

(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2015-501658
(P2015-501658A)

(43) 公表日 平成27年1月19日(2015.1.19)

(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 A	4 B 0 2 4
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 A	4 B 0 6 3
G O 1 N 33/543 (2006.01)	G O 1 N 33/543 5 9 7	
G O 1 N 33/53 (2006.01)	G O 1 N 33/53 Y	

審査請求 未請求 予備審査請求 未請求 (全 38 頁)

(21) 出願番号 特願2014-547393 (P2014-547393)
 (86) (22) 出願日 平成24年12月12日 (2012.12.12)
 (85) 翻訳文提出日 平成26年7月29日 (2014.7.29)
 (86) 国際出願番号 PCT/US2012/069310
 (87) 国際公開番号 W02013/090469
 (87) 国際公開日 平成25年6月20日 (2013.6.20)
 (31) 優先権主張番号 61/570,192
 (32) 優先日 平成23年12月13日 (2011.12.13)
 (33) 優先権主張国 米国 (US)

(71) 出願人 511110533
 シーケンタ 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 サウス
 サンフランシスコ イースト ジェイ
 ミー コート 400 スイート 301
 (74) 代理人 100102978
 弁理士 清水 初志
 (74) 代理人 100102118
 弁理士 春名 雅夫
 (74) 代理人 100160923
 弁理士 山口 裕孝
 (74) 代理人 100119507
 弁理士 刑部 俊
 (74) 代理人 100142929
 弁理士 井上 隆一

最終頁に続く

(54) 【発明の名称】 組織浸潤リンパ球の検出および測定方法

(57) 【要約】

本発明は、腫瘍または自己免疫疾患に罹患している組織などの固形組織中に浸潤しているリンパ球などの細胞の数、レベル、および/または比率を測定するための方法、ならびにこのような測定に基づいて患者の予後予測を行うための方法に関する。1つの局面において、本発明の方法は、末梢血などの入手しやすい組織に由来するリンパ球を、細胞傷害性T細胞および調節性T細胞などの機能的サブセットに選別する工程、ならびに各サブセットのクロノタイププロファイルを生成する工程を含む。入手しにくい疾患罹患組織をサンプリングし、かつ1つまたは複数のクロノタイププロファイルを生成する。後者のクロノタイププロファイルからは、疾患罹患組織での機能的サブセットの各々におけるリンパ球のレベルが、それらのクロノタイプによって決定されるが、それらのクロノタイプは、入手しやすい組織からサブセットに選別されたリンパ球より同定されたものである。

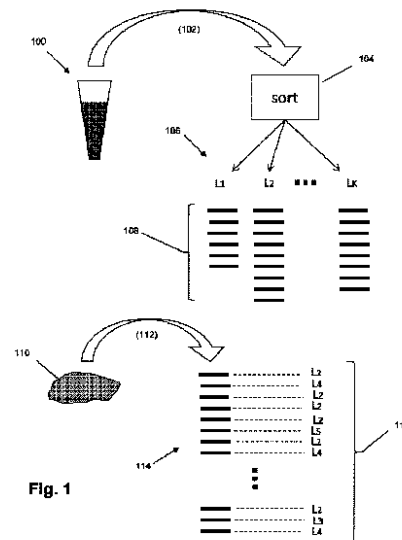


Fig. 1

【特許請求の範囲】

【請求項 1】

以下の工程を含む、固形組織に浸潤したリンパ球を同定するための方法：

個体の入手しやすい組織に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；

該入手しやすい組織に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；

固形組織の少なくとも1つの試料から、少なくとも1つのクロノタイププロファイルを生成する工程；および

該固形組織における各サブセットのリンパ球を、そのそれぞれのクロノタイプから検出する工程。

10

【請求項 2】

リンパ球の前記1つまたは複数のサブセットのうちの少なくとも1つが、B細胞、T細胞、細胞傷害性T細胞、ヘルパーT細胞、調節性T細胞、Th1 T細胞、Th2 T細胞、Th9 T細胞、Th17 T細胞、Tfh T細胞、抗原特異的B細胞の少なくとも1つのサブセット、および/または抗原特異的T細胞の少なくとも1つのサブセットを含む、請求項1に記載の方法。

【請求項 3】

検出する工程が、前記1つまたは複数のサブセットの各々のリンパ球の数、レベル、および/または比率を決定することをさらに含む、請求項2に記載の方法。

【請求項 4】

固形組織が固形腫瘍である、請求項3に記載の方法。

20

【請求項 5】

リンパ球の前記1つまたは複数のサブセットが、調節性T細胞および細胞傷害性T細胞を含む、請求項4に記載の方法。

【請求項 6】

検出する工程が、固形腫瘍中の調節性T細胞のレベル対細胞傷害性T細胞のレベルの比率を決定することを含む、請求項5に記載の方法。

【請求項 7】

リンパ球の前記1つまたは複数のサブセットが、抗原特異的T細胞の少なくとも1つのサブセットを含む、請求項3に記載の方法。

30

【請求項 8】

入手しやすい組織が個体の末梢血、骨髄、リンパ液、滑液、または脊髄液である、請求項1に記載の方法。

【請求項 9】

入手しやすい組織が個体の末梢血である、請求項8に記載の方法。

【請求項 10】

検出する工程が、前記サブセットの各々におけるリンパ球の数を、そのそれぞれのクロノタイプをカウントすることによって数えることをさらに含む、請求項1に記載の方法。

【請求項 11】

固形組織の試料が、該固形組織の異なる位置から取得された複数の組織試料を含む、請求項1に記載の方法。

40

【請求項 12】

リンパ球の前記1つまたは複数のサブセットが、ヘルパーT細胞および細胞傷害性T細胞を含む、請求項1に記載の方法。

【請求項 13】

ヘルパーT細胞がFoxP3+調節性T細胞である、請求項12に記載の方法。

【請求項 14】

細胞傷害性T細胞がCD8⁺マーカーによって同定され、かつ調節性T細胞がマーカーCD4⁺、CD25⁺(高)、およびCD127^(低)によって同定される、請求項13に記載の方法。

【請求項 15】

50

固形組織が固形腫瘍である、請求項14に記載の方法。

【請求項16】

リンパ球の前記1つまたは複数のサブセットが、1つまたは複数の細胞表面マーカーまたは細胞内マーカーに基づいて別々の容器に選別される、請求項1に記載の方法。

【請求項17】

リンパ球の前記1つまたは複数のサブセットを選別する工程が、蛍光活性化セルソーター (FACS) を用いて実行される、請求項16に記載の方法。

【請求項18】

クロノタイププロファイルのクロノタイプの各々が、T細胞受容体遺伝子の組換え部分または免疫グロブリン遺伝子の組換え部分を含む、請求項1に記載の方法。

10

【請求項19】

クロノタイププロファイルのクロノタイプの各々が、T細胞受容体鎖をコードする核酸のV(D)J領域の少なくとも一部分またはIgH鎖をコードする核酸のV(D)J領域の少なくとも一部分を含む、請求項18に記載の方法。

【請求項20】

クロノタイププロファイルの各々が、少なくとも30ヌクレオチドのクロノタイプを少なくとも1000種含む、請求項19に記載の方法。

【請求項21】

クロノタイププロファイルの各々が、少なくとも 10^4 種のクロノタイプを含む、請求項20に記載の方法。

20

【請求項22】

固形組織が、自己免疫疾患に罹患している固形組織である、請求項1に記載の方法。

【請求項23】

リンパ球の前記1つまたは複数のサブセットのうちの少なくとも1つが、B細胞、T細胞、細胞傷害性T細胞、ヘルパーT細胞、調節性T細胞、Th1 T細胞、Th2 T細胞、Th9 T細胞、Th17 T細胞、Tfh T細胞、抗原特異的B細胞、および抗原特異的T細胞より選択され、かつ検出する工程が、前記1つまたは複数のサブセットの各々のリンパ球の数、レベル、および/または比率を決定することをさらに含む、請求項22に記載の方法。

【請求項24】

リンパ球の前記1つまたは複数のサブセットが、調節性T細胞、Th1 T細胞、Th2 T細胞、Th9 T細胞、Th17 T細胞、Tfh T細胞、および/または抗原特異的T細胞の少なくとも1つのサブセットを含む、請求項23に記載の方法。

30

【請求項25】

固形組織が、結腸組織、小腸組織、皮膚、結合組織、皮下組織、肺組織、腎臓組織からなる群より選択される、請求項23に記載の方法。

【請求項26】

固形組織が正常組織であり、かつリンパ球の前記1つまたは複数のサブセットのうちの少なくとも1つが、B細胞、T細胞、細胞傷害性T細胞、ヘルパーT細胞、調節性T細胞、Th1 T細胞、Th2 T細胞、Th9 T細胞、Th17 T細胞、Tfh T細胞、抗原特異的B細胞、および抗原特異的T細胞より選択され、かつ検出する工程が、前記1つまたは複数のサブセットの各々のリンパ球の数、レベル、および/または比率を決定することをさらに含む、請求項1に記載の方法。

40

【請求項27】

以下の工程を含む、患者の固形腫瘍中へのリンパ球浸潤の状態から予後を判定するための方法：

患者の末梢血に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；

該末梢血に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；

固形腫瘍の少なくとも1つの試料から、少なくとも1つのクロノタイププロファイルを生

50

成する工程；ならびに

該1つまたは複数のサブセットの各々のリンパ球の数、レベル、および/または比率を決定する工程。

【請求項28】

前記1つまたは複数のサブセットが、調節性T細胞および細胞傷害性T細胞を含む、請求項27に記載の方法。

【請求項29】

決定する工程が、細胞傷害性T細胞のレベル対調節性T細胞のレベルの比率が前記患者の非腫瘍組織における該比率よりも高い場合に、平均を上回る患者生存率を示すことを含む、請求項28に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

相互参照

本出願は2011年12月13日に出願された同時係属中の米国仮特許出願第61/570,192号の優先権を主張し、この出願は全体として参照により本明細書に組み入れられる。

【背景技術】

【0002】

発明の背景

固形腫瘍などの疾患罹患組織中に浸潤している異なるリンパ球サブセットの数および比率は、その疾患の予後と関係している場合が多い；例えば、Deschoolmeester et al, BMC Immunology, 11:19 (2010) (非特許文献1)；Ohtani, Cancer Immunity, 7: 4 (2007) (非特許文献2)；Yu et al, Laboratory Investigation, 86: 231-245 (2006) (非特許文献3)；Diederichsen et al, Cancer Immunol. Immunother., 52: 423-428 (2003) (非特許文献4)など。残念ながら、免疫組織化学的検査またはフローサイトメトリーなどの利用可能な技術を用いたそのような量の測定は、難しく、大きな労働力を要し、かつ定常的な配備に適していない。

【0003】

それとは別に、DNA配列決定の塩基当たりのコストが下がったために、診断および予後予測適用において大規模DNA配列決定を使用することへの関心がますます高まっている。例えば、T細胞受容体もしくはB細胞受容体またはそれらの成分などの免疫分子をコードする核酸のプロファイルは、生物の健康または疾患の状態に関する豊富な情報を含み、そのため、このようなプロファイルを診断または予後予測の指標として使用することが、幅広い状態に対して提唱されている；例えば、Faham and Willis、米国特許出願公開第2010/0151471号 (特許文献1)；Freeman et al, Genome Research, 19: 1817-1824 (2009) (非特許文献5)；Boyd et al, Sci. Transl. Med., 1(12): 12ra23 (2009) (非特許文献6)；He et al, Oncotarget (March 8, 2011) (非特許文献7)。

【0004】

ハイスループット核酸配列決定における改善を用いて、組織浸潤リンパ球 (TIL) を測定するためのより簡便でかつより効果的なアッセイを提供することができるのであれば、医学分野および科学分野に非常に有用である。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】米国特許出願公開第2010/0151471号

【非特許文献】

【0006】

【非特許文献1】Deschoolmeester et al, BMC Immunology, 11:19 (2010)

【非特許文献2】Ohtani, Cancer Immunity, 7: 4 (2007)

【非特許文献3】Yu et al, Laboratory Investigation, 86: 231-245 (2006)

10

20

30

40

50

【非特許文献4】Diederichsen et al, Cancer Immunol. Immunother., 52: 423-428 (2003)

【非特許文献5】Freeman et al, Genome Research, 19: 1817-1824 (2009)

【非特許文献6】Boyd et al, Sci. Transl. Med., 1(12): 12ra23 (2009)

【非特許文献7】He et al, Oncotarget (March 8, 2011)

【発明の概要】

【0007】

本発明は、腫瘍などの固形組織中に浸潤しているリンパ球などの細胞の数、レベル、および/または比率を測定するための方法、ならびにこのような測定に基づいて患者の予後予測を行うことに関する。本発明は多くの実行および適用において例証されるが、そのうちのいくつかについて以下におよび本明細書を通して概要を述べる。

10

【0008】

1つの局面において、本発明は、以下の工程を含む、固形組織に浸潤したリンパ球を同定するための方法に関する：(a) 個体の入手しやすい組織に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；(b) 該入手しやすい組織に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；(c) 固形組織の少なくとも1つの試料から、少なくとも1つのクロノタイププロファイルを生成する工程；および (d) 該固形組織における各サブセットのリンパ球を、そのそれぞれのクロノタイプから検出する工程。

【0009】

別の局面において、本発明は、以下の工程を含む、患者の固形腫瘍中へのリンパ球浸潤の状態から予後を判定するための方法に関する：(a) 患者の末梢血に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；(b) 該末梢血に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；(c) 固形腫瘍の少なくとも1つの試料から、少なくとも1つのクロノタイププロファイルを生成する工程；ならびに (d) 該1つまたは複数のサブセットの各々のリンパ球の数、レベル、および/または比率を決定する工程。1つの態様において、固形腫瘍中へのリンパ球浸潤の状態とは、固形腫瘍内の選別された機能的サブセットのリンパ球の数、レベル、および/または比率を意味する。いくつかの態様において、固形腫瘍中へのリンパ球浸潤の状態はまた、固形腫瘍内でのまたは固形腫瘍に隣接しての、そのような値の空間分布を含み得る。

20

30

【0010】

本発明の、上記で特徴づけられたこれらの局面およびその他の局面は、説明がなされる多くの実行および適用において例証されるが、そのうちのいくつかを図面で示し、添付の特許請求の範囲において特徴づける。しかしながら、上記の概要は、本発明のそれぞれ説明がなされる態様またはすべての実行を記載することを意図するものではない。

【図面の簡単な説明】

【0011】

本発明の新規な特徴は、添付の特許請求の範囲に詳細に記載されている。本発明の特徴および利点のより良い理解は、本発明の原理が利用される例示的な態様を記載している以下の詳細な説明、および添付の図面を参照することによって得られる。

40

【0012】

【図1】本発明の1つの態様の工程を図示する。

【図2】図2A~2Cは、TCR 遺伝子を増幅するための二段階PCRスキームを示す。

【図3】図3Aは、図2CのPCR産物のヌクレオチド配列を決定する詳細を図示する。図3Bは、図2CのPCR産物のヌクレオチド配列を決定する別の態様の詳細を図示する。

【図4】図4Aは、単一反応においてIgH鎖から3つの配列決定鑄型を生成するためのPCRスキームを図示する。図4B~4Cは、3つの別々の反応においてIgH鎖から3つの配列決定鑄型を生成し、その後、得られたアンプリコンを混合して、P5プライマー結合部位およびP7プライマー結合部位を付加するための二次PCRを行うためのPCRスキームを図示する。図4Dは

50

、IgH鎖について生成された配列リードの位置を図示する。図4Eは、NDN領域における塩基コールを改善するための、V領域およびJ領域のコドン構造の使用を図示する。

【発明を実施するための形態】

【0013】

発明の詳細な説明

本発明の実施は、特記されない限り、当技術分野の技術の範囲内にある、分子生物学（組換え技法を含む）、バイオインフォマティクス、細胞生物学、および生化学の従来技法および説明を使用することができる。このような従来技法には、血液細胞のサンプリングおよび解析、核酸の配列決定および解析などが含まれるが、これらに限定されない。適切な技法の具体的な例は、本明細書における以下の実施例を参照することによって知ることができる。しかしながら、その他の同等な従来手順もまた、当然ながら用いることができる。このような従来技法および説明は、Genome Analysis: A Laboratory Manual Series (Vols. I-IV); PCR Primer: A Laboratory Manual; およびMolecular Cloning: A Laboratory Manual (いずれもCold Spring Harbor Laboratory Pressによる)などの標準的な実験室マニュアルにおいて見出すことができる。

10

【0014】

1つの局面において、本発明は、腫瘍、自己免疫疾患に罹患している組織、移植片対宿主病 (GVHD) に罹患している組織、正常組織などの固形組織中に浸潤しているリンパ球の型および数を決定する方法に関する。関心対象の固形組織は通常は疾患罹患固形組織であるが、いくつかの態様では、正常組織中のリンパ球の異なるサブセットのレベルおよび/または数および/または比率を用いて、個体の健康状態、および/または個体が疾患もしくは状態を患う傾向を判定することもできる。

20

【0015】

本発明の1つの態様の概要を図1に示す。末梢血などの簡単に入手しやすい組織 (100) から、個体のリンパ球のクロノタイプを決定する。任意で、末梢血単核細胞 (PBMC) を単離するなど、最小限の試料調製工程 (102) を実行してもよい。このような試料から、リンパ球をサブセット、 L_1 、 L_2 、... L_K (106) に選別するが (104)、これらは通常、別個の生物学的機能を有するリンパ球に対応する；このようなサブセットは場合により、本明細書においてリンパ球の「機能的サブセット」と称される。通常、選別は、このような機能的に別個のサブセットに特有の1つまたは複数の分子マーカーの有無に基づく。このようなマーカーは、細胞表面マーカーまたは細胞内マーカーであってよい。1つの態様において、このようなマーカーは細胞表面マーカーである。例示的な細胞表面マーカーには、CD3、CD4、CD8、CD19、CD20、CD25、CD45RO、CD117、CD127などが含まれるが、これらに限定されない。

30

【0016】

関心対象のリンパ球サブセットには、B細胞；T細胞；細胞傷害性T細胞；ヘルパーT細胞；調節性T細胞；Th1ヘルパーT細胞；Th2ヘルパーT細胞；Th9ヘルパーT細胞；Th17ヘルパーT細胞；TfhヘルパーT細胞；抗原特異的T細胞；および抗原特異的B細胞が含まれるが、これらに限定されない。固形組織が固形腫瘍である場合、特に関心が高いのは、細胞傷害性T細胞のサブセットおよび調節性T細胞のサブセットである。

40

【0017】

例えば不十分な選別技法のために重複するという理由で、または第2サブセットのメンバーがより大きな第1サブセット中に完全に含まれ得る（または入れ子になり得る）という理由で、いくつかのサブセットが他のサブセットのメンバーを含む場合がある；例えば、T細胞のサブセットは、完全に含まれる2つのサブセットとして、細胞傷害性T細胞およびヘルパーT細胞を含む。同様に、ヘルパーT細胞のサブセットも、言及した通り、完全に含まれるいくつかの他のサブセットを含む。典型的には、入れ子になったサブセット（すなわち、サブセットのサブセット）の細胞は、そのようなサブセットに特有の付加的なマーカーを用いることによって得られる。リンパ球のサブセットは通常は、多くの場合に市販のマーカーおよびキット（例えば、BD Biosciences、San Jose, CA）を用いて、従来の

50

アッセイ法を使用して、機能的におよび/または分子マーカーにより同定される。関心対象のいくつかのリンパ球に特徴的なマーカーは、以下の通りである：ヘルパーT細胞に対するCD4；細胞傷害性T細胞に対するCD8；調節性T細胞に対するCD4、CD25、および低CD127発現（またはその代わりに、調節性T細胞に対するCD4、CD25、および細胞内FoxP3発現）；ならびに記憶エフェクターT細胞に対するCD45RO+、CCR7-、CD28-、CD27-、CD8+など（ここで、「+」および「-」記号は、免疫学文献における慣習通りに、すなわち、それぞれ高発現および低発現（または発現無し）を示すために、用いられる）。以下に記載される、例えばFACSなどの選別技法により、このようなサブセットを単離するための抗体プローブが市販されている。上記の通り、いくつかの態様において、このようなサブセットのリンパ球の存在、非存在、および/またはレベルは、癌治療を受けている患者の生存期間などの予後情報を提供する。以下の表は、本発明に従ってリンパ球サブセットを同定するための表面マーカーおよび細胞内マーカーをまとめている。本発明の異なる態様は、本表のリンパ球サブセットの異なる組み合わせのクロノタイプを同定することを含む。

10

【0018】

(表I) FACS単離に有用なリンパ球サブセットの例示的な分子マーカー*

サブセット	細胞表面マーカー	細胞内マーカー
B細胞	Fc受容体, CD19+, CD20+, CD21+, CD22+, CD22+, CD23+	
T細胞	CD3+, CD4+, CD8+	
細胞傷害性T細胞	CD3+, CD4-, CD8+	
ヘルパーT細胞	CD3+, CD4+, CD8-	
Th1ヘルパーT細胞	CD4+, CXCR3	IFN- γ , IL-2, IL12, IL18, IL-27
Th2ヘルパーT細胞	CD4+, CCR4, Crth2	IL-4, IL-, IL-33
Th9ヘルパーT細胞	CD4+	IL-4, TGF- β
Th17ヘルパーT細胞	CD4+, CCR6	IL-17A, IL17F, IL-21, IL-22, IL-26, TNF, CCL20
TfhヘルパーT細胞	CD4+, CXCR5	IL-12, IL-6
調節性T細胞	CD4+, CD25+, CD127 ^{-/低}	FoxP3
抗原特異的B細胞	四量体技術によるBCR	
抗原特異的T細胞	四量体技術によるTCR	

20

30

*排他的または包括的リストであると意図するものではない。

【0019】

表面マーカーに基づいた細胞選別は、典型的には関心対象の細胞表面特性を特異的に認識しこれに結合する抗体または他の試薬を使用する、蛍光活性化細胞選別 (FACS)、磁気活性化細胞選別 (MACS)、パニング、リセットティング (resetting) などを含むがこれらに限定されない、1つまたは複数の技術によって行うことができる。1つの局面においては、細胞を固定し透過処理した後に、例えば、例えばPan et al, PlosOne, 6(3): e17536 (2011)に開示されているような細胞内マーカーに特異的な標識抗体を用いて染色することにより、細胞内マーカーに基づく細胞選別もまた、FACSを用いて行うことができる。このような選別技術およびそれらの適用は、以下の例示的な参考文献に開示されている：Recktenwald et al, editors, 「Cell Separation Methods and Applications」 (Marcel Dekker, 1998)；Kearse, editor, 「T Cell Protocols,」 Methods in Molecular Biology, Vol. 134 (Springer, 2000)；Miltenyi et al, Cytometry, 11: 231-238 (1990)；Davies, chapter 11, 「Cell sorting by flow cytometry,」 in Macey, Editor, Flow Cytometry

40

50

: Principles and Applications (Humana Press, Totowa, NJ)など。本発明にとって特に関心が高いのは、FACSを用いて、例えば、BD Biosciences FACS Aria IIIまたはBD Biosciences Influx (BD Biosciences, San Jose, CA)などの市販の装置ならびに製造業者のプロトコールおよびキットを用いて、リンパ球を関心対象のサブセットに選別することである。FACSを用いて調節性T細胞サブセットを単離することは、Boyce et al, 「Human regulatory T-cell isolation and measurement of function,」 BD Bioscience Application Note (March, 2010) に具体的に開示されており、これは参照により組み入れられる。T細胞受容体またはB細胞受容体のいずれかの抗原特異性に基づいてリンパ球を選別または単離することは、FACSを用いて、またはFACSをMACSなどの他の技術と併用して、行うことができる。抗原特異的T細胞またはB細胞を選別および/または単離するためにこのような技術を用いるための手引きは、以下の例示的な参考文献に開示されており、これらは参照により組み入れられる: Thiel et al, Clin. Immunol., 111(2): 155-161 (2004); Newman et al, J. Immunol. Meth., 272: 177-187 (2003); Hoven et al, J. Immunol. Meth., 117(2): 275-284 (1989); 米国特許第5,213,960号および第5,326,696号; Moody et al, Cytometry A, 73A: 1086-1092 (2008); Gratama et al, Cytometry A, 58A: 79-86 (2004); Davis et al, Nature Reviews Immunology, 11: 551-558 (2011); 米国特許第8,053,235号および第8,309,312号; Lee et al, Nature Medicine, 5(6): 677-685 (1999); Altman et al, Science, 274: 94-96 (1996); Leisner et al, PLoSOne 3(2): e1678 (2008); 「Pro5 MHC Pentamer Handbook,」 (ProImmune, Ltd., United Kingdom, 2012); および同様の参考文献。

10

20

【0020】

1つの態様において、入手しやすい組織(末梢血など)に由来する細胞の連続した試料は、2つの集団に選別することができる: (i) 単一の明確なサブセット(CD8+リンパ球; CD4+, CD25^(高); およびCD127^(低)など)、および(ii) 他のすべての細胞。集団(i)を収集し、例えば、核酸を抽出し、組換えDNAまたはRNA配列を増幅し、それらを配列決定し、クロノタイププロファイルを生成することによって、解析する。この手順は、異なるサブセット特異的プローブを用いて、所望のサブセットと同じ数だけ反復することができる。

【0021】

図1に戻り、選別されたサブセットの各々からDNAまたはRNAを抽出し、以下でより十分に記載される技法を用いて、各サブセットについてクロノタイププロファイル(108)を生成する。クロノタイププロファイルは、各サブセットにおけるクロノタイプ配列のリストを提供する。1つの態様において、選別されるリンパ球の数は、別個のクロノタイプを有する実質的にすべてのT細胞がクロノタイププロファイルにおいて同定され得るように、十分に多い。以下でより十分に論じられるように、いくつかの態様では、同定はサンプリングを含むため、「別個のクロノタイプを有する実質的にすべての細胞」とは、90パーセントまたは95パーセントなどの確率で、例えば0.0001などの所与の頻度またはそれ以上の頻度で存在するすべてのクロノタイプが決定されることを意味する。本発明に従って、リンパ球サブセットのクロノタイプ情報を用いて、固形腫瘍標本もしくは生検材料または自己免疫疾患に関与する組織などの、入手しやすさに劣る組織(110)中に浸潤したリンパ球の様々なサブセットの細胞の存在、非存在、数、および/またはレベルを同定する。これは、標本(110)からDNAまたはRNAを抽出し、クロノタイププロファイル(114)を生成すること(112)によって達成される。次に、プロファイル(114)の各クロノタイプは、そのようなクロノタイプの配列をサブセット特異的クロノタイププロファイル(108)において探すことにより、リンパ球サブセットと関連付けることができる(116); したがって、このような関連付けを行うことによって、固形組織に浸潤したメンバーを有するサブセットが同定される。その上、クロノタイプの多様性が高いことから、各サブセットのクロノタイプをカウントし、これにより、標本中に浸潤したサブセットのリンパ球の数の良好な概数が得られる。標本の量が判明しているかまたは決定可能である場合には、サブセットのリンパ球の密度を得ることができる。

30

40

50

【0022】

入手しにくい組織が固形腫瘍である場合などのいくつかの態様において、腫瘍の1つまたは複数の試料は、切除または外科的除去の前または後のいずれかに取得することができる。腫瘍除去の前に取得される試料は、針吸引またはその他の簡便な技法を用いて得ることができる。いくつかの態様では、例えば入手しにくい組織内でのリンパ球サブセットの空間分布を決定するために、複数の試料が得られる。いくつかの態様では、少なくとも1つは入手しにくい組織の表面または外部から、および少なくとも1つは入手しにくい組織の内部からという、少なくとも2つの試料を取得することができる。上述の通り、いくつかの態様において、入手しにくい組織は、図1中の標本(110)によって示されるような、患者から除去された固形腫瘍である。標本(110)に由来する試料は、切除後および固定後に得ることができる。固定組織試料からクロノタイププロファイルを生成することは、以下でより十分に記載される。

10

【0023】

本発明に従って、図1の態様を、以下の工程を用いて実行することができる：(a) 個体の入手しやすい組織に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；(b) 該入手しやすい組織に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；(c) 固形組織の試料からクロノタイププロファイルを生成する工程；および (d) 該固形組織における各サブセットのリンパ球を、そのそれぞれのクロノタイプから検出する工程。いくつかの態様において、工程(a)は、上記のように種々の技法により、例えば適切なマーカーに対する標識抗体プローブを用いるFACSにより、リンパ球を所望のまたは所定のサブセットに分離することによって、実行することができる。この工程の目的は、入手しにくい組織において同定されるクロノタイプからサブセットメンバーが誤ってコールされることを最小限にするために、入手しやすい組織からリンパ球の純粋なサブセットを濃縮する、または好ましくは単離することである。濃縮の程度は、使用する分離または選別技法、およびサブセットに対して利用可能なマーカーに依存する。いくつかの態様において、選別工程は、選別された集団の少なくとも50パーセントが標的リンパ球を含むような、標的リンパ球について濃縮された少なくとも1つのサブセットをもたらす。他の態様において、選別工程は、選別された集団の少なくとも80パーセントが標的リンパ球を含むような、標的リンパ球について濃縮された少なくとも1つのサブセットをもたらす。他の態様において、選別工程は、選別された集団の少なくとも90パーセントが標的リンパ球を含むような、標的リンパ球について濃縮された少なくとも1つのサブセットをもたらす。さらなる他の態様において、例えば、標的リンパ球が稀な細胞集団に属するか、または効率的なプローブが利用できない場合には、選別工程は、選別された集団のわずかに5パーセントのレベルにまでしか濃縮されていないサブセットをもたらす場合がある。このような場合、例えばMACSの後にFACSを行うなど、複数の選別技法を並行して用いることによって、さらなる濃縮が得られ得る。いくつかの態様では、以下に記載されるように、クロノタイププロファイルを生成する工程は、リンパ球に由来する組換え核酸を増幅し、得られたアンプリコンから単離された核酸を配列決定することによって、実行することができる。生成工程は、配列決定工程で得られた配列リードをクロノタイプに合体させることをさらに含み得る。また、生成工程は、得られたクロノタイプ配列のデータベース、例えばこのような配列を他のクロノタイププロファイルのクロノタイプ配列と比較するためのアルゴリズムを適用するといった解析に適しているデータベースを形成することをさらに含み得る。

20

30

40

【0024】

上記の通り、本発明は、以下の工程を含む、患者の固形腫瘍中へのリンパ球浸潤の状態から予後を判定する方法を含む：(a) 患者の末梢血に由来するリンパ球の試料を、1つまたは複数のサブセットに選別する工程；(b) 該末梢血に由来するリンパ球の該1つまたは複数のサブセットの各々について、クロノタイププロファイルを生成する工程；(c) 固形腫瘍の少なくとも1つの試料から、少なくとも1つのクロノタイププロファイルを生成する工程；ならびに (d) 該1つまたは複数のサブセットの各々のリンパ球の数、レベル、およ

50

び/または比率を決定する工程。本明細書で用いられる場合、「予後」とは、固形腫瘍などの入手しにくい組織におけるリンパ球の機能的サブセットの数、レベル、比率、および/または分布に基づいた、転帰の予測を意味する。転帰は、患者の生存、症状の回復の程度、腫瘍量の減少、または疾患状態の改善もしくは悪化の他の代替尺度であってよい。いくつかの態様において、予後は、測定によって改善または悪化は示されるが、改善の程度（例えば、追加の生存年数等）または悪化の程度は示されないという点で、定性的であってよい。いくつかの態様において、機能的サブセットのリンパ球のレベルは、例えばその患者の他の組織におけるレベルもしくは濃度（平均またはその他）、または個体の集団における平均レベルもしくは範囲と比較した相対値であってよい。1つの態様において、相対レベルは、患者の末梢血におけるリンパ球の機能的サブセットのレベルと比較したものである。

10

【0025】

試料

本発明に従って、入手しやすい組織に由来するリンパ球をサブセットに分離し、これを解析してクロノタイプを決定し、次にこのクロノタイプを用いて、入手しやすさに劣る組織における異なるサブセットのリンパ球の数および/またはレベルを決定する；したがって、大部分の態様において、少なくとも1つは入手しやすい組織から、および少なくとも1つは入手しにくい組織からという、少なくとも2つの種類の試料を得る。いくつかの態様において、そこから試料が取得される入手しやすい組織には、末梢血、骨髄、リンパ液、滑液などが含まれるが、これらに限定されない。いくつかの態様において、そこから試料が取得される入手しやすさに劣るまたは入手しにくい組織は、固形腫瘍、自己免疫疾患と関連した炎症組織などの固形組織である。入手しやすさに劣る試料がそこから取得される例示的な固形腫瘍には、メラノーマ、結腸直腸、卵巣、胃、乳房、肝細胞、尿路上皮などが含まれるが、これらに限定されない。特に関心が高いのは、結腸直腸腫瘍およびメラノーマである。自己免疫疾患と関連している例示的な固形組織には、結合組織、関節結合組織、筋肉組織、皮膚、肺組織、小腸組織、結腸組織などが含まれるが、これらに限定されない。他の態様において、入手しやすい組織は末梢血であり、入手しやすさに劣る組織は、サンプリングする際に患者にかなりの苦痛をもたらす任意の組織である。例えば、そのような態様において、入手しやすさに劣る組織には、骨髄、リンパ液、滑液など、および先に開示されたような固形組織が含まれ得る。

20

30

【0026】

クロノタイププロフィールは、多種多様な組織中に存在する免疫細胞の試料（入手しやすい組織中または入手しやすさに劣る組織中を問わない）から得られる。関心対象の免疫細胞には、T細胞および/またはB細胞が含まれる。T細胞（Tリンパ球）には、例えばT細胞受容体（TCR）を発現する細胞が含まれる。B細胞（Bリンパ球）には、例えばB細胞受容体（BCR）を発現する細胞が含まれる。T細胞には、ヘルパーT細胞（エフェクターT細胞またはTh細胞）、細胞傷害性T細胞（CTL）、記憶T細胞、および調節性T細胞が含まれ、これらは細胞表面マーカーによって識別され得る。1つの局面において、T細胞の試料は少なくとも1,000個のT細胞を含む；しかしより典型的には、試料は少なくとも10,000個のT細胞を含み、およびより典型的には少なくとも100,000個のT細胞を含む。別の局面において、試料は、細胞1000~1,000,000個の範囲内の数のT細胞を含む。免疫細胞の試料はまたB細胞を含み得る。B細胞には、例えば、形質B細胞、記憶B細胞、B1細胞、B2細胞、辺縁帯B細胞、および濾胞性B細胞が含まれる。B細胞は免疫グロブリン（抗体またはB細胞受容体とも称される）を発現し得る。上記のように、1つの局面において、B細胞の試料は少なくとも1,000個のB細胞を含む；しかしより典型的には、試料は少なくとも10,000個のB細胞を含み、およびより典型的には少なくとも100,000個のB細胞を含む。別の局面において、試料は、B細胞1000~1,000,000個の範囲内の数のB細胞を含む。

40

【0027】

本発明の方法において用いられる試料（場合により「組織試料」と称される）は、例えば、腫瘍組織、血液および血漿、リンパ液、脳および脊髄を取り囲む脳脊髄液、骨関節を

50

取り囲む滑液などを含む種々の組織に由来し得る。1つの態様において、試料は血液試料である。血液試料は、約0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、1.0、1.5、2.0、2.5、3.0、3.5、4.0、4.5、または5.0 mLであってよい。試料は腫瘍生検材料であってよい。生検材料は、例えば、脳、肝臓、肺、心臓、結腸、腎臓、または骨髄の腫瘍に由来し得る。対象から試料を単離するために、当業者によって用いられる任意の生検技法を用いることができる。例えば、生検は、全身麻酔を用いた開放生検であってよい。生検は、開放生検よりも小さな切開を行う閉鎖生検であってよい。生検は、組織の一部を除去するコア生検または切開生検であってよい。生検は、病変全体を除去することを試みる切除生検であってよい。生検は、針を用いて組織または液体の試料を得る細針吸引生検であってよい。いくつかの態様では、固形腫瘍の内部または周囲でのリンパ球サブセットの空間分布を決定するために、固形腫瘍から複数の試料を取得してもよい。いくつかの態様において、固形腫瘍からの試料の数は、2~10個の範囲内であってよく；他の態様において、そのような範囲は2~20個であってよい。

【0028】

試料または組織試料は、入手しやすいかまたは入手しやすさに劣るかを問わず、例えばDNA（例えば、ゲノムDNA）またはRNA（例えば、メッセンジャーRNA）といった核酸を含む。核酸は、例えば循環系から抽出された無細胞DNAまたはRNAであってよい；Vlassov et al, *Curr. Mol. Med.*, 10: 142-165 (2010)；Swarup et al, *FEBS Lett.*, 581: 795-799 (2007)。本発明の方法において、解析され得る対象由来のRNAまたはDNAの量は大きく異なる。クロノタイププロファイルを生成するためには、組織における個体の免疫受容体レパートリーの有用な表現（representation）を得るために、試料中に十分な核酸を得る。より具体的には、ゲノムDNAからクロノタイププロファイルを生成するためには、T細胞またはB細胞由来の少なくとも1 ngの全DNA（すなわち、約300個の二倍体ゲノム同等物）を試料から抽出し；別の態様においては、少なくとも2 ngの全DNA（すなわち、約600個の二倍体ゲノム同等物）を試料から抽出し；および別の態様においては、少なくとも3 ngの全DNA（すなわち、約900個の二倍体ゲノム同等物）を試料から抽出する。約1000種を超える独立したクロノタイプを含むクロノタイププロファイルを生成するためには、試料中のリンパ球の割合が減少するにつれて、DNAの前述の最小量を増加させなければならないことを当業者は認識するであろう。1つの態様において、RNAからクロノタイププロファイルを生成するためには、別個のTCR、BCR、またはその断片をコードする少なくとも1000種の転写物が得られるように、十分量のRNAを抽出する。この限界に相当するRNAの量は、試料中のリンパ球の割合、リンパ球の発達段階、サンプリング技法、組織の状態などに応じて試料ごとに大きく異なる。1つの態様においては、クロノタイププロファイルを生成するために、B細胞および/またはT細胞を含む組織試料から少なくとも100 ngのRNAを抽出し；別の態様においては、クロノタイププロファイルを生成するために、B細胞および/またはT細胞を含む組織試料から少なくとも500 ngのRNAを抽出する。本発明の方法において用いられるRNAは、組織試料から抽出された全RNA、または組織試料から抽出された全RNAから抽出されたもしくは組織試料から直接抽出されたポリA RNAのいずれかであってよい。上記の核酸抽出は、例えばInvitrogen (Carlsbad, CA)、Qiagen (San Diego, CA)、または同様の販売会社から市販されているキットを用いて行うことができる。RNAを抽出するための手引きは、Liedtke et al, *PCR Methods and Applications*, 4: 185-187 (1994)；および同様の参考文献において見出される。

【0029】

以下（定義）でより十分な議論がなされているように、リンパ球の試料は、別個のクロノタイプを有する実質的にすべてのT細胞またはB細胞がその中で表現（represent）され、それによって（この用語の本明細書における意味での）レパートリーが形成されるように、十分に多量である。1つの態様においては、ある集団の、0.001パーセントまたはそれ以上の頻度で存在するすべてのクロノタイプを、99パーセントの確率で含む試料が取得される。別の態様においては、ある集団の、0.0001パーセントまたはそれ以上の頻度で存在するすべてのクロノタイプを、99パーセントの確率で含む試料が取得される。1つの態様

10

20

30

40

50

において、B細胞またはT細胞の試料は少なくとも50万個の細胞を含み、別の態様においてはそのような試料は少なくとも100万個の細胞を含む。

【0030】

試料が取得される供給源物質が十分でない場合、例えば臨床研究試料等の場合、その物質からDNAが、パイアスのない技法、例えば総ゲノム増幅(WGA)、多置換増幅(MDA)；または同様の技法、例えばHawkins et al, *Curr. Opin. Biotech.*, 13: 65-67 (2002); Dean et al, *Genome Research*, 11: 1095-1099 (2001); Wang et al, *Nucleic Acids Research*, 32: e76 (2004); Hosono et al, *Genome Research*, 13: 954-964 (2003)の技法等により増幅され得る。

【0031】

血液試料は、入手しやすい試料として特に注目されており、従来技法、例えばInnis et al 編, *PCR Protocols* (Academic Press, 1990)等の技法を用いて得られ得る。例えば、白血球は、従来技法、例えばRosetteSepキット(Stem Cell Technologies, Vancouver, Canada)を用いて血液試料から分離され得る。血液試料は、100 μ Lから10 mLの範囲の容量であり得；1つの局面において、血液試料の容量は、100 μ Lから2 mLの範囲である。DNAおよび/またはRNAは、その後、本発明の方法において使用するために、そのような血液試料から、従来技法、例えばDNeasy Blood & Tissueキット(Qiagen, Valencia, CA)を用いて抽出され得る。任意で、白血球のサブセット、例えばリンパ球がさらに、従来技法、例えば蛍光活性化細胞ソーティング(FACS)(Becton Dickinson, San Jose, CA)、磁気活性化細胞ソーティング(MACS)(Miltenyi Biotec, Auburn, CA)等を用いて単離され得る。

【0032】

識別性のある組換えは各個体の適応免疫細胞のDNAおよびそれらに関連するRNA転写物に存在するので、RNAまたはDNAのいずれかが、提供される発明の方法において配列決定され得る。T細胞受容体もしくは免疫グロブリン分子またはそれらの一部分をコードするT細胞またはB細胞由来の組換え配列は、クロナタイプと称される。DNAまたはRNAは、T細胞受容体(TCR)遺伝子または抗体をコードする免疫グロブリン(Ig)遺伝子由来の配列に対応するものであり得る。例えば、DNAおよびRNAは、TCRの、 α 鎖、 β 鎖、または δ 鎖をコードする配列に対応するものであり得る。多数派のT細胞では、TCRは、 α 鎖および β 鎖からなるヘテロ二量体である。TCR α 鎖は、VJ組換えにより生じ、 β 鎖受容体はV(D)J組換えにより生じる。TCR α 鎖に関して、ヒトには48種のVセグメント、2種のDセグメントおよび13種のJセグメントが存在する。2つの接合部の各々ではいくつかの塩基が欠失または付加される(NおよびPヌクレオチドと呼ばれる)。少数派のT細胞では、TCRは、 α 鎖および δ 鎖からなる。TCR δ 鎖は、VJ組換えにより生じ、TCR β 鎖はV(D)J組換えにより生じる(Kenneth Murphy, Paul Travers, and Mark Walport, *Janeway's Immunology 7th edition*, Garland Science, 2007, その全体が参照により本明細書に組み入れられる)。

【0033】

本発明の方法において解析されるDNAおよびRNAは、定常領域(κ 、 λ 、 μ 、または δ)を有する重鎖免疫グロブリン(IgH)または定常領域 κ または λ を有する軽鎖免疫グロブリン(IgKまたはIgL)をコードする配列に対応するものであり得る。各抗体は、2つの同一の軽鎖および2つの同一の重鎖を有する。各鎖は、定常(C)領域および可変領域から構成される。重鎖に関して、可変領域は、可変(V)、多様(D)および結合(J)セグメントから構成される。これらのセグメントの各タイプをコードするいくつかの別個の配列がゲノム中に存在する。B細胞の発達の際に特定のVDJ組換えイベントが起こり、これはその細胞が特定の重鎖を生成することを示すものである。軽鎖における多様性は、D領域がなくVJ組換えのみがある点を除いて、同様の様式で生じる。組換え部位の付近では体細胞変異がしばしば起こり、それによっていくつかのヌクレオチドが付加または欠失し、これがB細胞によって生成される重鎖および軽鎖の多様性をさらに増大させる。B細胞により生成される抗体で生じ得る多様性は、異なる重鎖と軽鎖の掛け算である。重鎖および軽鎖の可変領域は、抗原認識(または結合)領域または部位を形成するのに寄与する。この多様

10

20

30

40

50

性には、あるエピトープに対して特異的な応答が惹起された後に起こり得る体細胞超変異が加わる。

【0034】

試料中のリンパ球の数が決定される1つの局面において、既知の量の既知の配列を有する特有の免疫受容体再構成分子、すなわち、既知の量の1つまたは複数の内部標準が、未知の量の試料由来のcDNAまたはゲノムDNAに添加される。既知の添加された配列について得られる分子の、同じ試料中の残りの配列について得られる分子に対する相対数をカウントすることによって、初期cDNA試料中の再構成免疫受容体分子の数を概算することができる。(そのような分子カウント技法は周知である;例えば、参照により本明細書に組み入れられるBrenner et al, 米国特許第7,537,897号)。添加した特有の配列の配列決定から得られるデータは、同時にリアルタイムPCR較正も使用される場合、異なる可能性を見分けるのに使用することができる;例えば、Faham and Willis(上記で引用)に開示されている。

10

【0035】

固定試料からの核酸の抽出

本発明と併用してそこから核酸が抽出される固定組織試料(例えば、切除された腫瘍組織などに由来する)は、典型的には、固形腫瘍などの疾患関連組織に由来する、化学的に固定された組織試料である。本発明において用いられる固定組織試料を作製するために用いられる化学固定液には、アルデヒド、アルコール、および同様の試薬が含まれる。典型的には、本発明において用いられる固定組織試料は、ホルムアルデヒドまたはグルタルアルデヒドで固定され、特にホルマリン固定パラフィン包埋(FFPE)組織試料として提供される。本発明と共に使用するための核酸抽出技法の手引きは、以下の参考文献に開示されており、これらは参照により組み入れられる: Dedhia et al, Asian Pacific J. Cancer Prev., 8: 55-59 (2007); Okello et al, Analytical Biochemistry, 400: 110-117 (2010); Bereczki et al, Pathology Oncology Research, 13(3): 209-214 (2007); Huijsmans et al, BMC Research Notes, 3: 239 (2010); Wood et al, Nucleic Acids Research, 38(14): e151 (2010); Gilbert et al, PLoSOne, 6: e537 (June 2007); Schweiger et al, PLoSOne, 4(5): e5548 (May 2009)。加えて、製造業者の説明書を用いて本発明と共に使用することができる、固定組織からの核酸抽出を行うためのいくつかの市販のキットが存在する: AllPrep DNA/RNA FFPE Kit (Qiagen, San Diego, CA); Absolutely RNA FFPE Kit (Agilent, Santa Clara, CA); QuickExtract FFPE DNA Extraction Kit (Epicentre, Madison, WI); RecoverAll Total Nucleic Acid Isolation Kit for FFPE (Ambion, Austin, TX) など。

20

30

【0036】

簡潔に説明すると、核酸抽出は以下の工程を含み得る: (i) 約20 μm厚またはそれ未満の切片に切断され、かつ約6 ngの増幅可能なDNAまたは約0.5~20 ngの逆転写可能および増幅可能なRNAをもたらすのに有効な量の固定試料を得る工程; (ii) 任意で、該固定試料を脱蠟する工程、例えば、キシレンおよびエタノール洗浄、d-リモネンおよびエタノール処理、マイクロ波処理などによる; (iii) 任意で、DNAの固定液誘導性架橋を戻すよう処理する工程、例えば、98 °Cで15分間のインキュベーションなど; (iv) 該固定試料の非核酸成分を消化する工程、例えば、従来の緩衝液、例えばTris-HCl、EDTA、NaCl、界面活性剤中のプロテイナーゼK、およびその後のプロテイナーゼKの熱変性、その後、得られた溶液は、任意で、クロノタイププロファイルを生成して相関クロノタイプを同定するために直接用いることができる; (v) ならびに任意で、核酸を抽出する工程、例えば、フェノール:クロロホルム抽出およびその後のエタノール沈殿;シリカカラムベースの抽出、例えばQIAamp DNA micro kit (Qiagen, CA) など。RNA単離に関しては、RNA特異的抽出のさらなる工程、例えば、RNase阻害物質処理、DNase処理、グアニジニウムチオシアネート/酸抽出などを行うことができる。付加的な任意の工程は、PCR阻害物質を除去するために、抽出された核酸試料を処理することを含んでよく、例えば、この目的のためにウシ血清アルブミンまたは同様の試薬を用いることができる;例えば、Sato et al, J. Clin. Micr

40

50

obiol., 36(11): 3423-3425 (1998)。

【 0 0 3 7 】

抽出された核酸の量および質は、種々の方法で測定することができ、このような方法には、PicoGreen Quantitation Assay (Molecular Probes, Eugene, OR) ; 2100 Bioanalyzer (Agilent, Santa Clara, CA) による解析 ; TBS-380 Mini-Fluorometer (Turner Biosystems, Sunnyvale, CA) などが含まれるが、これらに限定されない。1つの局面において、ある程度の核酸の質は、Van Dongen et al, Leukemia, 17: 2257-2317 (2003) に開示されているように、所定のサイズ、例えば、100、200、300、および400塩基対を有する内部標準遺伝子からの断片のセットを、例えばマルチプレックスPCRにおいて増幅することにより得ることができる。このような増幅後に、断片をサイズによって分離し、バンドを定量化して、抽出された核酸の断片のサイズ分布を反映するサイズ分布を提供する。

10

【 0 0 3 8 】

固定組織から抽出された核酸は、固定過程のために、典型的平均サイズが約200ヌクレオチドまたはそれ未満であるサイズ分布を有する。クロノタイプを含む断片は、100~400ヌクレオチドの範囲内にあると考えられるサイズを有する ; したがって、DNAを出発材料とするに当たり、抽出された核酸中に増幅可能なクロノタイプが存在することを確実にするためには、試料中のゲノム同等物の数は、所望されるクロノタイプの数をかなりの量、例えば3~6倍上回らなくてはならない。出発材料としてのRNAについても、同様の考慮がなされるべきである。固定による切断および/または付加物が、抽出された配列に沿ってランダムに分布している場合、N塩基対長の領域(例えば、クロノタイプを含む)が切断も付加物も含まない確率は、以下のように推定することができる。各ヌクレオチドが、切断または付加物を含む確率、pを有するとすると(例えば、pは、平均断片サイズの逆数である $1/200$ と考えてよい)、N bpのひと続きが切断も付加物も有しない確率の推定値は、 $(1-p)^N$ である ; 例えば、Ross, Introduction to Probability Models, Ninth Edition (Academic Press, 2006)。この量の逆数が、増幅可能な所望の断片の数を(平均して)得るためにサンプリングされなければならない、ゲノム同等物における係数増加である。例えば、少なくとも1000種の増幅可能なクロノタイプが所望されるのであれば、切断も、増幅阻害性の付加物または架橋も有しないクロノタイプ配列(例えば、300塩基対(bp)よりも大きい)を包含する少なくとも1000種の配列が存在しなくてはならない。N=300および $p=1/200$ 、 $(1-p)^N=0.22$ である場合、非固定組織からインタクトなDNAの約1000ゲノム同等物を得るために6 ngの試料が必要であるとすると、固定組織からは約 $(1/0.22) \times 6$ ngまたは25~30 ngが必要である。N=100および $p=1/200$ 、 $(1-p)^N=0.61$ である場合、非固定組織からインタクトなDNAの約1000ゲノム同等物を得るために6 ngの試料が必要であるとすると、固定組織からは約 $(1/0.61) \times 6$ ngまたは10 ngが必要である。1つの局面において、相関するクロノタイプを決定するためには、増幅可能なクロノタイプの数は1000~10000の範囲内にある。したがって、約50~100%のリンパ球を含む固定組織試料に関しては、10~500 ngの範囲内の量の、固定組織由来の核酸試料を得る。約1~10%のリンパ球を含む固定組織試料に関しては、1~50 μ gの範囲内の量の、固定組織由来の核酸試料を得る。

20

30

【 0 0 3 9 】

B細胞アイソタイプの同定

1つの態様において、本発明により、入手しにくい組織に浸潤するB細胞のアイソタイプの同定が可能になる。Bリンパ球によって産生される免疫グロブリンのアイソタイプは、免疫グロブリンの定常領域の一部をコードする核酸を含むように設計されたクロノタイプから決定され得る。したがって、本発明の1つの局面に従って、クロノタイプは、免疫グロブリン重鎖(IgH)をコードするヌクレオチドの配列リードから構築される。本発明のこのようなクロノタイプは、VDJコード領域の一部およびそれと関連する定常領域(またはC領域)の一部を含む。アイソタイプは、C領域の部分をコードするヌクレオチド配列から決定される。1つの態様において、C領域をコードする部分はVDJコード領域と隣接しており、その結果、参照により本明細書に組み入れられるFaham and Willis、米国特許出願公開第2011/0207134号によって開示されているように、ポリメラーゼ連鎖反応法(

40

50

PCR) などの簡便な技法によって、単一の連続配列が増幅され得る。C領域をコードするクロノタイプの部分は、特徴的なアレルの存在によってアイソタイプを同定するために用いられる。1つの態様では、8~100個のC領域コードヌクレオチドがクロノタイプ中に含まれる；別の態様では、8~20個のC領域コードヌクレオチドがクロノタイプ中に含まれる。1つの態様において、このようなC領域コード部分は、以下でより十分に記載されるように、IgHコード配列の増幅中に取り込まれる。このような増幅では、上記の範囲内の数のC領域コードヌクレオチドが、結果として得られるアンプリコン中に取り込まれるように、1種または複数種のC領域プライマーを配置する。

【0040】

哺乳動物Ig重鎖には、ギリシャ文字 α 、 β 、 γ 、 δ 、および μ で示される5つの型がある。存在する重鎖の型によって抗体のクラスが規定される；これらの鎖は、それぞれIgA、IgD、IgE、IgG、およびIgM抗体中に見出される。別種の重鎖はサイズおよび組成が異なる； α および β はおよそ450アミノ酸を含むのに対して、 μ および δ はおよそ550アミノ酸を有する。各重鎖は2つの領域、定常領域および可変領域を有する。定常領域は、同じアイソタイプの抗体すべてにおいて同一であるが、異なるアイソタイプの抗体では異なる。重鎖 α 、 β 、および δ は、3つの直列の（一列に並んだ）Igドメイン、および可動性を与えるためのヒンジ領域から構成される定常領域を有し；重鎖 μ および δ は、4つの免疫グロブリンドメインから構成される定常領域を有する。重鎖の可変領域は、異なるB細胞によって産生される抗体において異なるが、単一B細胞またはB細胞クローンによって産生されるすべての抗体については同じである。各重鎖の可変領域はおよそ110アミノ酸の長さであり、単一のIgドメインから構成される。ヒト（および他の）IgH C領域のヌクレオチド配列は、<http://www.imgt.org>における国際免疫遺伝情報システム（IMGT）などの、公的に利用可能なデータベースから得ることができる。

【0041】

上記のように、いくつかの態様において、本発明の方法は、アイソタイプ情報を含む免疫グロブリンのクロノタイプの生成を提供する。このような方法は、以下の工程を用いて実行することができる：(a) 個体のリンパ球から核酸の試料を得る工程であって、該試料が、B細胞受容体のC遺伝子セグメントの少なくとも一部分を各々含む組換え配列を含む、工程；(b) 該組換え配列からアンプリコンを生成する工程であって、該アンプリコンの各配列がC遺伝子セグメントの一部分を含む、工程；(c) クロノタイプのプロファイルを生成するために、該アンプリコンを配列決定する工程であって、各クロノタイプが、B細胞受容体のVDJ領域の少なくとも一部分およびC遺伝子セグメントの少なくとも一部分を含む、工程。後者の工程からは、サンプリングされたBリンパ球のアイソタイプが、そのクロノタイプのC遺伝子セグメントの配列を調べることにより決定される。1つの態様において、C遺伝子セグメントは、B細胞受容体のIgH鎖をコードするヌクレオチド配列に由来する。典型的には、C遺伝子セグメントはクロノタイプの一方向の端に存在し、かつ特有の組換え配列部分、例えばVDJ部分は、クロノタイプのもう一方の端に存在する。いくつかの態様において、クロノタイプの特有の部分はVDJ領域の少なくとも一部分を含む。

【0042】

核酸集団の増幅

標的核酸集団のアンプリコンは、様々な増幅技法により生成され得る。本発明の1つの局面においては、マルチプレックスPCRが、核酸の混合物、特に組換え免疫分子、例えばT細胞受容体またはそれらの一部分を含む混合物のメンバーを増幅するのに使用される。そのような免疫分子のマルチプレックスPCRを実施するための手引きは、参照により組み入れられる以下の参考文献において見出される：Faham and Willis、米国特許出願公開第2011/0207134号；Morley、米国特許第5,296,351号；Gorski、米国特許第5,837,447号；Dau、米国特許第6,087,096号；Von Dongen et al、米国特許出願公開第2006/0234234号；欧州特許公報EP 1544308B1等。

【0043】

ゲノムからのDNA増幅（またはRNAの逆転写によるcDNA形式での核酸増幅）の後、個々の

核酸分子は、単離され、任意で再増幅され、次いで個別に配列決定され得る。例示的な増幅プロトコールは、参照により組み入れられるvan Dongen et al, Leukemia, 17: 2257-2317 (2003)またはvan Dongen et al、米国特許出願公開第2006/0234234号に見出され得る。簡潔に説明すると、例示的なプロトコールは以下のようなものである：反応緩衝液：ABI Buffer IIまたはABI Gold Buffer (Life Technologies, San Diego, CA)；50 μ Lの最終反応容量；100 ngの試料DNA；10 pmolの各プライマー（以下に記載されるように増幅のバランスをとるために調整される）；終濃度200 μ MのdNTP；終濃度1.5 mMのMgCl₂（標的配列およびポリメラーゼに依存して最適化される）；Taqポリメラーゼ（1~2U/チューブ）；サイクル条件：95 での予備活性化7分間；60 でのアニール；サイクル時間：30秒間の変性；30秒間のアニール；30秒間の伸長。本発明の方法における増幅に使用することができるポリメラーゼは市販されており、例えば、Taqポリメラーゼ、AccuPrimeポリメラーゼまたはPfuを含む。使用するポリメラーゼの選択は、忠実性と効率性のどちらが好ましいかに基づくものであり得る。

10

【0044】

リアルタイムPCR、ピコグリーン染色、ナノ流体電気泳動（例えば、LabChip）またはUV吸収測定は、初期工程において、増幅可能な物質の関数的な量を判断するのに使用することができる。

【0045】

1つの局面において、マルチプレックス増幅は、出発集団における配列の相対量が増幅集団またはアンプリコンにおけるそれと実質的に同じになるように実行される。すなわち、マルチプレックス増幅は、試料集団のメンバー配列間の増幅バイアスが最小となるように実行される。1つの態様において、そのような相対量は、アンプリコンにおける各相対量が出発試料におけるその値の5倍以内である場合、実質的に同じとされる。別の態様において、そのような相対量は、アンプリコンにおける各相対量が出発試料におけるその値の2倍以内である場合、実質的に同じとされる。以下でより十分な議論がなされているように、PCRにおける増幅バイアスは、任意の試料においてバイアスのない増幅を提供するPCRプライマーセットが既定のレパートリーのために選択され得る従来技法を用いて、検出および補正され得る。

20

【0046】

TCRまたはBCR配列に基づく多くのレパートリーに関して、マルチプレックス増幅は、任意で、すべてのVセグメントを使用する。その反応は、異なるVセグメントプライマーにより増幅される配列の相対存在量を維持する増幅となるよう最適化される。プライマーのいくつかは関連するものであり、したがってプライマーの多くは「クロストーク」し、それと完全には一致しない鋳型を増幅し得る。その条件は、どのプライマーがそれを増幅するかによらず各鋳型が同様の様式で増幅され得るように最適化される。換言すると、2つの鋳型が存在する場合、1,000倍増幅後には両方の鋳型がおよそ1,000倍増幅され得るが、鋳型の一方において増幅産物の半分がクロストークのために異なるプライマーを保持していることは問題とならない。その後の配列決定データの解析において、プライマー配列は解析から除外され、したがって鋳型が均等に増幅されている限りどのプライマーが増幅に使用されたかは問題とならない。

30

40

【0047】

1つの態様において、増幅のバイアスは、標的配列と非相補的な尾部を有するプライマーを用いて第1または一次段階において少数の増幅サイクルを実施する二段階増幅（上記で引用したFaham and Willis に記載されている）を実施することによって回避され得る。この尾部は、一次アンプリコンの配列の末端に付加されるプライマー結合部位を含み、そのような部位は、1つのみの順方向プライマーおよび1つのみの逆方向プライマーを用いる第2段階の増幅において使用され、それによって増幅のバイアスの主たる原因が除かれる。好ましくは、一次PCRは、異なるプライマーによる差次的増幅を最小限にするよう十分少ないサイクル数（例えば、5~10）を有する。二次増幅は一对のプライマーを用いて実行され、したがって差次的増幅の問題は非常に小さい。一次PCRの1パーセントが二次PC

50

Rに直接利用される。2つの増幅間で使用される35サイクル（100倍希釈工程のない場合の約28サイクルに相当）は、サイクルの内訳によらずに、一次で1サイクル、二次で34サイクルであろうが、一次で25サイクル、二次で10サイクルであろうが、堅調な増幅を示すのに十分であった。理論的には一次PCRにおいて1サイクルのみ実施すれば増幅のバイアスが小さくなる可能性があるが、それ以外にも考慮すべきことがある。この1つの局面は、表現である。これは、出発投入量が最終的に得られるリード数に対して過剰でない場合に影響力がある。例えば、1,000,000のリードが得られ、1,000,000の投入分子で開始する場合、100,000の分子からの表現のみを二次増幅に移すのでは、当初試料中の異なる種の相対存在量の概算の精度が下がる。2つの工程の間での100倍の希釈は、一次PCR増幅が100より有意に多い分子を生成していない限り、その表現が縮小することを意味する。これは、最小で8サイクル（256倍）であるが余裕をもって10サイクル（約1,000倍）が使用され得ることを示している。その代案は、一次PCRの1%超を二次に利用することであるが、一次PCRにおいて使用されるプライマー濃度は高いため、これらのプライマーが増幅において干渉し配列間の増幅バイアスを悪化させないことを確実にするために、大きな希釈係数が使用され得る。別の代案は、精製または酵素処理工程を追加して一次PCR由来のプライマーを排除し、その希釈率を小さくすることである。この実施例では、一次PCRは10サイクル、二次は25サイクルであった。

10

【0048】

クロノタイプのための配列リードの生成

本発明の方法では、任意のハイスループット核酸配列決定技法を使用することができる。好ましくは、このような技法は、そこから少なくとも1000種のクロノタイプが決定され得る、および好ましくはそこから少なくとも10,000~1,000,000種のクロノタイプが決定され得る大量の配列データを、費用に対して効果の高い方法で生成する能力を有する。DNA配列決定技法は、標識されたターミネーターまたはプライマーおよびスラブまたはキャピラリー中でのゲル分離を用いるジデオキシ配列決定反応（サンガー法）、可逆的に停止される標識ヌクレオチドを用いる合成による配列決定、パイロシークエンシング、454配列決定、標識オリゴヌクレオチドプローブのライブラリーに対するアレル特異的ハイブリダイゼーション、標識クローンのライブラリーに対するアレル特異的ハイブリダイゼーションおよびその後のライゲーションを用いる合成による配列決定、重合工程における標識ヌクレオチドの組み込みのリアルタイムモニタリング、ポロニーシークエンシング、ならびにSOLiD配列決定を含む。分離された分子の配列決定は、最近になって、ポリメラーゼまたはリガーゼを用いる連続的または単回の伸長反応によって、およびプローブのライブラリーを用いる単回または連続的なディファレンシャルハイブリダイゼーションによって実証されている。これらの反応は、多くのクローン配列に対して並列で実施され、現在の商業利用においては1億超の配列の並列化が実現している。したがってこれらの配列決定アプローチは、T細胞受容体（TCR）および/またはB細胞受容体（BCR）のレパートリーの研究に使用することができる。

20

30

【0049】

本発明の1つの局面においては、個々の分子を固相表面上で空間的に単離し、その表面上で配列決定を並列で行う工程を含むハイスループットの配列決定法が使用される。そのような固相表面は、非多孔性表面（例えば、Solexa配列決定におけるようなもの、例えばBentley et al, Nature, 456: 53-59 (2008)、またはComplete Genomics配列決定、例えばDrmanac et al, Science, 327: 78-81 (2010)）、ビーズまたは粒子に結合された鋳型を含み得るウェルのアレイ（例えば、454と共に用いるもの、例えばMargulies et al, Nature, 437: 376-380 (2005)、またはIon Torrent配列決定、米国特許出願公開第2010/0137143号もしくは第2010/0304982号）、微細加工膜（例えば、SMRT配列決定と共に用いるもの、例えばEid et al, Science, 323: 133-138 (2009)）、またはビーズアレイ（SOLiD配列決定またはポロニーシークエンシングと共に用いるもの、例えばKim et al, Science, 316: 1481-1414 (2007)）を含み得る。

40

【0050】

50

別の局面において、そのような方法は、単離された分子を、それらを固相表面上で空間的に単離する前または後のいずれかにおいて増幅する工程を含む。先行増幅は、エマルジョンベースの増幅、例えばエマルジョンPCR、またはローリングサークル増幅を含み得る。特に関心対象となるものは、参照により組み入れられる、Bentley et al (上記で引用) および製造元の説明書 (例えば、TruSeq (商標) Sample Preparation Kit and Data Sheet, Illumina, Inc., San Diego, CA, 2010) ; さらに以下の参考文献、米国特許第6,090,592号 ; 同第6,300,070号 ; 同第7,115,400号 ; およびEP0972081B1に記載されるような、個々の鋳型分子を固相表面上で空間的に単離し、その後それらをブリッジPCRにより並列で増幅して個別のクローン集団またはクラスターを形成し、次いで配列決定する、Solexaベースの配列決定である。1つの態様において、固相表面上に配置され増幅される個々の分子は、1 cm²あたり少なくとも10⁵クラスターの密度 ; または1 cm²あたり少なくとも5 × 10⁵の密度 ; または1 cm²あたり少なくとも10⁶クラスターの密度のクラスターを形成する。1つの態様においては、比較的高いエラー率を有する配列決定化学が使用される。そのような態様において、そのような化学によりもたらされる平均品質スコアは、配列リード長の単調に下降する関数である。1つの態様において、そのような下降は、配列リードの0.5パーセントが1~75位に少なくとも1つのエラーを有する ; 配列リードの1パーセントが76~100位に少なくとも1つのエラーを有する ; および配列リードの2パーセントが101~125位に少なくとも1つのエラーを有することに相当する。

10

【0051】

1つの局面において、個体の配列ベースのクロノタイププロファイルは、以下の工程を用いて得られる : (a) 個体のT細胞および / またはB細胞から核酸試料を得る工程 ; (b) そのような核酸試料に由来する個々の分子を空間的に単離する工程であって、個々の分子が、該試料中の核酸から生成されかつ体細胞再構成領域またはその一部分を含む少なくとも1つの鋳型を含み、個々の分子が各々、少なくとも1つの配列リードを作成することができる、工程 ; (c) 空間的に単離された個々の分子を配列決定する工程 ; ならびに (d) クロノタイププロファイルを生成するために、該核酸試料に由来する核酸分子の異なる配列の存在量を決定する工程。別の態様において、配列ベースのクロノタイププロファイルは、以下の工程によって生成することができる : (a) T細胞および / またはB細胞を含む患者由来の試料を得る工程 ; (b) 該試料のT細胞および / またはB細胞に由来する核酸の分子を増幅する工程であって、該核酸の分子が、T細胞受容体遺伝子または免疫グロブリン遺伝子に由来する組換え配列を含む、工程 ; (c) クロノタイププロファイルを形成するために、増幅された該核酸の分子を配列決定する工程 ; ならびに (d) 以前に記録されていない任意のその系統発生的クロノタイプを含む、1つまたは複数の患者特異的クロノタイプの存在、非存在、および / またはレベルを、参照により本明細書に組み入れられるFaham and Willis、米国特許出願公開第2011/0207134号によって教示されているように決定する工程。

20

30

【0052】

1つの態様において、体細胞再構成領域の各々は、V領域およびJ領域を含む。別の態様において、配列決定する工程は、前記空間的に単離された個々の分子の各々を双方向に配列決定して、少なくとも1つの順方向配列リードおよび少なくとも1つの逆方向配列リードを作成することを含む。後者の態様に関してさらに、順方向配列リードの少なくとも1つと逆方向配列リードの少なくとも1つとが重複領域を有し、そのような配列リード間の逆相補関係によってそのような重複領域の塩基が決定されるようにする。さらに別の態様において、体細胞再構成領域の各々は、V領域およびJ領域を含み、配列決定する工程はさらに、その順方向配列リードの1つまたは複数と、J領域内のある位置から始まりそれと連結したV領域の方向へと延びる少なくとも1つの逆方向配列リードとから、個々の核酸分子の各々の配列を決定することを含む。別の態様において、個々の分子は、完全IgH分子、不完全IgH分子、完全IgK分子、IgK不活性分子、TCR 分子、TCR 分子、完全TCR 分子および不完全TCR 分子からなる群より選択される核酸を含む。別の態様において、配列決定する工程は、単調に下降する品質スコアを有する配列リードを生成することを含む。後者

40

50

の態様に関してさらに、単調に下降する品質スコアは、その配列リードが下記以上のエラー率を有するものである：配列リードの0.2パーセントが、塩基1～50位において少なくとも1つのエラーを含み、配列リードの0.2から1.0パーセントが、51～75位において少なくとも1つのエラーを含み、配列リードの0.5から1.5パーセントが、76～100位において少なくとも1つのエラーを含む。別の態様において、上記方法は以下の工程を含む：(a) 個体のT細胞および/またはB細胞から核酸試料を得る工程；(b) そのような核酸試料に由来する個々の分子を空間的に単離する工程であって、個々の分子が、該試料中の核酸から各々生成されかつ体細胞再編成領域またはその一部分を各々含む鋳型の入れ子セットを含み、各入れ子セットが、同じ方向に各々延びかつ該入れ子セットが生成された核酸上の異なる位置から各々始まる複数の配列リードを作成することができる、工程；(c) 空間的に単離された個々の分子を配列決定する工程；ならびに (d) クロノタイププロファイルを生成するために、該核酸試料に由来する核酸分子の異なる配列の存在量を決定する工程。1つの態様において、配列決定する工程は、入れ子セットの各々について複数の配列リードを作成することを含む。別の態様において、体細胞再編成領域の各々はV領域およびJ領域を含み、かつ複数の配列リードの各々は該V領域内の異なる位置から始まりそれと連結したJ領域の方向へと延びる。

10

20

30

40

50

【0053】

1つの局面において、個体由来の試料ごとに、本発明の方法において使用される配列決定技法は、1回あたり少なくとも1000種のクロノタイプの配列を生成し；別の局面において、そのような技法は、1回あたり少なくとも10,000種のクロノタイプの配列を生成し；別の局面において、そのような技法は、1回あたり少なくとも100,000種のクロノタイプの配列を生成し；別の局面において、そのような技法は、1回あたり少なくとも500,000種のクロノタイプの配列を生成し；そして別の局面において、そのような技法は、1回あたり少なくとも1,000,000種のクロノタイプの配列を生成する。さらに別の局面において、そのような技法は、個々の試料につき、1回あたり100,000から1,000,000種の間のクロノタイプの配列を生成する。

【0054】

提供される発明の方法において使用される配列決定技法は、1リードあたり約30 bp、約40 bp、約50 bp、約60 bp、約70 bp、約80 bp、約90 bp、約100 bp、約110 bp、約120 bp、1リードあたり約150 bp、約200 bp、約250 bp、約300 bp、約350 bp、約400 bp、約450 bp、約500 bp、約550 bpまたは約600 bpを生成することができる。

【0055】

配列データからのクロノタイプの決定

配列リードデータからのクロノタイプの構築は、一部、そのようなデータを生成するのに使用された配列決定法に依存しており、これは、方法が異なると、予想リード長およびデータ品質が異なるためである。1つのアプローチにおいて、解析用の配列リードデータの生成にSolexaシーケンサーが使用される。1つの態様において、少なくとも100万の鋳型分子をもたらすための少なくとも $0.5 \sim 1.0 \times 10^6$ 個のリンパ球を提供する試料が得られ、鋳型分子は、任意の増幅後に、対応する100万またはそれ以上の鋳型分子クローン集団（またはクラスター）をもたらす得る。Solexaアプローチを含む最もハイスループットな配列決定アプローチについては、配列決定の精度を増大させるべく高い冗長度で各鋳型配列が決定されるように、そのようなクラスターレベルでの過剰サンプリングが望ましい。Solexaベースでの実施について、好ましくは、各々独立した鋳型の配列は10回またはそれ以上決定される。予想リード長およびデータ品質が異なる他の配列決定アプローチについては、同等の配列決定精度のために異なる冗長度レベルが使用され得る。当業者は、上記のパラメータ、例えば試料サイズ、冗長度等が、特定の用途に関連する設計上の選択事項であることを認識している。

【0056】

本発明の1つの局面において、クロノタイプの配列（IgH、TCR、TCR、TCR、TCR、および/またはIgL（IgK）由来のものを含むがこれらに限定されない）は、1つまた

は複数の配列リードからの情報を、例えば選択された鎖のV(D)J領域に沿って組み合わせることにより決定され得る。別の局面において、クロノタイプの配列は、複数の配列リードからの情報を組み合わせることにより決定される。このような複数の配列リードは、センス鎖に沿う1つまたは複数の配列リード（すなわち、「順方向」配列リード）およびその相補鎖に沿う1つまたは複数の配列リード（すなわち、「逆方向」配列リード）を含み得る。同じ鎖に沿って複数の配列リードを生成する場合は、最初に、該配列リードの異なる位置に対して選択されたプライマーを用いて試料分子を増幅することにより、別々の鋳型が生成される。このコンセプトは、図4Aに図示されており、ここではプライマー（404、406および408）が、1回の反応でアンプリコン（それぞれ410、412および414）を生成するのに使用されている。このような増幅は、同じ反応において実施してもよく、または別々の反応において実施してもよい。1つの局面において、PCRが使用される場合は、別々の鋳型を生成するために別々の増幅反応が使用され、これらはその後組み合わせられ、同じ鎖に沿って複数の配列リードを生成するのに使用される。この後者のアプローチは、複数の鋳型の均等増幅を実現するためにプライマー濃度（および/またはその他の反応パラメータ）のバランスをとる必要がない点で好ましい（本明細書において、「バランスのとれた増幅」または「非バイアス増幅」と称されることがある）。別々の反応による鋳型の生成については、図4B~4Cに図示されている。その中で、IgHを含む試料（400）が3等分され（470、472および474）、これらがJ領域プライマー（401）およびV領域プライマー（それぞれ404、406および408）を用いる別々のPCRに添加され、アンプリコン（それぞれ420、422および424）が作製されている。後者のアンプリコンはその後、P5およびP7プライマーを用いる二次PCR（480）において組み合わせられ（478）、ブリッジPCRおよびIllumina GAシーケンサーまたは同等の機器における配列決定のための鋳型（482）が調製される。

【0057】

本発明の配列リードは、様々な長さを有するものであり、それは一部、使用される配列決定技法に依存する。例えば、いくつかの技法では、その実施の中でいくつかのトレードオフ、例えば、(i) 鋳型あたりの配列リードの数と長さおよび(ii) 配列決定作業の費用と時間、が発生し得る。1つの態様において、配列リードは、20から400ヌクレオチドの範囲であり；別の態様において、配列リードは、30から200ヌクレオチドの範囲であり；さらに別の態様において、配列リードは、30から120ヌクレオチドの範囲である。1つの態様においては、1つから4つの配列リードが、各クロノタイプの配列の決定のために生成され；別の態様においては、2つから4つの配列リードが、各クロノタイプの配列の決定のために生成され；そして別の態様においては、2つから3つの配列リードが、各クロノタイプの配列の決定のために生成される。上記の態様において、示されている数は、異なる個体由来の試料を同定するのに使用される配列リードを除いたものである。以下に記載される態様において使用される様々な配列リードの長さもまた、そのリードによって取り込むことが求められる情報により変化し得；例えば、配列リードの出発位置および長さは、NDN領域の長さおよびそのヌクレオチド配列を提供するよう設計され得；したがって全NDN領域に及ぶ配列リードが選択される。他の局面においては、Dおよび/またはNDN領域を組み合わせ（しかし別々にではなく）含む1つまたは複数の配列リードで十分である。

【0058】

上述したように、様々なアルゴリズムを使用して配列リードをクロノタイプに変換することができる。1つの態様において、クロノタイプの配列は、一部には、配列リードを1つまたは複数のV領域参照配列および1つまたは複数のJ領域参照配列とアラインすることにより、および、一部には、例えば、高可変性のNDN領域に関しては、参照配列とのアラインメントを用いない塩基決定により、決定される。様々なアラインメントアルゴリズムが、配列リードおよび参照配列に適用され得る。例えば、アラインメント法を選択する手引きは、参照により組み入れられるBatzoglu, Briefings in Bioinformatics, 6: 6-22 (2005)において利用可能である。1つの局面において、(上述のように)VリードまたはCリードがVおよびJ領域参照配列に対してアラインされる場合、ツリー検索アルゴリズムが使用される；例えば、Gusfield(上記で引用)およびCormen et al, Introduction to Algo

10

20

30

40

50

rithms, Third Edition (The MIT Press, 2009)に概説されている。

【 0 0 5 9 】

別の態様において、少なくとも1つの順方向リードの末端および少なくとも1つの逆方向リードの末端は、重複領域（例えば、図3Aでは308）で重複しており、そのためそれらのリードの塩基は、相互に対して逆相補関係にある。したがって、例えば、重複領域における順方向リードが「5'-acgttgc」である場合、同じ重複領域内の、逆相補関係にある逆方向リードは「5'-gcaacgt」である。1つの局面において、そのような重複領域内の塩基は、少なくとも一部、そのような逆相補関係から決定される。すなわち、予想される重複領域における塩基コール（または関連する品質スコア）の確度は、もしそれが2つの配列リード間の逆相補関係を保存しているまたはそれと符合している場合、高くなるのである。1つの局面において、（図3Aに図示される）TCR およびIgH鎖のクロノタイプは、そのJ領域から始まりそれと連結したV領域の方向へと延びる少なくとも1つの配列リード（本明細書において「Cリード」（304）と称される）およびそのV領域から始まりそれと連結したJ領域の方向へと延びる少なくとも1つの配列リード（本明細書において「Vリード」（306）と称される）により決定される。重複領域（308）は、図3Aに示されるようにNDN領域（315）を含んでいる場合とない場合がある。重複領域（308）は、全体がJ領域内にある、全体がNDN領域内にある、全体がV領域内にある場合もあるし、または、それはJ領域-NDN領域の境界もしくはV領域-NDN領域の境界またはそのような境界の両方を含む場合がある（図3Aに図示されている）。典型的には、そのような配列リードは、合成による配列決定反応においてポリメラーゼにより配列決定プライマー、例えば図3Aの（302）および（310）を伸長することによって生成される；例えば、Metzger, Nature Reviews Genetics, 11: 31-46 (2010); Fuller et al, Nature Biotechnology, 27: 1013-1023 (2009)。プライマー（302）および（310）の結合部位は、それらが配列リードの初期のアラインメントおよび解析のための出発点または投錨点を提供することができるよう、予め決定されている。1つの態様において、Cリードは、例えば図3Aおよび3Bに図示されているように、それがTCR またはIgH鎖のDおよび/またはNDN領域を網羅し、かつ隣接するV領域の一部を含むように位置決めされる。1つの局面において、V領域におけるVリードとCリードの重複は、これらのリードを互いにアラインするのに使用される。他の態様においては、そのような配列リードのアラインメントは必要ではなく、例えば、TCR 鎖では、Vリードがクロノタイプの特定のV領域を同定するのに十分に長いというだけであり得る。この後者の局面は、図3Bに図示されている。配列リード（330）は、V領域を同定するのに使用され、これは別の配列リードと重複しているまたはしておらず、そして別の配列リード（332）は、NDN領域にかかるものであり、その配列を決定するのに使用される。配列リード（332）の、V領域へと延びる部分（334）は、配列リード（332）の配列情報を配列リード（330）のそれと関連付け、クロノタイプを決定するのに使用される。いくつかの配列決定法、例えばSolexa配列決定法のような塩基単位（base-by-base）のアプローチでは、解析における配列決定サイクルの数を最小限にすることによって、配列決定に要する時間および試薬の費用が削減される。任意で、図3Bに示されるように、アンプリコン（300）は、異なる生物学的試料、例えば異なる患者に由来するクロノタイプを区別するための試料タグ（312）を含むように作製される。試料タグ（312）は、プライマーをプライマー結合領域（316）にアニールさせ、それを伸長させて（314）、タグ（312）にかけて配列リードを作成し、そこから試料タグ（312）をデコードすることによって同定され得る。

【 0 0 6 0 】

IgH鎖は、少なくとも2つの要因から、TCR 鎖よりも解析が難しい：i) 体細胞変異の存在がマッピングまたはアラインメントをより困難にしている、およびii) NDN領域が大きいいため、多くの場合CリードにVセグメントの一部をマッピングすることができない。本発明の1つの局面において、この問題は、Vリードを生成するために、V領域に沿う異なる位置に配置される複数のプライマーセットを使用することによって、好ましくは、プライマー結合部位が重複せず間隔を空けて配置され、そして少なくとも1つのプライマー結合部位がNDN領域に隣接するように、例えば、1つの態様ではV-NDN接合部から5から50塩基、

または別の態様ではV-NDN接合部から10から50塩基となるよう、複数のプライマーセットを使用することによって解消される。複数のプライマーセットの冗長度は、体細胞変異による影響を受ける結合部位を有する1つまたは2つのプライマーの不具合によるクロノタイプの検出の失敗の危険を最小限にする。さらに、NDN領域に隣接するプライマー結合部位が少なくとも1つ存在することにより、VリードがCリードと重複する可能性が高くなり、したがってCリードの長さが効果的に伸長される。これにより、すべてのサイズのNDN領域を網羅しかつVおよびJ領域の実質的に全体をそのNDN領域の両方の側にマッピングすることもできる連続配列を生成することが可能となる。そのようなスキームを実施する態様は、図4Aおよび4Dに図示されている。図4Aにおいて、IgH鎖を含む試料(400)は、単一セットのJ領域プライマー(401)および複数(示されているのは3)セットのV領域(402)プライマー(404、406、408)を用いて鎖を増幅し、すべてが同じNDN領域を含みかつV領域(402)の段階的に大きくなる部分(411、413、415)を含む異なる長さを有する複数の入れ子アンプリコン(例えば410、412、414)を作製することにより、各鎖につき複数のアンプリコンを生成することによって配列決定される。入れ子セットのメンバーは、それらのそれぞれのNDN、Jおよび/またはC領域の同一性(実質的同一性)を確認することにより、配列決定後にひとつにグループ化され得、それによって、リード長および/または配列決定品質が限定される他の配列決定プラットフォームの場合よりも長いV(D)Jセグメントの再編成が実現される。1つの態様において、複数のプライマーセットは、2から5の範囲の数であり得る。別の態様において、複数の数は2~3であり;さらに別の態様では複数の数は3である。複数のプライマーの濃度および位置は、様々変化し得る。V領域プライマーの濃度は、同じである場合もそうでない場合もある。1つの態様において、NDN領域に最も近いプライマーは、例えばNDN領域を含むアンプリコンが、得られるアンプリコンにおいて表現されることを確実にするために、その複数の中の他のプライマーよりも高い濃度を有する。複数の3つのプライマーを使用する特定の態様において、60:20:20の濃度比が用いられる。NDN領域(444)に隣接する1つまたは複数のプライマー(例えば、図4Dでは435および437)は、J領域プライマー(432)によって生成される配列リード(442)と重複する1つまたは複数の配列リード(例えば、434および436)を生成するのに使用され得、それによって重複領域(440)における塩基コールの質が改善される。複数のプライマーからの配列リードは、隣接する下流のプライマー結合部位および/または隣接する下流の配列リードと重複している場合もそうでない場合もある。1つの態様においては、NDN領域に隣接する配列リード(例えば、436および438)が、クロノタイプに関連する特定のV領域を同定するのに使用され得る。そのような複数のプライマーは、プライマー結合部位の1つが免疫グロブリンの発達時に超変異している場合に増幅が不完全または不成功となる可能性を低下させる。それはまた、V領域の超変異により導入された多様性がクロノタイプ配列に取り込まれる可能性を高める。二次PCRは、配列決定のための入れ子アンプリコンを調製するために、例えば図示されているようなP5(401)およびP7(404、406、408)プライマーを用いる増幅によりアンプリコン(420、422および424)を作製することによって実施され得、それらは固相表面上に単一分子として配分され得、さらにブリッジPCRまたは同様の技法により増幅される。

【0061】

(特に、IgH鎖の)NDN領域における塩基コールは、図4Eに図示されているように、隣接するJおよびV領域のコドン構造を使用することによって改善することができる。(本明細書において使用される場合、「コドン構造」は、NDN領域の外側のTCRまたはBCR転写物または遺伝子のセグメント、例えばV領域、J領域等の天然のリーディングフレームのコドンを意味する。)図4Bのアンプリコンの拡大図であるアンプリコン(450)は、上側には、Cリード(442)および隣接するVリード(434)の相対位置が、下側には、V領域(430)およびJ領域(446)のそれぞれのコドン構造(452および454)が、示されている。本発明のこの局面によれば、コドン構造(452および454)が従来のVおよびJ参照配列に対するアラインメントにより同定された後、NDN領域(456)の塩基は、配列リード(434)および(442)を用いて、1度に1塩基ずつ、J領域(446)からV領域(430)に向かっておよび反

10

20

30

40

50

対のV領域(430)からJ領域(446)に向かって移動しつつコール(または同定)される。通常の生物学的条件下では、V領域からNDN領域を通してJ領域までのインフレームコドン有する組換えTCRまたはIgH配列のみがタンパク質として発現される。すなわち、体細胞により生成されるパリアントのうち、発現されるのは、そのJ領域およびV領域のコドンフレームが互いに対してインフレームでありかつNDN領域を通じてインフレームの状態にあるもののみである。(ここでは、VおよびJ領域の正確なフレームは、参照配列から決定される。)フレーム外(out-of-frame)配列が1つまたは複数の低品質の塩基コールに基づき同定される場合、その対応するクロノタイプは、再評価のためまたは潜在的な疾患関連異常としてフラグを立てられる。同定された配列がインフレームでありかつ高品質の塩基コールに基づいている場合、そこにはその対応するクロノタイプが正確にコールされている高い信頼性がある。したがって、1つの局面において、本発明は、双方向配列リードからV(D)Jベースのクロノタイプを決定する方法であって、(a)J領域から始まりNDN領域へと延びる少なくとも1つのJ領域配列リードおよびV領域から始まりNDN領域へと延びる少なくとも1つのV領域配列リードを、そのJ領域配列リードおよびV領域配列リードが重複領域で重複しておりかつJ領域およびV領域の各々がコドン構造を有するように、生成する工程; (b)NDN領域へと延びるJ領域のコドン構造がNDN領域へと延びるV領域のコドン構造に対してインフレームであるかどうかを判定する工程を含む前記方法を含む。さらなる態様において、生成する工程は、V領域から始まり、NDN領域を通してJ領域まで延びる少なくとも1つのV領域配列リードを、J領域配列リードおよびV領域配列リードが重複領域で重複するように生成することを含む。

10

20

【0062】

いくつかの態様において、体細胞超変異を起こしたIgHベースのクロノタイプは、以下のようにして決定され得る。体細胞変異は、(関連セグメントの、通常はV、JまたはCの)対応する参照配列の塩基と異なり、かつ統計的に有意な数のリードに存在する、配列決定された塩基と定義される。1つの態様においては、Cリードが、マッピングされたJセグメントに関する体細胞変異を発見するのに使用され得、同様に、Vセグメントに関してはVリードが使用され得る。JまたはVセグメントに直接マッピングされるかまたはNDN境界までのクロノタイプ伸長物の内側であるかのいずれかのCおよびVリードのみが使用される。このようにして、NDN領域は回避され、以前にクロノタイプの決定に使用された同じ「配列情報」が、変異の発見に使用されることはない(実際は異なる組換えNDN領域であるにすぎないのに誤って変異ヌクレオチドとして分類されるのを回避するため)。セグメントタイプごとに、マッピングされたセグメント(優性のアレル)がスキャホールドとして使用され、リードのマッピング段階でこのアレルにマッピングされたすべてのリードが考慮される。少なくとも1つのリードがマッピングされている参照配列の各位置が、体細胞変異について解析される。1つの態様において、非参照塩基を有効な変異として受諾する基準は、以下のものを含む: 1) 所定の変異塩基を有する少なくともN個のリード、2) 少なくとも所定の分数 N/M のリード(Mはこの塩基位置にマッピングされたリードの総数である); および3) 2項分布、変異塩基におけるN個のリードの平均Qスコアおよび非変異塩基を有するリードの数(M-N)に基づく統計的な切り捨て。好ましくは、上記のパラメータは、クロノタイプ単位での変異の誤発見率が1000中1未満、より好ましくは10000中1未満となるように選択される。

30

40

【0063】

配列タグベースの方法は、配列データからクロノタイプを構築するための上記アプローチに代わるものである。配列データは典型的には、配列リードの大量の収集物、すなわち免疫分子を解析するために用いられるDNAシーケンサーからの塩基コールの配列および関連する品質スコアの大量の収集物を含む。クロノタイププロファイルを構築する上での重要な課題は、抽出工程、配列決定化学、増幅化学などの非生物学的原因によるエラーを含む配列リードから、真の相違を含む配列リードを迅速かつ正確に識別することである。クロノタイプを生成するための1つのアプローチでは、特有の配列タグを試料中の各クロノタイプに付着させて、そのようなコンジュゲートの配列リードが増幅または配列決定の前

50

と同じ元のクロノタイプに由来するのかどうかの判定に役立てることができる。配列タグを体細胞組換え核酸分子に付着させて、タグ-分子コンジュゲートを形成することができる。このようなコンジュゲートの各組換え核酸は特有の配列タグを有する。通常、このような付着は、T細胞および/またはB細胞を含む試料から核酸分子が抽出された後に行われる。好ましくは、このような特有の配列タグは、ハミング距離などの、配列の従来の距離尺度により決定して、互いに大きく異なる；したがって、タグ-分子コンジュゲート中の各配列タグのコピーは、本発明の工程によって導入される配列決定エラーまたは増幅エラーが高率であっても、任意の他の特有のタグ配列よりもその祖先のタグ配列に対してはるかに近いままである。例えば、16merの配列タグが使用され、ある1セットのクロノタイプ上のこのような各タグが、該クロノタイプ上の他のすべての配列タグとの、少なくとも8ヌクレオチドという50パーセントのハミング距離を有する場合、配列タグの誤リード（および誤った配列タグを有するクロノタイプの配列リードの不正確なグループ化）として1つのこのようなタグを別のものに変換するには、少なくとも8個の配列決定エラーまたは増幅エラーが必要である。1つの態様において、配列タグは、タグ-分子コンジュゲートを形成するように組換え核酸分子に付着させた後に、タグ-分子コンジュゲートのタグ間のハミング距離が、このような配列タグの全長の少なくとも25パーセントの数となるように、選択される（すなわち、各配列タグは、そのヌクレオチドの少なくとも25パーセントにおいて、他のすべてのそのようなタグと配列が異なる）；別の態様において、各配列タグ間のハミング距離は、このような配列タグの全長の少なくとも50パーセントの数である。

10

20

30

40

50

【0064】

1つの局面において、上記のアプローチは以下の工程によって実行される：(a) T細胞および/またはB細胞を含む個体由来の試料を得る工程；(b) タグ-分子コンジュゲートを形成するために、配列タグをT細胞および/またはB細胞のT細胞受容体遺伝子または免疫グロブリン遺伝子の組換え核酸の分子に付着させる工程であって、該タグ-分子コンジュゲートの実質的にすべての分子がそれぞれ特有の配列タグを有する、工程；(c) 該タグ-分子コンジュゲートを増幅する工程；(d) 該タグ-分子コンジュゲートを配列決定する工程；および (e) 同様の配列タグの配列リードをアラインして、そのレパートリーの同じクロノタイプに対応する配列リードを決定する工程。B細胞またはT細胞を含む試料は、以下でより十分に記載されるような従来の技法を用いて得られる。配列タグを付着させる工程において、好ましくは、配列タグは特有であるばかりでなく、多数の配列決定エラーまたは増幅エラーでさえも、1つの配列タグを別のものに変換する可能性がゼロに近くなるように、互いに十分異なる。配列タグの付着後、大部分の配列決定技術に関して、タグ-分子コンジュゲートの増幅が必要である；しかしながら、1分子配列決定技術が使用される場合、増幅工程は任意である。1分子配列決定技術には、1分子リアルタイム (SMRT) 配列決定、ナノポア配列決定などが含まれるが、これらに限定されない；例えば、米国特許第7,313,308号；第8,153,375号；第7,907,800号；第7,960,116号；第8,137,569号；Manrao et al, Nature Biotechnology, 4(8): 2685-2693 (2012)など。

【0065】

別の局面において、本発明は、特有の配列タグをカウントすることにより、試料中のリンパ球数を決定する方法を含む。配列タグがなくても、TCR またはIgH遺伝子のクロノタイプ、特にV(D)J領域を含むものは、リンパ球およびそのクローンに特有のマーカを提供する。ゲノムDNAから組換え核酸が得られる場合、試料中のリンパ球数は、配列決定後にカウントされた特有のクロノタイプの数によって推定され得る。同じクロノタイプと関連している同一リンパ球の著しいクローン集団が存在する場合には、このアプローチは失敗に終わる。配列タグの使用はこの欠点を克服し、リンパ腫または白血病などの多くのリンパ系障害に罹患している患者においてリンパ球数を提供するのに特に有用である。本発明の1つの局面に従って、白血病の場合のように、大きな支配的なクローンが存在するかどうかにかかわらず、配列タグを用いて、試料中のリンパ球の絶対数を得ることができる。このような方法は、以下の工程によって実行することができる：(a) リンパ球を含む個体由来の試料を得る工程；(b) タグ-分子コンジュゲートを形成するために、配列タグを

リンパ球のT細胞受容体遺伝子または免疫グロブリン遺伝子の組換え核酸の分子に付着させる工程であって、該タグ-分子コンジュゲートの実質的にすべての分子がそれぞれ特有の配列タグを有する、工程；(c) 該タグ-分子コンジュゲートを増幅する工程；(d) 該タグ-分子コンジュゲートを配列決定する工程；および (e) 別個の配列タグの数をカウントして、試料中のリンパ球数を決定する工程。

【0066】

いくつかの態様において、配列タグは、例えば、参照により本明細書に組み入れられる Brenner et al、米国特許第5,846,719号；Brenner et al、米国特許第7,537,897号；Macevicz、国際公開公報WO 2005/111242などに開示されているように、サンプリングによる標識化によって、試料の組換え核酸分子に付着させる。サンプリングによる標識化では、標識される（または一意的にタグ化される）集団のポリヌクレオチドを用いて、はるかにより大きな集団の配列タグを（付着、連結などによって）サンプリングする。すなわち、ポリヌクレオチドの集団がK個のメンバー（同じポリヌクレオチドの複製物を含む）を有し、かつ配列タグの集団がN個のメンバーを有するとすると、 $N \gg K$ である。1つの態様において、本発明と共に用いられる配列タグの集団のサイズは、試料中のクロノタイプの集団のサイズの少なくとも10倍である；別の態様において、本発明と共に用いられる配列タグの集団のサイズは、試料中のクロノタイプの集団のサイズの少なくとも100倍である；および別の態様において、本発明と共に用いられる配列タグの集団のサイズは、試料中のクロノタイプの集団のサイズの少なくとも1000倍である。他の態様において、配列タグ集団のサイズは、例えば、連結反応などの付着反応、増幅反応などにおいて、このようなクロノタイプがこのような配列タグ集団と組み合わせられた場合に、試料中の実質的にすべてのクロノタイプがそれぞれ特有の配列タグを有するように、選択される。いくつかの態様において、実質的にすべてのクロノタイプとは、このようなクロノタイプの少なくとも90パーセントが特有の配列タグを有することを意味する；他の態様において、実質的にすべてのクロノタイプとは、このようなクロノタイプの少なくとも99パーセントが特有の配列タグを有することを意味する；他の態様において、実質的にすべてのクロノタイプとは、このようなクロノタイプの少なくとも99.9パーセントが特有の配列タグを有することを意味する。多くの組織試料または生検材料において、T細胞またはB細胞の数は、最大100万個または約100万個の細胞であってよい；したがって、このような試料を使用する本発明のいくつかの態様において、サンプリングによる標識化において使用される特有の配列タグの数は、少なくとも 10^8 種であり、または他の態様においては少なくとも 10^9 種である。

【0067】

最大100万個のクロノタイプがサンプリングによって標識されるこのような態様において、配列タグの大きなセットは、例えば、参照により本明細書に組み入れられるChurch、米国特許第5,149,625号に開示されているように、合成反応の各付加工程において、全4種のヌクレオチド前駆体の混合物を反応させることにより、コンビナトリアル合成によって効率的に作製され得る。この結果は、「 $N_1N_2 \dots N_k$ 」という構造を有する配列タグのセットであり、式中、各 $N_i = A, C, G$ 、または T であり、かつ k はタグ中のヌクレオチドの数である。このようなコンビナトリアル合成によって作製される配列タグ1セットにおける配列タグの数は、 4^k である。したがって、サンプリングによる標識化により、 10^6 個のメンバーの分子集団に配列タグを付着させるためには、少なくとも14という k 、または約14~18の範囲内の k を有するこのような配列タグのセットが適している。

【0068】

種々の異なる付着反応を用いて、それぞれ特有のタグを試料中の実質的にすべてのクロノタイプに付着させることができる。1つの態様において、このような付着は、2つの分子集団のメンバーがランダムに混合し、例えば共有結合的に会合または連結されるように、組換え核酸分子を含む（同様に、クロノタイプ配列を含む）試料を配列タグの集団またはライブラリーと混合することによって、達成される。このようなタグ付着反応において、クロノタイプ配列は直鎖状の一本鎖または二本鎖ポリヌクレオチドを含み、配列タグは、PCRプライマーなどの増幅プライマー、連結アダプター、環状化可能プローブ、プラスミ

ドなどの試薬によって運ばれる。配列タグ集団を運ぶことができるいくつかのこのような試薬は、参照により本明細書に組み入れられるMacevicz、米国特許第8,137,93号；Faham et al、米国特許第7,862,999号；Landegren et al、米国特許第8,053,188号；Unrau and Deugau, Gene, 145: 163-169 (1994)；Church、米国特許第5,149,625号などに開示されている。

【0069】

TCR レパートリー解析

この実施例では、TCR 鎖を解析する。解析は、TCR 配列の増幅、配列決定、および解析を含む。1つのプライマーは、C₁およびC₂の共通配列に相補的であり、全48種のVセグメントを増幅することができる34種のVプライマーが存在する。C₁またはC₂は、J/C結合部から10位および14位の位置で互いに異なっている。C₁またはC₂のプライマーは16bpの位置で終了し、C₁またはC₂に対する優先性はない。34種のVプライマーを、Van Dongen等の米国特許出願公開第2006/0234234号（参照により本明細書に組み入れられる）に開示の元のプライマーセットから改変する。改変されたプライマーは、参照により本明細書に組み入れられるFahamらの米国特許出願公開第2010/0151471号に開示されている。

【0070】

Illuminaゲノムアナライザーを使用して、上記プライマーにより作製されたアンプリコンを配列決定する。図2A~2Bに記載のようにメッセンジャーRNA転写物（200）の二段階増幅を実施する。第1段階は上記プライマーを用い、第2段階にはブリッジ増幅および配列決定用の共通プライマーを加える。図2Aに示すように、3'末端がJ/C結合部（204）から16塩基であり、C₁（203）および2つのC₂アレルに完全に相補的である20bpプライマー（202）を片側に使用して、一次PCRを実施する。RNA転写物（200）のV領域（206）において、異なるV領域配列に相補的なプライマー配列を含有するプライマーセット（212）が提供される（一態様においては34種）。プライマーセット（212）は、また、P7プライマー（220）に特異的なプライマー結合部位（218）を有するアンプリコン（216）を作製する非相補的テール（214）を含有する。従来のマルチプレックスPCRの後、二次増幅のためのmRNA転写物の多様なJ（D）V領域（206、208および210）および共通プライマー結合部位（203および218）を含有するアンプリコン（216）が形成され、ブリッジPCRによるクラスター形成のための試料タグ（221）およびプライマー（220および222）を加える。二次PCRにおいて、鋳型の同じ側で、J/C結合部に最も近い10塩基の配列をその3'末端に有し、J/C結合部から15~31位の17bpの配列が続く、P5配列（224）がそれに続くプライマー（図2Bの222、本明細書において「C10-17-P5」と称される）を使用する。P5は、Solexa配列決定でのブリッジPCRによるクラスター形成において役割を果たす。（C10-17-P5プライマー（222）が第1PCRから作製された鋳型にアニーリングすると、J/C結合部に最も近い10塩基およびJ/C結合部から15~31位の塩基の配列にプライマーがハイブリダイズするために、鋳型に4bpループ（11~14位）が生じる。11~14位のループ形成により、C₁またはC₂を有する鋳型の差次的増幅が排除される。次に、J/C結合部に最も近い10塩基およびJ/C結合部から15~31位の塩基の配列に相補的なプライマー（このプライマーをC'と呼ぶ）を用いて、配列決定を行う。すべての増幅された物質が、クラスター形成において効率的に使用することができるインタクトな末端を有するようにするために、C10-17-P5プライマーをHPLC精製することができる。）

【0071】

図2Aにおいて、Vプライマー（212）のオーバーハングの長さは14bpが好ましい。一次PCRはより短いオーバーハング（214）により支援される。あるいは、二次PCRのためには、二次PCRがこの配列からプライミングされるので、一次PCRではVプライマーのオーバーハングはできるだけ長いものが使用される。効率的な二次PCRを支援するオーバーハング（214）の最小サイズを調査した。2bp刻みで10~30のオーバーハングサイズを有するVプライマーを二系列（2つの異なるVセグメント用に）作製した。適切な合成配列を用いて、これらの系列における各プライマーにより第1PCRを実施し、増幅されたものすべてを示すため

10

20

30

40

50

にゲル電気泳動を実施した。

【0072】

図2Aに示されるように、一次PCRは、RNA鋳型(200)のV領域(206)にアニーリングし、共通の14bpのオーバーハングを5'テールに含有する34種の異なるVプライマー(212)を使用する。14bpは、Illumina配列プライマーの1つ(Read2プライマーと呼ぶ)の部分配列である。同じ側における二次増幅プライマー(220)は、P7配列、タグ(221)およびRead2プライマー配列(223)を含む(このプライマーはRead2_タグX_P7と呼ぶ)。P7配列はクラスター形成に使用される。Read2プライマーおよびその相補体は、それぞれVセグメントおよびタグの配列決定に用いられる。1~96番のタグを有する96種のこれらのプライマーのセットを作製する(下記参照)。すべての増幅された物質が、クラスター形成で効率的に使用することができるインタクトな末端を有するようにするため、これらのプライマーはHPLC精製される。

10

【0073】

上記のように、第2段階のプライマーであるC-10-17-P5(222、図2B)は第1段階のPCRで生成された鋳型に対して中断のある相同性を有する。このプライマーを用いた増幅の効率を検証した。CsegP5と呼ばれるC-10-17-P5の代わりにプライマーは、第1段階のCプライマーおよびP5を有する5'テールに完全な相同性を有する。第1段階のPCRの鋳型を増幅する場合のC-10-17-P5およびCsegP5を用いた効率は、リアルタイムPCRを実施することにより比較した。数回の繰り返しで、C-10-17-P5プライマーを用いたPCRは、CsegP5プライマーを用いたPCRと比較して効率にほとんど違いが認められなかった。

20

【0074】

図2A~2Cに示される二段階増幅から生じるアンプリコン(230)は、図2Cに示されるような、Illuminaシーケンサーで典型的に用いられる構造を有する。分子の最も外側の部分にアニールする2種のプライマー、IlluminaプライマーP5およびP7は、分子の固相増幅(クラスター形成)に用いられる。1分子当たり3件の配列リードを行う。100 bpの第1リードは、Illumina配列決定過程に適している融解温度を有するC'プライマーを用いて行う。第2リードは6 bp長のみであり、単に試料タグを同定するためのものである。これは、製造業者(Illumina)によって提供されるタグプライマーを用いて生成される。最終リードは、やはり製造業者(Illumina)によって提供されるRead 2プライマーである。このプライマーを用いて、第1PCRのVプライマー配列から開始する、Vセグメントにおける100 bpリードが生成される。

30

【0075】

いくつかの特定の態様例を参照して本発明を説明してきたが、本発明の精神および範囲から逸脱することなく、それらに対して多くの変更がなされ得ることを、当業者は認識するであろう。本発明は、上記のものに加えて、様々なセンサー実現およびその他の主題に適用可能である。

【0076】

定義

本明細書において他に具体的に規定されない限り、本明細書で用いられる核酸化学、生化学、遺伝学、および分子生物学の用語および記号は、当分野における標準的な論文および教科書、例えば、Kornberg and Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999); Abbas et al, Cellular and Molecular Immunology, 6th edition (Saunders, 2007)の用語および記号に従っている。

40

【0077】

「アラインすること」とは、配列リードなどの試験配列を1つまたは複数の参照配列と比較して、どの参照配列が、または参照配列のどの部分が、何らかの配列距離尺度に基づいて最も近いかを判定する方法を意味する。ヌクレオチド配列をアラインする例示的な方法は、Smith Watermanアルゴリズムである。距離尺度には、ハミング距離、レーヴェンシ

50

ユタイン距離などが含まれ得る。距離尺度は、比較される配列のヌクレオチドの品質値に関連した成分を含み得る。

【0078】

「アンプリコン」とは、ポリヌクレオチド増幅反応の産物；すなわち、一本鎖または二本鎖であってよく、1つまたは複数の出発配列から複製されるポリヌクレオチドのクローン集団を意味する。1つもしくは複数の出発配列は、同じ配列の1つもしくは複数のコピーであってもよいし、またはそれらは異なる配列の混合物であってもよい。好ましくは、アンプリコンは、単一の出発配列の増幅によって形成される。アンプリコンは、その反応産物が1つまたは複数の出発核酸または標的核酸の複製物を含む、様々な増幅反応によって作製され得る。1つの局面において、アンプリコンを作製する増幅反応は、ヌクレオチドまたはオリゴヌクレオチドのいずれかである反応物の塩基対形成が、反応産物の創出に必要な相補体を鋳型ポリヌクレオチドにおいて有するという点で、「鋳型駆動型」である。1つの局面において、鋳型駆動型反応は、核酸ポリメラーゼによるプライマー伸長、または核酸リガーゼによるオリゴヌクレオチド連結である。このような反応には、参照により本明細書に組み入れられる以下の参考文献、Mullis et al、米国特許第4,683,195号；第4,965,188号；第4,683,202号；第4,800,159号(PCR)；Gelfand et al、米国特許第5,210,015号(「taqman」プローブによるリアルタイムPCR)；Wittwer et al、米国特許第6,174,670号；Kacian et al、米国特許第5,399,491号(「NASBA」)；Lizardi、米国特許第5,854,033号；Aono et al、特開平4-262799(ローリングサークル増幅)などに開示されているポリメラーゼ鎖反応(PCR)、線状ポリメラーゼ反応、核酸配列ベースの増幅(NASBA)、ローリングサークル増幅などが含まれるが、これらに限定されない。1つの局面において、本発明のアンプリコンはPCRによって作製される。増幅反応の進行に伴った反応産物の測定を可能にする検出化学が利用できるのであれば、増幅反応は「リアルタイム」増幅であってよく、例えば、以下に記載される「リアルタイムPCR」、またはLeon et al, *Nucleic Acids Research*, 26: 2150-2155 (1988)、および同様の参考文献に記載されているような「リアルタイムNASBA」であってよい。本明細書で用いられる「増幅すること」という用語は、増幅反応を行うことを意味する。「反応混合物」とは、反応を行うために必要な反応物をすべて含む溶液を意味し、この反応物には、反応中にpHを選択されたレベルに維持するための緩衝剤、塩、補因子、スカベンジャーなどが含まれ得るが、これらに限定されない。

10

20

30

【0079】

「クロノタイプ」とは、T細胞受容体(TCR)もしくはB細胞受容体(BCR)またはその一部分をコードする、T細胞またはB細胞の組換えヌクレオチド配列を意味する。1つの局面において、ある個体のあるリンパ球集団の別個のクロノタイプすべての収集物は、そのような集団の1つのレパートリーである；例えば、Arstila et al, *Science*, 286: 958-961 (1999)；Yassai et al, *Immunogenetics*, 61: 493-502 (2009)；Kedzierska et al, *Mol. Immunol.*, 45(3): 607-618 (2008)など。本発明で用いられる「クロノタイププロファイル」または「レパートリープロファイル」とは、レパートリーのクロノタイプおよびそれらの相対存在量の実質的にすべてを含めた、T細胞および/またはB細胞の試料(そのような細胞を含む末梢血試料など)のクロノタイプの集計である。「クロノタイププロファイル」、「レパートリープロファイル」、および「レパートリー」は、本発明において互換的に用いられる。(すなわち、以下にさらに詳述する「レパートリー」という用語は、リンパ球の試料から測定されるレパートリーを意味する)。本発明の1つの局面において、クロノタイプは、免疫グロブリン重鎖(IgH)またはTCR鎖の一部を含む。本発明の他の局面において、クロノタイプは、免疫グロブリン軽鎖もしくはTCR鎖またはそれらの一部分などのその他の組換え分子に基づき得る。

40

【0080】

「相補性決定領域」(CDR)とは、免疫グロブリン(すなわち、抗体)またはT細胞受容体の領域を意味するものであり、この領域において分子は抗原の高次構造を補完し、それによって該分子の特異性を決定し、特定の抗原と接触する。T細胞受容体および免疫グロブリンはそれぞれ、3つのCDRを有する：CDR1およびCDR2は可変(V)ドメイン中に見出され、CDR

50

3は、Vの一部、多様部(D)(重鎖のみ)および結合部(J)のすべて、ならびに定常(C)ドメインの一部を含む。

【0081】

参照配列と別の配列(「比較配列」と)の比較に関して用いられる「%相同」、「%同一」、または同様の用語は、この2つの配列の間の最適なアラインメントにおいて、比較配列が、示されるパーセンテージに等しいサブユニット位置の数において参照配列と同一であることを意味し、このサブユニットは、ポリヌクレオチド比較に関してはヌクレオチドであり、またはポリペプチド比較に関してはアミノ酸である。本明細書で用いられる、比較される配列の「最適なアラインメント」とは、サブユニット間の一致を最大にし、かつアラインメントの構築において使用されるギャップの数を最小にするアラインメントである。%同一性は、Needleman and Wunsch, *J. Mol. Biol.*, 48: 443-453 (1970) ("GAP" program of Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, WI) などによって記載されているようなアルゴリズムの商用の実行によって決定され得る。アラインメントを構築し、同一性のパーセンテージまたは類似性の他の尺度を算出するための、当技術分野におけるその他のソフトウェアパッケージには、Smith and Waterman, *Advances in Applied Mathematics*, 2: 482-489 (1981) (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, WI) のアルゴリズムに基づく「BestFit」プログラムが含まれる。言い換えれば、例えば、参照ヌクレオチド配列と少なくとも95パーセント同一であるヌクレオチド配列を有するポリヌクレオチドを得るためには、参照配列中のヌクレオチドの5パーセントまでが欠失されるか、もしくは別のヌクレオチドで置換されてよく、または参照配列中の全ヌクレオチド数の5パーセントまでのヌクレオチド数が参照配列中に挿入されてよい。

10

20

【0082】

「フローシステム」とは、(i) 粒子または細胞に関するマルチパラメータデータを収集する1つまたは複数の検出ステーションによって、またはこれを通過して、粒子または細胞を液流内の同一線上経路で移動させることができ、かつ(ii) 収集されたマルチパラメータデータに基づいて、このような粒子の数を数えるまたはこのような粒子を選別することができる、任意の機器または装置を意味する。参照により組み入れられる以下の文献、Shapiro, *Practical Flow Cytometry*, Fourth Edition (Wiley-Liss, 2003)、Bonner et al, *Rev Sci Instruments*, 43 404 (1972)、Huh et al, *Physiol Meas*, 26 R73-98 (2005)、Ateya et al, *Anal Bioanal Chem*, 391 1485-1498 (2008)、Bohm et al、米国特許第7, 157,274号; Wang et al、米国特許第7,068,874号などによって例証されるように、フローシステムは、多種多様な形態を有し、このような機能を達成するために多種多様な技法を使用する。フローシステムは、構成要素を有する流体力学システムを含んでよく、ここでは、例えば蛍光活性化細胞選別(FACS)機器と同様に、試料液流はシース液流中に挿入されて、試料液中の粒子または細胞は同一線上経路で移動するように制約されるが、これはキュベット、検出ステーションとして機能するその他のチャンバー、またはノズルもしくはstream-in-airジェットを作るその他の構造中で行われるものであり、これは次いで電氣的に操作され得る。フローシステム、フローサイトメーター、およびフローソーター、ならびにそれらの一般的な適用は、以下の参考文献の1つまたは複数において開示されており、これらは参照により組み入れられる: Robinson et al (Editors) *Current Protocols in Cytometry* (John Wiley & Sons, 2007); Shapiro, *Practical Flow Cytometry*, Fourth Edition (Wiley-Liss, 2003); Owens et al (Editors), *Flow Cytometry Principles for Clinical Laboratory Practice: Quality Assurance for Quantitative Immunophenotyping* (Wiley-Liss, 1994); Ormerod (Editor) *Flow Cytometry: A Practical Approach* (Oxford University Press, 2000) など。

30

40

【0083】

「ポリメラーゼ連鎖反応」または「PCR」とは、DNAの相補鎖の同時プライマー伸長による、特定のDNA配列のインピット増幅のための反応を意味する。言い換えると、PCRは、プライマー結合部位によって挟まれた標的核酸の複数のコピーまたは複製物を作製するため

50

の反応であり、このような反応は以下の工程の1回または複数回の反復を含む：(i) 標的核酸を変性させる工程、(ii) プライマーをプライマー結合部位にアニーリングさせる工程、および(iii) ヌクレオシド三リン酸の存在下で核酸ポリメラーゼによりプライマーを伸長させる工程。通常、反応は、サーマルサイクラー装置において、各工程に最適化された異なる温度の間で繰り返される。特定の温度、各工程における持続時間、および工程間の変化速度は、例えば参考文献、McPherson et al, editors, PCR: A Practical Approach and PCR2: A Practical Approach (IRL Press, Oxford、それぞれ1991および1995)によって例示される、当業者に周知の多くの要因に依存する。例えば、Taq DNAポリメラーゼを用いる従来のPCRでは、 >90 の温度で二本鎖標的核酸が変性され得、 $50\sim75$ の範囲の温度でプライマーがアニーリングされ得、そして $72\sim78$ の範囲の温度でプライマーが伸長され得る。「PCR」という用語は、RT-PCR、リアルタイムPCR、ネステッドPCR、定量的PCR、マルチプレックスPCRなどを含むがこれらに限定されない、該反応の派生形態を含む。反応容量は、数百ナノリットル、例えば $200\text{ nL}\sim$ 数百 μL 、例えば $200\mu\text{L}$ の範囲である。「逆転写PCR」または「RT-PCR」とは、標的RNAを相補的な一本鎖DNAへと変換する逆転写反応が先行し、次いで増幅が行われるPCRを意味する；例えば、参照により本明細書に組み入れられるTecott et al、米国特許第5,168,038号。「リアルタイムPCR」とは、反応の進行と共に反応産物、すなわちアンプリコンの量がモニターされるPCRを意味する。リアルタイムPCRの多くの形態が存在し、主に反応産物のモニタリングに用いられる検出化学が異なる；例えば、参照により本明細書に組み入れられるGelfand et al、米国特許第5,210,015号(「taqman」)；Wittwer et al、米国特許第6,174,670号および第6,569,627号(インターカレート色素)；Tyagi et al、米国特許第5,925,517号(分子ビーコン)。リアルタイムPCRのための検出化学は、同様に参照により本明細書に組み入れられるMackay et al, Nucleic Acids Research, 30: 1292-1305 (2002)において概説されている。「ネステッドPCR」とは、第1PCRのアンプリコンが、プライマーの新たなセットを用いる第2PCRのための試料となる二段階PCRを意味し、そのセットのうちの少なくとも一方は第1アンプリコンの内部の位置に結合する。ネステッド増幅反応に関する、本明細書で用いられる「初期プライマー」とは、第1アンプリコンを生成するために用いられるプライマーを意味し、「二次プライマー」とは、第2または入れ子アンプリコンを生成するために用いられる1つまたは複数のプライマーを意味する。「マルチプレックスPCR」とは、複数の標的配列(または単一の標的配列および1つまたは複数の参照配列)が同じ反応混合物中で同時に行われるPCRを意味する；例えば、Bernard et al, Anal. Biochem., 273: 221-228 (1999)(2色リアルタイムPCR)。通常、増幅される各配列について、プライマーの別個のセットが用いられる。典型的には、マルチプレックスPCRにおける標的配列の数は、 $2\sim50$ 、または $2\sim40$ 、または $2\sim30$ の範囲である。「定量的PCR」とは、試料または標本中の1つまたは複数の特定の標的配列の存在量を測定するために設計されたPCRを意味する。定量的PCRは、このような標的配列の絶対的な定量および相対的な定量の両方を含む。定量的な測定は、標的配列と別々にまたは共にアッセイされ得る1つまたは複数の参照配列または内部標準を用いてなされる。参照配列は、試料または標本にとって内因性であっても外因性であってもよく、後者の場合には、1つまたは複数の競合鋳型を含み得る。典型的な内因性の参照配列には、以下の遺伝子の転写物のセグメントが含まれる：アクチン、GAPDH、 α_2 ミクログロブリン、リボソームRNAなど。定量的PCRのための技法は、参照により組み入れられる以下の参考文献において例証されるように、当業者に周知である：Freeman et al, Biotechniques, 26: 112-126 (1999)；Becker-Andre et al, Nucleic Acids Research, 17: 9437-9447 (1989)；Zimmerman et al, Biotechniques, 21: 268-279 (1996)；Diviacco et al, Gene, 122: 3013-3020 (1992)；Becker-Andre et al, Nucleic Acids Research, 17: 9437-9446 (1989)など。

【0084】

「プライマー」とは、ポリヌクレオチド鋳型との二重鎖の形成に際して、核酸合成の開始点として働くことができ、伸長された二重鎖が形成されるように、鋳型に沿ってその3'末端から伸長され得る、天然または合成のいずれかのオリゴヌクレオチドを意味する。プ

10

20

30

40

50

ライマーの伸長は、通常、DNAポリメラーゼまたはRNAポリメラーゼなどの核酸ポリメラーゼを用いて行われる。伸長過程において付加されるヌクレオチドの配列は、鋳型ポリヌクレオチドの配列によって決定される。通常、プライマーはDNAポリメラーゼによって伸長される。プライマーは通常、14~40ヌクレオチドの範囲、または18~36ヌクレオチドの範囲の長さを有する。プライマーは、様々な核酸増幅反応、例えば、単一のプライマーを用いる線状増幅反応、または2種もしくはそれ以上のプライマーを用いるポリメラーゼ連鎖反応において用いられる。特定の適用に関してプライマーの長さおよび配列を選択するための指針は、参照により組み入れられる以下の参考文献によって明らかのように、当業者に周知である：Dieffenbach, editor, PCR Primer: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Press, New York, 2003)。

10

【0085】

「品質スコア」とは、特定の配列位置における塩基の割り当てが正しいという確率についての尺度である。異なる配列決定化学、検出システム、塩基コールアルゴリズムなどの結果としてコールされる塩基に関するような特定の状況について、品質スコアを算出するための様々な方法が当業者に周知である。一般的に、品質スコア値は、正しい塩基コールの確率に単調に関連している。例えば、10という品質スコアまたはQは、塩基が正しくコールされる可能性が90パーセントあることを意味し得、20というQは、塩基が正しくコールされる可能性が99パーセントあることを意味し得るなどである。いくつかの配列決定プラットフォーム、特に合成による配列決定化学を用いる配列決定プラットフォームでは、平均品質スコアは配列リード長に応じて低下し、その結果、配列リードの始めの品質スコアは配列リードの終わりの品質スコアよりも高く、そのような低下は、不完全な伸長、繰り返し伸長、鋳型の減少、ポリメラーゼの減少、キャップ形成の障害、脱保護の障害などのような現象に起因する。

20

【0086】

「レパートリー」または「免疫レパートリー」とは、ある個体のあるリンパ球集団中の、T細胞受容体(TCR)もしくはB細胞受容体(BCR)またはそれらの断片それぞれをコードする別個の組換えヌクレオチド配列のセットを意味し、該セットのヌクレオチド配列は、該集団のリンパ球の実質的にすべてについて、別個のリンパ球またはそれらのクローン亜集団と1対1の対応を有する。1つの局面において、レパートリーが決定されるリンパ球集団は、1つまたは複数の血液試料などの1つまたは複数の組織試料から取得される。レパートリーのメンバーヌクレオチド配列は、本明細書において「クロノタイプ」と称される。1つの局面において、レパートリーのクロノタイプは、TCRまたはBCRの発達中に体細胞組換えを起こしたT細胞またはB細胞集団に共通する核酸の任意のセグメントを含み、これには正常なまたは異常な(例えば、癌と関連する)その前駆体分子が含まれ、以下のもののうちのいずれかが含まれるがこれらに限定されない：免疫グロブリン重鎖(IgH)またはそのサブセット(例えば、IgH可変領域、CDR3領域など)、不完全なIgH分子、免疫グロブリン軽鎖またはそのサブセット(例えば、可変領域、CDR領域など)、T細胞受容体鎖またはそのサブセット、T細胞受容体鎖またはそのサブセット(例えば、可変領域、CDR3、V(D)J領域など)、CDR(TCRもしくはBCRのいずれかのCDR1、CDR2、もしくはCDR3、またはそのようなCDRの組み合わせを含む)、TCRまたはBCRのいずれかのV(D)J領域、IgH可変領域の超変異領域など。1つの局面において、レパートリーのクロノタイプを規定する核酸セグメントは、それらの多様性(すなわち、セット中の別個の核酸配列の数)が十分に高く、結果として個体中の実質的にすべてのT細胞もしくはB細胞またはそのクローンがそのようなレパートリーの特有の核酸配列を保有するように、選択される。すなわち、本発明に従って、実施者は、T細胞またはB細胞の集団の完全な多様性を反映しない、TCRまたはBCRをコードする組換え核酸の特定のセグメントまたは領域を、クロノタイプを規定するために選択してもよい；しかしながら、好ましくは、クロノタイプは、それらの由来元であるT細胞および/またはB細胞の集団の多様性を反映するように規定される。すなわち、好ましくは、試料のそれぞれ異なるクローンは異なるクロノタイプを有する。(当然ながら、いくつかの適用においては、白血病またはリンパ腫患者由来の試料の場合のように、プロファイル内には

30

40

50

1つまたは複数の特定のクロナタイプの複数のコピーが存在する)。本発明の他の局面において、レパトリーに相当するリンパ球集団は、循環B細胞であってよく、または循環T細胞であってよく、またはCD4+ T細胞、もしくはCD8+ T細胞、もしくは細胞表面マーカーによって規定されるその他の亜集団などを含むがこれらに限定されない、前述の集団のいずれかの亜集団であってよい。そのような亜集団は、特定の組織、例えば骨髄もしくはリンパ節などから試料を取得することによって、または1つもしくは複数の細胞表面マーカー、サイズ、形態などに基づいて試料(末梢血など)から細胞を選別もしくは濃縮することによって獲得され得る。さらなる他の局面において、レパトリーに相当するリンパ球集団は、腫瘍組織、感染組織などの罹患組織に由来し得る。1つの態様において、ヒトTCR鎖またはその断片を含むレパトリーは、 $0.1 \times 10^6 \sim 1.8 \times 10^6$ 種の範囲、または $0.5 \times 10^6 \sim 1.5 \times 10^6$ 種の範囲、または $0.8 \times 10^6 \sim 1.2 \times 10^6$ 種の範囲内の数の別個のヌクレオチド配列を含む。別の態様において、ヒトIgH鎖またはその断片を含むレパトリーは、 $0.1 \times 10^6 \sim 1.8 \times 10^6$ 種の範囲、または $0.5 \times 10^6 \sim 1.5 \times 10^6$ 種の範囲、または $0.8 \times 10^6 \sim 1.2 \times 10^6$ 種の範囲内の数の別個のヌクレオチド配列を含む。特定の態様において、本発明のレパトリーは、IgH鎖のV(D)J領域の実質的にすべてのセグメントをコードするヌクレオチド配列のセットを含む。1つの局面において、本明細書で用いられる「実質的にすべての」とは、0.001パーセントもしくはそれ以上の相対存在量を有するすべてのセグメントを意味し；または別の局面において、本明細書で用いられる「実質的にすべての」とは、0.0001パーセントもしくはそれ以上の相対存在量を有するすべてのセグメントを意味する。別の特定の態様において、本発明のレパトリーは、TCR鎖のV(D)J領域の実質的にすべてのセグメントをコードするヌクレオチド配列のセットを含む。別の態様において、本発明のレパトリーは、25~200ヌクレオチドの範囲の長さを有しかつTCR鎖のV、D、およびJ領域のセグメントを含むヌクレオチド配列のセットを含む。別の態様において、本発明のレパトリーは、25~200ヌクレオチドの範囲の長さを有し、かつIgH鎖のV、D、およびJ領域のセグメントを含むヌクレオチド配列のセットを含む。別の態様において、本発明のレパトリーは、別個のIgH鎖を発現するリンパ球の数と実質的に等しい多数の別個のヌクレオチド配列を含む。別の態様において、本発明のレパトリーは、別個のTCR鎖を発現するリンパ球の数と実質的に等しい多数の別個のヌクレオチド配列を含む。さらなる別の態様において、「実質的に等しい」とは、ヌクレオチド配列のレパトリーが、ある個体のある集団の、0.001パーセントまたはそれ以上の頻度で存在するすべてのリンパ球によって保有または発現されるIgHもしくはTCR鎖またはその一部分をコードするヌクレオチド配列を、99パーセントの確率で含むことを意味する。さらなる別の態様において、「実質的に等しい」とは、ヌクレオチド配列のレパトリーが、0.0001パーセントまたはそれ以上の頻度で存在するすべてのリンパ球によって保有または発現されるIgHもしくはTCR鎖またはその一部分をコードするヌクレオチド配列を、99パーセントの確率で含むことを意味する。前述の2文に記載されるクロナタイプのセットは、本明細書において、IgHおよび/またはTCR鎖の「完全なレパトリー」を表現すると見なされる場合がある。上記のように、クロナタイププロファイル(またはレパトリープロファイル)を測定または生成する場合には、そのようなプロファイルが、特定の適用に対してレパトリーのかなり正確な表現を提供するように、十分に多量のリンパ球の試料を得る。1つの局面において、特に1~10 mLの末梢血試料から得られる場合、 $10^5 \sim 10^7$ 個のリンパ球を含む試料が用いられる。

【0087】

「配列リード」とは、配列決定技法によって生成された一連のデータまたはデータストリームから決定されたヌクレオチドの配列を意味し、この決定は、例えば、この技法と関連した塩基コーリングソフトウェア、例えばDNA配列決定プラットフォームの商業的供給業者からの塩基コーリングソフトウェアを用いてなされる。配列リードは通常、配列内の各ヌクレオチドに関する品質スコアを含む。典型的には、配列リードは、例えばDNAポリメラーゼまたはDNAリガーゼを用いて、鋳型核酸に沿ってプライマーを伸長させることによって作られる。データは、このような伸長に付随する光学的、化学的(例えば、pH変化

10

20

30

40

50

)、または電氣的シグナルなどのシグナルを記録することによって生成される。このような初期データが配列リードに変換される。

【0088】

「配列タグ」(または「タグ」)または「バーコード」とは、ポリヌクレオチドまたは鑄型分子に付着させ、反応または一連の反応において該ポリヌクレオチドまたは鑄型を同定および/または追跡するために用いられるオリゴヌクレオチドを意味する。場合により本明細書において「タグ化ポリヌクレオチド」、または「タグ化鑄型」、または「タグ-ポリヌクレオチドコンジュゲート」、「タグ-分子コンジュゲート」などと称される線状コンジュゲートを形成させるために、配列タグをポリヌクレオチドもしくは鑄型の3'末端もしくは5'末端に付着させることができ、またはこれをそのようなポリヌクレオチドまたは鑄型の内部に挿入することができる。配列タグは、大きさおよび組成が大きく異なってよく；参照により本明細書に組み入れられる以下の参考文献は、特定の態様に適した配列タグのセットを選択するための指針を提供する：Brenner、米国特許第5,635,400号；Brenner and Macevicz、米国特許第7,537,897号；Brenner et al, Proc. Natl. Acad. Sci., 97: 1665-1670 (2000)；Church et al、欧州特許出願公開第0 303 459号；Shoemaker et al, Nature Genetics, 14: 450-456 (1996)；Morris et al、欧州特許出願公開第0799897A1号；Wallace、米国特許第5,981,179号など。配列タグの長さおよび組成は大きく異なってよく、特定の長さおよび/または組成の選択は、例えばハイブリダイゼーション反応によるか、もしくは配列決定のように酵素反応によるなど、リードを生成するためにタグがどのように用いられるか；それらが例えば蛍光色素などによって標識されるかどうか；ポリヌクレオチドのセットなどを明確に同定するために必要な、識別可能なオリゴヌクレオチドタグの数、および例えばクロスハイブリダイゼーションもしくは配列決定エラーによる誤同定がないことなど、信頼できる同定を確実にするために、セットのタグがどれだけ異なるべきかを非限定的に含む、いくつかの要因に依存する。1つの局面において、配列タグはそれぞれ、それぞれ2~36ヌクレオチド、または4~30ヌクレオチド、または8~20ヌクレオチド、または6~10ヌクレオチドの範囲内の長さを有し得る。1つの局面においては、セットの各配列タグが、同じセットの他のすべてのタグのヌクレオチド配列と少なくとも2塩基だけ異なる特有のヌクレオチド配列を有する、配列タグのセットが用いられ；別の局面においては、セットの各タグの配列が、同じセットの他のすべてのタグの配列と少なくとも3塩基だけ異なる、配列タグのセットが用いられる。

10

20

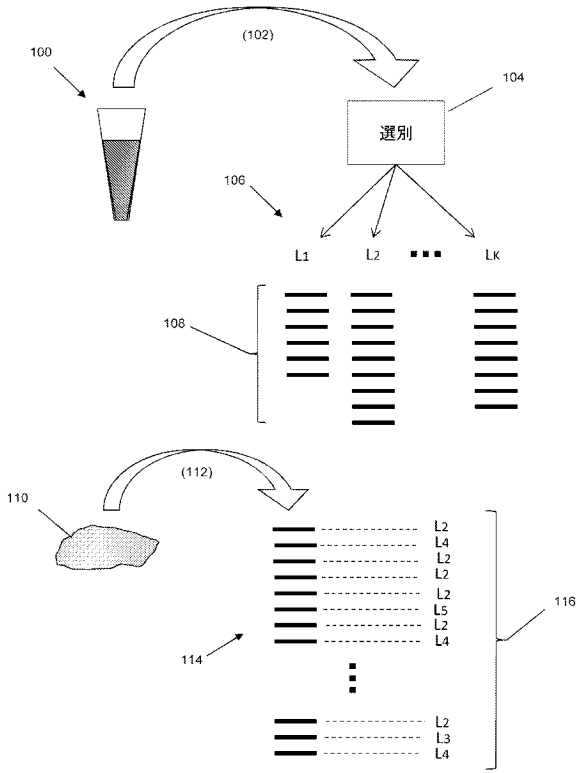
30

【0089】

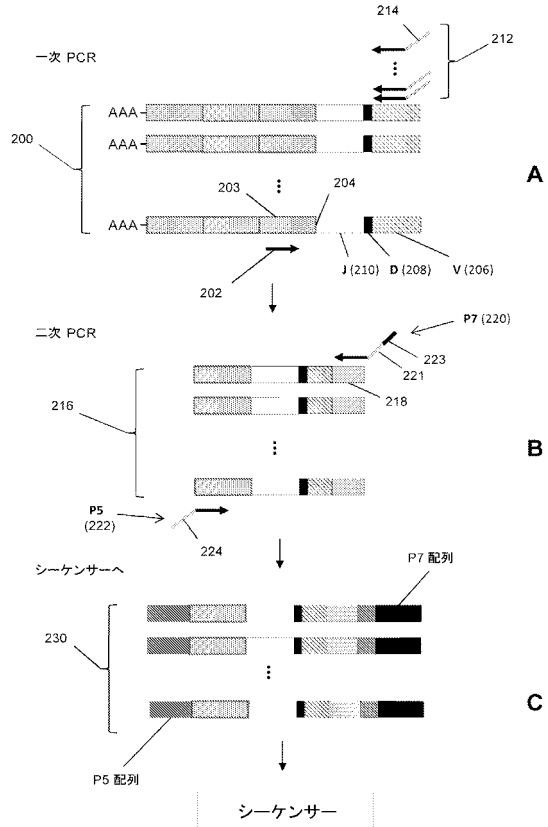
「配列木」とは、ヌクレオチド配列を表すための木データ構造を意味する。1つの局面において、本発明の木データ構造は、サイクルも循環経路も含まないノードおよびエッジを含む根付き有向木である。本発明の木データ構造のノードからのエッジは、通常は順序付けられている。ノードおよび/またはエッジは、値を有し得るかまたは値を伴い得る構造である。木内の各ノードは0個またはそれ以上の子ノードを有し、子ノードは慣例により木内でそれよりも下方に示される。子を有するノードは、子の親ノードと称される。ノードは最大で1つの親を有する。いかなる子も有しないノードは、葉ノードと称される。木内の最上位ノードは、根ノードと称される。最上位ノードであるため、根ノードは親を有しない。木において演算が通常始まるのは、このノードである(しかしながら、いくつかのアルゴリズムは葉ノードから始まり、徐々に上がっていき根で終わる)。エッジまたはリンクをたどることにより、そこから他のすべてのノードに到達することができる。

40

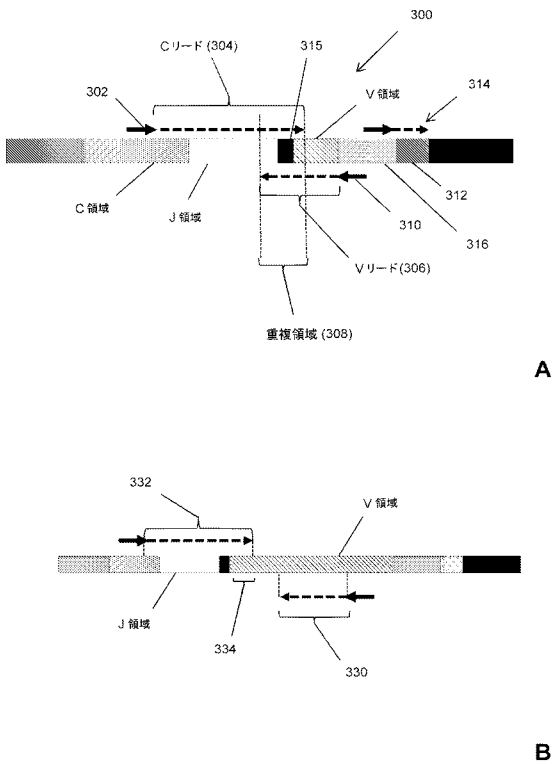
【図1】



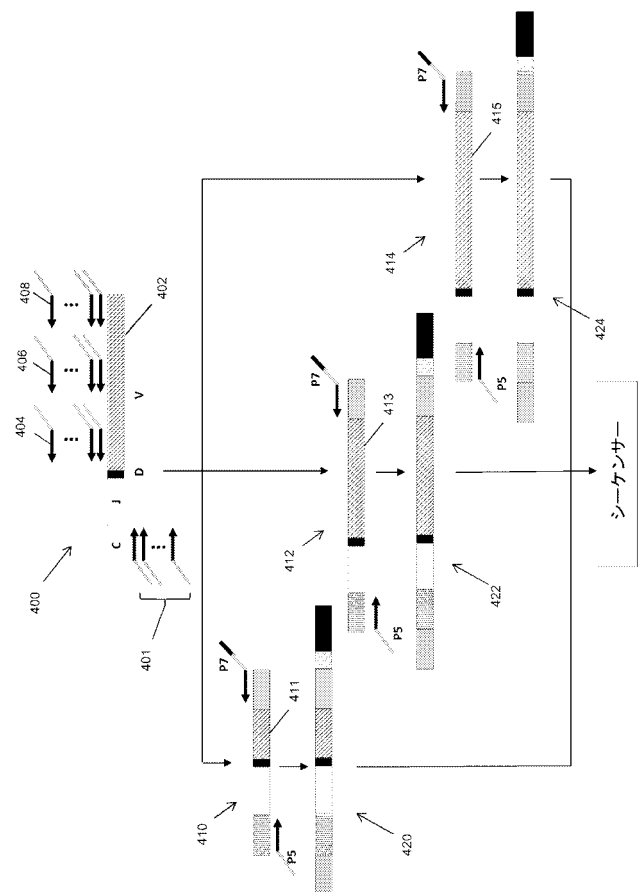
【図2】



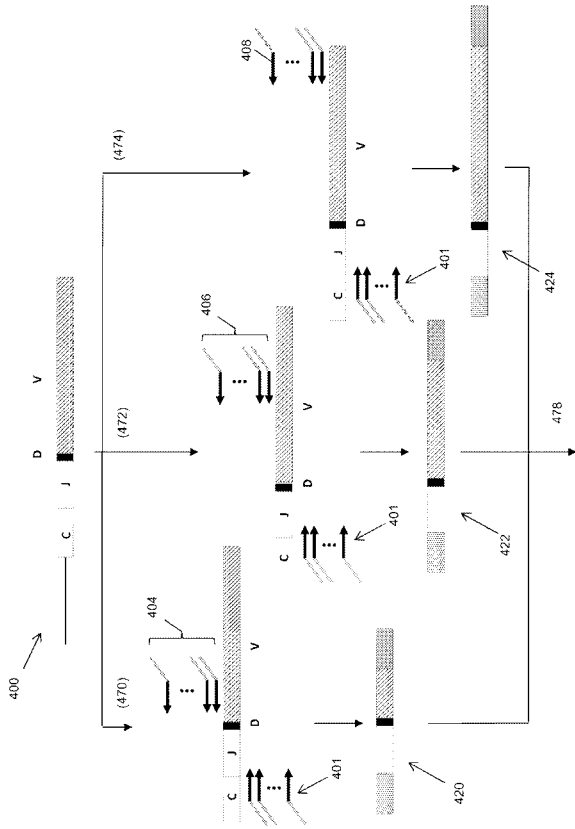
【図3】



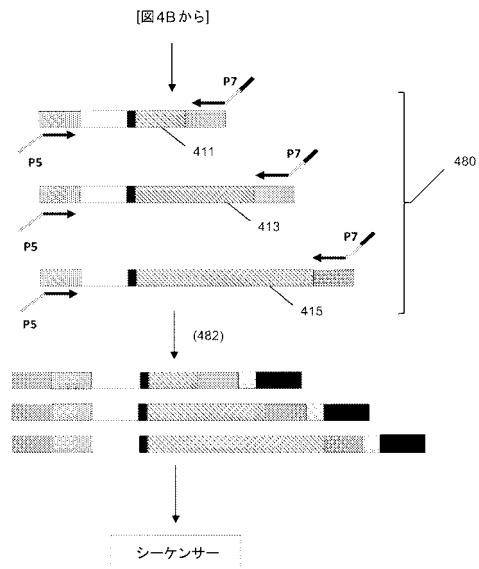
【図4A】



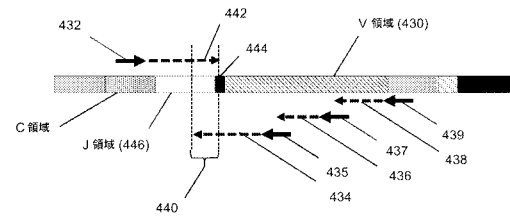
【図 4 B】



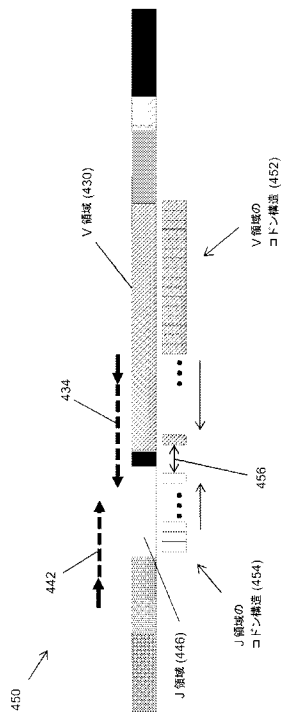
【図 4 C】



【図 4 D】



【図 4 E】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 12/69310
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - C12Q 1/68 (2013.01) USPC - 435/6.11 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) USPC - 435/6.11 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC - 435/6.11, 6.16, 612, 911, 912 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) *** Databases: WEST; Patbase; Google *** Search Terms Used: Sequentia, Faham, Klinger, clonotype, profile, lymphocyte, solid tumor, peripheral blood, FoxP3+, fluorescence-activated cell sorter, FACS, helper, cytotoxic, T cells, infiltrate, infiltrated, infiltration		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011/0207134 A1 (Faham et al.) 25 August 2011 (25.08.2011), especially para [0017], [0043], [0052]-[0055], [0060]-[0065], [0071], [0095], [0121], [0232],	1-29
Y	Jochems et al., 'Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity' (12 April 2011). Experimental Biology and Medicine, Vol. 236, Pg. 567-579; especially pg 570, col 1, para 2; pg 571, col 1, para 1; pg 571, col 2, para 3-4; pg 572, Table 2;	1-29
A	US 2010/0008920 A1 (Schneck et al.) 14 January 2010 (14.01.2010), entire document	1-29
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "G" document member of the same patent family		
Date of the actual completion of the international search 08 February 2013 (08.02.2013)		Date of mailing of the international search report 26 FEB 2013
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

フロントページの続き

(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, 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, 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, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC

(74)代理人 100148699
弁理士 佐藤 利光

(74)代理人 100128048
弁理士 新見 浩一

(74)代理人 100129506
弁理士 小林 智彦

(74)代理人 100130845
弁理士 渡邊 伸一

(74)代理人 100114340
弁理士 大関 雅人

(74)代理人 100114889
弁理士 五十嵐 義弘

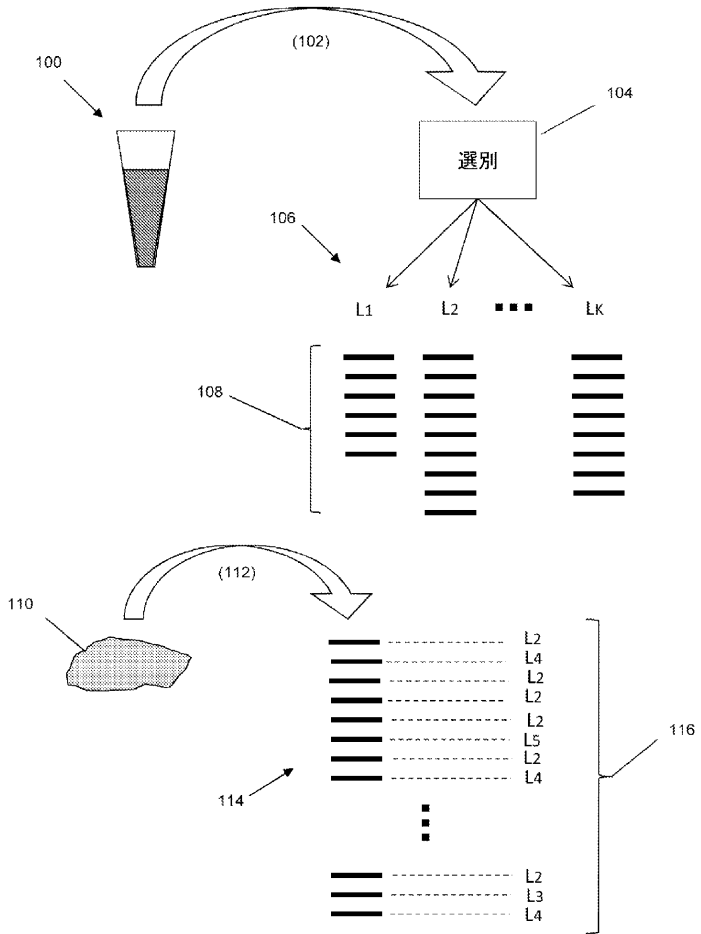
(74)代理人 100121072
弁理士 川本 和弥

(72)発明者 ファハム マレク
アメリカ合衆国 カリフォルニア州 サウス サンフランシスコ イースト ジェイミー コート
400 スイート 301

(72)発明者 クリンガー マーク
アメリカ合衆国 カリフォルニア州 サウス サンフランシスコ イースト ジェイミー コート
400 スイート 301

Fターム(参考) 4B024 AA12 CA09 HA12
4B063 QA01 QA19 QQ08 QQ43 QR55 QR62 QS25 QS34 QX02

【要約の続き】



专利名称(译)	检测和测量组织浸润淋巴细胞的方法		
公开(公告)号	JP2015501658A	公开(公告)日	2015-01-19
申请号	JP2014547393	申请日	2012-12-12
[标]申请(专利权)人(译)	SEQUENTA		
申请(专利权)人(译)	Shikenta公司		
[标]发明人	ファハムマレク クリンガーマーク		
发明人	ファハム マレク クリンガー マーク		
IPC分类号	C12Q1/68 C12N15/09 G01N33/543 G01N33/53		
CPC分类号	C12Q1/6881 C12Q1/06 G01N33/5091 G01N33/5094		
FI分类号	C12Q1/68.A C12N15/00.A G01N33/543.597 G01N33/53.Y		
F-TERM分类号	4B024/AA12 4B024/CA09 4B024/HA12 4B063/QA01 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B063/QX02		
代理人(译)	清水初衷 井上隆一 佐藤俊光 小林智彦 渡边真一 正人大关 五十嵐弘		
优先权	61/570192 2011-12-13 US		
其他公开文献	JP6343563B2 JP2015501658A5		
外部链接	Espacenet		

摘要(译)

本发明提供确定渗入实体组织例如肿瘤或患有自身免疫疾病的组织的细胞例如淋巴细胞的数量，水平和/或比率的方法，以及这样的方法。一种基于测量结果进行患者预后的方法。一方面，本发明的方法包括以下步骤：将得自可及组织例如外周血的淋巴细胞分类为功能亚组，例如细胞毒性T细胞和调节性T细胞，以及包括产生克隆型谱的步骤。对难以进入的受疾病影响的组织进行采样，并生成一种或多种克隆型谱。根据后者的克隆型谱，患病组织中每个功能亚群中的淋巴细胞水平由其克隆型决定，这些克隆型可从容易获得的组织中分为几类。它是从淋巴细胞中鉴定出来的。

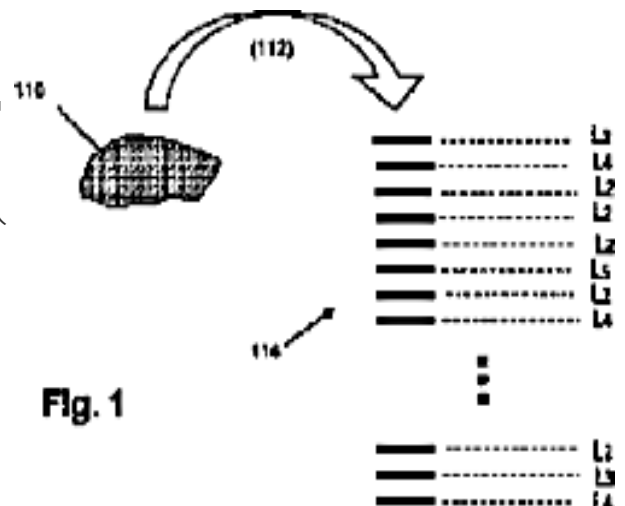


Fig. 1