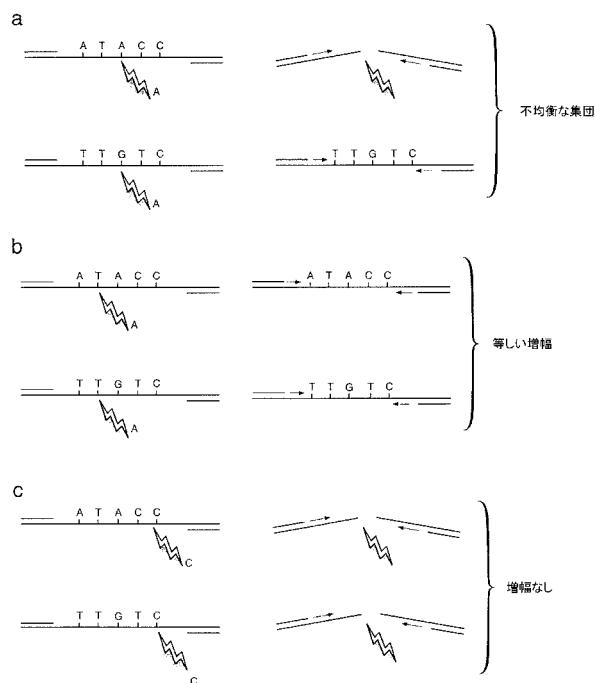


【図6-2】

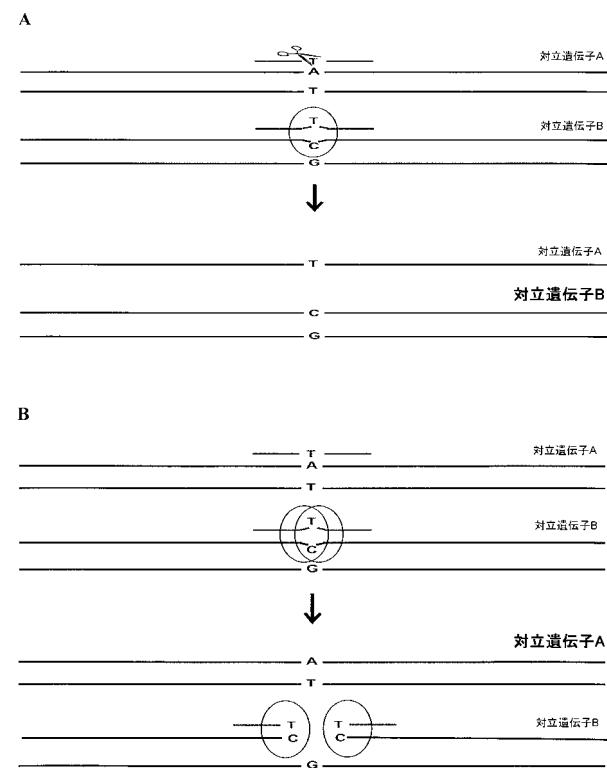


【図6-3】

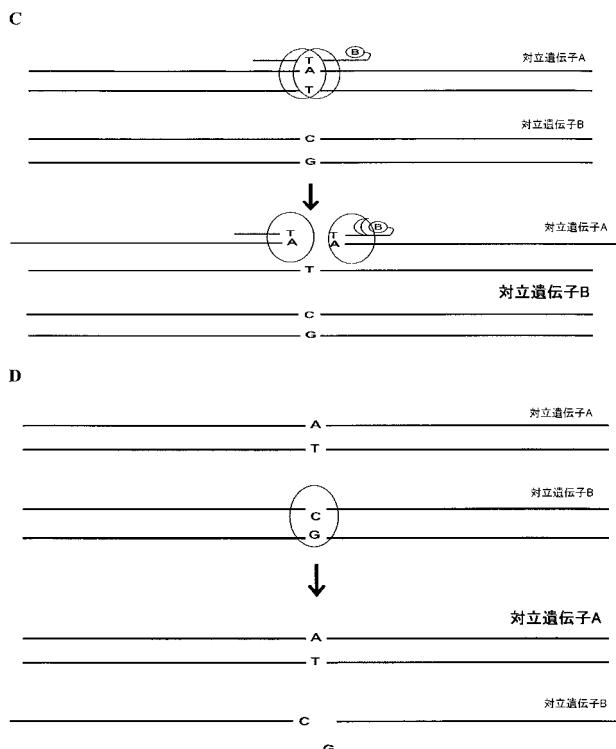
【図7】



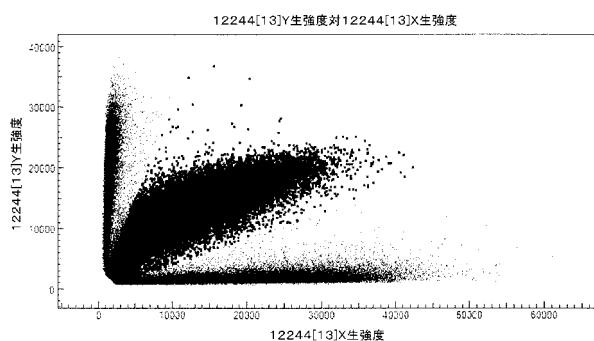
【図8-1】



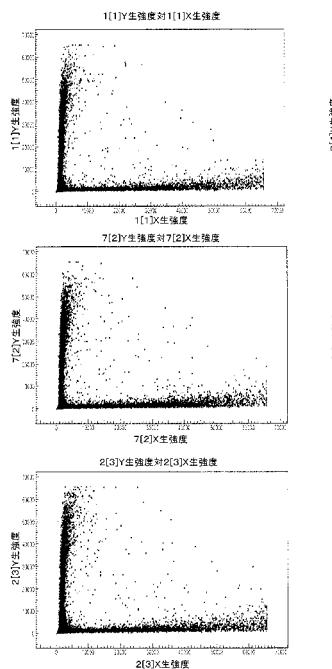
【図8-2】



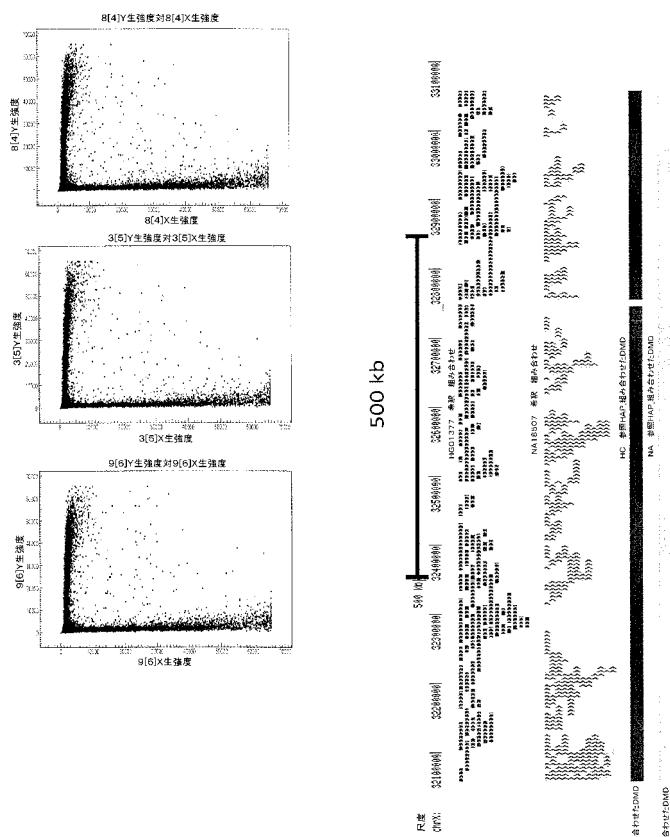
【図9】



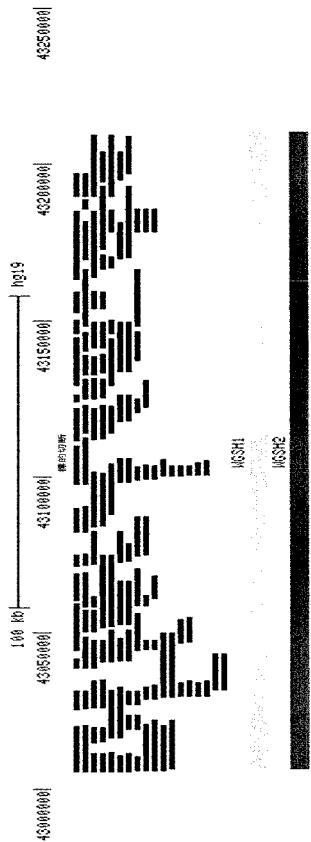
【図10】



【 図 1 1 】



【 図 1 2 】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2012/026623
A. CLASSIFICATION OF SUBJECT MATTER		
<i>C12Q 1/68(2006.01)i, C12N 15/11(2006.01)i</i>		
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) C12Q 1/68, C12M 1/00; NotA vailable		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2005-082110 A2 (ILLUMINA INC.) 09 September 2005 See abstract and claims	1-56
A	US 2008-0318235 A1 (HANDYSIDE ALAN) 25 December 2008 See abstract and claims	1-56
A	US 2010-0099092 A1 (SONG QING et al.) 22 April 2010 See abstract and claims	1-56
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* 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 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		
Date of the actual completion of the international search 26 DECEMBER 2012 (26.12.2012)		Date of mailing of the international search report 27 DECEMBER 2012 (27.12.2012)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer JEONG, Sei Joon Telephone No. 82-42-481-5592
		

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/US2012/026623

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2005-082110 A2	09.09.2005	US 2005-266432 A1 WO 2005-082110 A3	01.12.2005 12.01.2006
US 2008-0318235 A1	25.12.2008	AT 497021 T DE 602006019898 D1 EP 1951897 A1 EP 1951897 B1 EP 1951897 B8 ES 2360085 T3 GB 0523276 D0 JP 2009-515544 A JP 2009-515544 T WO 2007-057647 A1	15.02.2011 10.03.2011 06.08.2008 26.01.2011 23.03.2011 31.05.2011 21.12.2005 16.04.2009 16.04.2009 24.05.2007
US 2010-0099092 A1	22.04.2010	CA 2740205 A1 CN 102203285 A EP 2344667 A1 WO 2010-044923 A1	22.04.2010 28.09.2011 20.07.2011 22.04.2010

フロントページの続き

(51) Int.CI.

F I

テーマコード(参考)

G 0 1 N 33/50

P

(81) 指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,R0,RS,SE,SI,SK,SM,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,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,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,PE,PG,PH,PL,PT,QA,RO,RS,RW,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN

(72) 発明者 フィッシャー, ジェフリー エス.

アメリカ合衆国, カリフォルニア州 92122, サンディエゴ, 5200 イルミナ ウェイ

(72) 発明者 ケイパー, フィオナ

アメリカ合衆国, カリフォルニア州 92122, サンディエゴ, 5200 イルミナ ウェイ

F ターム(参考) 2G045 AA35 DA13 FB12

4B024 AA11 AA20 CA01 HA12 HA14 HA20

4B063 QA01 QA17 QA18 QQ02 QQ08 QQ42 QR32 QR40 QR56 QR62

QR72 QR77 QS24 QS34 QS36 QX02