

(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2017-506500

(P2017-506500A)

(43) 公表日 平成29年3月9日(2017.3.9)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 Z N A A	4 B 0 6 3
G O 1 N 33/53 (2006.01)	G O 1 N 33/53 M	4 C 0 8 7
G O 1 N 33/543 (2006.01)	G O 1 N 33/53 U	
C 1 2 N 15/09 (2006.01)	G O 1 N 33/543 5 4 1 A	
A 6 1 K 35/12 (2015.01)	C 1 2 N 15/00 A	
審査請求 未請求 予備審査請求 未請求 (全 32 頁) 最終頁に続く		

(21) 出願番号	特願2016-538772 (P2016-538772)	(71) 出願人	516114503
(86) (22) 出願日	平成26年12月10日 (2014.12.10)		コネクシオ ゲノミクス ピーティーワイ
(85) 翻訳文提出日	平成28年8月5日 (2016.8.5)		リミテッド
(86) 国際出願番号	PCT/AU2014/001114		オーストラリア国、ウェスタン オースト
(87) 国際公開番号	W02015/085350		ラリア州 6 1 5 8、イースト フリーマ
(87) 国際公開日	平成27年6月18日 (2015.6.18)		ントル、ウィンザー ロード 3 6
(31) 優先権主張番号	2013904801	(74) 代理人	100080791
(32) 優先日	平成25年12月10日 (2013.12.10)		弁理士 高島 一
(33) 優先権主張国	オーストラリア (AU)	(74) 代理人	100125070
			弁理士 土井 京子
		(74) 代理人	100136629
			弁理士 鎌田 光宜
		(74) 代理人	100121212
			弁理士 田村 弥栄子
		最終頁に続く	

(54) 【発明の名称】 遺伝子アリルを同定するための方法及びプローブ

(57) 【要約】

本発明は、高度多型性核酸をジェノタイピングする方法に関する。特に、本開示は、ハイスループットシーケンシング技術を用いて、HLAアリルなどの高度多型性遺伝子アリルをジェノタイピングするための方法に関する。とりわけ、本発明は、非コード配列へとキャプチャープローブを標的化することにより、高度多型性遺伝子におけるアリルを同定するのに適したキャプチャープローブ方法に関する。

【選択図】 なし

【特許請求の範囲】

【請求項 1】

対象における遺伝子アリルを同定するための方法であって、

a) 対象由来の核酸サンプルをオリゴヌクレオチドプローブと接触させることであって、ここで、該オリゴヌクレオチドプローブが、該核酸サンプル中の遺伝子標的配列とハイブリダイズするものである、こと；

b) オリゴヌクレオチドプローブとハイブリダイズされる核酸について、オリゴヌクレオチドプローブと結合されていない核酸からそれを分離することにより、濃縮すること；及び

c) 該濃縮核酸をシーケンシングして、1つ以上の遺伝子アリルを同定すること
を含み、ここで、該遺伝子標的配列が、遺伝子の非コード領域にある、方法。 10

【請求項 2】

遺伝子が高度多型性遺伝子である、請求項1に記載の方法。

【請求項 3】

遺伝子がHLA遺伝子である、請求項2に記載の方法。

【請求項 4】

オリゴヌクレオチドプローブと結合される核酸を増幅することを含む、請求項1～3のいずれか一項に記載の方法。

【請求項 5】

HLA遺伝子エクソンをシーケンシングすることを含む、請求項3又は4に記載の方法。 20

【請求項 6】

HLA遺伝子全体をシーケンシングすることを含む、請求項5に記載の方法。

【請求項 7】

オリゴヌクレオチドプローブがキャプチャータグを含む、請求項1～6のいずれか一項に記載の方法。

【請求項 8】

キャプチャータグが、ビオチン又はストレプトアビジンである、請求項7に記載の方法。

【請求項 9】

キャプチャータグを結合剤と接触させることを含む、請求項7又は8に記載の方法。 30

【請求項 10】

結合剤が、ビオチン又はストレプトアビジンである、請求項9に記載の方法。

【請求項 11】

オリゴヌクレオチドプローブと接触させる核酸サンプルが、一本鎖核酸を含む、請求項1～10のいずれか一項に記載の方法。

【請求項 12】

核酸サンプルが、オリゴヌクレオチドプローブと接触させる前にフラグメント化される、請求項1～11のいずれか一項に記載の方法。

【請求項 13】

約100 bp超の平均長を有する核酸フラグメントをオリゴヌクレオチドプローブと接触させることを含む、請求項12に記載の方法。 40

【請求項 14】

核酸サンプルが、オリゴヌクレオチドプローブと接触させた後にフラグメント化される、請求項1～11のいずれか一項に記載の方法。

【請求項 15】

HLA遺伝子の非コード領域がイントロンである、請求項3～14のいずれか一項に記載の方法。

【請求項 16】

非HLA遺伝子の1つ以上のアリルを検出することを含む、請求項3～15のいずれか一項に記載の方法。 50

【請求項 17】

シーケンシングが、ハイスループットシーケンシングである、請求項1～16のいずれか一項に記載の方法。

【請求項 18】

ハイスループットシーケンシングが、次世代シーケンシング技術である、請求項17に記載の方法。

【請求項 19】

対象における遺伝子アリルを同定するための方法であって、

- a) 対象から核酸サンプルを得ること；
 - b) 該核酸サンプルをフラグメント化して、少なくとも約2 kb長の核酸フラグメントを得ること；
 - c) 該核酸サンプルを、キャプチャータグを含むオリゴヌクレオチドプローブと接触させることであって、ここで、該オリゴヌクレオチドプローブが、該核酸サンプル中の遺伝子標的配列とハイブリダイズするものである、こと；
 - d) 該キャプチャータグを結合剤と接触させ、オリゴヌクレオチドとハイブリダイズされる核酸について濃縮すること；及び
 - e) 該核酸をシーケンシングして、1つ以上の遺伝子アリルを同定すること
- を含み、ここで、該遺伝子標的配列が、遺伝子の非コード領域にある、方法。

【請求項 20】

核酸サンプルが、オリゴヌクレオチドプローブと接触させる前にフラグメント化される、請求項19に記載の方法。

【請求項 21】

核酸サンプルが、オリゴヌクレオチドプローブと接触させた後にフラグメント化される、請求項19に記載の方法。

【請求項 22】

遺伝子アリルがHLA遺伝子アリルである、請求項19～21のいずれか一項に記載の方法。

【請求項 23】

オリゴヌクレオチドプローブの1つ以上が、配列番号1～8又は10～71のいずれか1つと少なくとも95%同一である核酸配列を含む、請求項3～33のいずれか一項に記載の方法。

【請求項 24】

移植を必要とするレシピエントのために移植ドナーを同定する方法であって、

- a) 請求項3～23のいずれか一項に記載の方法を実施して、移植を必要とするレシピエントにおける1つ以上のHLA遺伝子アリルを同定すること；並びに
 - b) 移植ドナー及び移植レシピエントの両方のHLA遺伝子における1つ以上のアリルの存在に基づき移植ドナーを同定すること
- を含む、方法。

【請求項 25】

移植レシピエントが移植片対宿主病を発症する可能性を低下させる方法であって、

- a) 請求項3～24のいずれか一項に記載の方法を実施して、移植を必要とするレシピエントにおける1つ以上のHLA遺伝子アリルを同定すること；並びに
 - b) 移植ドナー及び移植レシピエントの両方における1つ以上のHLA遺伝子アリルの存在に基づき移植ドナーを同定すること
- を含み、ここで、移植レシピエント及び移植ドナーの両方における1つ以上のHLA遺伝子アリルの存在が、移植ドナー由来の移植片を移植後に移植レシピエントが移植片対宿主病を発症する可能性の低下を示す、方法。

【請求項 26】

移植ドナーにおける1つ以上のHLA遺伝子アリルを同定することを含む、請求項24又は25に記載の方法。

【請求項 27】

レシピエントに同種異系移植片を移植する方法であって、

- a) 請求項24～26のいずれか一項に記載の方法を実施すること；及び
- b) 移植ドナーからドナー移植片を取り出すこと；及び
- c) 該移植片をレシピエントに移植すること

を含む、方法。

【請求項 28】

対象における遺伝子アリルを同定するためのキットであって、遺伝子の非コード領域における遺伝子標的配列とハイブリダイズするオリゴヌクレオチドプローブを含む、キット。

【請求項 29】

遺伝子が高度多型性遺伝子である、請求項28に記載のキット。

10

【請求項 30】

遺伝子がHLA遺伝子である、請求項29に記載のキット。

【請求項 31】

HLA遺伝子の非コード領域がイントロンである、請求項30に記載のキット。

【請求項 32】

オリゴヌクレオチドプローブがキャプチャータグを含む、請求項28～31のいずれか一項に記載のキット。

【請求項 33】

キャプチャータグが、ビオチン又はストレプトアビジンである、請求項32に記載のキット。

20

【請求項 34】

結合剤を含む、請求項33に記載のキット。

【請求項 35】

結合剤が基材と結合される、請求項34に記載のキット。

【請求項 36】

基材が磁気基材である、請求項35に記載のキット。

【請求項 37】

オリゴヌクレオチドプローブの1つ以上が、配列番号1～8又は10～71のいずれか1つと少なくとも95%同一である核酸配列を含む、請求項28～36のいずれか一項に記載のキット。

【発明の詳細な説明】

30

【技術分野】

【0001】

分野

本発明は、高度多型性 (highly polymorphic) 核酸をジェノタイピングする方法に関する。特に、本開示は、ハイスループットシーケンシング技術を用いて、HLAアリル (alleles) などの高度多型性遺伝子アリルをジェノタイピングするための方法に関する。

【背景技術】

【0002】

背景

いわゆる次世代シーケンシング (next generation sequencing、NGS) 技術は、遺伝子研究を加速させ、これまで未診断の遺伝的欠陥を迅速に特徴づけることができ、その結果、早期の疾患診断及び健康アウトカムの改善をもたらされている。次世代シーケンシング技術の鍵となる特徴は、数百もの単一DNA分子を同時にシーケンシングする能力である。シーケンシングの並行性及びスピードは、現在では、わずか数日でヒトゲノム全体をシーケンシングすることが可能なほどである。NGSにより複雑なゲノム全体をシーケンシングできることに加えて、多くのアプリケーションが、NGSの臨床的可能性、及び特定の標的化領域における遺伝的バリエーションの検出に重点を置いている。

40

【0003】

標的化遺伝子シーケンシング方法の一つは、標的領域と結合する標識化プローブの使用を伴い、該領域はその後ゲノムDNAから抽出することができる。

50

【0004】

使用されるプローブは、天然のRNA又はDNAであり得、通常、ほぼ100ヌクレオチド長程度である。

【0005】

キャプチャープローブの利用は、健常な個体においてあまり多型性を示す傾向にない遺伝領域に有用である。すなわち、健常な個体間での遺伝子配列の差異は、ほぼ<1%程度である。プローブと標的DNA間に配列上の差異が存在して、プローブが標的配列とアニーリングすることを妨げる場合、重複プローブの高いリダンダンシーを用いることで、該遺伝的差異と隣接する他のプローブが存在することとなり、バリエーションDNAフラグメントの抽出及びシーケンシングが可能となることが保証される。

10

【0006】

キャプチャープローブの使用は広範な適用を有するものの、その使用は普遍的ではなく、HLA系遺伝子など高度多型性遺伝子のジェノタイピングには不適切であると考えられる。HLAは、病原体に対する免疫反応の中心を担う。HLA分子は細胞表面に存在し、フラグメント化された病原体ペプチドを免疫系に提示する。個体の大半は、ユニークなHLAタイプを有しており、それにより、事実上いずれの病原体の攻撃に対しても種が防御し得ることが可能となる。

【0007】

我々の種に対するHLA多型性の利点は明確であるが、対照的な面として、造血幹細胞（HSC）を含む移植臓器上のHLA抗原が、レシピエントの免疫系により「外来」と認識され、結果として臓器拒絶反応がもたらされる。HSC移植の場合、移植される幹細胞は、新たな宿主を標的とし、その結果、移植片対宿主病（GvHD）がもたらされ、HSC移植について潜在的に致命的な帰結となる。

20

【0008】

従って、HSCドナーと患者間の適合性を保証するために、HLAジェノタイピングが必要とされる。複数のHLA遺伝子が存在しており、それらはすべて、ヒト第6染色体の短腕上に位置する主要組織適合遺伝子複合体（Major Histocompatibility Complex、MHC）として知られる領域にておよそ4百万塩基の範囲内に存在する。当該遺伝子は、2セットに分けることができる。HLA-A、HLA-B及びHLA-Cを含むクラスI遺伝子と、HLA-DRB1、HLA-DQB1及びHLA-DPB1を含むクラスII遺伝子である。移植目的のために通常タイピングされるのはこれらの遺伝子であるが、MHCはまた、さらなるクラスI HLA遺伝子も含み、それらはあまり多型性を示さず、病原体由来ペプチド提示以外の機能を有する。これらの遺伝子には、HLA-E、-F及び-Gが含まれる。さらに、多数のクラスI及びクラスII様偽遺伝子及び遺伝子フラグメントが存在する。

30

【0009】

古典的クラスI遺伝子は配列が類似している。それらはすべて、7~8のエクソンと小さなイントロンを特徴とするおよそ3,200塩基対の比較的小さな遺伝子である。クラス2遺伝子はより大きなイントロンを有し、ほぼ20,000塩基対程度のものもある。

【0010】

高レベルのHLA多型は、同一及び異なる座位のアリルから配列モチーフを「共有（sharing）」することにより生成される。比較的稀なユニークな多型及びSNPsが存在する。換言すれば、新たに同定されるアリルは、同じ座位の他のアリルで見られる、又は他の関連する座位のアリルから見られるモチーフのユニークな組み合わせを含むだろう。モチーフの共有は、多型性の急速な進化に有効なメカニズムであり、病原体と遭遇した際に種に都合よく作用する。アリル及び座位間で共有されるモチーフのサイズは変化する。いくつかの場合、モチーフは20 bp未満であり得るが、他のアリルは、遺伝子全体の半分を含み得る。高レベルの多型性の結果として、いくつかの場合、HLA適合HSCドナーを同定することが不可能となる。

40

【0011】

HLAについてのNGSは、従来のサンガーベースのシーケンシング技術を超える多大な恩恵

50

を提供する。従来のHLAシーケンシング技術は、ポリメラーゼ連鎖反応により増幅されるDNAのシーケンシングを必要とする。増幅されたDNAは、複数の配列分子としてシーケンシングされ、隣接するヌクレオチドが同一又は異なる（母系又は父系）染色体由来であるかどうかを決定することができない。配列をフェージング（phase）できないことは、高頻度で、不正確なタイピングへとつながり、不明瞭なタイプを解決するために、その後の実験が必要となる。

【0012】

NGSは、フェーズインフォメーション（phase information）を備えた単一分子配列データを生成し、これはAssign MPSなどの配列解析ソフトウェアに従って、明瞭なHLAタイプを生じる可能性があり得る。しかしながら、PCR産物をシーケンシングする場合、PCRにより組み込まれたエラー及びキメラ分子もシーケンシングされ、配列解析が困難なものとなる。

10

【0013】

しかしながら、従来のエクソンベースのプロープキャプチャー技術は、HLAタイプがモチーフシャッフリングにより広範に生成されること、及び配列の差異の程度が、1つのアリルから次のものへとかなりのもの（substantial）となり得ることから、信頼できない可能性が高い。

【発明の概要】

【発明が解決しようとする課題】

【0014】

20

従って、HLA遺伝子などの高度多型性遺伝子のハイスループットシーケンシングを可能にする方法が必要とされている。

【課題を解決するための手段】

【0015】

要旨

本発明者らは、非コード配列へとキャプチャープロープを標的化することにより、高度多型性遺伝子におけるアリルを同定するのに適したキャプチャープロープ方法を記述する。一般にタイピングされるHLAクラスI及びクラスII分子の非コード領域は、多型性エクソンよりも多型性が低く、異なるHLAタイプで潜在的に代替される配列の数は、通常のバリエーション生成プロセスを介して生じるランダムSNPsを除いて、比較的限りがあるようである。従って、少数のプロープを単一領域で設計でき、これは、すべてのアリル（将来的に記載される可能性の高いものを含む）のキャプチャーを可能にするだろう。キャプチャープロープは、多型性エクソンを含むDNAフラグメントをキャプチャーし、キャプチャーDNAフラグメントのシーケンシングにより、多型性フラグメントの配列が生成され、HLAタイプの推論（deduction）が可能となるだろう。

30

【0016】

従って、一側面において、本発明は、対象における遺伝子アリルを同定するための方法であって、

a) 対象由来の核酸サンプルをオリゴヌクレオチドプロープと接触させることであって、ここで、該オリゴヌクレオチドプロープが、該核酸サンプル中の遺伝子標的配列とハイブリダイズするものである、こと；

40

b) オリゴヌクレオチドプロープとハイブリダイズされる核酸について、オリゴヌクレオチドプロープと結合されていない核酸からそれを分離することにより、濃縮（enriching）すること；及び

c) 該核酸をシーケンシングして、1つ以上の遺伝子アリルを同定することを含み、ここで、該遺伝子標的配列が、遺伝子の非コード領域にある、方法を提供する。

【0017】

本発明の方法は、高度多型性遺伝子におけるアリルを同定するために特に有用である。従って、一実施態様において、遺伝子は、高度多型性遺伝子である。特定の一実施態様において、遺伝子はHLA遺伝子である。

50

【0018】

当業者ならば、シーケンシングのためにキャプチャーされる核酸を調製する際に、キャプチャーされる核酸を増幅することが望ましい場合があることを理解するだろう。従って、一実施態様において、当該方法は、オリゴヌクレオチドプローブと結合される核酸 (nucleic) を増幅することを含む。

【0019】

一実施態様において、当該方法は、HLA遺伝子エクソンをシーケンシングすることを含む。

【0020】

別の実施態様において、当該方法は、HLA遺伝子全体をシーケンシングすることを含む。

10

【0021】

関心のある核酸の濃縮を容易にするために、一実施態様において、オリゴヌクレオチドプローブはキャプチャータグを含む。

【0022】

一実施態様において、キャプチャータグはハイブリダイゼーションタグである。

【0023】

別の実施態様において、キャプチャータグは、ビオチン又はストレプトアビジンである。

【0024】

さらに別の実施態様において、当該方法は、キャプチャータグを結合剤と接触させることを含む。

20

【0025】

一実施態様において、結合剤はビオチン又はストレプトアビジンである。

【0026】

さらに別の実施態様において、結合剤は基材と結合 (coupled) される。特定の一実施態様において、結合剤は、磁気基材 (magnetic substrate)、例えば磁気ビーズなどと結合される。

【0027】

一実施態様において、オリゴヌクレオチドプローブと接触させる核酸サンプルは、一本鎖核酸を含む。

30

【0028】

さらに別の実施態様において、核酸サンプルは、オリゴヌクレオチドプローブと接触させる前にフラグメント化される。

【0029】

一実施態様において、当該方法は、約100 bp超の平均長を有する核酸フラグメントをオリゴヌクレオチドプローブと接触させることを含む。

【0030】

別の実施態様において、当該方法は、約2 kb超の平均長を有する核酸フラグメントをオリゴヌクレオチドプローブと接触させることを含む。

40

【0031】

或いは、核酸サンプルは、オリゴヌクレオチドプローブと接触させた後にフラグメント化される。

【0032】

オリゴヌクレオチドプローブは、遺伝子の非コード領域、例えば、イントロン及び遺伝子間領域などを標的化する。一実施態様において、遺伝子 (例えば、HLA遺伝子) の非コード領域はイントロンである。

【0033】

別の実施態様において、当該方法は、非HLA遺伝子の1つ以上のアレルを検出することをさらに含む。

50

【0034】

さらに別の実施態様において、シーケンシングは、ハイスループットシーケンシングである。

【0035】

一実施態様において、ハイスループットシーケンシングは、次世代シーケンシング技術である。

【0036】

第二の側面において、本発明は、対象における遺伝子アリルを同定するための方法であって、

a) 対象から核酸サンプルを得ること；

b) 該核酸サンプルをフラグメント化して、少なくとも約2 kb長の核酸フラグメントを得ること；

c) 該一本鎖核酸を、キャプチャータグを含むオリゴヌクレオチドプローブと接触させることであって、ここで、該オリゴヌクレオチドプローブが、該核酸サンプル中のHLA標的配列とハイブリダイズするものである、こと；

d) 該キャプチャータグを結合剤と接触させることにより、オリゴヌクレオチドプローブとハイブリダイズされる核酸について濃縮すること；及び

e) 該核酸をシーケンシングして、1つ以上の遺伝子アリルを同定すること
を含み、ここで、該HLA標的配列が、遺伝子の非コード領域にある、方法を提供する。

【0037】

本発明の方法で使用される核酸は、予め得られたサンプル由来であってよいが、一実施態様において、当該方法は、対象から得られた生物学的サンプルから核酸を単離することを含む。

【0038】

一実施態様において、核酸サンプルは、オリゴヌクレオチドプローブと接触させる前にフラグメント化される。

【0039】

別の実施態様において、核酸サンプルは、オリゴヌクレオチドプローブと接触させた後にフラグメント化される。

【0040】

さらに別の実施態様において、遺伝子アリルは、HLA遺伝子アリルである。

【0041】

第一又は第二の側面の一実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと少なくとも95%同一である核酸配列を含む。

【0042】

別の実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと少なくとも98%同一である核酸配列を含む。

【0043】

特定の一実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと同一である核酸配列を含む。

【0044】

第三の側面において、本発明は、移植を必要とするレシピエントのために移植ドナーを同定する方法であって、

a) 本発明の方法を実施して、移植を必要とするレシピエントにおける1つ以上のHLA遺伝子アリルを同定すること；並びに

b) 移植ドナー及び移植レシピエントの両方のHLA遺伝子における1つ以上のアリルの存在に基づき移植ドナーを同定すること
を含む、方法を提供する。

【0045】

本発明はさらに、移植レシピエントが移植片対宿主病を発症する可能性を低下させる方

10

20

30

40

50

法であって、

a) 本発明の方法を実施して、移植を必要とするレシピエントにおける1つ以上のHLA遺伝子アリルを同定すること；並びに

b) 移植ドナー及び移植レシピエントの両方における1つ以上のHLA遺伝子アリルの存在に基づき移植ドナーを同定すること

を含み、ここで、移植レシピエント及び移植ドナーの両方における1つ以上のHLA遺伝子アリルの存在が、移植ドナー由来の移植片を移植後に移植レシピエントが移植片対宿主病を発症する可能性の低下を示す、方法を提供する。

【0046】

一実施態様において、当該方法は、移植ドナーにおける1つ以上のHLA遺伝子アリルを同定することを含む。

10

【0047】

第四の側面において、本発明は、レシピエントに同種異系移植片 (allogeneic graft) を移植する方法であって、

a) 本発明の方法を実施すること；及び

b) 移植ドナーからドナー移植片を取り出すこと；及び

c) 該移植片をレシピエントに移植すること

を含む、方法を提供する。

【0048】

第五の側面において、本発明は、対象における遺伝子アリルを同定するためのキットであって、遺伝子の非コード領域における遺伝子標的配列とハイブリダイズするオリゴヌクレオチドプローブを含む、キットを提供する。

20

【0049】

一実施態様において、遺伝子は、高度多型性遺伝子である。

【0050】

特定の一実施態様において、遺伝子はHLA遺伝子である。

【0051】

さらに別の実施態様において、HLA遺伝子の非コード領域はイントロンである。

【0052】

一実施態様において、オリゴヌクレオチドプローブはキャプチャータグを含む。

30

【0053】

一実施態様において、キャプチャータグは、ビオチン又はストレプトアビジンである。

【0054】

さらに別の実施態様において、当該キットは結合剤を含む。

【0055】

特定の一実施態様において、結合剤は、ストレプトアビジン又はビオチンである。

【0056】

さらに別の実施態様において、結合剤は基材と結合される。特定の一実施態様において、基材は磁気基材である。特定の一実施態様において、基材は磁気ビーズである。

【0057】

さらに別の実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと少なくとも98%同一である核酸配列を含む。

40

【0058】

別の実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと少なくとも98%同一である核酸配列を含む。

【0059】

特定の一実施態様において、オリゴヌクレオチドプローブの1つ以上は、配列番号1～8又は10～71のいずれか1つと同一である核酸配列を含む。

【0060】

明らかなように、本発明の一側面の好ましい特徴及び特性は、本発明の多くの他の側面

50

に適用可能である。

【0061】

本明細書を通して、語「含む (comprise)」、又は「含む (comprises)」若しくは「含むこと (comprising)」などの変形は、記載される要素、整数若しくは工程、又は要素、整数若しくは工程のグループの包含を意味するが、任意の他の、要素、整数若しくは工程、又は要素、整数若しくは工程のグループの除外を意味するものではないことが理解されるだろう。

【0062】

以下の非限定的な実施例により、また添付の図面を参照して、本発明を以下に記載する。

10

【図面の簡単な説明】

【0063】

【図1-1】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する (例えば、A*01:01:01:01)。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

20

【図1-2】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する (例えば、A*01:01:01:01)。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

30

【図1-3】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する (例えば、A*01:01:01:01)。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一の

40

50

プローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

【図1-4】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

10

【図1-5】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

20

【図1-6】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

30

40

【図1-7】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し

50

、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

【図1-8】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

10

【図1-9】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

20

【図1-10】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

30

40

【図1-11】異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメント。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化 (synonymous changes) を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N

50

=Null又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

【図2】多型性コード領域をキャプチャー及びシーケンシングするための、低多型性非コード領域に対するキャプチャープローブの使用。

【図3】HLA-A遺伝子構造であって、原理実験の証明におけるエクソンとキャプチャープローブ位置とを示す。

【図4】標的化キャプチャーシーケンシングプロトコルを用いて作成された配列データの解析 (色は示さず)。1: HLAについての遺伝子構造を示し、その下に、遺伝子にわたる配列のデプス (depth) を要約する白色から赤色の陰影付き領域を示す。白色領域は、配列リードデプスが<50であることを示し、陰影が赤色に近づくほど配列リードデプスは少なくなる。黒色のブロックは、この領域にまたがる配列ファイルがないことを示す。2: これは、サンプル名 (Sample 01) 及び解析されている座位 (HLA-A=gA) を示す。3: この領域はデータを要約する。灰色の縦棒は各々、その位置での配列の量を示す。このスケールにより示される最大カバレッジ (maximum coverage) は100である。着色ボックスの横線はコンセンサス配列であり、ここでは、赤色=T、黒色=G、青色=C及び緑色=Aである。下部の着色ボックスは、その位置で見られる他の配列を示し、それらが灰色縦棒上にある場合、その位置にその塩基を有するリードのパーセンテージを示す。これらの配列は、通常、ベースコールエラーを示すが、ヘテロ接合性の配列も示し得る。黒色の横線は、ベースコールのパーセンテージを示し、従って、最初の横線上の青色ボックスは、その位置にてベースコールの1%がCであることを示す。スケールは、2番目の横線がリードの10%を示すよう、対数で表される。4: ソフトウェアインターフェース内のこの領域は、コンセンサス配列とHLA-Aライブラリー間で最も適合するジェノタイプを示す。最初の2つのカラムは、ジェノタイプを示す。3番目のカラムは、HLA-Aのコード配列との配列ミスマッチ数を示す。次のカラムは、HLA-Aの非コード配列とのミスマッチ数を示し、5番目及び6番目のカラムは、配列データをフェージングした場合の配列ミスマッチを示す。正しいジェノタイプは、カラム3~6において0個のミスマッチを有するだろう。このケースでは、最も適合するジェノタイプは、HLA-A*02:01:01:01+A*02:01:01:01を有する。

【図5】サンガーによってもシーケンシングされた3サンプルのキャプチャープローブ濃縮後の次世代シーケンシング結果。

【0064】

配列表の要点

配列番号1~8 - オリゴヌクレオチドキャプチャープローブ

配列番号9 - HLA Aイントロン1 (HLA A Intron 1)

配列番号10~71 - HLAキャプチャープローブ配列

【発明を実施するための形態】

【0065】

詳細な説明

一般的技術及び定義

特に別に定義されない限り、本明細書で使用する全ての技術的及び科学的用語は、当技術分野 (例えば、分子遺伝学、生化学及び免疫学) の当業者によって通常理解されるのと同じ意味を有すると解釈されるべきである。

【0066】

別に指示がない限り、本発明において利用される分子遺伝学、生化学及び免疫学的技術は当業者に周知の標準的な手順である。そのような技術は、例えば以下のようなソースの文献を通して記述及び説明される: J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook and Russell., Molecular Cloning: A Laboratory Manual, 3rd edn, Cold Spring Harbour Laboratory Press (2001), R. Scope

10

20

30

40

50

s, Protein Purification - Principals and Practice, 3rdedn, Springer (1994) , T.A . Brown (editor) , Essential Molecular Biology : A Practical Approach, Volumes 1 and 2, IRL Press (1991) , D.M. Glover and B.D. Hames (editors) , DNA Cloning : A Practical Approach, Volumes 1-4, IRL Press (1995及び1996) , 及びF.M. Ausubel et al. (editors) , Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience (1988, 現在までのすべての改訂を含む) , Ed Harlow and David Lane (editors) Antibodies : A Laboratory Manual, Cold Spring Harbour Laboratory, (1988) , 及びJ.E. Coligan et al. (editors) Current Protocols in Immunology, John Wiley & Sons (現在までのすべての改訂を含む) 。

【0067】

10

用語「核酸」又は「核酸配列」又は「核酸分子」とは、一本鎖又は二本鎖形態のいずれかである、デオキシリボヌクレオチド又はリボヌクレオチド及びそのポリマーをいう。核酸との用語は、遺伝子、相補的DNA (cDNA)、メッセンジャーRNA (mRNA)、オリゴヌクレオチド及びポリヌクレオチドと互換的に使用される。

【0068】

「単離された核酸分子」とは、その天然の状態（それが仮に天然に存在する場合）で関連又は関係するヌクレオチド配列から、通常、分離されている核酸分子を意味する。好ましくは、単離された核酸は、それが通常関連する他の成分を少なくとも60%含まず、より好ましくは少なくとも75%含まず、及びより好ましくは少なくとも90%含まない。核酸は、任意の適切な既知の技術を用いて、生物学的サンプルから単離されてよい。例えば、全ゲノムDNAは、当技術分野で既知の方法及び／又は市販のキットを用いて、例えば、QIAGENにより提供されるQIAamp DNA blood Mini Kit若しくはDNeasy Blood & Tissue Kitを用いることにより、又はフェノール／クロロホルム抽出及びエタノール沈殿などの方法を用いることにより、細胞から抽出されてよい。

20

【0069】

本明細書で使用する場合、用語「約」とは、逆の記載がない限り、指定される値の+/- 20%、より好ましくは+/- 10%、又はより好ましくは約+/- 5%をいう。

【0070】

遺伝子アリルを同定するための方法

本発明者らは現在、標的化キャプチャー及びハイスループットシーケンシング技術に特に適した、遺伝子アリルを同定する方法を開発した。本発明の方法は、DNAサンプルから関心のあるゲノム領域を分離するために、対象由来の核酸サンプルをオリゴヌクレオチドプローブと接触させることを含む。先行技術の標的化キャプチャー方法は、関心のある遺伝子のコード領域をキャプチャーするオリゴヌクレオチドに頼っているが、これらの方法は、高度多型性遺伝子のジェノタイピングに対しては、現在記載されるすべてのアリルを同定すること及び新規アリルを同定することが可能なエクソンベースのキャプチャープローブの設計が困難であることに起因して、適していない。対照的に、本発明は、非コード領域に位置する遺伝子標的配列とハイブリダイズするオリゴヌクレオチドキャプチャープローブを利用しており、従って、高度多型性遺伝子アリルのハイスループットシーケンシングが可能となる。

30

40

【0071】

図1は、異なるHLAアリルが、同一及び予測可能なHLAイントロン配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメントを示す。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A*03、A*11及びA*30アリルすべてをキャプチャーするはずである。

【0072】

キャプチャープローブと相補的な配列の、キャプチャーされるDNAフラグメントにおけ

50

る位置、及び多型性エクソンとのその相対位置を示すダイアグラムを図2に示す。図3は、シーケンシングのために多型性エクソンをキャプチャーするために、どのようにキャプチャープローブを設計し配置するかを実証する。

【0073】

本明細書で使用する場合、用語「非コード配列」又は「非コード領域」とは、アリルが同定される遺伝子のタンパク質コード配列の一部でない核酸配列をいう。従って、当該用語には、イントロン、5'及び3'非翻訳領域、プロモーター、転写制御エレメント並びに／又は遺伝子間DNAへの言及が含まれる。従って、非コード領域は、関心のある遺伝子に対して内部であるか、又は関心のある遺伝子に対して外部であるかのいずれかであってよく、例えば、遺伝子間領域にあってよい。一実施態様において、オリゴヌクレオチドプローブは、関心のある遺伝子のイントロンにある遺伝子標的配列とハイブリダイズする。

【0074】

核酸の調製

本発明の方法は、当技術分野で既知の任意の適切な技術を用いて、対象から既に得られている核酸サンプルに対して実施されてよい。或いは、一実施態様において、当該方法は、対象から得られた生物学的サンプルから核酸サンプルを得ることを含む。本明細書で使用する場合、「生物学的サンプル」は、例えば、リンパ球、全血、口腔スワブ、生検サンプル若しくは凍結組織、又はゲノムDNAを含む任意の他のサンプルであってよい。ほとんどの組織ソースも分子ジェノタイピングのために使用できるものの、例えば、末梢血由来のリンパ球が最もよく使用される。非侵襲的手段を通じて得られたサンプル、例えば、チークスワブ又は唾液ベースのDNA採取によって得られたサンプルを利用することも可能である。そのようなソースからDNAを抽出するのに適した様々な方法が当技術分野で既知である。これらは、有機溶媒抽出から、シリカコートビーズ及びアニオン交換カラム上への吸収 (absorption) まで、多岐にわたる。DNA抽出のための自動システムも市販されており、良好な質、高純度のDNAを提供し得る。

【0075】

本発明の方法で使用される核酸は、一本鎖及び／又は二本鎖のゲノムDNAであってよい。いくつかの実施態様において、核酸には、少なくとも約1 kb、少なくとも約2 kb、少なくとも約5 kb、少なくとも約10 kb、若しくは少なくとも約20 kb又はそれ以上の長さを含む長い核酸が含まれてよい。長い核酸は、当技術分野で周知の様々な方法により、ソースから調製され得る。生物学的サンプルを得て、その後、そのようなサンプルから核酸を単離する方法であって、核酸分子のインテグリティが保持される (すなわち、切断又はせん断が最小化される) 方法が好ましい。例示的な方法として、更なる精製を伴わない溶解方法 (例えば、洗剤、有機溶媒、アルカリ、及び／又はプロテアーゼを用いた、化学的又は酵素的溶解方法)、更なる核酸精製を伴った又は伴わない核単離 (nuclei isolation)、沈殿ステップを用いた単離方法、固体マトリックス (例えば、シリカ系メンブレン、ビーズ、又は核酸分子と結合する改変表面 (modified surfaces))、ゲル様マトリックス (例えば、アガロース) 又は粘性溶液を用いた核酸単離方法、並びに密度勾配を用いて核酸分子を濃縮する方法が挙げられるが、これらに限定されない。

【0076】

一実施態様において、本発明の方法で使用される核酸は、所望の平均フラグメントサイズを得るためにフラグメント化される。核酸フラグメントの所要の長さは、使用されるシーケンシング技術によって決まることが当業者に理解されるだろう。例えば、Ion Torrent及びMisSeqプラットフォームが、およそ100 bp~200 bp長のフラグメントを利用する一方、Pacific Biosciences NGSプラットフォームは、最大20 kb長の核酸フラグメントを利用できる。核酸は、物理的せん断、超音波処理、制限酵素消化、又は当技術分野で既知の他の適切な技術によりフラグメント化されてよい。核酸のフラグメント化は、所望の平均長を有する核酸フラグメントを生成するために実施され得る。例えば、核酸フラグメントの長さ又は平均長は、少なくとも約100 bp、少なくとも約200 bp、少なくとも約300 bp、少なくとも約400 bp、少なくとも約500 bp、少なくとも約600 bp、少なくとも約700 bp、

少なくとも約800 bp、少なくとも約900 bp、少なくとも約1 kb、少なくとも約2 kb、少なくとも約3 kb、少なくとも約4 kb、少なくとも約5 kb、少なくとも約6 kb、少なくとも約7 kb、少なくとも約8 kb、少なくとも約9 kb、少なくとも約10 kb、少なくとも約11 kb、少なくとも約12 kb、少なくとも約15 kb、又は少なくとも約20 kbであってよい。

【0077】

一実施態様において、核酸サンプルは、該サンプルをオリゴヌクレオチドプローブと接触させる前に、一本鎖核酸を生成するため、又は一本鎖領域を含む核酸を生成するために処理される。核酸は、当技術分野で既知の技術を用いて一本鎖とされることができ、それには、例えば、既知のハイブリダイゼーション技術及び市販のキット、例えば、ReadyAmp™ Genomic DNA Purification System (Promega) などが含まれる。或いは、適切なニッ

10

【0078】

オリゴヌクレオチドプローブ

本発明に従う用語「プローブ」又は「オリゴヌクレオチドプローブ」とは、関心のある核酸と特異的にハイブリダイズするように設計されるオリゴヌクレオチドをいう。好ましくは、プローブは、標的キャプチャー技術を用いたハイスループット (next-gen) シーケンシングのための核酸調製に使用するのに適している。従って、一実施態様において、プローブは、標的キャプチャーに適してよく、およそ60~120ヌクレオチド長であってよい。或いは、プローブは約10~25ヌクレオチドであってよい。ある実施態様において、プローブの長さは、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24又は25ヌクレオチドである。本発明で使用されるようなオリゴヌクレオチドプローブは、リボヌクレオチド、デオキシリボヌクレオチド、及び修飾ヌクレオチド、例えば、イノシン、又はそれらのハイブリダイゼーション特性を本質的に変化させない修飾基を含有するヌクレオチドなどであってよい。

20

【0079】

或いは、オリゴヌクレオチドプローブは、ヌクレオチドアナログを含んでよい。用語「ヌクレオチドアナログ」とは、別のところで一般に記載されるように、修飾ヌクレオチド塩基部分、修飾ペントース部分、及び／又は修飾ホスフェート部分を有し、ポリヌクレオチドの場合には修飾ヌクレオチド間結合を有する、合成アナログをいう (例えば、Scheit, Nucleotide Analogs, John Wiley, New York, (1980); Agarwal, Protocols for Poly nucleotides and Analogs, Humana Press, (1994))。一般に、修飾ホスフェート部分は、リン原子が+5酸化状態であり、1つ以上の酸素原子が非酸素部分 (例えば、硫黄) と置き換えられたホスフェートアナログを含む。例示的なホスフェートアナログとして、ホスホロチオエート、ホスホロジチオエート、ホスホロセレノエート (phosphoroselenoate)、ホスホロジセレノエート (phosphorodiselenoate)、ホスホロアニロチオエート (phosphoroanilothioate)、ホスホルアニリデート (phosphoranilidate)、ホスホルアミデート (phosphoramidate)、ボロノホスフェート (boronophosphates) (関連するカウンターイオン (例えば、 H^+ 、 NH_4^+ 、 Na^+) が存在する場合には、そのようなカウンターイオンを含む) が挙げられるが、これらに限定されない。例示的な修飾ヌクレオチド塩基部分として、5-メチルシトシン (5mC); C-5-プロピニルアナログ (C-5プロピニル-C及びC-5プロピニル-Uを含むが、これらに限定されない); 2,6-ジアミノプリン (2-アミノ アデニン又は2-アミノ-dAとしても知られる); ヒポキサンチン、プソイドウリジン、2-チオピリミジン、イソシトシン (isoC)、5-メチル isoC、及びイソグアニン (isoG; 例えば、米国特許第5,432,272号を参照) が挙げられるが、これらに限定されない。例示的な修飾ペントース部分として、ロックド核酸 (locked nucleic acid、LNA) アナログ (Bz-A-LNA、5-Me-Bz-C-LNA、dmf-G-LNA、及びT-LNAを含むが、限定されない)、並びに2'位又は3'位が、水素、ヒドロキシ、アルコキシ (例えば、メトキシ、エトキシ、アリルオキシ、イソプロポキシ、ブトキシ、イソブトキシ、及びフェノキシ)、アジド、アミノ、アルキルアミノ、フルオロ、クロロ、又はプロモである2'又は3'修飾物が挙げられるが、これらに限定されない。修飾ヌクレオチド間結合には、ホスフェートアナログ、アキラル及び非荷

30

40

50

電サブユニット間結合を有するアナログ、並びにアキラルサブユニット間結合を有する非荷電モルホリノベースのポリマー（例えば、米国特許第5,034,506号を参照）が含まれる。いくつかのヌクレオチド間結合アナログには、モルホリデート（morpholidate）、アセタール、及びポリアミド結合複素環が含まれる。ペプチド核酸（偽相補的（pseudocomplementary）ペプチド核酸を含む）（「PNA」）として知られるヌクレオチドアナログの1つのクラスでは、従来の糖及びヌクレオチド間結合が、2-アミノエチルグリシンアミド骨格ポリマーと置き換えられている。

【0080】

ハイブリダイゼーション

本明細書で使用する場合、用語「ハイブリダイゼーション」とは、オリゴヌクレオチドプローブが標的核酸と非共有結合的に結合して、安定な二本鎖ポリヌクレオチドを形成するプロセスをいう。ハイブリダイゼーション条件は、典型的には、約1 M未満、より一般的には約500 mM未満の塩濃度を含み、約200 mM未満であってよい。ハイブリダイゼーションバッファーは、5%SSPEなどの緩衝塩溶液、又は当技術分野で既知の他のそのようなバッファーを含む。ハイブリダイゼーション温度は、5℃まで低い温度であり得るが、典型的には22℃より高く、より典型的には約30℃より高く、典型的には37℃を超える。ハイブリダイゼーションは、通常、ストリンジェントな条件下、すなわち、プローブが、それと相補的なその標的配列とはハイブリダイズするが、他の非相補的な配列とはハイブリダイズしない条件下で実施される。

【0081】

本明細書で使用する場合、用語「相補的」及び文法上の同義語（equivalents）は、1つの核酸と別の核酸（修飾核酸及び核酸アナログを含む）とのヌクレオチド塩基対相互作用であって、その結果、二重、三重又は他の高次構造の形成をもたらされるものをいう。一次相互作用は、典型的には、ヌクレオチド塩基特異的、例えば、A:T、A:U、及びG:Cなどであり、ワトソン-クリック及びフーグスティーン型水素結合による。オリゴヌクレオチドプローブが、標的核酸の相補的又は実質的に相補的な領域とアニーリングする条件は、当技術分野で周知である。

【0082】

当技術分野で理解されるように、ストリンジェントなハイブリダイゼーション条件は配列依存的であり、状況が異なれば条件も異なる。例えば、より長いフラグメントは、短いフラグメントよりも、特異的ハイブリダイゼーションのために高いハイブリダイゼーション温度を必要とし得る。相補鎖の塩基組成及び長さ、有機溶媒の存在、並びに塩基ミスマッチの程度を含む他の要因もハイブリダイゼーションのストリンジェンシーに影響を及ぼし得るので、いずれか1つのパラメータ単独の絶対値よりもパラメータの組み合わせの方が重要である。一般に、ストリンジェントな条件は、所定のイオン強度及びpHにおけるその特異的配列についてのT_mよりも約5℃下回るように選択される。例示的なストリンジェントな条件には、少なくとも25℃の温度、及び約7.0～約8.3のpHにて、少なくとも0.01 Mから1 M以下のナトリウムイオン濃度（又は他の塩）の塩濃度が含まれる。例えば、5×SSPE（750 mM NaCl、50 mMリン酸ナトリウム、5 mM EDTA、pH 7.4）及び30℃の温度の条件が、アレル特異的プローブハイブリダイゼーションに適している。

【0083】

キャプチャタグ及び結合剤

一実施態様において、本発明の方法で使用されるオリゴヌクレオチドプローブは、サンプル中、他の核酸配列から、オリゴヌクレオチドプローブと結合した関心のある核酸を濃縮することを容易にするために、キャプチャタグを含む。他の核酸配列から、関心のある核酸について濃縮するために、キャプチャタグは、適切な結合剤と結合する。当技術分野で理解されるように、語句「核酸について濃縮する」とは、サンプル中の標的核酸配列の量を、オリゴヌクレオチドプローブと結合していない核酸と比較して増加させることをいう。それにより、サンプル中、対応する非標的核酸に対する標的配列の比率が増加する。

10

20

30

40

50

【0084】

一実施態様において、キャプチャータグは、「ハイブリダイゼーションタグ」である。本明細書で使用する場合、用語「ハイブリダイゼーションタグ」及び文法上の同義語は、結合剤（すなわち「結合タグ」）として作用する別の核酸配列の少なくとも一部と相補的な配列を含む核酸をいい得る。ハイブリダイゼーションタグと、対応する結合タグ配列との間の相補性の程度は、その適用によって変化し得る。いくつかの実施態様において、ハイブリダイゼーションタグは、結合タグ又はその一部と相補的又は実質的に相補的であり得る。例えば、ハイブリダイゼーションタグは、対応する結合タグと少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約90%、及び少なくとも約99%の相補性を有する配列を含み得る。いくつかの実施態様において、ハイブリダイゼーションタグは、対応する結合タグと100%の相補性を有する配列を含み得る。

10

【0085】

いくつかの実施態様において、キャプチャープローブは、複数のハイブリダイゼーションタグを含むことができ、それらについて、対応する結合タグは、同一の核酸又は異なる核酸中に位置する。ある実施態様において、RNAを含むハイブリダイゼーションタグが、過剰レベルのそのようなタグを効率的に除去するために有利であり得る。

【0086】

ある実施態様において、ハイブリダイゼーションタグは、少なくとも約5ヌクレオチド、少なくとも約10ヌクレオチド、少なくとも約15ヌクレオチド、少なくとも約20ヌクレオチド、少なくとも約25ヌクレオチド、少なくとも約30ヌクレオチド、少なくとも約35ヌクレオチド、少なくとも約40ヌクレオチド、少なくとも約45ヌクレオチド、少なくとも約50ヌクレオチド、少なくとも約55ヌクレオチド、少なくとも約60ヌクレオチド、少なくとも約65ヌクレオチド、少なくとも約70ヌクレオチド、少なくとも約75ヌクレオチド、少なくとも約80ヌクレオチド、少なくとも約85ヌクレオチド、少なくとも約90ヌクレオチド、少なくとも約95ヌクレオチド、及び少なくとも約100ヌクレオチドを含み得る。

20

【0087】

別の実施態様において、キャプチャータグは、「アフィニティタグ」を含んでよい。本明細書で使用する場合、用語「アフィニティタグ」及び文法上の同義語は、多成分複合体（multi-component complex）の成分であって、ここで、該多成分複合体の成分が、互いに特異的に相互作用又は結合するものをいい得る。例えば、アフィニティタグは、ストレプトアビジンと結合し得るビオチンを含み得る。多成分アフィニティタグ複合体の他の例として以下が挙げられる：リガンド及びその受容体、例えば、アビジン-ビオチン、ストレプトアビジン-ビオチン、及びビオチン、ストレプトアビジン又はアビジンの誘導體（2-イミノビオチン、デスチオビオチン、NeutrAvidin (Molecular Probes, Eugene, Oreg.)、CaptAvidin (Molecular Probes) などを含むが、これらに限定されない）；結合タンパク質／ペプチド（マルトース-マルトース結合タンパク質（maltose-maltose binding protein, MBP）、カルシウム-カルシウム結合タンパク質／ペプチド（calcium-calcium binding protein/peptide, CBP）を含む）；抗原-抗体（c-MYC、HA及びFLAG TagTMを含むエピトープタグ、及びその対応する抗エピトープ抗体を含む）；ハプテン、例えば、ジニトロフェニル及びジゴキシゲニン、並びにその対応する抗体；アプタマー及びその対応する標的；フルオロフォア及び抗フルオロフォア抗体など。一実施態様において、アフィニティタグは、ハイブリダイゼーションタグを含む標的核酸のバルク分離（bulk separation）のために使用される。

30

40

【0088】

従って、本発明の方法で使用される結合剤は、本明細書に記載するようなアフィニティタグと結合することができ、サンプル中、他の核酸配列から、関心のある核酸を分離することを容易にする。例えば、一実施態様において、アフィニティタグはビオチンを含み、結合剤はストレプトアビジンを含む。結合剤は典型的には基材（substrate）上にある。基材の例として、ビーズ、マイクロスフェア、平面、カラム、ウェルなどが挙げられる。用語「マイクロスフェア」若しくは「ビーズ」若しくは「粒子」又は文法上の同義語は当

50

技術分野で理解されており、小さな離散型 (discrete) 粒子をいう。基材の組成は、適用によって変化するだろう。適切な組成には、ペプチド、核酸及び有機部分合成で利用されるものが含まれ、これには、プラスチック、セラミック、ガラス、ポリスチレン、メチルスチレン、アクリルポリマー、常磁性材料、トリアゾル (thoria sol)、カーボングラファイト、二酸化チタン、ラテックス又は架橋デキストラン、例えばSepharose、セルロース、ナイロン、架橋ミセル (cross-linked micelles) 及びTeflonが含まれ (それらに限定されない)、それらはすべて使用されてよく、Bangs Laboratories, Fishers Ind.の「Microsphere Detection Guide」は役立つ手引き書である。ビーズは、必ずしも球形である必要はない; 不規則な粒子を使用してよい。いくつかの実施態様において、基材は、金属組成 (例えば、鉄) を含み得、また磁気特性を含んでもよい。磁気ビーズを利用する例となる実施態様には、ストレプトアビジンがコーティングされた磁気ビーズを含むキャプチャープローブが含まれる。さらに、ビーズは多孔性であってよく、それにより、キャプチャープローブと関連するために利用可能なビーズの表面積が増加する。ビーズのサイズは、ナノメートル、すなわち100 nmからミリメートル、すなわち1 mmの範囲に及び、約0.2 μm ~ 約200 μm 、又は約0.5 ~ 約5 μm のビーズであるものの、とはいえ、いくつかの実施態様において、より小さなビーズが使用されてもよい。

10

【0089】

ハイスループットシーケンシング

本発明の方法は、次世代シーケンシング技術などのハイスループットシーケンシング技術を用いたシーケンシングのために核酸を調製するのに特に適している。次世代シーケンシング (NGS) 技術には、装置稼働ごとに 10^{14} キロベースペア (kbp) 超のDNAをシーケンシングすることが可能な装置が含まれる。シーケンシングにより、典型的には、多数の独立したリード (reads) が作成され、それぞれが核酸の10~1000塩基の間のどこかを示している。核酸は通常、カバレッジと称される単位領域あたりの複製物 (すなわち、「10×カバレッジ」又は「100×カバレッジ」) により、確実性のために重複して (redundantly) シーケンシングされる。次世代シーケンシング方法は当技術分野で既知であり、例えば、Metzker (2010) などに記載される。

20

【0090】

従って、本明細書で使用する場合、用語「次世代シーケンシング」又は「NGS」又は「NGシーケンシング」とは、個々の核酸分子 (例えば、単一分子シーケンシングにて)、又は個々の核酸分子についてクローン増殖された代用物 (clonally expanded proxies) のいずれかのヌクレオチド配列を、ハイスループット様式 (例えば、 10^3 、 10^4 、 10^5 又はそれ以上を超える分子が同時にシーケンシングされる) で決定する任意のシーケンシング方法をいう。

30

【0091】

次世代シーケンシングのためのプラットフォームには、Roche/454のGenome Sequencer (GS) FLX System、Illumina/SolexaのGenome Analyzer (GA)、Life/APGのSupport Oligonucleotide Ligation Detection (SOLiD) システム、PolonatorのG.007システム、Helicos BioSciencesのHeliScope Gene Sequencingシステム、及びPacific BiosciencesのPacBio RSシステムが含まれるが、これらに限定されない。

40

【0092】

高度多型性遺伝子及びHLAタイピング

本発明の方法は、高度多型性遺伝子におけるアリの同定に特に適している。本明細書で使用する場合、用語「高度多型性遺伝子 (highly polymorphic gene)」には、遺伝子のコード領域において、その非コード領域と比較して、より高レベルの多型性を有する遺伝子への言及が含まれる。例えば、高度多型性遺伝子は、遺伝子の非コード配列1 kbあたりの多型数と比較した場合、コード配列1 kbあたり、より多くの多型数を有し得る。高度多型性遺伝子のよく知られた例は、ヒト白血球抗原 (HLA) 遺伝子である。

【0093】

従って、本発明の方法は、HLA遺伝子アリの同定及びHLAタイピングに適している。HL

50

A分子のコード領域は、病原体の脅威に反応して進化するため正の選択圧 (positive select pressure) 下にあると考えられることから、高度多型性である。HLAの非コード領域は、そのような選択圧下になく、同程度の多型性を共有しない。HLAクラスIの非コード領域は多型性であるものの、該多型は、これらの領域にわたってランダムに分布せず、密接に関連しており、コード配列類似性により、同一の非コード配列を有する。

【0094】

クラスI HLAの非コード領域からの配列アラインメントの解析により、ユニークな配列セットは限定的であることが示される。本発明者らは、Next Genシーケンシングアッセイにてその後使用され得るDNAフラグメントをキャプチャーするために、ユニークな配列セットそれぞれに対するプローブを設計可能であることを決定した。

10

【0095】

従って、エクソームからの重複プローブ（すなわち、エクソン領域で設計されるプローブ）を利用するアプローチの代わりに、本発明者らは、HLAの非コード領域に対して明確に設計されるプローブの使用を記載する。図1は、異なるHLAアリルが、同一及び予測可能なHLA配列を有することを実証する、HLA-Aイントロン1 (HLA-A intron 1) の配列アラインメントを示す。HLAアリル命名法は、コロンにより分けられる最大4つの区域を利用する（例えば、A*01:01:01:01）。第1区域は、抗体により検出され得る一般的なアミノ酸配列の差異を示す。第2区域は、ヌクレオチドの差異であって、その結果、アミノ酸変化がもたらされるが、抗体の使用では検出できない場合があるものを示す。第3区域は、同義変化を示し、第4区域は、非コード領域における変化を示す。接尾語のレターコードの使用、例えば、A*01:01:01:02NにおけるNなどの使用は、アリルが発現バリエーションであることを示す。例えば、N=NULL又は非発現 (not expressed) である。図に示すように、すべてのA*01アリルは、A*30アリルとヌクレオチドを共有する2 (A*01:02及びA*01:20) を除いて、同一の配列を有し、A*03及びA*11アリルはまた、A*01アリルの大多数と同一である。プローブが、単一塩基対ミスマッチを有するDNAを同様にキャプチャーするであろうことを考慮すると、単一のプローブは、A*01、A03、A*11及びA*30アリルすべてをキャプチャーするはずである。

20

【0096】

一実施態様において、本発明の方法は、1つ以上のHLA-A、HLA-B、HLA-C、HLA-DPA1、HLA-DPB1、HLA-DQA1、HLA-DQB1、HLA-DRA及び／又はHLA-DRB1遺伝子におけるアリルを同定

30

【0097】

いくつかの実施態様において、本発明の方法は、移植ドナー及びレシピエントの両方における遺伝子アリル（例えば、HLA遺伝子アリル）をマッチさせることにより、移植レシピエントに対して適切なドナーを決定するのに、及び／又は患者における移植片対宿主病 (GVHD) 若しくは急性GVHD (aGVHD) の発症の可能性を決定するために、有用である。

【0098】

さらに、移植候補においてaGVHDの発症の可能性を決定することは、移植が、実際にその病態のための最も適切な治療形態であるかどうかを決定するのに影響し得る。いくつかの例では、患者に良好な予後を提供する、利用可能な代替的療法形態が存在する場合がある。さらに、aGVHDの可能性の予測は、治療レジーム (regimes) の必要性を示し、或いは場合により、本来推奨されるものよりも積極的な治療レジームの必要性を示す。そのような積極的な療法は、多くの場合、望ましくない副作用を有し、そのため好ましくは、予後が必要性を示さない限り使用されない。例えば、中和抗TNF-アルファモノクローナル抗体での患者の治療は、aGVHDの改善をもたらし得るが、感染症のリスクも増加し得る。様々な積極的抗炎症療法も利用可能である。

40

【0099】

本明細書で使用する場合、用語「移植する (transplant)」又は「移植 (transplanting)」とは、ある個体 (ドナー) から得られた組織又は細胞を別の個体 (レシピエント) の身体中又は身体上へと移植 (grafting) 又は導入することをいう。ドナーから取り出さ

50

れ、レシピエントへと移植される細胞又は組織は、「移植片 (graft)」と称される。一般的に移植される組織の例は、骨髄、造血幹細胞 (hematopoietic stem cells)、器官、例えば、肝臓、心臓、皮膚、膀胱、肺、腎臓、角膜、脾臓、脾島、脳組織、骨、及び腸である。

【0100】

当業者は、用語「ハプロタイプ」が、染色体上近接して又は隣接する座位に位置し、一緒に遺伝するアレルの組合せ、又は統計学的に関連する、染色体ペアの単一の染色体上の一連の一塩基多型の組み合わせをいうことを理解するだろう。

【0101】

キット

本発明はさらに、本発明の方法に従って遺伝子アレルを同定するためのキットを提供する。当該キットは、関心のある遺伝子の非コード領域における遺伝子標的配列とハイブリダイズできるオリゴヌクレオチドプローブを含む。一実施態様において、オリゴヌクレオチドプローブは、キャプチャータグ、例えば、ビオチン又はストレプトアビジンなどを含む。当該キットはさらに、キャプチャータグと結合可能な結合剤を含んでよい。結合剤は、適切な基材、例えば、マイクロスフェア又はビーズなどにコーティング又は付着されてよい。いくつかの実施態様において、基材は、関心のある標的核酸の濃縮を容易にするために磁気であってよい。

【実施例】

【0102】

実施例

実施例1. DNAライブラリーの調製

Illumina TruSeq DNA Nanoプロトコルを用いて、200 ngのゲノムDNAからIlluminaライブラリーを調製し、550 bpのターゲットインサートサイズを選択した。

【0103】

実施例2. インترون特異的プローブを用いたHLAのキャプチャー

概念実証として、イントロン配列を標的とするカスタムオリゴヌクレオチドプローブプール (配列番号1~8) を1.5 pmol/ μ lの濃度に希釈した。最初のハイブリダイゼーション手順のために、ヒトCot DNA、増幅ライブラリー (上記のように調製)、ユニバーサルブロッキングオリゴTS-p7 (6nt)、及びユニバーサルブロッキングオリゴTS-p7 (8nt) をチューブに添加し、60°Cに設定した減圧濃縮器 (vacuum concentrator) にておよそ1時間ドライダウンした。次いで、ドライダウンしたサンプルを、Nimblegen 2X Hybridisationバッファー、Nimblegen Hybridisation Component A及びヌクレアーゼフリー水に再懸濁した。次いで、再懸濁したサンプルをボルテックスし、95°Cで10分間変性させた。変性後、2 μ lのプローブ (3 pmol) をサンプルに添加し、72°Cで4時間インキュベートした。

【0104】

プローブがハイブリダイズしたDNAを洗浄及び回収するために、Nimblegenウォッシュキットの洗浄溶液を製造者のプロトコルに従って希釈した。ハイブリダイズしたサンプルをストレプトアビジンビーズと結合させ、ボルテックスし、72°Cで45分間インキュベートした。インキュベーション後、結合したDNAをウォッシュ (wash) I及びストリンジェントウォッシュで72°Cにて洗浄し、それに続いてウォッシュII及びIIIで室温にて洗浄した。次いで、ストレプトアビジンビーズと結合した、ハイブリダイズ化サンプルをマグネット上に置き、0.125N NaOHを用いて、結合したDNAを溶出 (eluted off) させた。次いで、Ampure XPビーズを用いて、ハイブリダイズ化サンプルを精製し、次いで、KAPA HiFi HotStart ReadyMixを用いて12サイクル増幅させた。PCRに使用したプライマーは、ライゲートされたDNAにおけるIlluminaアダプター及びバーコード配列と相補的であった (Illumina P5及びP7)。Ampure XPビーズでサンプルを精製した。

【0105】

実施例3. ハイブリダイズしたDNAのシーケンシング

増幅させたハイブリダイズ化サンプルを、Illumina 2x300 bpシーケンシングプロトコ

10

20

30

40

50

ルを用いて、MiSeq又はHiSeqにてシーケンシングした。

【0106】

実施例4. 結果

作成された配列データをAssign MPS v 1.0配列解析ソフトウェア (Conexio Genomics, Fremantle, Australia) で解析した。標的化キャプチャーシーケンシングプロトコルを用いて作成された配列データの解析を図4に示す。Assign MPS v 1.0ソフトウェアを用いて、サンプルについて最も適合するジェノタイプは、HLA-A*02:01:01:01+A*02:01:01:01であることが決定された。

【0107】

実施例5. サンガーによってもシーケンシングされた3サンプルのキャプチャープローブ濃縮後の次世代シーケンシング

10

サンガーアッセイは、HLA-A及び-Bについて5'非翻訳領域からイントロン4の末端まで、並びにHLA-Cの5'非翻訳領域から3'非翻訳領域まで及ぶアンプリコンにて、HLA-A、-B及び-C配列を増幅する。ヘテロ接合体のアンビギュイティ (ambiguities) はすべて、アリル特異的シーケンシングプライマーを用いて解決されていた。

【0108】

図3に記載され、配列番号10~71として提供されるプローブを用いて、上記のプロトコルにより、キャプチャープローブアッセイを実施した。3サンプルについて、報告されるジェノタイプに関する完全な一致が存在した。アッセイの結果を図5に示す。

【0109】

20

サンプル1 HLA-Bは、リファレンス配列が報告されていないアリルを含有する。このアリルは、B*15:01:01:01の配列と比較して、イントロン1に位置する、塩基444における単一ミスマッチを有する。サンガーを用いた座位特異的増幅により、このサンプルは、新規アリルが、未記載の、B*15:01:01:01アリルのB*08:01:01アリルであるかを決定するために、B*08:01:01及びB*15:01:01アリル特異的シーケンシングプライマーを用いた更なる実験を必要とする。NGSは、遺伝子全体にわたる多型のフェージングを可能とし、新規アリルがB*15:01:01サブタイプであることを容易に同定できる。このサンプルはまた、新規C*04:01:01アリルも含有する。当該多型はイントロン5に位置する。アリル特異的シーケンシングプライマーの使用は、この領域では一般に使用されず、サンガーシーケンシングによる特性化は、アリルをPCRにより分離し、それに続いてサンガーシーケンシングすること

30

を必要とする。しかしながら、NGSは、新規アリルがC*04:01:01サブタイプであることを正確に同定した。

【0110】

このデータは、キャプチャープローブの使用により新規アリルの同定が可能であること、また配列多型のフェーズ関連性を決定するNGSの能力を実証し、従って、新規アリルの正確な特性化を可能とすることを実証する。

【0111】

当業者ならば、特定の実施態様で示されるような本発明に対して、広範に記載される本発明の範囲から逸脱することなく、多数の変更及び／又は改変がなされてよいことを理解するであろう。従って、本実施態様は、あらゆる点で例示的であり制限的なものではない

40

と考えらえるべきである。

【0112】

本明細書において論述及び／又は参照されたすべての刊行物は、その全体が本明細書に組み込まれる。

【0113】

本出願は、AU 2013904801の優先権を主張するものであり、その全体の内容が参照により本明細書に組み込まれる。

【0114】

本明細書に含まれている、文献、行為、材料、装置、商品などについてのあらゆる議論は、本発明に関する事情を提供することのみを目的とする。それが本出願の各請求項の優

50

【 0 1 1 5 】

Metzker (2010) Nat Rev Genet, 11 (1) : 31-46

[illegible][illegible]

【图 1-4】

[illegible]

【图 1-6】

[illegible]

【 図 1 - 7 】

[illegible]

【図 1 - 9】

[illegible]

【图 1-8】

[illegible]

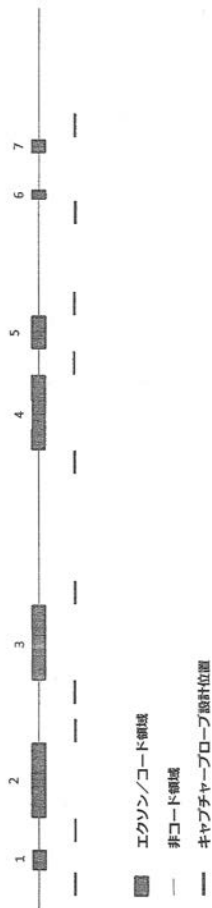
【 図 1 - 1 0 】

A*29:01.01:01	-T-----C-	-----	-----TG-----	-----C-----	-----
A*29:01.01:02N	-A-----C-	-----	-----TG-----	-----C-----	-----
A*29:02.01:01	-T*29:02.01:01	-----	-----TG-----	-----C-----	-----
A*29:02.01:02	-T-----C-	-----	-----TG-----	-----C-----	-----
A*29:46	-A-----C-	-----	-----TG-----	-----C-----	-----
A*31:01.02	-A*31:01.02	-----	-----G-A-----	-----C-----	-----
A*31:01.23	-T-----C-	-----	-----G-A-----	-----C-----	-----
A*31:04	-A-----C-	-----	-----G-A-----	-----C-----	-----
A*31:14N	-C-----C-	-----	-----G-A-----	-----C-----	-----
A*31:46	-A-----C-	-----	-----G-A-----	-----C-----	-----
A*33:01.01	-T-----C-	-----	-----G-A-----	-----C-----	-----
A*33:03.01	-A*33:03.01	-----	-----G-A-----	-----C-----	-----
A*33:07	-A-----C-	-----	-----G-A-----	-----C-----	-----
A*32:01.01	-T-----C-	-----	-----G-----	-----C-----	-----
A*32:06	-C-----C-	-----	-----G-----	-----C-----	-----
A*74:01	-A-----C-	-----	-----G-----	-----C-----	-----
A*74:02.01:01	-T*74:02.01:01	-----	-----G-----	-----C-----	-----
A*74:02.01:02	-T-----C-	-----	-----G-----	-----C-----	-----
A*34:01.01	-A-----C-	T-----	-----G-----	-----C-----	-----
A*34:02.01	-C-----C-	T-----	-----G-----	-----C-----	-----
A*43:01	-C-----C-	T-----	-----G-----	-----C-----	-----
A*66:01.01	-A-----C-	T-----	-----G-----	-----C-----	-----
A*66:17	-A*66:17	T-----	-----G-----	-----C-----	-----
A*68:01.01:01	-A-----C-	T-----	-----G-----	-----C-----	-----
A*68:01.01:02	-A*68:01.01:02	T-----	-----G-----	-----C-----	-----
A*68:01.02:01	-A-----C-	T-----	-----G-----	-----C-----	-----

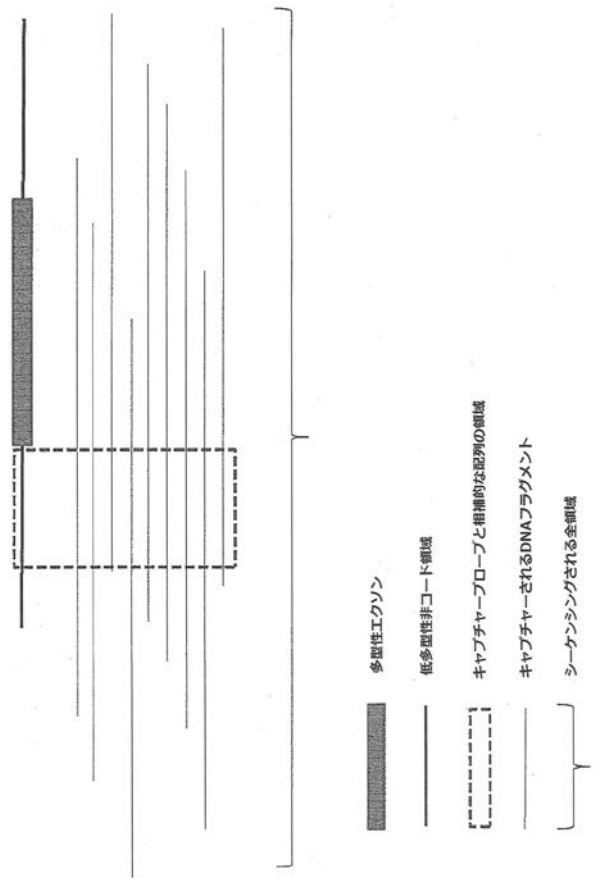
【図 1 - 1 1】

A*68:01:02:02	-C--	-A--	-T--	-G--	-C--	-G--
A*68:02:01:01	-C--	-A--	-T--	-G--	-C--	-G--
A*68:02:01:02	-C--	-A--	-T--	-G--	-C--	-G--
A*68:02:01:03	-C--	-A--	-T--	-G--	-C--	-G--
A*68:02:02	-C--	-A--	-T--	-G--	-C--	-G--
A*68:03:01	-C--	-A--	-T--	-G--	-C--	-G--
A*68:08:01	-C--	-A--	-T--	-G--	-C--	-G--
A*68:17	-C--	-A--	-T--	-G--	-C--	-G--
A*68:18N	-C--	-A--	-T--	-G--	-C--	-G--
A*68:22	-C--	-A--	-T--	-G--	-C--	-G--
A*68:71	-C--	-A--	-T--	-G--	-C--	-G--
A*68:113	-C--	-A--	-T--	-G--	-C--	-G--
A*69:01	-C--	-A--	-T--	-G--	-C--	-G--
A*80:01:01:01	-C--	-A--	-T--	-G--	-C--	-G--
A*80:01:01:02	-C--	-A--	-T--	-G--	-C--	-G--

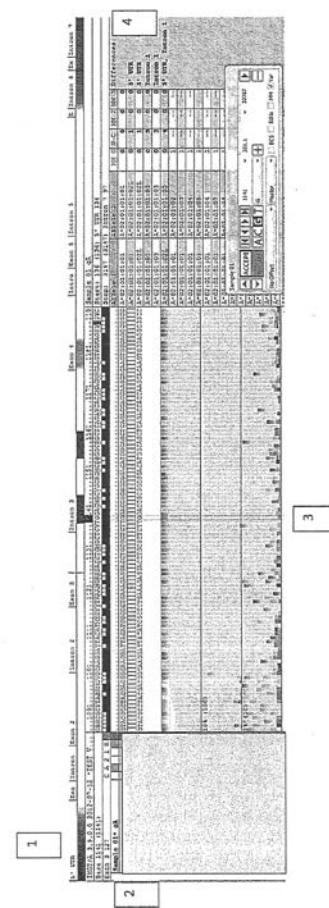
【図 3】



【図 2】



【図 4】



【図 5】

サンプル	座位	報告されるジェノタイプ	
		サンガー	キャプチャーロープ
サンプル 1	HLA A	01:01:01:01	01:01:01:01
	HLA A	24:02:01:01	24:02:01:01
	HLA B	08:01:01	08:01:01
	HLA B	15:01:01*	15:01:01*
	HLA C	04:01:01*	04:01:01*
	HLA C	07:01:01:01	07:01:01:01
サンプル 2	HLA A	02:01:01:01	02:01:01:01
	HLA A	31:01:02	31:01:02
	HLA B	27:05:02	27:05:02
	HLA B	40:01:02	40:01:02
	HLA C	02:02:02:01	02:02:02:01
	HLA C	03:04:01:01	03:04:01:01
サンプル 3	HLA A	03:01:01:01	03:01:01:01
	HLA A	24:03:01	24:03:01
	HLA B	39:06:02	39:06:02
	HLA B	50:01:01	50:01:01
	HLA C	12:03:01:01	12:03:01:01
	HLA C	15:04	15:04

【配列表】

2017506500000001.app

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/AU2014/001114
A. CLASSIFICATION OF SUBJECT MATTER C12Q 1/68 (2006.01)		
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)		
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)		
MEDLINE, HCA, BIOSIS, EMBASE, EPODOC, WPI (oligonucleotide, probe, target, non coding, intron, intergenic, untranslated, UTR, leader, trailer, enhancer, response element, insulator, promoter, transcriptional regulatory element, assaying, identifying, determining, type, typing, genotyping, HLA, human leukocyte antigen, sequencing, allele, allelomorph, variant, haplotype, fragmenting, next generation, high throughput, genome analyzer, support oligonucleotide ligation detection, heliscope gene sequencing, PacBio RS system, polonator, G.007, genome sequencer, FLX, polymorphic, SNP, highly, very, greater, many, enriched; and other synonyms and like terms) PUBMED, ESPACENET (Applicant and Inventor Search) GENOMEQUEST (SEQ ID NO's 1-8 or 10-71)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	
☒ Further documents are listed in the continuation of Box C ☒ See patent family annex		
* "A"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance	"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
"E"	earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O"	document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"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 13 February 2015		Date of mailing of the international search report 13 February 2015
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaaustralia.gov.au		Authorised officer Damian Triffett AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 62832845

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/AU2014/001114
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 1997/023645 A1 (SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH) 03 July 1997 See whole document especially Page 1, lines 15-19; Page 9, lines 25-29; Page 22, lines 28-30; Claims 1, 3, 11 and 32	1-15, 17-37
X	US 2004/0197775 A1 (Simons) 07 October 2004 See whole document especially Paragraphs [0139], [0141] and [0148]; Claims 1 and 12	1-5, 7-15, 17-37
X	EP 0892069 A2 (BIOTEST AG) 20 January 1999 See whole document especially Paragraphs, [0004], [0089] and [0090]; Claims 1, 6 and 18	1-5, 7-15, 17-37
X	WO 2005/108624 A2 (UNIVERSITY OF CHICAGO, UCTECH) 17 November 2005 See whole document especially Claims 1, 2, 5, 7, 13 and 20	1-4, 7-14, 17-30, 32-37
X	WO 1999/007883 A1 (VISIBLE GENETICS INC. et al.) 18 February 1999 See whole document especially Page 14, second to last paragraph, Pages 41 and 42, 8.2 and 8.3; Claims 1 and 14; SEQ ID NO's 40-46, 86, 247, 279 and 285	1-5, 7-15, 17-37
X	LIND, C. et al., 'Complete genomic sequencing of HLA-C*04:09n using next-generation sequencing technology', Human Immunology, September 2010, Vol. 71, Suppl. 1, page S129 See whole document	1-15, 17-37
X	WO 2012/171990 A1 (PROGENIKA BIOPHARMA, S. A.) 20 December 2012 See whole document especially Page 18, lines 11-15; Page 30, lines 1-21; Claims 1, 33 and 44	1, 2, 4, 7-14, 16-21, 28, 29, 32-36
A	DE SANTIS, D. et al., '16(th) IHIW : review of HLA typing by NGS', International Journal of Immunogenetics, February 2013, Vol. 40, No. 1, pages 72-76 See whole document	1-37
X	WO 2014/116729 A2 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 31 July 2014 See whole document especially Paragraphs [0086], [00113] and [00114]; Claims 1, 3, 7 and 50	3-37
P,X	WO 2014/116729 A2 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 31 July 2014 See whole document especially Paragraphs [0086], [00113] and [00114]; Claims 1, 3, 7 and 50	1-37
Form PCT/ISA/210 (fifth sheet) (July 2009)		

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/AU2014/001114	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 1997/023645 A1	03 July 1997		
US 2004/0197775 A1	07 October 2004	AU 654111 B2	27 Oct 1994
		AU 672519 B2	03 Oct 1996
		AU 6131990 A	28 Feb 1991
		AU 7285094 A	24 Nov 1994
		CA 2023888 A1	26 Feb 1991
		CA 2066211 A1	12 Mar 1991
		EP 0414469 A2	27 Feb 1991
		EP 0414469 B1	30 Oct 1996
		EP 0491831 A1	01 Jul 1992
		HK 1008053 A1	30 Apr 1999
		IL 95467 A	31 Jul 1995
		JP H03139300 A	13 Jun 1991
		JP 3206812 B2	10 Sep 2001
		JP H05500368 A	28 Jan 1993
		JP 2001309796 A	06 Nov 2001
		NZ 235051 A	29 Jan 1992
		SG 47747 A1	17 Apr 1998
		US 4997977 A	05 Mar 1991
		US 5192659 A	09 Mar 1993
		US 5612179 A	18 Mar 1997
		US 5789568 A	04 Aug 1998
		US 2003119003 A1	26 Jun 2003
		WO 9103457 A1	21 Mar 1991
		ZA 9006765 A	26 Jun 1991
EP 0892069 A2	20 January 1999	EP 0892069 B1	28 Jul 2010
		EP 0887423 A1	30 Dec 1998
		EP 2186911 A1	19 May 2010
		EP 2186911 B1	21 Jan 2015
WO 2005/108624 A2	17 November 2005	US 2005287556 A1	29 Dec 2005
WO 1999/007883 A1	18 February 1999	AU 8846898 A	01 Mar 1999
		CA 2300369 A1	18 Feb 1999
		EP 1003917 A1	31 May 2000
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

Form PCT/ISA/210 (Family Annex)(July 2009)

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/AU2014/001114	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		JP 2001512694 A	28 Aug 2001
WO 2012/171990 A1	20 December 2012	EP 2721171 A1	23 Apr 2014
		US 2014220567 A1	07 Aug 2014
WO 2014/116729 A2	31 July 2014	US 2014206547 A1	24 Jul 2014
WO 2014/116729 A2	31 July 2014	US 2014206547 A1	24 Jul 2014
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			

フロントページの続き

(51)Int.Cl.		F I		テーマコード (参考)
A 6 1 P 43/00	(2006.01)	A 6 1 K 35/12		
		A 6 1 P 43/00	1 2 1	

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, 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, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, 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, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(特許庁注：以下のものは登録商標)

1. T E F L O N

(74)代理人 100163658

弁理士 小池 順造

(74)代理人 100174296

弁理士 當麻 博文

(74)代理人 100137729

弁理士 赤井 厚子

(74)代理人 100151301

弁理士 戸崎 富哉

(72)発明者 セイヤー、デヴィッド チャールズ

オーストラリア国、ウェスタン オーストラリア州 6 1 5 8、イースト フリーマントル、ウィンザー ロード 7

(72)発明者 デサンティス、ディエアン

オーストラリア国、ウェスタン オーストラリア州 6 0 1 6、マウント ホウソーン、アンザック ロード 2 0

Fターム(参考) 4B063 QA01 QA13 QQ42 QR32 QR56 QR62 QS10 QS14 QS15 QS25

QS34 QS39

4C087 AA01 BB63 NA06 ZC75