

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6460343号
(P6460343)

(45) 発行日 平成31年1月30日 (2019. 1. 30)

(24) 登録日 平成31年1月11日 (2019. 1. 11)

(51) Int. Cl.

F I

C 1 2 Q 1/6844 (2018. 01)
C 4 O B 40/08 (2006. 01)C 1 2 Q 1/6844 Z N A Z
C 4 O B 40/08

請求項の数 9 (全 39 頁)

(21) 出願番号 特願2015-557201 (P2015-557201)
 (86) (22) 出願日 平成26年2月11日 (2014. 2. 11)
 (65) 公表番号 特表2016-506750 (P2016-506750A)
 (43) 公表日 平成28年3月7日 (2016. 3. 7)
 (86) 国際出願番号 PCT/US2014/015841
 (87) 国際公開番号 W02014/124451
 (87) 国際公開日 平成26年8月14日 (2014. 8. 14)
 審査請求日 平成29年1月27日 (2017. 1. 27)
 (31) 優先権主張番号 61/763, 451
 (32) 優先日 平成25年2月11日 (2013. 2. 11)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 517311482
 アイレパトリリー インコーポレイテッド
 アメリカ合衆国 35806 アラバマ州
 ハンツビル ゲノム ウェイ 601
 スイート 3005
 (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】 免疫レパトリリーを評価するための方法

(57) 【特許請求の範囲】

【請求項 1】

ヒトまたは動物対象の免疫プロファイルを、正常と同定および特徴づける、または、特定の疾患と関連付けられた細胞のタイプおよび数の存在を特徴とすると同定および特徴づけるために、半定量的配列情報を解析する方法であって、以下の工程を含む方法：

(a) 対象の免疫プロファイルとデータベース中に保存された疾患ライブラリー由来の累積的な免疫プロファイルとの間で共有される1つまたは複数の別個の相補性決定領域3 (CD R3) 配列を同定する工程；

(b) 該共有される別個のCDR3配列に対応する、対象の検出された配列の総数を合計する工程；

(c) 対象の免疫プロファイルと疾患ライブラリーとの間で共有されるそれらの別個のCDR3を表す、検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、1つまたは複数のオリジナル・シェアリング・インデックス (original sharing indice) を作成する工程；

(d) データベース中に保存されたパブリックライブラリーから配列を無作為に選択してサブライブラリーを形成する工程であって、該サブライブラリーが、疾患ライブラリー中の別個のCDR3配列の数にほぼ等しい複数の別個のCDR3配列を含む、工程；

(e) 対象の免疫プロファイルとサブライブラリーとの間で共有される1つまたは複数の別個のCDR3配列を同定する工程；

(f) 該共有されるCDR3配列に対応する検出された配列の総数を合計し、かつ対象の免疫

プロファイルとサブライブラリーとの間で共有される検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、サンプリング・シェアリング・インデックス (sampling sharing index) を作成する工程；

(g)工程(d)～(f)を、対象の免疫プロファイルを、正常と同定および特徴づける、または、特定の疾患と関連付けられた細胞のタイプおよび数の存在を特徴とすると同定および特徴づけるための、統計的に有意な結果をもたらすのに十分な回数繰り返す工程；および

(h)サンプリング・シェアリング・インデックスが、対象の免疫プロファイルと疾患ライブラリーとの間のオリジナル・シェアリング・インデックスに等しいかまたはそれを超える回数の割合としてP値を推定する工程。

【請求項 2】

工程(g)が、工程(d)～(f)を少なくとも1000回またはそれ以上繰り返すことを含む、請求項1記載の方法。

【請求項 3】

該対象からの別個のCDR3の半定量的増幅および配列決定が工程(a)に先行し、該半定量的増幅および配列決定が以下の工程を含む、請求項1または請求項2記載の方法：

(i)対象から単離された白血球の亜集団から、RNAを単離する工程；

(ii)ネステッドプライマーを用いてアンプリコンを生成するために第1増幅反応においてRT-PCRを用いてRNAを増幅する工程であって、ネステッドプライマーの少なくとも一部分が、結果として生じるアンプリコン中に共用プライマーに対する結合部位を組み込むための追加のヌクレオチドを含む、工程；

(iii)第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離する工程；

(iv)第2増幅反応において共用プライマーを添加することによって、共用プライマーに対する少なくとも1つの結合部位を有する第1増幅反応のアンプリコンを増幅する工程であって、第2増幅反応の産物が別個のCDR3を含むポリヌクレオチドである工程；ならびに

(v)細胞の亜集団における抗体および/または受容体の再構成を同定するために、第2増幅反応のアンプリコンを配列決定する工程；

ここで、工程(iv)および(v)の増幅および配列決定が、該対象からの別個のCDR3についての半定量的配列情報をもたらす、方法。

【請求項 4】

白血球の亜集団の単離が、フローサイトメトリーによって行われる、請求項3記載の方法。

【請求項 5】

白血球の亜集団がT細胞を含む、請求項3記載の方法。

【請求項 6】

T細胞が、ナイーブT細胞、成熟T細胞および記憶T細胞からなる群より選択される、請求項5記載の方法。

【請求項 7】

白血球の亜集団がB細胞を含む、請求項3記載の方法。

【請求項 8】

B細胞が、ナイーブB細胞、成熟B細胞および記憶B細胞からなる群より選択される、請求項7記載の方法。

【請求項 9】

細胞の亜集団における再構成が、B細胞免疫グロブリン重鎖 (IgH)、B細胞 軽鎖、T細胞受容体 、T細胞 およびT細胞 の再構成からなる群より選択される、請求項3記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

配列表

10

20

30

40

50

本出願は、ASCIIフォーマットで電子的に提出された配列表を含み、参照によりその全体が本明細書に組み入れられる。2014年2月5日に作成された、該ASCIIコピーは、15892-0005_SL.txtという名称であり、サイズは93,776バイトである。

【0002】

関連出願の相互参照

本出願は、「Method for Evaluating an Immunorepertoire」と題される、2013年2月11日に出願された、米国仮出願第61/763,341号の継続出願であり、その優先権を主張し、これは参照により本明細書に組み入れられる。

【0003】

発明の分野

本発明は、細胞の集団中のT細胞受容体抗体を同定するための方法、およびその情報を用いて患者の免疫状態を測定し、患者がどの疾患を有し得る可能性があるかを予測するための方法に関する。

【背景技術】

【0004】

発明の背景

科学者は、ある疾患が特定の遺伝子または遺伝子突然変異と関連付けられることを長年にわたって知っている。しかし、遺伝的因果関係は、ヒトにおいて診断される疾患の一部のみを占める。多くの疾患が、感染および環境要因に対する免疫系の反応と何らかの関係があるようであるが、免疫系が、癌、アルツハイマー病、肋軟骨炎、線維筋痛症、狼瘡、および他の疾患などの疾患においてどのような役割を果たすかは、依然として不明のままである。

【0005】

ヒトゲノムは、7つの主要座に局在する、一倍体ゲノム当たり総数567~588個のIG（免疫グロブリン）ならびにTR（T細胞受容体）遺伝子（339~354個のIGおよび228~234個のTR）を含む。それらは405~418個のV、32個のD、105~109個のJおよび25~29個のC遺伝子を含む。機能的IGおよびTR遺伝子の数は一倍体ゲノム当たり321~353個である。それらは187~216個のV、28個のD、86~88個のJおよび20~21個のC遺伝子を含む（<http://imgt.cines.fr>）。これらの遺伝子の再構成によって、推定 2.5×10^7 個の可能性のある抗体またはT細胞受容体が作製され得る。

【0006】

現在まで、わずかな疾患しか、共通抗原（Prinz, J. et al., Eur. J. Immunol. (1999) 29(10): 3360-3368, 「Selection of Conserved TCR VDJ Rearrangements in Chronic Psoriatic Plaques Indicates a Common Antigen in Psoriasis Vulgaris」（非特許文献1））および/または特定のVDJ再構成（Tamaru, J. et al., Blood (1994) 84(3): 708-715, 「Hodgkin's Disease with a B-cell Phenotype Often Shows a VDJ Rearrangement and Somatic Mutations in the VH Genes」（非特許文献2））に対する身体の反応と関連付けられていない。必要とされるのは、ヒト免疫反応細胞における変化を評価し、それらの変化と特定の疾患とを関連付けるためのよりよい方法である。

【先行技術文献】

【非特許文献】

【0007】

【非特許文献1】Prinz, J. et al., Eur. J. Immunol. (1999) 29(10): 3360-3368, 「Selection of Conserved TCR VDJ Rearrangements in Chronic Psoriatic Plaques Indicates a Common Antigen in Psoriasis Vulgaris」

【非特許文献2】Tamaru, J. et al., Blood (1994) 84(3): 708-715, 「Hodgkin's Disease with a B-cell Phenotype Often Shows a VDJ Rearrangement and Somatic Mutations in the VH Genes」

【発明の概要】

【0008】

本発明は、免疫反応細胞集団における変化を評価し、それらの変化と特定の疾患とを関連付けるための方法に関する。本発明の一面において、方法は、(a)少なくとも1つのヒトまたは動物対象から白血球の亜集団を単離する工程、(b)細胞の亜集団からRNAを単離する工程、(c)ネステッドプライマーを用いてアンプリコンを生成するために第1増幅反応においてRT-PCRを用いてRNAを増幅する工程であって、ネステッドプライマーの少なくとも一部分が、結果として生じるアンプリコン中に共用プライマーに対する結合部位を組み込むための追加のヌクレオチドを含む、工程、(d)第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離する工程、(e)第2増幅反応において共用プライマーを添加することによって、共用プライマーに対する少なくとも1つの結合部位を有する第1増幅反応のアンプリコンを増幅する工程、ならびに(f)細胞の亜集団における抗体および/または受容体の再構成を同定するために、第2増幅反応のアンプリコンを配列決定する工程を含む。一態様において、亜集団は、全血集団または別の混合集団のサンプルを含み得る。

10

【0009】

一態様において、白血球の亜集団を単離する工程を、ナイーブB細胞、成熟B細胞、記憶B細胞、ナイーブT細胞、成熟T細胞、および記憶T細胞を分離するために、フローサイトメトリーによって行ってもよい。方法の様々な態様において、細胞の亜集団における組換えは、B細胞免疫グロブリン重鎖(IgH)、および/または軽鎖(IgK、IgL)、T細胞受容体、の再構成である。

【0010】

20

本発明の別の局面において、方法は、以下を含む追加の工程を任意で含み得る：(g)ワクチンが投与された個体の集団について同定された再構成と、ワクチンが投与されなかった個体の集団について同定された再構成とを比較し、免疫反応を生じさせることにおけるワクチンの有効性を評価する工程。

【0011】

方法はまた、以下の追加の工程を任意で含み得る：(g)正常な個体の集団について同定された再構成と、疾患を有すると診断された個体の集団について同定された再構成とを比較して、特定の再構成または再構成のセットと疾患との間に相関関係があるかどうかを決定する工程。

【0012】

30

様々な局面において、方法は、T細胞またはB細胞内の遺伝子再構成から生じ、特定の抗原についての抗体および/またはT細胞受容体の親和性および特異性の原因である相補性決定領域3(CDR3)を含むポリヌクレオチドの半定量的増幅をもたらすことができる。半定量的増幅は、特定のCDR3配列の存在を検出するだけでなく、それらのCDR3配列を生成するために必要な組換え事象をもたらした細胞の相対存在量を決定するための方法を提供する。

【0013】

本発明の一面は、従って、半定量的配列情報を解析してヒトまたは動物についての1つまたは複数の免疫状態レポートを提供するための方法に関する。免疫状態レポートを与えるための方法は、以下の工程を含む：(a)対象の免疫プロファイルとデータベース中に保存された疾患ライブラリー由来の累積的な免疫プロファイルとの間で共有される1つまたは複数の別個のCDR3配列を同定し、該共有される別個のCDR3配列に対応する、対象の検出された配列の総数を合計し、そして対象の免疫プロファイルと疾患ライブラリーとの間で共有されるそれらの別個のCDR3を表す、検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、1つまたは複数のオリジナル・シェアリング・インデックス(original sharing indice)を作成する工程；(b)データベース中に保存されたパブリックライブラリーから配列を無作為に選択して、疾患ライブラリー中の別個のCDR3配列の数にほぼ等しい複数の配列を含むサブライブラリーを形成し、対象の免疫プロファイルとサブライブラリーとの間で共有される1つまたは複数の別個のCDR3配列を同定し、該共有されるCDR3配列に対応する検出された配列の総数を合計し、そして対象の免疫プ

40

50

ロファイルとサブライブラリーとの間で共有される検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、サンプリング・シェアリング・インデックス (sampling sharing index) を作成する工程；(c)工程(b)を少なくとも1000回またはそれ以上繰り返す工程；および(d)サンプリング・シェアリング・インデックスが、患者の免疫プロファイルと疾患ライブラリーとの間のオリジナル・シェアリング・インデックスに等しいかまたはそれを超える回数の割合としてP値を推定する工程。

【 0 0 1 4 】

[本発明1001]

免疫反応細胞集団における変化を評価し、それらの変化と特定の疾患とを関連付けるための方法であって、以下の工程を含む方法：

(a)少なくとも1つのヒトまたは動物対象から白血球の亜集団を単離する工程；

(b)細胞の該亜集団からRNAを単離する工程；

(c)ネステッドプライマーを用いてアンプリコンを生成するために第1増幅反応においてRT-PCRを用いてRNAを増幅する工程であって、ネステッドプライマーの少なくとも一部分が、結果として生じるアンプリコン中に共用プライマーに対する結合部位を組み込むための追加のヌクレオチドを含む、工程；

(d)第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離する工程；

(e)第2増幅反応において共用プライマーを添加することによって、共用プライマーに対する少なくとも1つの結合部位を有する第1増幅反応のアンプリコンを増幅する工程；なら

びに
(f)細胞の亜集団における抗体および/または受容体の再構成を同定するために、第2増幅反応のアンプリコンを配列決定する工程。

[本発明1002]

第2増幅反応の産物が、相補性決定領域3 (CDR3) を含むポリヌクレオチドである、本発明1001の方法。

[本発明1003]

白血球の亜集団を単離する工程が、フローサイトメトリーによって行われる、本発明1001の方法。

[本発明1004]

白血球の亜集団がT細胞を含む、本発明1001の方法。

[本発明1005]

T細胞が、ナイーブT細胞、成熟T細胞および記憶T細胞からなる群より選択される、本発明1004の方法。

[本発明1006]

白血球の亜集団がB細胞を含む、本発明1001の方法。

[本発明1007]

B細胞が、ナイーブB細胞、成熟B細胞および記憶B細胞からなる群より選択される、本発明1006の方法。

[本発明1008]

細胞の亜集団における再構成が、B細胞免疫グロブリン重鎖 (IgH)、B細胞 軽鎖、T細胞受容体、T細胞 およびT細胞 の再構成からなる群より選択される、本発明1001の方法。

[本発明1009]

(g)ワクチンが投与された個体の集団について同定された再構成と、ワクチンが投与されなかった個体の集団について同定された再構成とを比較する工程；および

(h)免疫反応を生じさせることにおけるワクチンの有効性を評価する工程をさらに含む、本発明1001の方法。

[本発明1010]

(g)正常な個体の集団について同定された再構成と、疾患を有すると診断された個体の

10

20

30

40

50

集団について同定された再構成とを比較する工程；

(h)特定の再構成または再構成のセットと疾患との間に相関関係があるかどうかを決定する工程

をさらに含む、本発明1001の方法。

[本発明1011]

ヒトまたは動物についての1つまたは複数の免疫状態レポートを提供するために、半定量的配列情報を解析する方法であって、以下の工程を含む方法：

(a)対象の免疫プロファイルとデータベース中に保存された疾患ライブラリー由来の累積的な免疫プロファイルとの間で共有される1つまたは複数の別個のCDR3配列を同定する工程；

(b)該共有される別個のCDR3配列に対応する、対象の検出された配列の総数を合計する工程；

(c)対象の免疫プロファイルと疾患ライブラリーとの間で共有されるそれらの別個のCDR3を表す、検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、1つまたは複数のオリジナル・シェアリング・インデックス(original sharing indice)を作成する工程；

(d)データベース中に保存されたパブリックライブラリーから配列を無作為に選択してサブライブラリーを形成する工程であって、該サブライブラリーが、疾患ライブラリー中の別個のCDR3配列の数にほぼ等しい複数の別個のCDR3配列を含む、工程；

(e)対象の免疫プロファイルとサブライブラリーとの間で共有される1つまたは複数の別個のCDR3配列を同定する工程；

(f)該共有されるCDR3配列に対応する検出された配列の総数を合計し、かつ対象の免疫プロファイルとサブライブラリーとの間で共有される検出された配列の総数の対象の免疫プロファイルにおけるパーセンテージを計算して、サンプリング・シェアリング・インデックス(sampling sharing index)を作成する工程；

(g)工程(d)～(f)を少なくとも1000回またはそれ以上繰り返す工程；および

(h)サンプリング・シェアリング・インデックスが、患者の免疫プロファイルと疾患ライブラリーとの間のオリジナル・シェアリング・インデックスに等しいかまたはそれを超える回数の割合としてP値を推定する工程。

[本発明1012]

個人の免疫レポートリーのデータベースを開発するための方法であって、以下の工程を含む方法：

(a)1つまたは複数の個体に由来する白血球の亜集団由来の1つまたは複数のRNAを増幅および配列決定する工程；

(b)配列をデータベースへ入力して、コンピューター、サーバー、または他の電子記憶装置上に保存され得るデータを提供する工程；

(c)1つまたは複数のRNAの配列に対応する個体についての識別情報および特徴を、コンピューター、サーバー、または他の電子記憶装置上に保存され得るデータとして入力する工程、ならびに

(d)1つまたは複数の個体について工程(b)および工程(e)のデータを評価して、1つまたは複数のRNA配列と配列に対応する個体の1つまたは複数の特徴との間に相関関係が存在するかどうかを決定する工程。

[本発明1013]

識別情報が、患者識別番号、患者のHLA型を含むコード、下されている場合がある1つまたは複数の臨床診断を含む疾患コード、サンプルの日付を含む「ステージング・コード」、RNAが増幅および配列決定された細胞亜集団の型を含む細胞型コード、ならびにサンプルについて同定された配列を含む1つまたは複数の配列コードからなる群より選択される、本発明1012の方法。

[本発明1014]

白血球の亜集団がT細胞を含む、本発明1012の方法。

[本発明1015]

T細胞が、ナイーブT細胞、成熟T細胞および記憶T細胞からなる群より選択される、本発明1014の方法。

[本発明1016]

白血球の亜集団がB細胞を含む、本発明1012の方法。

[本発明1017]

B細胞が、ナイーブB細胞、成熟B細胞および記憶B細胞からなる群より選択される、本発明1016の方法。

本開示は、下記の図面を参照してよりよく理解することができる。図面の要素は、必ずしも互いに一定の縮尺ではなく、代わりに、本開示の原理を明確に説明することを目的として強調されている。さらに、同様の参照番号は、いくつかの図にわたって対応する部分を示す。

【図面の簡単な説明】**【0015】**

【図1】図1aおよび1bは、本明細書に開示されるプライマーを用いて本発明の方法によって得られた増幅産物の存在を示すゲルの写真である。

【図2】図2aおよび2bは、疾患を有する個体および全般的に健康である個体における免疫プロファイル間の観察された多様性の差異を示す図であり、各黒丸は別個のCDR3配列を示し、丸のサイズは、別個のCDR3配列が免疫プロファイルにおいて見出される回数を示す。

【図3】パブリックライブラリーを作成するための方法を示す図である。

【図4】疾患ライブラリーを作成するための方法を示す図である。

【図5】患者免疫プロファイルと疾患ライブラリーとを比較し、これら2つに共有される患者免疫プロファイル中の各別個のCDR3についてのパーセンテージを計算し、そしてそれらのパーセンテージを合算して合計またはシェアリング・インデックスを作ることによって得られた結果を示す。

【図6】患者免疫プロファイルとパブリックライブラリーのサブセットとを比較し、これら2つに共有される各別個のCDR3についてのパーセンテージを計算し、そして患者免疫プロファイル中のそれらのパーセンテージを合算して合計またはシェアリング・インデックスを作ることによって得られた結果を示す。

【図7】本発明の方法を示すグラフであり、ここで、曲線下面積は、パブリックライブラリーのサブセット（サブライブラリー）について得られたトータル・シェアリング・インデックス（total sharing indice）を示し、P値が推定され、個体の免疫プロファイルおよび1つまたは複数の疾患ライブラリーの比較についてのシェアリング・インデックスは、垂線（DL₁、DL₂など）によって示される。

【発明を実施するための形態】**【0016】**詳細な説明

本発明者らは、多数の細胞からの抗体およびT細胞受容体の再構成を評価するための方法を開発し、この方法は、個体の集団において同定された再構成を比較して、特定の再構成または再構成のセットと疾患または疾患のある症状との間に相関関係があるかどうかを決定することについて有用である。この方法はまた、感染および/または環境要因に応答しての個体（単数または複数）の免疫反応の履歴を確立することについて、ならびにワクチンの有効性を評価することについて有用である。

【0017】

本発明は、免疫反応細胞集団における変化を評価し、それらの変化と特定の疾患とを関連付けるための方法に関する。本発明の一面において、方法は、(a)少なくとも1つのヒトまたは動物対象から白血球の亜集団を単離する工程、(b)細胞の亜集団からRNAを単離する工程、(c)ネステッドプライマーを用いてアンプリコンを生成するために第1増幅反応においてRT-PCRを用いてRNAを増幅する工程であって、ネステッドプライマーの少なくとも一部分が、結果として生じるアンプリコン中に共用プライマーに対する結合部位を組み込

むための追加のヌクレオチドを含む、工程、(d)第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離する工程、(e)第2増幅反応において共用プライマーを添加することによって、共用プライマーに対する少なくとも1つの結合部位を有する第1増幅反応のアンプリコンを増幅する工程、ならびに(f)細胞の亜集団における抗体および/または受容体の再構成を同定するために、第2増幅反応のアンプリコンを配列決定する工程を含む。一態様において、亜集団は、全血集団または別の混合集団サンプルを含み得る。

【0018】

一態様において、末梢血サンプルを患者から採取し、白血球の亜集団を単離する工程を、ナイーブB細胞、成熟B細胞、記憶B細胞、ナイーブT細胞、成熟T細胞、および記憶T細胞を分離するために、フローサイトメトリーによって行ってもよい。細胞の亜集団における組換えは、B細胞免疫グロブリン重鎖(IgH)、および/または軽鎖(IgK、IgL)、T細胞受容体、またはの再構成である。

【0019】

本発明の第2局面において、方法は、以下の追加の工程を含み得る：(g)正常な個体の集団について同定された再構成と、疾患を有すると診断された個体の集団について同定された再構成とを比較して、特定の再構成または再構成のセットと疾患との間に相関関係があるかどうかを決定する工程。

【0020】

本発明の別の局面において、方法は、以下を含む追加の工程を含み得る：(g)ワクチンが投与された個体の集団について同定された再構成と、ワクチンが投与されなかった個体の集団について同定された再構成とを比較し、免疫反応を生じさせることにおけるワクチンの有効性を評価する工程。

【0021】

ある態様において、第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離する工程を省略してもよく、2つの増幅反応を同じ反応チューブにおいて行ってもよい。

【0022】

本発明者は、公開番号WO2005/038039(この開示は参照によりその全体が本明細書に組み入れられる)に記載した、tem-PCRとして公知のPCR法を以前開発した。より最近になって、本発明者は、米国仮特許出願第61/042,259号(この開示は参照によりその全体が本明細書に組み入れられる)に記載した、arm-PCRと呼ばれる方法を開発した。さらに、サンプル中の標的ポリヌクレオチドを検出するための装置を記載し、この装置は以下を含む：ネステッドプライマーを用いてアンプリコンを生成するために1つまたは複数の標的ポリヌクレオチドを増幅するためのサーモサイクリング用の第1増幅チャンバであって、ネステッドプライマーの少なくとも一部分が、結果として生じるアンプリコン中に共用プライマーに対する結合部位を組み込むための追加のヌクレオチドを含む、第1増幅チャンバ；第1増幅反応由来のアンプリコンと第1増幅反応由来の1つまたは複数の未使用プライマーとを分離するための手段；および、第2増幅反応において共用プライマーを添加することによって、第1増幅反応の間に生成された1つまたは複数のアンプリコンを増幅するためのサーモサイクリング用の第2増幅チャンバであって、第1増幅反応のアンプリコンが、少なくとも1つの共用プライマーに対する少なくとも1つの結合部位を有する、第2増幅チャンバ。

【0023】

廃棄物リザーバおよび第2 PCRチャンバの両方に流体接続された第1 PCRチャンバを含むPCRチップも記載し、廃棄物リザーバおよび第2 PCRチャンバは各々、さらに少なくとも1つの電極を含み、電極は、第1 PCRチャンバから生成されたアンプリコンを分離するための手段を含む。第2 PCRチャンバは、ハイブリダイゼーションおよび検出チャンバに流体接続されており、ハイブリダイゼーションおよび検出チャンバは、ビーズの物理的な位置が、チップによって分析されるサンプル中の特定の標的ポリヌクレオチドの存在の指標と

10

20

30

40

50

なるように配置された、マイクロスフィアまたはビーズを含む。

【0024】

tem-PCR法、および特にarm-PCR法は、1つの反応で多数のポリヌクレオチドの半定量的増幅を提供する。さらに、arm-PCRは上乘せされた感度を提供する。両方とも、1つの反応で多数のポリヌクレオチドを増幅する能力を提供し、これは、例えば、様々なT細胞およびB細胞のレパートリーが非常に大きいために、本発明の方法において有利である。増幅反応における共用プライマー結合部位の添加、および共用プライマーを用いた標的分子のその後の増幅は、定量的、または半定量的結果を提供する - 患者血液サンプル中の様々な再構成を含む細胞の相対量を決定することを可能にする。抗原の認識に起因するクローン増殖は、その抗原を認識する、より大きな集団の細胞をもたらす、それらの相対数によって細胞を評価することは、抗原曝露が抗体産生B細胞または受容体保有T細胞の増殖に影響を与えたかどうかを決定するための方法を提供する。これは、例えば、特定の疾患を有すると診断された個体において、優勢である特定の細胞の集団が存在し得るかどうかを評価するために有用であり、また、ワクチンが与えられた個体においてワクチンが所望の免疫反応を達成したか否かを評価するために特に有用であり得る。

【0025】

いくつかの市販のハイスループットシーケンシング技術、例えば、Roche Life Sciencesの454シーケンシングがある。454シーケンシング法において、454Aおよび454Bプライマーが、PCR中にPCR産物上へ連結されるかまたはPCR反応後にライゲーションされる。tem-PCRまたはarm-PCRと協力して行われる場合、454Aおよび454Bプライマーは増幅反応において共用プライマーとして使用され得る。PCR産物、通常、異なる配列の混合物は、1 μ l当たり約200コピーへ希釈される。「エマルジョンPCR」反応において（半固体ゲル様環境）、希釈されたPCR産物は、マイクロビーズの表面上においてプライマー（454Aまたは454B）によって増幅される。PCRテンプレートは非常に希薄であるので、通常、たった1つのビーズが1つのテンプレートに隣接しており、半固体環境中に閉じ込められ、増幅はビーズ上および周囲においてのみ生じる。次いで、ビーズが溶出され、特別に設計されたウェルを有するプレート上へ置かれる。各ウェルは1つのビーズのみを保持することができる。次いで、試薬がウェルへ添加され、パイロシーケンシングが行われる。光ファイバー検出器が、各ウェルからのシーケンシング反応を読み取るために使用され得、データがコンピューターによって並行して収集される。1つのそのようなハイスループット反応が、6000万個までの読み取り（6000万個のビーズ）を生じさせることができ、各読み取りは、約300bp配列を生じさせることができる。

【0026】

本発明の一面は、「個人の免疫レパートリー」または免疫プロファイルのデータベースの開発を含み、その結果、各個体は、ベースラインを確立し、長年にわたって、既知および未知の両方の抗原に対する免疫反応の発達を追跡できる。この情報は、情報が多数の個体から集められる場合、特に、癌および心臓疾患などの疾患の発症に関して、価値ある情報をもたらすであろう疫学的データベースを提供する可能性があり、これらの疾患は、形質転換細胞またはウイルスもしくは他の感染要因への曝露から生じることが多いと考えられており、それらの多くはまだ確認されていない。本発明の方法についての1つの特に重要な使用は、感染症、環境要因、またはワクチンが自閉症の原因であり得るかどうかを決定するための小児の評価を含む。例えば、ワクチン投与は自閉症の発症を誘発し得ると多くの人が仮定している。しかし、多くの人はまた、その可能性のある相関関係を、ワクチン中のチメロサルなどの薬剤の使用によるものとし、研究は、チメロサルはその疾患の原因因子ではないようであることを実証した。カクテルワクチンの開発が、自閉症の症例数の増加と相関関係があるという推測が依然として存在するが、データを集めて多数の抗原についての潜在的な因果関係を評価することは、極めて困難である。本発明の方法は、そのプロセスを簡略化し、自閉症および他の疾患のよりよい理解のための重要な情報を提供し得、ここで、異なる個体の免疫反応は、因子または因子の群に曝露された一部の個体における疾患のディファレンシャルな発症についての説明を提供し得、一方、同様に

曝露された他の人はその疾患を発症しない。

【0027】

感染によって誘発される、免疫プロファイルの不均衡は、癌、白血病、神経疾患（アルツハイマー病、多発性硬化症、パーキンソン病、自閉症など）、自己免疫疾患、および代謝疾患を含む、多くの疾患に至り得る。これらの疾患は、免疫プロファイル疾患と呼ばれ得る。次の2つの免疫プロファイル疾患形態が存在し得る：(1)「機能喪失」形態、および(2)「機能獲得」形態。「機能喪失」形態において、人は疾患にかかりやすく、何故ならば、制限されたおよび/または限定された免疫プロファイルは、最も効果的かつ必要なIGおよびTRを産生する細胞を欠いているためである。「機能獲得」形態において、人は疾患にかかりやすく、何故ならば、免疫プロファイルは、通常はそこに存在するべきでないIGおよびTRを産生する細胞を獲得したためである。「機能喪失」(LOF)免疫プロファイル疾患において、個体は、疾患と戦うために適切な機能性B細胞またはT細胞を有しない。HLAタイピングは、それらの細胞が、免疫細胞成熟プロセスの初期段階の間に排除されることを明らかにしており、それらの細胞は、自己タンパク質に対して強く反応するために、一般に排除される。

10

【0028】

本発明の一局面はまた、以下を含む方法を提供する：(a)1つまたは複数の個体由来のT細胞および/またはB細胞由来の1つまたは複数のRNAを増幅および配列決定する工程、(b)配列をデータベースへ入力して、コンピューター、サーバー、または他の電子記憶装置上に保存され得るデータを提供する工程、(c)1つまたは複数のRNAの配列に対応する個体についての識別情報および特徴を、コンピューター、サーバー、または他の電子記憶装置上に保存されてもよいデータとして入力する工程、ならびに(d)1つまたは複数の個体について工程(b)および工程(e)のデータを評価して、1つまたは複数のRNA配列と配列に対応する個体の1つまたは複数の特徴との間に相関関係が存在するかどうかを決定する工程。識別情報は、例えば、患者識別番号、患者のHLA型を含むコード、下されている場合がある1つまたは複数の臨床診断を含む疾患コード、サンプルの日付を含む「ステージング・コード」、RNAが増幅および配列決定された細胞亜集団の型を含む細胞型コード、ならびにサンプルについて同定された配列を含む1つまたは複数の配列コードを含み得る。

20

【0029】

記載の方法は、免疫レパートリー全体の増幅を可能にするだけでなく、高度に多重化された様式でかつ半定量的に増幅を可能にする、新規のプライマー設計を含む。多重増幅は、ごく少数のPCRまたはRT-PCR反応が必要とされることを要求する。例えば、全てのIGが1つの反応で増幅され得、または、それをIgH、IgLまたはIgKについて2つまたは3つの反応に分割することができる。同様に、TRA、TRB、TRD、およびTRGを含む、T細胞受容体(TR)は、たった1つの反応で増幅され得るか、または少数の反応で増幅され得る。半定量的増幅は、多重反応における全ての標的が独立して増幅され、その結果、増幅産物のエンド・ポイント分析は、それらの標的における元の内部比率を反映することを意味する。

30

【0030】

様々な局面において、方法は、T細胞またはB細胞内の遺伝子再構成から生じ、特定の抗原についての抗体および/またはT細胞受容体の親和性および特異性の原因である相補性決定領域(CDR)を含むポリヌクレオチドの半定量的増幅をもたらすことができる。半定量的増幅は、特定のCDR3配列の存在を検出するだけでなく、それらのCDR3配列を生成するために必要な組換え事象をもたらした細胞の相対数を決定するための方法を提供する。

40

【0031】

本発明の一局面は、従って、半定量的配列情報を解析してヒトまたは動物についての1つまたは複数の免疫状態レポートを提供するための方法に関する。免疫状態レポートを作成するための方法は、以下の工程を含む：(a)対象の免疫プロファイルとデータベース中に保存された疾患ライブラリーとの間で共有される1つまたは複数の別個のCDR3配列を同定し、該共有されるCDR3配列の総数を合計し、そして対象の免疫プロファイルと疾患ライブラリーとの間で共有される、対象の免疫プロファイル中の配列の総数のパーセンテージ

50

を計算して、1つまたは複数のオリジナル・シェアリング・インデックスを作成する工程；(b)データベース中に保存されたパブリックライブラリーから配列を無作為に選択して、疾患ライブラリー中の別個のCDR3配列の数にほぼ等しい複数の配列を含むサブライブラリーを形成し、対象の免疫プロファイルとサブライブラリーとの間で共有される1つまたは複数の別個のCDR3配列を同定し、該共有されるCDR3配列の総数を合計し、そして対象の免疫プロファイルとサブライブラリーとの間で共有される、対象の免疫プロファイル中の配列の総数のパーセンテージを計算して、サンプリング・シェアリング・インデックスを作成する工程；(c)工程(b)を少なくとも1000回またはそれ以上繰り返す工程；および(d)サンプリング・シェアリング・インデックスが、患者の免疫プロファイルと疾患ライブラリーとの間のオリジナル・シェアリング・インデックスに等しいかまたはそれを超える回数の割合としてP値を推定する工程。

10

【0032】

本発明者らは、例えば、癌、自己免疫疾患などといった、ある疾患を有する個体の免疫プロファイルは、1つまたは複数の免疫細胞集団における多様性の欠如を特徴とし得ることを発見した。図1は、例えば、癌患者（図1a）由来の血液サンプル中に存在するT細胞の別個のタイプおよび数と、健康な患者（図1b）由来のものとの間で観察され得る差異を示す図であり、ここで、各円は、T細胞受容体の相補性決定領域（例えば、CDR3）の、増幅および配列決定された組換えcDNAによって示されるT細胞の別個のタイプと、同じCDR3配列を共有するとPCR増幅および配列決定によって決定された細胞の相対数とを示す。図1aが示すように、CDR3配列によって示されるように、異なる特異性の別個の細胞はより少ないが、ある特定の特異性の細胞の数はより多い場合がある。図1bは、異なる細胞がより多いが、同じCDR3配列を共有する各タイプの細胞の数がより少ない正常プロファイルを示している。

20

【0033】

各別個のCDR3を発現する細胞のリスト、およびヒトまたは動物由来の血液または組織サンプル中に示されるそのような細胞の数は、そのヒトまたは動物についての免疫プロファイルを構成し得る。ヒトのある群、例えば、健康な個体と様々な異なる疾患を有する個体との両方を含む群からの免疫プロファイルをコンパイルすることは、正常な集団において見られる多様性のタイプを表す「パブリックライブラリー」を提供し得る（図2）。同様に、特定の疾患を有すると臨床的に診断された個体の群の免疫プロファイルをコンパイルすることは、拡大された細胞集団の特定のCDR3など、多様性の欠如を表す「疾患ライブラリー」を提供し得る（図3）。これらの免疫プロファイルは、例えば、インターネットへのコンピューターアクセスによってアクセス可能なデータベース中に保存され得、それにより、情報は、患者の免疫状態を分析するために本発明の方法において使用され得る。

30

【0034】

別個のCDR3を発現する細胞（「別個のCDR3」、これらの細胞は特有のCDR3配列を共有する）のリスト、および個体由来の血液または組織サンプル中に存在する各別個のCDR3の数を含む、免疫プロファイルを、個々の患者について作成してもよい。患者の免疫プロファイルを、特定の疾患を有すると診断された患者の群の、組み合わせられた免疫プロファイル（データベース中に保存された、疾患ライブラリー）と比較する。これは、一連の疾患ライブラリーについて行われ得、図4に示す。

40

【0035】

何百万もの可能性のある組み合わせが、パブリックライブラリーについて可能であり、それらの個体の大部分の免疫系は、一般に、特定の疾患を有すると診断された個体の群のそれよりも増加した多様性を示す。従って、本発明者らは、本発明の方法についての正確な評価および比較が、パブリックライブラリー中の別個のCDR3のリストおよびそれらの数から無作為にサンプリング/選択することによってサブライブラリーを作成する工程によって促進されると判断した。CDR3断片の特有のペプチド配列によって示される、別個のCDR3の数は、疾患ライブラリーにおいて同定された別個のCDR3の数、または2つ以上の疾患ライブラリーから計算された平均値とほぼ同じであるべきである。パブリックライブラリ

50

ーから無作為にサンプリングすることによって作成された、有意な数のサブライブラリー、例えば、1000またはそれ以上のサブライブラリーを作成することは、様々な別個のCDR3の存在を増加させ、個々の患者の免疫プロファイルを、正常（「健康」）と同定および特徴づける、または、特定の疾患と関連付けられた細胞のタイプおよび数の存在を特徴とすると同定および特徴づけるために、統計的に有意に有効である結果をもたらす。

【0036】

本発明の方法において、患者は、そこから別個のCDR3が半定量的に増幅および配列決定される、例えば血液または組織を含む臨床サンプルを供給する。これは、別個のCDR3の全てについての各CDR3の識別および相対存在量を提供する。この情報は、少なくとも1つのパブリックライブラリーおよび1つまたは複数の疾患ライブラリーを含有するデータベースにアクセスするプログラムへ入力され得る。データ入力および/または解析のために使用されるソフトウェアは、データベースへのインターネットアクセスを介してアクセスされ得るか、またはデータベース中の配列情報へのインターネットアクセスを伴って、個々のパーソナル・コンピュータ上に置かれ得る。個体の免疫プロファイルと様々なライブラリーおよびサブライブラリーとの間で比較を得、図4および図5に一般に示されるように結果を作成し、ここで、特定のCDR3配列が検出され、検出されたそれらの別個のCDR3配列の数がカウントされ、そしてその特定の別個のCDR3が個体の免疫プロファイルおよび特定のライブラリーの両方に存在する（即ち、その特定の別個のCDR3が個体と該ライブラリーとの間で「共有される」）か否かについて判定がなされる。共有されると判定されたCDR3の数を示すパーセンテージを一緒に合算して、個体の免疫プロファイルと特定のライブラリーとの間で共有される個体の免疫プロファイル中のCDR3から構成される総数の割合から構成される合計を作成する（即ち、「シェアリング・インデックス」）。サブライブラリーについて得られた結果から、P値を、ランダムパーセンテージが特定の疾患ライブラリーについて示されるパーセンテージに等しいかまたはそれを超える確率として計算し、サンプリング・シェアリング・インデックスが特定のライブラリーについてのオリジナル・シェアリング・インデックスを超える回数の割合が例えば0.01未満である場合、有意な結果が示される。そのシェアリング・インデックスが個体の免疫プロファイルと疾患ライブラリーとの関係を示す場合、個体は、個体/患者が、特定の疾患ライブラリーによって示される疾患を有する可能性について知らされ得る。全ての疾患ライブラリーに対して算出されたP値が0.01を超える場合、個体のレポートは、免疫プロファイルが正常そうであり、疾患状態は検出されなかったことを示し得る。

【0037】

配列データはコンパイルされ、個体の複数の集団についての1つまたは複数のデータベース中に保存されるので、さらに、あるシェアリング・インデックスを、ある疾患の前提条件または素因を有する集団を示すライブラリーと関連付けることが可能であり得る。免疫系は事前対応（proactive）および事後対応（reactive）の両方であり、免疫プロファイルに反映される免疫系の変化は、素因、前提条件、またはさらには確立された疾患が存在するという、最初の、場合によっては唯一のシグナルを提供し得る。本発明者らは本方法を利用して、あるタイプの癌、炎症性腸疾患、およびあるウイルス感染症が、存在するCDR3を示す免疫レパートリーのサブセットを作成するためにARM-PCR法を使用してCDR3を配列決定することによって得られた患者および確立された疾患ライブラリー間のシェアリング・インデックスを決定することによって検出され得ることを実証した。

【0038】

フィルターが配列データへ適用される場合、結果は信頼性がさらにより高くなる。例えば、本発明者らは、信頼性が有意にさらに高い結果の作成を助ける、配列データについての「SMART」フィルターを開発した。これを実施例においてさらに説明する。

【0039】

さらなる説明のために、下記の例は本発明の方法を説明し得る。血液サンプルを任意のワクチンの投与前に小児から採取してもよく、各小児についてのそれらの血液サンプルは「ベースライン」を確立し、それから将来のサンプルが評価され得る。各小児について、

将来のサンプルは、疾患と相関することが公知の細胞の集団を拡大させた作用因子への曝露が存在したかどうかを決定するために利用され得、これは、将来的にその疾患を発症するリスクについての「マーカー」として役立ち得る。そのように同定された個体は、次いで、より厳密にモニターされ得、その結果、早期発見が可能となり、任意の利用可能な治療選択肢が、疾患過程のより早い段階で提供され得る。

【0040】

別の例によると、血液サンプルを任意のワクチンの投与前に小児から採取してもよく、各小児由来のそれらの血液サンプルは「ベースライン」を確立し、それから将来のサンプルが評価され得る。各小児についておよび研究における小児の集団全体について、それらのベースラインは、ワクチン投与後、抗体およびT細胞受容体を増幅させるために標的特異的プライマーを使用したT細胞およびB細胞のRNAシーケンシングの結果と比較され得る。比較は、症状、診断された疾患、ならびに各個体について対応する抗体およびT細胞受容体配列と関連付けられる他の情報に関するデータの評価をさらに含み得る。ワクチンの投与と特定の疾患の発症との間に関係が存在する場合、その疾患の症状を示す個体はまた、例えば、対応する抗体もしくはT細胞受容体、または対応する抗体もしくはT細胞受容体のセットを共有し得る。

【0041】

本発明の方法は、例えば、自己免疫疾患を有する個体間の共通点を同定するために特に有用であり得、感染および環境因子と、心臓疾患、アテローム性動脈硬化症、糖尿病、および癌などの疾患との相関関係をよりよく説明する疫学的データを提供し得る - 疾患の存在、または疾患を発症する傾向のいずれかを示す「バイオマーカー」を提供する。

【0042】

本方法はまた、受動免疫療法の開発のために有用であり得る。例えば、感染要因への曝露後、ある抗体産生B細胞および/またはT細胞が増殖される。本発明の方法は、例えば、防御抗体の同定を可能にし、それらの抗体は、そのような療法が必要とされる状況において受動免疫療法を提供するために利用され得る。

【0043】

本発明の方法はまた、望ましくない再構成を有する細胞の標的化除去を達成する能力を提供し得、この方法は、そのような細胞再構成が同定され得る手段を提供する。

【0044】

本発明者は、本発明の方法における使用のための標的特異的プライマーを同定および開発した。T細胞特異的プライマーを表1に示し、抗体特異的プライマーを表2に示す。本発明のさらなる態様は、表1または表2のプライマーのいずれか1つまたは組み合わせを使用して、血液サンプル由来のRNAを増幅する、より具体的には、細胞の集団中の抗体、T細胞受容体、およびHLA分子を同定する方法である。

【0045】

arm-PCRまたはtem-PCRは、本発明者によって以前記載された増幅方法において免疫グロブリンスーパーファミリー分子をコードする遺伝子を増幅するために使用され得る (Han et al., 2006. Simultaneous Amplification and Identification of 25 Human Papillomavirus Types with Templex Technology. J. Clin. Micro. 44(11). 4157-4162)。tem-PCR反応において、ネステッド遺伝子特異的プライマーは、最初のPCRサイクリングの間、標的を濃縮するように設計されている。その後、ユニバーサルな「スーパー」プライマーを、全ての標的を増幅するために使用する。プライマーを F_0 (フォワード・アウト)、 F_i (フォワード・イン)、 R_i (リバース・イン)、 R_0 (リバース・アウト)、FS (フォワード・スーパー・プライマー) およびRS (リバース・スーパー・プライマー) と呼び、スーパープライマーは、標的特異的プライマーの末端でのそれらのプライマーに対する結合部位の付加に起因して様々な分子に共通である。遺伝子特異的プライマー (F_0 、 F_i 、 R_i 、および R_0) を極めて低い濃度で使用する。異なるプライマーは、3つの主な段階の各々でtem-PCRプロセスに関与する。まず、「濃縮」段階で、低濃度遺伝子特異的プライマーに、テンプレートを見つけるために十分な時間を与える。各々の意図される標的について、どの

プライマーが使用されるかに依存して、4つの可能性のある産物が生じ得る： F_o/R_o 、 F_i/R_o 、 F_i/R_i 、および F_o/R_i 。濃縮段階は典型的に10サイクル行われる。第二または「タグ化」段階において、アニーリング温度を72℃へ上昇させ、長い40ヌクレオチドのインサイド・プライマー（ F_i および R_i ）のみが作用する。このタグ化段階の10サイクル後、全てのPCR産物がユニバーサル・スーパー・プライマー配列で「タグ化」される。次いで、第三「増幅」段階で、高濃度スーパープライマーは、全ての標的を増幅し、プロセス中にビオチンでPCR産物を標識するように、効率的に作用する。共有結合によって特定のプローブをLuminexカラー被覆ビーズと連結させてもよい。

【0046】

免疫グロブリンスーパーファミリー分子をコードする遺伝子を増幅するために、本発明者は、パブリック・ドメイン中の配列情報に基づいてネステッドプライマーを設計した。B細胞およびT細胞でのVDJ再構成を研究するために、本発明者は、再構成されそして発現されたRNAを増幅するためのプライマーを設計した。一般に、一对のネステッド・フォワード・プライマーはV遺伝子から設計され、リバーズ・ネステッド・プライマーのセットはJまたはC遺伝子から設計される。平均アンプリコンサイズは250～350bpである。例えば、IgHV遺伝子について、7個の異なるファミリーに分類され得る123個の遺伝子が存在し、本発明のプライマーをファミリー特異的となるように設計する。しかし、増幅されたcDNA配列を配列決定する場合は、同じファミリー内の遺伝子間のさらなる区別を可能にするために十分な配列多様性が存在する。MHC遺伝子座について、意図はゲノムDNAを増幅することである。

【実施例】

【0047】

シェアリング・インデックスの計算

Sは対象の免疫プロファイル（IP）であり、これは、N個の特有のCDR3配列である $CDR3_1$ 、 $CDR3_2$ 、... $CDR3_n$ によって示され、各CDR3はそれ自体の頻度、 s_1 、 s_2 、... s_n を有すると仮定する。

【0048】

Dは疾患ライブラリーであり、これは、M個の特有のCDR3を有するある数の患者の免疫プロファイルの合計である。疾患ライブラリー中の全ての患者は、同じ疾患を有すると診断された。

【0049】

Pはパブリックライブラリーであり、これは、多数の対照の免疫プロファイルの合計である。

【0050】

シェアリング・インデックス（SharingIndex）は、 s_x 、 s_y 、... s_z の合計と定義され、ここで、 $CDR3_x$ 、 $CDR3_y$ 、... $CDR3_z$ は、対象の免疫プロファイルとライブラリーとにおいて共有される。 s_x 、 s_y 、... s_z は対象の免疫プロファイルにおけるCDR3の頻度であって、ライブラリーにおけるものではないことに注意のこと。

【0051】

疾患ライブラリー（D）中よりもパブリックライブラリー（P）中により多くの特有のCDR3が常に存在すると仮定して、パブリックライブラリー中のM個の特有のCDR3を無作為に選択し、使用して、サブライブラリーP1を作成し、対象とサブライブラリーとの間のシェアリング・インデックス（ SI_{p1} ）を、上記の式に従って算出する。サンプリング手順を1000回またはそれ以上繰り返し、1000個またはそれ以上の SI_{px} を算出する。

【0052】

対象と疾患ライブラリーとの間のシェアリング・インデックス SI_d を同じ様式で算出する。P値を、全ての SI （ SI_{p1} 、 SI_{p2} 、... SI_{px} 、 SI_d （ SI_d が含まれることに注意のこと））の分率と定義し、これは、 SI_d に等しいかまたはそれを超える。パブリックライブラリー中のCDR3をサンプリングする場合、x個の対照の免疫プロファイル中に見られるCDR3は、x回のサンプリングされる機会が与えられることに注意のこと。

【 0 0 5 3 】

T細胞またはB細胞再構成部位の増幅

全てのオリゴ (oligo) を、1x TEを用いて再懸濁した。454Aおよび454Bを除く全てのオリゴを100 pmol / μ Lの濃度へ再懸濁した。454Aおよび454Bを1000 pmol / μ Lの濃度へ再懸濁した。454Aおよび454Bは、以前に記載された共用プライマーと機能的に同じであり、ハイスループット配列決定手順をフォローアップするために、異なる配列を使用した。

【 0 0 5 4 】

3つの異なるプライマーミックスを作製した。アルファ・デルタ・プライマーミックスは82個のプライマー (TRAV-C + TRDV-Cの全て) を含み、ベータ・ガンマ・プライマーミックスは79個のプライマー (TRBVCおよびTRGV-Cの全て) を含み、B細胞プライマーミックスは合計70個のプライマーを含んだ。F₀、F_iおよびR_iプライマーを1 pmol / μ Lの濃度で使用した。R₀プライマーを5 pmol / μ Lの濃度で使用した。454Aおよび454Bを30 pmol / μ Lの濃度で使用した。

【 0 0 5 5 】

3つの異なるRNAサンプルをALLCELLS(www.allcells.com) に注文した。全てのサンプルを最終濃度4 ng/uLまで希釈した。注文したサンプルは以下であった。

細胞型 :	供給源 :
ALL-PB-MNC	急性リンパ芽球性白血病を有する患者
NPB-Pan T 細胞	正常T細胞
NPB-B 細胞	正常B細胞

【 0 0 5 6 】

Qiagen One-Step RT-PCRキットを用いてRT-PCRを行った。各サンプルは以下を含有した :

- 10 μ LのQiagenバッファー
- 2 μ LのDNTP
- 2 μ Lの酵素
- 23.5 μ LのdH₂O
- 10 μ Lの適切なプライマーミックス
- 2.5 μ Lの適切なテンプレート (10ngのRNAトータル)

【 0 0 5 7 】

下記のサイクリング条件を使用して、サンプルを実行した :

- 50 で30分間
- 95 で15分間
- 94 で30秒間
- 以下を15サイクル
 - 55 で1分間
 - 72 で1分間
 - 94 で15秒間
- 以下を6サイクル
 - 70 で1分30秒間
 - 94 で15秒間
- 以下を30サイクル
 - 55 で15秒間
 - 72 で15秒間
 - 72 で3分間
 - 4 で維持。

【 0 0 5 8 】

10

20

30

40

50

図1aに示されるゲル中に配置されたサンプルの順序は以下であった：(1)ラダー（500bpが最大のワーキング（working）であり、20bpずつ低下し、図1aの中央の鮮明なバンドは200bpである）；(2) + プライマーミックスおよび10ng Pan T細胞テンプレート；(3) + プライマーミックスおよび10ng Pan T細胞テンプレート；(4)B細胞プライマーミックスおよび10ng B細胞テンプレート；(5)B細胞プライマーミックスおよび10ng ALL細胞テンプレート；(6) + プライマーミックスおよび10ng ALL細胞テンプレート；(7) + プライマーミックスおよび10ng ALL細胞テンプレート；8. + プライマーミックスブランク；(9) + プライマーミックスブランク；(10)B細胞プライマーミックスブランク；(11)ランニングバッファブランク。これらのサンプルを、1X ClearPAGE（登録商標）DNAネーティブランニングバッファを用いて、プレキャストClearPAGE（登録商標）SDS 10%ゲル上において実行した。

10

【 0 0 5 9 】

最初の実験は、テンプレートが含まれたPCR反応からスミアが生じることを示した。スミアは、異なるサイズのPCR産物が生じたことを示しており、これは、異なるVDJ再構成の混合物を示した。B細胞反応からいくらかのバックグラウンド増幅が存在する。そのプライマーミックスに対するさらなる改善が、反応をクリーンアップするために必要とされた。

【 0 0 6 0 】

PCR産物が異なるVDJ再構成を実際に含むかどうかを決定するために、単クローンを単離して配列決定することが必要であった。ルーチンなクローニング手順を使用する代わりに、本発明者は異なる戦略を用いた。アルファ・デルタ・ミックスおよびベータ・ガンマ・ミックスから生成されたPCR産物（図1a中のレーン2および3）を1:1000に希釈し、2 μ l アリコートを下記の反応においてPCRテンプレートとして使用した。次いで、レパートリー全体を標的化するプライマーの混合物を使用する代わりに、一対の特定のFiおよびRiプライマーを使用して（各々5 pmol）、たった1つの特定のPCR産物を増幅した。下記のサイクリング条件を使用して、サンプルを増幅した：

20

95 で5分間

以下を30サイクル

94 で30秒間

72 で1分間

72 で3分間

4 で維持。

30

【 0 0 6 1 】

Qiagen PCRキットを使用して産物を増幅した。PCRのために使用したマスターミックスは以下を含有した。

	1反応当たり	マスターミックス x 12
10x PCR バッファ	5 μ L	60 μ L
dNTP	1 μ L	12 μ L
HotStartTaq Plus	0.25 μ L	3 μ L
H ₂ O	39.75 μ L	477 μ L

40

【 0 0 6 2 】

図1bにおけるゲルの写真は以下の反応のPCR産物を示す：(1)ラダー；(2)TRAV1F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(3)TRAV2F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(4)TRAV3F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(5)TRAV4F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(6)TRAV5F_i+TRACR_iおよびアルファ・デルタP

50

an T PCR産物；(7)TRAV1F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(8)TRAV2F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(9)TRAV3F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(10)TRAV4F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(11)TRAV5F_i+TRACR_iおよびアルファ・デルタPan T PCR産物；(12)PCRブランク。F_iとして列挙されるプライマーは「フォワード・インナー」プライマーであり、F_oとして列挙されるプライマーは「フォワード・アウター」プライマーであり、R_iおよびR_oは、それぞれ、「リバーシ・インナー」および「リバーシ・アウター」プライマーを示す。

【 0 0 6 3 】

図1bによって示されるように、単一のPCR産物が各反応から生成された。異なるサイズバンドが異なる反応から生じた。このPCRクローニングアプローチは、次の2つの主な理由のために成功している - (1)この反応において使用したPCRテンプレートは、全ての可能性のあるVDJ再構成を増幅するためのプライマーミックスを使用した（例えば、T細胞受容体

10

および 遺伝子を増幅するために合計82個のプライマーを含んだプライマーミックスを使用した）以前の反応の、希釈したPCR産物(1:1000)であった、および(2)特定のV遺伝子を標的化する、たった1対のプライマーを、この「クローニング」実験の間、各反応において使用する。これらの産物の一部をゲル精製し、配列決定した。以下は、上述のプロトコルから得られる例示的配列である。各ケースにおいて、単一クローンが得られ、F_iプライマーにマッチした特定のT細胞受容体V遺伝子が同定された。

【 0 0 6 4 】

TRAV1テンプレート + シークエンシングプライマーとしての454A :

20

NNNNNNNNNNCNTANTCGGTCTAAGGGTACNGNTACCTCCTTTTGAAGGAGCT
CCAGATGAAAGACTCTGCCTCTTACCTCTGTGCTGTGAGAGATANCAACNATCA
CTTAATCTTGGGCGCTGGGAGCAGACTAATTATAATGCCAGATATCCACAACCC
TGACCCTGCCGCGTACCAGCTGAAAGACTATGAACAGGATGGGGAGGCAGNAG
NAGNAG (SEQ ID NO. 1)

【 0 0 6 5 】

TRAV1テンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNNNGNANGNNCAGGGTTCTGGATATTTGGTTTNACAATTAGCTTGGT
CCCTGCTCCAAAGATTAATTTGTAGTTGCTATCCCTCACAGCACAGAGGTAAGA
GGAAGAGTATTTCTTCTGGAGCTCCTTCAACAGGAGGAACTGTACCCTTTATA
CCTACTAAGGAATGAAGA (SEQ ID NO. 2)

30

【 0 0 6 6 】

TRAV2テンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNNNNNTNNCGGTTCTCTTNNTCGCTGCTCATCCTCCAGGTGCGGGA
GGCAGATGCTGCTGTTTACTACTGTGCTGTGNANNANGGCANNGACAACAACCT
CNTCTTTGGTGGAGGNACCCTACTNNTGGTTATNCCNAATANCCANAACCCTGA
CCCTGCCGAGNAGCAGCANAAAACTNNNAGGGGGGTGGAGAAGNANNNNN
(SEQ ID NO. 3)

【 0 0 6 7 】

TRAV3テンプレート + シークエンシングプライマーとしての454A :

40

NNNNNNNNNNNNNNNGGNNNGGNAGCTATGGCTTTGAAGCTGAATTTAACAAGA
GCCAAACCTCCTTCCACCTGAAGAAACCATCTGCCCTTGTGAGCGACTCCGCTT
TGTAATTTCTGTGCTGTGAGAGACATCAACGCTGCCGGCAACAACCTAACTTTTG
GAGGAAGAACCATGGTGCTAGTTAAACCAAATATCCATAACCCTGACGCTGCCG
TGTAACAGCTGAAAGACTCTGAGGGGGCTGGAGAGGNAGGNG (SEQ ID NO. 4)

【 0 0 6 8 】

TRAV4テンプレート + シークエンシングプライマーとしての454A :

NNNNNANNGGNNNNNGTTTATCCCTGCCGACAGAAAGTCCAGCACTCTGAGCC
TGCCCCGGGTTTCCCTGAGCGACACTGCTGTGTACTACTGCCTCGTGGGTGAC
CGGTCTGGAAACAGCGATGAAATTTTCATCTTAGGAAGAAGAACGCTTCTAGTC
ATCCANCCCAACATCCACAACCCTGCCGCGGAGNAGCACCAGAAAAAAGATGA
TGAGGGGGANGNAGNAGNANNNN (SEQ ID NO. 5)

【 0 0 6 9 】

TRAV5テンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNNNNNNNNNTCNCTGNTCTATTGAATAAAAAGGATAAACATCTGTCT
CTGCGCATTGCAGACACCCAGACTGGGGACTCAGCTATCTACTTCTGTGCAGA
GAGCCCCGGTGGCGGCAGCAACTTCTTCTTTGGTGGAGGAGCANTACTACTAG
TCGTTCTACATANCCACAACCATGATNCCGCCGAGTACNTGCTGAAAAAATATG
ATGAGGATGGAGAAGAAGNAGCATNAN (SEQ ID NO. 6)

10

【 0 0 7 0 】

TRBV19Fiテンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNCTGAGGGGTANNCGTCTCTCGGGAGAAGAAGGAATCCTTTCCTCTC
ACTGTGACATCGGCCCAAAGAACCCGACAGCTTTCTATCTCTGTGCCAGTAGT
ATGGGGGGGGGGGGCCTACAATGAGNACGGCGGGGGGAGGGACNNTGCTC
GTCGTGGAGGAGGACATGAAGGTCTTGCCCGCNCNGAGGAAGNTGNANANG
AACCATAAAAATGCGCTGGCTGAANN (SEQ ID NO. 7)

【 0 0 7 1 】

20

TRBV20Fiテンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNNNGCTCNNNNNNNCNATACGAGCAAGGCGTCGAGAAGGACAAG
TTTCTCACAAACCATGCAAGCCTGACCTTGTCCACTCTGACAGTGACCAAGTGCCC
ATCCTGAAGACAGCAGCTTCTACATCTGCAGTGCTAGAGGGGGGGGGGGGA
CGACTACTACTACTTCGGCGGGGGGGGGCATGCTGATCGTGGAGGAGGAGGAC
ATGNAGCTCCTCCCCGCCGCCGAGGTTGTTGTGTNTNNANCATCATACTGNTG
GTGGAGNAGNAGNAGCN (SEQ ID NO. 8)

【 0 0 7 2 】

TRBV21Fiテンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNNNNNNNNNGNNNNNNNNNNNTACTTTCNGAATGAAGAACTTATTCA
GAAAGCAGAAATAATCAATGAGCGATTTTTAGCCCAATGCTCCAAAACTCATCC
TGTACCTTGAGTTCCAGTCCACGGAGTCAGGGGACACAGCACTGTATTTCTGT
GCCAGCAGCA (SEQ ID NO. 9)

30

【 0 0 7 3 】

TRBV23Fiテンプレート + シークエンシングプライマーとしての454A :

NNNGNNNNNNNANNGGANANGCACAAGAAGCGATTCTCATCTCAATGCCCCAA
GAACGCACCCTGCAGCCTGGCAATCCTGTCTCAGAACCAGGAGACACGGCAC
TGTATCTCTGCGCCAGCAGTCAATCGGGGGGGGGGGGGAGGGCCGTCCGCAG
CGGGGGGGGGGGGGGGGGCGGGGGACGGTCCCAAAGAGAAAGAAAACCTGCCC
CCCGCGCTCGGGCGGTGTGATTGAGCGAAACAGACAGGAAGGNAAGNAAAAA
NNNNANCNNCNCCTCNN (SEQ ID NO. 10)

40

【 0 0 7 4 】

TRBV24Fiテンプレート + シークエンシングプライマーとしての454A :

NNNNNNNNNGNNANNTCTGATGGANACAGTGTCTCTCGACAGGCACAGGCTAA
ATTCTCCCTGTCCCTAGAGTCTGCCATCCCCAACAGACAGCTCTTACTTCTGT
GCCACCAGTGANGCGGGGGGGCGGGGACCACTACTTCGGGGGGGGGGAGGCGG
ACCAGGGTGCTGGTGCAGAGAAAAAGGAGTCCCCCCCCGCCGCCGCTGTGG
TTGTTGCTTCATAATAATCAGGNNGGNGAGGNAGNAGNANN (SEQ ID NO. 11)

【 0 0 7 5 】

TCR トランスクリプトームの全体的レパートリー分析に対する人工産物の影響を調べ

50

るために、本発明者らは、化学合成したTCR CDR3テンプレートを用いて対照実験を行った。このために、本発明者らは、4つの別個のクローンを化学合成し、各クローンをクローン的に精製し、アンプリコン・レスキュー・マルチプレックス (amplicon rescue multiplex) (ARM)-PCR用のテンプレートとして4つの構築物の異なるミックスを調製した。2つの異なる反応混合物を、2つの独立したARM-PCR反応へ供し、プールされたPCR産物を、Illumina HiSeq2000 (登録商標) を使用して両末端から100bpの長さで配列決定した。本発明者らは、先ず、改変Needleman-Wunschアルゴリズムを用いてオーバーラッピング・アライメントによってペアド-エンド読み取り (paired-end read) を共に連結し、次いで、この結合配列を生殖細胞系V、DおよびJリファレンス配列へマッピングした。

【0076】

クリーニング無しで、本発明者らは、テンプレートミックスIから合計5,729,613個の配列を得、これらをTCR V、DおよびJセグメントへマッピングすることができた。驚いたことに、配列読み取りは、その称するところでは、合計36,439個の特有のCDR3変異体を示した。従って、たった4つの別個のCDR3変異体がテンプレート混合物中に存在していたことを考えると、同定されたCDR3変異体のうちの事実上全てが、非真正 (non-authentic) であるはずである。同様の結果が第2テンプレートミックスについて得られ、ここでは、合計9,131,681個のVDJマッピングされた配列が同定され、これらは、50,354個の特有のTCR CDR3変異体の存在を模倣した。本発明者らの独立した配列決定実験は、ごく僅かな別個のCDR3テンプレート変異体が人工的なレパートリー多様性を作ることができ、これは実際のテンプレート多様性を遥かに上回ることを示しており、従って、本発明者らは、これらの人工産物を排除しようと試みた。

【0077】

3'末端Illumina配列決定読み取りの質は一般に低いと考えられている。レパートリー配列決定の文脈において、これは厄介であり、何故ならば、PCRプライマーは、プライマー-テンプレートミスマッチングに起因する悪影響を回避するために超可変V(D)J結合部から十分に遠位に配置される必要があるためである。結果として、関心対象のCDR3セグメントは、増加した配列決定エラー率を有する領域である、配列決定読み取りの3'末端のより近くへ一般に「シフトされる」。注目に値する別の技術的な問題は、配列決定エラーがコンテキスト特異的 (context-specific) であり、従ってストランド特異的 (strand-specific) であるという観察である。従って、順方向読み取りにおける配列決定エラーが対応する逆方向読み取りにおける配列決定エラーと同時に起こる確率は稀であると想定することが現実的である。

【0078】

このことを考慮して、本発明者らは、Illumina (登録商標) 技術に基づいて完全TCR CDR3セグメントの二本鎖配列決定を与えるペアド-エンド戦略を発明した。このアプローチにおいて、フォワードおよびリバースシーケンシングプライマーを、それぞれ、フレームワーク領域3におよびTCR J領域またはC領域の5'末端に配置する。Illumina配列読み取りの平均の長さ (現在100~150 bp) を考慮して、この設計は、CDR3セグメントを規定する両方の鎖の完全な配列決定を可能にする。第2の工程において、順方向および逆方向読み取りは、次いで、配列ミスマッチについて分析され、両方の鎖の非同一性を示すCDR3配列は、新しく開発されたペアド-エンド・フィルタリング・アルゴリズムを使用して排除される。

【0079】

テンプレートミックスIについて得られた5,729,613個のCDR3配列へこの配列決定エラーフィルターを適用することによって、本発明者らは、逆鎖において矛盾する配列情報を含む合計2,751,131個 (48%) のCDR3配列を同定した。これらの配列を捨てることによって、35,455個 (97.2%) の別個の人工的なCDR3変異体が排除された。これと一致して、ペアド-エンドフィルターは、テンプレートミックスIIから4,308,020個 (47%) のCDR3配列を除去し、49,063個 (97.4%) の人工的なCDR3変異体が排除された。それぞれ、合計973および1271の特有のCDR3変異体が、フィルターを通過した。これらの結果は、ペアド-エン

ド配列決定およびフィルタリングが、非真正の特有のCDR3配列の総数をほぼ2桁減らすことを示している。

【0080】

配列決定エラーフィルター後の非真正CDR3変異体の頻度分布の詳細な解析によって、両方の混合物中に、全ての人工産物のうちのおよそ50%がシングルコピー配列であったことが明らかとなった。これらの人工的なCDR3のうち約10%が、100を超えるコピー数を示し、全ての人工的なCDR3変異体の80%超を占めた。可変TCR遺伝子は体細胞超突然変異を受けないことを考慮して、本発明者らは、マッピングされた生殖細胞系V、DおよびJリファレンス配列に対してヌクレオチドミスマッチを示すCDR3配列読み取りを同定および除去するリファレンスアルゴリズムを開発した。何故ならば、これらは、PCR増幅または配列決定

10

【0081】

このフィルタリングアルゴリズムをテンプレートミックスIの「ペアド-エンドフィルタリングされた」配列へ適用することによって、合計29,804個の配列（これは609個の特有のCDR3変異体に対応する）が除去された。テンプレートミックスIIについて、54,516個の人工的な配列（831個の特有のCDR3変異体）が同定された。従って、リファレンス配列フィルターの使用は非真正の別個のCDR3配列の60%の減少をもたらす。リファレンスフィルターは、V-JおよびD-J結合部で無効である。何故ならば、体細胞組換えの間にこれらの領域中に無作為に付加されたヌクレオチドはマッピングすることができないためである。従って、本発明者らは、次の4つの変数をよりよく理解するために、コンピュータシミュレーション実験後にPCRフィルターを実行した：総エンド-ポイントエラー率に対する、最初のテンプレート数、各サイクルの複製効率、サイクル数(n)、およびDNAポリメラーゼエラー率(μ)の影響。対照的に、本発明者らは、PCRポリメラーゼエラー率が、累積誤差数に対して顕著な影響を有することを見出した。

20

【0082】

本発明者らの対照配列決定実験において、PCR増幅を、Taqポリメラーゼを使用して、第1および第2反応において15サイクルおよび45サイクルで行った。より現実的にARM-PCR反応中の誤差累積をシミュレートするために、PCR効率を最初の25サイクルについて1サイクル当たり5%および残りのサイクルについて1サイクル当たり10%減少するように設定した。それぞれの新たなPCRについてPCR効率を1.0へリセットした。さらに、本発明者らは、第2

30

【0083】

エンドポイントエラー率に対する最初のテンプレート数およびPCR効率の影響は小さいため、多重PCR反応において異なる最初の数および異なる複製効率を有する分子に対してシミュレーション実験から推定される同じエンド-ポイントエラー率を適用することが安全であるはずである。カットオフエラー率(μ)を、各カテゴリーについて9999番目の10000分位点でエラー率として経験的に設定した。3つ未満の位置が異なる、頻度NAおよびNB($NA \gg NB$)の、2つの同様のCDR3配列AおよびBについて、 $NA \cdot \mu \gg NB$ の場合（ここで、 μ は対応するカットオフエラー率である）、CDR3配列Bが除外される。テンプレートミックスIの「リファレンスフィルタリングされた」配列に対してこのフィルタリングアルゴリズムを適用することによって、合計22,369個の配列（これは281個の特有のCDR3変異体に対応した）が除去された。テンプレートミックスIIについて、39,920個の人工的な配列（

40

50

348個の特有のCDR3変異体)が同定された(表1)。従って、PCR増幅エラーフィルターの使用は、約80%の非真正の別個のCDR3配列のさらなる減少をもたらす。

【0084】

上記のフィルターを通過した配列のプールにおいて、本発明者らは、いくつかの高存在量CDR3変異体を同定し、これらは、多数の位置で、それらの最も類似するインプットテンプレート配列とは異なっていた。CDR3断片の多数の位置でのPCR置換および/またはindel突然変異の発生はシミュレーション実験によれば極めて稀であるため、それらのCDR3変異体は、人工産物の他の供給源から生じるはずである。興味深いことに、本発明者らは、これらの配列の一部は2つの別個のインプットテンプレートの断片から構成され、明確なブレイクポイントを示したことを見出し、それらをキメラと同定した。キメラ配列は、PCRの間の不完全なプライマー伸長またはテンプレートスイッチングから生じるPCR人工産物であり、モザイク様構造を形成する。この予想外のPCR人工産物を踏まえて、本発明者らは、コンピューター「モザイクフィルター」を開発した。このフィルタリングアルゴリズムを使用して、本発明者らは、テンプレート混合物IおよびIIにおいて、それぞれ、合計17および15個のキメラ配列を同定した。注目すべきことに、これらのCDR3キメラの一部は1000を超える配列コピー数を示し、このことは、フィルターについての本発明者らのアルゴリズムが、高存在量キメラCDR3配列を同定することができることを示している。

10

【0085】

フィルタリングアルゴリズムの適用は、たった4つの規定のTCR CDR3テンプレートのハイスループット配列決定によって生じた非真正の特有のCDR3配列の99.8%を排除した。それぞれ、たった62個および73個の人工的なCDR3配列が、全てのフィルターを通過した。これらの中で、2つの最も豊富なCDR3配列は、混合実験の両方において同一であった。最も考えられることとして、それらは、正確にブレイクポイントに配置された単一ヌクレオチド置換のためにフィルタリングを逃れたキメラ人工産物を示す。残っている間違いのCDR3の中で、それぞれ、85% (n=53)および75% (n=55)が単一読み取りであった。人工産物のこの小さな分率を排除するために、本発明者らは、TCR免疫レパトリーのハイストリンジェンシーデータ解析は、シングルコピーCDR3読み取りを除去する追加のフィルター(頻度閾値フィルター)を含むべきであることを提唱する。

20

【0086】

(表1)

30

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
TRAV-C	TRAV1Fo	TGCACGTACC AGACATCTGG	12	TGCACGTACCA GACATCTGG	12
	TRAV1Fi	AGGTCGTTTT TCTTCATTCC	13	GCCTCCCTCGC GCCATCAGAGG TCGTTTTTCTTC ATTCC	14
	TRAV2Fo	TCTGTAATCA CTCTGTGTCC	15	TCTGTAATCACT CTGTGTCC	15
	TRAV2Fi	AGGGACGATA CAACATGACC	16	GCCTCCCTCGC GCCATCAGAGG GACGATACAAC ATGACC	17
	TRAV3Fo	CTATTCAGTC TCTGGAAACC	18	CTATTCAGTCT CTGGAAACC	18
	TRAV3Fi	ATACATCACA GGGGATAACC	19	GCCTCCCTCGC GCCATCAGATA CATCACAGGGG ATAACC	20
	TRAV4Fo	TGTAGCCACA ACAACATTGC	21	TGTAGCCACAA CAACATTGC	21
	TRAV4Fi	AAAGTTACAA ACGAAGTGGC	22	GCCTCCCTCGC GCCATCAGAAA GTTACAAACGA AGTGGC	23
	TRAV5Fo	GCACTTACAC AGACAGCTCC	24	GCACTTACACA GACAGCTCC	24
	TRAV5Fi	TATGGACATG AAACAAGACC	25	GCCTCCCTCGC GCCATCAGTAT GGACATGAAAC AAGACC	26
	TRAV6Fo	GCAACTATAC AAACTATTCC	27	GCAACTATACA AACTATTCC	27
	TRAV6Fi	GTTTTCTTGC TACTCATACG	28	GCCTCCCTCGC GCCATCAGGTT TTCTTGCTACTC ATACG	29
	TRAV7Fo	TGCACGTACT CTGTCAGTCG	30	TGCACGTACTC TGTCAGTCG	30
	TRAV7Fi	GGATATGAGA AGCAGAAAGG	31	GCCTCCCTCGC GCCATCAGGGA TATGAGAAGCA GAAAGG	32
	TRAV8Fo	AATCTCTTCT GGTATGTSCA	33	AATCTCTTCTG GTATGTSCA	33
	TRAV8Fi	GGYTTTGAGG CTGAATTTA	34	GCCTCCCTCGC GCCATCAGGGY	35

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
				TTTGAGGCTGA ATTTA	
	TRAV9Fo	GTCCAATATC CTGGAGAAG G	36	GTCCAATATCC TGGAGAAGG	36
	TRAV9Fi	AACCACTTCT TTCCAATTGG	37	GCCTCCCTCGC GCCATCAGAAC CACTTCTTTCCA CTTGG	38
	TRAV10Fo	AATGCAATTA TACAGTGAGC	39	AATGCAATTATA CAGTGAGC	39
	TRAV10Fi	TGAGAACACA AAGTCGAACG	40	GCCTCCCTCGC GCCATCAGTGA GAACACAAAGT CGAACG	41
	TRAV11Fo	TCTTAATTGTA CTTATCAGG	42	TCTTAATTGTAC TTATCAGG	42
	TRAV11Fi	TCAATCAAGC CAGAAGGAG C	43	GCCTCCCTCGC GCCATCAGTCA ATCAAGCCAGA AGGAGC	44
	TRAV12Fo	TCAGTGTTCC AGAGGGAGC C	45	TCAGTGTTCCA GAGGGAGCC	45
	TRAV12Fi	ATGGAAGGTT TACAGCACAG	46	GCCTCCCTCGC GCCATCAGATG GAAGGTTTACA GCACAG	47
	TRAV13Fo	ACCCTGAGTG TCCAGGAGG G	48	ACCCTGAGTGT CCAGGAGGG	48
	TRAV13Fi	TTATAGACAT TCGTTCAAAT	49	GCCTCCCTCGC GCCATCAGTTA TAGACATTCGT TCAAAT	50
	TRAV14Fo	TGGACTGCAC ATATGACACC	51	TGGACTGCACA TATGACACC	51
	TRAV14Fi	CAGCAAAATG CAACAGAAGG	52	GCCTCCCTCGC GCCATCAGCAG CAAAATGCAAC AGAAGG	53
	TRAV16Fo	AGCTGAAGTG CAACTATTCC	54	AGCTGAAGTGC AACTATTCC	54
	TRAV16Fi	TCTAGAGAGA GCATCAAAGG	55	GCCTCCCTCGC GCCATCAGTCT AGAGAGAGCAT CAAAGG	56
	TRAV17Fo	AATGCCACCA TGAAGTGCAG	57	AATGCCACCAT GAAGTGCAG	57
	TRAV17Fi	GAAAGAGAGA AACACAGTGG	58	GCCTCCCTCGC GCCATCAGGAA	59

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
				AGAGAGAAACA CAGTGG	
	TRAV18Fo	GCTCTGACAT TAACTGCAC	60	GCTCTGACATT AACTGCAC	60
	TRAV18Fi	CAGGAGACG GACAGCAGA GG	61	GCCTCCCTCGC GCCATCAGCAG GAGACGGACAG CAGAGG	62
	TRAV19Fo	ATGTGACCTT GGACTGTGTG	63	ATGTGACCTTG GACTGTGTG	63
	TRAV19Fi	GAGCAAAATG AAATAAGTGG	64	GCCTCCCTCGC GCCATCAGGAG CAAAATGAAAT AAGTGG	65
	TRAV20Fo	ACTGCAGTTA CACAGTCAGC	66	ACTGCAGTTAC ACAGTCAGC	66
	TRAV20Fi	AGAAAGAAAG GCTAAAAGCC	67	GCCTCCCTCGC GCCATCAGAGA AAGAAAGGCTA AAAGCC	68
	TRAV21Fo	ACTGCAGTTT CACTGATAGC	69	ACTGCAGTTTC ACTGATAGC	69
	TRAV21Fi	CAAGTGGAAG ACTTAATGCC	70	GCCTCCCTCGC GCCATCAGCAA GTGGAAGACTT AATGCC	71
	TRAV22Fo	GGGAGCCAAT TCCACGCTGC	72	GGGAGCCAATT CCACGCTGC	72
	TRAV22Fi	ATGGAAGATT AAGCGCCAC G	73	GCCTCCCTCGC GCCATCAGATG GAAGATTAAGC GCCACG	74
	TRAV23Fo	ATTTCAATTAT AACTGTGC	75	ATTTCAATTATA AACTGTGC	75
	TRAV23Fi	AAGGAAGATT CACAATCTCC	76	GCCTCCCTCGC GCCATCAGAAG GAAGATTCACA ATCTCC	77
	TRAV24Fo	GCACCAATTT CACCTGCAGC	78	GCACCAATTTT ACCTGCAGC	78
	TRAV24Fi	AGGACGAATA AGTGCCACTC	79	GCCTCCCTCGC GCCATCAGAGG ACGAATAAGTG CCACTC	80
	TRAV25Fo	TCACCACGTA CTGCAATTCC	81	TCACCACGTAC TGCAATTCC	81
	TRAV25Fi	AGACTGACAT TTCAGTTTGG	82	GCCTCCCTCGC GCCATCAGAGA CTGACATTTCA GTTTGG	83
	TRAV26Fo	TCGACAGATT	84	TCGACAGATTC	84

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
		CMCTCCCAGG		MCTCCCAGG	
	TRAV26Fi	GTCCAGYACC TTGATCCTGC	85	GCCTCCCTCGC GCCATCAGGTC CAGYACCTTGA TCCTGC	86
	TRAV27Fo	CCTCAAGTGT TTTTTCCAGC	87	CCTCAAGTGTT TTTTCCAGC	87
	TRAV27Fi	GTGACAGTAG TTACGGGTGG	88	GCCTCCCTCGC GCCATCAGGTG ACAGTAGTTAC GGGTGG	89
	TRAV29Fo	CAGCATGTTT GATTATTTCC	90	CAGCATGTTTG ATTATTTCC	90
	TRAV29Fi	ATCTATAAGT TCCATTAAGG	91	GCCTCCCTCGC GCCATCAGATC TATAAGTTCCAT TAAGG	92
	TRAV30Fo	CTCCAAGGCT TTATATTCTG	93	CTCCAAGGCTT TATATTCTG	93
	TRAV30Fi	ATGATATTAC TGAAGGGTG G	94	GCCTCCCTCGC GCCATCAGATG ATATTACTGAA GGGTGG	95
	TRAV34Fo	ACTGCACGTC ATCAAAGACG	96	ACTGCACGTCA TCAAAGACG	96
	TRAV34Fi	TTGATGATGC TACAGAAAGG	97	GCCTCCCTCGC GCCATCAGTTG ATGATGCTACA GAAAGG	98
	TRAV35Fo	TGAACTGCAC TTCTTCAAGC	99	TGAACTGCACT TCTTCAAGC	99
	TRAV35Fi	CTTGATAGCC TTATATAAGG	100	GCCTCCCTCGC GCCATCAGCTT GATAGCCTTAT ATAAGG	101
	TRAV36Fo	TCAATTGCAG TTATGAAGTG	102	TCAATTGCAGT TATGAAGTG	102
	TRAV36Fi	TTTATGCTAA CTTCAAGTGG	103	GCCTCCCTCGC GCCATCAGTTT ATGCTAACTTC AAGTGG	104
	TRAV38Fo	GCACATATGA CACCAGTGAG	105	GCACATATGAC ACCAGTGAG	105
	TRAV38Fi	TCGCCAAGAA GCTTATAAGC	106	GCCTCCCTCGC GCCATCAGTCG CCAAGAAGCTT ATAAGC	107
	TRAV39Fo	TCTACTGCAA TTATTCAACC	108	TCTACTGCAATT ATTCAACC	108
	TRAV39Fi	CAGGAGGGA	109	GCCTCCCTCGC	110

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
		CGATTAATGG C		GCCATCAGCAG GAGGGACGATT AATGGC	
	TRAV40Fo	TGAACTGCAC ATACACATCC	111	TGAACTGCACA TACACATCC	111
	TRAV40Fi	ACAGCAAAAA CTTCGGAGGC	112	GCCTCCCTCGC GCCATCAGACA GCAAAAACTTC GGAGGC	113
	TRAV41Fo	AACTGCAGTT ACTCGGTAGG	114	AACTGCAGTTA CTCGGTAGG	114
	TRAV41Fi	AAGCATGGAA GATTAATTGC	115	GCCTCCCTCGC GCCATCAGAAG CATGGAAGATT AATTGC	116
	TRACRo	GCAGACAGAC TTGTCACTGG	117	GCAGACAGACT TGTCACCTGG	117
	TRACRi	AGTCTCTCAG CTGGTACACG	118	GCCTTGCCAGC CCGCTCAGAGT CTCTCAGCTGG TACACG	119
TRBV-C	TRBV1Fo	AATGAAACGT GAGCATCTGG	120	AATGAAACGTG AGCATCTGG	120
	TRBV1Fi	CATTGAAAAC AAGACTGTGC	121	GCCTCCCTCGC GCCATCAGCAT TGAAAACAAGA CTGTGC	122
	TRBV2Fo	GTGTCCCAT CTCTAATCAC	123	GTGTCCCATC TCTAATCAC	123
	TRBV2Fi	TGAAATCTCA GAGAAGTCTG	124	GCCTCCCTCGC GCCATCAGTGA AATCTCAGAGA AGTCTG	125
	TRBV3Fo	TATGTATTGG TATAAACAGG	126	TATGTATTGGTA TAAACAGG	126
	TRBV3Fi	CTCTAAGAAA TTTCTGAAGA	127	GCCTCCCTCGC GCCATCAGCTC TAAGAAATTTCT GAAGA	128
	TRBV4Fo	GTCTTTGAAA TGTGAACAAC	129	GTCTTTGAAAT GTGAACAAC	129
	TRBV4Fi	GGAGCTCATG TTTGTCTACA	130	GCCTCCCTCGC GCCATCAGGGA GCTCATGTTTG TCTACA	131
	TRBV5Fo	GATCAAAACG AGAGGACAG C	132	GATCAAAACGA GAGGACAGC	132

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
	TRBV5aFi	CAGGGGCCC CAGTTTATCT T	133	GCCTCCCTCGC GCCATCAGCAG GGGCCCCAGTT TATCTT	134
	TRBV5bFi	GAAACARAGG AAACTTCCCT	135	GCCTCCCTCGC GCCATCAGGAA ACARAGGAAAC TTCCCT	136
	TRBV6aFo	GTGTGCCCAG GATATGAACC	137	GTGTGCCCAGG ATATGAACC	137
	TRBV6bFo	CAGGATATGA GACATAATGC	138	CAGGATATGAG ACATAATGC	138
	TRBV6aFi	GGTATCGACA AGACCCAGG C	139	GCCTCCCTCGC GCCATCAGGGT ATCGACAAGAC CCAGGC	140
	TRBV6bFi	TAGACAAGAT CTAGGACTGG	141	GCCTCCCTCGC GCCATCAGTAG ACAAGATCTAG GACTGG	142
	TRBV7Fo	CTCAGGTGTG ATCCAATTTTC	143	CTCAGGTGTGA TCCAATTTTC	143
	TRBV7aFi	TCTAATTTACT TCCAAGGCA	144	GCCTCCCTCGC GCCATCAGTCT AATTTACTTCCA AGGCA	145
	TRBV7bFi	TCCCAGAGTG ATGCTCAACG	146	GCCTCCCTCGC GCCATCAGTCC CAGAGTGATGC TCAACG	147
	TRBV7cFi	ACTTACTTCA ATTATGAAGC	148	GCCTCCCTCGC GCCATCAGACT TACTTCAATTAT GAAGC	149
	TRBV7dFi	CCAGAATGAA GCTCAACTAG	150	GCCTCCCTCGC GCCATCAGCCA GAATGAAGCTC AACTAG	151
	TRBV9Fo	GAGACCTCTC TGTGTAAGG	152	GAGACCTCTCT GTGTAAGG	152
	TRBV9Fi	CTCATTCAGT ATTATAATGG	153	GCCTCCCTCGC GCCATCAGCTC ATTCAGTATTAT AATGG	154
	TRBV10Fo	GGAATCACCC AGAGCCCAAG	155	GGAATCACCCA GAGCCCAAG	155
	TRBV10Fi	GACATGGGCT GAGGCTGATC	156	GCCTCCCTCGC GCCATCAGGAC ATGGGCTGAGG CTGATC	157
	TRBV11Fo	CCTAAGGATC	158	CCTAAGGATCG	158

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
	TRBV11Fi	GATTTTCTGC	159	ATTTTCTGC	160
		ACTCTCAAGA TCCAGCCTGC		GCCTCCCTCGC GCCATCAGACT CTCAAGATCCA GCCTGC	
	TRBV12Fo	AGGTGACAGA GATGGGACAA	161	AGGTGACAGAG ATGGGACAA	161
	TRBV12aFi	TGCAGGGACT GGAATTGCTG	162	GCCTCCCTCGC GCCATCAGTGC AGGGACTGGAA TTGCTG	163
	TRBV12bFi	GTACAGACAG ACCATGATGC	164	GCCTCCCTCGC GCCATCAGGTA CAGACAGACCA TGATGC	165
	TRBV13Fo	CTATCCTATC CCTAGACACG	166	CTATCCTATCC CTAGACACG	166
	TRBV13Fi	AAGATGCAGA GCGATAAAGG	167	GCCTCCCTCGC GCCATCAGAAG ATGCAGAGCGA TAAAGG	168
	TRBV14Fo	AGATGTGACC CAATTTCTGG	169	AGATGTGACCC AATTTCTGG	169
	TRBV14Fi	AGTCTAAACA GGATGAGTCC	170	GCCTCCCTCGC GCCATCAGAGT CTAAACAGGAT GAGTCC	171
	TRBV15Fo	TCAGACTTTG AACCATAACG	172	TCAGACTTTGA ACCATAACG	172
	TRBV15Fi	AAAGATTTTA ACAATGAAGC	173	GCCTCCCTCGC GCCATCAGAAA GATTTTAAACAAT GAAGC	174
	TRBV16Fo	TATTGTGCCC CAATAAAAGG	175	TATTGTGCCCC AATAAAAGG	175
	TRBV16Fi	AATGTCTTTG ATGAAACAGG	176	GCCTCCCTCGC GCCATCAGAAT GTCTTTGATGA AACAGG	177
	TRBV17Fo	ATCCATCTTC TGGTCACATG	178	ATCCATCTTCT GGTCACATG	178
	TRBV17Fi	AACATTGCAG TTGATTCAGG	179	GCCTCCCTCGC GCCATCAGAAC ATTGCAGTTGA TTCAGG	180
	TRBV18Fo	GCAGCCCAAT GAAAGGACAC	181	GCAGCCCAATG AAAGGACAC	181
	TRBV18Fi	AATATCATAG ATGAGTCAGG	182	GCCTCCCTCGC GCCATCAGAAT ATCATAGATGA GTCAGG	183

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
	TRBV19Fo	TGAACAGAAT TTGAACCACG	184	TGAACAGAATT TGAACCACG	184
	TRBV19Fi	TTTCAGAAAG GAGATATAGC	185	GCCTCCCTCGC GCCATCAGTTT CAGAAAGGAGA TATAGC	186
	TRBV20Fo	TCGAGTGCCG TTCCCTGGAC	187	TCGAGTGCCGT TCCCTGGAC	187
	TRBV20Fi	GATGGCAACT TCCAATGAGG	188	GCCTCCCTCGC GCCATCAGGAT GGCAACTTCCA ATGAGG	189
	TRBV21Fo	GCAAAGATGG ATTGTGTTCC	190	GCAAAGATGGA TTGTGTTCC	190
	TRBV21Fi	CGCTGGAAGA AGAGCTCAAG	191	GCCTCCCTCGC GCCATCAGCGC TGGAAGAAGAG CTCAAG	192
	TRBV23Fo	CATTTGGTCA AAGGAAAAGG	193	CATTTGGTCAA AGGAAAAGG	193
	TRBV23Fi	GAATGAACAA GTTCTTCAAG	194	GCCTCCCTCGC GCCATCAGGAA TGAACAAGTTC TTCAAG	195
	TRBV24Fo	ATGCTGGAAT GTTCTCAGAC	196	ATGCTGGAATG TTCTCAGAC	196
	TRBV24Fi	GTCAAAGATA TAAACAAAGG	197	GCCTCCCTCGC GCCATCAGGTC AAAGATATAAA CAAAGG	198
	TRBV25Fo	CTCTGGAATG TTCTCAAACC	199	CTCTGGAATGT TCTCAAACC	199
	TRBV25Fi	TAATTCCACA GAGAAGGGA G	200	GCCTCCCTCGC GCCATCAGTAA TTCCACAGAGA AGGGAG	201
	TRBV26Fo	CCCAGAATAT GAATCATGTT	202	CCCAGAATATG AATCATGTT	202
	TRBV26Fi	ATTCACCTGG CACTGGGAG C	203	GCCTCCCTCGC GCCATCAGATT CACCTGGCACT GGGAGC	204
	TRBV27Fo	TTGTTCTCAG AATATGAACC	205	TTGTTCTCAGA ATATGAACC	205
	TRBV27Fi	TGAGGTGACT GATAAGGGAG	206	GCCTCCCTCGC GCCATCAGTGA GGTGACTGATA AGGGAG	207
	TRBV28Fo	ATGTGTCCAG GATATGGACC	208	ATGTGTCCAGG ATATGGACC	208
	TRBV28Fi	AAAAGGAGAT	209	GCCTCCCTCGC	210

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
		ATTCCTGAGG		GCCATCAGAAA AGGAGATATTC CTGAGG	
	TRBV29Fo	TCACCATGAT GTTCTGGTAC	211	TCACCATGATG TTCTGGTAC	211
	TRBV29Fi	CTGGACAGAG CCTGACACTG	212	GCCTCCCTCGC GCCATCAGCTG GACAGAGCCTG AACTG	213
	TRBV30Fo	TGTGGAGGG AACATCAAAC C	214	TGTGGAGGGAA CATCAAACC	214
	TRBV30Fi	TTCTACTCCG TTGGTATTGG	215	GCCTCCCTCGC GCCATCAGTTC TACTCCGTTGG TATTGG	216
	TRBCRo	GTGTGGCCTT TTGGGTGTGG	217	GTGTGGCCTTT TGGGTGTGG	217
	TRBCRi	TCTGATGGCT CAAACACAGC	218	GCCTTGCCAGC CCGCTCAGTCT GATGGCTCAAA CACAGC	219
TRDV-C	TRDV1Fo	TGTATGAAAC AAGTTGGTGG	220	TGTATGAAACA AGTTGGTGG	220
	TRDV1Fi	CAGAATGCAA AAAGTGGTCG	221	GCCTCCCTCGC GCCATCAGCAG AATGCAAAAAG TGGTCG	222
	TRDV2Fo	ATGAAAGGAG AAGCGATCGG	223	ATGAAAGGAGA AGCGATCGG	223
	TRDV2Fi	TGGTTTCAAA GACAATTTCC	224	GCCTCCCTCGC GCCATCAGTGG TTTCAAAGACA ATTTCC	225
	TRDV3Fo	GACACTGTAT ATTCAAATCC	226	GACACTGTATA TTCAAATCC	226
	TRDV3Fi	GCAGATTTTA CTCAAGGACG	227	GCCTCCCTCGC GCCATCAGGCA GATTTTACTCAA GGACG	228
	TRDCRo	AGACAAGCGA CATTTGTTCC	229	AGACAAGCGAC ATTTGTTCC	229
	TRDCRi	ACGGATGGTT TGGTATGAGG	230	GCCTTGCCAGC CCGCTCAGACG GATGGTTTGGT ATGAGG	231

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	配列	SEQ ID NO.
TRGV-C	TRGV1-5Fo	GGGTCATCTG CTGAAATCAC	232	GGGTCATCTGC TGAAATCAC	232
	TRGV1-5,8Fi	AGGAGGGGA AGGCCCCACA G	233	GCCTCCCTCGC GCCATCAGAGG AGGGGAAGGC CCCACAG	234
	TRGV8Fo	GGGTCATCAG CTGTAATCAC	235	GGGTCATCAGC TGTAATCAC	235
	TRGV5pFi	AGGAGGGGA AGACCCCACA G	236	GCCTCCCTCGC GCCATCAGAGG AGGGGAAGACC CCACAG	237
	TRGV9Fo	AGCCCGCCT GGAATGTGTG G	238	AGCCCGCCTGG AATGTGTGG	238
	TRGV9Fi	GCACTGTCAG AAAGGAATCC	239	GCCTCCCTCGC GCCATCAGGCA CTGTCAGAAAG GAATCC	240
	TRGV10Fo	AAGAAAAGTA TTGACATACC	241	AAGAAAAGTAT TGACATACC	241
	TRGV10Fi	ATATTGTCTC AACAAAATCC	242	GCCTCCCTCGC GCCATCAGATA TTGTCTCAACA AAATCC	243
	TRGV11Fo	AGAGTGCCCA CATATCTTGG	244	AGAGTGCCCA CATATCTTGG	244
	TRGV11Fi	GCTCAAGATT GCTCAGGTG G	245	GCCTCCCTCGC GCCATCAGGCT CAAGATTGCTC AGGTGG	246
	TRGCRo	GGATCCCAGA ATCGTGTTGC	247	GGATCCCAGAA TCGTGTTGC	247
	TRGCRI	GGTATGTTCC AGCCTTCTGG	248	GCCTTGCCAGC CCGCTCAGGGT ATGTTCCAGCC TTCTGG	249

10

20

30

座	プライマー名	配列	SEQ ID NO.	注文 (ordered)	SEQ ID NO.
IgHV-J	IgHV1aFo	AGTGAAGGTCTC CTGCAAGG	250	AGTGAAGGTCTC CTGCAAGG	250
	IgHV1bFo	AGTGAAGGTTTC CTGCAAGG	251	AGTGAAGGTTTC CTGCAAGG	251
	IgHV1aFi	AGTTCCAGGGCA GAGTCAC	252	GCCTCCCTCGCG CCATCAGAGTTC CAGGGCAGAGTC AC	253
	IgHV1bFi	AGTTTCAGGGCA GGGTCAC	254	GCCTCCCTCGCG CCATCAGAGTTT CAGGGCAGGGTC AC	255
	IgHV1cFi	AGTTCCAGGAAA GAGTCAC	256	GCCTCCCTCGCG CCATCAGAGTTC CAGGAAAGAGTC AC	257
	IgHV1dFi	AATTCCAGGACA GAGTCAC	258	GCCTCCCTCGCG CCATCAGAATTC CAGGACAGAGTC AC	259
	IgHV2Fo	TCTCTGGGTTCT CACTCAGC	260	TCTCTGGGTTCT CACTCAGC	260
	IgHV2Fi	AAGGCCCTGGAG TGGCTTGC	261	GCCTCCCTCGCG CCATCAGAAGGC CCTGGAGTGGCT TGC	262
	IgHV3aFo	TCCCTGAGACTC TCCTGTGC	263	TCCCTGAGACTC TCCTGTGC	263
	IgHV3bFo	CTCTCCTGTGCA GCCTCTGG	264	CTCTCCTGTGCA GCCTCTGG	264
	IgHV3cFo	GGTCCCTGAGAC TCTCCTGT	265	GGTCCCTGAGAC TCTCCTGT	265
	IgHV3dFo	CTGAGACTCTCC TGTGTAGC	266	CTGAGACTCTCC TGTGTAGC	266
	IgHV3aFi	CTCCAGGGAAGG GGCTGG	267	GCCTCCCTCGCG CCATCAGCTCCA GGGAAGGGGCT GG	268
	IgHV3bFi	GGCTCCAGGCAA GGGGCT	269	GCCTCCCTCGCG CCATCAGGGCTC CAGGCAAGGGGC T	270
	IgHV3cFi	ACTGGGTCCGCC AGGCTCC	271	GCCTCCCTCGCG CCATCAGACTGG GTCCGCCAGGCT CC	272
	IgHV3dFi	GAAGGGGCTGGA GTGGGT	273	GCCTCCCTCGCG CCATCAGGAAGG GGCTGGAGTGGG T	274

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	注文	SEQ ID NO.
	IgHV3eFi	AAAAGGTCTGGA GTGGGT	275	GCCTCCCTCGCG CCATCAGAAAAG GTCTGGAGTGGG T	276
	IgHV4Fo	AGACCCTGTCCC TCACCTGC	277	AGACCCTGTCCC TCACCTGC	277
	IgHV4Fi	AGGGVCTGGAGT GGATTGGG	278	GCCTCCCTCGCG CCATCAGAGGGV CTGGAGTGGATT GGG	279
	IgHV5Fo	GCGCCAGATGCC CGGGAAAG	280	GCGCCAGATGCC CGGGAAAG	280
	IgHV5Fi	GGCCASGTCACC ATCTCAGC	281	GCCTCCCTCGCG CCATCAGGGCCA SGTCACCATCTC AGC	282
	IgHV6Fo	CCGGGGACAGTG TCTCTAGC	283	CCGGGGACAGTG TCTCTAGC	283
	IgHV6Fi	GCCTTGAGTGGC TGGGAAGG	284	GCCTCCCTCGCG CCATCAGGCCTT GAGTGGCTGGGA AGG	285
	IgHV7Fo	GTTTCCTGCAAG GCTTCTGG	286	GTTTCCTGCAAG GCTTCTGG	286
	IgHV7Fi	GGCTTGAGTGGA TGGGATGG	287	GCCTCCCTCGCG CCATCAGGGCTT GAGTGGATGGGA TGG	288
	IgHJRo	ACCTGAGGAGAC GGTGACC	289	ACCTGAGGAGAC GGTGACC	289
	IgHJ1Ri	CAGTGCTGGAAG TATTCAGC	290	GCCTTGCCAGCC CGCTCAGCAGTG CTGGAAGTATTC AGC	291
	IgHJ2Ri	AGAGATCGAAGT ACCAGTAG	292	GCCTTGCCAGCC CGCTCAGAGAGA TCGAAGTACCAG TAG	293
	IgHJ3Ri	CCCCAGATATCA AAAGCATC	294	GCCTTGCCAGCC CGCTCAGCCCCA GATATCAAAAGC ATC	295
	IgHJ4Ri	GGCCCCAGTAGT CAAAGTAG	296	GCCTTGCCAGCC CGCTCAGGGCCC CAGTAGTCAAAG TAG	297
	IgHJ5Ri	CCCAGGGGTCGA ACCAGTTG	298	GCCTTGCCAGCC CGCTCAGCCCAG GGGTCGAACCAG TTG	299

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	注文	SEQ ID NO.
	IgHJ6Ri	CCCAGACGTCCA TGTAGTAG	300	GCCTTGCCAGCC CGCTCAGCCCAG ACGTCCATGTAG TAG	301
IgKV-C	IgKV1Fo	TAGGAGACAGAG TCACCATC	302	TAGGAGACAGAG TCACCATC	302
	IgKV1Fi	TTCAGYGRCACT GGATCTGG	303	GCCTCCCTCGCG CCATCAGTTCAG YGRCACTGGATC TGG	304
	IgKV2Fo	GGAGAGCCGGC CTCCATCTC	305	GGAGAGCCGGC CTCCATCTC	305
	IgKV2aFi	TGGTACCTGCAG AAGCCAGG	306	GCCTCCCTCGCG CCATCAGTGGTA CCTGCAGAAGCC AGG	307
	IgKV2bFi	CTTCAGCAGAGG CCAGGCCA	308	GCCTCCCTCGCG CCATCAGCTTCA GCAGAGGCCAGG CCA	309
	IgKV3-7Fo	GCCTGGTACCAG CAGAAACC	310	GCCTGGTACCAG CAGAAACC	310
	IgKV3Fi	GCCAGGTTCACT GGCAGTGG	311	GCCTCCCTCGCG CCATCAGGCCAG GTTCACTGGCAG TGG	312
	IgKV6-7Fi	TCGAGGTTCACT GGCAGTGG	313	GCCTCCCTCGCG CCATCAGTCGAG GTTCACTGGCAG TGG	314
	IgKV4-5Fi	GACCGATTCAGT GGCAGCGG	315	GCCTCCCTCGCG CCATCAGGACCG ATTCAGTGGCAG CGG	316
	IgKCRo	TTCAACTGCTCAT CAGATGG	317	TTCAACTGCTCAT CAGATGG	317
IgLV-C	IgKCRi	ATGAAGACAGAT GGTGCAGC	318	GCCTTGCCAGCC CGCTCAGATGAA GACAGATGGTGC AGC	319
	IgLV1aFo	GGGCAGAGGGTC ACCATCTC	320	GGGCAGAGGGTC ACCATCTC	320
	IgLV1bFo	GGACAGAAGGTC ACCATCTC	321	GGACAGAAGGTC ACCATCTC	321
	IgLV1aFi	TGGTACCAGCAG CTCCCAGG	322	GCCTCCCTCGCG CCATCAGTGGTA CCAGCAGCTCCC AGG	323

10

20

30

40

座	プライマー名	配列	SEQ ID NO.	注文	SEQ ID NO.
	IgLV1bFi	TGGTACCAGCAG CTTCCAGG	324	GCCTCCCTCGCG CCATCAGTGGTA CCAGCAGCTTCC AGG	325
	IgLV2Fo	CTGCACTGGAAC CAGCAGTG	326	CTGCACTGGAAC CAGCAGTG	326
	IgLV2Fi	TCTCTGGCTCCA AGTCTGGC	327	GCCTCCCTCGCG CCATCAGTCTCT GGCTCCAAGTCT GGC	328
	IgLV3aFo	ACCAGCAGAAGC CAGGCCAG	329	ACCAGCAGAAGC CAGGCCAG	329
	IgLV3bFo	GAAGCCAGGACA GGCCCCTG	330	GAAGCCAGGACA GGCCCCTG	330
	IgLV3aFi	CTGAGCGATTCT CTGGCTCC	331	GCCTCCCTCGCG CCATCAGCTGAG CGATTCTCTGGC TCC	332
	IgLV3bFi	TTCTCTGGGTCC ACCTCAGG	333	GCCTCCCTCGCG CCATCAGTTCTCT GGGTCCACCTCA GG	334
	IgLV3cFi	TTCTCTGGCTCC AGCTCAGG	335	GCCTCCCTCGCG CCATCAGTTCTCT GGGTCCAGCTCA GG	336
	IgLV4Fo	TCGGTCAAGCTC ACCTGCAC	337	TCGGTCAAGCTC ACCTGCAC	337
	IgLV4Fi	GGGCTGACCGCT ACCTCACC	358	GCCTCCCTCGCG CCATCAGGGGCT GACCGCTACCTC ACC	338
	IgLV5Fo	CAGCCTGTGCTG ACTCAGCC	339	CAGCCTGTGCTG ACTCAGCC	339
	IgLV5Fi	CCAGCCGCTTCT CTGGATCC	340	GCCTCCCTCGCG CCATCAGCCAGC CGCTTCTCTGGA TCC	341
	IgLV6Fo	CCATCTCCTGCA CCCGCAGC	342	CCATCTCCTGCA CCCGCAGC	342
	IgLV7-8Fo	TCCCCWGGAGG GACAGTCAC	343	TCCCCWGGAGG GACAGTCAC	343
	IgLV9,11Fo	CTCMCCTGCACC CTGAGCAG	344	CTCMCCTGCACC CTGAGCAG	344
	IgLV10Fo	AGACCGCCACAC TCACCTGC	345	AGACCGCCACAC TCACCTGC	345
	IgLV6,8Fi	CTGATCGSTTCTC TGGCTCC	346	GCCTCCCTCGCG CCATCAGCTGAT CGSTTCTCTGGC TCC	347
	IgLV7Fi	CTGCCCCGGTTCT	348	CTGCCCCGGTTCT	348

10

20

30

40

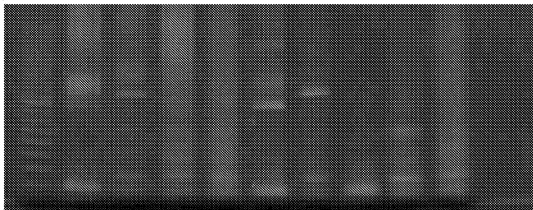
座	プライマー名	配列	SEQ ID NO.	注文	SEQ ID NO.
		CAGGCTCC		CAGGCTCC	
	IgLV9Fi	ATCCAGGAAGAG GATGAGAG	349	GCCTCCCTCGCG CCATCAGATCCA GGAAGAGGATGA GAG	350
	IgLV10-11Fi	CTCCAGCCTGAG GACGAGGC	351	GCCTCCCTCGCG CCATCAGCTCCA GCCTGAGGACGA GGC	352
	IgLC1-7Ro	GCTCCCGGGTAG AAGTCACT	353	GCTCCCGGGTAG AAGTCACT	353
	IgLC1-7Ri	AGTGTGGCCTTG TTGGCTTG	354	GCCTTGCCAGCC CGCTCAGAGTGT GGCCTTGTTGGC TTG	355
	454A	GCCTCCCTCGCG CCATCAG	356	GCCTCCCTCGCG CCATCAG	356
	454B	GCCTTGCCAGCC CGCTCAG	357	GCCTTGCCAGCC CGCTCAG	357

10

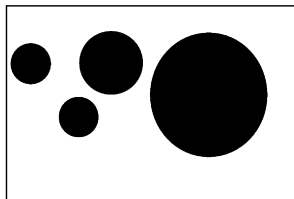
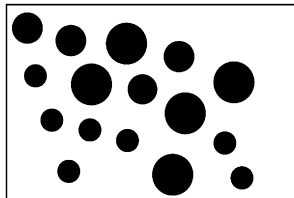
20

【図 1】

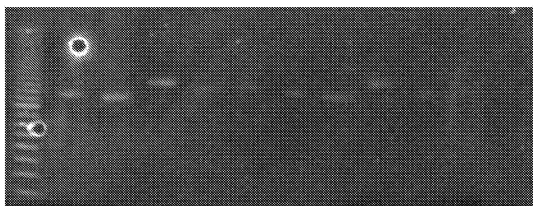
1 2 3 4 5 6 7 8 9 10 11

*a*

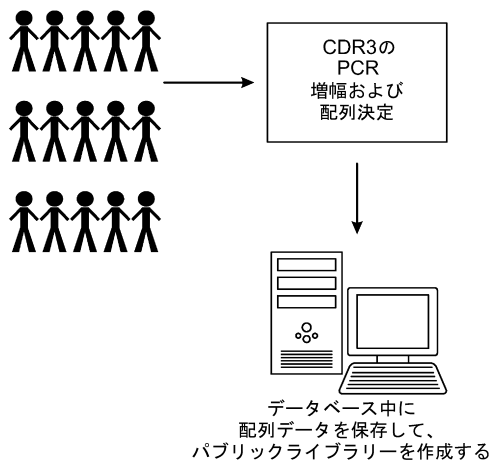
【図 2】

*a**b*

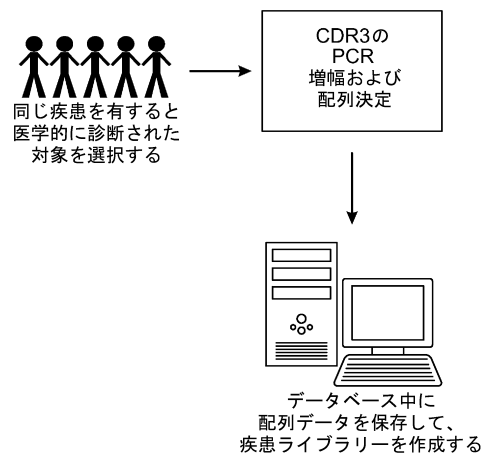
1 2 3 4 5 6 7 8 9 10 11 12

*b*

【図 3】

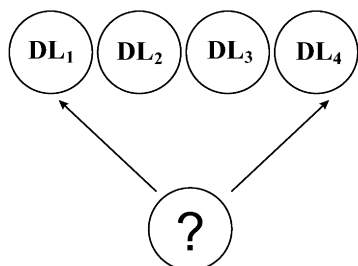


【図 4】

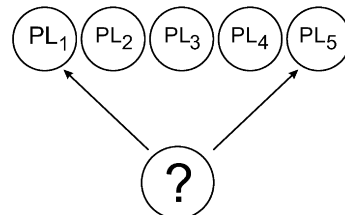


【図 5】

CDR3	読み取り カウント	パーセンテージ	共有される？
CDR3 ₁	120345	0.0602%	される
CDR3 ₂	1542	0.0008%	されない
CDR3 ₃	4530	0.0023%	されない
CDR3 ₄	8762	0.0044%	される
CDR3 ₅	689	0.0003%	されない
CDR3 ₆	325	0.0002%	されない
CDR3 ₇	8452	0.0042%	される
CDR3 ₈	23540	0.0118%	される
CDR3 ₉	3841	0.0019%	されない
CDR3 _n	20	0.0000%	されない
合計			0.0805495%

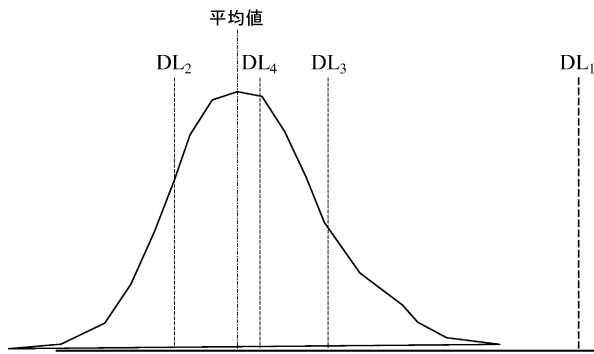


【図 6】



CDR3	読み取り カウント	パーセンテージ	共有される？
CDR3 ₁	120345	0.0602%	されない
CDR3 ₂	1542	0.0008%	される
CDR3 ₃	4530	0.0023%	される
CDR3 ₄	8762	0.0044%	されない
CDR3 ₅	689	0.0003%	される
CDR3 ₆	325	0.0002%	される
CDR3 ₇	8452	0.0042%	されない
CDR3 ₈	23540	0.0118%	されない
CDR3 ₉	3841	0.0019%	されない
CDR3 _n	20	0.0000%	される
合計			0.003553%

【図 7】



【配列表】

0006460343000001.app

フロントページの続き

- (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 100205707
弁理士 小寺 秀紀
- (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 ワン チュンリン
アメリカ合衆国 カリフォルニア州 メンロ パーク ロブレ アベニュー 735 #3
- (72)発明者 ハン ジャン
アメリカ合衆国 アラバマ州 ハンツビル ドニゴール ドライブ サウスイースト 7712

審査官 田名部 拓也

- (56)参考文献 米国特許出願公開第2010/0021896(US,A1)
米国特許出願公開第2012/0183969(US,A1)
特表2012-508011(JP,A)
国際公開第2011/139371(WO,A1)

(58)調査した分野(Int.Cl., DB名)

C12N 15/09

CAPLUS/MEDLINE/EMBASE/BIOSIS/WPIDS(STN)

JSTPLUS/JMEDPLUS/JST7580(JDreamIII)