

(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2020-532323
(P2020-532323A)

(43) 公表日 令和2年11月12日(2020.11.12)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/6872 (2018.01)	C 1 2 Q 1/6872 Z	4 B 0 6 3
C 1 2 Q 1/6862 (2018.01)	C 1 2 Q 1/6862 Z N A Z	4 B 0 6 5
C 1 2 Q 1/6813 (2018.01)	C 1 2 Q 1/6813 Z	
C 1 2 N 5/0783 (2010.01)	C 1 2 N 5/0783	
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/09 Z	

審査請求 未請求 予備審査請求 未請求 (全 214 頁) 最終頁に続く

(21) 出願番号 特願2020-534819 (P2020-534819)
 (86) (22) 出願日 平成30年9月5日 (2018.9.5)
 (85) 翻訳文提出日 令和2年4月28日 (2020.4.28)
 (86) 国際出願番号 PCT/US2018/049614
 (87) 国際公開番号 W02019/050994
 (87) 国際公開日 平成31年3月14日 (2019.3.14)
 (31) 優先権主張番号 62/644,191
 (32) 優先日 平成30年3月16日 (2018.3.16)
 (33) 優先権主張国・地域又は機関 米国 (US)
 (31) 優先権主張番号 62/579,734
 (32) 優先日 平成29年10月31日 (2017.10.31)
 (33) 優先権主張国・地域又は機関 米国 (US)

(71) 出願人 518213307
 グリットストーン オンコロジー インコーポレイテッド
 アメリカ合衆国 94608 カリフォルニア州 エメリービル ホートン ストリート 5959 스위트 300
 (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

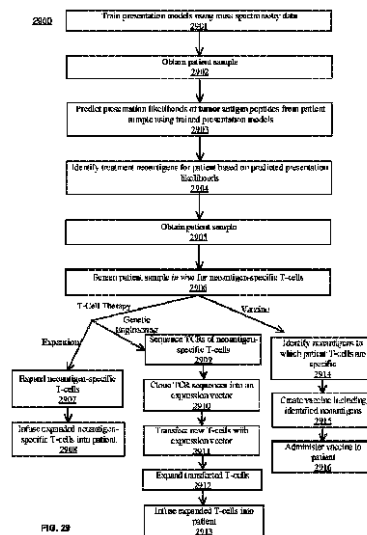
最終頁に続く

(54) 【発明の名称】 T細胞療法用の新生抗原の特定法

(57) 【要約】

対象の腫瘍細胞の表面上に提示される可能性の高い少なくとも1種類の新生抗原に対して抗原特異的なT細胞を特定するための方法を提供する。腫瘍新生抗原のペプチド配列は、対象の腫瘍細胞をシーケンシングすることにより取得される。新生抗原が対象の腫瘍細胞の表面上のMHCアレルによって提示される尤度をそれぞれが表す、腫瘍新生抗原についての提示尤度を生成するために、機械学習させた提示モデルにこれらのペプチド配列を入力する。新生抗原のサブセットが、提示尤度に基づいて選択される。サブセット中の新生抗原のうちの少なくとも1つに対して抗原特異的なT細胞が特定される。T細胞療法で使用するためにこれらのT細胞を増殖させることができる。これらの特定されたT細胞のTCRをシーケンシングし、新たなT細胞にクローニングしてT細胞療法に使用することもできる。

【選択図】 図 2 9



【特許請求の範囲】

【請求項 1】

対象の 1 つ以上の腫瘍細胞に由来し腫瘍細胞の表面上に提示される可能性の高い少なくとも 1 種類の新生抗原に対して抗原特異的な 1 つ以上の T 細胞を特定するための方法であって、

前記対象の前記腫瘍細胞及び正常細胞からエクソーム、トランスクリプトーム、または全ゲノムのヌクレオチドシーケンシングデータのうちの少なくとも 1 つを取得する工程であって、前記ヌクレオチドシーケンシングデータが、前記腫瘍細胞由来の前記ヌクレオチドシーケンシングデータと前記正常細胞由来の前記ヌクレオチドシーケンシングデータを比較することにより特定された新生抗原のセットのそれぞれのペプチド配列を表すデータを取得するために用いられ、各新生抗原のペプチド配列が、前記ペプチド配列を前記対象の前記正常細胞から特定された対応する野生型のペプチド配列とは異なるものとする少なくとも 1 つの変化を含む、前記工程；

10

前記新生抗原のそれぞれの前記ペプチド配列を対応する数値ベクトルにコード化する工程であって、各数値ベクトルが、前記ペプチド配列を構成する複数のアミノ酸、及び前記ペプチド配列における前記アミノ酸の位置のセットに関する情報を含む、前記工程；

前記新生抗原のセットについて提示尤度のセットを生成するために、機械学習させた提示モデルに前記数値ベクトルをコンピュータプロセッサを使用して入力する工程であって、前記セット内の各提示尤度が、対応する新生抗原が 1 つ以上の MHC アレルによって前記対象の前記腫瘍細胞の前記表面上に提示される尤度を表し、前記機械学習させた提示モデルが、

20

複数の試料における各試料について、前記試料中に存在するものとして特定された MHC アレルのセット内の少なくとも 1 つの MHC アレルに結合したペプチドの存在を測定する質量分析によって得られた標識、ならびに

前記試料のそれぞれについて、前記ペプチドを構成する複数のアミノ酸、及び前記ペプチドにおける前記アミノ酸の位置のセットに関する情報を含む数値ベクトルとしてコードされた訓練ペプチド配列

を含む訓練データセットに少なくとも基づいて特定される複数のパラメータと、

入力として受け取られた前記数値ベクトルと、前記数値ベクトル及び前記パラメータに基づいた出力として生成された前記提示尤度との間の関係を表す関数と

30

を含む、前記工程；

選択された新生抗原のセットを生成するために、前記新生抗原のセットのサブセットを、前記提示尤度のセットに基づいて選択する工程；

前記サブセット中の前記新生抗原のうちの少なくとも 1 つに対して抗原特異的である 1 つ以上の T 細胞を特定する工程；ならびに

前記 1 つ以上の特定された T 細胞を返す工程

を含む、前記方法。

【請求項 2】

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、

前記 1 つ以上の MHC アレルのそれぞれについて、前記ペプチド配列の特定の位置の特定のアミノ酸に基づいて前記 MHC アレルが前記新生抗原を提示するかどうかを示す依存性スコアを生成するために、前記機械学習させた提示モデルを前記新生抗原の前記ペプチド配列に適用すること

40

を含む、請求項 1 に記載の方法。

【請求項 3】

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、

各 MHC アレルについて、対応する MHC アレルが対応する新生抗原を提示する尤度を示す対応するアレル毎尤度を生成するために、前記依存性スコアを変換すること、及び

前記新生抗原の前記提示尤度を生成するために、前記アレル毎尤度を組み合わせることをさらに含む、請求項 2 に記載の方法。

50

【請求項 4】

前記依存性スコアを変換することが、前記新生抗原の提示を前記 1 つ以上の M H C アレルにわたって相互排他的なものとしてモデル化する、請求項 3 に記載の方法。

【請求項 5】

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、
前記提示尤度を生成するために、前記依存性スコアの組み合わせを変換することであって、前記依存性スコアの組み合わせを変換することが、前記新生抗原の提示を前記 1 つ以上の M H C アレル間で干渉するものとしてモデル化する、前記変換することをさらに含む、請求項 2 に記載の方法。

【請求項 6】

前記提示尤度のセットが、少なくとも 1 つ以上のアレル非相互作用特性によってさらに特定され、かつ、

前記アレル非相互作用特性に基づいて、前記対応する新生抗原のペプチド配列が提示されるかどうかを示す、前記アレル非相互作用特性についての依存性スコアを生成するために、前記機械学習させた提示モデルを前記アレル非相互作用特性に適用することをさらに含む、請求項 2 ~ 5 のいずれか 1 項に記載の方法。

【請求項 7】

前記 1 つ以上の M H C アレルにおける各 M H C アレルについての前記依存性スコアを、前記アレル非相互作用特性についての前記依存性スコアと組み合わせること、

対応する M H C アレルが対応する新生抗原を提示する尤度を示す、各 M H C アレルについてのアレル毎尤度を生成するために、各 M H C アレルについての前記組み合わせられた依存性スコアを変換すること、及び

前記提示尤度を生成するために、前記アレル毎尤度を組み合わせること
をさらに含む、請求項 6 に記載の方法。

【請求項 8】

前記 M H C アレルのそれぞれについての前記依存性スコアと、前記アレル非相互作用特性についての前記依存性スコアとを組み合わせること、及び

前記提示尤度を生成するために、前記組み合わせられた依存性スコアを変換すること
をさらに含む、請求項 6 に記載の方法。

【請求項 9】

前記 1 つ以上の M H C アレルが 2 つ以上の異なる M H C アレルを含む、請求項 1 ~ 8 のいずれか 1 項に記載の方法。

【請求項 10】

前記ペプチド配列が、アミノ酸 9 個以外の長さを有するペプチド配列を含む、請求項 1 ~ 9 のいずれか 1 項に記載の方法。

【請求項 11】

前記ペプチド配列をコード化する工程が、ワン・ホットエンコーディングスキームを用いて前記ペプチド配列をコード化することを含む、請求項 1 ~ 10 のいずれか 1 項に記載の方法。

【請求項 12】

前記複数の試料が、
(a) 単一の M H C アレルを発現するように操作された 1 つ以上の細胞株、
(b) 複数の M H C アレルを発現するように操作された 1 つ以上の細胞株、
(c) 複数の患者から得られた、または複数の患者に由来する 1 つ以上のヒト細胞株、
(d) 複数の患者から得られた新鮮なまたは凍結された腫瘍試料、及び
(e) 複数の患者から得られた新鮮なまたは凍結された組織試料
のうちの少なくとも 1 つを含む、請求項 1 ~ 11 のいずれか 1 項に記載の方法。

【請求項 13】

前記訓練データセットが、
(a) 前記ペプチドのうちの少なくとも 1 つについてのペプチド - M H C 結合親和性の測

10

20

30

40

50

定値に関連するデータ、及び

(b) 前記ペプチドのうちの少なくとも1つについてのペプチド - MHC 結合安定性の測定値に関連するデータ

のうちの少なくとも1つをさらに含む、請求項1～12のいずれか1項に記載の方法。

【請求項14】

前記提示尤度のセットが、RNA-seqまたは質量分析により測定される、前記対象における前記1つ以上のMHCアレルの少なくとも発現レベルによってさらに特定される、請求項1～13のいずれか1項に記載の方法。

【請求項15】

前記提示尤度のセットが、

(a) 前記新生抗原のセット内の新生抗原と前記1つ以上のMHCアレルとの間の予測される親和性、及び

(b) 前記新生抗原コード化ペプチド - MHC複合体の予測される安定性

のうちの少なくとも1つを含む特性によってさらに特定される、請求項1～14のいずれか1項に記載の方法。

【請求項16】

前記数値的尤度のセットが、

(a) そのソースタンパク質配列内の、前記新生抗原コード化ペプチド配列に隣接するC末端配列、及び

(b) そのソースタンパク質配列内の、前記新生抗原コード化ペプチド配列に隣接するN末端配列

のうちの少なくとも1つを含む特性によってさらに特定される、請求項1～15のいずれか1項に記載の方法。

【請求項17】

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記腫瘍細胞表面上に提示される尤度が増大している新生抗原を選択することを含む、請求項1～16のいずれか1項に記載の方法。

【請求項18】

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記対象において腫瘍特異的な免疫応答を誘導することができる尤度が増大している新生抗原を選択することを含む、請求項1～17のいずれか1項に記載の方法。

【請求項19】

前記選択された新生抗原のセットを選択する工程が、前記提示モデルに基づいて、選択されない新生抗原と比較してプロフェッショナル抗原提示細胞(APC)によってナイーブT細胞に提示されることができる尤度が増大している新生抗原を選択することを含み、任意で、前記APCが樹状細胞(DC)である、請求項1～18のいずれか1項に記載の方法。

【請求項20】

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して中枢性寛容または末梢性寛容によって阻害される尤度が減少している新生抗原を選択することを含む、請求項1～19のいずれか1項に記載の方法。

【請求項21】

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記対象において正常組織に対する自己免疫応答を誘導することができる尤度が減少している新生抗原を選択することを含む、請求項1～20のいずれか1項に記載の方法。

【請求項22】

前記1つ以上の腫瘍細胞が、肺癌、メラノーマ、乳癌、卵巣癌、前立腺癌、腎臓癌、胃

10

20

30

40

50

癌、結腸癌、精巣癌、頭頸部癌、膵臓癌、脳癌、B細胞リンパ腫、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、T細胞リンパ球性白血病、非小細胞肺癌、及び小細胞肺癌からなる群から選択される、請求項1～21のいずれか1項に記載の方法。

【請求項23】

前記選択された新生抗原のセットから個別化がんワクチンを構築するための出力を生成することをさらに含む、請求項1～22のいずれか1項に記載の方法。

【請求項24】

前記個別化がんワクチン用の出力が、前記選択された新生抗原のセットをコードした少なくとも1つのペプチド配列または少なくとも1つのヌクレオチド配列を含む、請求項23に記載の方法。

10

【請求項25】

前記機械学習させた提示モデルが、ニューラルネットワークモデルである、請求項1～24のいずれか1項に記載の方法。

【請求項26】

前記ニューラルネットワークモデルがMHCアレルについての複数のネットワークモデルを含み、それぞれのネットワークモデルが複数のMHCアレルのうちの対応するMHCアレルに割り当てられ、かつ1つ以上の層に配置された一連のノードを含む、請求項25に記載の方法。

【請求項27】

前記ニューラルネットワークモデルが、前記ニューラルネットワークモデルのパラメータを更新することによって訓練され、少なくとも2つのネットワークモデルの前記パラメータが、少なくとも1回の訓練イテレーションについて一緒に更新される、請求項26に記載の方法。

20

【請求項28】

前記機械学習させた提示モデルが、ノードの1つ以上の層を含むディープラーニングモデルである、請求項25～27のいずれか1項に記載の方法。

【請求項29】

前記1つ以上のT細胞を特定する工程が、前記1つ以上のT細胞を増殖させる条件下で前記1つ以上のT細胞を前記サブセット中の前記新生抗原のうちの1つ以上と共培養することを含む、請求項1～28のいずれか1項に記載の方法。

30

【請求項30】

前記1つ以上のT細胞を特定する工程が、前記T細胞とMHCマルチマーとの結合を可能とする条件下で前記1つ以上のT細胞を前記サブセット中の前記新生抗原のうちの1つ以上を含む前記MHCマルチマーと接触させることを含む、請求項1～29のいずれか1項に記載の方法。

【請求項31】

前記1つ以上の特定されたT細胞の1つ以上のT細胞受容体(TCR)を特定することをさらに含む、請求項1～30のいずれか1項に記載の方法。

【請求項32】

前記1つ以上のT細胞受容体を特定することが、前記1つ以上の特定されたT細胞のT細胞受容体配列をシーケンシングすることを含む、請求項31に記載の方法。

40

【請求項33】

請求項1～32のいずれか1項に記載の前記サブセット中の少なくとも1つの選択された新生抗原に対して抗原特異的な単離T細胞。

【請求項34】

前記1つ以上の特定されたT細胞受容体のうちの少なくとも1つを発現するように複数のT細胞を遺伝子操作する工程、

前記複数のT細胞を増殖させる条件下で前記複数のT細胞を培養する工程、及び前記増殖させたT細胞を前記対象に注入する工程

をさらに含む、請求項32に記載の方法。

50

【請求項 35】

前記 1 つ以上の特定された T 細胞受容体のうちの少なくとも 1 つを発現するように前記複数の T 細胞を遺伝子操作する工程が、

前記 1 つ以上の特定された T 細胞の T 細胞受容体配列を発現ベクターにクローニングすること、及び

前記複数の T 細胞のそれぞれに前記発現ベクターをトランスフェクトすることを含む、請求項 34 に記載の方法。

【請求項 36】

前記 1 つ以上の特定された T 細胞を増殖させる条件下で前記 1 つ以上の特定された T 細胞を培養する工程、及び

前記増殖させた T 細胞を前記対象に注入する工程をさらに含む、請求項 1 ~ 35 のいずれか 1 項に記載の方法。

【請求項 37】

前記サブセット中の前記新生抗原のうちの少なくとも 1 つに対して抗原特異的である前記 1 つ以上の T 細胞が、前記対象からの 5 ~ 30 mL の全血を使用して特定される、請求項 1 ~ 36 のいずれか 1 項に記載の方法。

【請求項 38】

前記新生抗原のサブセットが最大で 20 種類の新生抗原を含み、前記 1 つ以上の特定された T 細胞が、前記新生抗原のサブセット中の少なくとも 2 種類の新生抗原を認識する、請求項 1 ~ 37 のいずれか 1 項に記載の方法。

【請求項 39】

前記 1 つ以上の MHC アレルがクラス I MHC アレルである、請求項 1 ~ 38 のいずれか 1 項に記載の方法。

【発明の詳細な説明】

【背景技術】

【0001】

腫瘍特異的な新生抗原に基づいた治療用ワクチン及び T 細胞療法は、次世代の個別化がん免疫療法として極めて有望である¹⁻³。非小細胞肺癌 (NSCLC) 及びメラノーマなどの高い遺伝子変異量を有するがんは、新生抗原を生じる可能性が比較的高いことから、かかる治療法の特に有望な標的である^{4, 5}。初期の証拠により、新生抗原に基づいたワクチン接種が T 細胞応答を誘発し⁶、新生抗原を標的とした T 細胞療法が、選択された患者において腫瘍退縮を引き起こし⁷ことが示されている。MHC クラス I 及び MHC クラス II はいずれも T 細胞の応答に影響を及ぼす⁷⁰⁻⁷¹。

【0002】

しかしながら、新生抗原及び新生抗原認識 T 細胞の特定は、腫瘍応答を評価し^{77, 110}、腫瘍進化を調べ¹¹¹、次世代の個別化療法を設計する¹¹²うえで中心的な課題となっている。現在の新生抗原特定法は、時間と手間がかかるか^{84, 96}、または精度が充分とはいえない^{87, 91-93}。新生抗原認識 T 細胞は、TIL の主成分であり^{84, 96, 113, 114}、がん患者の末梢血中を循環している¹⁰⁷ことが近年、示されたが、新生抗原反応性 T 細胞を特定するための現在の方法は、(1) TIL^{97, 98}または白血球除去療法¹⁰⁷などの取得困難な臨床検体に頼ったものである、(2) 非実用的に大きなペプチドのライブラリーをスクリーニングする必要がある、または(3) 少数の MHC アレルについてのみ実用上、利用可能であり得る MHC アレルに頼ったものである、という制約条件のいくつかの組み合わせを有している。

【0003】

さらに、次世代のシーケンシング、RNA 遺伝子発現、及び新生抗原ペプチドの MHC 結合親和性の予測を用いた、変異に基づいた分析を取り入れた初期の方法が提案されている⁸。しかしながら、これらの提案されている方法では、遺伝子発現及び MHC 結合以外の多くの段階 (例えば、TAP 輸送、プロテアソーム切断、MHC 結合、ペプチド-MHC 複合体の細胞表面への輸送、及び/または MHC-I の TCR による認識; エンドサ

10

20

30

40

50

イトーシスまたはオートファジー、細胞外またはリソソームプロテアーゼ（例えばカテプシン）による切断、HLA-DMにより触媒されるHLA結合に対するCLIPペプチドとの競合、ペプチド-MHC複合体の細胞表面への輸送、及び/またはMHC-IIのTCRによる認識）を含む⁹ エピトープ生成プロセスの全体をモデル化することはできない。したがって、既存の方法は、陽性適中率（PPV）が低くなるという問題を有する傾向にある（図1A）。

【0004】

実際、複数の群によって実施された、腫瘍細胞により提示されるペプチドの分析は、遺伝子発現及びMHC結合親和性を用いて提示されることが予測されたペプチドの5%未満しか腫瘍表面のMHC上に見られないことを示している^{10, 11}（図1B）。結合予測とMHC提示との間のこのような低い相関は、変異の数単独に対してチェックポイント阻害剤反応について結合に制限された新生抗原の予測精度の向上が認められないことによってさらに指示されている¹²。

10

【0005】

提示を予測するための既存の方法のこのような低い陽性適中率（PPV）は、新生抗原に基づいたワクチンの設計における、また、新生抗原に基づいたT細胞療法における問題を提示する。PPVの低い予測を用いてワクチンが設計される場合、大部分の患者で治療用新生抗原が投与される可能性は低くなり、複数の新生抗原が投与される患者はさらに少なくなるものと考えられる（提示されるペプチドのすべてが免疫原性であると仮定したとしても）。同様に、治療用のT細胞がPPVの低い予測に基づいて設計される場合、大部分の患者で腫瘍新生抗原に対する反応性を有するT細胞が投与される可能性が低くなり、予測後に下流の検査方法を用いて予測新生抗原を特定する時間及び物理的リソースのコストが不要に高くなり得る。したがって、現行の方法による新生抗原ワクチン接種及びT細胞療法は、腫瘍を有する対象の相当数において奏功する可能性は低い（図1C）。

20

【0006】

さらに、これまでのアプローチは、シス作用性の変異のみを用いて候補新生抗原を生成するものであり、複数の腫瘍タイプで生じ、多くの遺伝子で異常スプライシングにつながるスプライシング因子の変異¹³、及びプロテアーゼ切断部位を生じるかまたは除去する変異を含む、新生ORFのさらなるソースをほとんどの場合で考慮していなかった。

【0007】

最後に、腫瘍ゲノム及びトランスクリプトーム解析に対する標準的アプローチは、ライブラリー構築、エクソーム及びトランスクリプトームの捕捉、シーケンシング、またはデータ分析における最適に満たない条件のために、候補新生抗原を生ずる体細胞突然変異を見逃す可能性がある。同様に、標準的な腫瘍分析のアプローチでは、配列アーチファクトまたは生殖系列多型を新生抗原として誤って助長してしまう場合があり、それぞれワクチン容量の非効率的な利用または自己免疫のリスクにつながりうる。

30

【発明の概要】

【0008】

本明細書では、個別化がんワクチン用の、T細胞療法用の、またはその両方のための新生抗原を特定及び選択するための最適化されたアプローチが開示される。第1に、次世代新生抗原（NGS）を用いた新生抗原候補を特定するための最適化された腫瘍エクソーム及びトランスクリプトーム解析アプローチに対する取り組みを行う。これらの方法は、最も感度及び特異性の高い新生抗原候補がすべてのクラスのゲノム変化にわたって発展されるように、NGSによる腫瘍解析の標準的アプローチに立脚したものである。第2に、特異性の問題を克服し、ワクチン添加用に、及び/またはT細胞療法の標的として開発される新生抗原が抗腫瘍免疫をより誘発しやすくするために高PPVの新生抗原選択に対する新規なアプローチが提供される。これらのアプローチには、実施形態に応じて、ペプチド-アレルマッピングを共にモデル化する訓練された統計学的回帰または非線形ディープラーニングモデル、ならびに異なる長さのペプチドにわたって統計学的効力を共有する、複数の長さのペプチドについてのアレルごとのモチーフが含まれる。特に非線形ディープラ

40

50

ーニングモデルは同じ細胞内の異なるMHCアレルを独立したものとして扱うように設計及び訓練することができるため、線形モデル同士が互いに干渉する線形モデルに伴う問題が解決される。最後に、新生抗原に基づいた個別化ワクチンの設計及び製造に関する、また、T細胞療法用の個別化新生抗原特異的なT細胞の製造におけるさらなる懸案事項が解決される。

【0009】

本明細書で開示されるモデルは、結合親和性で訓練された最新の予測ツール及びMSペプチドデータに基づいた初期の予測ツールの性能を最大で一桁上回る。ペプチドの提示をより高い信頼性で予測することにより、本モデルは、限られた量の患者の末梢血を使用し、患者当たりにつきスクリーニングするペプチドの数が少なく、MHCマルチマーに必ずしも頼らない、臨床的実用性のあるプロセスを用いた個別化療法のための新生抗原特異的または腫瘍抗原特異的なT細胞のより時間効率かつ費用効率の高い特定を可能とする。しかしながら、別の実施形態では、本明細書に開示されるモデルを用いることで、新生抗原または腫瘍抗原特異的なT細胞を特定するためにスクリーニングする必要がある、MHCマルチマーに結合したペプチドの数が少なくなることによって、MHCマルチマーを用いた腫瘍抗原特異的なT細胞のより時間効率及び費用効率の高い特定が可能となる。

10

【0010】

TIL新生エピートデータセット及び予想される新生抗原反応性T細胞の特定タスクにおける本明細書に開示されるモデルの予測性能は、HLAのプロセッシング及び提示をモデル化することによって、治療に有用な新生エピートの予測を得ることが今や可能であることを示す。要約すると、この作業は、抗原標的化免疫療法のための実用的なインシリコの抗原特定を可能とすることにより、患者の治療に向かう進行を加速するものである。

20

【図面の簡単な説明】

【0011】

本発明のこれらの特徴、態様、及び側面、ならびに他の特徴、態様、及び側面は、以下の説明文及び添付の図面に関してより深い理解が得られるであろう。

【0012】

【図1A】新生抗原の特定に対する現在の臨床的アプローチを示す。

【図1B】予測された結合ペプチドのうち、腫瘍細胞上に提示されるものは5%未満であることを示す。

30

【図1C】新生抗原予測の特異性の問題の影響を示す。

【図1D】結合予測が、新生抗原の特定に充分ではないことを示す。

【図1E】ペプチド長の関数としてのMHC-I提示の確率を示す。

【図1F】Promega社のダイナミックレンジ標準から生成された、例示的なペプチドスペクトルを示す。

【図1G】特性の追加が、いかにモデルの陽性適中率を増大させるかを示す。

【図2A】一実施形態による、患者におけるペプチド提示の尤度を特定するための環境の概略である。

【図2B】一実施形態による、提示情報を取得する方法を説明する。

【図2C】一実施形態による、提示情報を取得する方法を説明する。

40

【図3】一実施形態による、提示特定システムのコンピュータ論理構成要素を説明する、ハイレベルブロック図である。

【図4】一実施形態による、訓練データの例示的なセットを説明する。

【図5】MHCアレルに関連した例示的なネットワークモデルを説明する。

【図6A】一実施形態による、MHCアレルによって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。

【図6B】別の実施形態による、MHCアレルによって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。

【図7】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

50

【図 8】例示的なネットワークモデルを用いた、MHC アレルに関連したペプチドの提示尤度の生成を説明する。

【図 9】例示的なネットワークモデルを用いた、MHC アレルに関連したペプチドの提示尤度の生成を説明する。

【図 10】例示的なネットワークモデルを用いた、MHC アレルに関連したペプチドの提示尤度の生成を説明する。

【図 11】例示的なネットワークモデルを用いた、MHC アレルに関連したペプチドの提示尤度の生成を説明する。

【図 12】例示的なネットワークモデルを用いた、MHC アレルに関連したペプチドの提示尤度の生成を説明する。

【図 13 A】NSCLC 患者における変異負荷の標本度数分布を示す。

【図 13 B】一実施形態による、患者が最小変異負荷を満たすかどうかの選択基準に基づいて選択された患者に対してシミュレートしたワクチン中の提示新生抗原の数を示す。

【図 13 C】一実施形態による、提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、従来技術のモデルによって特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。

【図 13 D】HLA - A * 02 : 01 についての単一アレル毎提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、HLA - A * 02 : 01 及び HLA - B * 07 : 02 についてのアレル毎提示モデルの両方に基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。一実施形態によれば、ワクチン容量は、 $v = 20$ 個のエピトープに設定される。

【図 13 E】一実施形態による、変異負荷に基づいて選択された患者と、期待効用値スコアにより選択された患者との間でシミュレートしたワクチン中の提示新生抗原の数を比較したものである。

【図 14 A】各試験試料が 1 : 2500 の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む 5 種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全 MS モデル」、「ペプチド MS モデル」、及び $TPM > 0, 1$ 、及び 2 の 3 つの異なる遺伝子発現閾値を有する MHC Flurry 1.2.0 結合親和性モデルの再現率 40% における陽性適中率 (PPV) を比較したものである。

【図 14 B】各試験試料が 1 : 10, 000 の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む 15 種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全 MS モデル」、「ペプチド MS モデル」、及び $TPM > 0, 1$ 、及び 2 の 3 つの異なる遺伝子発現閾値を有する MHC Flurry 1.2.0 結合親和性モデルの再現率 40% における PPV を比較したものである。

【図 14 C】それぞれの試験試料が少なくとも 1 つの既存の T 細胞応答を示す患者から採取された 12 種類の異なる試験試料からなる試験セットについて、「完全 MS モデル」、「ペプチド MS モデル」、及び $TPM > 0, 1$ 、及び 2 の 3 つの異なる遺伝子発現閾値を有する MHC Flurry 1.2.0 結合親和性モデルによって特定された、上位 5、10、及び 20 位にランクされた体細胞変異について、T 細胞 (例えば、既存の T 細胞応答) によって認識された体細胞変異の割合を比較したものである。

【図 14 D】それぞれの試験試料が少なくとも 1 つの既存の T 細胞応答を示す患者から採取された 12 種類の異なる試験試料からなる試験セットについて、「完全 MS モデル」、「ペプチド MS モデル」、及び $TPM > 0, 1$ 、及び 2 の 3 つの異なる遺伝子発現閾値を有する MHC Flurry 1.2.0 結合親和性モデルによって特定された、上位 5、10、及び 20 位にランクされた最小新生エピトープについて、T 細胞 (例えば、既存の T 細胞応答) によって認識された最小新生エピトープの割合を比較したものである。

【図 15 A】9 人の患者における患者特異的的新生抗原ペプチドプールに対する T 細胞応答

10

20

30

40

50

の検出を示す。

【図15B】4人の患者における個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図15C】患者CU04についてELISpotウェルの例示的画像を示す。

【図16】各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」及び「アンカー残基のみのMSモデル」の再現率40%における陽性適中率(PPV)を比較したものである。

【図17A】各モデルを図14Aからの試験試料0で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

10

【図17B】各試験試料が1:5,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む15種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの再現率40%におけるPPVを比較したものである。

【図17C】各モデルを図14Aからの試験試料0で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

20

【図17D】各モデルを図14Aからの試験試料1で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17E】各モデルを図14Aからの試験試料2で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

30

【図17F】各モデルを図14Aからの試験試料3で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17G】各モデルを図14Aからの試験試料4で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17H】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*01:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

40

【図17I】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17J】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:03細胞株の試験データセットからの除外ペプチドで各モデルを試

50

験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17K】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:07細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17L】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*03:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17M】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*24:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17N】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*29:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17O】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*31:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17P】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*68:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17Q】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*35:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17R】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*44:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17S】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*44:03細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17T】1:10, 000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*51:01細胞株の試験データセットからの除外ペプチドで各モデルを試

10

20

30

40

50

験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHC Flurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17U】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*54:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHC Flurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17V】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*57:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHC Flurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図18】各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる図14Aの試験セットで各モデルを試験する場合のヒト腫瘍におけるHLA提示ペプチドをモデル化するためのMSモデルの異なるバージョン及び初期のアプローチ²⁹の再現率40%における陽性適中率(PPV)を比較したものである。

【図19A-1】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図19A-2】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図19B-1】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図19B-2】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図20】図15Aに示される各ドナー及び各インビトロ増殖についてPHA陽性対照に対するT細胞応答の検出を示す。

【図21A】患者CU04におけるプール#2中のそれぞれの個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図21B】患者CU04の3回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図21C】患者CU04の2回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的新生抗原ペプチドに対する、及び患者特異的新生抗原ペプチドプールに対するT細胞応答の検出を示す。

【図22】図15Aの患者について2つの患者特異的新生抗原ペプチドプール及びDMSO陰性対照に対するT細胞応答の検出を示す。

【図23】「MSモデル」、HLA-DRB1*15:01及びHLA-DRB5*01:01にわたって最も低いNetMHCIIpanパーセントイルランクを採用する、「NetMHCIIpan rank」: NetMHCIIpan 3.1⁷⁷、及びHLA-DRB1*15:01及びHLA-DRB5*01:01にわたってnM単位の最も強い親和性を採用する、「NetMHCIIpan nM」: NetMHCIIpan 3.1の、HLA-DRB1*15:01 / HLA-DRB5*01:01試験データセット内でのペプチドのランク付けにおける予測性能を比較したものである。

【図24-1】NSCLC患者の末梢血由来の新生抗原特異的メモリーT細胞のTCRをシーケンシングするための方法を示す。

【図24-2】NSCLC患者の末梢血由来の新生抗原特異的メモリーT細胞のTCRをシーケンシングするための方法を示す。

10

20

30

40

50

【図25】TCRをレシピエント細胞に導入するためのTCRコンストラクトの例示的な実施形態を示す。

【図26】治療法の開発のためにTCRを発現系にクローニングするための例示的なP526コンストラクト骨格のヌクレオチド配列を示す。

【図27】治療法の開発のために患者の新生抗原特異的TCRのクロータイプ1を発現系にクローニングするための例示的なコンストラクト配列を示す。

【図28】治療法の開発のために患者の新生抗原特異的TCRのクロータイプ3を発現系にクローニングするための例示的なコンストラクト配列を示す。

【図29】一実施形態により、患者に個別化された新生抗原特異的治療を行うための方法のフローチャートである。

【図30】図1及び3に示した実体を実施するための例示的なコンピュータを説明する。

【発明を実施するための形態】

【0013】

I. 定義

一般に、特許請求の範囲及び明細書において使用される用語は、当業者により理解される通常の意味を有するものとして解釈されるものとする。特定の用語を、さらなる明確性を与えるために下記に定義する。通常の意味と与えられる定義との間に矛盾が存在する場合、与えられる定義が用いられるものとする。

【0014】

本明細書で使用するところの「抗原」という用語は、免疫反応を誘導する物質のことである。

【0015】

本明細書で使用するところの「新生抗原」という用語は、例えば、腫瘍細胞の変異、または腫瘍細胞に特異的な翻訳後修飾によって、抗原に対応する野生型の親抗原とは異なるものとする少なくとも1つの変化を有する抗原のことである。新生抗原は、ポリペプチド配列またはヌクレオチド配列を含んでよい。変異は、フレームシフトもしくは非フレームシフト挿入欠失(indel)、ミスセンスもしくはナンセンス置換、スプライス部位変化、ゲノム再編成もしくは遺伝子融合、または、新生ORFを生じる任意のゲノム変化もしくは発現変化を含むことができる。変異はまた、スプライス変異体も含むことができる。腫瘍細胞に特異的な翻訳後修飾は、異常リン酸化を含むことができる。腫瘍細胞に特異的な翻訳後修飾はまた、プロテアソームによって生成されるスプライス抗原も含むことができる。Liepe et al., A large fraction of HLA class I ligands are proteasome-generated spliced peptides; Science. 2016 Oct 21; 354(6310): 354-358を参照されたい。

【0016】

本明細書で使用するところの「腫瘍新生抗原」という用語は、対象の腫瘍細胞または組織中に存在するが、対象に対応する正常細胞または組織中には存在しない新生抗原のことである。

【0017】

本明細書において使用される場合、「新生抗原ベースのワクチン」という用語は、1つ以上の新生抗原、例えば複数の新生抗原に基づいたワクチンコンストラクトのことである。

【0018】

本明細書において使用される場合、「候補新生抗原」という用語は、新生抗原を表しうる新たな配列を生じる変異、または他の異常のことである。

【0019】

本明細書において使用される場合、「コード領域」という用語は、タンパク質をコード化する遺伝子の部分のことである。

【0020】

10

20

30

40

50

本明細書において使用される場合、「コード変異」という用語は、コード領域で生じる変異のことである。

【0021】

本明細書において使用される場合、「ORF」という用語は、オープンリーディングフレームを意味する。

【0022】

本明細書において使用される場合、「新生ORF」という用語は、変異またはスプライシングなどの他の異常により生じる腫瘍特異的なORFのことである。

【0023】

本明細書において使用される場合、「ミスセンス変異」という用語は、1つのアミノ酸から別のアミノ酸への置換を引き起こす変異である。

【0024】

本明細書において使用される場合、「ナンセンス変異」という用語は、アミノ酸から終止コドンへの置換を引き起こす変異である。

【0025】

本明細書において使用される場合、「フレームシフト変異」という用語は、タンパク質のフレームに変更を引き起こす変異である。

【0026】

本明細書において使用される場合、「挿入欠失」という用語は、1つ以上の核酸の挿入または欠失である。

【0027】

本明細書において使用される場合、2つ以上の核酸またはポリペプチドの配列との関連での「同一率」(%)という用語は、下記の配列比較アルゴリズム(例えば、BLASTP及びBLASTN、または当業者が利用可能な他のアルゴリズム)のうちの1つを用いて、または目視検査により測定される、最大の一致について比較し、整列させた場合に、ヌクレオチドまたはアミノ酸残基の特定の比率(%)が同じである2つ以上の配列または部分配列のことを指す。用途に応じて、「同一率」(%)は、比較される配列の領域にわたって、例えば、機能ドメインにわたって存在するか、あるいは、比較される2つの配列の完全長にわたって存在することができる。

【0028】

配列比較では、一般的に、1つの配列が、試験配列が比較される参照配列として機能する。配列比較アルゴリズムを用いる場合、試験配列及び参照配列をコンピュータに入力し、必要な場合には部分配列座標を指定し、配列アルゴリズムプログラムのパラメータを指定する。次いで、配列比較アルゴリズムが、指定されたプログラムパラメータに基づいて、参照配列に対する試験配列の配列同一率(%)を算出する。あるいは、配列の類似性または相違性は、選択された配列位置(例えば、配列モチーフ)における特定のヌクレオチドの、または翻訳後の配列ではアミノ酸の有無の組み合わせによって確立することもできる。

【0029】

比較を行うための配列の最適なアラインメントは、例えば、Smith & Waterman, Adv. Appl. Math. 2:482 (1981)の局所相同性アルゴリズムによって、Needleman & Wunsch, J. Mol. Biol. 48:443 (1970)の相同性アラインメントアルゴリズムによって、Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988)の類似性の探索法によって、これらのアルゴリズムのコンピュータ処理による実行(Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.におけるGAP、BESTFIT、FASTA、及びTFASTA)によって、または目視検査によって実施することができる(一般的には、下記のAusubel et al.を参照)。

10

20

30

40

50

【0030】

配列同一率(%)及び配列類似率(%)を決定するのに適したアルゴリズムの1つの例として、Altschul et al., J. Mol. Biol. 215:403-410 (1990)に記載されるBLASTアルゴリズムがある。BLAST解析を行うためのソフトウェアは、National Center for Biotechnology Informationを通して公に入手可能である。

【0031】

本明細書において使用される場合、「ノンストップまたはリードスルー」という用語は、天然の終止コドンの除去を引き起こす変異のことである。

【0032】

本明細書において使用される場合、「エピトープ」という用語は、抗体またはT細胞受容体が一般的に結合する、抗原の特異的な部分のことである。

【0033】

本明細書において使用される場合、「免疫原性」という用語は、例えば、T細胞、B細胞、またはその両方を介して免疫応答を誘発する能力のことである。

【0034】

本明細書において使用される場合、「HLA結合親和性」、「MHC結合親和性」という用語は、特異的な抗原と特異的なMHCアレルとの結合の親和性を意味する。

【0035】

本明細書において使用される場合、「ベイト」という用語は、DNAまたはRNAの特異的な配列を試料から濃縮するために使用される核酸プローブのことである。

【0036】

本明細書において使用される場合、「変異」という用語は、対象の核酸と、対照として使用される参照ヒトゲノムとの差である。

【0037】

本明細書において使用される場合、「変異コール」という用語は、典型的にはシーケンシングからの、変異の存在のアルゴリズム的決定である。

【0038】

本明細書において使用される場合、「多型」という用語は、生殖細胞系列変異、すなわち、個体のすべてのDNA保有細胞において見出される変異である。

【0039】

本明細書において使用される場合、「体細胞変異」という用語は、個体の非生殖系列細胞において生じる変異である。

【0040】

本明細書において使用される場合、「アレル」という用語は、遺伝子の1つのバージョンまたは遺伝子配列の1つのバージョンまたはタンパク質の1つのバージョンのことである。

【0041】

本明細書において使用される場合、「HLA型」という用語は、HLA遺伝子アレルの相補体のことである。

【0042】

本明細書において使用される場合、「ナンセンス変異依存分解機構」または「NMD」という用語は、未成熟な終止コドンに起因する細胞によるmRNAの分解のことである。

【0043】

本明細書において使用される場合、「トランカル変異(truncal mutation)」という用語は、腫瘍の発生の初期に生じ、腫瘍の細胞の大部分に存在する変異である。

【0044】

本明細書において使用される場合、「サブクローナル変異」という用語は、腫瘍の発生において後期に生じ、腫瘍の細胞の一部のみに存在する変異である。

10

20

30

40

50

【 0 0 4 5 】

本明細書において使用される場合、「エクソーム」という用語は、タンパク質をコードするゲノムのサブセットである。エクソームは、ゲノムの集合的なエクソンでありうる。

【 0 0 4 6 】

本明細書において使用される場合、「ロジスティック回帰」という用語は、従属変数が 1 に等しい確率のロジットが従属変数の線形関数としてモデル化される、統計からのバイナリデータ用の回帰モデルである。

【 0 0 4 7 】

本明細書において使用される場合、「ニューラルネットワーク」という用語は、多層の線形変換に続いて一般的に確率的勾配降下法及び逆伝搬により訓練された要素ごとの非線形変換を行うことからなる分類または回帰のための機械学習モデルである。

10

【 0 0 4 8 】

本明細書において使用される場合、「プロテオーム」という用語は、細胞、細胞の群、または個体によって発現される、及び/または翻訳されるすべてのタンパク質のセットのことである。

【 0 0 4 9 】

本明細書において使用される場合、「ペプチドーム」という用語は、細胞表面上の MHC - I または MHC - II によって提示されるすべてのペプチドのセットのことである。ペプチドームは、細胞または細胞の集合の性質を指す場合もある（例えば、腫瘍ペプチドームは、腫瘍を含むすべての細胞のペプチドームの和集合を意味する）。

20

【 0 0 5 0 】

本明細書において使用される場合、「ELISPOT」という用語は、ヒト及び動物において免疫応答を観察するための一般的な方法である、酵素結合免疫吸着スポットアッセイを意味する。

【 0 0 5 1 】

本明細書において使用される場合、「デキサトラマー」という用語は、フローサイトメトリーにおいて抗原特異的 T 細胞染色に使用される、デキストランベースのペプチド - MHC マルチマーである。

【 0 0 5 2 】

本明細書において使用される場合、「MHC マルチマー」という用語は、複数のペプチド - MHC モノマー単位からなるペプチド - MHC 複合体である。

30

【 0 0 5 3 】

本明細書において使用される場合、「MHC テトラマー」という用語は、4 つのペプチド - MHC モノマー単位からなるペプチド - MHC 複合体である。

【 0 0 5 4 】

本明細書において使用される場合、「寛容または免疫寛容」という用語は、1 つ以上の抗原、例えば、自己抗原に対する免疫不応答の状態のことである。

【 0 0 5 5 】

本明細書において使用される場合、「中枢性寛容」という用語は、自己反応性 T 細胞クローンを欠失させること、または自己反応性 T 細胞クローンの免疫抑制性制御性 T 細胞 (Treg) への分化を促進することのいずれかにより、胸腺において与えられる寛容である。

40

【 0 0 5 6 】

本明細書において使用される場合、「末梢性寛容」という用語は、中枢性寛容を生き延びた自己反応性 T 細胞を下方制御もしくはアネルギー化すること、またはこれらの T 細胞の Treg への分化を促進することにより、末梢系において与えられる寛容である。

【 0 0 5 7 】

「試料」という用語は、静脈穿刺、排泄、射精、マッサージ、生検、針吸引、洗浄試料、擦過、外科的切開、もしくは介入、または当技術分野において公知の他の手段を含む手段によって対象から採取された、単一細胞、または複数の細胞、または細胞の断片、また

50

は体液のアリコートを含むことができる。

【0058】

「対象」という用語は、インビボ、エクスピボ、またはインビトロ、雄または雌のいずれかの、細胞、組織、または生物体、ヒトまたは非ヒトを包含する。対象という用語は、ヒトを含む哺乳動物を含める。

【0059】

「哺乳動物」という用語は、ヒト及び非ヒトの両方を包含し、ヒト、非ヒト霊長類、イヌ、ネコ、マウス、ウシ、ウマ、及びブタを含むが、それらに限定されない。

【0060】

「臨床的因子」という用語は、対象の状態、例えば、疾患の活性または重症度の測定を指す。「臨床的因子」は、非試料マーカーを含む、対象の健康状態のすべてのマーカー、ならびに/または、非限定的に年齢及び性別などの、対象の他の特徴を包含する。臨床的因子は、対象または所定の条件下の対象由来の試料（または試料の集団）の評定から取得され得るスコア、値、または値のセットであることができる。臨床的因子はまた、マーカー、及び/または遺伝子発現代替物などの他のパラメータによっても予測することができる。臨床的因子は、腫瘍タイプ、腫瘍サブタイプ、及び喫煙歴を含むことができる。

【0061】

略語：MHC：主要組織適合性複合体；HLA：ヒト白血球抗原、またはヒトMHC遺伝子座；NGS：次世代シーケンシング；PPV：陽性適中率；TSNA：腫瘍特異的新生抗原；FFPE：ホルマリン固定パラフィン包埋；NMD：ナンセンス変異依存分解機構；NSCLC：非小細胞肺癌；DC：樹状細胞。

【0062】

本明細書及び添付の特許請求の範囲において使用される場合、単数形「a」、「an」、及び「the」は、文脈によってそうでない旨が明示されない限り、複数の指示物を含む点に留意されたい。

【0063】

本明細書において直接定義されていない用語は、本発明の技術分野の範囲内で理解されるような、一般的にそれらに付随する意味を有するものとして理解されるべきである。本発明の態様の組成物、装置、方法など、ならびにそれらの製造または使用法を説明するうえで実施者にさらなる手引きを与える目的で特定の用語が本明細書で検討される。同じものについて複数の言い方がなされる点は認識されるであろう。したがって、代替的な語及び同義語が、本明細書で検討される用語の任意の1つ以上について用いられる場合がある。本明細書においてある用語が詳述または検討されているか否かに重きが置かれるべきではない。いくつかの同義語または代用可能な方法、材料などが提供される。1つまたは数個の同義語または均等物の記載は、明確に述べられない限り、他の同義語または均等物の使用を除外しない。用語の例を含む例の使用は、あくまで説明を目的としたものにすぎず、本明細書における発明の態様の範囲及び意味を限定しない。

【0064】

本明細書の本文において引用されるすべての参考文献、発行特許、及び特許出願は、あらゆる目的でそれらの全容を参照により本明細書に援用するものである。

【0065】

II. 新生抗原を特定する方法

本明細書では、腫瘍細胞の表面上に提示される確率の高い、対象の腫瘍細胞由来の新生抗原に対して抗原特異的であるT細胞を特定するための方法を開示する。本方法は、対象の腫瘍細胞及び正常細胞からエクソーム、トランスクリプトーム、及び/または全ゲノムのヌクレオチドシーケンシングデータを取得することを含む。このヌクレオチドシーケンシングデータを用いて新生抗原のセットの中の各新生抗原のペプチド配列が取得される。新生抗原のセットは、腫瘍細胞からのヌクレオチドシーケンシングデータと、正常細胞からのヌクレオチドシーケンシングデータと比較することによって特定される。具体的には、新生抗原のセットの中の各新生抗原のペプチド配列は、ペプチド配列を対象の

10

20

30

40

50

正常細胞から特定された対応する野生型の親ペプチド配列とは異なるものとする少なくとも1つの変化を含む。本方法は、新生抗原のセットの中の各新生抗原のペプチド配列を対応する数値ベクトルにコード化することをさらに含む。各数値ベクトルは、ペプチド配列を構成するアミノ酸及びペプチド配列内のアミノ酸の位置を記述する情報を含む。本方法は、機械学習させた提示モデルに数値ベクトルを入力することにより新生抗原のセットの中の各新生抗原について提示尤度を生成することをさらに含む。各提示尤度は、対応する新生抗原が対象の腫瘍細胞の表面上のMHCアレルによって提示される尤度を表す。機械学習させた提示モデルは、複数のパラメータ及び関数を含む。複数のパラメータは、訓練データセットに基づいて特定される。訓練データセットは、複数の試料の中のそれぞれの試料について、その試料中に存在するものとして特定されたMHCアレルのセットの中の少なくとも1つのMHCアレルに結合したペプチドの存在を測定する質量分析によって得られた標識と、ペプチドを構成するアミノ酸及び/またはペプチド内のアミノ酸の位置を記述する情報を含む数値ベクトルとしてコード化された訓練ペプチド配列とを含む。関数は、機械学習させた提示モデルによって入力として受け取られた数値ベクトルと、数値ベクトル及び複数のパラメータに基づいて機械学習させた提示モデルによって出力として生成された提示尤度との間の関係を表す。本方法は、選択された新生抗原のセットを生成するために、新生抗原のセットのサブセットを、提示尤度に基づいて選択することをさらに含む。本方法は、サブセット中の新生抗原のうちの少なくとも1つに対して抗原特異的なT細胞を特定することと、これらの特定されたT細胞を返すこととをさらに含む。

10

20

30

40

50

【0066】

いくつかの実施形態では、機械学習させた提示モデルに数値ベクトルを入力することは、機械学習させた提示モデルを新生抗原のペプチド配列に適用してMHCアレルのそれぞれについて依存性スコアを生成することを含む。あるMHCアレルについての依存性スコアは、ペプチド配列の特定の位置の特定のアミノ酸に基づいてそのMHCアレルが新生抗原を提示するかどうかを示す。さらなる実施形態では、機械学習させた提示モデルに数値ベクトルを入力することは、各MHCアレルについて、対応するMHCアレルが対応する新生抗原を提示する尤度を示す、対応するアレル毎尤度を生成するために、依存性スコアを変換すること、及び新生抗原の提示尤度を生成するために、アレル毎尤度を組み合わせること、をさらに含む。いくつかの実施形態では、依存性スコアを変換することは、新生抗原の提示をMHCアレルにわたって相互排他的なものとしてモデル化する。代替的な実施形態では、機械学習させた提示モデルに数値ベクトルを入力することは、依存性スコアの組み合わせを変換して提示尤度を生成することをさらに含む。かかる実施形態では、依存性スコアの組み合わせを変換することは、新生抗原の提示をMHCアレル間の干渉としてモデル化する。

【0067】

いくつかの実施形態では、提示尤度のセットは、1つ以上のアレル非相互作用特性によってさらに特定される。かかる実施形態では、本方法は、機械学習させた提示モデルをアレル非相互作用特性に適用することによってアレル非相互作用特性に対する依存性スコアを生成することをさらに含む。依存性スコアは、対応する新生抗原のペプチド配列が、アレル非相互作用特性に基づいて提示されるかどうかを示す。いくつかの実施形態では、本方法は、各MHCアレルについての依存性スコアをアレル非相互作用特性についての依存性スコアと組み合わせること、各MHCアレルについてアレル毎尤度を生成するために、各MHCアレルについての組み合わせられた依存性スコアを変換すること、及び提示尤度を生成するためにアレル毎尤度を組み合わせること、をさらに含む。あるMHCアレルについてのアレル毎尤度は、そのMHCアレルが対応する新生抗原を提示する湯戸を示す。代替的な実施形態では、本方法は、MHCアレルについての依存性スコアとアレル非相互作用特性についての依存性スコアとを組み合わせることと、組み合わせられた依存性スコアを変換して提示尤度を生成することと、をさらに含む。

【0068】

いくつかの実施形態では、MHCアレルは2つ以上の異なるMHCアレルを含む。

【0069】

いくつかの実施形態では、ペプチド配列は、アミノ酸9個以外の長さを有するペプチド配列を含む。

【0070】

いくつかの実施形態では、ペプチド配列をコード化することは、ワン・ホットエンコーディングスキームを用いてペプチド配列をコード化することを含む。

【0071】

特定の実施形態では、複数の試料は、単一のMHCアレルを発現するように操作された細胞株、複数のMHCアレルを発現するように操作された細胞株、複数の患者から得られたまたは由来するヒト細胞株、複数の患者から得られた新鮮なまたは凍結された組織試料のうち少なくとも1つを含む。

10

【0072】

いくつかの実施形態では、訓練データセットは、ペプチドのうち少なくとも1つについてのペプチド-MHC結合親和性の測定値に関連したデータ、及び、ペプチドのうち少なくとも1つについてのペプチド-MHC結合安定性の測定値に関連したデータのうちの少なくとも1つをさらに含む。

【0073】

いくつかの実施形態では、提示尤度のセットは、RNA-seqまたは質量分析により測定される、対象におけるMHCアレルの発現レベルによってさらに特定される。

【0074】

いくつかの実施形態では、提示尤度のセットは、新生抗原のセットの中の新生抗原とMHCアレルとの間の予測される親和性、及び、新生抗原コード化ペプチド-MHC複合体の予測される安定性少なくとも一方を含む特性によってさらに特定される。

20

【0075】

いくつかの実施形態では、数値的尤度のセットは、そのソースタンパク質配列内の新生抗原コード化ペプチド配列に隣接するC末端配列、及びそのソースタンパク質配列内の新生抗原コード化ペプチド配列に隣接するN末端配列のうち少なくとも1つを含む特性によってさらに特定される。

【0076】

いくつかの実施形態では、選択された新生抗原のセットを選択することは、機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して腫瘍細胞表面上に提示される尤度が増大している新生抗原を選択することを含む。

30

【0077】

いくつかの実施形態では、選択された新生抗原のセットを選択することは、機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して対象において腫瘍特異的な免疫応答を誘導することができる尤度が増大している新生抗原を選択することを含む。

【0078】

いくつかの実施形態では、選択された新生抗原のセットを選択することは、提示モデルに基づいて、選択されない新生抗原と比較してプロフェッショナル抗原提示細胞(APC)によってナイーブT細胞に提示されることができる尤度が増大している新生抗原を選択することを含む。かかる実施形態では、APCは任意で樹状細胞(DC)である。

40

【0079】

いくつかの実施形態では、選択された新生抗原のセットを選択することは、機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して中枢性寛容または末梢性寛容によって阻害される尤度が減少している新生抗原を選択することを含む。

【0080】

いくつかの実施形態では、選択された新生抗原のセットを選択することは、機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して対象において正常組織に対する自己免疫応答を誘導することができる尤度が減少している新生抗原を選択することを含む。

50

【0081】

いくつかの実施形態では、1つ以上の腫瘍細胞は、肺癌、メラノーマ、乳癌、卵巣癌、前立腺癌、腎臓癌、胃癌、結腸癌、精巣癌、頭頸部癌、膵臓癌、脳癌、B細胞リンパ腫、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、及びT細胞リンパ球性白血病、非小細胞肺癌、及び小細胞肺癌からなる群から選択される。

【0082】

いくつかの実施形態では、本方法は、選択された新生抗原のセットから個別化がんワクチンを構築するための出力を生成することをさらに含む。かかる実施形態では、個別化がんワクチン用の出力は、選択された新生抗原のセットをコードした少なくとも1つのペプチド配列または少なくとも1つのヌクレオチド配列を含むことができる。

10

【0083】

いくつかの実施形態では、機械学習させた提示モデルは、ニューラルネットワークモデルである。かかる実施形態では、ニューラルネットワークモデルは、それぞれのネットワークモデルが複数のMHCアレルのうちに対応するMHCアレルに割り当てられ、かつ1つ以上の層に配置された一連のノードを含む、MHCアレルについての複数のネットワークモデルを含み得る。かかる実施形態では、ニューラルネットワークモデルは、ニューラルネットワークモデルのパラメータを更新することによって訓練することができ、少なくとも2つのネットワークモデルのパラメータが少なくとも1回の訓練イテレーションについて一緒に更新される。いくつかの実施形態では、機械学習させた提示モデルは、ノードの1つ以上の層を含むディープラーニングモデルとすることができる。

20

【0084】

いくつかの実施形態では、T細胞を特定することは、T細胞を増殖させる条件下でT細胞をサブセットの中の新生抗原のうち1つ以上と共培養することを含む。

【0085】

いくつかの実施形態では、T細胞を特定することは、T細胞を、サブセットの中の新生抗原のうち1つ以上を含むMHCマルチマーと、T細胞とMHCマルチマーとの結合が可能な条件下で接触させることを含む。

【0086】

いくつかの実施形態では、本方法は、特定されたT細胞のT細胞受容体(TCR)を特定することをさらに含む。かかる実施形態では、T細胞受容体を特定することは、特定されたT細胞のT細胞受容体配列をシーケンシングすることを含む。かかる実施形態では、本方法は、1つ以上の特定されたT細胞受容体のうちの少なくとも1つを発現するようにT細胞を遺伝子操作することと、T細胞が増殖する条件下でT細胞を培養することと、増殖させたT細胞を対象に注入することと、をさらに含むことができる。かかる実施形態では、特定されたT細胞受容体のうちの少なくとも1つを発現するようにT細胞を遺伝子操作することは、特定されたT細胞のT細胞受容体配列を発現ベクターにクローニングすることと、T細胞のそれぞれに発現ベクターをトランスフェクトすることと、を含むことができる。

30

【0087】

いくつかの実施形態では、本方法は、特定されたT細胞を、特定されたT細胞が増殖する条件下で培養することと、増殖させたT細胞を対象に注入することと、をさらに含むことができる。

40

【0088】

いくつかの実施形態では、サブセット中の新生抗原のうち少なくとも1つに対して抗原特異的であるT細胞が、対象からの5~30mLの全血を使用して特定される。

【0089】

いくつかの実施形態では、新生抗原のサブセットは、最大で20種の新生抗原を含み、特定されたT細胞は新生抗原のサブセット中の少なくとも2種類の新生抗原を認識する。

【0090】

いくつかの実施形態では、MHCアレルは、クラスI MHCアレルである。

50

【0091】

本明細書では、上記に述べた新生抗原のサブセット中の少なくとも1つの選択された新生抗原に対して抗原特異的である単離されたT細胞も開示される。

【0092】

III. 新生抗原における腫瘍特異的変異の特定

また、ある特定の変異（例えば、がん細胞中に存在する変異またはアレル）の特定のための方法も、本明細書に開示する。特に、これらの変異は、がんを有する対象のがん細胞のゲノム、トランスクリプトーム、プロテオーム、またはエクソーム中に存在し得るが、対象由来の正常組織には存在し得ない。

【0093】

腫瘍における遺伝子変異は、それらが腫瘍において排他的にタンパク質のアミノ酸配列における変更をもたらす場合、腫瘍の免疫学的ターゲティングに有用と考えることができる。有用な変異は、以下を含む：（1）タンパク質において異なるアミノ酸をもたらす非同義変異；（2）C末端に新規の腫瘍特異的配列を有する、より長いタンパク質の翻訳をもたらす、終止コドンが修飾されているかまたは欠失しているリードスルー変異；（3）成熟mRNAにおけるイントロンの包含、したがって固有の腫瘍特異的タンパク質配列をもたらす、スプライス部位変異；（4）2種類のタンパク質の接合部に腫瘍特異的配列を有するキメラタンパク質を生じる、染色体再編成（すなわち、遺伝子融合）；（5）新規の腫瘍特異的タンパク質配列を有する新たなオープンリーディングフレームをもたらす、フレームシフト変異または欠失。変異はまた、非フレームシフト挿入欠失、ミスセンスもしくはナンセンス置換、スプライス部位変化、ゲノム再編成もしくは遺伝子融合、または、新生ORFを生じる任意のゲノム変化もしくは発現変化のうちの1つ以上も含むことができる。

【0094】

例えば、腫瘍細胞におけるスプライス部位、フレームシフト、リードスルー、または遺伝子融合の変異から生じた、変異を有するペプチドまたは変異したポリペプチドは、腫瘍対正常細胞において、DNA、RNA、またはタンパク質をシーケンシングすることによって特定することができる。

【0095】

また、変異は、以前に特定された腫瘍特異的変異を含むことができる。公知の腫瘍変異は、Catalogue of Somatic Mutations in Cancer (COSMIC) データベースで見出すことができる。

【0096】

様々な方法を、個体のDNAまたはRNAにおいて特定の変異またはアレルの存在を検出するために利用可能である。この分野における進歩は、正確で、容易な、かつ安価な大規模SNP遺伝子型判定を提供している。例えば、動的アレル特異的ハイブリダイゼーション(DASH)、マイクロプレートアレイ対角線ゲル電気泳動(MADGE)、パイロシーケンシング、オリゴヌクレオチド特異的ライゲーション、TaqManシステム、及びAffymetrix SNPチップなどの種々のDNA「チップ」技術を含むいくつかの技法が、記載されている。これらの方法は、典型的にはPCRによる、標的遺伝子領域の増幅を利用する。さらに他の方法は、侵襲性切断による小さなシグナル分子の生成及びその後の質量分析、または、固定化されたパッドロックプローブ及びローリングサークル増幅に基づく。特異的な変異を検出するための、当技術分野において公知の方法のいくつかを、下記に要約する。

【0097】

PCRベースの検出手段は、多数のマーカーの多重増幅を同時に含むことができる。例えば、サイズがオーバーラップせず、同時に解析することができるPCR産物を生成するようにPCRプライマーを選択することが、当技術分野において周知である。あるいは、差次的にラベル化され、したがって、各々を差次的に検出することができるプライマーで異なるマーカーを増幅することが可能である。当然、ハイブリダイゼーションベースの検

10

20

30

40

50

出手段により、試料における複数のPCR産物の差次的な検出が可能になる。複数のマーカーの多重解析を可能にする他の技法が、当技術分野において公知である。

【0098】

いくつかの方法が、ゲノムDNAまたは細胞RNAにおける単一ヌクレオチド多型の解析を容易にするために開発されている。例えば、一塩基多型は、例えば、Mundy, C. R. (米国特許第4,656,127号)において開示されているような、特化されたエキソヌクレーゼ抵抗性ヌクレオチドを用いることによって検出することができる。この方法にしたがって、多型部位のすぐ3'のアレル配列に対して相補的なプライマーを、特定の動物またはヒトから取得された標的分子に対してハイブリダイズさせる。標的分子上の多型部位が、存在する特定のエキソヌクレーゼ抵抗性ヌクレオチド誘導体に対して相補的であるヌクレオチドを含有する場合、その誘導体は、ハイブリダイズされたプライマーの末端上に組み込まれる。そのような組み込みのために、プライマーはエキソヌクレーゼに対して抵抗性になり、それによりその検出が可能になる。試料のエキソヌクレーゼ抵抗性誘導体の同一性は既知であるため、プライマーがエキソヌクレーゼに対して抵抗性になったという知見により、標的分子の多型部位に存在するヌクレオチドが、反応において使用されたヌクレオチド誘導体のものに対して相補的であることが明らかになる。この方法は、多量の外来性配列データの決定を必要としないという利点を有する。

10

【0099】

多型部位のヌクレオチドの同一性を決定するために、溶液ベースの方法を使用することができる(Cohen, D. et al. (フランス国特許第2,650,840号; PCT出願第WO91/02087号)。米国特許第4,656,127号のMundyの方法におけるように、多型部位のすぐ3'のアレル配列に対して相補的であるプライマーを使用する。この方法は、多型部位のヌクレオチドに対して相補的である場合は、プライマーの末端上に組み込まれるようになる、ラベル化ジデオキシヌクレオチド誘導体を用いて、その部位のヌクレオチドの同一性を決定する。Genetic Bit AnalysisまたはGBAとして公知である代替的な方法が、Goellet, P. et al. (PCT出願第92/15712号)により記載されている。Goellet, P. et al.の方法は、ラベル化ターミネーターと、多型部位の3'の配列に対して相補的であるプライマーとの混合物を使用する。Goellet, P. et al.の方法は、ラベル化ターミネーターと、多型部位の3'の配列に対して相補的であるプライマーとの混合物を使用する。Cohen et al. (フランス国特許第2,650,840号; PCT出願第WO91/02087号)の方法とは対照的に、Goellet, P. et al.の方法は、プライマーまたは標的分子が固相に固定化される、不均一相アッセイであることができる。

20

30

【0100】

DNAにおいて多型部位をアッセイするための、いくつかのプライマーガイドヌクレオチド組み込み手順が、記載されている(Komher, J. S. et al., Nucl. Acids. Res. 17:7779-7784(1989); Sokolov, B. P., Nucl. Acids Res. 18:3671(1990); Syvanen, A.-C., et al., Genomics 8:684-692(1990); Kuppuswamy, M. N. et al., Proc. Natl. Acad. Sci. (U.S.A.) 88:1143-1147(1991); Prezant, T. R. et al., Hum. Mutat. 1:159-164(1992); Ugozzoli, L. et al., GATA 9:107-112(1992); Nyren, P. et al., Anal. Biochem. 208:171-175(1993))。これらの方法は、それらが、多型部位で塩基間を識別するためにラベル化デオキシヌクレオチドの組み込みを利用する点で、GBAとは異なる。そのような形式において、シグナルは、組み込まれたデオキシヌクレオチドの数に比例するため、同じヌクレオチドのランにおいて起こる多型は、ランの長さ按比例するシグナルを結果としてもたらしすることができる(Syvanen, A.-C., et al., Amer. J. Hum. Genet. 52:

40

50

46 - 59 (1993))。

【0101】

数多くのイニシアティブは、DNAまたはRNAの何百万もの個々の分子から並行して直接、配列情報を取得する。リアルタイムの単一分子の合成によるシーケンシング技術は、シーケンシングされる鋳型に対して相補的であるDNAの新生鎖の中に組み込まれる際の、蛍光ヌクレオチドの検出に依拠する。1つの方法において、長さが30~50塩基のオリゴヌクレオチドを、ガラスのカバーガラスに、5'端で共有結合性に固着させる。これらの固着した鎖は、2つの機能を果たす。第1に、それらは、鋳型が、表面結合オリゴヌクレオチドに対して相補的な捕捉尾部を有して構成されている場合に、標的鋳型鎖の捕捉部位として作用する。それらはまた、配列読み取りの基礎を形成する、鋳型指向性プライマー伸長のためのプライマーとしても作用する。捕捉プライマーは、複数サイクルの合成、検出、及び、色素を除去するための色素-リンカーの化学的切断を用いた、配列決定のための、固定された位置部位として機能する。各サイクルは、ポリメラーゼ/ラベル化ヌクレオチド混合物の添加、リンス、画像化、及び色素の切断からなる。代替的な方法において、ポリメラーゼは、蛍光ドナー分子で修飾されてスライドガラス上に固定化され、他方、各ヌクレオチドは、-ホスファートに付着したアクセプター蛍光部分で色分けされている。ヌクレオチドが、新規の鎖の中に組み込まれるようになる際に、システムが、蛍光タグ付加されたポリメラーゼと蛍光修飾されたヌクレオチドとの間の相互作用を検出する。他の合成によるシーケンシング技術もまた、存在する。

10

【0102】

任意の適している合成によるシーケンシングプラットフォームを、変異を特定するために使用することができる。上記のように、4種類の主要な合成によるシーケンシングプラットフォームを、現在利用可能である：Roche/454 Life Sciencesより販売されるGenome Sequencer、Illumina/Solexaより販売される1G Analyzer、Applied Biosystemsより販売されるSOLIDシステム、及びHelicos Bioscienceより販売されるHeliscopeシステム。合成によるシーケンシングプラットフォームはまた、Pacific Biosciences及びVisiGen Biotechnologiesによっても記載されている。いくつかの実施形態において、シーケンシングされる多数の核酸分子は、支持体（例えば、固体支持体）に結合している。核酸を支持体上に固定化するために、捕捉配列/万能プライミング部位を、鋳型の3'端及び/または5'端に付加することができる。核酸は、支持体に共有結合性に付着した相補的配列に対して捕捉配列をハイブリダイズすることによって、支持体に結合させることができる。捕捉配列（万能捕捉配列とも呼ばれる）は、万能プライマーとして二重に働き得る、支持体に付着した配列に対して相補的な核酸配列である。

20

30

【0103】

捕捉配列に対する代替物として、カップリングペア（例えば、抗体/抗原、受容体/リガンド、または、例えば米国特許出願第2006/0252077号に記載されているようなアビジン-ビオチンペアなど）のメンバーを、各断片に連結させて、そのカップリングペアのそれぞれの第2のメンバーでコーティングされた表面上に捕捉させることができる。

40

【0104】

捕捉に続いて、配列を、例えば、鋳型依存的な合成によるシーケンシングを含む、例えば、実施例及び米国特許第7,283,337号に記載されているような、単一分子検出/シーケンシングによって解析することができる。合成によるシーケンシングにおいて、表面に結合した分子は、ポリメラーゼの存在下で、多数のラベル化ヌクレオチド三リン酸に曝露される。鋳型の配列は、成長する鎖の3'端の中に組み込まれるラベル化ヌクレオチドの順序によって決定される。これは、リアルタイムで行うことができ、ステップ・アンド・リピートモードで行うことができる。リアルタイム解析のために、各ヌクレオチドに対して異なる光ラベルを組み込むことができ、複数のレーザーを、組み込まれた

50

ヌクレオチドの刺激のために利用することができる。

【0105】

シーケンシングはまた、他の大規模並列処理シーケンシング、または次世代シーケンシング (NGS) 技法及びプラットフォームも含むことができる。大規模並列処理シーケンシング技法及びプラットフォームの追加的な例は、Illumina HiSeq または MiSeq、ThermoPGM または Proton、PacBio RS II または Sequel、Qiagen の Gene Reader、及び Oxford Nanopore MinION である。追加的な類似した現在の大規模並列処理シーケンシング技術、及びこれらの技術の将来世代を、使用することができる。

【0106】

任意の細胞タイプまたは組織を利用して、本明細書に記載した方法における使用のための核酸試料を取得することができる。例えば、DNA または RNA 試料を、腫瘍または体液、例えば、公知の技法 (例えば、静脈穿刺) によって取得された血液、もしくは唾液から取得することができる。あるいは、核酸試験を、乾燥試料 (例えば、髪または皮膚) に対して行うことができる。加えて、試料を、シーケンシングのために腫瘍から取得することができ、別の試料を、正常組織が腫瘍と同じ組織タイプのものである場合に、シーケンシングのために正常組織から取得することができる。試料を、シーケンシングのために腫瘍から取得することができ、別の試料を、正常試料が腫瘍とは別個の組織タイプのものである場合に、シーケンシングのために正常組織から取得することができる。

【0107】

腫瘍は、肺癌、黒色腫、乳癌、卵巣癌、前立腺癌、腎臓癌、胃癌、結腸癌、精巣癌、頭頸部癌、膵臓癌、脳癌、B細胞リンパ腫、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、及びT細胞リンパ球性白血病、非小細胞肺癌、及び小細胞肺癌のうちの1つ以上を含むことができる。

【0108】

あるいは、タンパク質質量分析を使用して、腫瘍細胞上のMHCタンパク質に結合した変異したペプチドの存在を特定または実証することができる。ペプチドは、腫瘍細胞から、または腫瘍から免疫沈降させたHLA分子から酸溶出することができ、次いで、質量分析を用いて特定することができる。

【0109】

IV. 新生抗原

新生抗原は、ヌクレオチドまたはポリヌクレオチドを含むことができる。例えば、新生抗原は、ポリペプチド配列をコードするRNA配列であることができる。ワクチンにおいて有用な新生抗原は、したがって、ヌクレオチド配列またはポリペプチド配列を含むことができる。

【0110】

本明細書に開示する方法によって特定された腫瘍特異的変異を含む単離されたペプチド、公知の腫瘍特異的変異を含むペプチド、及び、本明細書に開示する方法によって特定された変異ポリペプチドまたはその断片を、本明細書に開示する。新生抗原ペプチドは、新生抗原が関連するポリペプチド配列をコードするヌクレオチド配列 (例えば、DNA または RNA) を含む場合に、それらのコード配列の文脈において記載することができる。

【0111】

新生抗原ヌクレオチド配列によってコードされる1つ以上のポリペプチドは、以下のうちの少なくとも1つを含むことができる: 1000nM未満のIC50値でのMHCとの結合親和性、MHCクラスIペプチドについてはアミノ酸8~15個、8、9、10、11、12、13、14、または15個の長さ、プロテアソーム切断を促進するペプチド内またはその近くの配列モチーフの存在、及び、TAP輸送を促進する配列モチーフの存在。MHCクラスIIのポリペプチドではアミノ酸6~30、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または30個の長さ、細胞外またはリソソームプロテア

10

20

30

40

50

ーゼ（例えば、カテプシン類）による切断またはHLA-DMにより触媒されるHLA結合を促進するペプチド内またはその近くの配列モチーフの存在。

【0112】

1つ以上の新生抗原は、腫瘍の表面上に存在することができる。

【0113】

1つ以上の新生抗原は、腫瘍を有する対象において免疫原性であることができ、例えば、対象においてT細胞応答またはB細胞応答を惹起することができる。

【0114】

対象において自己免疫応答を誘導する1つ以上の新生抗原は、腫瘍を有する対象のためのワクチン生成の文脈において、考察から排除することができる。

【0115】

少なくとも1つの新生抗原性ペプチド分子のサイズは、約5個、約6個、約7個、約8個、約9個、約10個、約11個、約12個、約13個、約14個、約15個、約16個、約17個、約18個、約19個、約20個、約21個、約22個、約23個、約24個、約25個、約26個、約27個、約28個、約29個、約30個、約31個、約32個、約33個、約34個、約35個、約36個、約37個、約38個、約39個、約40個、約41個、約42個、約43個、約44個、約45個、約46個、約47個、約48個、約49個、約50個、約60個、約70個、約80個、約90個、約100個、約110個、約120個、またはそれよりも多いアミノ分子残基、及びこれらの範囲から導出される任意の範囲を含むことができるが、それらに限定されない。具体的な実施形態において、新生抗原性ペプチド分子は、アミノ酸50個以下である。

【0116】

新生抗原性ペプチド及びポリペプチドは、MHCクラスIについては長さが15残基以下で、通常約8～約11残基の間からなり、特に9または10残基であることができ；MHCクラスIIについては、6～30残基であることができる。

【0117】

望ましい場合、より長いペプチドを、いくつかのやり方において設計することができる。1つの例において、HLAアレル上のペプチドの提示尤度が予測されるかまたは公知である場合、より長いペプチドは、(1)各々の対応する遺伝子産物のN末端及びC末端に向かって2～5アミノ酸の伸長を有する個々の提示されるペプチド；(2)各々について伸長した配列を有する、提示されるペプチドのいくつかまたはすべての連鎖のいずれかからなることができる。別の例において、シーケンシングにより、腫瘍中に存在する長い(10残基より長い)新生エピトープ配列(例えば、新規のペプチド配列をもたらすフレームシフト、リードスルー、またはイントロンの包含による)が明らかになる場合、より長いペプチドは、(3)新規の腫瘍特異的アミノ酸のストレッチ全体からなることになり、したがって、最強のHLAに提示されるより短いペプチドの計算的またはインビトロ試験ベースの選択の必要を回避する。いずれの例においても、より長いペプチドの使用によって、患者細胞による内因性のプロセッシングが可能になり、より有効な抗原提示及びT細胞応答の誘導がもたらされ得る。

【0118】

新生抗原性ペプチド及びポリペプチドは、HLAタンパク質上に提示されることができる。いくつかの態様において、新生抗原性ペプチド及びポリペプチドは、野生型ペプチドよりも強い親和性でHLAタンパク質上に提示される。いくつかの態様において、新生抗原性ペプチドまたはポリペプチドは、少なくとも5000nM未満、少なくとも1000nM未満、少なくとも500nM未満、少なくとも250nM未満、少なくとも200nM未満、少なくとも150nM未満、少なくとも100nM未満、少なくとも50nM未満、またはそれよりも小さいIC50を有することができる。

【0119】

いくつかの態様において、新生抗原性ペプチド及びポリペプチドは、対象に投与された場合に、自己免疫応答を誘導せず、及び/または免疫寛容を引き起こさない。

10

20

30

40

50

【0120】

また、少なくとも2種類以上の新生抗原性ペプチドを含む組成物も提供する。いくつかの実施形態において、組成物は、少なくとも2種類の異なるペプチドを含有する。少なくとも2種類の異なるペプチドは、同じポリペプチドに由来することができる。異なるポリペプチドとは、ペプチドが、長さ、アミノ酸配列、またはその両方において異なることを意味する。ペプチドは、腫瘍特異的変異を含有することが知られているか、または見出されている任意のポリペプチドに由来する。新生抗原性ペプチドが由来することができる、適しているポリペプチドは、例えば、COSMICデータベースにおいて見出すことができる。COSMICは、ヒトがんにおける体細胞性変異についての総合的な情報の管理を行う。ペプチドは、腫瘍特異的変異を含有する。いくつかの態様において、腫瘍特異的変異は、特定のがんタイプについてのドライバー変異である。

10

【0121】

望ましい活性または性質を有する新生抗原性ペプチド及びポリペプチドは、望ましいMHC分子に結合して適切なT細胞を活性化する非改変ペプチドの生物学的活性を増大させるかまたは実質的にそのすべてを少なくとも保持しつつ、特定の望ましい属性、例えば、改善された薬理学的特徴を与えるように改変することができる。例として、新生抗原性ペプチド及びポリペプチドを、保存的または非保存的のいずれかの置換などの、種々の改変にさらに供することができる。そのような改変は、改善されたMHC結合、安定性、または提示などの、それらの使用におけるある特定の利点を提供し得る。保存的置換とは、アミノ酸残基を、生物学的及び/または化学的に類似している別のもので、例えば、1つの疎水性残基を別の疎水性残基、または1つの極性残基を別の極性残基で置き換えることを意味する。置換は、Gly、Ala; Val、Ile、Leu、Met; Asp、Glu; Asn、Gln; Ser、Thr; Lys、Arg; 及びPhe、Tyrなどの組み合わせを含む。単一アミノ酸置換の効果はまた、D-アミノ酸を用いて探査してもよい。そのような改変は、例えば、Merrifield, Science 232:341-347 (1986), Barany & Merrifield, The Peptides, Gross & Meienhofer, eds. (N.Y., Academic Press), pp. 1-284 (1979); 及びStewart & Young, Solid Phase Peptide Synthesis, (Rockford, Ill., Pierce), 2d Ed. (1984)に記載されているように、周知のペプチド合成手順を用いて行うことができる。

20

30

【0122】

種々のアミノ酸模倣物または非天然アミノ酸でのペプチド及びポリペプチドの改変は、インビボでのペプチド及びポリペプチドの安定性の増大に特に有用である場合がある。安定性は多くの方法でアッセイすることができる。例として、ペプチダーゼ、ならびに、ヒト血漿及び血清などの種々の生物学的媒質が、安定性を試験するために使用されている。例えば、Verhoef et al., Eur. J. Drug Metab Pharmacokin. 11:291-302 (1986)を参照されたい。ペプチドの半減期は、25%ヒト血清(v/v)アッセイを用いて好都合に決定することができる。プロトコールは、概して以下のようなものである。プールしたヒト血清(タイプAB、非熱不活性化)を、使用前に遠心分離によって脱脂する。次いで、血清を、RPMI組織培養培地で25%に希釈し、ペプチド安定性を試験するために使用する。あらかじめ決定された時間間隔で、少量の反応溶液を取り出して、6%水性トリクロロ酢酸またはエタノールのいずれかに添加する。濁った反応試料を15分間冷却(4℃)し、次いで、スピンドして沈降血清タンパク質を沈殿させる。次いで、ペプチドの存在を、安定性特異的クロマトグラフィー条件を用いた逆相HPLCによって決定する。

40

【0123】

ペプチド及びポリペプチドを、改善された血清半減期以外の望ましい属性を提供するために修飾することができる。例として、CTL活性を誘導するペプチドの能力を、Tヘルパー細胞応答を誘導することができる少なくとも1つのエピトープを含有する配列への連

50

結によって増強することができる。免疫原性ペプチド/Tヘルパーコンジュゲートは、スペーサー分子によって連結することができる。スペーサーは、典型的には、生理学的条件下で実質的に無電荷である、アミノ酸またはアミノ酸模倣物などの相対的に小さな中性分子から構成される。スペーサーは、典型的には、例えば、Ala、Gly、または、非極性アミノ酸もしくは中性極性アミノ酸の他の中性スペーサーから選択される。任意で存在するスペーサーは、同じ残基から構成される必要はなく、したがって、ヘテロオリゴマーまたはホモオリゴマーであり得ることが、理解されるであろう。存在する場合、スペーサーは、通常、少なくとも1または2残基、より通常は、3~6残基であろう。あるいは、ペプチドを、スペーサーなしでTヘルパーペプチドに連結することができる。

【0124】

新生抗原性ペプチドは、ペプチドのアミノ末端またはカルボキシ末端のいずれかで、直接またはスペーサーを介してのいずれかでTヘルパーペプチドに連結することができる。新生抗原性ペプチドまたはTヘルパーペプチドのいずれかのアミノ末端を、アシル化することができる。例示的なTヘルパーペプチドは、破傷風毒素の830~843、インフルエンザの307~319、マラリアスポロゾイトの周囲382~398及び378~389を含む。

【0125】

タンパク質またはペプチドは、標準的な分子生物学的技法を通じたタンパク質、ポリペプチド、もしくはペプチドの発現、天然由来源からのタンパク質もしくはペプチドの単離、またはタンパク質もしくはペプチドの化学合成を含む、当業者に公知の任意の技法によって作製することができる。種々の遺伝子に対応する、ヌクレオチドならびにタンパク質、ポリペプチド及びペプチドの配列は、以前に開示されており、当業者に公知のコンピュータ処理されたデータベースで見出すことができる。1つのそのようなデータベースは、National Institutes of Healthのウェブサイトに位置する、National Center for Biotechnology InformationのGenbank及びGenPeptデータベースである。公知の遺伝子のコード領域は、本明細書に開示する技法を用いて、または当業者に公知であるように、増幅及び/または発現させることができる。あるいは、タンパク質、ポリペプチド、及びペプチドの種々の商業的調製物が、当業者に公知である。

【0126】

さらなる態様において、新生抗原は、新生抗原性ペプチドまたはその一部をコードする核酸(例えば、ポリヌクレオチド)を含む。ポリヌクレオチドは、例えば、DNA、cDNA、PNA、CNA、RNA(例えば、mRNA)、例えば、ホスホロチオアートバックボーンを有するポリヌクレオチドなどの、ポリヌクレオチドの一本鎖及び/もしくは二本鎖、または天然形態もしくは安定化形態のいずれか、または、それらの組み合わせであることができ、イントロンを含有してもよく、または含有しなくてもよい。またさらなる態様は、ポリペプチドまたはその一部を発現することができる発現ベクターを提供する。様々な細胞タイプ用の発現ベクターが、当技術分野において周知であり、過度の実験なしで選択することができる。概して、DNAを、プラスミドなどの発現ベクター中に、発現のための適正な方向及び正確なリーディングフレームで挿入する。必要な場合は、DNAを、望ましい宿主によって認識される適切な転写及び翻訳調節性制御ヌクレオチド配列に連結することができるが、そのような制御は、概して発現ベクターにおいて利用可能である。次いで、ベクターを、標準的な技法を通して宿主中に導入する。手引きは、例えば、Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.において見出すことができる。

【0127】

IV. ワクチン組成物

また、特異的な免疫応答、例えば、腫瘍特異的な免疫応答を生じることができる免疫原

10

20

30

40

50

性組成物、例えば、ワクチン組成物も、本明細書に開示する。ワクチン組成物は、典型的に、例えば、本明細書に記載した方法を用いて選択された多数の新生抗原を含む。ワクチン組成物はまた、ワクチンと呼ぶこともできる。

【0128】

ワクチンは、1～30種類のペプチド、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、もしくは30種類の異なるペプチド、6、7、8、9、10、11、12、13、もしくは14種類の異なるペプチド、または12、13、もしくは14種類の異なるペプチドを含有することができる。ペプチドは、翻訳後修飾を含むことができる。ワクチンは、1～100種類もしくはそれよりも多いヌクレオチド配列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100種類もしくはそれよりも多い異なるヌクレオチド配列、6、7、8、9、10、11、12、13、もしくは14種類の異なるヌクレオチド配列、または12、13、もしくは14種類の異なるヌクレオチド配列を含有することができる。ワクチンは、1～30種類の新生抗原配列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100種類もしくはそれよりも多い異なる新生抗原配列、6、7、8、9、10、11、12、13、もしくは14種類の異なる新生抗原配列、または12、13、もしくは14種類の異なる新生抗原配列を含有することができる。

10

20

30

【0129】

一実施形態では、異なるペプチド及び/もしくはポリペプチド、またはそれらをコードするヌクレオチド配列は、ペプチド及び/またはポリペプチドが、異なるMHCクラスI分子及び/または異なるMHCクラスII分子などの異なるMHC分子と結合することができるように選択される。いくつかの態様において、1つのワクチン組成物は、最も頻繁に存在するMHCクラスI分子及び/またはMHCクラスII分子と結合することができるペプチド及び/またはポリペプチドのコード配列を含む。したがって、ワクチン組成物は、少なくとも2種類の好ましい、少なくとも3種類の好ましい、または少なくとも4種類の好ましいMHCクラスI分子及び/またはMHCクラスII分子と結合することができる異なる断片を含むことができる。

40

【0130】

ワクチン組成物は、特異的な細胞傷害性T細胞応答、及び/または特異的なヘルパーT細胞応答を生じることができる。

【0131】

ワクチン組成物は、アジュバント及び/または担体をさらに含むことができる。有用なアジュバント及び担体の例を、本明細書の下記に示す。組成物は、例えば、タンパク質などの担体、または、例えば、T細胞に対してペプチドを提示することができる樹状細胞(DC)などの抗原提示細胞と結合することができる。

【0132】

50

アジュバントは、ワクチン組成物中へのその混合が、新生抗原に対する免疫応答を増大させるか、または別の方法で修飾する任意の物質である。担体は、新生抗原がそれに結合することができる足場構造、例えば、ポリペプチドまたは多糖であることができる。任意で、アジュバントは、共有結合性または非共有結合性にコンジュゲートされる。

【0133】

抗原に対する免疫応答を増大させるアジュバントの能力は、典型的に、免疫媒介性反応の有意なもしくは実質的な増大、または疾患症候の低減によって明示される。例えば、体液性免疫の増大は、典型的に、抗原に対して生じた抗体の力価の有意な増大によって明示され、T細胞活性の増大は、典型的に、細胞増殖、または細胞性細胞傷害、またはサイトカイン分泌の増大において明示される。アジュバントはまた、例えば、主として体液性またはTh応答を、主として細胞性またはTh応答へと変更することによって、免疫応答を変化させ得る。

10

【0134】

適しているアジュバントは、1018 ISS、アラム、アルミニウム塩、Amplivax、AS15、BCG、CP-870、893、CpG7909、CyaA、dSLIM、GM-CSF、IC30、IC31、イミキモド、ImuFact IMP321、IS Patch、ISS、ISCOMATRIX、JuvImmune、LipoVac、MF59、モノホスホリル脂質A、Montanide IMS 1312、Montanide ISA206、Montanide ISA 50V、Montanide ISA-51、OK-432、OM-174、OM-197-MP-EC、ONTAK、PepTelベクターシステム、PLGマイクロ粒子、レシキモド、SRL172、ピロソーム及び他のウイルス様粒子、YF-17D、VEGFトラップ、R848、グルカン、Pam3Cys、サポニンに由来するAquila's QS21 stimulon (Aquila Biotech, Worcester, Mass., USA)、マイコバクテリア抽出物及び合成細菌細胞壁模倣物、及びRibi's Detox. QuilまたはSuperfosなどの他の専売アジュバントを含むが、それらに限定されない。不完全フロインドまたはGM-CSFなどのアジュバントが、有用である。樹状細胞に特異的なくつかの免疫学的アジュバント(例えば、MF59)及びそれらの調製物が、以前に記載されている(Dupuis M, et al., Cell Immunol. 1998; 186(1): 18-27; Allison AC; Dev Biol Stand. 1998; 92: 3-11)。また、サイトカインを使用することもできる。いくつかのサイトカインは、リンパ組織に対する樹状細胞の遊走への影響(例えば、TNF-)、Tリンパ球に対する効率的な抗原提示細胞への樹状細胞の成熟の加速化(例えば、GM-CSF、IL-1、及びIL-4)(具体的にその全体が参照により本明細書に組み入れられる、米国特許第5,849,589号)、及び免疫アジュバントとしての作用(例えば、IL-12)に直接結び付けられている(Gabrilovich DI, et al., J Immunother Emphasis Tumor Immunol. 1996(6): 414-418)。

20

30

【0135】

CpG免疫刺激性オリゴヌクレオチドもまた、ワクチン設定においてアジュバントの効果を増強することが報告されている。TLR 7、TLR 8、及び/またはTLR 9に結合するRNAなどの他のTLR結合分子がまた、使用されてもよい。

40

【0136】

有用なアジュバントの他の例は、化学的に修飾されたCpG(例えば、CpR、Idera)、Poly(I:C)(例えば、polyi:C I2U)、非CpG細菌DNAまたはRNA、ならびに、治療的に及び/またはアジュバントとして作用し得る、シクロホスファミド、スニチニブ、ベバシズマブ、セレブレックス、NCX-4016、シルデナフィル、タダラフィル、バルデナフィル、ソラフィニブ、XL-999、CP-547632、パゾパニブ、ZD2171、AZD2171、イピリムマブ、トレメリムマブ、及びSC58175などの免疫活性小分子及び抗体を含むが、それらに限定されない。アジ

50

ュバント及び添加物の量及び濃度は、当業者が過度の実験なしで容易に決定することができる。追加的なアジュバントは、顆粒球マクロファージコロニー刺激因子（GM-CSF、サルグラモスチム）などのコロニー刺激因子を含む。

【0137】

ワクチン組成物は、1種類よりも多い異なるアジュバントを含むことができる。さらに、治療用組成物は、上記の任意またはそれらの組み合わせを含む、任意のアジュバント物質を含むことができる。ワクチン及びアジュバントを、任意の適切な配列において、一緒にまたは別々に投与できることもまた、企図される。

【0138】

担体（または賦形剤）は、アジュバントから独立して存在することができる。担体の機能は、例えば、活性または免疫原性を増大させるため、安定性を与えるため、生物学的活性を増大させるため、または血清半減期を増大させるために、特に変異体の分子量を増大させることであり得る。さらに、担体は、T細胞に対してペプチドを提示するのを助けることができる。担体は、当業者に公知の任意の適している担体、例えば、タンパク質または抗原提示細胞であることができる。担体タンパク質は、キーホールリンペットヘモシアニン、トランスフェリンなどの血清タンパク質、ウシ血清アルブミン、ヒト血清アルブミン、サイログロブリンもしくはオボアルブミン、免疫グロブリン、またはインスリンなどのホルモン、またはパルミチン酸であることができるが、それらに限定されない。ヒトの免疫化のためには、担体は概して、ヒトに許容されかつ安全な、生理学的に許容される担体である。しかし、破傷風トキソイド及び/またはジフテリアトキソイドは、適している担体である。あるいは、担体は、デキストラン、例えばセファロースであることができる。

10

20

【0139】

細胞傷害性T細胞（CTL）は、無傷の外来抗原自体よりも、MHC分子に結合したペプチドの形態において抗原を認識する。MHC分子自体は、抗原提示細胞の細胞表面に位置する。したがって、CTLの活性化は、ペプチド抗原、MHC分子、及びAPCの三量体複合体が存在する場合に可能である。対応して、ペプチドがCTLの活性化のために使用される場合だけではなく、追加的にそれぞれのMHC分子を有するAPCが添加される場合に、それは免疫応答を増強し得る。したがって、いくつかの実施形態において、ワクチン組成物は、追加的に、少なくとも1つの抗原提示細胞を含有する。

30

【0140】

新生抗原はまた、ワクシニア、鶏痘、自己複製アルファウイルス、マラバウイルス、アデノウイルス（例えば、Tatsis et al., *Adenoviruses, Molecular Therapy* (2004) 10, 616-629を参照されたい）、または、第2、第3、もしくはハイブリッド第2/第3世代のレンチウイルス、及び特異的な細胞タイプもしくは受容体を標的とするように設計された任意の世代の組換えレンチウイルスを含むがそれらに限定されないレンチウイルス（例えば、Hu et al., *Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases, Immunol Rev.* (2011) 239(1): 45-61、Sakuma et al., *Lentiviral vectors: basic to translational, Biochem J.* (2012) 443(3): 603-18、Cooper et al., *Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, Nucl. Acids Res.* (2015) 43(1): 682-690、Zufferey et al., *Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery, J. Virol.* (1998) 72(12): 9873-9880を参照されたい)などの、ウイルスベクターベースの

40

50

ワクチンプラットフォームに含めることもできる。上述のウイルスベクターベースのワクチンプラットフォームのパッケージング能力に依存して、このアプローチは、1つ以上の新生抗原ペプチドをコードする1つ以上のヌクレオチド配列を送達することができる。配列は、非変異配列が隣接していてもよく、リンカーによって分離されていてもよく、または、細胞内区画を標的とする1つもしくは複数の配列が先行していてもよい(例えば、Gros et al., *Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients*, *Nat Med.* (2016) 22(4): 433-8、Stronen et al., *Targeting of cancer neoantigens with donor-derived T cell receptor repertoires*, *Science.* (2016) 352(6291): 1337-41、Lu et al., *Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions*, *Clin Cancer Res.* (2014) 20(13): 3401-10を参照されたい)。宿主中への導入時に、感染した細胞は、新生抗原を発現し、それにより、ペプチドに対する宿主免疫(例えば、CTL)応答を惹起する。免疫化プロトコールにおいて有用なワクチンベクター及び方法は、例えば、米国特許第4,722,848号に記載されている。別のベクターは、BCG(カルメット・ゲラン桿菌)である。BCGベクターは、Stover et al. (*Nature* 351: 456-460 (1991))に記載されている。新生抗原の治療的投与または免疫化に有用な、多種多様の他のワクチンベクター、例えば、チフス菌(*Salmonella typhi*)ベクターなどが、本明細書における記載から当業者に明らかであろう。

【0141】

IV.A. ワクチン設計及び製造のさらなる考慮事項

IV.A.1. すべての腫瘍サブクローンのカバーするペプチドのセットの決定

すべての、または大部分の腫瘍サブクローンによって提示されるものを意味するトランカルペプチド(truncal peptide)が、ワクチン中への包含について優先される⁵³。任意で、高い確率で提示されかつ免疫原性であることが予測されるトランカルペプチドがない場合、または、高い確率で提示されかつ免疫原性であることが予測されるトランカルペプチドの数が、追加的な非トランカルペプチドをワクチンに含めることができるほど少ない場合には、腫瘍サブクローンの数及び同一性を推定すること、及びワクチンによってカバーされる腫瘍サブクローンの数を最大化するようにペプチドを選ぶことによって、さらなるペプチドを優先順位付けすることができる⁵⁴。

【0142】

IV.A.2. 新生抗原の優先順位決定

上記の新生抗原フィルターのすべてを適用した後、ワクチン技術が対応できるよりも多くの候補新生抗原が、依然としてワクチン包含に利用可能である可能性がある。追加的に、新生抗原解析の種々の態様についての不確定度が残っている可能性があり、候補ワクチン新生抗原の様々な性状の間にトレードオフが存在する可能性がある。したがって、選択プロセスの各段階でのあらかじめ決定されたフィルターの代わりに、少なくとも以下の軸を有する空間に候補新生抗原を置き、積分アプローチを用いて選択を最適化する、積分多次元モデルを考えることができる。

1. 自己免疫または寛容のリスク(生殖細胞系列のリスク)(より低い自己免疫のリスクが、典型的に好ましい)
2. シークエンシングアーチファクトの確率(より低いアーチファクトの確率が、典型的に好ましい)
3. 免疫原性の確率(より高い免疫原性の確率が、典型的に好ましい)
4. 提示の確率(より高い提示の確率が、典型的に好ましい)

5. 遺伝子発現（より高い発現が、典型的に好ましい）

6. HLA遺伝子のカバレッジ（新生抗原のセットの提示に関与する、より多い数のHLA分子は、腫瘍が、HLA分子の下方制御または変異を介して免疫攻撃を回避する確率を低くする可能性がある）

7. HLAクラスのカバレッジ（HLA-I及びHLA-IIの両方をカバーすることで、治療応答の確率が高まり、腫瘍の免疫回避の確率が低くなる可能性がある）

【0143】

V. 治療及び製造方法

本明細書に開示する方法を用いて特定された複数の新生抗原などの1つ以上の新生抗原を対象に投与することにより、対象に腫瘍特異的な免疫応答を誘導し、腫瘍に対するワクチン接種を行い、対象のがんの症状を治療及び/または緩和する方法も提供される。

10

【0144】

いくつかの態様において、対象は、がんと診断されているか、またはがんを発症するリスクにある。対象は、ヒト、イヌ、ネコ、ウマ、または、腫瘍特異的な免疫応答が望ましい任意の動物であることができる。腫瘍は、乳、卵巣、前立腺、肺、腎臓、胃、結腸、精巣、頭頸部、膵臓、脳、黒色腫、及び他の組織器官の腫瘍などの、任意の固形腫瘍、ならびに、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、T細胞リンパ球性白血病、及びB細胞リンパ腫を含むリンパ腫及び白血病などの、血液腫瘍であることができる。

【0145】

20

新生抗原は、CTL応答を誘導するのに十分な量で投与することができる。

【0146】

新生抗原は、単独で、または他の治療用物質との組み合わせで投与することができる。治療用物質は、例えば、化学療法剤、放射線、または免疫療法である。特定のがんのための任意の適している治療的処置を、施すことができる。

【0147】

加えて、対象に、チェックポイント阻害因子などの抗免疫抑制性/免疫刺激性物質をさらに投与することができる。例えば、対象に、抗CTLA抗体または抗PD-1または抗PD-L1をさらに投与することができる。抗体によるCTLA-4またはPD-L1の遮断は、患者においてがん性細胞に対する免疫応答を増強することができる。特に、CTLA-4遮断は、ワクチン接種プロトコルを採用した場合に有効であることが示されている。

30

【0148】

ワクチン組成物に含まれるべき各新生抗原の最適量、及び最適投薬レジメンを、決定することができる。例えば、新生抗原またはその変異体は、静脈内(i.v.)注射、皮下(s.c.)注射、皮内(i.d.)注射、腹腔内(i.p.)注射、筋肉内(i.m.)注射のために調製することができる。注射の方法は、s.c.、i.d.、i.p.、i.m.、及びi.v.を含む。DNAまたはRNA注射の方法は、i.d.、i.m.、s.c.、i.p.、及びi.v.を含む。ワクチン組成物の投与の他の方法は、当業者に公知である。

40

【0149】

ワクチンは、組成物中に存在する新生抗原の選択、数、及び/または量が、組織、がん、及び/または患者に特異的であるように編集することができる。例として、ペプチドの厳密な選択は、所定の組織における親タンパク質の発現パターンによって手引きされ得る。選択は、がんの特異的なタイプ、疾患の状態、より早期の処置レジメン、患者の免疫状態、及び当然、患者のHLAハロタイプに依存し得る。さらに、ワクチンは、特定の患者の個人的な必要にしたがって、個別化された構成要素を含有することができる。例は、特定の患者における新生抗原の発現にしたがって新生抗原の選択を変えること、または、処置の第1のラウンドまたはスキームの後の二次的処置についての調整を含む。

【0150】

50

がんのためのワクチンとして使用されるべき組成物について、正常組織において多量に発現している類似した正常な自己ペプチドを有する新生抗原は、本明細書に記載した組成物において、避けられるか、または少量で存在することができる。他方で、患者の腫瘍が、多量のある特定の新生抗原を発現することが公知である場合、このがんの処置のためのそれぞれの薬学的組成物は、多量に存在することができ、及び/または、この特定の新生抗原もしくはこの新生抗原の経路に特異的な1種類よりも多い新生抗原を含めることができる。

【0151】

新生抗原を含む組成物を、既にごんを患っている個体に投与することができる。治療的適用において、組成物は、腫瘍抗原に対する有効なCTL応答を惹起し、かつ、症候及び/または合併症を治癒するかまたは少なくとも部分的に停止するのに十分な量で、患者に投与される。これを達成するのに妥当な量を、「治療的有效用量」として定義する。この用途のために有効な量は、例えば、組成物、投与の様式、処置される疾患の病期及び重症度、患者の体重及び健康の全身状態、ならびに処方医の判断に依存するであろう。組成物は、概して、重篤な疾患状態、すなわち、命に関わるか、または潜在的に命に関わる状況、特にがんが転移している場合に使用できることを、心に留めるべきである。そのような例において、外来性物質の最小化、及び新生抗原の相対的な非毒性の性質を考慮して、実質的過剰量のこれらの組成物を投与することが、可能であり、かつ処置する医師が望ましいと感じることができる。

10

【0152】

治療的用途のために、投与は、腫瘍の検出または外科的除去時に始めることができる。これに、少なくとも症候が実質的に減ずるまで、及びその後ある期間にわたって、ブースト用量が続く。

20

【0153】

治療的処置のための薬学的組成物（例えば、ワクチン組成物）は、非経口、局部、経鼻、経口、または局所投与について意図される。薬学的組成物は、非経口的に、例えば、静脈内、皮下、皮内、または筋肉内に投与することができる。組成物は、腫瘍に対する局所免疫応答を誘導するために、外科的切除の部位に投与することができる。新生抗原の溶液を含む非経口投与用の組成物を、本明細書に開示し、ワクチン組成物は、許容される担体、例えば、水性担体に溶解または懸濁される。様々な水性担体、例えば、水、緩衝水、0.9%食塩水、0.3%グリシン、ヒアルロン酸などを使用することができる。これらの組成物は、従来周知の滅菌技法によって滅菌することができ、または滅菌濾過することができる。結果として生じた水溶液を、そのまま使用のためにパッケージングするか、または凍結乾燥することができ、凍結乾燥調製物は、投与前に滅菌溶液と組み合わせられる。組成物は、pH調整剤及び緩衝剤、等張化剤、湿潤剤など、例えば、酢酸ナトリウム、乳酸ナトリウム、塩化ナトリウム、塩化カリウム、塩化カルシウム、ソルビタンモノラウレート、トリエタノールアミンオレートなどのような、生理学的条件に近づけるために必要とされる、薬学的に許容される補助物質を含有してもよい。

30

【0154】

新生抗原はまた、それらをリンパ組織などの特定の細胞組織にターゲティングする、リポソームを介して投与することもできる。リポソームはまた、半減期を増大させるのにも有用である。リポソームは、エマルジョン、フォーム、ミセル、不溶性単層、液晶、リン脂質分散物、ラメラ層などを含む。これらの調製物において、送達されるべき新生抗原は、単独で、または、CD45抗原に結合するモノクローナル抗体などの、例えば、リンパ系細胞の間で優性な受容体に結合する分子、または他の治療用組成物もしくは免疫原性組成物と共に、リポソームの一部として組み込まれる。したがって、所望の新生抗原で満たされたリポソームは、リンパ系細胞の部位へ方向付けられることができ、そこで、リポソームは次いで、選択された治療用/免疫原性組成物を送達する。リポソームは、概して、中性及び負電荷を有するリン脂質、及びコレステロールなどのステロールを含む、標準的な小胞形成脂質から形成され得る。脂質の選択は、概して、例えば、リポソームサイズ、

40

50

酸不安定性、及び血流におけるリポソームの安定性の考慮により手引きされる。例えば、Szoka et al., Ann. Rev. Biophys. Bioeng. 9; 467 (1980)、米国特許第4,235,871号、第4,501,728号、第4,501,728号、第4,837,028号、及び第5,019,369号に記載されているように、様々な方法を、リポソームを調製するために利用可能である。

【0155】

免疫細胞へのターゲティングのために、リポソーム中に組み込まれるべきリガンドは、例えば、所望の免疫系細胞の細胞表面決定基に特異的な抗体またはその断片を含むことができる。リポソーム懸濁液は、とりわけ、投与の様式、送達されるペプチド、及び処置される疾患の病期にしたがって変動する用量で、静脈内、局所、局部などに投与することができる。

10

【0156】

治療目的または免疫化目的で、本明細書に記載したペプチド、及び任意でペプチドの1つ以上をコードする核酸をまた、患者に投与することもできる。数多くの方法が、核酸を患者に送達するために好都合に使用される。例として、核酸を、「裸のDNA」として直接送達することができる。このアプローチは、例として、Wolf et al., Science 247:1465-1468 (1990)、ならびに米国特許第5,580,859号及び第5,589,466号に記載されている。核酸はまた、例として、米国特許第5,204,253号に記載されているような弾道送達を用いて投与することもできる。単にDNAからなる粒子を、投与することができる。あるいは、DNAを、金粒子などの粒子に接着させることができる。核酸配列を送達するためのアプローチは、エレクトロポレーションを伴うかまたは伴わない、ウイルスベクター、mRNAベクター、及びDNAベクターを含むことができる。

20

【0157】

核酸はまた、カチオン性脂質などのカチオン性化合物に複合体化させて送達することもできる。脂質媒介性遺伝子送達法は、例として、9618372WOAWO 96/18372; 9324640WOAWO 93/24640; Mannino & Gould-Fogerite, BioTechniques 6(7): 682-691 (1988); 米国特許第5,279,833号 Rose、米国特許第5,279,833号; 9106309WOAWO 91/06309; 及び Felgner et al., Proc. Natl. Acad. Sci. USA 84: 7413-7414 (1987)に記載されている。

30

【0158】

新生抗原はまた、ワクシニア、鶏痘、自己複製アルファウイルス、マラバウイルス、アデノウイルス(例えば、Tatsis et al., Adenoviruses, Molecular Therapy (2004) 10, 616-629を参照されたい)、または、第2、第3、もしくはハイブリッド第2/第3世代のレンチウイルス、及び特異的な細胞タイプもしくは受容体を標的とするように設計された任意の世代の組換えレンチウイルスを含むがそれらに限定されないレンチウイルス(例えば、Hu et al., Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases, Immunol Rev. (2011) 239(1): 45-61、Sakuma et al., Lentiviral vectors: basic to translational, Biochem J. (2012) 443(3): 603-18、Cooper et al., Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, Nucl. Acids Res. (2015) 43(1): 682-690、Zufferey et al., Self-Inactivating Lentivirus Vector f

40

50

or Safe and Efficient In Vivo Gene Delivery, *J. Virol.* (1998) 72 (12): 9873-9880を参照されたい)などの、ウイルスベクターベースのワクチンプラットフォームに含めることもできる。上述のウイルスベクターベースのワクチンプラットフォームのパッケージング能力に依存して、このアプローチは、1つ以上の新生抗原ペプチドをコードする1つ以上のヌクレオチド配列を送達することができる。配列は、非変異配列が隣接していてもよく、リンカーによって分離されていてもよく、または、細胞内区画を標的とする1つもしくは複数の配列が先行していてもよい(例えば、Gros et al., *Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients*, *Nat Med.* (2016) 22 (4): 433-8、Stronen et al., *Targeting of cancer neoantigens with donor-derived T cell receptor repertoires*, *Science.* (2016) 352 (6291): 1337-41、Lu et al., *Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions*, *Clin Cancer Res.* (2014) 20 (13): 3401-10を参照されたい)。宿主中への導入時に、感染した細胞は、新生抗原を発現し、それにより、ペプチドに対する宿主免疫(例えば、CTL)応答を惹起する。免疫化プロトコールにおいて有用なワクシニアベクター及び方法は、例えば、米国特許第4,722,848号に記載されている。別のベクターは、BCG(カルメット・ゲラン桿菌)である。BCGベクターは、Stover et al. (*Nature* 351: 456-460 (1991))に記載されている。新生抗原の治療的投与または免疫化に有用な、多種多様の他のワクチンベクター、例えば、チフス菌ベクターなどが、本明細書における記載から当業者に明らかであろう。

【0159】

核酸を投与する手段は、1つ以上のエピトープをコードするミニ遺伝子構築物を使用する。ヒト細胞における発現のための、選択されたCTLエピトープをコードするDNA配列(ミニ遺伝子)を作製するために、エピトープのアミノ酸配列を逆翻訳する。各アミノ酸に対するコドン選択を手引きするために、ヒトコドン使用頻度表を使用する。これらのエピトープをコードするDNA配列を、直接隣り合わせて、連続的なポリペプチド配列を作製する。発現及び/または免疫原性を最適化するために、追加の要素を、ミニ遺伝子設計中に組み入れることができる。逆翻訳して、ミニ遺伝子配列に含めることができるアミノ酸配列の例は、ヘルパーTリンパ球エピトープ、リーダー(シグナル)配列、及び小胞体保持シグナルを含む。加えて、CTLエピトープのMHC提示は、CTLエピトープに近接した合成の(例えば、ポリアラニン)または天然に存在するフランキング配列を含むことによって、改善することができる。ミニ遺伝子配列は、ミニ遺伝子のプラス鎖及びマイナス鎖をコードするオリゴヌクレオチドをアSEMBLすることによって、DNAに変換される。オーバーラップするオリゴヌクレオチド(30~100塩基長)を、周知の技法を用いて適切な条件下で、合成し、リン酸化し、精製し、アニーリングする。オリゴヌクレオチドの端は、T4 DNAリガーゼを用いて連結する。CTLエピトープポリペプチドをコードするこの合成ミニ遺伝子を、次いで、望ましい発現ベクター中にクローニングすることができる。

【0160】

精製プラスミドDNAは、様々な製剤を用いて、注射のために調製することができる。これらのうちでもっとも単純なものは、滅菌リン酸緩衝食塩水(PBS)における凍結乾燥DNAの再構成である。様々な方法が記載されており、新たな技法が利用可能になり得る。上記で言及したように、核酸は、カチオン性脂質で好都合に製剤化される。加えて、糖脂質、融合性リポソーム、ペプチド、及び保護的、相互作用的、非縮合性(PINC)

と集合的に呼ばれる化合物もまた、精製プラスミドDNAと複合体化させて、安定性、筋肉内分散、または特異的な器官もしくは細胞タイプへの輸送などの変数に影響を及ぼすことができる。

【0161】

また、本明細書に開示する方法の工程を行うこと；及び、多数の新生抗原または多数の新生抗原のサブセットを含む腫瘍ワクチンを生産する工程を含む、腫瘍ワクチンを製造する方法も、本明細書に開示する。

【0162】

本明細書に開示する新生抗原は、当技術分野において公知の方法を用いて製造することができる。例えば、本明細書に開示する新生抗原またはベクター（例えば、1つ以上の新生抗原をコードする少なくとも1つの配列を含むベクター）を生産する方法は、新生抗原またはベクターを発現するのに適している条件下で宿主細胞を培養する工程であって、宿主細胞が、新生抗原またはベクターをコードする少なくとも1つのポリヌクレオチドを含む工程、及び、新生抗原またはベクターを精製する工程を含むことができる。標準的な精製法は、クロマトグラフィー技法、電気泳動技法、免疫学的技法、沈降技法、透析技法、濾過技法、濃縮技法、及びクロマトフォーカシング技法を含む。

10

【0163】

宿主細胞は、チャイニーズハムスター卵巣（CHO）細胞、NS0細胞、酵母、またはHEK293細胞を含むことができる。宿主細胞は、本明細書に開示する新生抗原またはベクターをコードする少なくとも1つの核酸配列を含む、1つ以上のポリヌクレオチドで形質転換することができ、任意で、単離されたポリヌクレオチドは、新生抗原またはベクターをコードする少なくとも1つの核酸配列に機能的に連結されたプロモーター配列をさらに含む。ある特定の実施形態において、単離されたポリヌクレオチドは、cDNAであることができる。

20

【0164】

VI. 新生抗原の特定

VI. A. 新生抗原候補の特定

腫瘍及び正常のエクソーム及びトランスクリプトームのNGS解析のための研究法を、新生抗原の特定のスペースに記載し、適用している^{6, 14, 15}。下記の例は、臨床設定における新生抗原の特定について、より大きな感度及び特異性のためのある特定の最適化を考慮している。これらの最適化は、実験室プロセスに関連するもの及びNGSデータ解析に関連するものの、2つの区域にグループ化することができる。

30

【0165】

VI. A. 1. 実験室プロセスの最適化

本明細書に提示したプロセスの改善は、標的とされるがんパネルにおける信頼できるがんドライバー遺伝子の評価について開発された概念¹⁶を、新生抗原の特定のために必要な全エクソーム設定及び全トランスクリプトーム設定に拡大することによって、低い腫瘍含量及び少ない体積の臨床標本からの高精度の新生抗原の発見における難題に対処する。具体的には、これらの改善は、以下を含む：

1. 低い腫瘍含量またはサブクローン状態のいずれかにより、低い変異体アレル頻度で存在する変異を検出するための、腫瘍エクソームにわたる深い（500xよりも大きい）固有の平均カバレッジのターゲティング。

40

2. 可能性のある新生抗原の見逃しが最も少ないように、100x未満でカバーされる塩基が5%未満である、例として、

a. 個々のプローブQCを有するDNAベースの捕捉プローブの使用¹⁷

b. 十分にカバーされていない領域についての追加的なベイトの包含

3. 可能性のある新生抗原が体細胞性/生殖細胞系列ステータスについて分類されていないままである（したがってTSNAとして使用可能ではない）ことが最も少ないように、20x未満でカバーされる塩基が5%未満である、正常エクソームにわたる均一カバレッジのターゲティング。

50

4. 必要とされるシーケンシングの総量を最小化するために、配列捕捉プローブは、非コードRNAは新生抗原を生じることができないことから、遺伝子のコード領域のみについて設計される。追加的な最適化は、以下を含む：

a. GCリッチであり、標準的なエクソームシーケンシングでは十分に捕捉されないHLA遺伝子についての補充的プローブ¹⁸。

b. 不十分な発現、プロテアソームによる最適に満たない消化、または異例の配列特性などの要因により、候補新生抗原を少ししかまたは全く生成しないと予測される遺伝子の排除。

5. 変異検出、遺伝子及びスプライス変異体（「アイソフォーム」）発現の定量、ならびに融合物検出を可能にするために、腫瘍RNAが同様に、高深度（100Mリードよりも大きい）でシーケンシングされる。FFPE試料由来のRNAは、DNAにおいてエクソームを捕捉するために使用されるのと同じまたは類似したプローブで、プローブベース濃縮¹⁹を用いて抽出される。

【0166】

VI. A. 2. NGSデータ解析の最適化

解析法の改善は、一般的な研究変異コーリングアプローチの最適に満たない感度及び特異性に対処し、具体的には、臨床設定における新生抗原の特定のために関連するカスタマイズ化を考慮する。これらは、以下を含む：

1. アラインメントのための、HG38参照ヒトゲノムまたはより後のバージョンの使用（それが、以前のゲノムリリースとは対照的に、集団多型をより良好に反映する複数のMHC領域アセンブリを含有するため）。

2. 様々なプログラム⁵からの結果をマージすることによる、単一変異コーラー20の限界の克服。

a. 単一ヌクレオチド変異及び挿入欠失は、以下を含む一連のツールで、腫瘍DNA、腫瘍RNA、及び正常DNAから検出される：Strelka²¹及びMutect²²などの、腫瘍及び正常DNAの比較に基づくプログラム；ならびに、低純度の試料において特に有利である²³、UNCeqRなどの、腫瘍DNA、腫瘍RNA、及び正常DNAを組み入れるプログラム。

b. 挿入欠失は、Strelka及びABRA²⁴などの、局所リアセンブリを行うプログラムで決定される。

c. 構造的再編成は、Pindel²⁵またはBreakseq²⁶などの専用のツールを用いて決定される。

3. 試料スワップを検出して阻止するために、同じ患者についての試料由来の変異コーラーが、選ばれた数の多型部位で比較される。

4. 例として、以下による、人工的コールの広範囲のフィルタリングが行われる：

a. 潜在的に、低いカバレッジの例においては緩やかな検出パラメータで、及び挿入欠失の例においては許容的な近接基準での、正常DNAにおいて見出される変異の除去。

b. 低いマッピング品質または低い塩基品質による変異の除去²⁷。

c. たとえ対応する正常において観察されないとしても、再出現するシーケンシングアーチファクトから生じる変異の除去²⁷。例は、主として1本の鎖上に検出される変異を含む。

d. 無関連の対照のセットにおいて検出される変異の除去²⁷。

5. seq2HLA²⁸、ATHLATES²⁹、またはOptitypeのうちの1つを使用する、かつまた、エクソーム及びRNAシーケンシングデータを組み合わせる²⁸、正常エクソームからの正確なHLAコーリング。追加的な潜在的最適化は、ロングリードDNAシーケンシングなどの、HLAタイピングのための専用アッセイの採用³⁰、または、RNA断片を連結して連続性を保持するための方法の適応³¹を含む。

6. 腫瘍特異的スプライス変異体から生じた新生ORFの堅牢な検出は、CLASS³²、Bayessembler³³、StringTie³⁴、またはそのリファレンスガイドモードにおける類似したプログラム（すなわち、各実験からそれらの全体の転写産物

10

20

30

40

50

を再作製するように試みるよりもむしろ、公知の転写産物構造を用いる)を用いて、RNA-seqデータから転写産物をアセンブルすることによって、行われる。Cufflinks³⁵が、この目的で一般的に使用されるが、それは頻繁に、信じ難いほど多数のスプライス変異体を産生し、それらの多くは、完全長遺伝子よりもはるかに短く、単純な陽性対照をリカバーすることができない場合がある。コード配列及び潜在的なナンセンス変異依存分解機構は、変異体配列を再導入した、SpliCeR³⁶及びMAMBA³⁷などのツールで決定される。遺伝子発現は、Cufflinks³⁵またはExpress (Roberts and Pachter, 2013)などのツールで決定される。野生型及び変異体特異的な発現カウント及び/または相対レベルは、ASE³⁸またはHTSeq³⁹などの、これらの目的で開発されたツールで決定される。潜在的なフィルタリング段階は、以下を含む：

10

a. 不十分に発現されていると考えられる候補新生ORFの除去。

b. ナンセンス変異依存分解機構(NMD)を引き起こすと予測される候補新生ORFの除去。

7. 腫瘍特異的と直接検証することができない、RNAにおいてのみ観察される候補新生抗原(例えば、新生ORF)は、例として以下を考慮することにより、追加的なパラメータにしたがって、腫瘍特異的である可能性が高いとして分類される：

a. 腫瘍DNAのみのシス作用性フレームシフトまたはスプライス部位変異の支持の存在。

b. スプライシング因子における腫瘍DNAのみのトランス作用性変異の確証の存在。例として、R625変異体SF3B1での3つの独立して公開された実験において、最も差次的にスプライシングを呈する遺伝子は、1つの実験がブドウ膜黒色腫患者を検討し⁴⁰、第2の実験がブドウ膜黒色腫細胞株を検討し⁴¹、及び第3の実験が乳がん患者を検討した⁴²にもかかわらず、一致していた。

20

c. 新規のスプライシングアイソフォームについては、RNA-seqデータにおける「新規の」スプライス-ジャンクションリードの確証の存在。

d. 新規の再編成については、正常DNAには存在しない腫瘍DNAにおけるエクソン近傍リードの確証の存在。

e. GTEX⁴³などの遺伝子発現大要からの欠如(すなわち、生殖細胞系列起源の可能性をより低くする)。

30

8. アラインメント及びアノテーションベースのエラー及びアーチファクトを直接避けるために、アセンブルされたDNAの腫瘍及び正常リード(またはそのようなリード由来のkマー)を比較することによる、参照ゲノムアラインメントベースの解析の補完(例えば、生殖細胞系列変異またはリピートコンテクスト挿入欠失の近くに生じる体細胞性変異について)。

【0167】

ポリアデニル化RNAを有する試料において、RNA-seqデータにおけるウイルスRNA及び微生物RNAの存在は、患者の応答を予測し得る追加的因子の特定に向かって、RNA-COMPASS⁴⁴または類似した方法を用いて評価される。

【0168】

VI. B. HLAペプチドの単離及び検出

HLAペプチド分子の単離は、組織試料の溶解及び可溶化後に、古典的な免疫沈降(IP)法を用いて行った⁵⁵⁻⁵⁸。清澄化した溶解物を、HLA特異的IPに使用した。

【0169】

免疫沈降は、抗体がHLA分子に特異的である、ビーズにカップリングした抗体を用いて行った。汎クラスI HLA免疫沈降のためには、汎クラスI CR抗体を使用し、クラスII HLA-DRのためには、HLA-DR抗体を使用する。抗体を、一晚インキュベーション中に、NHS-セファロースビーズに共有結合で付着させる。共有結合性の付着後、ビーズを洗浄して、IPのために等分した^{59、60}。ビーズに共有結合されていない抗体を用いて免疫沈降を行うこともできる。一般的に、これは、抗体をカラムに保

40

50

持するために Protein A 及び / または Protein G でコーティングしたセファロースまたは磁気ビーズを使用して行われる。MHC / ペプチド複合体を選択的に濃縮するために使用することができるいくつかの抗体を下記に示す。

抗体名	特異性
W6/32	クラス I HLA-A, B, C
L243	クラス II - HLA-DR
Tu36	クラス II - HLA-DR
LN3	クラス II - HLA-DR
Tu39	クラス II - HLA-DR, DP, DQ

10

【 0 1 7 0 】

清澄化した組織溶解物を、免疫沈降のために抗体ビーズに添加する。免疫沈降後、ビーズを溶解物から除去し、追加的な IP を含む追加的な実験のために、溶解物を保存する。標準的な技法を用いて、IP ビーズを洗浄して非特異的結合を除去し、HLA / ペプチド複合体をビーズから溶出する。分子量スピンカラムまたは C 1 8 分画を用いて、タンパク質構成要素をペプチドから除去する。結果として生じたペプチドを、Speed Vac 蒸発によって乾燥させ、いくつかの場合には、MS 解析の前に - 2 0 で保存する。

【 0 1 7 1 】

乾燥したペプチドを、逆相クロマトグラフィーに適している HPLC 緩衝液において再構成し、Fusion Lumos 質量分析計 (Thermo) における勾配溶出のために、C - 1 8 マイクロキャピラリー HPLC カラム上にロードする。ペプチド質量 / 電荷 (m / z) の MS 1 スペクトルを、Orbitrap 検出器において高解像度で収集し、その後、MS 2 低解像度スキャンを、選択イオンの HCD フラグメンテーション後にイオントラップ検出器において収集した。追加的に、MS 2 スペクトルは、CID もしくは ETD フラグメンテーション法、または、ペプチドのより大きなアミノ酸カバレッジを獲得するための 3 つの技法の任意の組み合わせのいずれかを用いて、取得することができる。MS 2 スペクトルはまた、Orbitrap 検出器において高解像度質量精度で測定することもできる。

20

30

【 0 1 7 2 】

各解析由来の MS 2 スペクトルを、Comet^{6 1}、^{6 2} を用いてタンパク質データベースに対して検索し、ペプチド特定を、Percolator^{6 3 ~ 6 5} を用いてスコア化する。PEAKS studio (Bioinformatics Solutions Inc.) 及び他のサーチエンジンを用いてさらなるシーケンシングを行うか、またはスペクトルマッチング及びデノボシーケンシング^{7 5} を含むシーケンシング法を用いることができる。

【 0 1 7 3 】

VI . B . 1 . 総合的 HLA ペプチドシーケンシングのための MS 検出限界の研究

ペプチド YVYVADVA AK を用いて、何が検出の限界かを、LC カラム上にロードした様々な量のペプチドを用いて決定した。試験したペプチドの量は、1 pmol、1 0 0 fmol、1 0 fmol、1 fmol、及び 1 0 0 amol であった。(表 1) 結果を図 1 F に示す。これらの結果は、検出の最低限界 (LOD) がアトモルの範囲 ($1 0^{-1 8}$) にあること、ダイナミックレンジが 5 桁に及ぶこと、及び、シグナル対ノイズが、低いフェムトモル範囲 ($1 0^{-1 5}$) でシーケンシングに十分であるように見えることを示す。

40

【 0 1 7 4 】

ペプチド m/z	カラムへのロード	1e9個の細胞における コピー/細胞
566.830	1 pmol	600
562.823	100 fmol	60
559.816	10 fmol	6
556.810	1 fmol	0.6
553.802	100 amol	0.06

10

【 0 1 7 5 】

V I I . 提示モデル

V I I . A . システムの概要

図 2 A は、1つの実施形態にしたがう、患者におけるペプチド提示の尤度を特定するための環境 1 0 0 の概要である。環境 1 0 0 は、それ自体が提示情報記憶装置 1 6 5 を含む提示特定システム 1 6 0 を導入するコンテキストを提供する。

【 0 1 7 6 】

提示特定システム 1 6 0 は、図 3 0 に関して下記で議論されるようなコンピュータ計算システムにおいて具現化された、1つまたはコンピュータモデルであり、M H C アレルのセットに関連するペプチド配列を受け取り、ペプチド配列が、関連する M H C アレルのセットの1つ以上によって提示されるであろう尤度を決定する。提示特定システム 1 6 0 はクラス I 及びクラス I I M H C アレルの両方に適用することができる。これは、様々なコンテキストにおいて有用である。提示特定システム 1 6 0 の1つの具体的な用途の例は、患者 1 1 0 の腫瘍細胞由来の M H C アレルのセットに関連する候補新生抗原のヌクレオチド配列を受け取り、候補新生抗原が、腫瘍の関連する M H C アレルの1つ以上によって提示され、及び/または患者 1 1 0 の免疫系において免疫原性応答を誘導するであろう尤度を決定することができることである。システム 1 6 0 によって決定された際に高い尤度を有するそれらの候補新生抗原を、ワクチン 1 1 8 における包含のために選択することができ、そのような抗腫瘍免疫応答が、腫瘍細胞を提供する患者 1 1 0 の免疫系から誘発され得る。さらに、高い提示尤度を有する候補新生抗原に対する反応性を有する T C R を有する T 細胞を T 細胞療法で使用するために作製することが可能であり、これにより、患者 1 1 0 の免疫系から抗腫瘍免疫応答も誘発される。

20

30

【 0 1 7 7 】

提示特定システム 1 6 0 は、1つ以上の提示モデルを通して提示尤度を決定する。具体的には、提示モデルは、所定のペプチド配列が、関連する M H C アレルのセットについて提示されるかどうかの尤度を生成し、尤度は、記憶装置 1 6 5 に保存された提示情報に基づいて生成される。例えば、提示モデルは、ペプチド配列「Y V Y V A D V A A K」が、試料の細胞表面上のアレルのセット H L A - A * 0 2 : 0 1、H L A - A * 0 3 : 0 1、H L A - B * 0 7 : 0 2、H L A - B * 0 8 : 0 3、H L A - C * 0 1 : 0 4 について提示されるかどうかの尤度を生成し得る。提示情報 1 6 5 は、M H C アレルによってペプチドが提示されるようにこれらのペプチドが様々なタイプの M H C アレルに結合するかどうかについての情報を含み、これは、モデルにおいて、ペプチド配列中のアミノ酸の位置に応じて決定される。提示モデルは、提示情報 1 6 5 に基づいて、認識されていないペプチド配列が、M H C アレルの関連するセットと結合して提示されるかどうかを予測することができる。上記に述べたように、提示モデルはクラス I 及びクラス I I M H C アレルの両方に適用することができる。

40

【 0 1 7 8 】

V I I . B . 提示情報

図 2 は、1つの実施形態にしたがう、提示情報を取得する方法を説明する。提示情報 1

50

65は、2つの一般的部類の情報：アレル相互作用情報及びアレル非相互作用情報を含む。アレル相互作用情報は、MHCアレルのタイプに依存する、ペプチド配列の提示に影響を及ぼす情報を含む。アレル非相互作用情報は、MHCアレルのタイプに非依存的な、ペプチド配列の提示に影響を及ぼす情報を含む。

【0179】

VII.B.1. アレル相互作用情報

アレル相互作用情報は、主として、ヒト、マウスなど由来の1つ以上の特定されたMHC分子によって提示されていることが公知である、特定されたペプチド配列を含む。注目すべきことに、これは、腫瘍試料から取得されたデータを含んでもよく、または含まなくてもよい。提示されたペプチド配列は、単一のMHCアレルを発現する細胞から特定されてもよい。この例において、提示されたペプチド配列は、概して、あらかじめ決定されたMHCアレルを発現するように操作されてその後合成タンパク質に曝露された単一アレル細胞株から収集される。MHCアレル上に提示されたペプチドは、酸溶出などの技法によって単離され、質量分析により特定される。図2Bは、あらかじめ決定されたMHCアレルHLA-DRB1*12:01上に提示された例示的なペプチド

10

YEMFNDKSQRAPDDKMF

が単離され、質量分析により特定される、この例を示す。この状況においては、ペプチドが、単一のあらかじめ決定されたMHCタンパク質を発現するように操作された細胞を通して特定されるため、提示されたペプチドとそれが結合したMHCタンパク質との間の直接の関連が、決定的に既知である。

20

【0180】

提示されたペプチド配列はまた、複数のMHCアレルを発現する細胞から収集されてもよい。典型的にヒトにおいては、6種類の異なるタイプのMHC I分子及び最大で12種類の異なるタイプのMHC II分子が細胞で発現している。そのような提示されたペプチド配列は、複数のあらかじめ決定されたMHCアレルを発現するように操作されている複数アレル細胞株から特定されてもよい。そのような提示されたペプチド配列はまた、正常組織試料または腫瘍組織試料のいずれかの、組織試料から特定されてもよい。この例において特に、MHC分子は、正常組織または腫瘍組織から免疫沈降させることができる。複数のMHCアレル上に提示されたペプチドは、同様に、酸溶出などの技法によって単離され、質量分析により特定されることができる。図2Cは、6種類の例示的なペプチド

30

YEMFNDKSF, HROEIFSHDFJ, FJIEJFOESS, NEIOREIREI,
JFKSIFEMMSJDSSUIFLKSJFIEIFJ, 及び KNFLENFIESOFI

が、特定されたクラスI MHCアレルHLA-A*01:01、HLA-A*02:01、HLA-B*07:02、HLA-B*08:01、及びクラスII MHCアレルHLA-DRB1*10:01、HLA-DRB1:11:01上に提示されており、単離され、質量分析により特定される、この例を示す。単一アレル細胞株とは対照的に、結合したペプチドが、特定される前のMHC分子から単離されるため、提示されたペプチドとそれが結合したMHCタンパク質との間の直接の関連は、未知である可能性がある。

【0181】

40

アレル相互作用情報はまた、ペプチド-MHC分子複合体の濃度、及びペプチドのイオン化効率の両方に依存する、質量分析イオン電流も含むことができる。イオン化効率は、配列依存性様式で、ペプチドごとに変動する。概して、イオン効率は、およそ2桁にわたってペプチドごとに変動し、他方、ペプチド-MHC複合体の濃度は、それよりも大きい範囲にわたって変動する。

【0182】

アレル相互作用情報はまた、所定のMHCアレルと所定のペプチドとの間の結合親和性の測定値または予測値も含むことができる。1つ以上の親和性モデルが、そのような予測値を生成することができる(72, 73, 74)。例えば、図1Dに示した例に戻ると、提示情報165は、ペプチドYEMFNDKSFとクラスIアレルHLA-A*01:0

50

1 との間の 1000 nM の結合親和性予測値を含み得る。IC50 > 1000 nm であるペプチドはわずかしが、MHC によって提示されず、より低い IC50 値が、提示の確率を増大させる。提示情報 165 は、ペプチド KNFLNFIESOFI とクラス II アレル HLA - DRB1 : 11 : 01 との間の結合親和性予測値を含み得る。

【0183】

アレル相互作用情報はまた、MHC 複合体の安定性の測定値または予測値も含むことができる。1 つ以上の安定性モデルが、そのような予測値を生成することができる。より安定なペプチド - MHC 複合体（すなわち、より長い半減期を有する複合体）は、腫瘍細胞上、及びワクチン抗原に遭遇する抗原提示細胞上に高コピー数で提示される可能性がより高い。例えば、図 2C に示した例に戻ると、提示情報 165 は、クラス I 分子 HLA - A * 01 : 01 について 1 時間の半減期の安定性予測値を含み得る。提示情報 165 はクラス II 分子 HLA - DRB1 : 11 : 01 の半減期の安定性予測値も含み得る。

10

【0184】

アレル相互作用情報はまた、ペプチド - MHC 複合体の形成反応の、測定されたかまたは予測された速度も含むことができる。より速い速度で形成する複合体は、高濃度で細胞表面上に提示される可能性がより高い。

【0185】

アレル相互作用情報はまた、ペプチドの配列及び長さも含むことができる。MHC クラス I 分子は典型的に、8 ~ 15 ペプチドの長さを有するペプチドを提示することを好む。提示されたペプチドの 60 ~ 80 % は、長さ 9 を有する。MHC クラス II 分子は一般的にペプチド 6 ~ 30 個の長さを有するペプチドを提示する傾向にある。

20

【0186】

アレル相互作用情報はまた、新生抗原コード化ペプチド上のキナーゼ配列モチーフの存在、及び新生抗原コード化ペプチド上の特異的な翻訳後修飾の有無も含むことができる。キナーゼモチーフの存在は、MHC 結合を増強または干渉し得る、翻訳後修飾の確率に影響を及ぼす。

【0187】

アレル相互作用情報はまた、(RNA seq、質量分析、または他の方法によって測定されたかまたは予測された際の) 翻訳後修飾のプロセスに関するタンパク質、例えば、キナーゼの発現または活性レベルも含むことができる。

30

【0188】

アレル相互作用情報はまた、質量分析プロテオミクスまたは他の手段によって評価された際の、特定の MHC アレルを発現する他の個体由来の細胞における、類似した配列を有するペプチドの提示の確率も含むことができる。

【0189】

アレル相互作用情報はまた、(例えば、RNA - seq または質量分析によって測定された際の) 問題の個体における特定の MHC アレルの発現レベルも含むことができる。高レベルで発現している MHC アレルに最も強く結合するペプチドは、低レベルで発現している MHC アレルに最も強く結合するペプチドよりも、提示される可能性がより高い。

【0190】

アレル相互作用情報はまた、特定の MHC アレルを発現する他の個体における、特定の MHC アレルによる提示の、全体的な新生抗原コード化ペプチド配列非依存的確率も含むことができる。

40

【0191】

アレル相互作用情報はまた、他の個体における同じファミリーの分子（例えば、HLA - A、HLA - B、HLA - C、HLA - DQ、HLA - DR、HLA - DP）の MHC アレルによる提示の、全体的なペプチド配列に非依存的な確率も含むことができる。例えば、HLA - C 分子は典型的に、HLA - A または HLA - B 分子よりも低いレベルで発現しており、したがって、HLA - C によるペプチドの提示は、HLA - A または HLA - B II による提示よりも先験的に確率が低い。別の例として、HLA - DP は一般的

50

にHLA-DRまたはHLA-DQよりも低いレベルで発現されることから、HLA-DPによるペプチドの提示はHLA-DRまたはHLA-DQによる提示よりもより確率が低いものと推測される。

【0192】

アレル相互作用情報はまた、特定のMHCアレルのタンパク質配列も含むことができる。

【0193】

下記のセクションに列挙される任意のMHCアレル非相互作用情報もまた、MHCアレル相互作用情報としてモデル化することができる。

【0194】

VII.B.2. アレル非相互作用情報

アレル非相互作用情報は、そのソースタンパク質配列内の、新生抗原コード化ペプチドに隣接するC末端配列を含むことができる。MHC-Iでは、C末端フランキンゲン配列は、ペプチドのプロテアソームプロセッシングに影響を及ぼし得る。しかし、C末端フランキンゲン配列は、ペプチドが小胞体に輸送され、細胞の表面上のMHCアレルと遭遇する前に、プロテアソームによってペプチドから切断される。その結果、MHC分子は、C末端フランキンゲン配列についてのいかなる情報も受け取らず、したがって、C末端フランキンゲン配列の効果は、MHCアレルタイプに応じて変動することができない。例えば、図2Cに示した例に戻ると、提示情報165は、ペプチドのソースタンパク質から特定された、提示されたペプチドFJIEJFOESSのC末端フランキンゲン配列FOEIFNDKSLDKFJIを含み得る。

【0195】

アレル非相互作用情報はまた、mRNA定量測定値も含むことができる。例えば、mRNA定量データは、質量分析訓練データを提供する同じ試料について取得することができる。図13Hに関して後に記載するように、RNA発現は、ペプチド提示の強い予測因子であると特定された。一実施形態では、mRNA定量測定値は、ソフトウェアツールRSEMから特定される。RSEMソフトウェアツールの詳細な実行は、Bo Li and Colin N. Dewey. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 12:323, August 2011で見出すことができる。一実施形態では、mRNA定量は、100万個のマッピングされたリードあたりの転写産物のキログラムあたりの断片の単位(FPKM)で測定される。

【0196】

アレル非相互作用情報はまた、そのソースタンパク質配列内の、ペプチドに隣接するN末端配列も含むことができる。

【0197】

アレル非相互作用情報はペプチド配列のソース遺伝子も含むことができる。ソース遺伝子はペプチド配列のEnsemblタンパク質ファミリーとして定義することができる。他の例では、ソース遺伝子はペプチド配列のソースDNAまたはソースRNAとして定義することができる。ソース遺伝子は、例えば、タンパク質をコードするヌクレオチドのストリングとして表すか、またはその代わりに、特定のタンパク質をコードしていることが知られている既知のDNAまたはRNA配列の命名されたセットに基づいてよりカテゴリー化された形で表すことができる。別の例では、アレル非相互作用情報は、EnsemblまたはRefSeqのようなデータベースから抽出されたペプチド配列のソース転写産物もしくはアイソフォームまたは潜在的なソース転写産物もしくはアイソフォームのセットも含むことができる。

【0198】

アレル非相互作用情報はまた、ペプチド配列が由来する細胞の組織タイプ、細胞タイプ、または腫瘍タイプも含むことができる。

10

20

30

40

50

【0199】

アレル非相互作用情報はまた、(RNA-seqまたは質量分析によって測定された際の)任意で、腫瘍細胞における対応するプロテアーゼの発現にしがって重み付けされる、ペプチドにおけるプロテアーゼ切断モチーフの存在も含むことができる。プロテアーゼ切断モチーフを含有するペプチドは、プロテアーゼによってより容易に分解され、したがって細胞内で安定性がより低いことになるため、提示される可能性がより低い。

【0200】

アレル非相互作用情報はまた、適切な細胞タイプにおいて測定された際の、ソースタンパク質の代謝回転速度も含むことができる。より速い代謝回転速度(すなわち、より低い半減期)は提示の確率を増大させるが、類似していない細胞タイプにおいて測定された場合、この特性の予測力は低い。

10

【0201】

アレル非相互作用情報はまた、RNA-seqもしくはプロテオーム質量分析によって測定された際、または、DNAもしくはRNA配列データにおいて検出される生殖細胞系列もしくは体細胞性スプライシング変異のアノテーションから予測された際の、任意で、腫瘍細胞において最も高発現している特異的なスプライス変異体(「アイソフォーム」)を考慮する、ソースタンパク質の長さも含むことができる。

【0202】

アレル非相互作用情報はまた、(RNA-seq、プロテオーム質量分析、または免疫組織化学によって測定され得る)腫瘍細胞におけるプロテアソーム、イムノプロテアソーム、胸腺プロテアソーム、または他のプロテアーゼの発現のレベルも含むことができる。異なるプロテアソームは、異なる切断部位の好みを有する。より大きい重みが、その発現レベルに比例して、プロテアソームの各タイプの切断の好みに与えられる。

20

【0203】

アレル非相互作用情報はまた、(例えば、RNA-seqまたは質量分析によって測定された際の)ペプチドのソース遺伝子の発現も含むことができる。可能な最適化は、腫瘍試料内の間質細胞及び腫瘍浸潤リンパ球の存在を説明する、測定された発現を調整することを含む。より高発現している遺伝子由来のペプチドは、提示される可能性がより高い。検出不可能なレベルの発現を有する遺伝子由来のペプチドは、考察から排除することができる。

30

【0204】

アレル非相互作用情報はまた、新生抗原コード化ペプチドのソースmRNAが、ナンセンス変異依存分解機構のモデル、例えば、Rivas et al, Science 2015からのモデルによって予測されるようなナンセンス変異依存分解機構に供されるであろう確率も含むことができる。

【0205】

アレル非相互作用情報はまた、細胞周期の種々の段階の最中の、ペプチドのソース遺伝子の典型的な組織特異的発現も含むことができる。(RNA-seqまたは試料分析プロテオミクスによって測定された際に)全体的に低いレベルで発現しているが、細胞周期の特異的な段階の最中に高レベルで発現していることが公知である遺伝子は、非常に低いレベルで安定に発現している遺伝子よりも、より提示されるペプチドを産生する可能性が高い。

40

【0206】

アレル非相互作用情報はまた、例えば、UniProtまたはPDB <http://www.rcsb.org/pdb/home/home.do>において与えられるような、ソースタンパク質の特性の総合的なカタログも含むことができる。これらの特性は、とりわけ、タンパク質の二次構造及び三次構造、細胞内局在化11、遺伝子オントロジー(GO)用語を含み得る。具体的には、この情報は、タンパク質のレベルで作用するアノテーション、例えば、5'UTR長、及び特異的残基のレベルで作用するアノテーション、例えば、残基300~310のヘリックスモチーフを含有し得る。これらの特性はまた、ターンモチーフ、シートモチーフ、及

50

び無秩序残基も含むことができる。

【0207】

アレル非相互作用情報はまた、ペプチドを含有するソースタンパク質のドメインの性状を説明する特性、例えば、二次構造または三次構造（例えば、ヘリックス対シート）；選択的スプライシングも含むことができる。

【0208】

アレル非相互作用情報はまた、ペプチドのソースタンパク質におけるペプチドの位置での提示ホットスポットの有無を説明する特性も含むことができる。

【0209】

アレル非相互作用情報はまた、他の個体における問題のペプチドのソースタンパク質由来のペプチドの提示の確率（それらの個体におけるソースタンパク質の発現レベル、及びそれらの個体の様々なHLAタイプの影響を調整した後）も含むことができる。

10

【0210】

アレル非相互作用情報はまた、ペプチドが、技術的バイアスのために質量分析によって検出されないか、または過剰に表されるであろう確率も含むことができる。

【0211】

腫瘍細胞、間質、または腫瘍浸潤リンパ球（TIL）の状態について情報を与える、RNAseq、マイクロアレイ、Nanosttringなどの標的化パネルなどの、遺伝子発現アッセイ、または、RT-PCRなどのアッセイによって測定される遺伝子モジュールを代表する単一/複数遺伝子によって測定された際の、種々の遺伝子モジュール/経路の発現（ペプチドのソースタンパク質を含有する必要はない）。

20

【0212】

アレル非相互作用情報はまた、腫瘍細胞におけるペプチドのソース遺伝子のコピー数も含むことができる。例えば、腫瘍細胞においてホモ接合性欠失に供される遺伝子由来のペプチドは、ゼロの提示確率を割り当てることができる。

【0213】

アレル非相互作用情報はまた、ペプチドがTAPに結合する確率、または、測定されたかもしくは予測された、TAPに対するペプチドの結合親和性も含むことができる。TAPに結合する可能性がより高いペプチド、またはより高い親和性でTAPに結合するペプチドは、MHC-Iによって提示される可能性がより高い。

30

【0214】

アレル非相互作用情報はまた、（RNA-seq、プロテオーム質量分析、免疫組織化学によって測定され得る）腫瘍細胞におけるTAPの発現レベルも含むことができる。MHC-Iでは、より高いTAP発現レベルは、すべてのペプチドの提示の確率を増大させる。

【0215】

アレル非相互作用情報はまた、以下を含むがそれらに限定されない、腫瘍変異の有無も含むことができる：

i . EGFR、KRAS、ALK、RET、ROS1、TP53、CDKN2A、CDKN2B、NTRK1、NTRK2、NTRK3などの公知のがんドライバー遺伝子におけるドライバー変異。

40

ii . 抗原提示マシナリーに關与するタンパク質をコードする遺伝子（例えば、B2M、HLA-A、HLA-B、HLA-C、TAP-1、TAP-2、TAPBP、CALR、CNX、ERP57、HLA-DM、HLA-DMA、HLA-DMB、HLA-DO、HLA-DOA、HLA-DOBHLA-DP、HLA-DPA1、HLA-DPB1、HLA-DQ、HLA-DQA1、HLA-DQA2、HLA-DQB1、HLA-DQB2、HLA-DR、HLA-DRA、HLA-DRB1、HLA-DRB3、HLA-DRB4、HLA-DRB5、または、プロテアソームもしくはイムノプロテアソームの構成要素をコードする遺伝子のいずれか）におけるもの。その提示が、腫瘍において機能喪失変異の影響下にある抗原提示マシナリーの構成要素に依拠するペプチドは、提示

50

の確率が低減している。

【0216】

以下を含むがそれらに限定されない、機能的生殖細胞系列多型の有無：

i. 抗原提示マシナリーに關与するタンパク質をコードする遺伝子（例えば、B2M、HLA-A、HLA-B、HLA-C、TAP-1、TAP-2、TAPBP、CALR、CNX、ERP57、HLA-DM、HLA-DMA、HLA-DMB、HLA-DO、HLA-DOA、HLA-DOBHLA-DP、HLA-DPA1、HLA-DPB1、HLA-DQ、HLA-DQA1、HLA-DQA2、HLA-DQB1、HLA-DQB2、HLA-DR、HLA-DRA、HLA-DRB1、HLA-DRB3、HLA-DRB4、HLA-DRB5、または、プロテアソームもしくはイムノプロテアソームの構成要素をコードする遺伝子のいずれか）におけるもの。

10

【0217】

アレル非相互作用情報はまた、腫瘍タイプ（例えば、NSCLC、黒色腫）も含むことができる。

【0218】

アレル非相互作用情報はまた、例としてHLAアレル接尾辞によって反映されるような、HLAアレルの公知の機能性も含むことができる。例えば、アレル名HLA-A*24:09NにおけるNの接尾辞は、発現せず、したがってエピトープを提示する可能性が低いヌルアレルを示し；完全なHLAアレル接尾辞の命名法は、<https://www.ebi.ac.uk/ipd/imgt/hla/nomenclature/suffixes.html>に記載されている。

20

【0219】

アレル非相互作用情報はまた、臨床的腫瘍サブタイプ（例えば、扁平上皮肺癌対非扁平上皮）も含むことができる。

【0220】

アレル非相互作用情報はまた、喫煙歴も含むことができる。

【0221】

アレル非相互作用情報はまた、日焼け、日光曝露、または他の変異原に対する曝露の経歴も含むことができる。

【0222】

アレル非相互作用情報はまた、任意でドライバー変異によって層別化される、関連性のある腫瘍タイプまたは臨床的サブタイプにおけるペプチドのソース遺伝子の局所的発現も含むことができる。関連性のある腫瘍タイプにおいて典型的に高レベルで発現している遺伝子は、提示される可能性がより高い。

30

【0223】

アレル非相互作用情報はまた、すべての腫瘍における、または同じタイプの腫瘍における、または少なくとも1つの共有されたMHCアレルを有する個体由来の腫瘍における、または少なくとも1つの共有されたMHCアレルを有する個体中の同じタイプの腫瘍における、変異の頻度も含むことができる。

【0224】

変異した腫瘍特異的ペプチドの例において、提示の確率を予測するために使用される特性の一覧はまた、変異のアノテーション（例えば、ミスセンス、リードスルー、フレームシフト、融合など）、または、変異がナンセンス変異依存分解機構（NMD）を結果としてもたらすと予測されるかどうかも含み得る。例えば、ホモ接合性早期終止変異のために腫瘍細胞において翻訳されないタンパク質セグメント由来のペプチドは、ゼロの提示確率を割り当てることができる。NMDは、提示の確率を減少させる、mRNA翻訳の減少を結果としてもたらす。

40

【0225】

VII.C. 提示特定システム

図3は、1つの実施形態による、提示特定システム160のコンピュータ論理構成要素を説明する、ハイレベルブロック図である。この例示的实施形態において、提示特定シス

50

テム 160 は、データ管理モジュール 312、コード化モジュール 314、訓練モジュール 316、及び予測モジュール 320 を含む。提示特定システム 160 はまた、訓練データ記憶装置 170 及び提示モデル記憶装置 175 から構成される。モデル管理システム 160 のいくつかの実施形態は、本明細書に記載したものと異なるモジュールを有する。同様に、機能は、本明細書に記載したものは異なる様式で、モジュールの間に分配され得る。

【0226】

VII.C.1. データ管理モジュール

データ管理モジュール 312 は、提示情報 165 から訓練データ 170 のセットを生成する。各々の訓練データのセットは、多数のデータ例を含有し、各データ例 i は、少なくとも、提示されるかまたは提示されないペプチド配列 p^i と、ペプチド配列 p^i と結合した 1 つ以上の関連する MHC アレル a^i と、提示特定システム 160 が、独立変数の新たな値を予測することに関心があるという情報を表す従属変数 y^i とを含む、独立変数 z^i のセットを含有する。

10

【0227】

本明細書の残りの部分を通じて言及される 1 つの特定の実現形態において、従属変数 y^i は、ペプチド p^i が、1 つ以上の関連する MHC アレル a^i によって提示されたかどうかを示す、バイナリーラベルである。しかし、他の実現形態において、従属変数 y^i は、提示特定システム 160 が、独立変数 z^i に依存して予測することに関心があるという任意の他の種類の情報を表し得ることが、認識される。例えば、別の実現形態において、従属変数 y^i はまた、データ例について特定された質量分析イオン電流を示す数値であってもよい。

20

【0228】

データ例 i についてのペプチド配列 p^i は、 k_i 個のアミノ酸の配列であり、 k_i は、データ例 i の間で、ある範囲内で変動し得る。例えば、その範囲は、MHC クラス I については 8 ~ 15、または MHC クラス II については 6 ~ 30 であり得る。システム 160 の 1 つの具体的な実現形態において、訓練データセット中のすべてのペプチド配列 p^i は、同じ長さ、例えば 9 を有し得る。ペプチド配列中のアミノ酸の数は、MHC アレルのタイプ（例えば、ヒトにおける MHC アレルなど）に応じて変動し得る。データ例 i についての MHC アレル a^i は、どの MHC アレルが対応するペプチド配列 p^i と結合して存在したかを示す。

30

【0229】

データ管理モジュール 312 はまた、訓練データ 170 に含有されるペプチド配列 p^i 及び結合した MHC アレル a^i と共に、結合親和性 b^i 及び安定性 s^i の予測値などの追加的なアレル相互作用変数も含み得る。例えば、訓練データ 170 は、ペプチド p^i と、 a^i において示される結合した MHC 分子の各々との間の結合親和性予測値 b^i を含有し得る。別の例として、訓練データ 170 は、 a^i において示される MHC アレルの各々についての安定性予測値 s^i を含有し得る。

【0230】

データ管理モジュール 312 はまた、ペプチド配列 p^i と共に、C 末端フランキング配列及び mRNA 定量測定値などのアレル非相互作用変数 w^i も含み得る。

40

【0231】

データ管理モジュール 312 はまた、MHC アレルによって提示されないペプチド配列も特定して、訓練データ 170 を生成する。概して、これは、提示の前に、提示されるペプチド配列を含むソースタンパク質の「より長い」配列を特定することを含む。提示情報が、操作された細胞株を含有する場合、データ管理モジュール 312 は、細胞の MHC アレル上に提示されなかった、細胞がそれに対して曝露された合成タンパク質における一連のペプチド配列を特定する。提示情報が、組織試料を含有する場合、データ管理モジュール 312 は、提示されたペプチド配列の起源であるソースタンパク質を特定して、組織試料細胞の MHC アレル上に提示されなかった、ソースタンパク質における一連のペプチド

50

【 0 2 3 5 】

訓練データ170が、異なる長さのアミノ酸の配列を含有する場合、コード化モジュール314は、さらに、あらかじめ決定されたアルファベットを拡張するようにPAD文字を追加することによって、ペプチドを同等の長さのベクトルへとコード化し得る。例えば、これは、ペプチド配列の長さが、訓練データ170において最大の長さを有するペプチド配列に達するまで、ペプチド配列をPAD文字でレフトパディングすることによって行われ得る。したがって、最大の長さを有するペプチド配列が $k_{\text{最大}}$ 個のアミノ酸を有する場合、コード化モジュール314は、各配列を、 $(20+1) \cdot k_{\text{最大}}$ 個の要素の行ベクトルとして数値的に表す。例として、拡張されたアルファベット { P A D , A , C , D , E , F , G , H , I , K , L , M , N , P , Q , R , S , T , V , W , Y } 及び $k_{\text{最大}} = 5$ の最大アミノ酸長について、3個のアミノ酸の同じ例示的なペプチド配列 E A F は、10

05要素の行ベクトル

$$p^i = [10000000000$$

0000000000010000000000000000000000000000000000010000000000000000

000100]

によって表され得る。C末端フランキング配列 c^i または他の配列データは、同様に、上記のようにコード化することができる。したがって、ペプチド配列 p^i または c^i における各々の独立変数または列は、配列の特定の位置の特定のアミノ酸の存在を表す。

【 0 2 3 6 】

配列データをコード化する上記の方法は、アミノ酸配列を有する配列に関して記載したが、方法を、同様に、例えば、DNAまたはRNAの配列データなどの、他のタイプの配列データに拡張することができる。

【 0 2 3 7 】

コード化モジュール314はまた、データ例 i についての1つ以上のMHCアレル a^i を、 m 要素の行ベクトルへとコード化し、各要素 $h = 1, 2, \dots, m$ は、固有の特定されたMHCアレルに対応する。データ例 i について特定されたMHCアレルに対応する要素は、1の値を有する。その以外の、残りの要素は、0の値を有する。例として、 $m = 4$ の固有の特定されたMHCアレルタイプ { H L A - A * 0 1 : 0 1 , H L A - C * 0 1 : 0 8 , H L A - B * 0 7 : 0 2 , H L A - D R B 1 * 1 0 : 0 1 } の中の、複数アレル細胞株に対応するデータ例 i についてのアレル H L A - B * 0 7 : 0 2 及び H L A - D R B 1 * 1 0 : 0 1 は、4要素の行ベクトル $a^i = [0 \ 0 \ 1 \ 1]$ によって表され得、 $a_3^i = 1$ 及び $a_4^i = 1$ である。4種類の特定されたMHCアレルタイプでの例を、本明細書に記載するが、MHCアレルタイプの数は、実際には数百または数千であることができる。上記で述べたように、各データ例 i は、典型的に、ペプチド配列 p_i に関連して最大で6種類の異なるMHCアレルタイプを含む。

【 0 2 3 8 】

コード化モジュール314はまた、各データ例 i についてのラベル y_i を、 $\{0, 1\}$ のセットからの値を有するバイナリ変数としてコード化し、1の値は、ペプチド x^i が、関連するMHCアレル a^i のうちの1つによって提示されたことを示し、0の値は、ペプチド x^i が、関連するMHCアレル a^i のいずれによっても提示されなかったことを示す。従属変数 y_i が、質量分析イオン電流を表す場合、コード化モジュール314は、 $[0, \quad]$ の間のイオン電流値について $[- , \quad]$ の範囲を有する \log 関数などの種々の関数を用いて、値を追加的にスケール調整し得る。

【 0 2 3 9 】

コード化モジュール314は、ペプチド p_i 及び関連するMHCアレル h についてのアレル相互作用変数 x_h^i のペアを、アレル相互作用変数の数値的表示が次々に連結されている行ベクトルとして表し得る。例えば、コード化モジュール314は、 x_h^i を、 $[p^i]$ 、 $[p^i b_h^i]$ 、 $[p^i s_h^i]$ 、または $[p^i b_h^i s_h^i]$ と同等の行ベクトルとして表し得、 b_h^i は、ペプチド p_i 及び関連するMHCアレル h についての結合親

10

20

30

40

50

和性予測値であり、同様に、 s_{h^i} は、安定性についてのものである。あるいは、アレル相互作用変数の1つ以上の組み合わせは、個々に（例えば、個々のベクトルまたは行列として）保存されてもよい。

【0240】

1つの例において、コード化モジュール314は、結合親和性について測定されたかまたは予測された値をアレル相互作用変数 x_{h^i} に組み入れることによって、結合親和性情報を表す。

【0241】

1つの例において、コード化モジュール314は、結合安定性について測定されたかまたは予測された値をアレル相互作用変数 x_{h^i} に組み入れることによって、結合安定性情報を表す。

10

【0242】

1つの例において、コード化モジュール314は、結合オンレートについて測定されたかまたは予測された値をアレル相互作用変数 x_{h^i} に組み入れることによって、結合オンレート情報を表す。

【0243】

1つの例において、クラスI MHC分子によって提示されるペプチドについて、コード化モジュール314は、ペプチド長を、ベクトル

$$T_k = [1(L_k=8) 1(L_k=9) 1(L_k=10) 1(L_k=11)$$

$$1(L_k=12) 1(L_k=13) 1(L_k=14) 1(L_k=15)]$$

20

（ただし、

1

は指標関数であり、 L_k はペプチド p_k の長さを意味する）として表す。ベクトル T_k を、アレル相互作用変数 x_{h^i} に含めることができる。別の例では、クラスIIのMHC分子によって提示されるペプチドについて、コード化モジュール314はペプチド長をベクトル

$$T_k = [1(L_k=6) 1(L_k=7) 1(L_k=8) 1(L_k=9) 1(L_k=10) 1(L_k=11)$$

$$1(L_k=12) 1(L_k=13) 1(L_k=14) 1(L_k=15) 1(L_k=16) 1(L_k=17) 1(L_k=18) 1(L_k=19) 1(L_k=20) 1(L_k=21)$$

$$1(L_k=22) 1(L_k=23) 1(L_k=24) 1(L_k=25) 1(L_k=26) 1(L_k=27) 1(L_k=28) 1(L_k=29) 1(L_k=30)]$$

30

（ただし、

1

は指標関数であり、 L_k はペプチド p_k の長さを意味する）として表す。ベクトル T_k を、アレル相互作用変数 x_{h^i} に含めることができる。

【0244】

1つの例において、コード化モジュール314は、MHCアレルのRNA-seqベースの発現レベルをアレル相互作用変数 x_{h^i} に組み入れることによって、MHCアレルのRNA発現情報を表す。

【0245】

同様に、コード化モジュール314は、アレル非相互作用変数 w^i を、アレル非相互作用変数の数値的表示が次々に連鎖している行ベクトルとして表し得る。例えば、 w^i は、 $[c^i]$ または $[c^i m^i w^i]$ と同等の行ベクトルであってもよく、 w^i は、ペプチド p_i のC末端フランキング配列及びペプチドに関連するmRNA定量測定値 m^i に加えて任意の他のアレル非相互作用変数を表す、行ベクトルである。あるいは、アレル非相互作用変数の1つ以上の組み合わせは、個々に（例えば、個々のベクトルまたは行列として）保存されてもよい。

40

【0246】

1つの例において、コード化モジュール314は、代謝回転速度または半減期をアレル非相互作用変数 w^i に組み入れることによって、ペプチド配列についてのソースタンパク

50

質の代謝回転速度を表す。

【0247】

1つの例において、コード化モジュール314は、タンパク質長をアレル非相互作用変数 w^i に組み入れることによって、ソースタンパク質またはアイソフォームの長さを表す。

【0248】

1つの例において、コード化モジュール314は、 1_i 、 2_i 、 5_i サブユニットを含むイムノプロテアソーム特異的プロテアソームサブユニットの平均発現を、アレル非相互作用変数 w^i に組み入れることによって、イムノプロテアソームの活性化を表す。

【0249】

1つの例において、コード化モジュール314は、(RSEMなどの技法によってFPKM、TPMの単位で定量された)ペプチド、またはペプチドの遺伝子もしくは転写産物のソースタンパク質のRNA-seq存在量を、ソースタンパク質の存在量をアレル非相互作用変数 w^i に組み入れることによって表す。

【0250】

1つの例において、コード化モジュール314は、例えば、Rivas et al. Science, 2015におけるモデルによって推定されるような、ペプチドの起源の転写産物がナンセンス変異依存分解機構(NMD)を受けるであろう確率を、この確率をアレル非相互作用変数 w^i に組み入れることによって表す。

【0251】

1つの例において、コード化モジュール314は、RNA-seqを介して評価された遺伝子モジュールまたは経路の活性化状況を、例えば、経路における遺伝子の各々について、例えばRSEMを用いてTPMの単位で、経路における遺伝子の発現を定量すること、次いで、経路における遺伝子にわたる要約統計量、例えば平均値をコンピュータ計算することによって表す。平均を、アレル非相互作用変数 w^i に組み入れることができる。

【0252】

1つの例において、コード化モジュール314は、ソース遺伝子のコピー数を、コピー数をアレル非相互作用変数 w^i に組み入れることによって表す。

【0253】

1つの例において、コード化モジュール314は、(例えば、ナノモル単位での)測定されたかまたは予測されたTAP結合親和性をアレル非相互作用変数 w^i に含むことによって、TAP結合親和性を表す。

【0254】

1つの例において、コード化モジュール314は、RNA-seqによって測定され(かつ、例えばRSEMによってTPMの単位で定量された)TAP発現レベルをアレル非相互作用変数 w^i に含むことによって、TAP発現レベルを表す。

【0255】

1つの例において、コード化モジュール314は、腫瘍変異を、アレル非相互作用変数 w^i における指標変数のベクトル(すなわち、ペプチド p^k がKRAS G12D変異を有する試料に由来するならば $d^k = 1$ 、それ以外は0)として表す。

【0256】

1つの例において、コード化モジュール314は、抗原提示遺伝子における生殖細胞系列多型を、指標変数のベクトル(すなわち、ペプチド p^k がTAPにおいて特異的な生殖細胞系列多型を有する試料に由来するならば $d^k = 1$)として表す。これらの指標変数を、アレル非相互作用変数 w^i に含めることができる。

【0257】

1つの例において、コード化モジュール314は、腫瘍タイプを、腫瘍タイプ(例えば、NSCLC、黒色腫、大腸癌など)のアルファベットについての長さ1のワン・ホットエンコーディングされたベクトルとして表す。これらのワン・ホットエンコーディングされた変数を、アレル非相互作用変数 w^i に含めることができる。

10

20

30

40

50

【0258】

1つの例において、コード化モジュール314は、MHCアレル接尾辞を、4桁のHLAアレルを様々な接尾辞で処理することによって表す。例えば、HLA-A*24:09Nは、モデルの目的で、HLA-A*24:09とは異なるアレルと考えられる。あるいは、N接尾辞で終わるHLAアレルは発現しないため、N接尾辞のMHCアレルによる提示の確率は、すべてのペプチドについてゼロに設定することができる。

【0259】

1つの例において、コード化モジュール314は、腫瘍サブタイプを、腫瘍サブタイプ（例えば、肺腺癌、肺扁平上皮細胞癌など）のアルファベットについての長さ1のワン・ホットエンコーディングされたベクトルとして表す。これらのワン・ホットエンコーディングされた変数を、アレル非相互作用変数 w^i に含めることができる。

10

【0260】

1つの例において、コード化モジュール314は、喫煙歴を、アレル非相互作用変数 w^i に含めることができる、バイナリー指標変数（患者が喫煙歴を有するならば $d^k = 1$ 、それ以外は0）として表す。あるいは、喫煙歴を、喫煙の重症度のアルファベットについての長さ1のワン・ホットエンコーディングされた変数としてコード化することができる。例えば、喫煙状況を、1が非喫煙者を示し、5が現在の大量喫煙者を示す、1~5のスケールに査定することができる。喫煙歴は、主として肺腫瘍と関連性があるため、複数の腫瘍タイプに対するモデルを訓練する場合、この変数は、患者が喫煙の経歴を有し、かつ腫瘍タイプが肺腫瘍であるならば1と同等であり、それ以外はゼロであると定義することもできる。

20

【0261】

1つの例において、コード化モジュール314は、日焼け歴を、アレル非相互作用変数 w^i に含めることができる、バイナリー指標変数（患者が重症の日焼けの経歴を有するならば $d^k = 1$ 、それ以外は0）として表す。重症の日焼けは、主として黒色腫と関連性があるため、複数の腫瘍タイプに対するモデルを訓練する場合、この変数は、患者が重症の日焼けの経歴を有し、かつ腫瘍タイプが黒色腫であるならば1と同等であり、それ以外はゼロであると定義することもできる。

【0262】

1つの例において、コード化モジュール314は、ヒトゲノムにおける各遺伝子または転写産物についての特定の遺伝子または転写産物の発現レベルの分布を、TCGAなどの参照データベースを用いることによって、発現レベルの分布の要約統計量（例えば、平均値、中央値）として表す。具体的には、腫瘍タイプ黒色腫を有する試料におけるペプチド p^k について、ペプチド p^k の起源の遺伝子または転写産物の、測定された遺伝子または転写産物の発現レベルをアレル非相互作用変数 w^i に含むことができるだけでなく、TCGAによって測定された際の、黒色腫におけるペプチド p^k の起源の遺伝子または転写産物の、平均値及び/または中央値の遺伝子または転写産物発現も含むことができる。

30

【0263】

1つの例において、コード化モジュール314は、変異タイプを、変異タイプ（例えば、ミスセンス、フレームシフト、NMD誘導性など）のアルファベットについての長さ1のワン・ホットエンコーディングされた変数として表す。これらのワン・ホットエンコーディングされた変数を、アレル非相互作用変数 w^i に含めることができる。

40

【0264】

1つの例において、コード化モジュール314は、タンパク質のタンパク質レベルの特性を、ソースタンパク質のアノテーション（例えば、5'UTR長）の値として、アレル非相互作用変数 w^i において表す。別の例において、コード化モジュール314は、ペプチド p^i についてのソースタンパク質の残基レベルのアノテーションを、ペプチド p^i がヘリックスモチーフとオーバーラップするならば1と同等であり、それ以外は0であるか、または、ペプチド p^i がヘリックスモチーフ内に完全に含有されるならば1と同等である指標変数を、アレル非相互作用変数 w^i に含むことによって表す。別の例において、ヘリ

50

ックスモチーフアノテーション内に含有されるペプチド p^i における残基の割合を表す特性を、アレル非相互作用変数 w^i に含めることができる。

【0265】

1つの例において、コード化モジュール314は、ヒトプロテオームにおけるタンパク質またはアイソフォームのタイプを、ヒトプロテオームにおけるタンパク質またはアイソフォームの数と同等の長さを有する指標ベクトル o^k として表し、対応する要素 o^k_i は、ペプチド p^k がタンパク質 i に由来するならば1であり、それ以外は0である。

【0266】

1つの例において、コード化モジュール314は、ペプチド p^i のソース遺伝子 $G = gene(p^i)$ を L 個の可能なカテゴリーを有するカテゴリー変数として表す（ただし、 L は添え字を付したソース遺伝子の数の上限 $1, 2, \dots, L$ を示す）。

10

【0267】

1つの例において、コード化モジュール314は、ペプチド p^i の組織タイプ、細胞タイプ、腫瘍タイプ、または腫瘍組織学タイプ $T = 組織(p^i)$ を M 個の可能なカテゴリーを有するカテゴリー変数として表す（ただし、 M は添え字を付したタイプ $1, 2, \dots, M$ の数の上限を示す）。組織のタイプとしては、例えば、肺組織、心臓組織、腸組織、神経組織などを挙げる事ができる。細胞のタイプとしては、樹状細胞、マクロファージ、CD4 T細胞などを挙げる事ができる。肺腺癌、肺扁平上皮癌、メラノーマ、非ホジキンリンパ腫などを挙げる事ができる。

【0268】

コード化モジュール314はまた、ペプチド p^i 及び関連するMHCアレル h についての変数 z^i の全体的なセットを、アレル相互作用変数 x^i 及びアレル非相互作用変数 w^i の数値的表示が次々に連鎖している行ベクトルとしても表し得る。例えば、コード化モジュール314は、 z_h^i を、 $[x_h^i w^i]$ または $[w_i x_h^i]$ と同等の行ベクトルとして表し得る。

20

【0269】

VIII. 訓練モジュール

訓練モジュール316は、ペプチド配列に関連するMHCアレルによってペプチド配列が提示されるかどうかの尤度を生成する、1つ以上の提示モデルを構築する。具体的には、ペプチド配列 p^k 及びペプチド配列 p_k に関連するMHCアレル a^k のセットを与えられ、各提示モデルは、ペプチド配列 p^k が、関連するMHCアレル a^k のうちの1つ以上によって提示されるであろう尤度を示す、推定値 u_k を生成する。

30

【0270】

VIII. A. 概要

訓練モジュール316は、165に保存された提示情報から生成された、記憶装置170に保存された訓練データセットに基づいて、1つ以上の提示モデルを構築する。概して、提示モデルの具体的なタイプに関わらず、提示モデルのすべては、損失関数が最小化されるように、訓練データ170における独立変数と従属変数との間の依存性を捕捉する。具体的には、損失関数

40

$$\ell(y_{i \in S}, u_{i \in S}, \theta)$$

は、訓練データ170における1つ以上のデータ例 S についての従属変数 $y_{i \in S}$ の値と、提示モデルによって生成されたデータ例 S についての推定された尤度 $u_{i \in S}$ との間の矛盾を表す。本明細書の残りの部分を通じて言及される1つの特定の実現形態において、損失関数

$$\ell(y_{i \in S}, u_{i \in S}, \theta)$$

は、以下のような等式(1a)によって与えられる負のlog尤度関数である。

$$\ell(y_{i \in S}, u_{i \in S}; \theta) = \sum_{i \in S} (y_i \log u_i + (1 - y_i) \log(1 - u_i)) \quad (1a)$$

しかし、実際には、別の損失関数が使用されてもよい。例えば、質量分析イオン電流について予測がなされる場合、損失関数は、以下のような等式1bによって与えられる平均二

50

乗損失である。

$$\ell(y_{i \in S}, u_{i \in S}; \theta) = \sum_{i \in S} (\|y_i - u_i\|_2^2) \quad (1b)$$

【 0 2 7 1 】

提示モデルは、1つ以上のパラメータが、独立変数と従属変数との間の依存性を数学的に明記する、パラメトリックモデルであり得る。典型的に、損失関数 $(y_{i \in S}, u_{i \in S}, \theta)$

を最小化するパラメトリックタイプの提示モデルの種々のパラメータは、例えば、バッチ勾配アルゴリズム、確率的勾配アルゴリズムなどの、勾配ベースの数値的最適化アルゴリズムを通して決定される。あるいは、提示モデルは、モデル構造が、訓練データ170から決定され、固定されたパラメータのセットに厳密には基づかない、ノンパラメトリックモデルであり得る。

【 0 2 7 2 】

V I I I . B . アレル毎モデル

訓練モジュール316は、アレルごとベースでペプチドの提示尤度を予測するための提示モデルを構築し得る。この例において、訓練モジュール316は、単一のMHCアレルを発現する細胞から生成された訓練データ170におけるデータ例Sに基づいて、提示モデルを訓練し得る。

【 0 2 7 3 】

一実現形態では、訓練モジュール316は、

$$u_k^h = \Pr(p^k \text{ 提示}; \text{MHCアレル} h) = f(g_h(x_h^k; \theta_h)) \quad (2)$$

によって、特定のアレルhについてのペプチドp^kの推定提示尤度u^kをモデル化し、ただし、ペプチド配列x_h^kは、ペプチドp^k及び対応するMHCアレルhについてのコード化されたアレル相互作用変数を意味し、f(・)は、任意の関数であり、記載の便宜上、本明細書中を通して変換関数と呼ばれる。さらに、g_h(・)は、任意の関数であり、記載の便宜上、本明細書中を通して依存性関数と呼ばれ、MHCアレルhについて決定されたパラメータ_hのセットに基づいて、アレル相互作用変数x_h^kについての依存性スコアを生成する。各MHCアレルhについてのパラメータ_hのセットの値は、_hに関する損失関数を最小化することによって決定することができ、ここでiは、単一のMHCアレルhを発現する細胞から生成された訓練データ170のサブセットSにおける各例である。

【 0 2 7 4 】

依存性関数g_h(x_h^k; _h)の出力は、MHCアレルhが、少なくともアレル相互作用特性x_h^kに基づいて、及び特に、ペプチドp^kのペプチド配列のアミノ酸の位置に基づいて、対応する新生抗原を提示するかどうかを示す、MHCアレルhについての依存性スコアを表す。例えば、MHCアレルhについての依存性スコアは、MHCアレルhが、ペプチドp^kを提示する可能性が高い場合に、高い値を有し得、提示の可能性が高くない場合に、低い値を有し得る。変換関数f(・)は、入力を変換し、より具体的には、この例においてg_h(x_h^k; _h)によって生成された依存性スコアを、ペプチドp^kがMHCアレルによって提示されるであろう尤度を示す適切な値に変換する。

【 0 2 7 5 】

本明細書の残りの部分を通じて言及される1つの特定の実現形態において、f(・)は、適切なドメイン範囲について[0, 1]内の範囲を有する関数である。1つの例において、f(・)は、

$$f(z) = \frac{\exp(z)}{1 + \exp(z)} \quad (4)$$

によって与えられるexpit関数である。

別の例として、f(・)はまた、ドメインzの値が0以上である場合、

$$f(z) = \tanh(z) \quad (5)$$

10

20

30

40

50

によって与えられる双曲線正接関数であることもできる。あるいは、予測が、範囲 $[0, 1]$ の外側の値を有する質量分析イオン電流についてなされる場合、 $f(\cdot)$ は、例えば、恒等関数、指数関数、 \log 関数などの任意の関数であることができる。

【0276】

したがって、ペプチド配列 p^k が MHC アレル h によって提示されるであろうアレル毎尤度は、MHC アレル h についての依存性関数 $g_h(\cdot)$ をペプチド配列 p^k のコード化されたバージョンに適用して、対応する依存性スコアを生成することによって、生成することができる。依存性スコアは、ペプチド配列 p^k が MHC アレル h によって提示されるであろうアレル毎尤度を生成するように、変換関数 $f(\cdot)$ によって変換されてもよい。

【0277】

V I I I . B . 1 アレル相互作用変数についての依存性関数

本明細書を通して言及される1つの特定の實現形態において、依存性関数 $g_h(\cdot)$ は、 x_h^k における各アレル相互作用変数を、関連する MHC アレル h について決定されたパラメータ θ_h のセットにおける対応するパラメータと線形結合する、

$$g_h(x_h^i; \theta_h) = x_h^i \cdot \theta_h \quad (6)$$

によって与えられるアフィン関数である。

【0278】

本明細書を通して言及される別の特定の實現形態において、依存性関数 $g_h(\cdot)$ は、1つ以上の層において配置された一連のノードを有するネットワークモデル $NN_h(\cdot)$ によって表される、

$$g_h(x_h^i; \theta_h) = NN_h(x_h^i; \theta_h) \quad (7)$$

によって与えられるネットワーク関数である。ノードは、パラメータ θ_h のセットにおける関連するパラメータを各々有する接続を通して、他のノードに接続され得る。1つの特定のノードでの値は、特定のノードに関連する活性化関数によってマッピングされた関連するパラメータによって重み付けられた、特定のノードに接続されたノードの値の和として表され得る。アフィン関数と対照的に、ネットワークモデルは、提示モデルが非線形性、及び異なる長さのアミノ酸配列を有するプロセスデータを組み入れることができるため、有利である。具体的には、非線形モデリングを通して、ネットワークモデルは、ペプチド配列中の異なる位置のアミノ酸間の相互作用、及びこの相互作用がペプチド提示にいか

【0279】

概して、ネットワークモデル $NN_h(\cdot)$ は、人工ニューラルネットワーク (ANN)、畳み込みニューラルネットワーク (CNN)、深層ニューラルネットワーク (DNN) などのフィードフォワードネットワーク、及び/または、長・短期記憶ネットワーク (LSTM)、双方向再帰型ネットワーク、深層双方向再帰型ネットワークなどの再帰型ネットワークなどとして、構造化され得る。

【0280】

本明細書の残りの部分を通じて言及される1つの例において、 $h = 1, 2, \dots, m$ における各 MHC アレルは、別々のネットワークモデルに関連し、 $NN_h(\cdot)$ は、MHC アレル h に関連するネットワークモデルからの出力を意味する。

【0281】

図5は、任意の MHC アレル $h = 3$ に関連した例示的なネットワークモデル $NN_3(\cdot)$ を説明する。図5に示すように、MHC アレル $h = 3$ についてのネットワークモデル $NN_3(\cdot)$ は、層 $l = 1$ での3種類の入力ノード、層 $l = 2$ での4種類のノード、層 $l = 3$ での2種類のノード、及び層 $l = 4$ での1種類の出力ノードを含む。ネットワークモデル $NN_3(\cdot)$ は、10種類のパラメータ $\theta_3(1), \theta_3(2), \dots, \theta_3(10)$ のセットに関連している。ネットワークモデル $NN_3(\cdot)$ は、MHC アレル $h = 3$ についての3種類のアレル相互作用変数 $x_3^k(1), x_3^k(2)$ 、及び $x_3^k(3)$ についての入力値 (コード化されたポリペプチド配列データ及び使用される任意の他の訓練データを含む、個々のデータ例) を受け取り、値 $NN_3(x_3^k)$ を出力する。ネットワ

10

20

30

40

50

ーク関数は、異なるアレル相互作用変数をそれぞれが入力として取る1つ以上のネットワークモデルを含んでもよい。

【0282】

別の例において、特定されたMHCアレル $h = 1, 2, \dots, m$ は、単一ネットワークモデル $NN_H(\cdot)$ に関連しており、 $NN_h(\cdot)$ は、MHCアレル h に関連する単一ネットワークモデルの1つ以上の出力を意味する。そのような例において、パラメータ θ_h のセットは、単一ネットワークモデルについてのパラメータのセットに対応し得、したがって、パラメータ θ_h のセットは、すべてのMHCアレルによって共有され得る。

【0283】

図6Aは、MHCアレル $h = 1, 2, \dots, m$ によって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。図6Aに示すように、ネットワークモデル $NN_H(\cdot)$ は、MHCアレルに各々対応する、 m 個の出力ノードを含む。ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、MHCアレル $h = 3$ に対応する値 $NN_3(x_3^k)$ を含む、 m 個の値を出力する。

10

【0284】

さらに別の例において、単一ネットワークモデル $NN_H(\cdot)$ は、MHCアレル h のアレル相互作用変数 x_h^k 及びコード化されたタンパク質配列 d_h を与えられて依存性スコアを出力する、ネットワークモデルであり得る。そのような例において、パラメータ θ_h のセットは、再び、単一ネットワークモデルについてのパラメータのセットに対応し得、したがって、パラメータ θ_h のセットは、すべてのMHCアレルによって共有され得る。したがって、そのような例において、 $NN_h(\cdot)$ は、単一ネットワークモデルに対して入力 $[x_h^k, d_h]$ を与えられた、単一ネットワークモデル $NN_H(\cdot)$ の出力を意味する。そのようなネットワークモデルは、訓練データにおいて未知であったMHCアレルについてのペプチド提示確率を、単にそれらのタンパク質配列を特定することによって正しく予測することができるため、有利である。

20

【0285】

図6Bは、MHCアレルによって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。図6Bに示すように、ネットワークモデル $NN_H(\cdot)$ は、MHCアレル $h = 3$ のアレル相互作用変数及びタンパク質配列を入力として受け取り、MHCアレル $h = 3$ に対応する依存性スコア $NN_3(x_3^k)$ を出力する。

30

【0286】

さらに別の例において、依存性関数 $g_h(\cdot)$ は、

$$g_h(x_h^k; \theta_h) = g'_h(x_h^k; \theta'_h) + \theta_h^0$$

として表すことができ、式中、 $g'_h(x_h^k; \theta'_h)$ は、パラメータ θ'_h のセットを伴うアフィン関数、ネットワーク関数などであり、MHCアレル h についての提示のベースライン確率を表す、MHCアレルのアレル相互作用変数についてのパラメータのセットにおけるバイアスパラメータ θ_h^0 を伴う。

【0287】

別の実現形態において、バイアスパラメータ θ_h^0 は、MHCアレル h の遺伝子ファミリーにしたがって共有されてもよい。すなわち、MHCアレル h についてのバイアスパラメータ θ_h^0 は、遺伝子 (h) と同等であり得、遺伝子 (h) は、MHCアレル h の遺伝子ファミリーである。例えば、クラスI MHCアレルHLA-A*02:01、HLA-A*02:02、及びHLA-A*02:03は、「HLA-A」の遺伝子ファミリーに割り当てられてもよく、これらのMHCアレルの各々についてのバイアスパラメータ θ_h^0 が共有されてもよい。別の例として、クラスII MHCアレルHLA-DRB1:10:01、HLA-DRB1:11:01、及びHLA-DRB3:01:01を「HLA-DRB」の遺伝子ファミリーに割り当て、これらのMHCアレルのそれぞれのバイアスパラメータ θ_h^0 を共有することができる。

40

【0288】

例として、等式(2)に戻ると、アフィン依存性関数 $g_h(\cdot)$ を用いた、 $m = 4$ の異

50

なる特定されたMHCアレルの中でMHCアレル $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k^3 = f(x_3^k \cdot \theta_3)$$

によって生成することができ、式中、 x_3^k は、MHCアレル $h = 3$ について特定されたアレル相互作用変数であり、 θ_3 は、損失関数最小化を通してMHCアレル $h = 3$ について決定されたパラメータのセットである。

【0289】

別の例として、別々のネットワーク変換関数 $g_h(\cdot)$ を用いた、 $m = 4$ の異なる特定されたMHCアレルの中でMHCアレル $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k^3 = f(NN_3(x_3^k; \theta_3))$$

によって生成することができ、式中、 x_3^k は、MHCアレル $h = 3$ について特定されたアレル相互作用変数であり、 θ_3 は、MHCアレル $h = 3$ に関連するネットワークモデル $NN_3(\cdot)$ について決定されたパラメータのセットである。

【0290】

図7は、例示的なネットワークモデル $NN_3(\cdot)$ を用いた、MHCアレル $h = 3$ に関連したペプチド p^k の提示尤度の生成を説明する。図7に示すように、ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成する。出力は、関数 $f(\cdot)$ によってマッピングされて、推定提示尤度 u_k を生成する。

【0291】

VIII.B.2. アレル非相互作用変数を伴うアレルごと

—実現形態では、訓練モジュール316は、アレル非相互作用変数を組み入れて、

$$u_k^h = \Pr(p^k \text{ 提示}) = f(g_w(w^k; \theta_w) + g_h(x_h^i; \theta_h)) \quad (8)$$

によって、ペプチド p^k の推定提示尤度 u_k をモデル化し、式中、 w^k は、ペプチド p^k についてのコード化されたアレル非相互作用変数を意味し、 $g_w(\cdot)$ は、アレル非相互作用変数について決定されたパラメータ θ_w のセットに基づく、アレル非相互作用変数 w^k についての関数である。具体的には、各MHCアレル h についてのパラメータ θ_h のセット及びアレル非相互作用変数についてのパラメータ θ_w のセットの値を、 θ_h 及び θ_w に関する損失関数を最小化することによって決定することができ、 i は、単一のMHCアレルを発現する細胞から生成された訓練データ170のサブセット S における各例である。

【0292】

依存性関数 $g_w(w^k; \theta_w)$ の出力は、アレル非相互作用変数の影響に基づいて、1つ以上のMHCアレルによってペプチド p^k が提示されるかどうかを示す、アレル非相互作用変数についての依存性スコアを表す。例えば、アレル非相互作用変数についての依存性スコアは、ペプチド p^k の提示に正の影響を及ぼすことが公知であるC末端フランキング配列とペプチド p^k が結合している場合は、高い値を有し得、ペプチド p^k の提示に負の影響を及ぼすことが公知であるC末端フランキング配列とペプチド p^k が結合している場合は、低い値を有し得る。

【0293】

等式(8)によると、ペプチド配列 p^k がMHCアレル h によって提示されるであろうアレル毎尤度は、MHCアレル h についての関数 $g_h(\cdot)$ を、ペプチド配列 p^k のコード化されたバージョンに適用して、アレル相互作用変数について対応する依存性スコアを生成することによって、生成することができる。アレル非相互作用変数についての関数 $g_w(\cdot)$ もまた、アレル非相互作用変数についての依存性スコアを生成するように、アレル非相互作用変数のコード化されたバージョンに適用される。両方のスコアが組み合わせられ、組み合わせられたスコアが、MHCアレル h によってペプチド配列 p^k が提示されるであろうアレル毎尤度を生成するように、変換関数 $f(\cdot)$ によって変換される。

10

20

30

40

50

【 0 2 9 4 】

あるいは、訓練モジュール 3 1 6 は、等式 (2) においてアレル非相互作用変数 w^k をアレル相互作用変数 x_h^k に付加することにより、予測におけるアレル非相互作用変数 w^k を含んでもよい。したがって、提示尤度は、

$$u_k^h = \Pr(p^k \text{ 提示}; \text{アレル } h) = f(g_h([x_h^k w^k]; \theta_h)) \quad (9)$$

によって与えられ得る。

【 0 2 9 5 】

V I I I . B . 3 アレル非相互作用変数についての依存性関数

アレル相互作用変数についての依存性関数 $g_h(\cdot)$ と同様に、アレル非相互作用変数についての依存性関数 $g_w(\cdot)$ は、アフィン関数、または別々のネットワークモデルがアレル非相互作用変数 w^k に関連しているネットワーク関数であり得る。 10

【 0 2 9 6 】

具体的には、依存性関数 $g_w(\cdot)$ は、 w^k におけるアレル非相互作用変数を、パラメータ θ_w のセットにおける対応するパラメータと線形結合する、

$$g_w(w^k; \theta_w) = w^k \cdot \theta_w$$

によって与えられるアフィン関数である。

【 0 2 9 7 】

依存性関数 $g_w(\cdot)$ はまた、パラメータ θ_w のセットにおける関連するパラメータを有するネットワークモデル $NN_w(\cdot)$ によって表される、 20

$$g_w(w^k; \theta_w) = NN_w(w^k; \theta_w)$$

によって与えられるネットワーク関数である。ネットワーク関数は、異なるアレル非相互作用変数をそれぞれが入力として取る 1 つ以上のネットワークモデルを含んでもよい。

【 0 2 9 8 】

別の例において、アレル非相互作用変数についての依存性関数 $g_w(\cdot)$ は、

$$g_w(w^k; \theta_w) = g'_w(w^k; \theta'_w) + h(m^k; \theta_w^m) \quad (10)$$

によって与えられ得、式中、 $g'_w(w^k; \theta'_w)$ は、アレル非相互作用パラメータ θ'_w のセットを伴うアフィン関数、ネットワーク関数などであり、 m^k は、ペプチド p^k についての m RNA 定量測定値であり、 $h(\cdot)$ は、定量測定値を変換する関数であり、かつ θ_w^m は、 m RNA 定量測定値についての依存性スコアを生成するように m RNA 定量測定値と組み合わせられる、アレル非相互作用変数についてのパラメータのセットにおけるパラメータである。本明細書の残りの部分を通じて言及される 1 つの特定の実施形態において、 $h(\cdot)$ は \log 関数であるが、実際には、 $h(\cdot)$ は、様々な異なる関数のうちのいずれか 1 つであり得る。 30

【 0 2 9 9 】

さらに別の例において、アレル非相互作用変数についての依存性関数 $g_w(\cdot)$ は、

$$g_w(w^k; \theta_w) = g'_w(w^k; \theta'_w) + \theta_w^o \cdot o^k \quad (11)$$

によって与えられ、式中、 $g'_w(w^k; \theta'_w)$ は、アレル非相互作用パラメータ θ'_w のセットを伴うアフィン関数、ネットワーク関数などであり、 o^k は、ペプチド p^k についてヒトプロテオームにおけるタンパク質及びアイソフォームを表す、セクション V I I . C . 2 で述べた指標ベクトルであり、かつ θ_w^o は、指標ベクトルと組み合わせられるアレル非相互作用変数についてのパラメータのセットにおける、パラメータのセットである。1 つのバリエーションにおいて、 o^k 及びパラメータ θ_w^o のセットの次元が有意に高い場合、 40

$$\lambda \cdot \|\theta_w^o\|$$

(

$\|\cdot\|$

は、 L_1 ノルム、 L_2 ノルム、組み合わせなどを表す) などのパラメータ正則化項を、パラメータの値を決定する時に損失関数に加えることができる。ハイパーパラメータの最適値を、適切な方法を通して決定することができる。 50

【0300】

さらに別の例において、アレル非相互作用変数に対する依存性関数 $g_w(\cdot)$ は下式により与えられる。すなわち、

$$g_w(\mathbf{w}^k; \boldsymbol{\theta}_w) = g'_w(\mathbf{w}^k; \boldsymbol{\theta}'_w) + \sum_{l=1}^L \mathbb{1}(\text{遺伝子}(p^k) = l) \cdot \boldsymbol{\theta}_w^l \quad (12)$$

ただし、 $g'_w(\mathbf{w}^k; \boldsymbol{\theta}'_w)$ は、アレル非相互作用パラメータ $\boldsymbol{\theta}'_w$ のセットを伴うアフィン関数、ネットワーク関数などであり、

 $\mathbb{1}(\text{遺伝子}(p^k) = l)$

は、ペプチド p^k がアレル非相互作用変数に関して上記に述べたソース遺伝子 l に由来するものである場合に 1 に等しいインジケータ関数であり、 $\boldsymbol{\theta}_w^l$ はソース遺伝子 l の「抗原性」を示すパラメータである。1つのバリエーションにおいて、 L が十分に大きく、したがって、パラメータの数 $\boldsymbol{\theta}_w^{l=1, 2, \dots, L}$ が十分に大きい場合、

$$\lambda \cdot \|\boldsymbol{\theta}_w^l\|$$

のようなパラメータ正則化項（ただし、

$$\|\cdot\|$$

は、 L_1 ノルム、 L_2 ノルム、組み合わせなど）をパラメータの値を決定する際に損失関数に加えることができる。ハイパーパラメータの最適値は適当な方法によって決定することができる。

【0301】

さらに別の例において、アレル非相互作用変数に対する依存性関数 $g_w(\cdot)$ は下式により与えられる。すなわち、

$$g_w(\mathbf{w}^k; \boldsymbol{\theta}_w) = g'_w(\mathbf{w}^k; \boldsymbol{\theta}'_w) + \sum_{m=1}^M \sum_{l=1}^L \mathbb{1}(\text{遺伝子}(p^k) = l, \text{組織}(p^k) = m) \cdot \boldsymbol{\theta}_w^{lm} \quad (12b)$$

ただし、 $g'_w(\mathbf{w}^k; \boldsymbol{\theta}'_w)$ は、アレル非相互作用パラメータ $\boldsymbol{\theta}'_w$ のセットを伴うアフィン関数、ネットワーク関数などであり、

 $\mathbb{1}(\text{遺伝子}(p^k) = l, \text{組織}(p^k) = m)$

は、アレル非相互作用変数に関して上記に述べたようにペプチド p^k がソース遺伝子 l に由来するものである場合、かつペプチド p^k が組織タイプ m に由来するものである場合に 1 に等しいインジケータ関数であり、 $\boldsymbol{\theta}_w^{lm}$ はソース遺伝子 l と組織タイプ m との組み合わせの抗原性を示すパラメータである。詳細には、組織タイプ m の遺伝子 l の抗原性は、組織タイプ m の細胞が、RNA発現及びペプチド配列コンテキストについての調節後に遺伝子 l 由来のペプチドを提示する残留傾向を示し得る。

【0302】

1つのバリエーションにおいて、 L または M が十分に大きく、したがって、パラメータの数 $\boldsymbol{\theta}_w^{l=1, 2, \dots, L, M}$ が十分に大きい場合、

$$\lambda \cdot \|\boldsymbol{\theta}_w^{lm}\|$$

のようなパラメータ正則化項（ただし、

$$\|\cdot\|$$

は、 L_1 ノルム、 L_2 ノルム、組み合わせなど）をパラメータの値を決定する際に損失関数に加えることができる。ハイパーパラメータの最適値は適当な方法によって決定することができる。別のバリエーションにおいて、同じソース遺伝子に対する係数が組織タイプ間で大きく異ならないように、パラメータの値を決定する際にパラメータ正則化項を損失関数に加えることができる。例えば、以下のようなペナルティ項：

$$\lambda \cdot \sum_{l=1}^L \sqrt{\sum_{m=1}^M (\boldsymbol{\theta}_w^{lm} - \bar{\boldsymbol{\theta}}_w^l)^2}$$

（ただし、

10

20

30

40

50

$\overline{\theta_w^l}$

はソース遺伝子 l の組織タイプにわたった平均の抗原性である)は、損失関数中の異なる組織タイプにわたった抗原性の標準偏差にペナルティを付加することができる。

【0303】

実際には、式(10)、(11)、(12a)及び(12b)のいずれかの追加項を組み合わせることによってアレル非相互作用変数に関する依存性関数 $g_w(\cdot)$ を生成することができる。例えば、式(10)のmRNA定量測定値を示す項 $h(\cdot)$ と式(12)のソース遺伝子の抗原性を示す項とを他の任意のアフィン関数またはネットワーク関数とともに互いに加え合わせることで、アレル非相互作用変数に関する依存性関数を生成することができる。

10

【0304】

例として、等式(8)に戻ると、アフィン変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m=4$ の異なる特定されたMHCアレルの中でMHCアレル $h=3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k^3 = f(w^k \cdot \theta_w + x_3^k \cdot \theta_3)$$

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル非相互作用変数であり、 w は、アレル非相互作用変数について決定されたパラメータのセットである。

【0305】

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m=4$ の異なる特定されたMHCアレルの中でMHCアレル $h=3$ によってペプチド p^k が提示されるであろう尤度は、

20

$$u_k^3 = f(NN_w(w^k; \theta_w) + NN_3(x_3^k; \theta_3))$$

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル相互作用変数であり、 w は、アレル非相互作用変数について決定されたパラメータのセットである。

【0306】

図8は、例示的なネットワークモデル $NN_3(\cdot)$ 及び $NN_w(\cdot)$ を用いた、MHCアレル $h=3$ に関連したペプチド p^k の提示尤度の生成を説明する。図8に示すように、ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h=3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成する。ネットワークモデル $NN_w(\cdot)$ は、ペプチド p^k についてのアレル非相互作用変数 w^k を受け取り、出力 $NN_w(w^k)$ を生成する。出力は、組み合わせられ、関数 $f(\cdot)$ によってマッピングされて、推定提示尤度 u_k を生成する。

30

【0307】

VIII.C. 複数アレルモデル

訓練モジュール316はまた、2つ以上のMHCアレルが存在する複数アレル設定においてペプチドの提示尤度を予測するための提示モデルを構築し得る。この例において、訓練モジュール316は、単一のMHCアレルを発現する細胞、複数のMHCアレルを発現する細胞、またはそれらの組み合わせから生成された訓練データ170におけるデータ例Sに基づいて、提示モデルを訓練し得る。

40

【実施例】

【0308】

VIII.C.1. 実施例1：アレル毎モデルの最大値

一実現形態では、訓練モジュール316は、複数のMHCアレル H のセットに関連したペプチド p^k の推定提示尤度 u_k を、等式(2)~(11)と共に上記で説明したような、単一アレルを発現する細胞に基づいて決定されたセット H におけるMHCアレル h の各々について決定された提示尤度

$$u_k^{h \in H}$$

の関数としてモデル化する。具体的には、提示尤度 u_k は、

50

$u_k^{h \in H}$

の任意の関数であることができる。一実現形態では、等式(12)に示すように、関数は最大値関数であり、提示尤度 u_k は、セット H における各 MHC アレル h についての提示尤度の最大値として決定することができる。

$$u_k = \Pr(p^k \text{ 提示}; \text{アレル } H) = \max(u_k^{h \in H})$$

【0309】

V I I I . C . 2 . 実施例 2 . 1 : 和の関数モデル

一実現形態では、訓練モジュール 316 は、ペプチド p^k の推定提示尤度 u_k を、

$$u_k = \Pr(p^k \text{ 提示}) = f\left(\sum_{h=1}^m a_h^k \cdot g_h(x_h^k; \theta_h)\right) \quad (13)$$

10

によってモデル化し、式中、要素 a_h^k は、ペプチド配列 p^k に関連する複数の MHC アレル H について 1 であり、 x_h^k は、ペプチド p^k 及び対応する MHC アレルについてのコード化されたアレル相互作用変数を意味する。各 MHC アレル h についてのパラメータ θ_h のセットの値は、 θ_h に関する損失関数を最小化することによって決定することができる。 i は、単一の MHC アレルを発現する細胞及び/または複数の MHC アレルを発現する細胞から生成された訓練データ 170 のサブセット S における各例である。依存性関数 g_h は、セクション V I I I . B . 1 . において上記で導入された依存性関数 g_h のいずれかの形態であり得る。

【0310】

20

等式(13)によると、ペプチド配列 p^k が 1 つ以上の MHC アレル h によって提示されるであろう提示尤度は、依存性関数 $g_h(\cdot)$ を、MHC アレル H の各々についてペプチド配列 p^k のコード化されたバージョンに適用して、アレル相互作用変数についての対応するスコアを生成することによって、生成することができる。各 MHC アレル h についてのスコアが組み合わされて、ペプチド配列 p^k が MHC アレル H のセットによって提示されるであろう提示尤度を生成するように変換関数 $f(\cdot)$ によって変換される。

【0311】

等式(13)の提示モデルは、各ペプチド p^k についての関連するアレルの数が 1 よりも大きいことができる点で、等式(2)のアレル毎モデルとは異なる。換言すると、 a_h^k における 1 つよりも多い要素が、ペプチド配列 p^k に関連する複数の MHC アレル H について 1 の値を有することができる。

30

【0312】

例として、アフィン変換関数 $g_h(\cdot)$ を用いた、 $m = 4$ の異なる特定された MHC アレルの中で MHC アレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(x_2^k \cdot \theta_2 + x_3^k \cdot \theta_3)$$

によって生成することができ、式中、 x_2^k 、 x_3^k は、MHC アレル $h = 2$ 、 $h = 3$ について特定されたアレル相互作用変数であり、 θ_2 、 θ_3 は、MHC アレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

【0313】

40

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定された MHC アレルの中で MHC アレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(NN_2(x_2^k; \theta_2) + NN_3(x_3^k; \theta_3))$$

によって生成することができ、式中、 $NN_2(\cdot)$ 、 $NN_3(\cdot)$ は、MHC アレル $h = 2$ 、 $h = 3$ について特定されたネットワークモデルであり、 θ_2 、 θ_3 は、MHC アレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

【0314】

図 9 は、例示的なネットワークモデル $NN_2(\cdot)$ 及び $NN_3(\cdot)$ を用いた、MHC アレル $h = 2$ 、 $h = 3$ に関連したペプチド p^k の提示尤度の生成を説明する。図 9 に示す

50

ように、ネットワークモデル $NN_2(\cdot)$ は、MHCアレル $h = 2$ についてのアレル相互作用変数 x_2^k を受け取り、出力 $NN_2(x_2^k)$ を生成し、ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成する。出力は、組み合わせられ、関数 $f(\cdot)$ によってマッピングされて、推定提示尤度 u_k を生成する。

【0315】

VIII.C.3. 実施例2.2: アレル非相互作用変数を伴う和の関数モデル

— 実現形態では、訓練モジュール316は、アレル非相互作用変数を組み入れて、

$$u_k = \Pr(p^k \text{ 提示}) = f\left(g_w(w^k; \theta_w) + \sum_{h=1}^m a_h^k \cdot g_h(x_h^k; \theta_h)\right) \quad (14)$$

10

によって、ペプチド p^k の推定提示尤度 u_k をモデル化し、式中、 w^k は、ペプチド p^k についてのコード化されたアレル非相互作用変数を意味する。具体的には、各MHCアレル h についてのパラメータ θ_h のセット及びアレル非相互作用変数についてのパラメータ θ_w のセットの値を、 θ_h 及び θ_w に関する損失関数を最小化することによって決定することができ、 i は、単一のMHCアレルを発現する細胞及び/または複数のMHCアレルを発現する細胞から生成された訓練データ170のサブセット S における各例である。依存性関数 g_w は、セクションVIII.B.3.において上記で導入された依存性関数 g_w のいずれかの形態であり得る。

【0316】

20

したがって、等式(14)によると、1つ以上のMHCアレル H によってペプチド配列 p^k が提示されるであろう提示尤度は、関数 $g_h(\cdot)$ を、MHCアレル H の各々についてペプチド配列 p^k のコード化されたバージョンに適用して、各MHCアレル h のアレル相互作用変数について対応する依存性スコアを生成することによって、生成することができる。アレル非相互作用変数についての関数 $g_w(\cdot)$ もまた、アレル非相互作用変数についての依存性スコアを生成するように、アレル非相互作用変数のコード化されたバージョンに適用される。スコアが組み合わせられ、組み合わせられたスコアが、MHCアレル H によってペプチド配列 p^k が提示されるであろう提示尤度を生成するように、変換関数 $f(\cdot)$ によって変換される。

【0317】

30

等式(14)の提示モデルにおいて、各ペプチド p^k についての関連するアレルの数は、1よりも大きいことができる。換言すると、 a_h^k における1つよりも多い要素が、ペプチド配列 p^k に関連する複数のMHCアレル H について1の値を有することができる。

【0318】

例として、アフィン変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定されたMHCアレルの中でMHCアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(w^k \cdot \theta_w + x_2^k \cdot \theta_2 + x_3^k \cdot \theta_3)$$

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル非相互作用変数であり、 θ_w は、アレル非相互作用変数について決定されたパラメータのセットである。

40

【0319】

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定されたMHCアレルの中でMHCアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(NN_w(w^k; \theta_w) + NN_2(x_2^k; \theta_2) + NN_3(x_3^k; \theta_3))$$

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル相互作用変数であり、 θ_w は、アレル非相互作用変数について決定されたパラメータのセットである。

【0320】

50

図10は、例示的なネットワークモデル $NN_2(\cdot)$ 、 $NN_3(\cdot)$ 、及び $NN_w(\cdot)$ を用いた、MHCアレル $h = 2$ 、 $h = 3$ に関連したペプチド p^k の提示尤度の生成を説明する。図10に示すように、ネットワークモデル $NN_2(\cdot)$ は、MHCアレル $h = 2$ についてのアレル相互作用変数 x_2^k を受け取り、出力 $NN_2(x_2^k)$ を生成する。ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成する。ネットワークモデル $NN_w(\cdot)$ は、ペプチド p^k についてのアレル非相互作用変数 w^k を受け取り、出力 $NN_w(w^k)$ を生成する。出力は、組み合わせられ、関数 $f(\cdot)$ によってマッピングされて、推定提示尤度 u_k を生成する。

【0321】

あるいは、訓練モジュール316は、等式(15)においてアレル非相互作用変数 w^k をアレル相互作用変数 x_h^k に付加することにより、予測におけるアレル非相互作用変数 w^k を含んでもよい。したがって、提示尤度は、

$$u_k = \Pr(p^k \text{ 提示}) = f\left(\sum_{h=1}^m a_h^k \cdot g_h([x_h^k w^k]; \theta_h)\right) \quad (15)$$

によって与えられ得る。

【0322】

VIII.C.4.実施例3.1:暗黙のアレル毎尤度を用いたモデル

別の実現形態において、訓練モジュール316は、ペプチド p^k の推定提示尤度 u_k を

$$u_k = \Pr(p^k \text{ 提示}) = r\left(s(v = [a_1^k \cdot u_k^1(\theta) \dots a_m^k \cdot u_k^m(\theta)])\right) \quad (16)$$

によってモデル化し、式中、要素 a_h^k は、ペプチド配列 p^k に関連する複数のMHCアレル $h \in H$ について1であり、 u_k^h は、MHCアレル h についての暗黙のアレル毎提示尤度であり、ベクトル v は、要素 v_h が、 $a_h^k \cdot u_k^h$ に対応するベクトルであり、 $s(\cdot)$ は、 v の要素をマッピングする関数であり、かつ $r(\cdot)$ は、入力の値を所定の範囲中にクリップするクリッピング関数である。より詳細に下記に記載するように、 $s(\cdot)$ は、総和関数または二次関数であってもよいが、他の実施形態において、 $s(\cdot)$ は、最大値関数などの任意の関数であり得ることが認識される。暗黙のアレル毎尤度についてのパラメータのセットの値は、に関する損失関数を最小化することによって決定することができ、 i は、単一のMHCアレルを発現する細胞及び/または複数のMHCアレルを発現する細胞から生成された訓練データ170のサブセット S における各例である。

【0323】

等式(17)の提示モデルにおける提示尤度は、各々が、個々のMHCアレル h によってペプチド p^k が提示されるであろう尤度に対応する、暗黙のアレル毎提示尤度 u_k^h の関数としてモデル化される。暗黙のアレル毎尤度は、暗黙のアレル毎尤度についてのパラメータが、単一アレル設定に加えて、提示されるペプチドと対応するMHCアレルとの間の直接の関連が未知である複数アレル設定から学習され得る点で、セクションVIII.Bのアレル毎提示尤度とは異なる。したがって、複数アレル設定において、提示モデルは、ペプチド p^k が全体としてMHCアレル H のセットによって提示されるかどうかを推定できるだけでなく、どのMHCアレル h がペプチド p^k を提示した可能性が最も高いかを示す個々の尤度 u_k^h も提供することもできる。この利点は、提示モデルが、単一のMHCアレルを発現する細胞についての訓練データを伴わずに暗黙の尤度を生成できることである。

【0324】

本明細書の残りの部分を通じて言及される1つの特定の実現形態において、 $r(\cdot)$ は、範囲 $[0, 1]$ を有する関数である。例えば、 $r(\cdot)$ は、クリップ関数：

$$r(z) = \min(\max(z, 0), 1)$$

10

20

30

40

50

であってもよく、 z と 1 の間の最小値が、提示尤度 u_k として選ばれる。別の実現形態において、 $r(\cdot)$ は、

$$r(z) = \tanh(z)$$

として与えられる双曲線正接関数であり、ドメイン z の値は、0 以上である。

【0325】

V I I I . C . 5 . 実施例 3 . 2 : 関数の和モデル

1 つの特定の実現形態において、 $s(\cdot)$ は、総和関数であり、提示尤度は、暗黙のアレル毎提示尤度を総和することによって与えられる。

$$u_k = \Pr(\mathbf{p}^k \text{ 提示}) = r\left(\sum_{h=1}^m a_h^k \cdot u_k^h(\boldsymbol{\theta})\right) \quad (17)$$

10

【0326】

1 つの実現形態では、MHC アレル h についての暗黙のアレル毎提示尤度を、

$$u_k^h = f(g_h(x_h^k; \boldsymbol{\theta}_h)) \quad (18)$$

によって生成して、提示尤度が、

$$u_k = \Pr(\mathbf{p}^k \text{ 提示}) = r\left(\sum_{h=1}^m a_h^k \cdot f(g_h(x_h^k; \boldsymbol{\theta}_h))\right) \quad (19)$$

によって推定されるようにする。

20

【0327】

等式(19)によると、1 つ以上の MHC アレル H によってペプチド配列 \mathbf{p}^k が提示されるであろう提示尤度は、関数 $g_h(\cdot)$ を、MHC アレル H の各々についてペプチド配列 \mathbf{p}^k のコード化されたバージョンに適用して、アレル相互作用変数についての対応する依存性スコアを生成することによって、生成することができる。各依存性スコアは、最初に、暗黙のアレル毎提示尤度 u_k^h を生成するように、関数 $f(\cdot)$ によって変換される。アレル毎尤度 u_k^h が組み合わされ、組み合わされた尤度にクリッピング関数が、値を範囲 $[0, 1]$ 中にクリップするために適用されて、ペプチド配列 \mathbf{p}^k が MHC アレル H のセットによって提示されるであろう提示尤度が生成され得る。依存性関数 g_h は、セクション V I I I . B . 1 . において上記で導入された依存性関数 g_h のいずれかの形態であり得る。

30

【0328】

例として、アフィン変換関数 $g_h(\cdot)$ を用いた、 $m = 4$ の異なる特定された MHC アレルの中で MHC アレル $h = 2$ 、 $h = 3$ によってペプチド \mathbf{p}^k が提示されるであろう尤度は、

$$u_k = r\left(f(x_2^k \cdot \boldsymbol{\theta}_2) + f(x_3^k \cdot \boldsymbol{\theta}_3)\right)$$

によって生成することができ、式中、 x_2^k 、 x_3^k は、MHC アレル $h = 2$ 、 $h = 3$ について特定されたアレル相互作用変数であり、 $\boldsymbol{\theta}_2$ 、 $\boldsymbol{\theta}_3$ は、MHC アレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

40

【0329】

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定された MHC アレルの中で MHC アレル $h = 2$ 、 $h = 3$ によってペプチド \mathbf{p}^k が提示されるであろう尤度は、

$$u_k = r\left(f(NN_2(x_2^k; \boldsymbol{\theta}_2)) + f(NN_3(x_3^k; \boldsymbol{\theta}_3))\right)$$

によって生成することができ、式中、 $NN_2(\cdot)$ 、 $NN_3(\cdot)$ は、MHC アレル $h = 2$ 、 $h = 3$ について特定されたネットワークモデルであり、 $\boldsymbol{\theta}_2$ 、 $\boldsymbol{\theta}_3$ は、MHC アレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

【0330】

50

図 11 は、例示的なネットワークモデル $NN_2(\cdot)$ 及び $NN_3(\cdot)$ を用いた、MHC アレル $h = 2$ 、 $h = 3$ に関連したペプチド p^k の提示尤度の生成を説明する。図 9 に示すように、ネットワークモデル $NN_2(\cdot)$ は、MHC アレル $h = 2$ についてのアレル相互作用変数 x_2^k を受け取り、出力 $NN_2(x_2^k)$ を生成し、ネットワークモデル $NN_3(\cdot)$ は、MHC アレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成する。各出力は、関数 $f(\cdot)$ によってマッピングされ、組み合わせられて、推定提示尤度 u_k を生成する。

【0331】

別の実現形態において、予測が、質量分析イオン電流の \log についてなされる場合、 $r(\cdot)$ は \log 関数であり、 $f(\cdot)$ は指数関数である。

10

【0332】

VIII.C.6. 実施例 3.3: アレル非相互作用変数を伴う関数の和モデル

1つの実現形態では、MHC アレル h についての暗黙のアレル毎提示尤度を、

$$u_k^h = f(g_h(x_h^k; \theta_h) + g_w(w^k; \theta_w)) \quad (20)$$

によって生成して、提示尤度が、

$$u_k = \Pr(p^k \text{ 提示}) = r\left(\sum_{h=1}^m a_h^k \cdot f(g_w(w^k; \theta_w) + g_h(x_h^k; \theta_h))\right) \quad (21)$$

によって生成されるようにして、ペプチド提示に、アレル非相互作用変数の影響を組み入れる。

20

【0333】

等式(21)によると、1つ以上のMHCアレルHによってペプチド配列 p^k が提示されるであろう提示尤度は、関数 $g_h(\cdot)$ を、MHCアレルHの各々についてペプチド配列 p^k のコード化されたバージョンに適用して、各MHCアレルhのアレル相互作用変数について対応する依存性スコアを生成することによって、生成することができる。アレル非相互作用変数についての関数 $g_w(\cdot)$ もまた、アレル非相互作用変数についての依存性スコアを生成するように、アレル非相互作用変数のコード化されたバージョンに適用される。アレル非相互作用変数のスコアが、アレル相互作用変数の依存性スコアの各々に組み合わせられる。組み合わせられたスコアの各々が、暗黙のアレル毎提示尤度を生成するように、関数 $f(\cdot)$ によって変換される。暗黙の尤度が組み合わせられ、組み合わせられた出力にクリッピング関数が、値を範囲 $[0, 1]$ 中にクリップするために適用されて、MHCアレルHによってペプチド配列 p^k が提示されるであろう提示尤度が生成され得る。依存性関数 g_w は、セクションVIII.B.3.において上記で導入された依存性関数 g_w のいずれかの形態であり得る。

30

【0334】

例として、アフィン変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定されたMHCアレルの中でMHCアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = r\left(f(w^k \cdot \theta_w + x_2^k \cdot \theta_2) + f(w^k \cdot \theta_w + x_3^k \cdot \theta_3)\right)$$

40

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル非相互作用変数であり、 w は、アレル非相互作用変数について決定されたパラメータのセットである。

【0335】

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定されたMHCアレルの中でMHCアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = r\left(f(NN_w(w^k; \theta_w) + NN_2(x_2^k; \theta_2)) + f(NN_w(w^k; \theta_w) + NN_3(x_3^k; \theta_3))\right)$$

50

によって生成することができ、式中、 w^k は、ペプチド p^k について特定されたアレル相互作用変数であり、 w は、アレル非相互作用変数について決定されたパラメータのセットである。

【0336】

図12は、例示的なネットワークモデル $NN_2(\cdot)$ 、 $NN_3(\cdot)$ 、及び $NN_w(\cdot)$ を用いた、MHCアレル $h = 2$ 、 $h = 3$ に関連したペプチド p^k の提示尤度の生成を説明する。図12に示すように、ネットワークモデル $NN_2(\cdot)$ は、MHCアレル $h = 2$ についてのアレル相互作用変数 x_2^k を受け取り、出力 $NN_2(x_2^k)$ を生成する。ネットワークモデル $NN_w(\cdot)$ は、ペプチド p^k についてのアレル非相互作用変数 w^k を受け取り、出力 $NN_w(w^k)$ を生成する。出力は、組み合わせられ、関数 $f(\cdot)$ によってマッピングされる。ネットワークモデル $NN_3(\cdot)$ は、MHCアレル $h = 3$ についてのアレル相互作用変数 x_3^k を受け取り、出力 $NN_3(x_3^k)$ を生成し、これも、同じネットワークモデル $NN_w(\cdot)$ の出力 $NN_w(w^k)$ と組み合わせられ、関数 $f(\cdot)$ によってマッピングされる。両方の出力が組み合わせられて、推定提示尤度 u_k を生成する。

10

【0337】

別の実現形態では、MHCアレル h についての暗黙のアレル毎提示尤度を、

$$u_k^h = f(g_h([x_h^k w^k]; \theta_h)) \quad (22)$$

によって生成して、提示尤度が、

$$u_k = \Pr(p^k \text{ 提示}) = r \left(\sum_{h=1}^m a_h^k \cdot f(g_h([x_h^k w^k]; \theta_h)) \right)$$

20

によって生成されるようにする。

【0338】

VIII.C.7. 実施例4：二次モデル

一実現形態では、 $s(\cdot)$ は、二次関数であり、ペプチド p^k の推定提示尤度 u_k は、

$$u_k = \Pr(p^k \text{ 提示}) = \sum_{h=1}^m a_h^k \cdot u_k^h(\theta) - \sum_{h=1}^m \sum_{j < h} a_h^k \cdot a_j^k \cdot u_k^h(\theta) \cdot u_k^j(\theta) \quad (23)$$

によって与えられ、式中、要素 u_k^h は、MHCアレル h についての暗黙のアレル毎提示尤度である。暗黙のアレル毎尤度についてのパラメータのセットの値は、 θ に関する損失関数を最小化することによって決定することができ、 i は、単一のMHCアレルを発現する細胞及び/または複数のMHCアレルを発現する細胞から生成された訓練データ170のサブセット S における各例である。暗黙のアレル毎提示尤度は、上記の等式(18)、(20)、及び(22)において示すいずれかの形態であり得る。

30

【0339】

一態様において、等式(23)のモデルは、ペプチド配列 p^k が、2つのMHCアレルによって同時に提示されるであろう可能性が存在し、2つのHLAアレルによる提示は統計学的に独立していることを含意し得る。

【0340】

等式(23)によると、1つ以上のMHCアレル H によってペプチド配列 p^k が提示されるであろう提示尤度は、暗黙のアレル毎提示尤度を組み合わせること、及び、MHCアレル H によってペプチド配列 p^k が提示されるであろう提示尤度を生成するように、MHCアレルの各ペアがペプチド p^k を同時に提示するであろう尤度を総和から差し引くことによって、生成することができる。

40

【0341】

例として、アフィン変換関数 $g_h(\cdot)$ を用いた、 $m = 4$ の異なる特定されたHLAアレルの中でHLAアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(x_2^k \cdot \theta_2) + f(x_3^k \cdot \theta_3) - f(x_2^k \cdot \theta_2) \cdot f(x_3^k \cdot \theta_3)$$

50

によって生成することができ、式中、 x_2^k 、 x_3^k は、HLAアレル $h = 2$ 、 $h = 3$ について特定されたアレル相互作用変数であり、 θ_2 、 θ_3 は、HLAアレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

【0342】

別の例として、ネットワーク変換関数 $g_h(\cdot)$ 、 $g_w(\cdot)$ を用いた、 $m = 4$ の異なる特定されたHLAアレルの中でHLAアレル $h = 2$ 、 $h = 3$ によってペプチド p^k が提示されるであろう尤度は、

$$u_k = f(NN_2(x_2^k; \theta_2)) + f(NN_3(x_3^k; \theta_3)) - f(NN_2(x_2^k; \theta_2)) \cdot f(NN_3(x_3^k; \theta_3))$$

によって生成することができ、式中、 $NN_2(\cdot)$ 、 $NN_3(\cdot)$ は、HLAアレル $h = 2$ 、 $h = 3$ について特定されたネットワークモデルであり、 θ_2 、 θ_3 は、HLAアレル $h = 2$ 、 $h = 3$ について決定されたパラメータのセットである。

10

【0343】

IX. 実施例 5：予測モジュール

予測モジュール 320 は、配列データを受け取って、提示モデルを用いて配列データ中の候補新生抗原を選択する。具体的には、配列データは、患者の腫瘍組織細胞から抽出されたDNA配列、RNA配列、及び/またはタンパク質配列であってよい。予測モジュール 320 は、配列データを、MHC-I については 8 ~ 15 個のアミノ酸を有する、または MHC-II については 6 ~ 30 個のアミノ酸を有する複数のペプチド配列 p^k に処理する。例えば、予測モジュール 320 は、所定の配列

IEFROEIFJEF

20

を、9 個のアミノ酸を有する 3 種類のペプチド配列

“IEFROEIFJ,” “EFROEIFJE,” 及び “FROEIFJEF”

に処理することができる。一実施形態では、予測モジュール 320 は、患者の正常組織細胞から抽出された配列データをその患者の腫瘍組織細胞から抽出された配列データと比較して 1 つ以上の変異を有する部分を特定することによって、変異したペプチド配列である候補新生抗原を特定することができる。

【0344】

予測モジュール 320 は、提示モデルの 1 つ以上を処理されたペプチド配列に適用してペプチド配列の提示尤度を推定する。具体的には、予測モジュール 320 は、提示モデルを候補新生抗原に適用することによって、腫瘍HLA分子上に提示される可能性が高い 1 つ以上の候補新生抗原ペプチド配列を選択することができる。一実現形態では、予測モジュール 320 は、あらかじめ決定された閾値を上回る推定提示尤度を有する候補新生抗原配列を選択する。別の実現形態では、提示モデルは、最も高い推定提示尤度を有する v 個の候補新生抗原配列を選択する (v は、一般的に、ワクチン中で送達することができるエピトープの最大数である)。所定の患者について選択された候補新生抗原を含むワクチンを患者に注射して免疫応答を誘導することができる。

30

【0345】

X. 実施例 6：患者選択モジュール

患者選択モジュール 324 は、患者が選択基準を満たすかどうかに基づいてワクチン治療及び/またはT細胞療法に対する患者のサブセットを選択する。一実施形態では、選択基準は、提示モデルによって生成される患者の新生抗原候補の提示尤度に基づいて決定される。選択基準を調整することにより、患者選択モジュール 324 は、患者の新生抗原候補の提示尤度に基づいてワクチン投与及び/またはT細胞療法を受ける患者数を調整することができる。具体的には、厳密な選択基準では、ワクチン及び/またはT細胞療法によって治療される患者の数はより少なくなるが、有効な治療(例えば、1 つ以上の腫瘍特異的新生抗原(TSNA)及び/または1 つ以上の新生抗原応答性T細胞)を受けるワクチン及び/またはT細胞療法による治療患者の比率は高くなり得る。これに対して、緩い選択基準では、ワクチン及び/またはT細胞療法で治療される患者の数はより多くなるが、有効な治療を受けるワクチン及び/またはT細胞療法による治療患者の比率は低くなり得る。患者選択モジュール 324 は、治療を受ける患者の目標比率と、有効な治療を受ける

40

50

患者の比率との間の所望のバランスに基づいて選択基準を変更する。

【0346】

いくつかの実施形態では、ワクチン治療を受ける患者を選択するための選択基準は、T細胞療法を受ける患者を選択するための選択基準と同じである。しかしながら、代替的な実施形態では、ワクチン治療を受ける患者を選択するための選択基準は、T細胞療法を受ける患者を選択するための選択基準と異なり得る。以下のセクションX・A及びX・Bでは、ワクチン治療を受ける患者を選択するための選択基準、及びT細胞療法を受ける患者を選択するための選択基準についてそれぞれ検討する。

【0347】

X・A・ワクチン治療を行う患者の選択

10

一実施形態では、患者に、ワクチン容量 v を有するその患者に対する個別化ワクチンに潜在的に含ませることが可能な v 種類の新生抗原候補の対応する治療サブセットが関連付けられる。一実施形態では、ある患者に対する治療サブセットは、提示モデルによって決定される最も高い提示尤度を有する新生抗原候補である。例えば、ワクチンが $v = 20$ 個のエピトープを含み得る場合、ワクチンは、提示モデルによって決定される最も高い提示尤度を有する各患者の治療サブセットを含み得る。しかしながら、他の実施形態では、ある患者に対する治療サブセットは、他の方法に基づいて決定することもできる点は認識される。例えば、ある患者に対する治療サブセットは、その患者に対する新生抗原候補のセットからランダムに選択することができ、または、ペプチド配列の結合親和性もしくは安定性をモデル化する従来技術のモデル、または提示モデルから得られる提示尤度及びこれらのペプチド配列に関する親和性または安定性情報を含む特定の因子の組み合わせの一部に基づいて決定することができる。

20

【0348】

一実施形態では、患者選択モジュール324は、患者の腫瘍変異負荷が最小の変異負荷に等しいかまたはそれよりも高い場合に患者が選択基準を満たすものと判定する。ある患者の腫瘍変異負荷 (TMB) は、腫瘍エクソームの非同義変異の総数を示す。一実施形態では、患者選択モジュール324は、患者のTMBの絶対数が所定の閾値に等しいかまたはそれよりも高い場合にワクチン治療を行う患者を選択する。別の実現形態では、患者選択モジュール324は、患者のTMBが患者のセットについて決定されたTMB間の閾値パーセンタイル内にある場合にワクチン治療を行う患者を選択する。

30

【0349】

別の実施形態では、患者選択モジュール324は、患者の治療サブセットに基づく患者の効用値スコアが最小の効用値スコアに等しいかまたはそれよりも高い場合に患者が選択基準を満たすものと判定する。一実施形態では、効用値スコアは、治療サブセットからの提示抗原の推定数の尺度である。

【0350】

提示抗原の推定数は、新生抗原の提示を1つ以上の確率分布のランダム変数としてモデル化することによって予測することができる。一実現形態では、患者 i の効用値スコアは、治療サブセットからの提示新生抗原候補の期待数、またはその特定の関数である。例として、各新生抗原の提示は、提示 (成功) の確率が新生抗原候補の提示尤度によって与えられるベルヌーイのランダム変数としてモデル化することができる。詳細には、それぞれが最も高い提示尤度 u_{i1} 、 u_{i2} 、 \dots 、 u_{iv} を有する v 種類の新生抗原候補 p^{i1} 、 p^{i2} 、 \dots 、 p^{iv} の治療サブセット S_i について、新生抗原候補 p^{ij} の提示は、ランダム変数 A_{ij} によって与えられ、ここで、

40

$$P(A_{ij} = 1) = u_{ij}, \quad P(A_{ij} = 0) = 1 - u_{ij} \quad (24)$$

提示される新生抗原の期待数は、各新生抗原候補の提示尤度の総和により与えられる。換言すれば、患者 i の効用値スコアは、下式として表される：

$$\text{util}_i(S_i) = \mathbb{E} \left[\sum_{j=1}^v A_{ij} \right] = \sum_{j=1}^v u_{ij} \quad (25)$$

50

患者選択モジュール324は、ワクチン治療について最小効用値に等しいかまたはそれよりも高い効用値スコアを有する患者のサブセットを選択する。

【0351】

別の実現形態では、患者*i*の効用値スコアは、少なくとも閾値数の新生抗原*k*が提示される確率である。1つの例では、新生抗原候補の治療サブセット*S_i*内の提示抗原の数は、提示（成功）の確率がエピトープのそれぞれの提示尤度によって与えられるポアソン二項ランダム変数としてモデル化される。詳細には、患者*i*の提示抗原の数は、ランダム変数*N_i*によって与えることができる：

$$N_i = \sum_{j=1}^v A_{ij} \sim \text{PBD}(u_{i1}, u_{i2}, \dots, u_{iv}) \quad (26)$$

10

式中、PBD(・)は、ポアソン二項分布を示す。少なくとも閾値数の新生抗原*k*が提示される確率は、提示抗原の数*N_i*が*k*に等しいかまたはそれよりも大きい確率の操作によって与えられる。換言すれば、患者*i*の効用値スコアは、下式として表される：

$$\text{util}_i(S_i) = \mathbb{P}[N_i \geq k] = \sum_{m=1}^k \mathbb{P}[N_i = m] \quad (27)$$

患者選択モジュール324は、ワクチン治療について最小効用値に等しいかまたはそれよりも高い効用値スコアを有する患者のサブセットを選択する。

【0352】

20

別の実現形態では、患者*i*の効用値スコアは、1つ以上の患者のHLAアレルに対して固定閾値（例えば500nM）よりも低い結合親和性または予測される結合親和性を有する新生抗原候補の治療サブセット*S_i*内の新生抗原の数である。1つの例では、固定閾値は、1000nM～10nMの範囲である。任意で、効用値スコアは、RNA-seqによって発現されたものとして検出された新生抗原のみをカウントしてもよい。

【0353】

別の実現形態では、患者*i*の効用値スコアは、その患者の1つ以上のHLAアレルに対する結合親和性とそのHLAアレルに対するランダムなペプチドの結合親和性の閾値パーセントイル以下である、新生抗原候補の治療サブセット*S_i*内の新生抗原の数である。1つの例では、閾値パーセントイルは、10パーセントイル～0.1パーセントイルの範囲である。任意で、効用値スコアは、RNA-seqによって発現されたものとして検出された新生抗原のみをカウントしてもよい。

30

【0354】

式(25)及び(27)に関して説明した効用値スコアの例はあくまで例示的なものに過ぎず、患者選択モジュール324は他の統計学または確率分布を用いて効用値スコアを生成することもできる点は認識されよう。

【0355】

X . B . T細胞療法を行う患者の選択

別の実施形態では、ワクチン治療を受けることに代えて、またはそれに加えて、患者はT細胞療法を受けることができる。ワクチン治療と同様、患者がT細胞療法を受ける実施形態では、患者を上記に述べたような*v*種類の新生抗原候補の対応する治療サブセットと関連付けることができる。この*v*種類の新生抗原候補の治療サブセットを、*v*種類の新生抗原候補のうち1つ以上に対する反応性を有する、患者由来のT細胞のインビトロでの特定に用いることができる。次に、これらの特定されたT細胞を増殖させて個別化T細胞療法において患者に注入することができる。

40

【0356】

2つの異なる時点においてT細胞療法を受ける患者を選択することができる。第1の時点は、患者に対する*v*種類の新生抗原候補の治療サブセットがモデルを用いて予測された後であるが、予測された*v*種類の新生抗原候補の治療サブセットに対して特異的であるT細胞のインビトロスクリーニングを行う前である。第2の時点は、予測された*v*種類の新

50

生抗原候補の治療サブセットに対して特異的であるT細胞のインビトロスクリーニングを行った後である。

【0357】

最初に、その患者に対するv種類の新生抗原候補の治療サブセットが予測された後で、かつ予測されたv種類の新生抗原候補の治療サブセットに対して特異的である患者由来のT細胞のインビトロでの特定を行う前に、T細胞療法を受ける患者を選択することができる。詳細には、患者由来の新生抗原特異的なT細胞のインビトロスクリーニングはコストが嵩み得るため、患者が新生抗原特異的T細胞を有する可能性が高い場合にのみ、新生抗原特異的T細胞についてスクリーニングする患者を選択することが望ましいと考えられる。インビトロT細胞スクリーニング工程の前に患者を選択するには、ワクチン治療を行う患者を選択するために用いられるものと同じ基準を用いることができる。詳細には、いくつかの実施形態において、患者選択モジュール324は、患者の腫瘍変異負荷が上記に述べたような最小の変異負荷に等しいかまたはそれよりも高い場合にT細胞療法を受ける患者を選択することができる。別の実施形態では、患者選択モジュール324は、その患者についてのv種類の新生抗原候補の治療サブセットに基づく患者の効用値スコアが、上記に述べたような最小の効用値スコアに等しいかまたはそれよりも高い場合にT細胞療法を受ける患者を選択することができる。

10

【0358】

第2に、予測されたv種類の新生抗原候補の治療サブセットに対して特異的である患者由来のT細胞のインビトロ特定を行う前にT細胞療法を受ける患者を選択することに加えて、またはそれに代えて、予測されたv種類の新生抗原候補の治療サブセットに対して特異的であるT細胞のインビトロ特定を行った後でT細胞療法を受ける患者を選択することもできる。詳細には、ある患者は、新生抗原認識についての患者のT細胞のインビトロスクリーニングにおいてその患者について新生抗原特異的なTCRの少なくとも閾値量が特定された場合にT細胞療法を受けるものとして選択することができる。例えば、ある患者は、少なくとも2種類の新生抗原特異的TCRがその患者について特定された場合にのみ、または2つの異なる新生抗原に対して新生抗原特異的なTCRが特定された場合にのみ、T細胞療法を受けるものとして選択することができる。

20

【0359】

別の実施形態では、ある患者は、その患者に対するv種類の新生抗原候補の治療サブセットの閾値量の新生抗原が患者のTCRによって認識された場合にのみ、T細胞療法を受けるものとして選択することができる。例えば、ある患者は、その患者に対するv種類の新生抗原候補の治療サブセットのうち少なくとも1種類の新生抗原が患者のTCRによって認識された場合にのみ、T細胞療法を受けるものとして選択することができる。さらなる実施形態では、ある患者は、その患者に対するTCRの少なくとも閾値量が、特定のHLA拘束性クラスの新生抗原ペプチドに対して新生抗原特異的なものとして特定された場合にのみ、T細胞療法を受けるものとして選択することができる。例えば、ある患者は、その患者に対する少なくとも1つのTCRが、新生抗原特異的なHLAクラスI制限新生抗原ペプチドとして特定された場合にのみ、T細胞療法を受けるものとして選択することができる。

30

40

【0360】

いっそうさらなる実施形態では、ある患者は、特定のHLA拘束性クラスの新生抗原ペプチドの少なくとも閾値量が患者のTCRによって認識された場合にのみ、T細胞療法を受けるものとして選択することができる。例えば、ある患者は、少なくとも1つのHLAクラスI制限新生抗原ペプチドがその患者のTCRによって認識された場合にのみ、T細胞療法を受けるものとして選択することができる。別の例では、ある患者は、少なくとも2つのHLAクラスII制限新生抗原ペプチドがその患者のTCRによって認識された場合にのみ、T細胞療法を受けるものとして選択することができる。上記の基準の任意の組み合わせを、その患者について予測されたv種類の新生抗原候補の治療サブセットに対して特異的であるT細胞をインビトロで特定した後でT細胞療法を受ける患者を選択するた

50

めに用いることもできる。

【0361】

X I . 実施例 7 : 例示的な患者選択性能を示す実験結果

セクション X で述べた患者選択の妥当性を、質量分析データにおいてシミュレートした新生抗原のサブセットが提示されていることが分かっている、シミュレートした新生抗原候補の試験セットがそれぞれに関連付けられたシミュレートした患者のセットで患者の選択を行うことにより検証する。詳細には、試験セット内のそれぞれのシミュレートした新生抗原候補に、その新生抗原がパスナーニ - スターンバーグデータセット (データセット「D1」) (データは、www.ebi.ac.uk/pride/archive/projects/PXD0000394にみること

10

【0362】

同じ H L A アレルについてのアレル毎提示モデルを、I E D B データセット (データセット「D2」) (データは、http://www.iedb.org/doc/mhc_ligand_full.zipにみること

20

【0363】

以下の例に図 1 3 A ~ 1 3 E を参照して開示するように、ペプチド結合の予測について訓練された提示モデル及び従来技術のモデルなどの異なるモデルを、それぞれのシミュレートした患者に対する新生抗原候補の試験セットに適用することによって予測に基づき患者に対する異なる治療サブセットを特定する。ワクチン治療について選択基準を満たす患者を選択し、患者の治療サブセットにエピトープを含む個別化ワクチンに関連付ける。治療サブセットのサイズは、異なるワクチン容量に応じて異なる。提示モデルを訓練するために用いられる訓練セットとシミュレートした新生抗原候補の試験セットとの間の重複は導入されない。

30

【0364】

以下の例では、ワクチンに含まれるエピトープ間で少なくとも特定の数の提示新生抗原を有する選択された患者の比率を分析する。この統計は、患者に免疫応答を誘発する潜在的な新生抗原を送達するうえでのシミュレートしたワクチンの有効性を示すものである。詳細には、ある試験セット内のシミュレートした新生抗原は、その新生抗原が質量分析データセット D 2 において提示されている場合に提示される。提示された新生抗原を有する患者の高い比率は、免疫応答を誘導することによって新生抗原ワクチンによる治療の奏功の可能性を示す。

40

【0365】

X I . A . 実施例 7 A : N S C L C がん患者における変異負荷の度数分布

図 1 3 A は、N S C L C 患者における変異負荷の標本度数分布を示す。N S C L C を含む異なる腫瘍タイプにおける変異負荷及び変異は、例えば、がんゲノムアトラス (the cancer genome atlas) (TCGA) (<https://cancergenome.nih>

50

.gov) にみることができる。X 軸は各患者の非同義変異の数を表し、Y 軸は特定の数の非同義変異を有する標本患者の比率を表す。図 1 3 A の標本度数分布は、3 ~ 1 7 8 6 個の変異の範囲を示し、患者の 3 0 % は 1 0 0 個よりも少ない変異を有している。図 1 3 A には示されていないが、変異負荷は非喫煙者と比較して喫煙者でより高く、変異負荷が患者における新生抗原負荷の強力な指標となり得ることが研究によって示されている。

【 0 3 6 6 】

上記のセクション X I の冒頭で導入したように、シミュレートした患者の数のそれぞれに、新生抗原候補の試験セットが関連付けられる。各患者の試験セットは、各患者について図 1 3 A に示される度数分布から変異負荷 m_i をサンプリングすることによって生成される。各変異について、ヒトプロテオーム由来の 2 1 マーのペプチド配列を、シミュレートする変異配列を表すようにランダムに選択する。新生抗原候補配列の試験セットを、2 1 マー内の変異にわたった各 (8、9、1 0、1 1) マーのペプチド配列を特定することにより患者 i について生成する。各新生抗原候補に、新生抗原候補配列が質量分析 D 1 データセット内に存在するかどうかを示すラベルを関連付ける。例えば、データセット D 1 内に存在する新生抗原候補配列にはラベル「1」を関連付け、データセット D 1 内に存在しない配列にはラベル「0」を関連付けることができる。以下でより詳細に述べるように、図 1 3 B ~ 1 3 E は、試験セット内の患者の提示新生抗原に基づいた患者選択の実験結果を示している。

10

【 0 3 6 7 】

X I . B . 実施例 7 B : 変異負荷の選択基準に基づく新生抗原提示を有する選択された患者の比率

20

図 1 3 B は、患者が最小変異負荷を満たすかどうかの選択基準に基づいて選択された患者に対してシミュレートしたワクチン中の提示新生抗原の数を示す。対応する試験において少なくとも特定の数の提示新生抗原を有する選択された患者の比率を特定する。

【 0 3 6 8 】

図 1 3 B において、x 軸は、「最小数の変異」のラベルで示される、腫瘍変異負荷に基づいたワクチン治療から除外される患者の比率を示す。例えば、「最小数の変異」2 0 0 におけるデータポイントは、患者選択モジュール 3 2 4 が、少なくとも変異が 2 0 0 個の変異負荷を有するシミュレートした患者のサブセットのみを選択したことを示す。別の例として、「最小数の変異」3 0 0 におけるデータポイントは、患者選択モジュール 3 2 4 が、少なくとも 3 0 0 個の変異を有するより低い比率のシミュレートした患者を選択したことを示す。y 軸は、ワクチン容量 v を有さない試験セット内の少なくとも特定の数の提示された新生抗原が関連付けられた選択された患者の比率を示す。詳細には、上のプロットは、少なくとも 1 種類の新生抗原を提示する選択された患者の比率を示し、中間のプロットは、少なくとも 2 個の抗原を提示する選択された患者の比率を示し、下のプロットは、少なくとも 3 個の抗原を提示する選択された患者の比率を示す。

30

【 0 3 6 9 】

図 1 3 B に示されるように、提示された新生抗原を有する患者の比率は、変異負荷が高くなるほど顕著に増大している。これは、選択基準としての変異負荷が、新生抗原ワクチンが効果的な免疫反応を誘導する可能性の高い患者を選択するうえで効果的であり得ることを示している。

40

【 0 3 7 0 】

X I . C . 実施例 7 C : 提示モデル 対 従来技術モデルによって特定されたワクチンにおける新生抗原提示の比較

図 1 3 C は、提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、従来技術のモデルによって特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。左側のプロットは、限定的なワクチン容量として $v = 1 0$ を仮定しており、右側のプロットは限定的なワクチン容量として $v = 2 0$ を仮定している。患者は、提示された新生抗原の期待数を示す効用値スコアに基づいて選択される。

50

【0371】

図13Cにおいて、実線は、アレルHLA-A*02:01及びHLA-B*07:02に対する提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた患者を示す。各患者に対する治療サブセットは、試験セット内の配列に提示モデルのそれぞれを適用し、最も高い提示尤度を有する v 種類の新生抗原候補を特定することによって特定される。点線は、単一アレルHLA-A*02:01に対する従来技術のモデルNETMHCpanに基づいて特定された治療サブセットを含むワクチンに関連付けられた患者を示す。NETMHCpanについての実施の詳細は、<http://www.cbs.dtu.dk/services/NetMHCpan>に示されている。各患者に対する治療サブセットは、試験セット内の配列にNETMHCpanモデルを適用し、最も高い推定される結合親和性を有する v 種類の新生抗原候補を特定することによって特定される。両グラフの x 軸は、提示モデルに基づいて特定された治療サブセット内の提示新生抗原の期待数を示す期待効用値スコアに基づいてワクチン治療から除外された患者の比率を示す。期待効用値スコアは、セクションXにおいて式(25)に関連して述べたようにして決定される。 y 軸は、ワクチンに含まれる少なくとも特定の数の新生抗原(1、2、または3種類の新生抗原)を提示する選択された患者の比率を示す。

10

【0372】

図13Cに示されるように、提示モデルに基づいた治療サブセットを含むワクチンに関連付けられた患者には、従来技術のモデルに基づいた治療サブセットを含むワクチンに関連付けられた患者よりも有意に高い割合で提示新生抗原を含むワクチンが投与される。例えば、右側のグラフに示されるように、従来技術のモデルに基づいたワクチンに関連付けられた選択された患者のわずか40%と比較して、提示モデルに基づいたワクチンに関連付けられた選択された患者の80%に、ワクチン中で少なくとも1つの提示新生抗原が投与される。これらの結果は、本明細書に述べられる提示モデルは、腫瘍を治療するための免疫反応を誘発する可能性の高いワクチンの新生抗原候補を選択するうえで効果的であることを示している。

20

【0373】

XI.D. 実施例7D: 提示モデルにより特定されたワクチンの新生抗原提示に対するHLAカバレッジの影響

図13Dは、HLA-A*02:01についての単一アレル毎提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、HLA-A*02:01及びHLA-B*07:02についてのアレル毎提示モデルの両方に基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。ワクチン容量は、 $v=20$ 個のエピトープに設定する。各実験について、異なる治療サブセットに基づいて決定された期待効用値スコアに基づいて患者を選択する。

30

【0374】

図13Dにおいて、実線は、HLAアレルHLA-A*02:01及びHLA-B*07:02に対する提示モデルの両方に基づいた治療サブセットを含むワクチンに関連付けられた患者を示す。各患者に対する治療サブセットは、試験セット内の配列に提示モデルのそれぞれを適用し、最も高い提示尤度を有する v 種類の新生抗原候補を特定することによって特定される。点線は、HLAアレルHLA-A*02:01に対する単一提示モデルに基づいた治療サブセットを含むワクチンに関連付けられた患者を示す。各患者に対する治療サブセットは、試験セット内の配列に単一HLAアレルのみについての提示モデルを適用し、最も高い提示尤度を有する v 種類の新生抗原候補を特定することによって特定される。実線のプロットでは、 x 軸は、両方の提示モデルにより特定された治療サブセットに対する期待効用値スコアに基づいてワクチン治療から除外された患者の比率を示す。点線のプロットでは、 x 軸は、単一の提示モデルにより特定された治療サブセットに対する期待効用値スコアに基づいてワクチン治療から除外された患者の比率を示す。 y 軸は、少なくとも特定の数の新生抗原(1、2、または3種類の新生抗原)を提示する選択され

40

50

た患者の比率を示す。

【0375】

図13Dに示されるように、両方のHLAアレルに対する提示モデルより特定された治療サブセットを含むワクチンに関連付けられた患者は、単一の提示モデルにより特定された治療サブセットを含むワクチンに関連付けられた患者よりも有意に高い割合で新生抗原を提示する。これらの結果は、高いHLAアレルカバレッジを有する提示モデルを確立することの重要性を示すものである。

【0376】

XI.E. 実施例7E：変異負荷 対 提示新生抗原期待数によって選択された患者における新生抗原提示の比較

図13Eは、変異負荷に基づいて選択された患者と、期待効用値スコアにより選択された患者との間でシミュレートしたワクチン中の提示新生抗原の数を比較したものである。期待効用値スコアは、 $v = 20$ 個のエピトープのサイズを有する提示モデルにより特定された治療サブセットに基づいて決定する。

【0377】

図13Eにおいて、実線は、提示モデルにより特定された治療サブセットを含むワクチンに関連付けられた期待効用値スコアに基づいて選択された患者を示す。各患者に対する治療サブセットは、試験セット内の配列に提示モデルのそれぞれを適用し、最も高い提示尤度を有する $v = 20$ 種類の新生抗原候補を特定することによって特定される。治療効用値スコアは、セクションXで式(25)に基づいて特定された治療サブセットの提示尤度に基づいて決定される。点線は、提示モデルにより特定された治療サブセットを含むワクチンに関連付けられた変異負荷に基づいて選択された患者を示す。x軸は、実線のプロットの期待効用値スコアに基づいてワクチン治療から除外された患者の比率、及び点線のプロットの変異負荷に基づいて除外された患者の比率を示す。y軸は、少なくとも特定の数の提示新生抗原(1、2、または3種類の新生抗原)を含むワクチンが投与される選択された患者の比率を示す。

【0378】

図13Eに示されるように、期待効用値スコアに基づいて選択された患者には、変異負荷に基づいて選択された患者よりも高い割合で提示新生抗原を含むワクチンが投与される。しかしながら、変異負荷に基づいて選択された患者には、選択されない患者よりも高い割合で提示新生抗原を含むワクチンが投与される。したがって、変異負荷は効果的な新生抗原ワクチン遅漏における効果的な患者選択基準であるが、期待効用値スコアはより効果的である。

【0379】

XII. 実施例8：除外(Held-Out)質量分析データに対する質量分析訓練モデルの評価

腫瘍細胞によるHLAペプチド提示は抗腫瘍免疫における主要な必要条件であるため^{91, 96, 97}、ペアリングしたクラスI HLAペプチド配列、HLA型、及びトランスクリプトームRNA-seq(方法)を含むヒト腫瘍及び正常組織試料の大規模な($N = 74$ 人の患者)統合データセットを、ヒトのがんにおける抗原提示を予測するためにこれらのデータ及び公表されているデータ^{92, 98, 99}を用いて新規なディープラーニングモデルを訓練する¹⁰⁰ことを目的として生成した。免疫療法の開発用の対象とするいくつかの腫瘍タイプの中から組織の入手しやすさに基づいて試料を選択した。質量分析によって、試料当たり平均で3704種のペプチドが、0.1未満のペプチドレベルFDRで特定された(344~11301種の範囲)。これらのペプチドは、アミノ酸8~15個の長さであり、モーダル長さ9(ペプチドの56%)の特徴的なクラスI HLA長の分布に従った。従来報告と一致して、ペプチドの大部分(中央値79%)が、MHCflurryにより、標準的な500nMの親和性閾値で少なくとも1つの患者のHLAアレルに結合することが予測された⁹⁰が、試料間でかなりのばらつきがみられた(例えば、1つの試料ではペプチドの33%で予測された親和性は500nM超であった)。一

10

20

30

40

50

一般的に用いられる¹⁰¹「強い結合物質」の閾値である50 nMは、提示されたペプチドのわずか42%の中央値しか捕捉しなかった。トランスクリプトームシーケンシングでは、試料当たり平均で131Mの固有のリードが得られ、遺伝子の68%が少なくとも1つの試料で少なくとも1TPM (transcript per million) のレベルで発現され、これは、最大数の遺伝子の発現が観察されるように設定された大規模かつ多様な試料の価値を強調するものである。HLAによるペプチド提示は、mRNA発現と強い相関があった。RNA発現または配列のみにおける差異によって説明されるよりも大きな、顕著でかつ再現可能なペプチド提示の割合の差異が遺伝子間で観察された。観察されたHLA型は、主としてヨーロッパ系の祖先を有する患者群由来の試料についての予想と一致した。

10

【0380】

これらの、及び公表されているHLAペプチドデータ^{92, 98, 99}を用いて、HLA抗原提示を予測するためにニューラルネットワーク(NN)モデルを訓練した。腫瘍の質量分析データから、各ペプチドが6つのHLAアレルのいずれか1つによって提示されている可能性のあるアレル特異的モデルを学習するため、アレルペプチドマッピング及びアレル特異的提示モチーフを一緒に学習することが可能な新規なネットワークアーキテクチャ(方法)を開発した。各患者について、陽性として示されたデータポイントは、質量分析によって検出されたペプチドであり、陰性として示されたデータポイントは、その試料中で質量分析によって検出されなかった、参照プロテオーム(SwissProt)からのペプチドとした。データを、訓練、検証、及び試験セットに分けた(方法)。訓練セットは、101個の試料(69個の本研究で新たに記載されるもの、及び32個の従来公表されているもの)から得られた142,844種のHLA提示ペプチド(FDR<約0.02)で構成されるものとした。検証セット(早期終了に用いた)は、同じ101個の試料からの18,004種の提示ペプチドで構成されるものとした。以下の2つの質量分析データセットを試験に用いた。すなわち、(1)訓練データから外された5つのさらなる腫瘍試料(2個の肺、2個の結腸、1個の卵巣)から得た571種の提示ペプチドで構成される腫瘍試料試験セット、及び(2)訓練データに含まれる単一アレルペプチドの位置に隣接した(ただし、異なる)ゲノム位置ウインドウ(ブロック)からの2,128種の提示ペプチドで構成される単一アレル細胞株試験セット(訓練/試験分割に関するさらなる詳細は「方法」を参照)。

20

30

【0381】

訓練データは、53種のHLAアレルについて予測モデルを特定した。従来の研究^{92, 104}と異なり、これらのモデルは、複数の長さのペプチドにおける各配列位置に対するHLA提示の依存度を捕捉した。このモデルは、遺伝子RNA発現及び遺伝子特異的提示傾向に対する臨界依存度を適正に学習し、mRNA存在量及び学習された遺伝子当たりの提示の傾向を独立して組み合わせると、発現量が最低で最も提示されにくい遺伝子と発現量が最高で最も提示されやすい遺伝子との間の提示の割合に最大約60倍の差を生じた。このモデルは、予測された結合親和性の制御後であっても、IEDBにおけるHLA/ペプチド複合体の測定される安定性⁸⁸を予測する(10個のアレルに対して $p < 1e-10$)ことがさらに観察された(試験した10個のアレルのうち8個について $p < 0.05$)。これらをまとめると、これらの特性は、免疫原性のHLAクラスIペプチドの改善された予想の基礎をなすものである。

40

【0382】

除外質量分析試験セットに対するHLA提示の予測ツールとしてのこのNNモデルの性能を評価し、インビトロのHLA結合データで訓練したニューラルネットワークツールである最新の結合親和性予測ツールMHCF1urry⁹⁰(バージョン1.2.0, 方法)と比較した。HLA提示におけるmRNAレベルの重要性が強調された従来の報告に基づき、RNA-seq^{81, 92, 103}によってアッセイされる遺伝子発現に対する増大する閾値を取り入れた。

【0383】

50

図14A～Dは、「完全MSモデル」、「ペプチドMSモデル」、及び3つの遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの予測性能を比較したものである。「完全MSモデル」及び「ペプチドMSモデル」はいずれも、上記に述べた質量分析データで訓練されたニューラルネットワークモデルである。しかしながら、「完全MSモデル」が試料のすべての特性に基づいて訓練及び試験されるのに対して、「ペプチドMSモデル」は試料のHLA型及びペプチド配列のみに基づいて訓練及び試験される。TPM>0の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデル、TPM>1の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデル、及びTPM>2の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルのMHCFlurry1.2.0結合親和性モデルの3つの異なるバージョンを試験する。「ペプチドMSモデル」及びTPM>1の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルは、いずれも試料のHLA型及びペプチド配列のみに基づいて訓練及び試験され、また、いずれも同じRNA発現閾値を有することから、これら2つのモデルの性能を直接比較することで、結合親和性訓練データに対して質量分析データから学習されるペプチドモチーフの差に起因する予測の改善が定量化される。

10

20

30

40

50

【0384】

最初に図14Aを参照すると、図14Aは、各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%における陽性適中率(PPV)を比較したものである(方法)。図14Aはまた、5つの試験試料について「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%における平均PPVも示している。図14Aに示されるように、再現率40%における平均PPVは、「完全MSモデル」で0.54、TPM>2、1、及び0の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルでそれぞれ0.076、0.072、及び0.061であった。「完全MSモデル」とTPM>0の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルとの間の比較はいずれも、 $p < 1e-6$ で統計的に有意である。

【0385】

次に図14Bを参照すると、図14Bは、各試験試料が1:10,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む15種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%におけるPPVを比較したものである(方法)。図14Bはまた、15種類の試験試料について「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%における平均PPVも示している。図14Bに示されるように、再現率40%における平均PPVは、「完全MSモデル」で0.37、TPM>2、1、及び0の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルでそれぞれ0.094、0.090、及び0.071であった。「完全MSモデル」とTPM>0の遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルとの間の比較はいずれも、 $p = 1.6e-4$ であるHLA-A*01:01を含む試験試料を除いて、 $p < 1e-6$ で統計的に有意である。

【0386】

図16は、各モデルを図14Aに関して上記に述べた試験セットで試験する場合の「完全MSモデル」及び「アンカー残基のみのMSモデル」の再現率40%における陽性適中率(PPV)を比較したものである(方法)。図16はまた、5種類の試験試料について

「完全MSモデル」及び「アンカー残基のみのMSモデル」の再現率40%における平均のPPVも示している。「完全MSモデル」と同様、「アンカー残基のみのMSモデル」は、上記に述べた質量分析で訓練したニューラルネットワークモデルである。しかしながら、「アンカー残基のみのMSモデル」は試料中の全ペプチド配列に基づいて訓練及び試験される代わりに、「アンカー残基のみのMSモデル」は、試料のペプチド配列の「アンカー」残基（第1、第2、及び最後の残基）のみに基づいて訓練及び試験される。したがって、図16に示される結果は、モデルの予測性能に対するアンカー残基及び非アンカー残基の相対的な重要度を定量化している。図16に示されるように、「アンカー残基のみのMSモデル」の性能は、完全MSモデルと比較して大幅に低くなっている。アンカー残基のみのMSモデルにおける再現率40%での平均のPPVは、完全MSモデルにおける0.50と比較して0.13である。したがって、ペプチド配列の非アンカー配列を含むモデルの訓練及び試験は、モデルの予測改善につながると推論することができる。

10

【0387】

図17Aは、各モデルを図14Aからの試験試料0で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す(方法)。図17Aに示されるように、「完全MSモデル」及び「ペプチドMSモデル」は、TPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルと比較してより良好な性能を実現している。

20

【0388】

図17Bは、各試験試料が1:5,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む15種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%におけるPPVを比較したものである(方法)。図17Bはまた、15種類の試験試料について「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの再現率40%における平均PPVも示している。図14Bの結果(各試験試料は、1:10,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む)を図17Aの結果(各試験試料は、1:5,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む)と比較することにより、ペプチド提示の存在率は絶対PPVと強く相関していると推論することができる。一般に、予測される事象(例えば、提示)の存在率が低いほど、高いPPV予測を得ることはより困難となる。したがって、試験データにおける存在率を低下させる(上昇させる)ことで、すべてのモデルの絶対PPVは低下する(上昇する)。しかしながら、異なるモデルのPPV間の相対的な差は、予想される試験セット存在率の変化によって影響されない。

30

【0389】

図17C~Gは、各モデルを図14Aからの試験試料0~4で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す(方法)。

40

【0390】

図17H~Vは、各試験試料が1:10,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む図14Bの15種類の異なる試験試料で各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す(方法)。

【0391】

50

図18は、各モデルを図14Aの5種類の異なる試験試料で試験する場合のヒト腫瘍におけるHLA提示ペプチドをモデル化するためのMSモデルの異なるバージョン及び初期のアプローチ¹⁰⁴の再現率40%における陽性適中率(PPV)を比較したものである(方法)。図18はまた、5種類の試験試料について各モデルの再現率40%における平均のPPVも示している。図18で試験したモデルには、「完全MSモデル」、「MSモデル、フランキング配列なし」、「MSモデル、フランキング配列または遺伝子毎係数なし」、「ペプチドのみのMSモデル、すべての長さを一緒に訓練」、「ペプチドのみのMSモデル、すべての長さを別々に訓練」、「線形のペプチドのみのMSモデル」、「MixMHCPred 1.1」モデル、及び「結合親和性」モデルが含まれる。「完全MSモデル」、「MSモデル、フランキング配列なし」、「MSモデル、フランキング配列または遺伝子毎係数なし」、「ペプチドのみのMSモデル、すべての長さを一緒に訓練」、「ペプチドのみのMSモデル、すべての長さを別々に訓練」、及び「線形のペプチドのみのMSモデル」は、いずれも上記に述べた質量分析で訓練されたニューラルネットワークモデルである。しかしながら、各モデルは試料の異なる特性を用いて訓練及び試験される。「MixMHCPred 1.1」モデル及び「結合親和性」モデルは、HLA提示ペプチドをモデル化するための初期のアプローチである¹⁰⁴。

10

20

30

40

50

【0392】

全体として、NNモデルはHLAペプチド提示の有意に改善された予測を実現し、PPVは、腫瘍試験セットで標準的な結合親和性+遺伝子発現よりも最大で9倍高く(図14A)、単一アレルデータセットで最大で5倍高かった(図14B)。MSに基づいたNNモデルの大きなPPVの利点は、異なる再現率にわたって保たれ(図17A)、統計的に有意であった(図14A及び14Bのすべての腫瘍試料及び $p = 1.6e-4$ であるHLA-A*01:01を除く単一アレル試料において $p < 10^{-6}$)。HLAペプチド提示についての標準的な結合親和性+遺伝子発現の陽性適中率は、6%と低い値に達し、従来の推定値と一致した^{87, 93}。しかしながら、わずかな割合のペプチドしか提示されているものとして検出されないことから(例えば、腫瘍MS試験データセットでは2500のうち約1)、この約6%のPPVは、ベースライン存在率と比較して100倍超の改善を表す点に留意されたい。

【0393】

入力としてHLA型及びペプチド配列のみを用いた質量分析データで訓練された縮小モデル(「ペプチドMSモデル」、図14A~B、方法を参照)を完全MSモデルと比較することにより、結合親和性の予測に対するPPVの増加の約30%は、質量分析によって捕捉することができるが結合親和性アッセイによっては捕捉されないペプチドの外因性特性(RNA存在量、フランキング配列、遺伝子毎係数)のモデル化に起因するものと判断された(図14A~B、図17A及び18も参照されたい)。増加の残りの70%は、ペプチド配列の改善されたモデル化に起因した(図14A~B)。このモデル化は、ヒト腫瘍におけるHLA提示ペプチドのモデル化における初期のアプローチ¹⁰⁴よりも高い性能を示したことから、性能の改善に寄与したのは、訓練データセット(HLA提示ペプチド)の性質だけではなく、全体のモデルアーキテクチャであった(図18)。この新たなモデルアーキテクチャにより、結合親和性予測またはハードクラスタリングアプローチ^{104~106}を用いた、想定される提示アレルに対するペプチドの事前の割り当てを必要としない、エンド・ツー・エンド訓練プロセスによるアレル特異的モデルの学習が可能となった。重要な点として、新たなモデルアーキテクチャはまた、線形性などの逆畳み込みの前提条件としてアレル特異的サブモデルに精度を低下させる制約を課すことがなく、各ペプチド長を別に考慮する¹⁰⁴必要もなかった。完全モデルは、これらの制約を課すいくつかの単純化されたモデル及び従来、公表されているアプローチの性能を上回った(図18)。

【0394】

図18は、MS試験セットでのいくつかの単純化されたモデルの性能を示す。完全モデルに組み込まれるモデル化の改良点の相対的重要度を、モデル化の改良点を一度に一つず

つ除去してMS試験セットで予測性能を試験することによって定量化する。さらに、質量分析から溶出されるペプチドをモデル化する近年公表されているアプローチ(MixMHC Pred)と本明細書に開示される提示モデルとの比較を行った。MixMHC Predは、9個及び10個以外の長さのペプチドは現時点でモデル化していないため、9マー及び10マーのみを比較で用いた。各モデルは(左から右の順に)、「完全MSモデル」(「方法」に記載される完全NNモデル)、「MSモデル、フランキンク配列なし」(フランキンク配列特性を除いた点以外は完全NNモデルと同じ)、「MSモデル、フランキンク配列または遺伝子毎係数なし」(フランキンク配列特性及び遺伝子毎係数を除いた点以外は完全NNモデルと同じ)、「ペプチドのみのMSモデル、すべての長さを一緒に訓練」(使用した特性はペプチド配列及びHLA型のみである点以外は完全NNモデルと同じ)、「ペプチドのみのMSモデル、それぞれの長さを別々に訓練」(このモデルでは、9マー及び10マーについて別々のモデルを訓練した点以外、モデル構造は、ペプチドのみのMSモデルと同じ)、「線形のペプチドのみのMSモデル(アンサンプリングを伴う)」(ニューラルネットワークを用いてペプチド配列をモデル化する代わりに、完全モデルで使用され、「方法」に記載されるものと同じ最適化手順を用いて訓練された線形モデルのアンサンブルを用いた点以外は、それぞれのペプチド長さを別々に訓練したペプチドのみのMSモデルと同じ)；「MixMHC Pred 1.1」は、デフォルト設定のMixMHC Predである；「結合親和性」は、本文におけるものと同じMHCflurry 1.2.0である。最後の5つのモデル(「ペプチドのみのMSモデル、すべての長さを一緒に訓練」~「結合親和性」)は、ペプチド配列及びHLA型のみと同じ入力を有する。詳細には、最後の5つのモデルのいずれも予想を行ううえでRNA存在量を用いていない。最も性能の良いペプチドのみのモデル(「ペプチドのみのMSモデル、すべての長さを一緒に訓練」)が再現率40%において0.41の平均PPVを与えたのに対して、最も性能の低い質量分析データで訓練されたペプチドのみのモデル(「線形のペプチドのみのMSモデル(アンサンプリングを伴う)」)では、平均PPVはわずか28%であり(MixMHC predの18%の平均PPVよりもごくわずかに高かった)、ペプチド配列の改善されたNNモデル化の値を強調するものである。MixMHC predは、線形のペプチドのみのMSモデルとは異なるデータで訓練されるが、同じモデル化特性の多くを有する(例えば、それぞれのペプチド長が別々に訓練される線形モデルである)点に留意されたい。

【0395】

XIII. 実施例9：レトロスペクティブな新生抗原T細胞データアプリケーションのモデル評価

本発明者らは、HLAペプチド提示のこのような正確な予測が、ヒト腫瘍CD8 T細胞のエピトープ(すなわち、免疫療法の標的)を特定する能力につながるかどうかを評価した。この評価を行うための適当な試験データセットには、T細胞によって認識され、かつ腫瘍細胞表面上のHLAによって提示されるペプチドが含まれる。さらに、正式な性能評価は、陽性として示される(すなわち、T細胞に認識された)ペプチドだけでなく、十分な数の陰性標識された(すなわち、試験したが、認識されなかった)ペプチドも必要とする。質量分析データセットは、腫瘍提示に対応するがT細胞認識には対応せず、逆に、ワクチン接種後のプライミングまたはT細胞アッセイはT細胞前駆細胞の存在及びT細胞認識に対応するが腫瘍提示には対応しない。例えば、ソース遺伝子の腫瘍内での発現が低レベルである強力なHLA結合ペプチドは、こうしたペプチドは腫瘍によって提示されないため、治療上の有用性のない強い免疫後CD8 T細胞応答を生じ得る。

【0396】

適当なデータセットを得るため、公表されているCD8 T細胞エピトープを、必要な基準を満たす以下の4つの最近の研究から集めた。すなわち、研究A⁹⁶は、消化管腫瘍を有する9人の患者でTILを調べ、自家樹状細胞(DC)中でタンデムミニ遺伝子(TMG)法を用いてIFN-ELISPOTにより試験した、1,053個の体細胞SNV変異のうち12個のT細胞による認識を報告している。研究B¹⁰⁷もTMGを用い、

4人の黒色腫患者から得たCD8+PD-1+循環リンパ球による574個のSNVのうち6個のT細胞による認識を報告している。研究C⁹⁷は、パルス化ペプチド刺激を用いて3人の黒色腫患者から得たTILを評価し、381個の試験したSNV変異のうち5個に対する応答を見出している。研究D¹⁰⁸は、TMGアッセイの組み合わせを用い、最小エピトープペプチドでパルスして1人の乳がん患者から得たTILを評価し、62個のSNVのうち2個の認識を報告している。組み合わせたデータセットは、既存のT細胞反応を示す26種類の新生抗原を含む、17人の患者から得た2,009個のアッセイされたSNVからなるものとした。重要な点として、データセットは大部分が腫瘍浸潤リンパ球による新生抗原認識からなるため、効果的な予測は、従来文献^{81, 82, 97}に記載されるように、T細胞をプライミングできる新生抗原だけでなく、より厳密には、腫瘍によってT細胞に提示される新生抗原を特定する能力を示唆する。

10

【0397】

個別化免疫療法用の抗原の選択をシミュレートするため、「完全MSモデル」、「ペプチドMSモデル」、ならびにTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルを用いて、提示の確率の順に体細胞変異をランク付けした。抗原特異的な免疫療法は、標的とされる特異性の数において技術的に制約されているため（例えば、現在の個別化ワクチンは約10~20個の体細胞変異をコードしている^{80~82}）、予測方法を、少なくとも1つの既存のT細胞応答を示すそれぞれの患者について上位5、10、または20位にランクされた体細胞変異における既存のT細胞応答の数をカウントすることにより比較した。これらの結果を図14Cに示す。詳細には、図14Cは、それぞれの試験試料が少なくとも1つの既存のT細胞応答を示す患者から採取された12種類の異なる試験試料からなる試験セットについて、「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルによって特定された、上位5、10、及び20位にランクされた体細胞変異について、T細胞（例えば、既存のT細胞応答）によって認識された体細胞変異の割合を比較したものである。「完全MSモデル」とTPM>0の遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルとの間の比較はいずれも、 $p = 0.056$ である上位5位にランクされた体細胞変異を除いて、 $p < 0.005$ で統計的に有意である。

20

【0398】

予想されたように、結合親和性の予測は、優先順位付けされた変異間の既存のT細胞応答のわずかな部分、例えば、TPM>0の上位20位にランクされた変異間の全部で26個のうち9個（35%）のみを含んだ（補足の表1）。これに対して、既存のT細胞応答の大半（19/26、73%）が完全MSモデルによって上位20位にランクされ、平均は異なるランク及び遺伝子発現閾値にわたって保たれた（図14C、補足の表1）。患者レベルでは、完全MSモデルは、TPM>0の結合親和性におけるわずか0.69（ $p = 1.4e-4$ ）と比較して、少なくとも1つの既存のT細胞応答を示す13人の患者の上位20位の予測された変異において平均で1.54の既存の新生抗原T細胞反応を示した。

30

【0399】

次いで、本発明者らは、細胞療法用のT細胞/TCRを特定するうえで有用であり得るように、最小新生エピトープのレベル（すなわち、変異と重なる8~11マーが認識された）での変異を評価した。換言すれば、「完全MSモデル」、「ペプチドMSモデル」、ならびにTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCF1urry1.2.0結合親和性モデルを用いて、提示の確率の順に最小新生エピトープをランク付けした。上記に述べたように、抗原特異的な免疫療法は、標的とされる特異性の数において技術的に制約されているため、予測方法を、少なくとも1つの既存のT細胞応答を示すそれぞれの患者について上位5、10、または20位にランクされた最小新生エピトープにおける既存のT細胞応答の数をカウントすることにより比較した。陽性として示されるエピトープは、ペプチドベースのアッセイ（TMGベースのアッセイの代わり、また

40

50

はそれに加えて)によって免疫原性最小エピトープであることが確認されたものであり、陰性の例は、ペプチドベースのアッセイで認識されないすべてのエピトープ及び認識されないミニ遺伝子に含まれるすべての変異にわたったエピトープである。これらの結果を図14Dに示す。

【0400】

詳細には、図14Dは、それぞれの試験試料が少なくとも1つの既存のT細胞応答を示す患者から採取された12種類の異なる試験試料からなる試験セットについて、「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルによって特定された、上位5、10、及び20位にランクされた最小新生エピトープについて、T細胞(例えば、既存のT細胞応答)によって認識された最小新生エピトープの割合を比較したものである。「完全MSモデル」とTPM > 0の遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルとの間の比較はいずれも、 $p = 0.082$ である上位5位にランクされた最小新生エピトープを除いて、 $p < 0.05$ で統計的に有意である。すべてのパネルで、エラーバーは90%信頼区間を表す。

10

【0401】

図14Dに示されるように、TPM > 0での結合親和性に対するNNモデルの利点が、図14Cよりもさらに顕著に認められ、少なくとも4倍多い新生エピトープが上位にランクされた最小エピトープに含まれていた。注目される点として、研究A、B、及びDでは強い結合親和性を有するペプチドのみが個々に試験されたことから、この比較は結合親和性の予測に有利にバイアスされている点である。これらの研究ではアッセイされなかったがこのモデルによれば選択されていたであろう弱い予測HLA結合親和性を有するT細胞認識ペプチドが存在する可能性がある。そのようなペプチドがこの研究で観察され、図15A及び補足の表3に関して下記に詳細に考察する。

20

【0402】

システインを含むペプチドの検出における質量分析の既知の制約^{92, 104}にもかかわらず、NNモデルは、システイン含有T細胞認識エピトープでの結合親和性予測の性能を上回り、TPM > 0の遺伝子発現閾値を有する結合親和性における上位5位の7個のうちの1個と比較して、上位5位の7個のシステイン含有エピトープのうち3個(43%)がランク付けされた。質量分析試験セットと同様、質量分析訓練データに基づいてモデル化することができるさらなる特性(RNA、ランキング配列、遺伝子毎係数)は、予測性能の増大に大きく寄与したが、質量分析試験データにおけるように、ペプチドのみのMSモデルの予測性能は結合親和性予測に対して大幅に改善され、この改善の大部分はペプチド配列の改善されたモデル化に起因するものであることが示された(図14C~D、水色のバーと緑色のバーとを比較)。

30

【0403】

注目される点として、この改善は、現在のTILアッセイの制約による試験セット新生エピトープの偽陰性(すなわち、T細胞によって認識されることが可能な腫瘍により提示される新生エピトープで、T細胞応答が検出されなかった場合)の潜在的な増大にもかかわらず観察された。これらの制約としては、(a)免疫抑制性の腫瘍微小環境及び非効率的なT細胞プライミング、(b)新生エピトープ応答性のT細胞枯渇、(c)TILによるIFN γ 以外のサイトカインの産生、及び(d)使用される腫瘍画分の不均質性を挙げることができる。したがって、本明細書に記載される上位5~20位の免疫原性ペプチドの数による絶対予測性能は、例えば強力な新生抗原がんワクチンの投与などの他の状況に対して悲観的なものである可能性がある。

40

【0404】

XIII.A. データ

本発明者らは、Gros et al⁸⁴、Tran et al¹⁴⁰、Stronen et al¹⁴¹及びZacharakis et alの補足情報から変異コール、HLA型、及びT細胞認識データを得た。患者特異的RNA-seqデータは入手で

50

きなかった。腫瘍のRNA発現は異なる患者で同じ腫瘍型と相関しているものと推論し、TCGAより得た腫瘍型一致患者からのRNA-seqデータを代用し、これを結合親和性予測の前にニューラルネットワーク予測及びTPM>1でのRNA発現フィルタリングの両方で用いた。腫瘍型の一致したRNA-seqデータの追加によって予測性能が改善された(図14C~D)。

【0405】

変異レベルの分析(図14C)では、Gros et al、Tran et al及びZacharakis et alにおける陽性として示されるデータポイントは、TMGアッセイまたは最小エピトープペプチドパルスアッセイの両方において患者T細胞によって認識される変異とした。陰性として示されるデータポイントは、TMGアッセイで試験した他のすべての変異とした。Stronen et alでは、陽性として示される変異は、少なくとも1つの認識されるペプチドがまたがる変異とし、陰性のデータポイントは、試験されたがテトラマーアッセイにおいて認識されなかったすべての変異とした。変異した25マーTMGアッセイは変異にまたがるすべてのペプチドのT細胞による認識を試験するため、Gros、Tran及びZacharakisのデータとしては、変異を、変異にまたがるすべてのペプチドにわたって提示の確率を総和するかまたは最小結合親和性を取ることによってランク付けした。Stronenのデータとしては、変異をテトラマーアッセイで試験した変異にまたがるすべてのペプチドにわたって提示の確率を総和するかまたは最小結合親和性を取るによってランク付けした。変異及び特性の完全なリストを補足の表1に示す。

10

20

【0406】

エピトープレベルの分析では、陽性として示されるデータポイントは、ペプチドパルスアッセイまたはテトラマーアッセイにおいて患者T細胞によって認識されたすべての最小エピトープとし、陰性データポイントは、ペプチドパルスアッセイまたはテトラマーアッセイにおいてT細胞によって認識されなかったすべての最小エピトープ、及び患者T細胞によって認識されなかった、試験したTMGからの変異にまたがるすべてのペプチドとした。Gros et al、Tran et al及びZacharakis et alの場合では、ペプチドパルスアッセイによって試験しなかった、TMG分析において認識された変異にまたがる最小エピトープペプチドは、これらのペプチドのT細胞認識状態は実験的に調べられなかったため、分析から除外した。

30

【0407】

XIV. 実施例10：がん患者における新生抗原応答性T細胞の特定

本実施例では、改善された予測が通常患者試料からの新生抗原の特定を可能とすることを実証する。これを行うため、アーカイブしたFFPE腫瘍生検及び5~30mlの末梢血を、抗PD(L)1療法を行っている転移性NSCLCを有する9人の患者で分析した(補足の表2：図15A~Cで調べたN=9人の患者についての患者人口統計及び治療情報。主なフィールドは、腫瘍ステージ及びサブタイプ、行った抗PD1療法、及びNGSの結果の概要を含む)。腫瘍全エクソームシーケンシング、腫瘍トランスクリプトームシーケンシング、及びマッチド正常エクソームシーケンシングにより、患者当たり、平均で198個の体細胞変異(SNV及び短いindel)が得られ、そのうち、平均で118個が発現された(「方法」、補足の表2)。完全MSモデルを適用して既存の抗腫瘍T細胞応答に対して試験を行うために患者当たり20個の新生エピトープを優先順位付けした。可能性の高いCD8応答に集中して分析を行うため、優先順位付けしたペプチドを8~11マーの最小エピトープとして合成し(「方法」、次いで末梢血単核細胞(PBMC)を合成したペプチドと短期インビトロ刺激(IVS)培養中で培養して新生抗原応答性T細胞を増殖させた(補足の表3)。2週間後、抗原特異的T細胞の存在を優先順位付けした新生エピトープに対するIFN-ELISPOTを用いて評価した。十分なPBMCが利用可能な7人の患者で別々の実験をさらに行って認識された特異的抗原の完全または部分的な置み込みを行った。これらの結果を図15A~C及び図19A~22に示す。

40

50

【0408】

図15Aは、9人の患者における患者特異的新生抗原ペプチドプールに対するT細胞応答の検出を示す。それぞれの患者について、予測された新生抗原を、モデルランク付け及び任意の配列ホモロジーにそれぞれ従って10個のペプチドの2つのプールに組み合わせた(相同ペプチドは異なるプールに分割した)。次いで、それぞれの患者について、その患者についてインビトロ増殖させたPBMCをIFN-ELISpotで2つの患者特異的新生抗原ペプチドプールにより刺激した。図15Aのデータは、バックグラウンド(対応するDMSO陰性対照)を引いた播種細胞 10^5 個当たりのスポット形成単位(SFU)として示す。バックグラウンド測定値(DMSO陰性対照)を図22に示す。患者1-038-001、1-050-001、1-001-002、CU04、1-024-001、1-024-002及びCU05について、同族のペプチドプール#1及び#2に対する単一のウェル(患者1-038-001、CU02、CU03及び1-050-001)または平均及び標準偏差を含むレプリケート(他のすべての患者)の応答を示す。患者CU02及びCU03では、細胞数のために、特異的ペプチドプール#1に対してのみ試験が可能であった。バックグラウンドよりも2倍超高い増加倍率の値を有する試料は陽性とみなし、星印で示した(応答性ドナーには、患者1-038-001、CU04、1-024-001、1-024-002、及びCU02が含まれる)。非応答性ドナーには、患者1-050-001、1-001-002、CU05、及びCU03が含まれる。図15Cは、IFN-ELISpotにおいて、DMSO陰性対照、PHA陽性対照、CU04特異的新生抗原ペプチドプール#1、CU04特異的ペプチド1、CU04特異的ペプチド6、及びCU04特異的ペプチド8で刺激した患者CU04由来のインビトロ増殖させたPBMCを含むELISpotウェルの写真を示す。

10

20

【0409】

図19A~Bは、HLAを一致させた健康なドナーにおいて患者新生抗原を用いた対照実験の結果を示す。これらの実験の結果は、インビトロ培養条件は、インビトロでのデノボプライミングを可能とするのではなく、既存のインビボプライミングされたメモリーT細胞のみを増殖させたことを示す。

【0410】

図20は、図15Aに示される各ドナー及び各インビトロ増殖についてPHA陽性対照に対するT細胞応答の検出を示す。図15Aの各ドナー及び各インビトロ増殖について、インビトロ増殖させた患者PBMCを最大のT細胞活性化となるようにPHAで刺激した。図20のデータは、バックグラウンド(対応するDMSO陰性対照)を引いた播種細胞 10^5 個当たりのスポット形成単位(SFU)として示す。単一ウェルまたは生物学的レプリケートの応答を、患者1-038-001、1-050-001、1-001-002、CU04、1-024-001、1-024-002、CU05及びCU03について示す。患者CU02ではPHAによる試験は行わなかった。ペプチドプール#1に対する陽性応答(図15A)が生存可能かつ機能性のT細胞を示したため、患者CU02由来の細胞を分析に含めた。図15Aに示されるように、ペプチドプールに応答性を示したドナーには、患者1-038-001、CU04、1-024-001、及び1-024-002が含まれる。やはり図15Aに示されるように、ペプチドプールに応答性を示さなかったドナーには、患者1-050-001、1-001-002、CU05、及びCU03が含まれる。

30

40

【0411】

図21Aは、患者CU04におけるプール#2中のそれぞれの個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。図21Aは、患者CU04におけるPHA陽性対照に対するT細胞応答の検出も示す。(この陽性対照データは図20にも示される。)患者CU04では、その患者のインビトロ増殖させたPBMCをIFN-ELISpotにおいて患者CU04に対するプール#2からの患者特異的な個々の新生抗原ペプチドで刺激した。患者のインビトロ増殖させたPBMCをIFN-ELISpotにおいて陽性対照としてのPHAでも刺激した。データは、バックグラウンド(対応するDM

50

S O 陰性対照) を引いた播種細胞 10^5 個当たりのスポット形成単位 (S F U) として示す。

【0412】

図21Bは、患者CU04の3回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的の新生抗原ペプチドに対するT細胞応答の検出を示す。両方の患者において、その患者のインビトロ増殖させたPBM CをIFN - E L I S p o tにおいて患者特異的な個々の新生抗原ペプチドで刺激した。各患者について、各来院のデータは、バックグラウンド(対応するDMSO対照)を引いた播種細胞 10^5 個当たりの累積(総和)スポット形成単位(SFU)として示す。患者CU04のデータは、バックグラウンドを引いた3回の来院の累積SFUとして示す。患者CU04では、バックグラウンドを引いたSFUを、最初の来院(T0)ならびに最初の来院(T0)の2ヶ月(T0+2ヶ月)及び14ヶ月(T0+14ヶ月)後のその後の来院について示す。患者1-024-002のデータは、バックグラウンドを引いた2回の来院の累積SFUとして示す。患者1-024-002では、バックグラウンドを引いたSFUを、最初の来院(T0)及び最初の来院(T0)の1ヶ月(T0+1ヶ月)後のその後の来院について示す。バックグラウンドよりも2倍超高い増加倍率の値を有する試料を陽性とみなし、星印で示した。

10

【0413】

図21Cは、患者CU04の2回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的の新生抗原ペプチドに対する、及び患者特異的の新生抗原ペプチドプールに対するT細胞応答の検出を示す。両方の患者において、その患者のインビトロ増殖させたPBM CをIFN - E L I S p o tにおいて患者特異的な個々の新生抗原ペプチド及び患者特異的な新生抗原ペプチドプールで刺激した。具体的には、患者CU04では、患者CU04のインビトロ増殖させたPBM CをIFN - E L I S p o tにおいてCU04特異的な個々の新生抗原ペプチド6及び8ならびにCU04特異的な新生抗原ペプチドプールで、患者1-024-002では、患者1-024-002のインビトロ増殖させたPBM CをIFN - E L I S p o tにおいて1-024-002特異的な個々の新生抗原ペプチド16及び1-024-002特異的な新生抗原ペプチドプールで刺激した。図21Cのデータは、平均及び範囲を有する各テクニカルレプリケートについて、バックグラウンド(対応するDMSO対照)を引いた播種細胞 10^5 個当たりのスポット形成単位(SFU)として示す。患者CU04のデータは、バックグラウンドを引いた2回の来院のSFUとして示す。患者CU04では、バックグラウンドを引いたSFUを、最初の来院(T0、テクニカルトリプレート)及び最初の来院(T0)の2ヶ月(T0+2ヶ月、テクニカルトリプレート)後のその後の来院について示す。患者1-024-002のデータは、バックグラウンドを引いた2回の来院のSFUとして示す。患者1-024-002では、バックグラウンドを引いたSFUを、最初の来院(T0、テクニカルトリプレート)及び最初の来院(T0)の1ヶ月(T0+1ヶ月、患者1-024-002特異的な新生抗原ペプチドプールで刺激した試料を除くテクニカルデュプレート)後のその後の来院について示す。

20

30

40

【0414】

図22は、図15Aの患者について2つの患者特異的の新生抗原ペプチドプール及びDMSO陰性対照に対するT細胞応答の検出を示す。それぞれの患者について、その患者についてインビトロ増殖させたPBM CをIFN - E L I S p o tで2つの患者特異的の新生抗原ペプチドプールにより刺激した。各ドナー及び各インビトロ増殖について、インビトロ増殖させた患者PBM CをIFN - E L I S p o tにおいて陰性対照としてのDMSOでも刺激した。図22のデータは、患者特異的の新生抗原ペプチドプール及び対応するDMSOコントロールについてバックグラウンド(対応するDMSO陰性対照)を含めた播種細胞 10^5 個当たりのスポット形成単位(SFU)として示す。患者1-038-001、1-050-001、1-001-002、CU04、1-024-001、1-0

50

24-002及びCU05について、同族のペプチドプール#1及び#2に対する単一のウェル(患者1-038-001、CU02、CU03及び1-050-001)または生物学的デュプリケートの標準偏差を含む平均(他のすべての試料)の応答を示す。患者CU02及びCU03では、細胞数のために、特異的ペプチドプール#1に対してのみ試験が可能であった。バックグラウンドよりも2倍超高い増加倍率の値を有する試料は陽性とみなし、星印で示した(応答性ドナーには、患者1-038-001、CU04、1-024-001、1-024-002、及びCU02が含まれる)。非応答性ドナーには、患者1-050-001、1-001-002、CU05、及びCU03が含まれる。

【0415】

図19A~Bに関して上記で簡単に述べたように、インビトロ培養条件が、インビトロでのデノボプライミングを可能とするのではなく、既存のインビボプライミングされたメモリーT細胞のみを増殖させたことを確認するため、HLAを一致させた健康なドナーで新生抗原を用いて一連の対照実験を行った。これらの実験の結果を図19A~B及び補足の表5に示す。これらの実験の結果により、IVS培養法を用いて健康なドナーにデノボプライミングが生じないこと及び検出可能な新生抗原特異的T細胞応答が生じないことが確認された。

【0416】

これに対して、既存の新生抗原応答性T細胞が、IFN-ELISpotを用いて患者特異的ペプチドプールで試験した患者の大部分(5/9、56%) (図15A及び図20~22)で特定された。細胞数が個々の新生抗原同族ペプチドの完全または部分的な試験を可能とした7人の患者のうち、4人の患者が試験した新生抗原ペプチドのうち少なくとも1つに応答し、これらの患者のすべてが対応するプールに対する応答を示した(図15B)。個々の新生抗原で試験した残りの3人の患者(患者1-001-002、1-050-001及びCU05)は、単一のペプチドに対する検出可能な応答は示さず(データは示さず)、これらの患者で新生抗原プールに対して見られた応答がないことが確認された(図15A)。4人の応答患者のうち、1回の来院からの試料が、応答を示した2人の患者で得られており(患者1-024-001及び1-038-001)、複数回の来院からの試料が応答を示した残りの2人の患者で得られた(CU04及び1-024-002)。複数の来院からの試料がある2人の患者について、3回の来院(患者CU04)及び2回の来院(患者1-024-002)からの累積(総和)スポット形成単位(SFU)を図15Bに示し、来院ごとの内訳を図21Bに示す。同じ来院からのさらなるPBMC試料が患者1-024-002及びCU04でやはり得られており、反復IVS培養及びELISpotにより患者特異的新生抗原に対する応答が確認された(図21C)。

【0417】

全体として、図15Aの10種のペプチドのプールに対する応答によって示されるように少なくとも1つのT細胞認識新生エピトープが特定された患者のうち、認識された新生エピトープの数は患者当たり平均で少なくとも2個であった(1個の認識されるペプチドとして逆畳み込みできない認識されたプールをカウントし、5人の患者で最小で10個のエピトープが特定された)。ELISpotによってIFN-応答について試験を行うことに加え、培養上清をELISAによってグランザイムBについて、さらにMSDサイトカインマルチプレックスアッセイによりTNF-、IL-2、及びIL-5についても試験した。陽性ELISpotを示した5人の患者のうち、4人からの細胞がグランザイムBを含む3種類以上の被検物質を分泌し(補足の表4)、新生抗原特異的T細胞の多機能性を示した。重要な点として、予測及びIVS法の組み合わせは利用可能なMHCマルチマーの限定されたセットに頼らないことから、応答は制限HLAアレルにわたって広く試験された。さらに、このアプローチは、認識された変異を特定し、最小エピトープを特定するうえで別の逆畳み込み工程を必要とするタンデムミニ遺伝子スクリーニングと異なり、最小エピトープを直接特定する。全体として、新生抗原の特定収率は、アフレーシス試料を用いてすべての変異に対してTILを試験する従来最良の方法^{9 6}と同等で

10

20

30

40

50

あった一方で、通常の5～30mLの全血を用いてわずか20種の合成ペプチドをスクリーニングするだけでよい。

【0418】

XIV.A. ペプチド

特注の組換え凍結乾燥ペプチドをJPT Peptide Technologies (Berlin, Germany)またはGenscript (Piscataway, NJ, USA)より購入し、滅菌DMSO (VWR International, Pittsburgh, PA, USA)中、10～50mMで戻し、一定の分量に分けて-80で保存した。

【0419】

XIV.B. ヒト末梢血単核細胞 (PBMC)

健康なドナーからの凍結乾燥したHLAタイピングしたPBMC (HIV、HCV及びHBVについて血清反応陰性であることを確認したものを)、Precision for Medicine (Gladstone, NJ, USA)またはCellular Technology, Ltd. (Cleveland, OH, USA)より購入し、液体窒素中で使用時まで保存した。新鮮な血液試料をResearch Blood Components (Boston, MA, USA)より、leukopakをAllCells (Boston, MA, USA)より購入し、PBMCをFicoll-Paque密度勾配 (GE Healthcare Bio, Marlborough, MA, USA)により単離した後、凍結保存した。患者のPBMCを地域の臨床処理センターで地域臨床標準業務手順書 (SOP) 及びIRBにより承認されたプロトコルに従って処理した。承認IRBは、Quorum Review IRB、Comitato Etico Interaziendale A.O.U. San Luigi Gonzaga di Orbassano、及びComite Etico de la Investigacion del Grupo Hospitalario Quiron en Barcelonaであった。

【0420】

簡単に述べると、PBMCを密度勾配遠心分離によって単離し、洗浄、カウントし、CryoStor CS10 (STEMCELL Technologies, Vancouver, BC, V6A 1B6, Canada)中、 5×10^6 細胞/mlで凍結保存した。凍結保存した細胞をcryoportで発送し、到着後に移してLN₂中で保管した。患者の人口統計を補足の表2に示す。凍結保存細胞を解凍し、Benzonase (EMD Millipore, Billerica, MA, USA)を加えたOptmizer T-cell Expansion Basal Medium (Gibco, Gaithersburg, MD, USA)中で2回洗浄し、Benzonaseなしで1回洗浄した。細胞カウント及び生存率をGuava ViaCount試薬及びGuava easyCyte HTサイトメーター (EMD Millipore) 上のモジュールを用いて評価した。次いで、細胞をその後のアッセイに適した濃度でアッセイに適した培地中に再懸濁した (次のセクションを参照)。

【0421】

XIV.C. インビトロ刺激 (IVS) 培養

健康なドナーまたは患者試料から得た既存のT細胞を、Ott et alにより適用されたアプローチ^{8, 1}と同様のアプローチで同族ペプチド及びIL-2の存在下で増殖させた。簡単に述べると、解凍したPBMCを一晩休ませ、24ウェル組織培養プレート中で10IU/mlのrhIL-2 (R&D Systems Inc., Minneapolis, MN)を添加したImmunoCult (商標) - XF T-cell Expansion Medium (STEMCELL Technologies)中、ペプチドプール (ペプチド当たり10μM、プール当たり10種のペプチド)の存在下で14日間、刺激した。細胞を 2×10^6 細胞/ウェルで播種し、培地の2/3を2～3日ごとに交換することによって培養した。1つの患者試料はプロトコルからの逸脱を示し、

10

20

30

40

50

潜在的な偽陰性とみなすべきものと考えられた。患者CU03は、解凍後に十分な数の細胞を生じなかったため、細胞をペプチドプール当たり 2×10^5 細胞で播種した（プロトコルの記載よりも10倍少ない数）。

【0422】

XIV.D. IFN 酵素結合免疫スポット (ELISpot) アッセイ

IFN 産生T細胞の検出をELISpotアッセイ¹⁴²により行った。簡単に述べると、PBMC（エクスピボまたはインビトロ増殖後のもの）を回収し、無血清RPMI（VWR International）中で洗浄し、抗ヒトIFN 捕捉抗体（Mabtech, Cincinnati, OH, USA）をコーティングしたELISpot Multiscreenプレート（EMD Millipore）中、Optmizer T-cell Expansion Basal Medium（エクスピボ）または ImmunoCult（商標）-XF T-cell Expansion Medium（増殖させた培養物）中でコントロールまたは同族ペプチドの存在下で培養した。5% CO₂、37 °Cの加湿したインキュベーター内で18時間インキュベートした後、細胞をプレートから除去し、膜に結合したIFN を抗ヒトIFN 検出抗体（Mabtech）、Vectastain Avidinペルオキシダーゼ複合体（Vector Labs, Burlingame, CA, USA）及びAEC Substrate（BD Biosciences, San Jose, CA, USA）を用いて検出した。ELISpotプレートを乾燥させ、遮光して保存し、標準化された評価¹⁴³を行うためにZellnet Consulting, Inc., Fort Lee, NJ, USAに送った。データをプレートに播種した細胞の数当たりのスポット形成単位（SFU）として示す。

10

20

【0423】

XIV.E. グランザイムB ELISA及びMSDマルチプレックスアッセイ

ELISpot上清中に分泌されたIL-2、IL-5及びTNF- α の検出をトリプレックスアッセイであるMSD U-PLEX Biomarkerアッセイ（カタログ番号 K15067L-2）を使用して行った。アッセイは製造者の指示にしたがって行った。被検物質濃度（pg/ml）を、各サイトカインについて既知の標準物質の連続希釈を用いて計算した。データをグラフ化するため、標準曲線の最小範囲よりも低い値を0に等しくなるように示した。ELISpot上清中のグランザイムBの検出を、GranzymeB DuoSet（登録商標）ELISA（R&D Systems, Minneapolis, MN）を製造者の指示に従って使用して行った。簡単に述べると、ELISpot上清を試料希釈剤中、1:4に希釈し、グランザイムB標準の連続希釈液に並べて流して濃度（pg/ml）を計算した。データをグラフ化するため、標準曲線の最小範囲よりも低い値を0に等しくなるように示した。

30

【0424】

XIV.F. IVSアッセイの陰性対照実験 - 健康なドナーで試験した腫瘍細胞株由来の新生抗原

図19Aは、健康なドナーで試験した腫瘍細胞株由来の新生抗原についてのIVSアッセイの陰性対照実験を示す。健康なドナーのPBMCを、IVS培養中、陽性対照ペプチド（感染症に予め曝露したもの）、腫瘍細胞株由来のHLAを一致させた新生抗原（曝露しないもの）、及びドナーが血清反応陰性であった病原体に由来するペプチドを含むペプチドプールで刺激した。次いで、増殖させた細胞を、DMSO（陰性対照、黒い丸）、PHA及び一般的な感染症ペプチド（陽性対照、赤い丸）、新生抗原（非曝露、水色の丸）、またはHIV及びHCVペプチド（ドナーが血清反応陰性であったもの。濃紺色、A及びB）で刺激した後、IFN ELISpot（ 10^5 細胞/ウェル）により分析した。データを播種細胞 10^5 個当たりのスポット形成単位（SFU）として示す。平均及びSEMを含む生物学的レプリケートを示す。新生抗原またはドナーが曝露されていない病原体由来のペプチド（血清反応陰性）に対する応答は観察されなかった。

40

【0425】

50

XIV . G . I V S アッセイの陰性対照実験 - 健康なドナーで試験した患者由来の新生抗原

図19Aは、健康なドナーにおける応答性について試験した患者由来の新生抗原についてのIVSアッセイの陰性対照実験を示す。HLAを一致させた新生抗原ペプチドプールに対する健康なドナーにおけるT細胞応答の評価。左パネル：健康なドナーのPBMCを、エキスピボのIFN ELISpotにおいて対照(DMSO、CEF、及びPHA)またはHLA一致させた患者由来新生抗原ペプチドで刺激した。データを3重のウェルについてプレートに播種した細胞 2×10^5 個当たりのスポット形成単位(SFU)として示す。右パネル：新生抗原プールまたはCEFプールの存在下で増殖させたIVS培養後の健康なドナーのPBMCを、IFN ELISpotにおいて対照(DMSO、CEF、及びPHA)またはHLAを一致させた患者由来の新生抗原ペプチドプールで刺激した。データを3重のウェルについてプレートに播種した細胞 1×10^5 個当たりのSFUとして示す。健康なドナーにおいて新生抗原に対する応答は認められなかった。

10

【0426】

XIV . H . 補足の表3：NSCLC患者におけるT細胞認識について試験したペプチド
図15A～Cで調べたN=9人の患者で試験した新生抗原ペプチドの詳細(NSCLC患者由来の新生抗原応答性T細胞の特定)。主なフィールドは、ソース変異、ペプチド配列、ならびに観察されたプール及び個々のペプチド配列を含む。列「most__probable__restriction」は、モデルが予測したどのアレルが各ペプチドを提示する可能性が最も高かったかを示す。結合親和性予測(「方法」)により計算した各患者のすべての変異ペプチド間でのこれらのペプチドのランクも含まれる。

20

【0427】

4つのペプチドが、完全MSモデルによって高い順位にランクされ、予測された結合親和性が低いかまたは結合親和性予測によって低い順位にランクされたCD8T細胞によって認識された。

【0428】

これらのペプチドのうち3つでは、これはモデルとMHCflurry1.2.0との間のHLAカバレッジの差による。ペプチドYEHEDEVKEAは、MHCflurry1.2.0によってカバーされていないHLA-B*49:01によって提示されると予測される。同様に、ペプチドSSAAPFPL及びFVSTSDIKSMは、やはりMHCflurry1.2.0によってカバーされていないHLA-C*03:04により提示されると予測される。原理的にはすべてのアレルをカバーするpan特異的結合親和性予測ツールであるオンラインNetMHCpan4.0(BA)予測ツールは、SSAAPFPLをHLA-C*03:04に対する強い結合物質としてランクし(23.2nM、患者1-024-002で2位にランク)、FVSTSDIKSMのHLA-C*03:04に対する弱い結合(943.4nM、患者1-024-002で39位にランク)及びYEHEDEVKEAのHLA-B*49:01に対する弱い結合(3387.8nM)を、また、やはりこの患者に存在するがモデルによってカバーされていないHLA-B*41:01に対するより強い結合(208.9nM、患者1-038-001で11位にランク)を予測している。したがって、これら3つのペプチドのうち、FVSTSDIKSMは結合親和性予測によれば漏れていたであろうし、SSAAPFPLは捕捉されていたであろうし、YEHEDEVKEAのHLA拘束性は不明である。

30

40

【0429】

ペプチド特異的T細胞応答が逆畳み込みされた残りの5つのペプチドは、モデルによって判定された最も可能性の高い提示アレルがやはりMHCflurry1.2.0によってカバーされていた患者に由来するものであった。これら5つのペプチドのうち4つ(4/5)は標準的な500nMの閾値よりも強い予測結合親和性を有し、上位20位にランクされたが、モデルによるランクよりもいくぶん低いランクであった(ペプチド

DENITTIQF,

QDVSQVER, EVADAATLTM, DTVEYPYTSF

は、MHCflurryによって2、14、7、及び9位にランクされたのに対して、モデルによってそれぞれ0、4、5、7位にランクされた)。ペプチドGTTKKDVLKはCD8T細胞により認識され、モデルによって1位にランクされたが、MHCflurryによるランクは70位であり、予測結合親和性は2169nMであった。

【0430】

全体として、完全MSモデルにより高い順位にランクされた8つの個々に認識されたペプチドのうち6つ(6/8)が、結合親和性予測を用いた場合にもやはり高い順位にランクされ、予測結合親和性が500nm未満であったのに対して、8つの個々に認識されたペプチドのうち2つ(2/8)は、完全MSモデルの代わりに結合親和性予測が用いられていたならば漏れていたであろうと考えられる。

10

【0431】

XIV.I.補足の表4: NSCLC新生抗原ペプチドから得たELISpot上清に対するMSDサイトカインマルチプレックス及びELISAアッセイ

陽性ELISpot(IFN)ウェルから得た上清中で検出された被検物質をグランザイムB(ELISA)、TNF、IL-2及びIL-5(MSD)について示す。値はテクニカルレプリケートからの平均のpg/mlとして示す。陽性値を斜体で示す。グランザイムB ELISA: DMSOバックグラウンドよりも1.5倍以上の値を陽性とみなした。U-plex MSDアッセイ: DMSOバックグラウンドよりも1.5倍以上の値を陽性とみなした。

20

【0432】

XIV.J.補足の表5: IVS対照実験における新生抗原及び感染症エピトープ

図19A~Bに示されるIVS対照実験で試験した腫瘍細胞株新生抗原及びウイルスペプチドの詳細なフィールドには、ソース細胞株またはウイルス、ペプチド配列、及び予測された提示HLAアレルが含まれる。

【0433】

XIV.K.データ

予測モデルを訓練及び試験するために使用したMSペプチドデータセット(図14A~D)は、MassIVEアーカイブ(massive.ucsd.edu)、アクセス番号MSV000082648で取得可能である。ELISpot(FIGS. 15A-C and 19A-B)により試験される新生抗原ペプチドはマニユスクリプトとともに含まれている(補足の表3及び5)。

30

【0434】

XV.実施形態8~10の方法

XV.A.質量分析

XV.A.1.試料

質量分析による分析用のアーカイブされた凍結組織試料は、BioServe(Beltsville, MD)、ProteoGenex(Culver City, CA)、iSpecimen(Lexington, MA)、及びIndivumed(Hamburg, Germany)を含む販売元から入手した。試料のサブセットも、Comite de Protection des Personnes, Ile-de-France VIIによって承認されたりサーチプロトコールに基づき、Hopital Marie Lannelongue(Le Plessis-Robinson, France)で患者から予め採取した。

40

【0435】

XV.A.2.HLA免疫沈降

HLAペプチド分子の単離を、組織試料の溶解及び可溶化後^{87, 124-126}に免疫沈降(IP)法を用いて行った。新鮮な凍結組織を粉碎し(CryoPrep; Cov

50

aris, Woburn, MA)、溶解バッファー(1% CHAPS、20 mM Tris-HCl、150 mM NaCl、プロテアーゼ及びホスファターゼ阻害剤、pH = 8)を加えて組織を可溶化し、得られた溶液を4 で2時間遠心して破片をペレット化した。清澄化したライセートをHLA特異的IPに使用した。免疫沈降はこれまでに述べられているようにして抗体W6/32を使用して行った^{1 2 7}。ライセートを抗体ビーズに加え、4 で一晩回転し、免疫沈降を行った。免疫沈降後、ビーズをライセートから除去した。IPビーズを洗浄して非特異的結合を除去し、HLA/ペプチド複合体を2 N酢酸でビーズから溶出させた。タンパク質成分を分子量スピンカラムを使用してペプチドから除去した。得られたペプチドをSpeed Vac蒸発により乾燥状態とし、MS分析を行うまで-20 で保存した。

10

【0436】

XV.A.3 ペプチドシーケンシング

乾燥させたペプチドをHPLCバッファーA中で戻し、C-18マイクロキャピラリーHPLCカラムにロードして質量分析計に勾配溶出した。180分の0~40%B(溶媒A:0.1%ギ酸、溶媒B:80%アセトニトリル中0.1%ギ酸)の勾配を用いてペプチドをFusion Lumos質量分析計(Thermo)中に溶出した。ペプチドの質量/電荷(m/z)のMS1スペクトルを分解能120,000でOrbitrap検出器中に収集し、続いて選択したイオンのHCD断片化後に20回のMS2の低分解能スキャンをOrbitrapまたはイオントラップ検出器中に収集した。MS2イオンの選択は、データ依存性取得モード及びイオンのMS2選択後に30秒間の動的排除を用いて行った。MS1スキャンの自動利得制御(AGC)を 4×10^5 に設定し、MS2スキャンについては 1×10^4 に設定した。HLAペプチドのシーケンシングには、MS2断片化を行うために+1、+2及び+3の荷電状態を選択することができる。

20

【0437】

各分析からのMS2スペクトルをComet^{1 2 8}、^{1 2 9}を用いてタンパク質データベースに対して検索し、ペプチド特定をPercolator^{1 3 0 ~ 1 3 2}を使用してスコア付けした。

【0438】

XV.B.機械学習

XV.B.1.データのコード化

各試料について、訓練データポイントは、試料中で発現された正確に1つの遺伝子にマッピングされる参照プロテオームからのすべての8~11マー(包括する)ペプチドとした。全体の訓練データセットは、各訓練試料からの訓練データセットを連結することにより生成した。8~11個の範囲はすべてのHLAクラスI提示ペプチドの約95%を捕捉するため、この長さの範囲を選択したが、12~15個の長さを加えることも、計算需要の中度の増大を代償として同じ方法を用いて実現することが可能である。ペプチド及びフランキング配列を、ワン・ホットエンコーディングスキームを用いてベクトル化した。複数の長さ(8~11)のペプチドを、アミノ酸のアルファベットをパッド文字で増やし、すべてのペプチドを最大長さ11までパディングすることにより固定長さのベクターとして表した。訓練ペプチドのソースタンパク質のRNA存在量を、RSEM^{1 3 3}から得られたイソフォームレベルのTPM(transcripts per million)推定値の対数として表した。各ペプチドについて、ペプチドごとTPMを、ペプチドを含むイソフォームのそれぞれについてイソフォームごとTPM推定値の総和として計算した。0TPMで発現された遺伝子からのペプチドは訓練データから除外し、試験時に、非発現遺伝子からのペプチドに提示確率0を割り当てた。最後に、各ペプチドに、Ensemblタンパク質ファミリーIDを割り当て、各固有のEnsemblタンパク質ファミリーIDは遺伝子毎提示傾向切片に対応した(次のセクションを参照)。

40

【0439】

XV.B.2.モデルアーキテクチャの仕様

完全提示モデルは以下の関数形を有する。すなわち、

50

(式1) $Pr(\text{提示ペプチド } i) = \sum_{k=1}^m a_k^i \cdot Pr(\text{アレル } a \text{ による提示ペプチド } i)$

式中 k は、1 ~ m までのデータセット内の H L A アレルの添え字であり、

a_k^i

は、標識変数であり、アレル k がペプチド i が由来する試料中に存在する場合にその値は 1 であり、そうでない場合には 0 である。特定のペプチド i について、

a_k^i

のうちの最大 6 個 (ペプチド i の由来する試料の H L A 型に対応する 6 個) 以外のすべては、0 である。確率の総和は、例えば、

$$\epsilon = 10^{-6}$$

で、

$$1 - \epsilon$$

でクリップする。

【0440】

提示のアレル毎の確率を下記のようにモデル化する。すなわち、

$$Pr(\text{アレル } a \text{ による提示ペプチド } i) = \text{sigmoid} \{ NN_a(\text{ペプチド } i) + NN_{\text{ランキング}}(\text{ランキング } i) + NN_{RNA}(\log(TPM_i)) + \text{試料}(i) + \text{タンパク質}(i) \}$$

式中、変数は以下の意味を有する。sigmoid は、シグモイド (expit としても知られる) 関数であり、ペプチド i は、ペプチド i のワンホットエンコーディングされた中間パディングされたアミノ酸配列であり、 NN_a は、提示の確率に対するペプチド配列の寄与をモデル化する線形最終層活性化によるニューラルネットワークであり、ランキング i は、ソースタンパク質中のペプチド i のワンホットエンコーディングされたランキング配列であり、 $NN_{\text{ランキング}}$ は、提示の確率に対するランキング配列の寄与をモデル化する線形最終層活性化によるニューラルネットワークであり、 TPM_i は、TPM 単位内のペプチド i の mRNA のソース mRNA の発現であり、 $\text{試料}(i)$ は、ペプチド i の由来する試料 (すなわち患者) であり、 $\text{試料}(i)$ は試料ごと切片であり、 $\text{タンパク質}(i)$ は、ペプチド i のソースタンパク質であり、 $\text{タンパク質}(i)$ はタンパク質ごと切片である (提示の遺伝子毎傾向としても知られる)。

【0441】

結果のセクションに記載されるモデルでは、各コンポーネントのニューラルネットワークは以下のアーキテクチャを有する。すなわち、

- NN_a のそれぞれは、入力次元数 231 (11 残基 × 残基ごとに 21 個の可能な文字 (パッド文字を含む))、幅 256、隠れ層の整流線形単位 (ReLU) 活性化、出力層の線形活性化、及び訓練データセットの H L A アレル a ごとに 1 個の出力ノードを有する単一隠れ層の多層パーセプトロン (MLP) の 1 個の出力ノードである。

- $NN_{\text{ランキング}}$ は、入力次元数 210 (N 末端ランキング配列の 5 残基 + C 末端ランキング配列の 5 残基 × 残基ごとに 21 個の可能な文字 (パッド文字を含む))、幅 32、隠れ層の整流線形単位 (ReLU) 活性化、及び出力層の線形活性化を有する単一隠れ層の MLP である。

- NN_{RNA} は、入力次元数 1、幅 16、隠れ層の整流線形単位 (ReLU) 活性化、及び出力層の線形活性化を有する単一隠れ層の MLP である。

【0442】

モデルの一部のコンポーネント (例えば、 NN_a) は、特定の H L A に依存するが、多くのコンポーネント ($NN_{\text{ランキング}}$ 、 NN_{RNA} 、 $\text{試料}(i)$ 、 $\text{タンパク質}(i)$) は依存しない。前者は「アレル相互作用性」と呼ばれ、後者は「アレル非相互作用性」と呼ばれる。アレル相互作用性及びアレル相互作用性としてモデル化するための特性は、従来の生物学的知見に基づいて選択される。すなわち、H L A アレルはペプチドを見るため、ペプチド配列はアレル相互作用性としてモデル化されるべきであるが、ソースタンパク質、RNA 発現またはランキング配列に関する情報は H L A アレルに伝達されず (ペプチドは小胞体内で H L A と出会うまでにそのソースタンパク質から分離されているた

10

20

30

40

50

め)、したがってこれらの特性はアレル非相互作用性としてモデル化されるべきである。このモデルを、Keras v2.0.4¹³⁴及びTheano v0.9.0¹³⁵で実施した。

【0443】

ペプチドMSモデルは、完全MSモデル(式1)と同じ逆畳み込み手順を用いるが、提示のアレル毎の確率はペプチド配列及びHLAアレルのみを考慮した縮小アレル毎モデルを用いて生成した。

$$Pr(\text{アレル } a \text{ による提示ペプチド } i) = \text{sigmoid}\{NN_a(\text{ペプチド } i)\}$$

【0444】

ペプチドMSモデルは結合親和性予測と同じ特性を用いるが、モデルの重みは異なるデータタイプ(すなわち、HLAペプチド結合親和性データに対する質量分析データ)で訓練される。したがって、ペプチドMSモデルの予測性能を完全MSモデルと比較することで、全体の予測性能に対する非ペプチド特性(すなわち、RNA存在量、ランキング配列、遺伝子ID)の寄与が明らかとなり、ペプチドMSモデルの予測性能を結合親和性モデルと比較することで、全体の予測性能に対するペプチド配列の改善されたモデル化の重要度が明らかとなる。

10

【0445】

XV.B.3. 訓練/検証/試験の分割

本発明者らは、以下の手順を用いて訓練/検証/試験セットのうち2つ以上に現れるペプチドがないようにした。すなわち、最初に、2つ以上のタンパク質に現れるすべてのペプチドを参照プロテオームから除去し、次に、プロテオームを10個の隣接したペプチドに分配する。各ブロックを訓練、検証、及び試験セットに固有に割り当てる。これにより、訓練、検証、または試験データセットのうち2つ以上のデータセットに現れるペプチドはなくなる。検証セットは、早期終了のみに用いた。図14Aの腫瘍試料試験データは、訓練及び検証セットから完全に除外された5つの腫瘍試料からの試験セットペプチド(すなわち、試験セットに固有に割り当てられた隣接するペプチドのブロックからのペプチド)を表す。単一アレル試料からのペプチドが訓練データには含まれたが、訓練及び検証セットに組み込まれたペプチドのセット(提示及び非提示の療法)は、図14Bで試験データとして使用されるペプチドのセットとは互いに素である。

20

【0446】

XV.B.4. モデルの訓練

モデルを訓練するため、ペプチドごと損失が負のベルヌーイ対数尤度損失関数である(対数損失としても知られる)ものとして、すべてのペプチドを独立したものとしてモデル化した。正式には、ペプチド*i*の全体の損失に対する寄与は、

$Loss(i) = -\log(\text{Bernoulli}(y_i | Pr(\text{提示ペプチド } i)))$
ただし、 y_i は、ペプチド*i*のラベルである(すなわち、ペプチド*i*が提示される場合に $y_i = 1$ であり、そうでない場合には0であり、 $i.i.d.$ の2項観測ベクトル y が与えられるものとして、 $\text{Bernoulli}(y | p)$ は、パラメータ $p \in [0, 1]$ のベルヌーイ尤度を示す)。このモデルを損失関数を最小化することにより訓練した。

30

【0447】

訓練時間を低減するため、陰性標識された訓練データの90%を無作為に除去することによりクラスバランスを調整し、全体の訓練セットクラスバランスを約2000個の非提示ペプチドごとに1個の提示ペプチドとした。モデル重みを、Glorot均一手順61を用いて初期化し、標準的なパラメータを用いたADAM62確率論的オプティマイザを用いてNvidia Maxwell TITAN X GPUで訓練した。全データの10%からなる検証セットを早期終了に用いた。このモデルを1/4エポックごとに検証セットで評価し、モデル訓練を検証損失(すなわち、検証セットでの負のベルヌーイ対数尤度)が低下しなくなった最初の1/4エポック後に終了した。

40

【0448】

完全提示モデルは、アンサンブル内のすべてのモデルについてモデル重さの異なるラン

50

ダム初期化を行った同じ訓練データのシャッフルしたコピーで各レプリケートを独立して訓練した10個のモデルレプリケートのアンサンブルとした。試験時に、モデルレプリケートによって出力された確率の平均を取ることによって予測を生成した。

【0449】

XV.B.5. モチーフロゴ

モチーフロゴを、weblogolib Python API v3.5.0¹³⁸を用いて生成した。結合親和性ロゴを生成するため、mhc_ligand_full.csv ファイルを2017年7月にImmune Epitope Database (IEDB⁸⁸)からダウンロードし、以下の基準を満たすペプチドを保持した。すなわち、ナノモル(nM)単位の測定値、参照日が2000年以降であり、オブジェクトタイプが「直鎖状ペプチド」に等しく、ペプチド内のすべての残基が標準的な20文字のアミノ酸アルファベットから引用されるもの。測定された結合親和性が従来¹⁰の結合閾値である500nMよりも低いフィルタリングされたペプチドのサブセットを使用してロゴを生成した。IEDBの結合物質の数が少なすぎるアレルペアについてはロゴは生成されなかった。学習された提示モデルを表すロゴを生成するため、2,000,000種のランダムなペプチドについてモデル予測を各アレル及び各ペプチド長で予測した。各アレル及び各長さについて、学習された提示モデルにより上位1%(すなわち、上位の20,000種)にランク付けされたペプチドを用いてロゴを生成した。重要な点として、IEDBからのこの結合親和性データは、モデルの訓練または試験では使用せず、学習されたモチーフの比較にのみ使用した。²⁰

【0450】

XV.B.6. 結合親和性の予測

本発明者らは、モデルのNetMHCファミリーに匹敵する性能を有するオープンソースのGPU互換性HLAクラスI結合親和性予測ツールであるMHCflurry^{1.2.0¹³⁹}からの結合親和性のみの予測ツールを使用してペプチド-MHC結合親和性を予測した。複数のHLAアレルにわたって単一のペプチドについての結合親和性予測を組み合わせるため、最小の結合親和性を選択した。複数のペプチドにわたって結合親和性を組み合わせるため(すなわち、図14Cに示されるように、複数の変異ペプチドがまたがった変異をランク付けするため)、ペプチドにわたって最小の結合親和性を選択した。T細胞データセットに対するRNA発現の閾値を決定するため、TPM>1の閾値までのTCGAからの腫瘍型が一致したRNA-seqデータを使用した。初期のT細胞データセットのすべてを初期の公表中、TPM>0でフィルタリングしたため、TPM>0でフィルタリングされるTCGAのRNA-seqデータは使用されなかった。³⁰

【0451】

XV.B.7. 提示の予測

複数のHLAアレルにわたって単一のペプチドについての提示の確率を加え合わせるため、確率の総和を式1に示されるように特定した。複数のペプチドにわたって提示の確率を加え合わせるため(すなわち、図14Cに示されるように、複数の変異ペプチドがまたがった変異をランク付けするため)、提示の確率の総和を特定した。確率的には、ペプチドの提示がi.i.d.のベルヌーイランダム変数とみなされる場合、確率の総和は提示される変異ペプチドの予想数に対応する。すなわち、⁴⁰

$$E[\text{変異}i\text{にまたがる提示新生抗原の数}] = \sum_{j=1}^{n_i} \text{Pr}[\text{提示エピトープ}j]$$

ただし、Pr[提示エピトープj]は、エピトープjに訓練された提示モデルを適用することによって得られ、 n_i は、変異iにまたがる変異エピトープの数を示す。例えば、そのソース遺伝子の末端から遠いSNV_iについて、8個のまたがった8マー、9個のまたがった9マー、10個のまたがった10マー、及び11個のまたがった11マーが、全部で $n_i = 38$ 個のまたがった変異エピトープについて存在する。

【0452】

10

20

30

40

50

X V . C . 次世代シーケンシング

X V . C . 1 . 試料

凍結切除された腫瘍のトランスクリプトーム分析を行うため、MS分析に使用したのと同じ組織試料（腫瘍または隣接する正常組織）からRNAを得た。抗PD1療法を行っている患者の新生抗原エクソーム及びトランスクリプトーム分析を行うため、DNA及びRNAをアーカイブされたFFPE腫瘍生検から得た。隣接正常組織、一致血液、またはPBMCを用いて正常エクソーム及びHLAタイピングを行うための正常DNAを得た。

【0453】

X V . C . 2 . 核酸抽出及びライブラリーの構築

血液由来の正常/生殖系DNAを、Qiagen DNeasyカラム（Hilden, Germany）を製造者の推奨する手順に従って使用して単離した。組織試料からのDNA及びRNAをQiagen Allprep DNA/RNA単離キットを製造者の推奨する手順に従って使用して単離した。これらのDNA及びRNAをPicogreen及びRibogreen Fluorescence（Molecular Probes）によりそれぞれ定量し、収量が50ng超の試料をライブラリーの構築に進めた。DNAシーケンシングライブラリーを超音波せん断（Covaris, Woburn, MA）に続き、DNA Ultra II（NEB, Beverly, MA）ライブラリー調製キットを製造者の推奨するプロトコールに従って使用することにより作製した。RNA Ultra II（NEB）による熱断片化及びライブラリー構築により腫瘍RNAシーケンシングライブラリーを作製した。得られたライブラリーをPicogreen（Molecular Probes）により定量した。

【0454】

X V . C . 3 . 全エクソーム捕捉

DNA及びRNAシーケンシングライブラリーのエクソン濃縮を、xGEN Whole Exome Panel（Integrated DNA Technologies）を使用して行った。1~1.5µgの正常DNAまたは腫瘍DNAもしくはRNA由来ライブラリーを入力として用い、12時間よりも長い時間にわたってハイブリダイズさせた後、ストレプトアビジン精製を行った。捕捉されたライブラリーをPCRにより最小限増幅し、NEBNext Library Quant Kit（NEB）により定量した。捕捉された各ライブラリーを等モル濃度でプールし、c-bot（Illumina）を用いてクラスター化し、75塩基対の端部においてHiSeq4000（Illumina）で、500x超の腫瘍エクソーム、100x超の正常エクソーム、及び100M超のリードの腫瘍トランスクリプトームのターゲットユニークな平均カバレッジにまでシーケンシングした。

【0455】

X V . C . 4 分析

エクソームリード（FFPE腫瘍エクソーム及び一致させた正常エクソーム）を、BWA-MEM¹⁴⁴（v.0.7.13-r1126）を使用して参照ヒトゲノム（hg38）とアラインした。RNA-seqリード（FFPE及び凍結腫瘍組織試料）を、STAR（v.2.5.1b）を使用してゲノム及びGENCODE転写産物（v.25）とアラインした。RNA発現を、同じ参照転写産物でRSEM¹³³（v.1.2.31）を使用して定量した。Picard（v.2.7.1）を使用してデュプリケートアラインメントをマークし、アラインメントメトリックを計算した。GATK¹⁴⁵（v.3.5-0）による塩基のクオリティー補正（base quality score recalibration）後のFFPE腫瘍試料について、FreeBayes¹⁴⁶（1.0.2）によりペアリングした腫瘍-正常エクソームを用いて置換及び短いindel変異を検出した。フィルターには、アレル頻度>4%；塩基クオリティーの中央値>25、支持リードの最小マッピングクオリティー=30、及び十分なカバレッジが得られたとして正常の代替的リードカウント 2が含まれた。変異体はまた、両方の鎖で検出されなければならない。反復領域に生じる体細胞変異体は除外した。翻訳及びアノテ

10

20

30

40

50

ーションは、RefSeq転写産物を用いてsnpeff¹⁴⁷(v.4.2)により行った。腫瘍RNAアラインメントで確認された非同義、ノンストップ変異体を新生抗原予測に進めた。Optitype¹⁴⁸1.3.1を使用してHLA型を生成した。

【0456】

XV.C.5. 図19A~B: IVS対照実験用の腫瘍細胞株及び一致した正常細胞株
腫瘍細胞株H128、H122、H2009、H2126、Colo829及びそれらの正常なドナー一致した対照細胞株BL128、BL2122、BL2009、BL2126 and Colo829BLをすべて、ATCC (Manassas, VA)より購入し、販売業者の指示にしたがって $10^{8.3} \sim 10^{8.4}$ 個の細胞にまで増殖させた後、核酸抽出及びシーケンシング用にスナップ凍結した。NGSを、Mutect¹⁴⁹(3.1-0)を置換変異の検出にのみ用いた点以外は、概ね上記に述べたのと同様にして行った。IVS対照アッセイで使用したペプチドを、補足の表5に示す。

10

【0457】

XV.D. Class IIモデルの概念実証

本発明者らは、本明細書に開示される予測モデルがクラスII HLAペプチド予測にも適用可能かどうかを評価した。これを行うため、公開されているクラスIIの質量分析データを、それぞれが単一のHLAクラスIIアレルを発現する2つの細胞株について取得した。一方の細胞株はHLA-DRB1*15:01を発現し、他方はHLA-DRB5*01:01を発現した¹⁵⁰。これら2つの細胞株を訓練データ用に使用した。試験データ用には、クラスIIの質量分析データをHLA-DRB1*15:01及びHLA-DRB5*01:01の両方を発現する別の細胞株から取得した¹⁵¹。RNAシーケンシングデータは、訓練用及び試験用細胞株のいずれでも入手できなかったため、異なるB細胞株B721.221からのRNAシーケンシングデータ⁹²で代用した。

20

【0458】

ペプチドセットを、クラスIIのデータでは、9~20個の長さを有するペプチドを含めた点を除いてHLAクラスIIデータと同じ手順を用いて、訓練、検証、及び試験セットに分割した。訓練データは、HLA-DRB1*15:01により提示された330種のペプチド、及びHLA-DRB5*01:01により提示された103種のペプチドを含むものとした。検査データセットは、HLA-DRB1*15:01またはHLA-DRB5*01:01により提示された223種のペプチド及び4708種の非提示ペプチドを含むものとした。

30

【0459】

本発明者らは、10個のモデルのアンサンブルを訓練データセットで訓練してHLAクラスIIのペプチド提示を予測した。これらのモデルのアーキテクチャ及び訓練手順は、クラスIIのモデルでは入力ペプチドとして長さ11個ではなく長さ20個にワンホットエンコーディングされ、パディングが0の配列を採用した点を除き、クラスIIの提示を予測するために用いたものと同じとした。

【0460】

図23は、「MSモデル」、HLA-DRB1*15:01及びHLA-DRB5*01:01にわたって最も低いNetMHCIIpanパーセントイルランクを採用する、「NetMHCIIpan rank」: NetMHCIIpan 3.1⁷⁷、及びHLA-DRB1*15:01及びHLA-DRB5*01:01にわたってnM単位の最も強い親和性を採用する「NetMHCIIpan nM」: NetMHCIIpan 3.1の、HLA-DRB1*15:01 / HLA-DRB5*01:01試験データセット内のペプチドのランク付けにおける予測性能を比較したものである。「MSモデル」は、本明細書に開示されるMHCクラスIIの提示予測モデルである。

40

【0461】

詳細には、図23は、受信者操作特性(ROC)曲線、ならびにROC曲線下面積AUC(パネルA)及びAUC_{0.1}(パネルB)統計をこれらのランク付け法について示したものである。AUC_{0.1}は、エピトープ予測フィールドで一般的に考慮される、0~

50

0.1 FPR * 10 の AUC である¹⁹。NetMHCIIpan nM法とrank法の性能は同様であった。MSモデルの性能が最も高く、特にROC曲線の臨界高特異性領域において、比較した方法の性能を大幅に上回った (AUC_{0.1} が 0.27 に対して 0.41)。

【0462】

XVI. 実施例 11: NSCLC 患者の末梢血由来の新生抗原特異的メモリー T 細胞の TCR のシーケンシング

図 24 は、NSCLC 患者の末梢血由来の新生抗原特異的メモリー T 細胞の TCR をシーケンシングするための方法を示す。NSCLC 患者 CU04 からの末梢血単核細胞 (PBMC) (図 15A ~ 22 に関して上記に述べたもの) を ELISpot インキュベーション後に回収した。詳細には、上記に述べたように、患者 CU04 の 2 回の来院からインビトロ増殖させた PBMC を IFN ELISpot において CU04 特異的な個別の新生抗原ペプチド (図 21C)、CU04 特異的な新生抗原ペプチドプール (図 21C)、及び DMSO 陰性対照 (図 22) で刺激した。インキュベーション後、検出抗体の添加に先立って、PBMC を新しい培養プレートに移し、ELISpot アッセイが完了する間、インキュベーター内で維持した。陽性 (応答性) ウェルを ELISpot の結果に基づいて特定した。図 21 に示されるように、特定された陽性ウェルには、CU04 特異的な個別のペプチド 8 で刺激したウェル、及び CU04 特異的な新生抗原ペプチドプールで刺激したウェルが含まれる。これらの陽性ウェル及び陰性対照 (DMSO) ウェルからの細胞を加え合わせて、磁気標識抗体で CD137 について染色し、Miltenyi 磁気単離カラムを使用して濃縮を行った。

【0463】

上記に述べたように単離して増殖させた CD137 濃縮及び枯渇 T 細胞画分を 10x Genomics シングルセル・レゾリューション・ペアド免疫 TCR プロファイリングアプローチを用いてシーケンシングした。詳細には、生 T 細胞を、続くシングルセル cDNA 作製及び完全長 TCR のプロファイリング (5' UTR ~ 定常領域、 と を対合させる) を行うためにシングルセルエマルジョン中に分配した。1つのアプローチでは、転写産物の 5' 末端の分子バーコード付けしたテンプレート交換オリゴを利用し、第 2 のアプローチでは、3' 末端の分子バーコード付けした定常領域オリゴを利用し、第 3 のアプローチでは、RNA ポリメラーゼプロモーターを TCR の 5' 末端または 3' 末端に連結する。これらのアプローチはいずれも、シングルセルレベルでの 及び TCR ペアの特定及び逆畳み込みを可能とするものである。得られたバーコード付けした cDNA 転写産物に最適化した酵素及びライブラリー構築ワークフローを行ってバイアスを低減し、細胞のプール内のクロナタイプの正確な表現を確実にした。各ライブラリーを Illumina 社の MiSeq または HiSeq 4000 装置 (対合末端 150 サイクル) で細胞当たり約 5 千 ~ 5 万リードのターゲットシーケンシング深さに対してシーケンシングした。得られた TCR 核酸配列を補足の表 6 に示す。補足の表 6 に記載される TCR 及び TCR 鎖の存在を、オーソゴナルなアンカー PCR ベース TCR シーケンシングアプローチ (Archer) によって確認した。この特定のアプローチは、10x Genomics ベースの TCR シーケンシングと比較して限定された数の細胞を入力として使用し、用いられる酵素的操作の数が少ないという利点を有する。

【0464】

シーケンシングの出力を 10x ソフトウェア及びカスタムバイオインフォマティクスパイプラインを使用して分析して、やはり補足の表 6 に示されるような T 細胞受容体 (TCR) 及び 鎖を特定した。補足の表 6 は、 及び 可変 (V) 領域、連結 (J) 領域、定常 (C) 領域、及び 多様性 (D) 領域、ならびに大部分の優位な TCR クロナタイプの CDR3 アミノ酸配列をさらに示す。クロナタイプは、固有の CDR3 アミノ酸配列の、 鎖のペアとして定義される。クロナタイプを、2 細胞よりも高い頻度で存在する 1 個の 鎖と 1 個の 鎖のペアについてフィルタリングして患者 CU04 のターゲットペプチド当たりのクロナタイプの最終リストを得た (補足の表 6)。

【0465】

要約すると、図24に関して述べた方法を用いることで、セクションXIVの実施例10に関して上記に述べたようにして特定された患者CU04の腫瘍新生抗原に対して新生抗原特異的である患者CU04の末梢血由来のメモリーCD8+T細胞が特定された。これらの特定された新生抗原特異的T細胞のTCRをシーケンシングした。またさらに、上記の提示モデルにより特定された患者CU04の腫瘍新生抗原に対して新生抗原特異的であるシーケンシングしたTCRが特定された。

【0466】

XVII. 実施例12：T細胞療法の新生抗原特異的メモリーT細胞の使用

患者の腫瘍によって提示される新生抗原に対して新生抗原特異的であるT細胞及び/またはTCRが特定された後、これらの特定された新生抗原特異的T細胞及び/またはTCRを患者のT細胞療法に使用することができる。詳細には、これらの特定された新生抗原特異的T細胞及び/またはTCRを使用してT細胞療法において患者に注入するための治療量の新生抗原特異的T細胞を作製することができる。患者のT細胞療法で使用するための治療量の新生抗原特異的T細胞を作製するための2つの方法を、本明細書のセクションXVII.A.及びXVII.B.で考察する。第1の方法は、患者試料から特定された新生抗原特異的T細胞を増殖させることを含む(セクションXVII.A.)。第2の方法は、特定された新生抗原特異的T細胞のTCRをシーケンシングすることと、シーケンシングされたTCRを新たなT細胞にクローニングすることとを含む(セクションXVII.B.)。本明細書では明示的に触れていないT細胞療法で使用するための新生抗原特異的T細胞を作製するための代替的な方法を用いて、T細胞療法で使用するための治療量の新生抗原特異的T細胞を作製することもできる。これらの方法の1つ以上によって新生抗原特異的T細胞が得られた後、これらの新生抗原特異的T細胞をT細胞療法を行うために患者に注入することができる。

【0467】

XVII.A. T細胞療法用の患者試料からの新生抗原特異的メモリーT細胞の特定及び増殖

患者のT細胞療法に使用するための治療量の新生抗原特異的T細胞を作製するための第1の方法は、患者試料から特定された新生抗原特異的T細胞を増殖させることを含む。

【0468】

詳細には、新生抗原特異的T細胞を患者のT細胞療法に使用するための治療量にまで増殖させるため、患者のがん細胞によって提示される可能性が最も高い新生抗原ペプチドのセットを、上記に述べた提示モデルを使用して特定する。さらに、T細胞を含有する患者試料を患者から得る。患者試料は、患者の末梢血、腫瘍浸潤リンパ球(TIL)、またはリンパ節細胞を含み得る。

【0469】

患者試料が患者の末梢血を含む実施形態では、以下の方法を用いて新生抗原特異的T細胞を治療量にまで増殖させることができる。一実施形態では、プライミングを行うことができる。別の実施形態では、既に活性化されているT細胞を上記に述べた方法のうちの1つ以上を用いて特定することができる。別の実施形態では、プライミング及び既に活性化されているT細胞の特定の両方を行うことができる。プライミング及び既に活性化されているT細胞の特定の両方を行うことの利点は、表現される特異性の数が最大となることである。プライミング及び既に活性化されているT細胞の特定の両方を行うことの難点は、このアプローチは困難で時間がかかることである。別の実施形態では、必ずしも活性化されていない新生抗原特異的細胞を単離することができる。かかる実施形態では、これらの新生抗原特異的細胞の抗原特異的または非特異的増殖を行うこともできる。これらのプライミングされたT細胞を回収した後、プライミングされたT細胞を迅速増殖プロトコールに供することができる。例えば、いくつかの実施形態では、プライミングされたT細胞をRosenberg迅速増殖プロトコール(<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2978753/>、<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2305721/>)¹⁵³,¹

10

20

30

40

50

^{5 4} に供することができる。

【 0 4 7 0 】

患者試料が患者の T I L を含む実施形態では、以下の方法を用いて新生抗原特異的 T 細胞を治療量にまで増殖させることができる。一実施形態では、新生抗原特異的 T I L は、エクスピボで選別されたテトラマー/マルチマーとし、次いで選別された T I L を上記に述べたような迅速増殖プロトコールに供することができる。別の実施形態では、T I L の新生抗原非特異的増殖を行った後、新生抗原特異的 T I L をテトラマー選別し、その後、選別された T I L を上記に述べたような迅速増殖プロトコールに供することができる。別の実施形態では、T I L を迅速増殖プロトコールに供する前に抗原特異的培養を行うことができる (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4607110/>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/eji.201545849>) ^{1 5 5}, ^{1 5 6}。

10

【 0 4 7 1 】

いくつかの実施形態では、R o s e n b e r g 迅速増殖プロトコールを改変することができる。例えば、抗 P D 1 及び/または抗 4 1 B B を T I L 培養に加えてより迅速な増殖を刺激することができる (<https://jitc.biomedcentral.com/articles/10.1186/s40425-016-0164-7>) ^{1 5 7}。

【 0 4 7 2 】

X V I I . B . 新生抗原特異的 T 細胞の特定、特定された新生抗原特異的 T 細胞の T C R のシーケンシング、及びシーケンシングされた T C R の新たな T 細胞へのクローニング

20

患者に T 細胞療法を使用するための治療量の新生抗原特異的 T 細胞を作製するための第 2 の方法は、患者試料から新生抗原特異的 T 細胞を特定することと、特定された新生抗原特異的 T 細胞の T C R をシーケンシングすることと、シーケンシングされた T C R を新たな T 細胞をクローニングすることと、を含む。

【 0 4 7 3 】

最初に、新生抗原特異的 T 細胞を患者試料から特定し、特定された新生抗原特異的 T 細胞の T C R をシーケンシングする。T 細胞を単離することができる患者試料は、血液、リンパ節、または腫瘍のうちの 1 つ以上を含み得る。より詳細には、T 細胞を単離することができる患者試料は、末梢血単核細胞 (P B M C)、腫瘍浸潤細胞 (T I L)、解離腫瘍細胞 (D T C)、インビトロでプライミングされた T 細胞、及び/またはリンパ球から単離された細胞のうちの 1 つ以上を含み得る。これらの細胞は、新鮮なもの及び/または凍結されたものであってよい。P B M C 及びインビトロでプライミングされた T 細胞は、がん患者及び/または健康な対象から得ることができる。

30

【 0 4 7 4 】

患者試料が得られた後、試料を増殖させ、かつ/またはプライミングすることができる。さまざまな方法を実施して患者試料を増殖させ、プライミングすることができる。一実施形態では、新鮮な及び/または凍結した P B M C をペプチドまたはタンデムミニ遺伝子の存在下で刺激することができる。別の実施形態では、新鮮な及び/または凍結した単離された T 細胞をペプチドまたはタンデムミニ遺伝子の存在下で抗原提示細胞 (A P C) により刺激及びプライミングすることができる。A P C の例としては、B 細胞、単球、樹状細胞、マクロファージまたは人工抗原提示細胞 (関連する H L A 及び/または共刺激分子を提示する細胞またはビーズなど、<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929753> に概説されている)。別の実施形態では、P B M C、T I L、及び/または単離された T 細胞をサイトカインの存在下で刺激することができる (例えば、I L - 2、I L - 7、及び/または I L - 1 5)。別の実施形態では、T I L 及び/または単離された T 細胞を、最大刺激、サイトカイン (複数可)、及び/またはフィーダー細胞の存在下で刺激することができる。そのような実施形態では、T 細胞は、活性化マーカー及び/またはマルチマー (例えば、テトラマー) によって単離することができる。別の実施形態では、T I L 及び/または単離された T 細胞を、刺激性及び/または共刺激性マーカー (例えば、C D 3 抗体、C D 2 8 抗体) 及び/またはビーズ (例えば、D y n a B e a d s) で刺激す

40

50

ることができる。別の実施形態では、DTCを、リッチ培地中、高用量のIL-2でフィーダー細胞上で急速増殖プロトコルを用いて増殖させることができる。

【0475】

次に、新生抗原特異的T細胞を特定して単離する。いくつかの実施形態では、T細胞は、予め増殖させることなく患者試料からエキスピボで単離される。一実施形態では、セクションXVIに関連して上記に述べた方法を用いて患者試料から新生抗原特異的T細胞を特定することができる。代替的な実施形態では、単離は、ポジティブセレクションによって特定の細胞集団について濃縮するか、またはネガティブセレクションにより特定の細胞集団を枯渇させることによって行われる。いくつかの実施形態では、ポジティブまたはネガティブセレクションは、それぞれポジティブまたはネガティブに選択された細胞上で発現されるか（マーカー⁺）または比較的高いレベルで発現される（マーカー^高）1つ以上の表面マーカーに特異的に結合する1つ以上の抗体または他の結合物質と細胞を培養することによって行われる。

10

【0476】

いくつかの実施形態では、T細胞は、CD14などの、B細胞、単球、または他の白血球細胞などの非T細胞上で発現されるマーカーのネガティブセレクションによってPBMC試料から分離される。いくつかの態様では、CD4+またはCD8+の選択工程を用いてCD4+ヘルパー及びCD8+細胞傷害性T細胞を分離する。そのようなCD4+及びCD8+の集団は、1つ以上のナイーブ、メモリー、及び/またはエフェクターT細胞の亜集団上で発現される、または比較的高い程度で発現されるマーカーについてポジティブまたはネガティブセレクションによって亜集団にさらに分別することができる。

20

【0477】

いくつかの実施形態では、CD8+細胞は、それぞれの亜集団に関連付けられた表面抗原に基づいたポジティブまたはネガティブセレクションなどによって、ナイーブ、中枢メモリー、エフェクターメモリー、及び/または中枢メモリー幹細胞についてさらに濃縮されるか、または枯渇される。いくつかの実施形態では、中枢メモリーT細胞(TCM)細胞の濃縮を行うことで、例えば投与後の長期の生存、増殖、及び/または生着を改善させるなど、有効性が高められ、これはいくつかの態様で、かかる亜集団において特に安定的である(Terakura et al. (2012) Blood, 1:72-82; Wang et al. (2012) J Immunother, 35(9):689-701を参照)。いくつかの実施形態では、TCMを濃縮したCD8+T細胞とCD4+T細胞を組み合わせることでさらに有効性が高められる。

30

【0478】

複数の実施形態では、メモリーT細胞は、CD8+末梢血リンパ球のCD62L+及びCD62L-サブセットの両方に存在する。PBMCは、例えば抗CD8及び抗CD62L抗体を用いることなどにより、CD62L-CD8+及び/またはCD62L+CD8+画分を濃縮または枯渇させることができる。

【0479】

いくつかの実施形態では、中枢メモリーT(TCM)細胞の濃縮は、CD45RO、CD62L、CCR7、CD28、CD3、及び/またはCD127の陽性または高い表面発現に基づいたものであり、いくつかの態様では、CD45RA及び/またはグランザイムBを発現するか、または高度に発現する細胞のネガティブセレクションに基づいたものである。いくつかの態様では、TCM細胞について濃縮されたCD8+集団の単離は、CD4、CD14、CD45RAを発現する細胞の枯渇、及びCD62Lを発現する細胞のポジティブセレクションまたは濃縮によって行われる。一態様では、中枢メモリーT(TCM)細胞の濃縮は、CD4発現に基づいて選択された陰性画分から出発し、これにCD14及びCD45RAの発現に基づいてネガティブセレクションを、また、CD62Lに基づいてポジティブセレクションを行うことによって行われる。これらの選択はいくつかの態様では同時に行われ、他の態様では、いずれかの順序で順次行われる。いくつかの態様では、CD8+細胞集団または亜集団を調製するのに用いたのと同じCD4発現に基づ

40

50

く選択工程をさらに用いて、CD4+細胞集団または亜集団を生成することにより、CD4に基づいた選択からの陽性及び陰性画分が保持され、必要に応じて1つ以上のさらなるポジティブまたはネガティブセレクション工程を行った後、方法の次の工程で用いられる。

【0480】

特定の例では、PBMCの試料または他の白血球試料にCD4+細胞の選択を行う（陰性及び陽性画分の両方が保持される）。次に、陰性画分に、CD14及びCD45RAまたはROR1の発現に基づいたネガティブセレクションを、また、CD62LまたはCCR7などの中枢メモリーT細胞のマーカー特性に基づいたポジティブセレクションを行う（ポジティブセレクション及びネガティブセレクションはいずれかの順序で行う）。

10

【0481】

CD4+Tヘルパー細胞は、細胞表面抗原を有する細胞集団を特定することにより、ナイーブ、中枢メモリー、及びエフェクター細胞に分別される。CD4+リンパ球は標準的方法によって得ることができる。いくつかの実施形態では、ナイーブCD4+Tリンパ球は、CD45RO⁻、CD45RA⁺、CD62L⁺、CD4+T細胞である。いくつかの実施形態では、中枢メモリーCD4+細胞は、CD62L⁺かつCD45RO⁺である。いくつかの実施形態では、エフェクターCD4+細胞は、CD62L⁻かつCD45RO⁻である。

【0482】

一実施形態では、ネガティブセレクションによってCD4+細胞について濃縮するため、モノクローナル抗体カクテルは通常、CD14、CD20、CD11b、CD16、HLA-DR、及びCD8に対する抗体を含む。いくつかの実施形態では、抗体または結合パートナーは、磁性ビーズまたは常磁性ビーズのような固体支持体またはマトリックスに結合されることでポジティブ及び/またはネガティブセレクション用の細胞の分離を可能とする。例えば、いくつかの実施形態では、細胞及び細胞集団は免疫磁気（または親和性磁気）分離法を用いて分離または単離される（Methods in Molecular Medicine, vol. 58: Metastasis Research Protocols, Vol. 2: Cell Behavior In Vitro and In Vivo, p 17-25 Edited by: S.A. Brooks and U. Schumacher Humana Press Inc., Totowa, N.J.に概説されている）。

20

30

【0483】

いくつかの態様では、分離する細胞の試料または組成を、常磁性ビーズ（例えば、DynabeadsまたはMACSビーズなど）などの磁気応答性粒子または微粒子などの微小な磁化可能または磁気応答性材料とインキュベートする。磁気応答性材料、例えば粒子は、材料が分離することが望ましい、例えば材料がネガティブまたはポジティブに選択することが望ましい細胞、複数の細胞、または細胞の集団に存在する分子、例えば表面マーカーに特異的に結合する結合パートナー、例えば抗体に直接的または間接的に付着される。

【0484】

いくつかの実施形態では、磁性粒子またはビーズは、抗体または他の結合パートナーのような特定の結合要素に結合された磁気応答性材料を含む。磁気分離法に使用される多くの周知の磁気応答性材料が存在する。適当な磁性粒子としては、本明細書に参照により援用するところのMoldayによる米国特許第4,452,773号、及び欧州特許明細書EP 452342 B号に記載されるものがある。他の例としては、Owenによる米国特許第4,795,698号、及びLibertiらによる米国特許第5,200,084号に記載されるものなどのコロイド粒径の粒子がある。

40

【0485】

インキュベーションは、磁性粒子またはビーズに付着された、抗体または結合パートナーと特異的に結合する抗体もしくは結合パートナー、または分子（二次抗体または他の試

50

薬など)が、試料中の細胞上に存在する場合に細胞表面分子に特異的に結合するような条件下で一般的に行われる。

【0486】

いくつかの態様では、試料は磁場の中に置かれ、磁気応答性または磁化可能な粒子が付着された細胞が磁石に引き寄せられて非標識細胞から分離される。ポジティブセクションでは、磁石に引き寄せられた細胞が保持され、ネガティブセクションでは引き寄せられなかった細胞(非標識細胞)が保持される。いくつかの態様では、ポジティブセクションとネガティブセクションとの組み合わせが同じ選択工程において行われ、陽性及び陰性画分が保持され、さらに処理されるかまたはさらなる分離工程に供される。

【0487】

特定の実施形態では、磁気応答性粒子は、一次抗体もしくは他の結合パートナー、二次抗体、レクチン、酵素、またはストレプトアビジンでコーティングされる。特定の実施形態では、磁性粒子は1つ以上のマーカーに特異的な一次抗体のコーティングによって細胞に付着される。特定の実施形態では、ビーズではなく、細胞を一次抗体または結合パートナーで標識し、次いで、細胞型に特異的な二次抗体または他の結合パートナー(例えばストレプトアビジン)でコーティングした磁性ビーズを加える。特定の実施形態では、ストレプトアビジンでコーティングされた磁性ビーズがビオチン化された一次または二次抗体と組み合わせて使用される。

【0488】

いくつかの態様では、磁気応答性粒子は、後でインキュベート、培養、及び/または操作しようとする細胞に付着されたままとされ、いくつかの態様では、粒子は患者に投与される細胞に付着されたままとされる。いくつかの実施形態では、磁化可能な、または磁気応答性粒子は細胞から除去される。磁化可能な粒子を細胞から除去するための方法は周知のものであり、例えば、非標識抗体、磁化可能な粒子、または切断可能なリンカーと結合させた抗体などの使用が含まれる。いくつかの実施形態では、磁化可能な粒子は生分解性である。

【0489】

いくつかの実施形態では、親和性に基づく選択は、磁気活性化細胞選別(MACS)(Miltenyi Biotech, Auburn, Calif.)により行われる。磁気活性化細胞選別(MACS)システムは、磁化粒子が付着された細胞を高純度で選択することが可能である。いくつかの実施形態では、MACSは、外部磁場の印可後に非標的種と標的種が順次溶出されるモードで動作する。すなわち、磁化粒子に付着された細胞が保持される一方で付着していない種は溶出する。次いで、この第1の溶出工程が完了した後、磁場に捕捉され、溶出が防止された種が、溶出して回収できるような何らかの形で遊離される。特定の実施形態では、大型T細胞以外のT細胞が標識され、細胞の不均質集団から枯渇される。

【0490】

特定の実施形態では、単離または分離は、本方法の単離、細胞調製、分離、プロセッシング、インキュベーション、培養、及び/または配合工程のうちの1つ以上を行うシステム、機器、または装置を用いて行われる。いくつかの態様では、例えば、エラー、ユーザによる取り扱い、及び/または汚染を最小限に抑制するため、これらの工程のそれぞれを閉鎖または滅菌環境内で行うシステムが用いられる。1つの例では、システムは、国際公開第WO2009/072003号、または米国特許出願公開第US20110003380A1号である。

【0491】

いくつかの態様では、システムまたは装置は、単離、プロセッシング、及び配合工程のうちの1つ以上、例えばすべてを、統合された、または自給式システムで、及び/または自動化若しくはプログラム可能な形で行う。いくつかの態様では、システムまたは装置は、ユーザがプロセッシング、単離、操作、及び配合工程をプログラムし、制御し、その結果を評価し、及び/またはそのさまざまな側面を調節することを可能とする、システムまたは

10

20

30

40

50

装置と通信するコンピュータ及び/またはコンピュータプログラムを含む。

【0492】

いくつかの態様では、分離及び/または他の工程は、例えば、閉鎖及び滅菌システム内で臨床規模レベルの細胞の自動化分離を行うためのCliniMACSシステム(Miltenyi Biotec)を用いて行われる。構成要素としては、統合マイクロコンピュータ、磁気分離ユニット、蠕動ポンプ、及びさまざまなピンチバルブが含まれ得る。統合コンピュータは、いくつかの態様では、機器のすべての構成要素を制御し、システムに、繰り返し手順を標準化された順序で実行するように指示する。磁気分離ユニットは、いくつかの態様では、可動式永久磁石及び選択カラム用のホルダーを含む。蠕動ポンプはピンチバルブとともに、チューブセット全体にわたった流量を制御し、システムを通るパッ

10

【0493】

CliniMACSシステムは、いくつかの態様では、滅菌された非発熱性溶液中で供給される、抗体と結合された磁化可能な粒子を使用する。いくつかの実施形態では、磁性粒子で細胞を標識した後、細胞を洗浄して余分な粒子を除去する。次いで、細胞調製バッグをチューブセットに接続し、チューブセットをパッファーの入ったバッグと細胞回収バッグに接続する。チューブセットは、予備カラム及び分離カラムを含む予め組み立てられた滅菌チューブで構成され、1回使用のみである。分離プログラムの開始後、システムは細胞試料を分離カラムに自動的に適用する。標識細胞がカラムに保持されるのに対して、非標識細胞は一連の洗浄工程により除去される。いくつかの実施形態では、本明細書に記載される方法とともに使用される細胞集団は非標識細胞であり、カラムに保持されない。いくつかの実施形態では、本明細書に記載される方法とともに使用される細胞集団は標識細胞であり、カラムに保持される。いくつかの実施形態では、本明細書に記載される方法とともに使用される細胞集団は、磁場の除去後にカラムから溶出され、細胞回収バッグ内に回収される。

20

【0494】

いくつかの実施形態では、分離及び/または他の工程は、CliniMACS Prodigyシステム(Miltenyi Biotech)を使用して行われる。CliniMACS Prodigyシステムは、いくつかの実施形態では、遠心分離による細胞の自動洗浄及び分画化を可能とする細胞プロセッシングユニットを備えている。CliniMACS Prodigyシステムは、オンボードカメラ及びソース細胞産物のマクロスコピック層を判別することによって最適な細胞分画化エンドポイントを決定する画像認識ソフトウェアも含むことができる。例えば、末梢血を、赤血球、白血球、及び血漿層に自動的に分離することができる。CliniMACS Prodigyシステムは、例えば、細胞分化及び増殖、抗原ローディング、及び長期の細胞培養などの細胞培養プロトコルを行う統合型細胞培養チャンバも含むことができる。入力ポートによって培地の無菌的な除去及び補充が可能であり、統合型顕微鏡を使用して細胞を監視することができる(例えば、Klebanoff et al. (2012) J Immunother. 35(9): 651-660, Terakura et al. (2012) Blood. 1: 72-82, and Wang et al. (2012) J Immunother. 35(9): 689-701を参照)。

30

40

【0495】

いくつかの実施形態では、本明細書に記載される細胞集団は、複数の細胞表面マーカーについて染色された細胞が流体流で搬送されるフローサイトメトリーによって回収及び濃縮(または枯渇)される。いくつかの実施形態では、本明細書に記載される細胞集団は、分取スケール(FACS)ソーティングにより回収及び濃縮(または枯渇)される。特定の実施形態では、本明細書に記載される細胞集団は、FACSに基づいた検出システムと組み合わせた微小電気機械的システム(MEMS)チップの使用により回収及び濃縮(または枯渇)される(例えば、WO2010/033140, Cho et al. (2010) Lab Chip 10, 1567-1573; 及び Godin et al. (

50

2008) J Biophoton. 1(5): 355-376を参照)。いずれの場合も、細胞を複数のマーカーで標識することができ、明確に定義された高純度のT細胞サブセットの単離を可能とする。

【0496】

いくつかの実施形態では、ポジティブ及び/またはネガティブセレクションによる分離を容易とするため、抗体または結合パートナーは1つ以上の検出可能なマーカーで標識される。例えば、分離は、蛍光標識された抗体との結合に基づいたものとして行うことができる。いくつかの例では、1つ以上の細胞表面マーカーに対して特異的な抗体または他の結合パートナーの結合に基づいた細胞の分離は、例えばフローサイトメトリー検出システムと組み合わせた、分取スケール(FACS)及び/または微小電気機械的(MEMS)チップを含む、蛍光活性化セルソーティング(FACS)などにより、流体流で搬送される。かかる方法は、同時に複数のマーカーに基づいたポジティブ及びネガティブセレクションを可能とするものである。

10

【0497】

いくつかの実施形態では、調製法は、単離、インキュベーション、及び/または操作の前または後で細胞を凍結、例えば凍結保存する工程を含む。いくつかの実施形態では、凍結工程及びそれに続く解凍工程により、細胞集団中の顆粒球及びある程度の単球が除去される。いくつかの実施形態では、細胞は、血漿及び血小板を除去するために、例えば洗浄工程後に凍結溶液中に懸濁される。いくつかの態様では、各種の公知の凍結溶液及びパラメータのいずれをも使用することができる。1つの例では、20% DMSO及び8% ヒト血清アルブミン(HSA)を含有するPBSまたは他の適当な細胞凍結媒体を使用する。次いで、これを、媒体で1:1に希釈することにより、それぞれ、DMSO及びHSAの最終濃度を10%及び4%とする。他の例としては、Cryostor(登録商標)、CTL-Cryo(商標)ABC凍結媒などが挙げられる。その後、細胞を毎分1の速度で-80にまで凍結し、液体窒素貯蔵タンクの蒸気相中で保存する。

20

【0498】

いくつかの実施形態では、提供される方法には、培養(cultivation)、インキュベーション、培養(culture)、及び/または遺伝子操作の工程が含まれる。例えば、いくつかの実施形態では、枯渇させた細胞集団をインキュベート及び/または操作する方法、ならびに培養開始組成物が提供される。

30

【0499】

したがって、いくつかの実施形態では、細胞集団は培養開始組成物中でインキュベートされる。インキュベーション及び/または操作は、ユニット、チャンバー、ウェル、コラム、チューブ、チューブセット、バルブ、バイアル、培養皿、バッグ、または他の培養または細胞培養用の容器などの培養容器中で行われる。

【0500】

いくつかの実施形態では、細胞は、遺伝子操作に先立って、または遺伝子操作に関連してインキュベート及び/または培養される。インキュベーション工程は、培養(culture)、培養(cultivation)、刺激、活性化、及び/または増殖を含み得る。いくつかの実施形態では、組成物または細胞は、刺激条件または刺激剤の存在下でインキュベートされる。かかる条件としては、例えば、組換え抗原受容体を導入する目的などで、抗原曝露を模倣し、及び/または遺伝子操作を行うために細胞をプライミングするために、集団内の細胞の分裂、増殖、活性化、及び/または生存を誘導するように設計されたものが挙げられる。

40

【0501】

こうした条件には、特定の培地、温度、酸素含有量、二酸化炭素含有量、時間、作用物質、例えば、栄養素、アミノ酸、抗生物質、イオン、及び/または刺激因子、例えば、サイトカイン、ケモカイン、抗原、結合パートナー、融合タンパク質、組換え可溶性受容体、ならびに細胞を活性化するように設計された他の任意の作用物質のうちの1つ以上が含まれ得る。

50

【0502】

いくつかの実施形態では、刺激条件または刺激剤には、TCR複合体の細胞内シグナル伝達ドメインを活性化することができる1つ以上の作用物質、例えばリガンドが含まれる。いくつかの態様では、作用物質は、T細胞におけるTCR/CD3細胞内シグナル伝達カスケードをオンするか、または開始する。このような作用物質には、例えば、ビーズなどの固体支持体に結合した、例えば、抗CD3、抗CD28などの、TCR成分及び/または共刺激受容体に対して特異的な抗体などの抗体、及び/または1つ以上のサイトカインが含まれ得る。任意で、増殖方法は、抗CD3及び/または抗CD28抗体を（例えば、少なくとも約0.5 ng/mlの濃度で）培地に添加する工程をさらに含んでもよい。一部の態様において、刺激剤は、IL-2及び/またはIL-15、例えば、少なくとも約10単位/mlの濃度のIL-2が含まれる。

10

【0503】

いくつかの態様では、インキュベーションは、Riddellらへの米国特許第6,040,177号、Klebanoff et al. (2012) J Immunother. 35(9):651-660、Terakura et al. (2012) Blood. 1:72-82、及び/またはWang et al. (2012) J Immunother. 35(9):689-701に記載されるものなどの方法に従って行われる。

【0504】

いくつかの実施形態では、T細胞は、（例えば、結果として生じる細胞集団が、増殖させようとする初期集団中で各Tリンパ球について少なくとも約5個、10個、20個、または40個、またはそれよりも多いPBMCフィーダー細胞を含有するように）、例えば、非分裂末梢血単核球（PBMC）などのフィーダー細胞を培養開始組成物に添加し、培養物を（例えば、T細胞の数を増やすのに十分な時間にわたって）インキュベートすることによって増殖される。いくつかの態様では、非分裂フィーダー細胞は、線を照射したPBMCフィーダー細胞を含んでもよい。いくつかの実施形態では、PBMCに約3000~3600ラドの範囲の線を照射して細胞分裂を阻止する。いくつかの実施形態では、PBMCフィーダー細胞をマイトマイシンCで不活化する。いくつかの態様では、T細胞の集団を加える前にフィーダー細胞を培地に加える。

20

【0505】

いくつかの実施形態では、刺激条件は、ヒトTリンパ球の増殖に適した温度、例えば、少なくとも約25℃、一般的には少なくとも約30度、及び一般的には37℃または約37℃を含む。必要に応じて、インキュベーションは、フィーダー細胞として非分裂性のEBVで形質転換したリンパ芽球様細胞（LCL）を添加することをさらに含んでもよい。LCLに、約6000~10,000ラドの範囲の線を照射することができる。LCLフィーダー細胞はいくつかの態様では、LCLフィーダー細胞と初期Tリンパ球との比が少なくとも約10:1など、任意の適切な量で提供される。

30

【0506】

各実施形態において、抗原特異的T細胞、例えば、抗原特異的CD4+及び/またはCD8+T細胞は、ナイーブまたは抗原特異的なTリンパ球を抗原で刺激することによって得られる。例えば、サイトメガロウイルス抗原に対する抗原特異的T細胞株またはクローンを、感染させた対象からT細胞を単離し、この細胞をインビトロで同じ抗原で刺激することによって作製することができる。

40

【0507】

いくつかの実施形態では、新生抗原特異的T細胞は、機能性アッセイ（例えば、ELISPOT）による刺激後に特定及び/または単離される。いくつかの実施形態では、新生抗原特異的T細胞は、細胞内サイトカイン染色によって多機能性細胞によって単離される。いくつかの実施形態では、新生抗原特異的T細胞は、活性化マーカー（例えば、CD137、CD38、CD38/HLA-DR二重陽性、及び/またはCD69）を用いて特定及び/または単離される。いくつかの実施形態では、新生抗原特異的CD8+、ナチュ

50

ラルキラーT細胞、メモリーT細胞、及び/またはCD4+T細胞は、クラスIまたはクラスIIマルチマー及び/または活性化マーカーを用いて特定及び/または単離される。いくつかの実施形態では、新生抗原特異的CD8+及び/またはCD4+T細胞はメモリーマーカー（例えば、CD45RA、CD45RO、CCR7、CD27、及び/またはCD62L）を用いて特定及び/または単離される。いくつかの実施形態では、増殖中の細胞が特定及び/または単離される。いくつかの実施形態では、活性化されたT細胞が特定及び/または単離される。

【0508】

患者試料からの新生抗原特異的T細胞の単離後、特定された新生抗原特異的T細胞の新生抗原特異的TCRをシークエンシングする。新生抗原特異的TCRをシークエンシングするには、最初にTCRを特定する必要がある。T細胞の新生抗原特異的TCRを特定する1つの方法は、T細胞を少なくとも1種類の新生抗原を含むHLAマルチマー（例えばテトラマー）と接触させ、HLAマルチマーとTCRとの間の結合によりTCRを特定することを含むことができる。新生抗原特異的TCRを特定する別の方法は、TCRを有する1つ以上のT細胞を得ることと、1つ以上のT細胞を少なくとも1つの抗原提示細胞（APC）上に提示された少なくとも1種類の新生抗原により活性化することと、少なくとも1種類の新生抗原との相互作用により活性化された1つ以上の細胞を選択することによりTCRを特定することと、を含むことができる。

10

【0509】

新生抗原特異的TCRの特定後、TCRをシークエンシングすることができる。一実施形態では、セクションXVIに関連して上記に述べた方法を用いてTCRをシークエンシングすることができる。別の実施形態では、TCRのTCRa及びTCRbをバルクシークエンシングした後、頻度に基づいてペアリングすることができる。別の実施形態では、TCRは、Howie et al., Science Translational Medicine 2015 (doi: 10.1126/scitranslmed.aac5624)の方法を用いてシークエンシングし、ペアリングすることができる。別の実施形態では、TCRは、Han et al., Nat Biotech 2014 (PMID 24952902, doi 10.1038/nbt.2938)の方法を用いてシークエンシングし、ペアリングすることができる。別の実施形態では、ペアリングされたTCR配列を、<https://www.biorxiv.org/content/early/2017/05/05/134841> 及び/または<https://patents.google.com/patent/US20160244825A1/>により記載される方法を用いて得ることができる^{158, 159}。

20

30

【0510】

別の実施形態では、T細胞のクローン集団を、希釈を限定し、次いでT細胞のクローン集団のTCRa及びTCRbをシークエンシングすることによって得ることができる。さらに別の実施形態では、T細胞をウェルを有するプレート上でウェル1つ当たり1個のT細胞となるように選別し、次いで各T細胞のTCRa及びTCRbをシークエンシング及びペアリングすることができる。

【0511】

次に、新生抗原特異的T細胞を患者試料から特定し、特定された新生抗原特異的T細胞のTCRをシークエンシングした後、シークエンシングされたTCRを新たなT細胞にクローニングする。これらのクローニングされたT細胞は、新生抗原特異的受容体を含む（例えば、TCRを含む細胞外ドメインを含む）。そのような細胞の集団、及びそのような細胞を含む組成物も提供される。いくつかの実施形態では、組成物または集団は、そのような細胞について濃縮される（例えば、TCRを発現する細胞が、組成物中の全細胞、またはT細胞またはCD8+もしくはCD4+細胞などの特定の細胞の少なくとも1、5、10、20、30、40、50、60、70、80、90、91、92、93、94、95、96、97、98、99、または99%超を構成するように）。いくつかの実施形態では、組成物は、本明細書に開示されるTCRを有する少なくとも1つの細胞を含む。かかる組成物には、養子細胞療法などで投与するための医薬組成物及び製剤がある。細胞及

40

50

び組成物を対象、例えば患者に投与するための治療方法も提供される。

【0512】

したがって、TCR（複数可）を発現する遺伝子操作された細胞も提供される。細胞は、一般的に哺乳動物細胞などの真核細胞であり、通常はヒト細胞である。いくつかの実施形態では、細胞は、血液、骨髄、リンパ、またはリンパ器官に由来するものであり、自然免疫または獲得免疫などの免疫系の細胞である（例えば、リンパ球を含む骨髄細胞またはリンパ系細胞、典型的にはT細胞及びNK細胞）。他の例示的な細胞としては、人工多能性幹細胞（iPSC）を含む、多能性及び万能幹細胞などの幹細胞が含まれる。細胞は通常、対象から直接単離されるか、かつ/または対象から単離されて凍結されたものなどの初代細胞である。いくつかの実施形態では、細胞には、1つ以上のT細胞または全T細胞集団、CD4+細胞、CD8+細胞、ならびに機能、活性化状態、成熟度、分化能、増殖、再循環、局在化、及び/または持続能力、抗原特異性、抗原受容体の種類、特定の臓器またはコンパートメントにおける存在、マーカーまたはサイトカイン分泌プロファイル、及び/または分化の程度によって規定されるものなどのこれらの亜集団などの他の細胞タイプが含まれる。治療される対象に関連して、細胞は、同種及び/または自家であってよい。方法には、オフザシェルフ法が含まれる。オフザシェルフ法用などのいくつかの態様では、細胞は、例えば幹細胞、例えば人工多能性幹細胞（iPSC）である。いくつかの態様では、方法は、本明細書に記載されるようにして、対象から細胞を単離することと、それらの細胞を調製、プロセッシング、培養、及び/または操作することと、凍結保存の前または後でそれらの細胞を同じ患者に再導入することと、を含む。

10

20

【0513】

T細胞及び/またはCD4+及び/またはCD8+T細胞のサブタイプ及び亜集団としては、ナイーブT（TN）細胞、エフェクターT細胞（TEFF）、メモリーT細胞そのサブタイプ、例えば、幹細胞メモリーT（TSCM）、中枢メモリーT（TCM）、エフェクターメモリーT（TEM）、または最終分化したエフェクターメモリーT細胞、腫瘍浸潤リンパ球（TIL）、未成熟T細胞、成熟T細胞、ヘルパーT細胞、細胞傷害性T細胞、粘膜関連インバリアントT（MALT）細胞、天然及び適応制御性T（Treg）細胞、ヘルパーT細胞、例えばTH1細胞、TH2細胞、TH3細胞、TH17細胞、TH9細胞、TH22細胞、濾胞ヘルパーT細胞、 / T細胞、及び / T細胞がある。

30

【0514】

いくつかの態様において、細胞は、ナチュラルキラー（NK）細胞である。いくつかの態様において、細胞は、単球または顆粒球、例えば骨髄性細胞、マクロファージ、好中球、樹状細胞、マスト細胞、好酸球、及び/または好塩基球である。

【0515】

細胞は、内因性TCRの発現を低減するかまたはノックアウトするように遺伝子改変することができる。かかる改変は、本明細書に参照により援用するところのMol Ther Nucl Acids . 2012 Dec ; 1 (12) : e 6 3 ; Blood . 2011 Aug 11 ; 118 (6) : 1495 - 503 ; Blood . 2012 Jun 14 ; 119 (24) : 5697 - 5705 ; Torikai , Hiroki et al " HLA and TCR Knockout by Zinc Finger Nucleases : Toward " off - the - Shelf " Allogeneic T - Cell Therapy for CD19+ Malignancies . . " Blood 116 . 21 (2010) : 3766 ; Blood . 2018 Jan 18 ; 131 (3) : 311 - 322 . doi : 10 . 1182 / blood - 2017 - 05 - 787598 ; 及びWO2016069283に記載されている。

40

【0516】

細胞は、サイトカイン分泌を促進するために遺伝子改変することができる。かかる改変は、Hsu C、Hughes MS、Zheng Z、Bray RB、Rosenberg SA、Morgan RA、Primary human T lymphocytes engineered with a codon - optimized I

50

L - 15 gene resist cytokine withdrawal - induced apoptosis and persist long - term in the absence of exogenous cytokine . J Immunol . 2005 ; 175 : 7226 - 34 ; Quintarelli C , Vera JF , Savoldo B , Giordano Attianese GM , Pule M , Foster AE , Co - expression of cytokine and suicide genes to enhance the activity and safety of tumor - specific cytotoxic T lymphocytes . Blood . 2007 ; 110 : 2793 - 802 ; 及び、Hsu C , Jones SA , Cohen CJ , Zheng Z , Kerstann K , Zhou J , Cytokine - independent growth and clonal expansion of a primary human CD8+ T - cell clone following retroviral transduction with the IL - 15 gene . Blood . 2007 ; 109 : 5168 - 77 に記載されている。

10

【0517】

T細胞上のケモカイン受容体と腫瘍分泌ケモカインのミスマッチが、腫瘍微小環境内へのT細胞の不適切なトラフィックの主要因であることが示されている。治療効果を高めるため、腫瘍微小環境中でのケモカインの認識を高めるように遺伝子改変することができる。そのような改変の例は、Moon , EK Carpenito , CSun , JWang , LCKapoor , VPredina , J Expression of a functional CCR2 receptor enhances tumor localization and tumor eradication by retargeted human T - cells expressing a mesothelin - specific chimeric antibody receptor . Clin Cancer Res . 2011 ; 17 : 4719 - 4730 ; 及びCraddock , JALu , ABear , APule , MBrenner , MKRooney , CM et al . , Enhanced tumor trafficking of GD2 chimeric antigen receptor T - cells by expression of the chemokine receptor CCR2b . J Immunother . 2010 ; 33 : 780 - 788 に記載されている。

20

30

【0518】

細胞を、CD28及び41BBなどの共刺激/増強受容体の発現を増大させるために遺伝子改変することができる。

【0519】

T細胞療法の副作用には、サイトカイン放出症候群及び長期のB細胞枯渇が含まれ得る。レシピエント細胞への自殺/安全スイッチの導入によって細胞療法の安全性プロファイルを改善することができる。したがって、細胞を自殺/安全スイッチを含むように遺伝子改変することができる。自殺/安全スイッチは、遺伝子が発現される細胞に作用物質、例えば薬物に対する感度を与え、細胞が作用物質と接触する際に死なせる遺伝子であってよい。例示的な自殺/安全スイッチについては、Protein Cell . 2017 Aug ; 8 (8) : 573 - 589 に記載されている。自殺/安全スイッチはHSV - TKとすることができる。自殺/安全スイッチはシトシンデアミナーゼ、プリンヌクレオシドホスホリラーゼ、またはニトロレダクターゼであってもよい。自殺/安全スイッチは、米国特許出願公開第 US 20170166877A1号に記載されるRapaCIDE (商標) であってもよい。自殺/安全スイッチシステムは、Haematologica . 2009 Sep ; 94 (9) : 1316 - 1320 に記載されるCD20/リツキシマブであってもよい。これらの参考文献は、それらの全容を参照により援用するものである。

40

50

【0520】

TCRは、ヘテロ二量体化小分子の存在下でのみアSEMBルする分割受容体としてレシピエント細胞に導入することができる。かかるシステムについては、本明細書に参照により援用するところの *Science*, 2015 Oct 16; 350(6258): aab4077、及び米国特許第9,587,020号に記載されている。

【0521】

いくつかの実施形態では、細胞は1つ以上の核酸、例えば、本明細書に開示されるTCRをコードしたポリヌクレオチドを含み、かかるポリヌクレオチドは遺伝子操作により導入され、それにより本明細書に開示される組換えまたは遺伝子操作されたTCRを発現する。いくつかの実施形態では、核酸は、異種、すなわち、別の生物または細胞から得られたものなどの、細胞または細胞から得られる試料中に通常は存在していないものであって、例えば、操作される細胞及び/またはかかる細胞が由来する生物には通常みられないものである。いくつかの実施形態では、核酸は、複数の異なる細胞タイプからの異なるドメインをコードした核酸のキメラ複合体を含むものなど、自然界にはみられない核酸など、天然に存在しないものである。

10

【0522】

核酸は、コドン最適化されたヌクレオチド配列を含むことができる。特定の理論または機序に束縛されるものではないが、ヌクレオチド配列のコドン最適化はmRNA転写産物の翻訳効率を高めるものと考えられる。ヌクレオチド配列のコドン最適化では、天然のコドンを、同じアミノ酸をコードしているが細胞内でより利用されやすいtRNAによって翻訳されうる別のコドンに置換することにより、翻訳効率を高めることができる。ヌクレオチド配列の最適化は、翻訳を妨げる二次mRNA構造を低減させることで翻訳効率を高めることもできる。

20

【0523】

コンストラクトまたはベクターを用いてレシピエント細胞にTCRを導入することができる。例示的なコンストラクトを本明細書に記載する。TCRの鎖及び鎖をコードするポリヌクレオチドは、単一のコンストラクト内であってもよく、別々のコンストラクト内であってもよい。鎖及び鎖をコードするポリヌクレオチドは、プロモーター、例えば異種プロモーターと機能的に連結することができる。異種プロモーターは、例えば、EF1、CMV、PGK1、Ubc、アクチン、CAGプロモーターのような強力なプロモーターとすることができる。異種プロモーターは、弱いプロモーターであってもよい。異種プロモーターは誘導性プロモーターであってもよい。例示的な誘導性プロモーターとしては、これらに限定されるものではないが、TRE、NFAT、GAL4、LACなどが挙げられる。他の例示的な誘導性発現系は、本明細書にその全容を援用するところの米国特許第5,514,578号、同第6,245,531号、同第7,091,038号、及び欧州特許第0517805号に記載されている。

30

【0524】

TCRをレシピエント細胞に導入するためのコンストラクトは、シグナルペプチドをコードしたポリヌクレオチド(シグナルペプチドエレメント)を含んでもよい。シグナルペプチドは、導入されたTCRの表面トラフィックを促進することができる。例示的なシグナルペプチドとしては、これらに限定されるものではないが、CD8シグナルペプチド、免疫グロブリンシグナルペプチドが挙げられ、具体例としてはGM-CSF及びIgGが挙げられる。かかるシグナルペプチドは、本明細書に参照により援用するところの *Trends Biochem Sci*, 2006 Oct; 31(10): 563-71. Epub 2006 Aug 21; 及び An, et al. "Construction of a New Anti-CD19 Chimeric Antigen Receptor and the Anti-Leukemia Function Study of the Transduced T-cells." *Oncotarget* 7.9(2016): 10638-10649. PMC. Web. 16 Aug. 2018に記載されている。

40

50

【0525】

場合により、例えば、鎖及び鎖が単一のコンストラクトまたはオープンリーディングフレームから発現される場合、またはコンストラクトにマーカ遺伝子が含まれるような場合では、コンストラクトはリボソームスキップ配列を含むことができる。リボソームスキップ配列は、2Aペプチド、例えば、P2AまたはT2Aペプチドであってよい。例示的なP2AまたはT2Aペプチドは、本明細書に参照によりその全容を援用するところのScientific Reports volume 7, Article number: 2193 (2017)に記載されている。場合により、FURIN/PACE切断部位が2Aエレメントの上流に導入される。FURIN/PACE切断部位は、例えば、<http://www.nuolan.net/substrates.html>に記載されている。切断ペプチドは、第Xa因子の切断部位であってもよい。鎖及び鎖が単一のコンストラクトまたはオープンリーディングフレームから発現される場合、コンストラクトは配列内リボソーム進入部位(IRES)を含むことができる。

10

【0526】

コンストラクトは、1つ以上のマーカ遺伝子をさらに含むことができる。例示的なマーカ遺伝子としては、これらに限定されるものではないが、GFP、ルシフェラーゼ、HA、lacZが挙げられる。マーカは、当業者には周知の抗生物質耐性マーカ、重金属耐性マーカ、または殺生物剤耐性マーカなどの選択可能なマーカであってよい。マーカは、栄養要求性宿主で使用するための補完的マーカとすることができる。例示的な補完的マーカ及び栄養要求性宿主は、Gene. 2001 Jan 24; 263 (1-2): 159-69に記載されている。かかるマーカは、IRES、フレームシフト配列、2Aペプチドリンカー、TCRとの融合体によって発現させるか、または別のプロモーターから別に発現させることができる。

20

【0527】

TCRをレシピエント細胞に導入するための例示的なベクターとしては、これらに限定されるものではないが、アデノ随伴ウイルス、アデノウイルス、アデノウイルス+改変ワクシニア、アンカラウイルス(MVA)、アデノウイルス+レトロウイルス、アデノウイルス+センダイウイルス、アデノウイルス+ワクシニアウイルス、アルファウイルス(VEE)レプリコンワクチン、アンチセンスオリゴヌクレオチド、ピフィドバクテリウム・ロンガム、CRISPR-Cas9、E.coli、フラビウイルス、遺伝子銃、ヘルペスウイルス、単純ヘルペスウイルス、ラクトコッカス・ラクティス、電気穿孔法、レンチウイルス、リポフェクション、リステリア・モノサイトゲネス、麻疹ウイルス、改変ワクシニアアンカラウイルス(MVA)、mRNA電気穿孔法、ネイキッド/プラスミドDNA、ネイキッド/プラスミドDNA+アデノウイルス、ネイキッド/プラスミドDNA+改変ワクシニアアンカラウイルス(MVA)、ネイキッド/プラスミドDNA+RNAトランスファー、ネイキッド/プラスミドDNA+ワクシニアウイルス、ネイキッド/プラスミドDNA+水泡口炎ウイルス、ニューカッスル病ウイルス、非ウイルス性PiggyBac(商標)(PB)トランスポゾン、ナノ粒子ベースのシステム、ポリオウイルス、ポックスウイルス、ポックスウイルス+ワクシニアウイルス、レトルウイルス、RNAトランスファー、RNAトランスファー+ネイキッド/プラスミドDNA、RNAウイルス、サッカロマイセス・セレピシエ、サルモネラ・ティフィムリウム、セムリキ森林ウイルス、センダイウイルス、志賀赤痢菌、シミアンウイルス、siRNA、スリーピング・ビューティートランスポゾン、ストレプトコッカス・ミュータンス、ワクシニアウイルス、ベネズエラウマ脳炎ウイルスレプリコン、水泡口炎ウイルス、及びコレラ菌が挙げられる。

30

40

【0528】

好ましい実施形態では、TCRは、アデノ随伴ウイルス(AAV)、アデノウイルス、CRISPR-CAS9、ヘルペスウイルス、レンチウイルス、リポフェクション、mRNA電気穿孔法、PiggyBac(商標)(PB)トランスポゾン、レトロウイルス、RNAトランスファー、またはスリーピング・ビューティートランスポゾンによってレシ

50

ピエント細胞に導入される。

【0529】

いくつかの実施形態では、TCRをレシピエント細胞に導入するためのベクターは、ウイルスベクターである。例示的なウイルスベクターとしては、アデノウイルスベクター、アデノ随伴ウイルス(AAV)ベクター、レンチウイルスベクター、ヘルペスウイルスベクター、レトロウイルスベクターなどが挙げられる。かかるベクターは本明細書に記載される。

【0530】

TCRをレシピエント細胞に導入するためのTCRコンストラクトの例示的な実施形態を図25に示す。いくつかの実施形態では、TCRコンストラクトは、5'~3'の方向に以下のポリヌクレオチド配列、すなわち、プロモーター配列、シグナルペプチド配列、TCR可変(TCR_v)配列、TCR定常(TCR_c)配列、切断ペプチド(例えば、P2A)、シグナルペプチド配列、TCR可変(TCR_v)配列、及びTCR定常(TCR_c)配列を含む。いくつかの実施形態では、コンストラクトのTCR_c及びTCR_c配列は、1つ以上のマウス領域、例えば、本明細書に記載される完全マウス定常配列またはヒトマウスへのアミノ酸交換を含む。いくつかの実施形態では、コンストラクトはさらに、TCR_c配列の3'末端、切断ペプチド配列(例えば、T2A)に続くレポーター遺伝子を含む。一実施形態では、コンストラクトは、5'~3'の方向に以下のポリヌクレオチド配列、すなわち、プロモーター配列、シグナルペプチド配列、TCR可変(TCR_v)配列、1つ以上のマウス領域を含むTCR定常(TCR_c)配列、切断ペプチド(例えば、P2A)、シグナルペプチド配列、TCR可変(TCR_v)配列、及び1つ以上のマウス領域を含むTCR定常(TCR_c)配列、切断ペプチド配列(例えば、T2A)、及びレポーター遺伝子を含む。

10

20

【0531】

図26は、治療法の開発のためにTCRを発現系にクローニングするための例示的なP526コンストラクト骨格のヌクレオチド配列を示す。

【0532】

図27は、治療法の開発のために患者の新生抗原特異的TCRのクロナタイプ1を発現系にクローニングするための例示的なコンストラクト配列を示す。

【0533】

図28は、治療法の開発のために患者の新生抗原特異的TCRのクロナタイプ3を発現系にクローニングするための例示的なコンストラクト配列を示す。

30

【0534】

TCRをコードする単離された核酸、かかる核酸を含むベクター、及びかかるベクターを含む宿主細胞、ならびにTCRを作製するための組換え法も提供される。

【0535】

核酸は組換え体であってよい。組換え核酸は、天然または合成核酸セグメントを、生細胞内で複製可能な核酸分子またはその複製産物に連結することにより生細胞の外部で構築することができる。本明細書の目的では、複製はインビトロ複製またはインビボ複製であってよい。

40

【0536】

TCRの組換え作製を行うには、TCRをコードする核酸(複数可)を単離してさらなるクローニング(すなわち、DNAの増幅)または発現を行うための複製可能なベクターに挿入することができる。いくつかの態様では、核酸は、例えば本明細書に参照により援用するところの米国特許第5,204,244号に記載されるような相同組み換えによって作製することができる。

【0537】

多くの異なるベクターが当該技術分野では周知である。ベクターの構成要素は一般的に以下のもの、すなわち、例えば、本明細書にその全容を参照により援用するところの米国特許第5,534,615号に記載されるような、シグナル配列、複製起点、1つ以上の

50

マーカー遺伝子、エンハンサーエレメント、プロモーター、及び転写終結配列のうちの1つ以上を含む。

【0538】

TCR、抗体、またはその抗原結合フラグメントを発現させるのに適した例示的なベクターまたはコンストラクトとしては、例えば、pUCシリーズ (Fermentas Life Sciences)、pBluescriptシリーズ (Stratagene, La Jolla, CA)、pETシリーズ (Novagen, Madison, WI)、pGEXシリーズ (Pharmacia Biotech, Uppsala, Sweden)、及びpEXシリーズ (Clontech, Palo Alto, CA)が挙げられる。AGT10、AGT11、AZapII (Stratagene)、AEMBL4、及びANM1149などのバクテリオファージも本明細書に開示されるTCRを発現させるうえで適当である。

10

【0539】

XVII. 治療の概要のフローチャート

図29は、一実施形態による、患者に個別化された新生抗原特異的治療を行うための方法のフローチャートである。他の実施形態では、本方法は、図29に示されるもの以外の異なる及び/またはさらなる工程を含んでもよい。さらに、本方法の各工程は、異なる実施形態において、図29に関連して述べられる順序とは異なる順序で行うことができる。

【0540】

提示モデルを、上記に述べた質量分析データを用いて訓練する(2901)。患者試料を取得する(2902)。いくつかの実施形態では、患者試料は、腫瘍生検及び/または患者の末梢血を含む。工程2902で得られた患者試料をシーケンシングして、患者試料からの腫瘍抗原ペプチドが提示される尤度を予測するための提示モデルに入力するデータを特定する。工程2902で得られた患者試料からの腫瘍抗原ペプチドの提示尤度を、訓練された提示モデルを用いて予測する(2903)。予測された提示尤度に基づいて患者に対する治療抗原を特定する(2904)。次いで、別の患者試料を取得する(2905)。患者試料は、患者の末梢血、腫瘍浸潤リンパ球(TIL)、リンパ、リンパ節細胞、及び/または他のT細胞源を含み得る。工程2905で得られた患者試料を新生抗原特異的T細胞についてインビボでスクリーニングする(2906)。

20

【0541】

治療プロセスのこの時点において、患者は、T細胞療法及び/またはワクチン治療を受けることができる。ワクチン治療を受けるには、患者のT細胞が特異的である新生抗原を特定する(2914)。次いで、特定された新生抗原を含むワクチンを作製する(2915)。最後に、ワクチンを患者に投与する(2916)。

30

【0542】

T細胞治療を受けるには、新生抗原特異的T細胞を増殖させるか、及び/または新たな新生抗原特異的T細胞を遺伝子操作する。T細胞療法で使用するための新生抗原特異的T細胞を増殖させるには、細胞を単純に増殖させ(2907)、患者に注入する(2908)。

【0543】

T細胞療法用の新たな新生抗原特異的T細胞を遺伝子操作するには、インビボで特定された新生抗原特異的T細胞のTCRをシーケンシングする(2909)。次いで、これらのTCR配列を発現ベクターにクローニングする(2910)。次にこの発現ベクター2910を新たなT細胞にトランスフェクトする(2911)。トランスフェクトしたT細胞を増殖させる(2912)。そして最後に、増殖させたT細胞を患者に注入する(2913)。

40

【0544】

患者はT細胞療法及びワクチン療法の両方を受けることができる。一実施形態では、患者は最初にワクチン療法を受け、その後、T細胞療法を受ける。このアプローチの利点の1つは、ワクチン療法により、検出可能なレベルのT細胞によって認識される腫瘍特異的

50

T細胞の数及び新生抗原の数を増やすことができる点である。

【0545】

別の実施形態では、患者はT細胞療法を受けた後、ワクチン療法を受けることができ、その場合、ワクチンに含まれるエピトープのセットは、T細胞療法のターゲットとなるエピトープのうちの一つ以上を含む。このアプローチの利点の一つは、ワクチンの投与によって治療T細胞の増殖及び持続を促進することができる点である。

【0546】

XIX. 例示的なコンピュータ

図30は、図1及び図3に示した実体を実施するための例示的なコンピュータ3000を説明する。コンピュータ3000は、チップセット3004に連結された少なくとも一つのプロセッサ3002を含む。チップセット3004は、メモリコントローラハブ3020及び入力/出力(I/O)コントローラハブ3022を含む。メモリ3006及びグラフィックスアダプタ3012は、メモリコントローラハブ3020に連結されており、ディスプレイ3018は、グラフィックスアダプタ3012に連結されている。記憶デバイス3008、入力装置3014、及びネットワークアダプタ3016は、I/Oコントローラハブ3022に連結されている。コンピュータ3000の他の実施形態は、異なるアーキテクチャを有する。

【0547】

記憶デバイス3008は、ハードドライブ、コンパクトディスク読み出し専用メモリ(CD-ROM)、DVD、またはソリッドステートメモリ装置などの、非一時的なコンピュータ可読の記憶媒体である。メモリ3006は、プロセッサ3002によって使用される命令及びデータを保持する。入力インターフェイス3014は、タッチスクリーンインターフェイス、マウス、トラックボール、もしくは他のタイプのポインティングデバイス、キーボード、またはそれらのいくつかの組み合わせであり、データをコンピュータ3000中に入力するために使用される。いくつかの実施形態において、コンピュータ3000は、ユーザーからのジェスチャーを介して、入力インターフェイス3014からの入力(例えば、コマンド)を受け取るように構成されていてもよい。グラフィックスアダプタ3012は、ディスプレイ3018上に画像及び他の情報を表示する。ネットワークアダプタ3016は、コンピュータ3000を、一つ以上のコンピュータネットワークに連結する。

【0548】

コンピュータ3000は、本明細書に記載した機能性を提供するためのコンピュータプログラムモジュールを遂行するように適合している。本明細書において使用される場合、「モジュール」という用語は、特定の機能性を提供するために使用されるコンピュータプログラム論理を指す。したがって、モジュールは、ハードウェア、ファームウェア、及び/またはソフトウェアにおいて実行されることができる。一実施形態では、プログラムモジュールは、記憶デバイス3008に保存され、メモリ3006中にロードされ、プロセッサ3002によって遂行される。

【0549】

図1の実体によって使用されるコンピュータ3000のタイプは、実体によって必要とされる実施形態及びプロセッシングパワーに応じて変動することができる。例えば、提示特定システム160は、単一のコンピュータ3000、または、例えばサーバーファームにおいてネットワークを通して互いに通信する複数のコンピュータ3000において、起動することができる。コンピュータ3000は、グラフィックスアダプタ3012及びディスプレイ3018などの、上記の構成要素のうちいくつかを欠いてもよい。

【0550】

補足の表1

既存の応答による予測された変異のランク

10

20

30

40

変異 ID	患者 ID	MHCFlurry, TPM > 0	MHCFlurry, TPM > 1	MHCFlurry, TPM > 2	ペプチドMS モデル, TPM > 1	完全MS モデル
KARS_D356H	3942	81	44	36	26	5
NUP98_A359D	3942	13	8	7	0	0
CASP8_F67V	3971	13	3	2	3	1
KRAS_G12D	3995	36	21	18	2	2
RNF213_N1702S	3995	0	0	0	7	7
TUBGCP2_P293L	3995	2	2	2	8	6
H3F3B_A48T	4007	33	23	21	13	0
SKIV2L_R653H	4007	2	1	1	15	17
API5_R243Q	4032	52	31	27	10	1
PHLPP1_G566E	4032	54	33	29	72	67
RNF10_E572K	4032	43	23	22	46	46
ZFYVE27_R6H	4069	35	23	22	0	0
CADPS2_R1266H	4136	23	22	22	4	5
KIAA0368_S186F	4136	2	2	2	1	0
FLNA_R2049C	NCI-3784	91	85	81	31	5
KIF16B_L1009P	NCI-3784	22	21	19	74	69
SON_R1927C	NCI-3784	37	35	32	105	83
KIF1BP_P246S	NCI-3903	66	35	32	22	7
MAGEA6_E168K	NCI-3998	15	10	9	1	0
MED13_P1691S	NCI-3998	5	3	2	0	1
PD55A_Y1000F	NCI-3998	13	8	7	6	4
CDK4_R71L	患者1	56	23	20	5	0
DNAH17_H8302Y	患者1	42	80	59	112	77
GCN1_L2330P	患者1	59	25	22	3	1
BRWD1_R925W	患者2	80	62	58	74	75
PARG_Y427N	患者2	88	69	65	51	49
	中央値	35.5	23	21.5	9	5

【 0 5 5 1 】

補記の表 2

N S C L S 患者の人口統計

患者 ID	年齢範囲 (歳)	性別	人種	最初の(肺癌) 診断年	腫瘍ステージ (組入れ時)	原発性腫瘍の部位	組織学的タイプ	全身性 NSCLC-標的療法	現在の抗-PD(L)-1療法	HLA-A
1-001-002	81-90	男性	白人	2010	IIIB	肺	扁平上皮	カルボプラチン	ニボルマブ	A*01:01
1-024-001	81-90	男性	白人	2016	IV	肺	肉腫様肺癌	カルボプラチン, ドセタキセル, ペバシズマブ, ラムシルマブ, ペメトレキセド, ニナトリウム	ベムプロリズマブ	A*32:01
1-024-002	51-60	女性	白人	2016	IV	肺	腺癌	カルボプラチン, ドセタキセル, ペバシズマブ, ラムシルマブ, ペメトレキセド, ニナトリウム	ニボルマブ	A*68:01
1-038-001	61-70	男性	白人	2016	IV	肺	腺癌	ペメトレキセド, シスプラチン	ニボルマブ	A*69:01
1-050-001	71-80	女性	白人	2015	IIIB	肺	腺癌	エトボシド, シスプラチン	ニボルマブ	A*29:02
CU05	71-80	女性	白人	2013	IV	肺	肺扁平上皮	カルボプラチン+ペメトレキセド	ニボルマブ	A*24:02
CU04	61-70	女性	ヒスパニックまたはラテン系	2013	I	肺	腺癌		デュルバルマブ + トレメリムマブ	A*24:26
CU03	61-70	男性	黒人またはアフリカ系アメリカ人	2016	I	肺	肺扁平上皮		n/a	A*23:01
CU02	61-70	男性	白人	2016	I	肺	肺扁平上皮	カルボプラチン+ ゲムシタビン	n/a	A*02:01

10

20

30

40

HLA-A	HLA-B	HLA-B	HLA-C	HLA-C	HLA-C	発現された変異	非同義変異	正常DNAエクソンカバレッジ中央値	腫瘍DNAエクソンカバレッジ中央値	RNA PF固有リード (M)	既知のドライバー変異	可能性の高いドライバー変異	VAF中央値
A*01:01	B*08:01	B*51:01	C*01:02	C*01:02	C*07:01	122	232	145	552	173	KRAS_G12D, TP53_R213*	STK11_G52fs	0.22
A*03:01	B*27:05	B*27:05	C*02:02	C*02:02	C*02:02	83	143	165	508	131.9	KRAS_G12C, TP53_R280T	PML_E43*, NF2_R341*	0.093
A*68:01	B*40:02	B*40:27	C*03:04	C*03:04	C*03:04	38	69	190	454	114.4	KRAS_G12S, TP53_Q331*	STK11_E198*	0.182
A*01:02	B*41:01	B*49:01	C*17:01	C*17:01	C*07:01	158	265	158	983	311.8	KRAS_G12V	KDM5C_E303*	0.19
A*26:01	B*44:03	B*07:05	C*16:01	C*16:01	C*15:05	53	92	117	556	119			0.059
A*68:02	B*14:02	B*15:17	C*07:01	C*07:01	C*08:02	65	109	191	448	83.6			0.095
A*26:01	B*18:01	B*38:01	C*12:03	C*12:03	C*12:03	336	511	213	552	240.4	TP53_R158G	NFKBIE_G41fs, CDH1_Q346*, NF1_D2163fs, MED12_R730*	0.224
A*01:01	B*08:01	B*15:03	C*01:02	C*01:02	C*12:03	105	187	114	830	182.1			0.242
A*03:01	B*07:02	B*57:01	C*07:02	C*07:02	C*06:02	102	174	105	738	185.3	TP53_R175H	ATR_Q195*	0.32

10

20

30

40

50

補足の表 3

NSCLC患者においてT細胞認識について試験したペプチド

患者	ペプチド	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	IMHCFlurryによりカバーされる最も高確率の制限
1-001-002	HSPFATSL	N		1-001-002_pool_1	N	chr15_28215653_C_A	snp	HERC2	A2060S	41.9	HLA-C*01:02	0	95	5169.68205	偽
1-001-002	DPEEVIVTV	N		1-001-002_pool_1	N	chr17_59680958_C_T	snp	CLTC	S989L	272.1	HLA-B*51:01	1	61	3455.25069	真
1-001-002	ELDPDQLEY	N		1-001-002_pool_1	N	chr13_30210371_C_A	snp	KATNAL1	D407Y	12.81	HLA-A*01:01	2	1	24.2177849	真
1-001-002	TPLTKDVTL	N		1-001-002_pool_1	N	chr5_78100974_A_T	snp	AP3B1	S817T	44.4	HLA-B*08:01	3	2	48.9740194	真
1-001-002	DGVGKSAL	N		1-001-002_pool_1	N	chr12_25243550_C_T	snp	KRAS	G12D	40.75	HLA-B*08:01	4	89	4714.29522	真
1-001-002	YTTVRALTL	N		1-001-002_pool_1	N	chr17_28339664_G_T	snp	TNFAIP1	R48L	45.62	HLA-B*08:01	5	26	973.417701	真
1-001-002	TPSAAVKLI	N		1-001-002_pool_1	N	chr15_81319417_T_C	snp	STARSD5	M108V	1.95	HLA-B*51:01	6	39	2030.48603	真
1-001-002	WPVILLNV	N		1-001-002_pool_1	N	chr3_179025167_AAC_A	del_fs	ZMAT3	V240fs	14.99	HLA-B*51:01	7	16	600.564752	真
1-001-002	ELNARRCSF	N		1-001-002_pool_1	N	chr18_79943341_G_A	snp	PQLC1	R109C	33.89	HLA-B*08:01	8	5	62.0439997	真
1-001-002	QMKNPILFL	N		1-001-002_pool_1	N	chr9_127663287_G_T	snp	STXBP1	R171L	38.76	HLA-B*08:01	9	20	674.64733	真
1-001-002	LTEKVSLLK	N		1-001-002_pool_2	N	chr9_92719180_C_T	snp	BICD2	E489K	42.66	HLA-A*01:01	10	10	428.744925	真
1-001-002	SPETATSL	N		1-001-002_pool_2	N	chr15_28215653_C_A	snp	HERC2	A2060S	41.9	HLA-B*08:01	11	4	59.1155419	真
1-001-002	NVDMRTISF	N		1-001-002_pool_2	N	chr9_121353262_I_A	snp	STOM	K93N	360.6	HLA-B*08:01	12	30	1490.72261	真
1-001-002	TSIVVQTL	N		1-001-002_pool_2	N	chr4_39205691_C_T	snp	WDR19	A282V	18.12	HLA-B*08:01	13	176	9862.33009	真
1-001-002	HIKIEPVAL	N		1-001-002_pool_2	N	chr13_73062087_C_T	snp	KIF5	T163I	25.77	HLA-B*08:01	14	27	1122.27455	真
1-001-002	DSPDGSNGL	N		1-001-002_pool_2	N	chr20_44197575_C_T	snp	OSER1	S119N	20.7	HLA-C*01:02	15	471	21598.414	偽
1-001-002	YTAVHYAASY	N		1-001-002_pool_2	N	chr12_56248788_C_A	snp	ANKRD52	A559S	18.32	HLA-A*01:01	16	0	11.5906737	真
1-001-002	VGADGVGKSAL	N		1-001-002_pool_2	N	chr12_25243550_C_T	snp	KRAS	G12D	40.75	HLA-C*01:02	17	370	17985.3612	偽
1-001-002	MMPPPLPGI	N		1-001-002_pool_2	N	chr17_32369404_A_T	snp	ZNF207	Q409L	186.04	HLA-B*51:01	18	136	7609.76602	真

患者	ペプチド	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
1-001-002	FPVPGMTNQ	N		1-001-002_pool_2	N	chr5_109186272_G_T	snp	FER	C759F	67.36	HLA-B*51:01	19	38	1999.07208	真
1-024-001	VTHHAPLSW	N		1-024-001_pool_1	Y	chr3_125523270_C_A	snp	OS9PL11	G489W	24.12	HLA-A*32:01	0	7	77.009026	真
1-024-001	GTKKDVLIK	Y		1-024-001_pool_1	Y	chr20_56513366_G_A	snp	RTFDC1	E177K	61.32	HLA-A*03:01	1	70	2168.51668	真
1-024-001	GLNVPVQSNK	N		1-024-001_pool_1	Y	chr4_88390868_G_T	snp	HERC6	R218L	8.7	HLA-A*03:01	2	4	59.675168	真
1-024-001	VVVGACGVGK	N		1-024-001_pool_1	Y	chr12_25243351_C_A	snp	KRAS	G12C	40.05	HLA-A*03:01	3	11	133.648023	真
1-024-001	AQFAGKQIY	N		1-024-001_pool_1	Y	chr9_89045819_C_A	snp	SHC3	E376D	8.88	HLA-A*32:01	4	91	3715.42819	真
1-024-001	KVLPDVTISY	N		1-024-001_pool_1	Y	chr3_48591778_G_T	snp	COL7A1	R468S	25.42	HLA-A*32:01	6	85	3234.15772	真
1-024-001	MILMKNISTK	N		1-024-001_pool_1	Y	chr12_6959976_G_A	snp	PTPN6	E471K	105.39	HLA-A*03:01	7	0	12.2301919	真
1-024-001	DLAGGTFD	N		1-024-001_pool_1	Y	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-B*27:05	9	353	18290.7955	真
1-024-001	LIFDLAGGTF	N		1-024-001_pool_1	Y	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-C*02:02	11	57	1716.74204	偽
1-024-001	INVLFDLA	N		1-024-001_pool_1	Y	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-A*32:01	17	621	27984.1357	真
1-024-001	VVGACGVGK	N		1-024-001_pool_2	N	chr12_25243351_C_A	snp	KRAS	G12C	40.05	HLA-A*03:01	5	19	197.846108	真
1-024-001	VIMLNGTKK	N		1-024-001_pool_2	N	chr20_56513366_G_A	snp	RTFDC1	E177K	61.32	HLA-A*03:01	8	10	122.750322	真
1-024-001	LAGGTFD	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-C*02:02	10	632	28384.8834	偽
1-024-001	LRNSGGEVF	N		1-024-001_pool_2	N	chr14_80906012_TC_T	del_fs	CEP128	R102fs	11.31	HLA-B*27:05	12	46	1020.95087	真
1-024-001	VVLPDVTISY	N		1-024-001_pool_2	N	chr3_48591778_G_T	snp	COL7A1	R468S	25.42	HLA-A*32:01	13	62	1925.29397	真
1-024-001	IFDLAGGTF	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-C*02:02	14	427	21255.2074	偽
1-024-001	GLLDEAKRLLY	N		1-024-001_pool_2	N	chr19_57575861_G_T	snp	ZNF416	Q49K	11.89	HLA-A*03:01	15	24	354.82068	真
1-024-001	SVLLPENYTK	N		1-024-001_pool_2	N	chr11_122789248_G_T	snp	UBASH3B	G307V	12.11	HLA-A*03:01	16	23	228.127132	真

患者	ペプチド	個々のペプチド反応 (任意の時点)	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	変異遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFIurryのランク	IMHCFIurry (nM)	IMHCFIurryによりカバーされる最も高確率の制限
1-024-001	DLAGGTFDVS	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-A*32:01	18	487	23357.3292	真
1-024-001	IFDLAGGTFDV	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-C*02:02	19	563	25887.4267	偽
1-024-002	AEWRNGSTSSL	N		1-024-002_pool_1	Y	chr3_122703943_C_G	snp	PARP14	P1095A	129.52	HLA-A*68:01	0	8	126.397714	真
1-024-002	YWSEKDVISAK	N		1-024-002_pool_1	Y	chr2_43889858_G_A	snp	LRPPRC	T1335I	79.08	HLA-A*68:01	1	9	136.482978	真
1-024-002	EGSLGISHTR	N		1-024-002_pool_1	Y	chr18_62157782_C_A	snp	PIGN	W83L	20.74	HLA-A*68:01	2	6	88.2623459	真
1-024-002	IPASVSAPK	N		1-024-002_pool_1	Y	chr13_109784018_C_A	snp	IRS2	S679I	63.55	HLA-A*68:01	3	16	224.278982	真
1-024-002	QDVSQVQVER	Y		1-024-002_pool_1	Y	chr9_64411223_T_G	snp	ANKRD20A 4	M646R	8.92	HLA-A*68:01	4	14	193.974327	真
1-024-002	LVVVGASGVGK	N		1-024-002_pool_1	Y	chr12_25245351_C_T	snp	KRAS	G12S	72.77	HLA-A*68:01	6	41	1238.56407	真
1-024-002	RATVPEL	N		1-024-002_pool_1	Y	chr7_131463253_A_T	snp	MKLN1	D521V	84.08	HLA-C*03:04	7	266	16010.7063	偽
1-024-002	SSAAAPFPL	Y		1-024-002_pool_1	Y	chr6_13711102_T_A	snp	RANBP9	H135L	43.5	HLA-C*03:04	8	103	4565.97417	偽
1-024-002	GVSKIIIGGNPK	N		1-024-002_pool_1	Y	chr4_10116175_C_T	snp	WDR1	D26N	134.53	HLA-A*68:01	9	125	6797.60699	真
1-024-002	EQNFVSTSDIK	個々に試験せず		1-024-002_pool_1	Y	chr3_25791346_A_C	snp	OXSM	K109T	12.82	HLA-A*68:01	17	156	9099.70986	真
1-024-002	RTQDVSQVQVER	N		1-024-002_pool_2	Y	chr9_64411223_T_G	snp	ANKRD20A 4	M646R	8.92	HLA-A*68:01	5	53	1847.42359	真
1-024-002	EAGNNSRVPR	N		1-024-002_pool_2	Y	chr2_74046630_G_T	snp	TET3	G238V	56.35	HLA-A*68:01	10	13	161.242762	真
1-024-002	RYVLHVAAA	N		1-024-002_pool_2	Y	chr3_122703943_C_G	snp	PARP14	P1095A	129.52	HLA-A*68:01	11	176	10453.627	真
1-024-002	VSKIIIGGNPK	N		1-024-002_pool_2	Y	chr4_10116175_C_T	snp	WDR1	D26N	134.53	HLA-A*68:01	12	38	954.724495	真
1-024-002	QPSGVPVTSI	N		1-024-002_pool_2	Y	chr12_14478436_GG_T	mnp	ATF7IP	G1021L	123.21	HLA-A*68:01	13	139	7795.97025	真

患者	ペプチド	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点)	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	MHCFIurryによりカバーされる最も高確率の制限
1-024-002	DVSVQVER	N		1-024-002_pool_2	Y	chr9_64411223_I_G	snp	ANKRD20A	M646R	8.92	HLA-A*68:01	14	7	123.489687	真
1-024-002	FVSTSDIKSM	Y		1-024-002_pool_2	Y	chr3_25791346_A_C	snp	OXSM	K109T	12.82	HLA-C*03:04	15	128	7025.56581	偽
1-024-002	FPWVSHSL	N		1-024-002_pool_2	Y	chr1_116062776_G_C	snp	SLC22A15	A396P	8.57	HLA-C*03:04	16	155	9082.40652	偽
1-024-002	APPRLGDSAL	N		1-024-002_pool_2	Y	chr6_137111102_I_A	snp	RANBP9	H135L	43.5	HLA-A*68:01	18	196	11590.601	真
1-024-002	ATVPELNEI	N		1-024-002_pool_2	Y	chr7_131463253_A_T	snp	MKINI	D521V	84.08	HLA-A*68:01	19	365	19785.1419	真
1-038-001	QEFAPLGTV	N	プールの結果を参照	1-038-001_pool_1	Y	chr2_219501883_G_T	snp	GMPPA	G92V	21.6	HLA-B*49:01	0	31	3481.07375	偽
1-038-001	MNQVLHAY	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr14_100354547_C_G	snp	WARS	D148H	757.21	HLA-C*07:01	12	422	27180.1513	偽
1-038-001	HEDVKEAI	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	16	300	24830.2411	偽
1-038-001	GPYFVQAV	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr1_111242326_C_T	snp	CHI3L2	L379F	122.33	HLA-B*49:01	1	19	1176.97782	偽
1-038-001	YEHEDVKEAI	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	2	212	22559.0306	偽
1-038-001	EESVMLITV	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr1_155833354_CC_AG	mnp	AGMAT	G105L	1.03	HLA-B*49:01	3	109	17185.8013	偽
1-038-001	IEEDSAEKI	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr6_84215849_C_A	snp	CEP162	E82D	15.62	HLA-B*49:01	4	171	20568.515	偽
1-038-001	TEEDVKIKF	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr7_93105459_C_A	snp	SAMD9	M213I	68.23	HLA-B*49:01	5	226	22894.2742	偽
1-038-001	NEQSKILKV	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chrX_70375298_C_G	snp	KIF4A	L625V	19.51	HLA-B*49:01	6	141	19054.8385	偽
1-038-001	VDNIIQSI	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr20_2654879_G_T	snp	NOP56	M167I	89.39	HLA-B*49:01	7	119	17928.8022	偽

患者	ペプチド	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点)	プールID	プール反応 (任意の時点) 変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	IMHCFlurryによりカバーされる最も高確率の制限
1-038-001	YEHEDVKEA	Y		1-038-001_pool_2	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	9	250	23419.567	偽
1-038-001	YWSEVPVSV	個々に試験せず		1-038-001_pool_2	chr17_2330604_G_A	snp	TSR1	H561Y	48.21	HLA-C*17:01	10	0	6.07874308	偽
1-038-001	SELTVHQRI	個々に試験せず		1-038-001_pool_2	chr19_37564705_G_C	snp	ZNF571	L575V	19.07	HLA-B*49:01	11	159	19886.0407	偽
1-038-001	VGVGKSAI	個々に試験せず		1-038-001_pool_2	chr12_25245350_C_A	snp	KRAS	G12V	91.89	HLA-C*17:01	13	388	26432.7668	偽
1-038-001	DMNQVLHAY	個々に試験せず		1-038-001_pool_2	chr14_100354547_C_G	snp	WARS	D148H	757.21	HLA-C*07:01	14	64	10286.4383	偽
1-038-001	NEKGKALY	個々に試験せず		1-038-001_pool_2	chr17_51294040_G_T	snp	UTP18	M547I	63.21	HLA-C*07:01	15	339	25564.2874	偽
1-038-001	TEYKLVVGVAV	個々に試験せず		1-038-001_pool_2	chr12_25245350_C_A	snp	KRAS	G12V	91.89	HLA-B*49:01	17	233	23113.572	偽
1-038-001	QEFAPLGTVG	個々に試験せず		1-038-001_pool_2	chr2_219501883_G_T	snp	GMPPA	G92V	21.6	HLA-B*49:01	18	338	25558.5468	偽
1-038-001	QEVNRTLLNV	個々に試験せず		1-038-001_pool_2	chr17_4085728_C_A	snp	ZZEF1	G863V	63	HLA-B*49:01	19	124	18359.7482	偽
1-038-001	VEMLGLISC	個々に試験せず		1-038-001_pool_2	chr4_168427109_C_A	snp	DDX60L	A631S	44.71	HLA-B*49:01	8	267	23949.2398	偽
1-050-001	LFDMMVSY	N		1-050-001_pool_1	chr1_193097666_T_C	snp	GLRX2	N94S	17.92	HLA-A*29:02	0	1	44.54051	真
1-050-001	ISTFRQCAL	個々に試験せず		1-050-001_pool_1	chr17_80346815_G_T	snp	RNF213	R2827L	330.56	HLA-C*16:01	10	322	22721.4424	偽
1-050-001	YNTDIEFY	個々に試験せず		1-050-001_pool_1	chr15_26580447_G_T	snp	GABRB3	T185N	2.2	HLA-A*29:02	16	20	447.152559	真
1-050-001	EETPPFSNY	N		1-050-001_pool_1	chr21_31266125_T_A	snp	TIAM1	Y283F	13.99	HLA-B*44:03	1	26	537.02592	真

患者	ペプチド	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurryのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
1-050-001	QASGNHHVW	個々に試験せず		1-050-001_pool_1	N	chr22_30893501_T_C	snp	OSBP2	Y677H	7.86	HLA-B*44:03	19	109	7506.81856	真
1-050-001	EEVPIAI	個々に試験せず		1-050-001_pool_1	N	chr18_5419733_G_A	snp	EPB41L3	S495L	51.69	HLA-B*44:03	2	17	390.306194	真
1-050-001	IEHNI RN AKY	個々に試験せず		1-050-001_pool_1	N	chr3_52617347_T_G	snp	PBRM1	D578A	65.68	HLA-B*44:03	3	10	186.953378	真
1-050-001	AERLDVKAI	個々に試験せず		1-050-001_pool_1	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-B*44:03	5	34	1075.19965	真
1-050-001	LFQQGKDLQQY	個々に試験せず		1-050-001_pool_1	N	chr17_80346815_G_T	snp	RNF213	R2827L	330.56	HLA-A*29:02	6	54	2855.46701	真
1-050-001	DTSPVAVAL	個々に試験せず		1-050-001_pool_1	N	chr5_73074790_T_C	snp	FCHO2	L543S	43.6	HLA-A*26:01	8	91	5750.39585	真
1-050-001	AEETPPFSNY	N		1-050-001_pool_2	N	chr21_31266125_T_A	snp	TIAM1	Y283F	13.99	HLA-B*44:03	9	16	364.187996	真
1-050-001	AAKAALEDF	個々に試験せず		1-050-001_pool_2	N	chr3_47661451_C_G	snp	SMARCC1	E721D	39.53	HLA-C*16:01	11	307	22125.437	偽
1-050-001	EVTPIAIR	個々に試験せず		1-050-001_pool_2	N	chr18_5419733_G_A	snp	EPB41L3	S495L	51.69	HLA-A*26:01	12	125	9269.11767	真
1-050-001	DVKAI G PIV	個々に試験せず		1-050-001_pool_2	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-A*26:01	13	90	5692.75283	真
1-050-001	NETPVA LTI	個々に試験せず		1-050-001_pool_2	N	chr7_79453094_C_A	snp	MAGI2	G76V	2.29	HLA-B*44:03	14	13	253.431553	真
1-050-001	LFVVFQTVY	個々に試験せず		1-050-001_pool_2	N	chr1_159535913_A_T	snp	OR10J5	L32Q	0.9	HLA-A*29:02	15	9	139.510048	真
1-050-001	AEAERLDVKAI	個々に試験せず		1-050-001_pool_2	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-B*44:03	17	38	1465.22509	真

患者	ペプチド	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	IMHCFlurryによりカバーされる最も高確率の制限
1-050-001	ASGNHHW	個々に試験せず	個々に試験せず	1-050-001_pool_2	N	chr22_308993501_I_C	snp	OSBP2	Y677H	7.86	HLA-C*16:01	18	173	13216.9384	偽
1-050-001	KLFHDMNVS	個々に試験せず	個々に試験せず	1-050-001_pool_2	N	chr1_193097666_I_C	snp	GLRX2	N94S	17.92	HLA-A*29:02	4	21	453.621334	真
1-050-001	ETPPSNVNTL	個々に試験せず	個々に試験せず	1-050-001_pool_2	N	chr21_31266125_I_A	snp	TIAM1	Y283F	13.99	HLA-A*26:01	7	172	13162.6216	真
CU04	DENITIQF	Y	Y	CU04_pool_1	Y	chr4_22413213_C_A	snp	ADGRA3	C734F	20.67	HLA-B*18:01	0	2	8.27203164	真
CU04	MELKVESF	N	N	CU04_pool_1	Y	chr1_37874128_G_C	snp	INPP5B	Q606E	36.85	HLA-B*18:01	1	5	13.0510076	真
CU04	EHIPESAGF	N	N	CU04_pool_1	Y	chr3_9943508_G_C	snp	CRELD1	Q347H	29.9	HLA-B*38:01	2	103	4218.0095	真
CU04	YHGDPMPCL	N	N	CU04_pool_1	Y	chr12_7066530_C_T	snp	CLS	P295L	157.54	HLA-B*38:01	3	12	76.7416543	真
CU04	DEERIPVL	N	N	CU04_pool_1	Y	chr7_5752914_I_C	snp	RNF216	M45V	49.2	HLA-B*18:01	4	29	387.328968	真
CU04	EVADAATLTM	Y	Y	CU04_pool_1	Y	chr1_52268541_A_C	snp	ZFYVE9	K845T	70.08	HLA-A*26:01	5	7	38.7340629	真
CU04	IEVEVEI	N	N	CU04_pool_1	Y	chr7_135598004_C_G	snp	NUP205	L691V	42.37	HLA-B*18:01	6	21	209.301169	真
CU04	DTVEPYTSF	Y	Y	CU04_pool_1	Y	chr14_34713369_C_A	snp	CFL2	D66Y	16.65	HLA-A*26:01	7	9	42.7267485	真
CU04	VEIQLY	N	N	CU04_pool_1	Y	chr11_62827178_C_G	snp	STX5	E134Q	83.43	HLA-B*18:01	8	3	11.6727539	真
CU04	LEKAVHAY	N	N	CU04_pool_1	Y	chr7_138762364_G_T	snp	ATP6V0A4	P163H	47.21	HLA-B*18:01	9	0	3.63590379	真
CU04	EEADFLAY	N	N	CU04_pool_2	N	chr6_10556704_C_T	snp	GCNT2	P94L	25.19	HLA-B*18:01	10	1	6.48490966	真
CU04	ENITTIQFY	N	N	CU04_pool_2	N	chr4_22413213_C_A	snp	ADGRA3	C734F	20.67	HLA-A*26:01	11	16	135.44155	真
CU04	FHATNPLN	N	N	CU04_pool_2	N	chr14_75117203_C_G	snp	NEK9	D252H	20.29	HLA-B*38:01	12	8	39.1165673	真
CU04	VFKDLSVTL	N	N	CU04_pool_2	N	chrX_40597563_G_A	snp	ATP6AP2	E145K	88.26	HLA-B*38:01	13	45	1080.8332	真
CU04	QAAVAQKIL	N	N	CU04_pool_2	N	chr17_42104792_I_A	snp	DHX58	M513L	35.87	HLA-C*12:03	14	136	6872.44	真

患者	ペプチド	個々のペプチド反応 (任意の時点)	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFlurryのランク	IMHCFlurry (nM)	IMHCFlurryによりカバーされる最も高確率の制限
CU04	IQDQIQNCI	N		CU04_pool_2	N	chr2_67404159_G_C	snp	ETAA1	E493Q	38.47	HLA-B*38:01	15	59	1665.0162	真
CU04	VAKGFISRM	N		CU04_pool_2	N	chr2_85395579_C_T	snp	CAPG	E314K	151.69	HLA-C*12:03	16	107	5236.61406	真
CU04	QTKPASLLY	N		CU04_pool_2	N	chr2_32487684_AG_A	del/fs	BIRC6	G2619fs	111.74	HLA-A*26:01	17	47	1143.73481	真
CU04	DHFETIIKY	N		CU04_pool_2	N	chr1_220024376_C_G	snp	EPRS	M277I	76.64	HLA-B*18:01	18	6	29.8996386	真
CU04	VEPYTFS	N		CU04_pool_2	N	chr14_34713369_C_A	snp	CHL2	D66Y	16.65	HLA-B*18:01	19	4	12.3783994	真
CU05	SVSDISEYRV	N		CU05_pool_1	N	chr12_15670870_G_C	snp	EPS8	Q64E	52.56	HLA-A*68:02	0	1	6.0399624	真
CU05	YTFEIQVNGV	N		CU05_pool_1	N	chr1_22865138_C_G	snp	EPHB2	A410G	74.99	HLA-A*68:02	1	22	132.877429	真
CU05	IYTSSGQLQLF	N		CU05_pool_1	N	chr10_73293336_I_C	snp	CFAP70	E636G	30.45	HLA-A*24:02	2	17	46.3526841	真
CU05	FATPSLHTSV	N		CU05_pool_1	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	4	16	43.8761927	真
CU05	AVSKPGLDYEL	N		CU05_pool_1	N	chr14_77026556_I_A	snp	IRF2BP1	M413L	58.51	HLA-A*68:02	5	274	13566.6012	真
CU05	KYINKTIRV	N		CU05_pool_1	N	chr19_2328426_C_T	snp	LSM7	D20N	76.01	HLA-A*24:02	8	32	318.671051	真
CU05	ETTEEMIKYV	N		CU05_pool_1	N	chr6_80040624_G_A	snp	ITK	G804E	17.14	HLA-A*68:02	9	37	398.324158	真
CU05	VVSHPLVYW	N		CU05_pool_1	N	chr4_106232956_C_G	snp	TBCK	D478H	71.17	HLA-A*68:02	11	235	10875.8686	真
CU05	DIFQVVKAI	N		CU05_pool_1	N	chr1_198754369_C_A	snp	PTPRC	L1204I	104.6	HLA-A*68:02	13	36	394.198029	真
CU05	FAFVAVKPGI	N		CU05_pool_1	N	chr14_77026556_I_A	snp	IRF2BP1	M413L	58.51	HLA-A*68:02	18	65	1067.11951	真
CU05	SVSDISEYR	N		CU05_pool_2	N	chr12_15670870_G_C	snp	EPS8	Q64E	52.56	HLA-A*68:02	3	94	2050.45825	真
CU05	YTFEIQGV	N		CU05_pool_2	N	chr1_22865138_C_G	snp	EPHB2	A410G	74.99	HLA-A*68:02	6	11	26.6362167	真
CU05	ATPSLHTSV	N		CU05_pool_2	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	7	25	177.027506	真
CU05	DFATPSLHTSV	N		CU05_pool_2	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	10	185	7619.02631	真
CU05	KYINKTIRVKF	N		CU05_pool_2	N	chr19_2328426_C_T	snp	LSM7	D20N	76.01	HLA-A*24:02	12	42	538.209517	真

10

20

30

40

患者	ペプチド	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点)	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFIurryのランク	IMHCFIurry (nM)	IMHCFIurryによりカバーされる最も高確率の制限
CU05	SVKPHLCSL	N		CU05_pool_2	N	chr17_35363437_C_T	snp	SIFN11	R124H	91.5	HLA-A*68:02	14	88	1897.58723	真
CU05	DISEYRVEHL	N		CU05_pool_2	N	chr12_15670870_G_C	snp	EPS8	Q64E	52.56	HLA-A*68:02	15	59	885.161001	真
CU05	WVSHPHLV	N		CU05_pool_2	N	chr4_106232956_C_G	snp	TBCK	D478H	71.17	HLA-A*68:02	16	15	40.725305	真
CU05	KVFLGNKV	N		CU05_pool_2	N	chrX_24810777_G_A	snp	POLA1	E1017K	19.31	HLA-A*68:02	17	61	954.869111	真
CU05	VSRPGLDYEL	N		CU05_pool_2	N	chr14_77026556_T_A	snp	IRF2BP1	M413L	58.51	HLA-A*68:02	19	258	12457.5646	真
CU02	SPSKTSLTL	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr12_132750694_G_T	snp	ANKLE2	P266T	43.78	HLA-B*07:02	0	7	20.5140939	真
CU02	ASADGTGKLV	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr16_1977246_A_G	snp	TBL3	I545V	26.23	HLA-B*57:01	1	20	77.5504026	真
CU02	LVGPAQLSHW	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr8_143930249_G_A	snp	PLEC	P863L	528.48	HLA-B*57:01	4	42	287.473059	真
CU02	QTAANGVVK	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr7_77773271_A_G	snp	RSBN1L	T584A	25.89	HLA-A*03:01	5	19	76.1012011	真
CU02	FPSKTSITL	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr12_132750694_G_T	snp	ANKLE2	P266T	43.78	HLA-B*07:02	6	26	131.765585	真
CU02	SSTSNRSSTW	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr10_96604023_G_A	snp	PIK3AP1	R733W	9.84	HLA-B*57:01	7	30	162.029882	真
CU02	LVYGPLGAGK	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr13_33821175_C_T	snp	RFC3	S44L	9.76	HLA-A*03:01	8	2	8.21211585	真
CU02	HSYSEICTW	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr8_119802006_C_G	snp	TAF2	D194H	29.74	HLA-B*57:01	9	3	10.120376	真
CU02	VTLDVILER	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr9_108979413_T_G	snp	CTNNA1	E323D	32.44	HLA-B*57:01	10	136	2107.24068	真

患者	ペプチド	個々のペプチド反応 (任意の時点) 試験せず	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFIurryのランク	IMHCFIurry (nM)	IMHCFIurryによりカバーされる最も高確率の制限
CU02	HSKPEDTDAW	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr12_133057238_A_G	snp	ZNF84	T175A	29.84	HLA-B*57:01	11	23	90.7546185	真
CU03	IAASRSVVM	個々に試験せず		CU03_pool_1	N	chr1_230868472_G_A	snp	C1orf198	A14V	36.47	HLA-C*12:03	0	19	146.699014	真
CU03	AAIAASRSV	個々に試験せず		CU03_pool_1	N	chr1_230868472_G_A	snp	C1orf198	A14V	36.47	HLA-C*12:03	2	42	492.404622	真
CU03	AASRSVVM	個々に試験せず		CU03_pool_1	N	chr1_230868472_G_A	snp	C1orf198	A14V	36.47	HLA-C*12:03	6	116	3437.73836	真
CU03	EMDMHLSDY	個々に試験せず		CU03_pool_1	N	chr5_37180032_T_A	snp	CSorf42	I1908L	14.78	HLA-A*01:01	8	7	35.7275148	真
CU03	VENQKHSL	個々に試験せず		CU03_pool_1	N	chr12_30728769_C_T	snp	CAPRIN2	S554N	6.69	HLA-B*08:01	10	124	3970.47602	真
CU03	QYMDSSLVKI	個々に試験せず		CU03_pool_1	N	chr10_60788061_G_T	snp	CDK1	S107I	26.84	HLA-A*23:01	7	8	50.3301427	真
CU03	SASLHPATV	個々に試験せず		CU03_pool_1	N	chr2_25929006_C_T	snp	KIF3C	R785H	17.29	HLA-C*12:03	9	30	260.370195	真
CU03	VPDQSKQL	個々に試験せず		CU03_pool_1	N	chr6_63685063_T_G	snp	PHF3	N447K	47.53	HLA-B*08:01	13	130	4071.14261	真
CU03	IVFIATSEF	個々に試験せず		CU03_pool_1	N	chr11_65976483_A_T	snp	SART1	N554I	70.53	HLA-B*15:03	5	3	17.4168253	真
CU03	YPAQPPVL	個々に試験せず		CU03_pool_1	N	chr20_44066022_C_A	snp	TOX2	S382Y	11.56	HLA-B*08:01	11	101	2455.95947	真

【 0 5 5 3 】

10

20

30

40

50

補足の表 4

被検物質 (平均)	刺激	ドナー ID				
		1-038-001	CU04	1-024-001	1-024-002	CU02
グランザイム B (pg/ml)*	DMSO	1786.73	1383.53	2639.03	854.78	1449.74
	ペプチドプール1	1672.60	4269.64	2449.23	1281.54	1132.49
TNF α (pg/ml)#	DMSO	1874.02	3747.71	2382.01	626.20	n/a
	ペプチドプール2	3118.30	3191.90	2006.73	872.89	n/a
IL-2 (pg/ml)#	DMSO	37.58	34.64	21.76	38.07	1.22
	ペプチドプール1	53.02	217.57	42.05	57.13	7.44
IL-5 (pg/ml)#	DMSO	16.58	80.81	24.98	24.77	n/a
	ペプチドプール2	61.54	75.70	33.70	48.84	n/a
IL-2 (pg/ml)#	DMSO	1.78	3.86	4.24	0.23	6.67
	ペプチドプール1	15.53	9.88	7.75	0.00	0.00
IL-5 (pg/ml)#	DMSO	26.66	27.25	5.72	10.20	n/a
	ペプチドプール2	0.00	19.15	11.48	0.00	n/a
IL-5 (pg/ml)#	DMSO	26.47	5.20	20.92	11.96	18.91
	ペプチドプール1	10.48	14.65	26.72	9.42	17.64
陽性値は斜体で示す。* グランザイムB ELISA: DMSOバックグラウンドよりも1.5倍以上高い値は陽性とみなした。# U-Plex MSD アッセイ: DMSOバックグラウンドよりも1.5倍以上高い値は陽性とみなした。	DMSO	27.31	19.65	11.01	29.93	n/a
	ペプチドプール2	26.47	25.43	20.11	40.11	n/a

10247321

10

20

30

40

【 0 5 5 4 】

補足の表 5

I V S 対 照 実 験 に お け る T S N A 及 び 感 染 症 エ ン ド ト ー プ

ペプチド名	配列	由来(細胞株, 遺伝子)	予測されたHLA拘束性	予測された結合親和性(nM)	変異位置	変異
新生抗原_A1	APKKKSIKL	H2009 PPPIA3	B*07:02	125	chr19-49140014	C → T
新生抗原_A2	LLLEVVVHL	H128 FANCA	A*02:01	6	chr16-89808348	C → T
新生抗原_A3	FTDEKVKAY	H2122 PDE10A	A*01:01	41	chr6-165543564	G → T
新生抗原_A6	RTAKQNPLTK	H2122 GPR183	A*03:01	138	chr13-99295446	G → A
新生抗原_A7	FLAPTGVVPV	H128 NTM	A*02:01	8	chr11-131911555	T → C
新生抗原_A10	RLADAEKLFQL	H128 PLEKHG4	A*02:01	201	chr16-67284435	G → A
新生抗原_A11	RTAKQNPLTKK	H2122 GPR183	A*03:01	131	chr13-99295446	G → A
新生抗原_B2	IMYLTGMVNVK	H2009 GSPT1	A*03:01	33	chr16-11891120	G → A
新生抗原_B3	TLQELSHAL	H128 PRPF19	A*02:01	106	chr11-60902829	G → T
新生抗原_B6	VSQPVAPSY	Colo829 KIAA0319L	A*01:01	948	chr1-35479047	C → T
新生抗原_B7	RLFTPIISAGY	H2126 CYP26B1	A*03:01	157	chr2-72133060	G → C
新生抗原_B8	ITEEPILMTY	H2122 RP111	A*01:01	308	chr8-10611205	C → A
新生抗原_B10	KVTGHRWLK	H2009 BSG	A*03:01	51	chr19-579577	G → A
新生抗原_B12	KLSEQLKK	H2009 TLR5	A*03:01	39	chr1-223110532	C → G
新生抗原_C3	GTKPNPHVY	H2126 OAS3	A*03:01	7336	chr12-112961105	G → T
新生抗原_C4	QQQQVVTNK	H2126 LRP1	A*03:01	2361	chr12-57162861	G → T
新生抗原_C5	KVLGKGSFAK	H2126 PLK2	A*03:01	40	chr5-58459089	G → A
新生抗原_C6	SVQAPVPPK	H2009 ENGASE	A*03:01	279	chr17-79084548	C → G
EBV RAKF	RAKFKQLL	EBV BZLF-1	B*08:01	457	Nan	Nan
Flu CTCL	CTELKLSDY	Influenza NP	A*01:01	39	Nan	Nan
Flu ELRS	ELRSRYWAI	Influenza A	B*08:01	12	Nan	Nan
CMV NLVP	NLVPMTVATV	CMV pp65	A*02:01	45	Nan	Nan
Flu GILG	GILGFVFTL	Influenza MP	A*02:01	20	Nan	Nan
HCV KLVA	KLVALGINAV	HCV NS3	A*02:01	49	Nan	Nan
HIV ILKE	ILKEPVHGV	HIV pol	A*02:01	144	Nan	Nan
RSV NPKA	NPKASLLSL	RSV NP	B*07:02	60	Nan	Nan
*新生抗原配列中の 変異ペプチドを 下線で示す。 **腫瘍細胞株: Colo829, H128, H2009, H2122, H2126						

【 0 5 5 5 】

補足の表 6

T C R 核酸配列

10

20

30

40

補足の表 7
頻度及び C D R 3 a / b 配列

クロナタイプID	頻度	割合	cdr3s_aa	cdr3s_nt
クロナタイプ1	386	0.491719745	TRA:CAVTVTGRRALTF;TRB:CASNPPDAARGOETQYF	TRA:TGTGCTGTGACCGTACGGGCACGGGAGGAGAGCAGCATTACTTTT;IRB:TGTGCCAGCACA CCCCCGGACGCTGGGAGGGACAAAGAGACCAGTACTTC
クロナタイプ3	53	0.067515924	TRA:CALNARLMF;TRB:CASSYREYNTEAFF	TRA:TGTGCTCTAAATGCCAGACTCATGTTT;IRB:TGTGCCAGCAGTTACCCGGGAGT CAACACTGAAGCTTCTTT
クロナタイプ4	34	0.043312102	TRA:CATGLDLDKLI;TRA:CAVRWGGNQYF;TRB:CASSLAGTGGNVEQFF	TRA:TGTGCTACGGGCCTAGATTTGGACAAGCTCATCTTT;TRA:TGTGCTGTCAAGT GGCGGTAAACCAGTTCTATTTT;IRB:TGTGCCAGCAGCTTAGCCGGGACAGGGGGT AATTATGAGCAGTCTTTC
クロナタイプ6	10	0.012738854	TRA:CAVTVTGRRALTF;TRB:CASNPPDAARGOETQYF;TRB:CASSYREYNTEAFF	TRA:TGTGCTGTGACCGTACGGGCACGGGAGGAGAGCAGCATTACTTTT;IRB:TGTGCCAGCACA CCCCCGGACGCTGGGAGGGACAAAGAGACCAGTACTTC;IRB:TGTGCCAGCAGT TACCGGAGTACAACACTGAAGCTTCTTT
クロナタイプ9	7	0.008917197	TRA:CAVLDSDNYQLIW;TRB:CSATRHLNSNPQHF	TRA:TGTGCTGTGGTGTGGATAGCAACTATCAGTTAATCTGG;IRB:TGCCAGTGCAAC CAGGGGCAGCTTGAGCAATCAGCCCGCAGCATTT
クロナタイプ10	5	0.006369427	TRA:CATASRQGGSEKLVF;TRB:CASSRGGGTDTQYF	TRA:TGTGCTACGGCTAGCCGTACGGGCGGATCTGAAAAGCTGGTCTTT;IRB:TGTG CCAGCAGCCGAGGGGGGGGCACAGATACCGAGTATTTT

【 0 5 5 7 】

10

20

30

40

50

補足の表 8
配列及び V、D、J 遺伝子

クローノタイプID	コンセンサスID	長さ	鎖	V 遺伝子	D 遺伝子	J 遺伝子	C 遺伝子	完全長	プロダクティブ	cdri3	cdri3_nt	リード	Umis
0	クローノタイプ1	1263	TRA	TRAV8-4	None	TRAJ5	TRAC	真	真	CAVVTGRRALIF	TGTGCTGTGACCCGTACGGGCGAGGAGCAGCTACTTTT	271775	10562
1	クローノタイプ10	1332	TRB	TRBV2	TRBD2	TRBJ2-5	TRBC2	真	真	CASNPDAARGQETQYF	TGTGCCAGAACCCCGGACGCTCCGAGGGGACAAGACCAGTACTTC	1796338	7242
2	クローノタイプ10	604	TRA	TRAV17	None	TRAJ57	TRAC	真	真	CATASRGGGSEKLVF	TGTGCTACGGCTAGCCGCTCAGGGCGGATCTGAAAAGCTGGTCTTT	4006	13
3	クローノタイプ10	774	TRB	TRBV7-6	TRBD1	TRBJ2-3	TRBC2	真	真	CASSRGGGDTQYF	TGTGCCAGCAGCCGAGGGGGGACAGATACGCAGTATTT	9401	35
103	クローノタイプ3	1663	TRB	TRBV6-1	TRBD2	TRBJ1-1	TRBC1	真	真	CASSYREYNTEAFF	TGTGCCAGAGTTACCGGGAGTACAACACTGAAGCTTCTTT	243324	894
104	クローノタイプ3	608	TRA	TRAV6	None	TRAJ31	TRAC	真	真	CALNARLIF	TGTGCTCTAAATGCCAGACTCATGTT	87200	247
125	クローノタイプ4	683	TRA	TRAV17	None	TRAJ34	TRAC	真	真	CATGIDLKLIIF	TGTGCTAGGGGCTAGATTTGGACAAGCTCATCTTT	131803	385
126	クローノタイプ4	669	TRA	TRAV41	None	TRAJ49	TRAC	真	真	CAVRVGGNGYF	TGTGCTGTCCAGATGGGGCGGTACCAGTCTTATTTT	88320	307
127	クローノタイプ4	1315	TRB	TRBV7-9	TRBD1	TRBJ2-1	TRBC2	真	真	CASSIAGTGGNVEGQF	TGTGCCAGCAGCTTAGCCGGGACAGGGGTAAATATGAGCAGTCTTC	237200	830
171	クローノタイプ6	1007	TRB	TRBV6-1	TRBD2	TRBJ1-1	TRBC1	真	真	CASSYREYNTEAFF	TGTGCCAGAGTTACCGGGAGTACAACACTGAAGCTTCTTT	23803	105
172	クローノタイプ6	713	TRA	TRAV8-4	None	TRAJ5	TRAC	真	真	CAVVTGRRALIF	TGTGCTGTGACCCGTACGGGCGAGGAGCAGTACTTTT	24293	106
242	クローノタイプ9	568	TRA	TRAV22	None	TRAJ33	TRAC	真	真	CAVLDSNYQLIW	TGTGCCAGCAACCCCGGACGCTGCCAGGGGACAAGACCAGTACTTC	34437	132
243	クローノタイプ9	1102	TRB	TRBV20-1	TRBD1	TRBJ1-5	TRBC1	真	真	CSATRGHLSNQPQHF	TGCAGTGCAACAGGGGGGCACTTGAGCAATCAGCCCGCAGCATTTT	9883	29

【 0 5 5 8 】

10

20

30

40

50

参考文献

1. Desrichard, A., Snyder, A. & Chan, T. A. Cancer Neoantigens and Applications for Immunotherapy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* (2015). doi:10.1158/1078-0432.CCR-14-3175
2. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69–74 (2015).
3. Gubin, M. M., Artyomov, M. N., Mardis, E. R. & Schreiber, R. D. Tumor neoantigens: building a framework for personalized cancer immunotherapy. *J. Clin. Invest.* 125, 3413–3421 (2015).
4. Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128 (2015).
5. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189–2199 (2014). 10
6. Carreno, B. M. et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T-cells. *Science* 348, 803–808 (2015).
7. Tran, E. et al. Cancer immunotherapy based on mutation-specific CD4+ T-cells in a patient with epithelial cancer. *Science* 344, 641–645 (2014).
8. Hacohen, N. & Wu, C. J.-Y. United States Patent Application: 0110293637 - COMPOSITIONS AND METHODS OF IDENTIFYING TUMOR SPECIFIC NEOANTIGENS. (A1). at <
17. Target Capture for NextGen Sequencing - IDT. at <[<http://www.idtdna.com/pages/products/nextgen/target-capture>](http://www.idtdna.com/pages/products/nextgen/target-capture)
18. Shukla, S. A. et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol.* 33, 1152–1158 (2015).
19. Cieslik, M. et al. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res.* 25, 1372–1381 (2015). 40
20. Bodini, M. et al. The hidden genomic landscape of acute myeloid leukemia: subclonal structure revealed by undetected mutations. *Blood* 125, 600–605 (2015).
21. Saunders, C. T. et al. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinforma. Oxf. Engl.* 28, 1811–1817 (2012).

22. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31, 213–219 (2013).
23. Wilkerson, M. D. et al. Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. *Nucleic Acids Res.* 42, e107 (2014).
24. Mose, L. E., Wilkerson, M. D., Hayes, D. N., Perou, C. M. & Parker, J. S. ABRA: improved coding indel detection via assembly-based realignment. *Bioinforma. Oxf. Engl.* 30, 2813–2815 (2014).
25. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinforma. Oxf. Engl.* 25, 2865–2871 (2009).
26. Lam, H. Y. K. et al. Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library. *Nat. Biotechnol.* 28, 47–55 (2010).
27. Frampton, G. M. et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* 31, 1023–1031 (2013).
28. Boegel, S. et al. HLA typing from RNA-Seq sequence reads. *Genome Med.* 4, 102 (2012).
29. Liu, C. et al. ATHLATES: accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Res.* 41, e142 (2013).
30. Mayor, N. P. et al. HLA Typing for the Next Generation. *PloS One* 10, e0127153 (2015).
31. Roy, C. K., Olson, S., Graveley, B. R., Zamore, P. D. & Moore, M. J. Assessing long-distance RNA sequence connectivity via RNA-templated DNA-DNA ligation. *eLife* 4, (2015).
32. Song, L. & Florea, L. CLASS: constrained transcript assembly of RNA-seq reads. *BMC Bioinformatics* 14 Suppl 5, S14 (2013).
33. Maretty, L., Sibbesen, J. A. & Krogh, A. Bayesian transcriptome assembly. *Genome Biol.* 15, 501 (2014).
34. Perteza, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295 (2015).
35. Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinforma. Oxf. Engl.* (2011). doi:10.1093/bioinformatics/btr355
36. Vitting-Seerup, K., Porse, B. T., Sandelin, A. & Waage, J. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics* 15, 81 (2014).
37. Rivas, M. A. et al. Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* 348, 666–669 (2015).
38. Skelly, D. A., Johansson, M., Madeoy, J., Wakefield, J. & Akey, J. M. A powerful and flexible statistical framework for testing hypotheses of allele-specific gene expression from RNA-seq data. *Genome Res.* 21, 1728–1737 (2011).
39. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinforma. Oxf. Engl.* 31, 166–169 (2015).
40. Furney, S. J. et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov.* (2013). doi:10.1158/2159-8290.CD-13-0330
41. Zhou, Q. et al. A chemical genetics approach for the functional assessment of novel cancer genes. *Cancer Res.* (2015). doi:10.1158/0008-5472.CAN-14-2930
42. Maguire, S. L. et al. SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J. Pathol.* 235, 571–580 (2015).
43. Carithers, L. J. et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation Biobanking* 13, 311–319 (2015).
44. Xu, G. et al. RNA CoMPASS: a dual approach for pathogen and host transcriptome analysis of RNA-seq datasets. *PloS One* 9, e89445 (2014).
45. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinforma. Oxf. Engl.* (2015). doi:10.1093/bioinformatics/btv639

46. Jørgensen, K. W., Rasmussen, M., Buus, S. & Nielsen, M. NetMHCstab - predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. *Immunology* 141, 18–26 (2014).
47. Larsen, M. V. et al. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. *Eur. J. Immunol.* 35, 2295–2303 (2005).
48. cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* 57, 33–41 (2005).
49. Boisvert, F.-M. et al. A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells. *Mol. Cell. Proteomics* 11, M111.011429–M111.011429 (2012).
50. Duan, F. et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J. Exp. Med.* 211, 2231–2248 (2014). 10
51. Janeway's Immunobiology: 9780815345312: Medicine & Health Science Books @ Amazon.com. at <<http://www.amazon.com/Janeways-Immunobiology-Kenneth-Murphy/dp/0815345313>>
52. Calis, J. J. A. et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* 9, e1003266 (2013).
53. Zhang, J. et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 346, 256–259 (2014)
54. Walter, M. J. et al. Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* 366, 1090–1098 (2012).
55. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, Cox AL, Appella E, Engelhard VH. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* 1992. 255: 1261-1263. 20
56. Zarring AL, Polefrone JM, Evans AM, Mikesch LM, Shabanowitz J, Lewis ST, Engelhard VH, Hunt DF. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci U S A.* 2006 Oct 3;103(40):14889-94.
57. Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ, Mann M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol Cell Proteomics.* 2015 Mar;14(3):658-73. doi: 10.1074/mcp.M114.042812.
58. Abelin JG, Trantham PD, Penny SA, Patterson AM, Ward ST, Hildebrand WH, Cobbold M, Bai DL, Shabanowitz J, Hunt DF. Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. *Nat Protoc.* 2015 Sep;10(9):1308-18. doi: 10.1038/nprot.2015.086. Epub 2015 Aug 6
59. Barnstable CJ, Bodmer WF, Brown G, Galfre G, Milstein C, Williams AF, Ziegler A. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell.* 1978 May;14(1):9-20. 30
60. Goldman JM, Hibbin J, Kearney L, Orchard K, Th'ng KH. HLA-DR monoclonal antibodies inhibit the proliferation of normal and chronic granulocytic leukaemia myeloid progenitor cells. *Br J Haematol.* 1982 Nov;52(3):411-20.
61. Eng JK, Jahan TA, Hoopmann MR. Comet: an open-source MS/MS sequence database search tool. *Proteomics.* 2013 Jan;13(1):22-4. doi: 10.1002/pmic.201200439. Epub 2012 Dec 4.
62. Eng JK, Hoopmann MR, Jahan TA, Egertson JD, Noble WS, MacCoss MJ. A deeper look into Comet--implementation and features. *J Am Soc Mass Spectrom.* 2015 Nov;26(11):1865-74. doi: 10.1007/s13361-015-1179-x. Epub 2015 Jun 27.
63. Lukas Käll, Jesse Canterbury, Jason Weston, William Stafford Noble and Michael J. MacCoss. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature Methods* 4:923 – 925, November 2007 40
64. Lukas Käll, John D. Storey, Michael J. MacCoss and William Stafford Noble. Assigning confidence measures to peptides identified by tandem mass spectrometry. *Journal of Proteome Research*, 7(1):29-34, January 2008

65. Lukas Käll, John D. Storey and William Stafford Noble. Nonparametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. *Bioinformatics*, 24(16):i42-i48, August 2008
66. Bo Li and C. olin N. Dewey. RSEM: accurate transcript quantification from RNA-Seq data with or without a referenfe genome. *BMC Bioinformatics*, 12:323, August 2011
67. Hillary Pearson, Tariq Daouda, Diana Paola Granados, Chantal Durette, Eric Bonneil, Mathieu Courcelles, Anja Rodenbrock, Jean-Philippe Laverdure, Caroline Côté, Sylvie Mader, Sébastien Lemieux, Pierre Thibault, and Claude Perreault. MHC class I-associated peptides derive from selective regions of the human genome. *The Journal of Clinical Investigation*, 2016,
68. Juliane Liepe, Fabio Marino, John Sidney, Anita Jeko, Daniel E. Bunting, Alessandro Sette, Peter M. Kloetzl, Michael P. H. Stumpf, Albert J. R. Heck, Michele Mishto. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science*, 21, October 2016. 10
69. Mommen GP., Marino, F., Meiring HD., Poelen, MC., van Gaans-van den Brink, JA., Mohammed S., Heck AJ., and van Els CA. Sampling From the Proteome to the Human Leukocyte Antigen-DR (HLA-DR) Ligandome Proceeds Via High Specificity. *Mol Cell Proteomics* 15(4): 1412-1423, April 2016.
70. Sebastian Kreiter, Mathias Vormehr, Niels van de Roemer, Mustafa Diken, Martin Löwer, Jan Diekmann, Sebastian Boegel, Barbara Schrörs, Fulvia Vascotto, John C. Castle, Arbel D. Tadmor, Stephen P. Schoenberger, Christoph Huber, Özlem Türeci, and Ugur Sahin. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 520, 692-696, April 2015.
71. Tran E., Turcotte S., Gros A., Robbins P.F., Lu Y.C., Dudley M.E., Wunderlich J.R., Somerville R.P., Hogan K., Hinrichs C.S., Parkhurst M.R., Yang J.C., Rosenberg S.A. Cancer immunotherapy based on mutation-specific CD4+ T-cells in a patient with epithelial cancer. *Science* 344(6184) 641-645, May 2014. 20
72. Andreatta M., Karosiene E., Rasmussen M., Stryhn A., Buus S., Nielsen M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics* 67(11-12) 641-650, November 2015.
73. Nielsen, M., Lund, O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* 10:296, September 2009.
74. Nielsen, M., Lundegaard, C., Lund, O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8:238, July 2007.
75. Zhang, J., et al. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Molecular & Cellular Proteomics*. 11(4):1-8. 1/2/2012.
76. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189–2199 (2014). 30
77. Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128 (2015).
78. Gubin, M. M., Artyomov, M. N., Mardis, E. R. & Schreiber, R. D. Tumor neoantigens: building a framework for personalized cancer immunotherapy. *J. Clin. Invest.* 125, 3413–3421 (2015).
79. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* 348, 69–74 (2015).
80. Carreno, B. M. et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T-cells. *Science* 348, 803–808 (2015).
81. Ott, P. A. et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 547, 217–221 (2017).
82. Sahin, U. et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* 547, 222–226 (2017). 40
83. Tran, E. et al. T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *N. Engl. J. Med.* 375, 2255–2262 (2016).
84. Gros, A. et al. Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat. Med.* 22, 433–438 (2016).

85. The problem with neoantigen prediction. *Nat. Biotechnol.* 35, 97–97 (2017).
86. Vitiello, A. & Zanetti, M. Neoantigen prediction and the need for validation. *Nat. Biotechnol.* 35, 815–817 (2017).
87. Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L. J. & Mann, M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell. Proteomics MCP* 14, 658–673 (2015).
88. Vita, R. et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* 43, D405–412 (2015).
89. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinforma. Oxf. Engl.* 32, 511–517 (2016).
90. O'Donnell, T. J. et al. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst.* (2018). doi:10.1016/j.cels.2018.05.014 10
91. Bassani-Sternberg, M. et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat. Commun.* 7, 13404 (2016).
92. Abelin, J. G. et al. Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. *Immunity* 46, 315–326 (2017).
93. Yadav, M. et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 515, 572–576 (2014).
94. Stranzl, T., Larsen, M. V., Lundegaard, C. & Nielsen, M. NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* 62, 357–368 (2010).
95. Bentzen, A. K. et al. Large-scale detection of antigen-specific T-cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat. Biotechnol.* 34, 1037–1045 (2016).
96. Tran, E. et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* 350, 1387–1390 (2015). 20
97. Stronen, E. et al. Targeting of cancer neoantigens with donor-derived T-cell receptor repertoires. *Science* 352, 1337–1341 (2016).
98. Trolle, T. et al. The Length Distribution of Class I-Restricted T-cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference. *J. Immunol. Baltim. Md* 1950 196, 1480–1487 (2016).
99. Di Marco, M. et al. Unveiling the Peptide Motifs of HLA-C and HLA-G from Naturally Presented Peptides and Generation of Binding Prediction Matrices. *J. Immunol. Baltim. Md* 1950 199, 2639–2651 (2017).
100. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning.* (MIT Press, 2016).
101. Sette, A. et al. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T-cell epitopes. *J. Immunol. Baltim. Md* 1950 153, 5586–5592 (1994). 30
102. Fortier, M.-H. et al. The MHC class I peptide repertoire is molded by the transcriptome. *J. Exp. Med.* 205, 595–610 (2008).
103. Pearson, H. et al. MHC class I-associated peptides derive from selective regions of the human genome. *J. Clin. Invest.* 126, 4690–4701 (2016).
104. Bassani-Sternberg, M. et al. Deciphering HLA-I motifs across HLA peptidomes improves neoantigen predictions and identifies allostery regulating HLA specificity. *PLoS Comput. Biol.* 13, e1005725 (2017).
105. Andreatta, M., Lund, O. & Nielsen, M. Simultaneous alignment and clustering of peptide data using a Gibbs sampling approach. *Bioinforma. Oxf. Engl.* 29, 8–14 (2013).
106. Andreatta, M., Alvarez, B. & Nielsen, M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res.* (2017). doi:10.1093/nar/gkx248 40
107. Gros, A. et al. Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat. Med.* 22, 433–438 (2016).
108. Zacharakis, N. et al. Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. *Nat. Med.* 24, 724–730 (2018).

109. Chudley, L. et al. Harmonisation of short-term in vitro culture for the expansion of antigen-specific CD8+ T-cells with detection by ELISPOT and HLA-multimer staining. *Cancer Immunol. Immunother.* 63, 1199–1211 (2014).
110. Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207–211 (2015).
111. Anagnostou, V. et al. Evolution of Neoantigen Landscape during Immune Checkpoint Blockade in Non-Small Cell Lung Cancer. *Cancer Discov.* 7, 264–276 (2017).
112. Carreno, B. M. et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T-cells. *Science* 348, 803–808 (2015).
113. Stevanović, S. et al. Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer. *Science* 356, 200–205 (2017). 10
114. Pasetto, A. et al. Tumor- and Neoantigen-Reactive T-cell Receptors Can Be Identified Based on Their Frequency in Fresh Tumor. *Cancer Immunol. Res.* 4, 734–743 (2016).
115. Gillette, M. A. & Carr, S. A. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat. Methods* 10, 28–34 (2013).
116. Boegel, S., Löwer, M., Bukur, T., Sahin, U. & Castle, J. C. A catalog of HLA type, HLA expression, and neo-epitope candidates in human cancer cell lines. *Oncoimmunology* 3, e954893 (2014).
117. Johnson, D. B. et al. Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. *Nat. Commun.* 7, 10582 (2016).
118. Robbins, P. F. et al. A Pilot Trial Using Lymphocytes Genetically Engineered with an NY-ESO-1-Reactive T-cell Receptor: Long-term Follow-up and Correlates with Response. *Clin. Cancer Res.* 21, 1019–1027 (2015). 20
119. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189–2199 (2014).
120. Calis, J. J. A. et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput. Biol.* 9, e1003266 (2013).
121. Duan, F. et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J. Exp. Med.* 211, 2231–2248 (2014).
122. Glanville, J. et al. Identifying specificity groups in the T-cell receptor repertoire. *Nature* 547, 94–98 (2017).
123. Dash, P. et al. Quantifiable predictive features define epitope-specific T-cell receptor repertoires. *Nature* 547, 89–93 (2017).
124. Hunt, D. F. et al. Pillars article: Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* 1992, 255: 1261–1263. *J. Immunol. Baltim. Md 1950* 179, 2669–2671 (2007). 30
125. Zarling, A. L. et al. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc. Natl. Acad. Sci. U. S. A.* 103, 14889–14894 (2006).
126. Abelin, J. G. et al. Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. *Nat. Protoc.* 10, 1308–1318 (2015).
127. Barnstable, C. J. et al. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens—new tools for genetic analysis. *Cell* 14, 9–20 (1978).
128. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics* 13, 22–24 (2013).
129. Eng, J. K. et al. A deeper look into Comet—implementation and features. *J. Am. Soc. Mass Spectrom.* 26, 1865–1874 (2015). 40
130. Käll, L., Storey, J. D., MacCoss, M. J. & Noble, W. S. Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. *J. Proteome Res.* 7, 29–34 (2008).

131. Käll, L., Storey, J. D. & Noble, W. S. Non-parametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. *Bioinforma. Oxf. Engl.* 24, i42-48 (2008).
132. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* 4, 923–925 (2007).
133. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323 (2011).
134. Chollet, F. & others. Keras. (2015).
135. Bastien, F. et al. Understanding the difficulty of training deep feedforward neural networks. *Proc. Thirteenth. Int. Conf. Artif. Intell. Stat.* 249–256 (2010).
136. Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* 249–256 (2010). 10
137. Kingma, D. & Ba, J. Adam: A method for stochastic optimization. *ArXiv Prepr. ArXiv14126980* (2014).
138. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18, 6097–6100 (1990).
139. Rubinsteyn, A., O'Donnell, T., Damaraju, N. & Hammerbacher, J. Predicting Peptide-MHC Binding Affinities With Imputed Training Data. *bioRxiv* (2016). doi:<https://doi.org/10.1101/054775>
140. Tran, E. et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* 350, 1387–1390 (2015).
141. Stronen, E. et al. Targeting of cancer neoantigens with donor-derived T-cell receptor repertoires. *Science* 352, 1337–1341 (2016). 20
142. Janetzki, S., Cox, J. H., Oden, N. & Ferrari, G. Standardization and validation issues of the ELISPOT assay. *Methods Mol. Biol. Clifton NJ* 302, 51–86 (2005).
143. Janetzki, S. et al. Guidelines for the automated evaluation of Elispot assays. *Nat. Protoc.* 10, 1098–1115 (2015).
144. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* 25, 1754–1760 (2009).
145. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498 (2011).
146. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* (2012).
147. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly (Austin)* 6, 80–92 (2012). 30
148. Szolek, A. et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinforma. Oxf. Engl.* 30, 3310–3316 (2014).
149. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31, 213–219 (2013).
150. Scholz, E. M. et al. Human Leukocyte Antigen (HLA)-DRB1*15:01 and HLA-DRB5*01:01 Present Complementary Peptide Repertoires. *Front. Immunol.* 8, 984 (2017).
151. Ooi, J. D. et al. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T-cells. *Nature* 545, 243–247 (2017).
152. Karosiene, E. et al. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics* 65, 711–724 (2013). 40
153. Dudley ME, Gross CA, Langhan MM, et al. CD8+ enriched “young” tumor infiltrating lymphocytes can mediate regression of metastatic melanoma. *Clinical cancer research : an official*

journal of the American Association for Cancer Research. 2010;16(24):6122-6131. doi:10.1158/1078-0432.CCR-10-1297.

154. Dudley ME, Wunderlich JR, Shelton TE, Even J, Rosenberg SA. Generation of Tumor-Infiltrating Lymphocyte Cultures for Use in Adoptive Transfer Therapy for Melanoma Patients. *Journal of immunotherapy (Hagerstown, Md: 1997)*. 2003;26(4):332-342.

155. Cohen CJ, Gartner JJ, Horovitz-Fried M, et al. Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *The Journal of Clinical Investigation*. 2015;125(10):3981-3991. doi:10.1172/JCI82416.

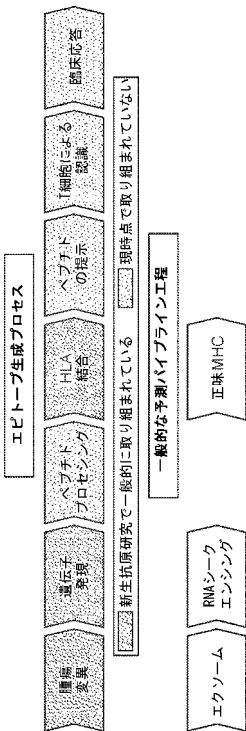
156. Kelderman, S., Heemskerk, B., Fanchi, L., Philips, D., Toebes, M., Kvistborg, P., Buuren, M. M., Rooij, N., Michels, S., Germeroth, L., Haanen, J. B. and Schumacher, N. M. (2016), Antigen-specific TIL therapy for melanoma: A flexible platform for personalized cancer immunotherapy. *Eur. J. Immunol.*, 46: 1351-1360. doi:10.1002/eji.201545849.

157. Hall M, Liu H, Malafa M, et al. Expansion of tumor-infiltrating lymphocytes (TIL) from human pancreatic tumors. *Journal for Immunotherapy of Cancer*. 2016;4:61. doi:10.1186/s40425-016-0164-7.

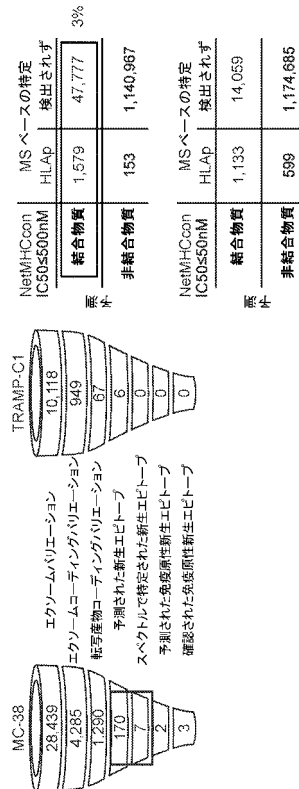
158. Briggs A, Goldfless S, Timberlake S, et al. Tumor-infiltrating immune repertoires captured by single-cell barcoding in emulsion. *bioRxiv*. 2017. doi.org/10.1101/134841.

159. US Patent Application No. 20160244825A1.

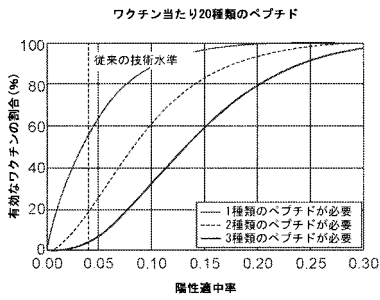
【図 1 A】



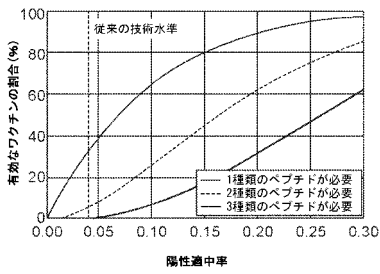
【図 1 B】



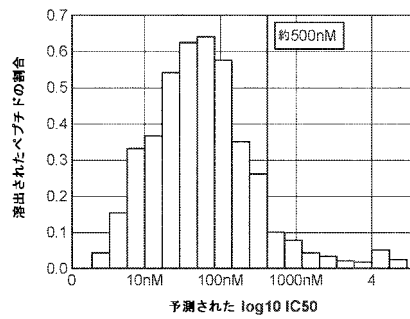
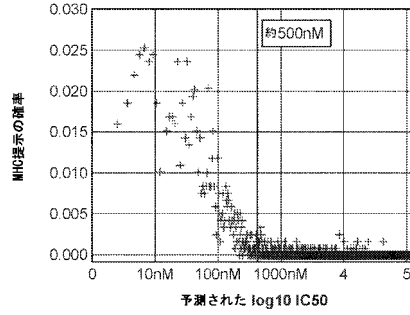
【図 1 C】



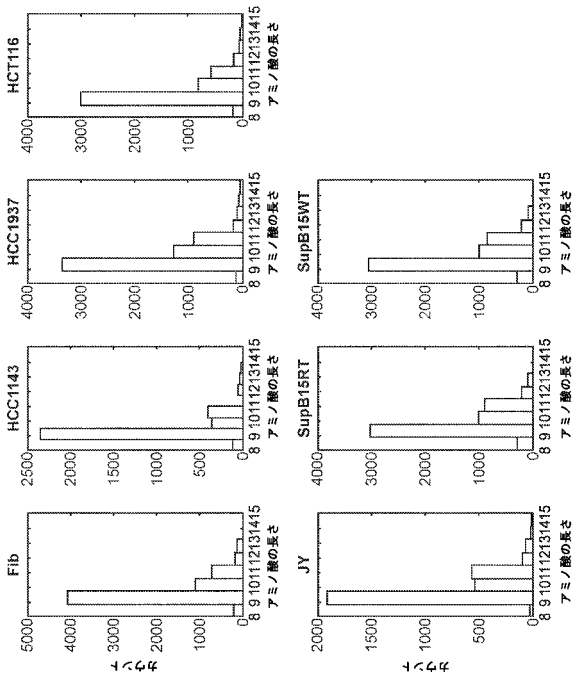
ワクチン当たり10種類のペプチド



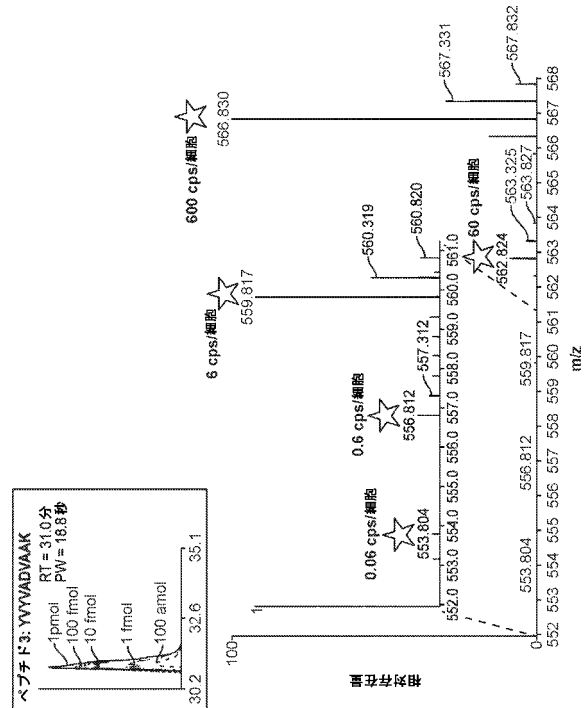
【図 1 D】



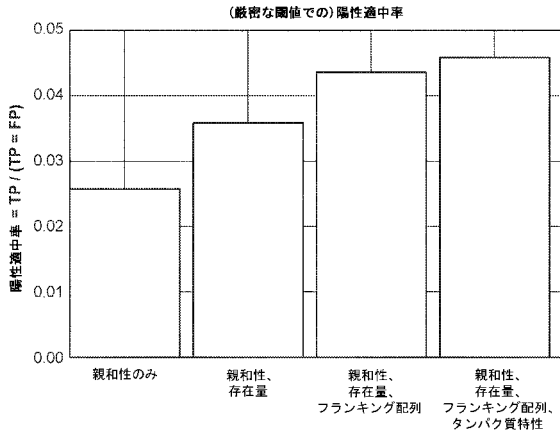
【図 1 E】



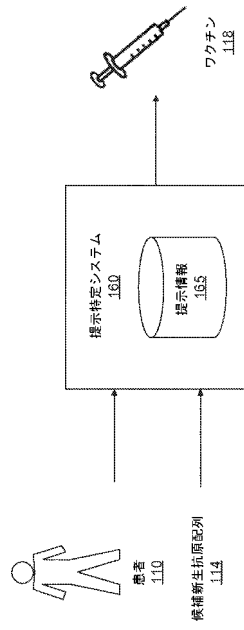
【図 1 F】



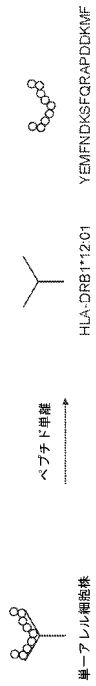
【 図 1 G 】



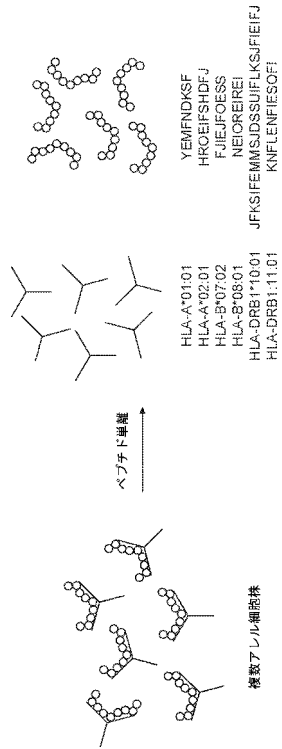
【 図 2 A 】



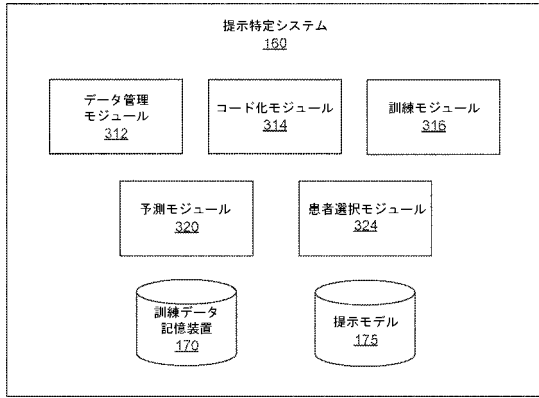
【 図 2 B 】



【 図 2 C 】



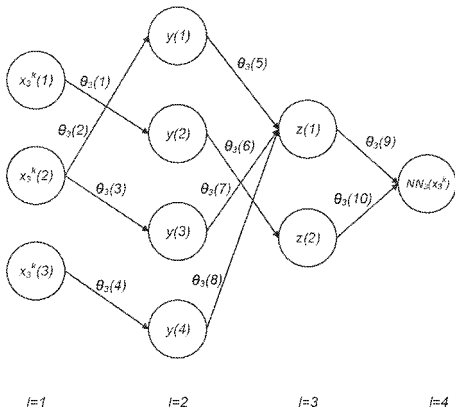
【 図 3 】



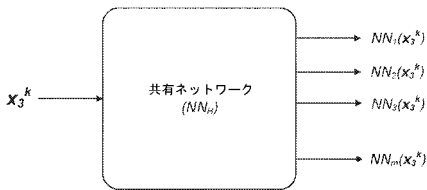
【 図 4 】

訓練データ 170A	アレル依存性 (x)		アレル非依存性 (w)		mRNA Q. (m・TPM)	ラベル (y)
	親和性 (p-nm)	安定性 (s-n)	アレル (p)	アレル (q)		
GCEIOWAREFLKEIGJ	1000	1	HLA-DRB3*01:01	FJELFISBOSJFIE	10 ²	提示されない
FIEUHFVI	1500	15	HLA-C*01:03	FEGRKUOCI	10 ³	提示される
FEWRHRJTRUUR	650	20	HLA-C*01:03	PJFIOEJOUJGEIO	10 ¹	提示される
QIEJQIEJIE	500	1	HLA-B*07:02	PJFIOEJOUJGEIO	1	提示される
	600	14	HLA-C*01:03			
	1200	7	HLA-A*01:01			

【 図 5 】



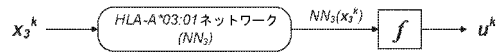
【 図 6 A 】



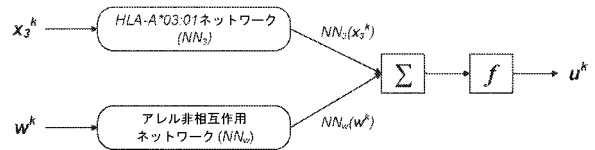
【 図 6 B 】



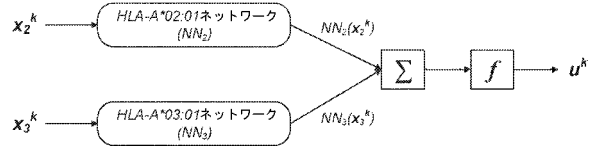
【 図 7 】



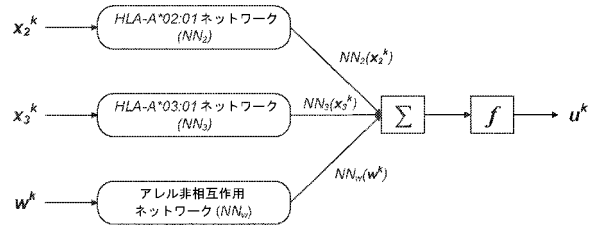
【 図 8 】



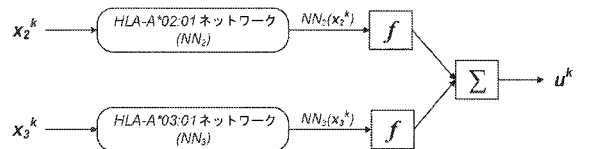
【 図 9 】



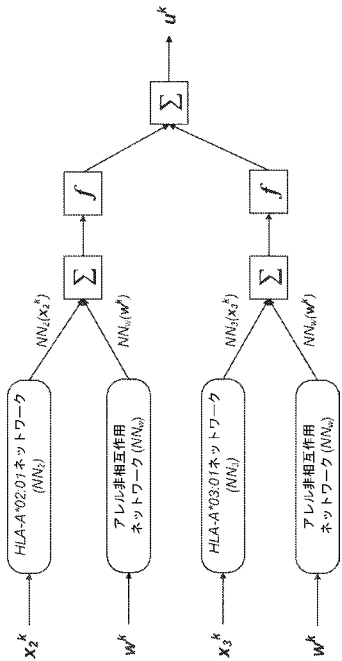
【 図 10 】



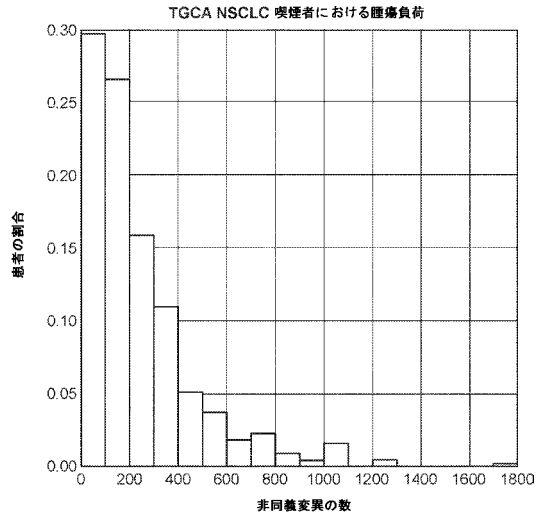
【 図 11 】



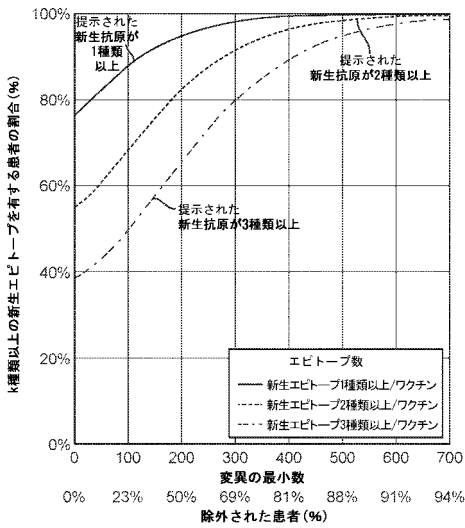
【図12】



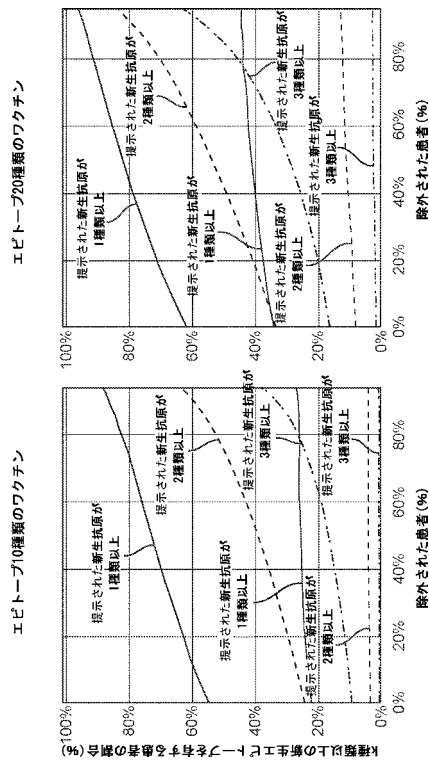
【図13A】



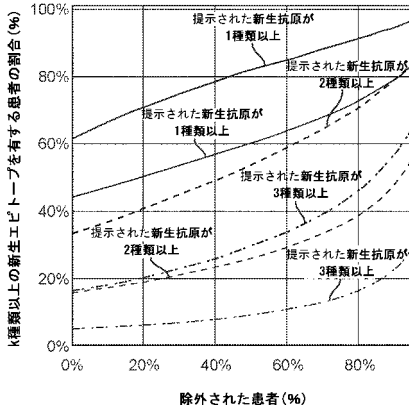
【図13B】



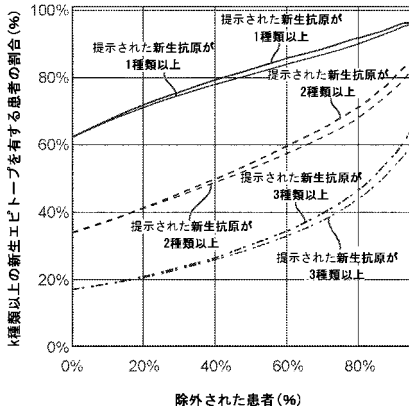
【図13C】



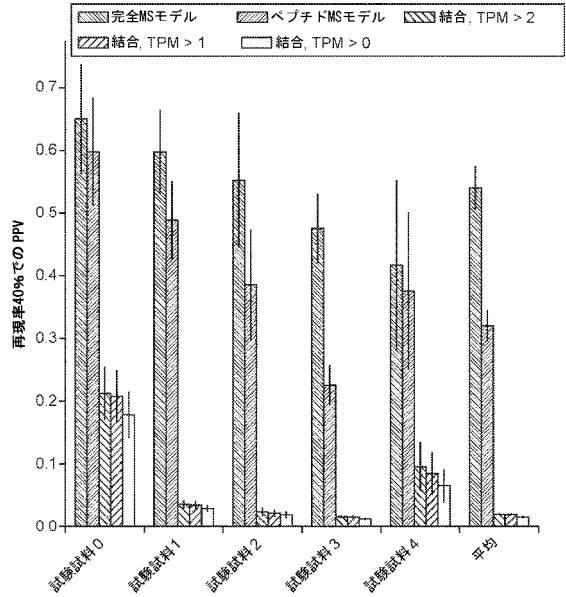
【図 1 3 D】



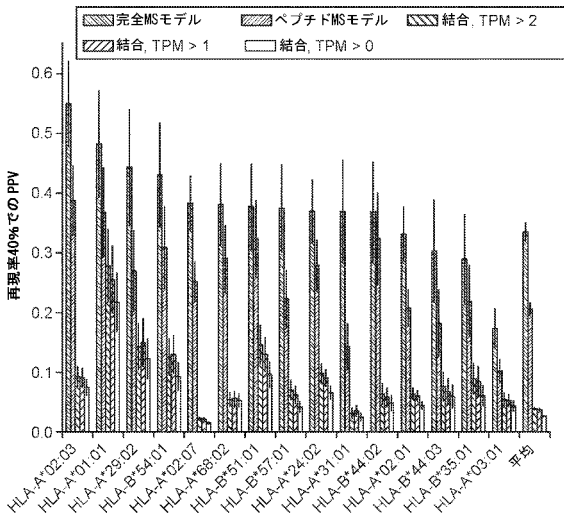
【図 1 3 E】



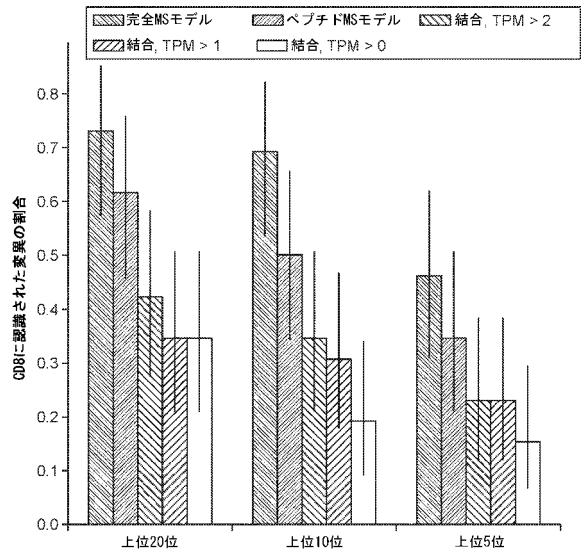
【図 1 4 A】



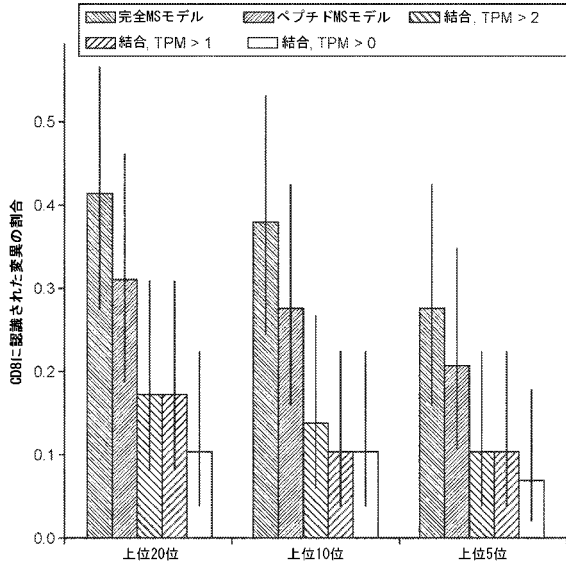
【図 1 4 B】



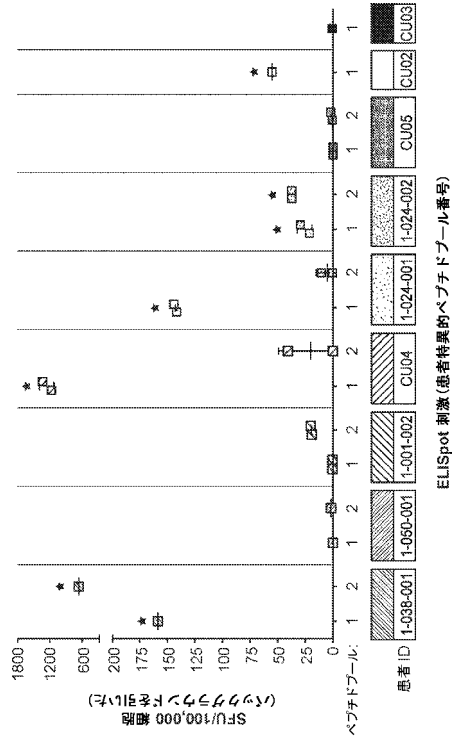
【図 1 4 C】



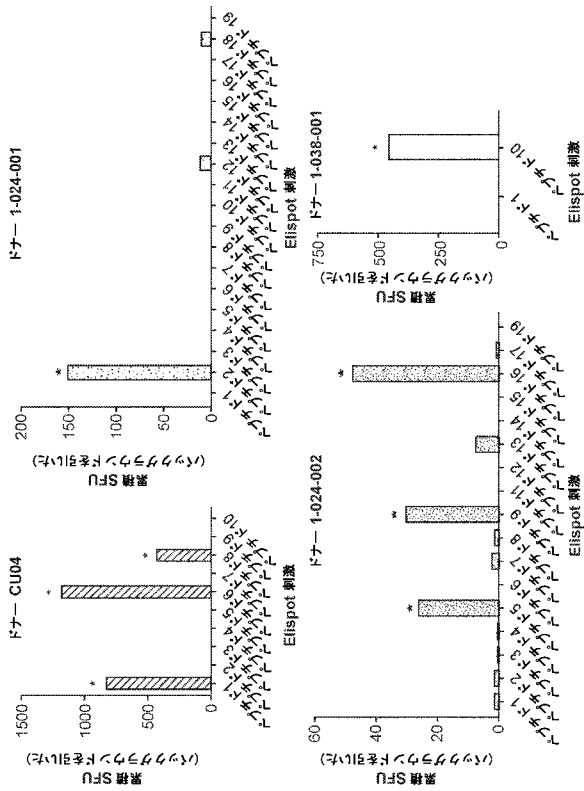
【図 14 D】



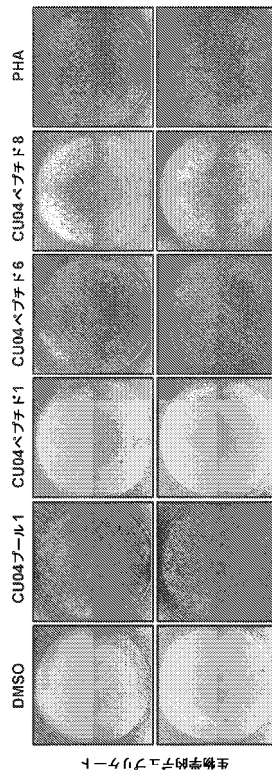
【図 15 A】



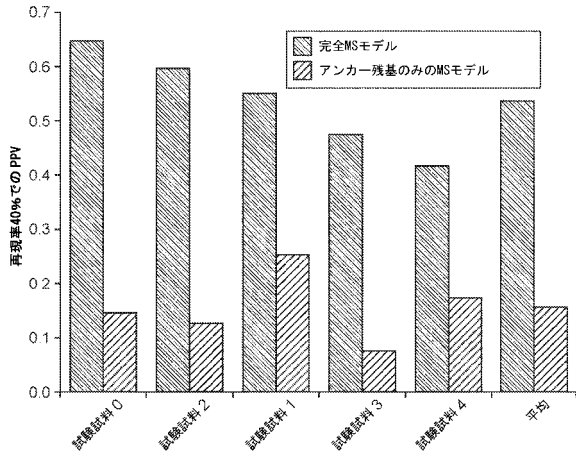
【図 15 B】



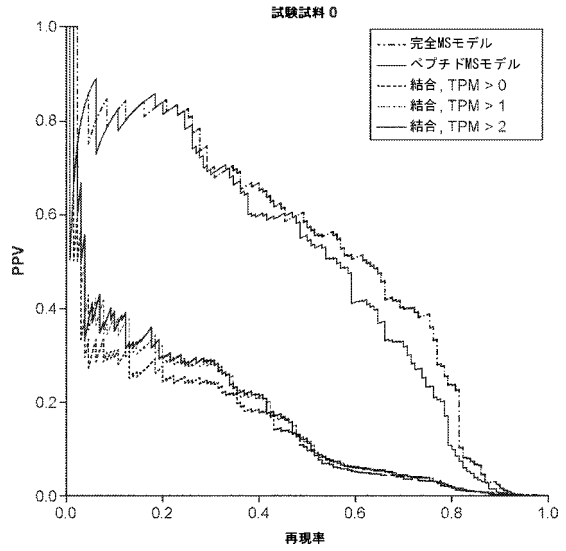
【図 15 C】



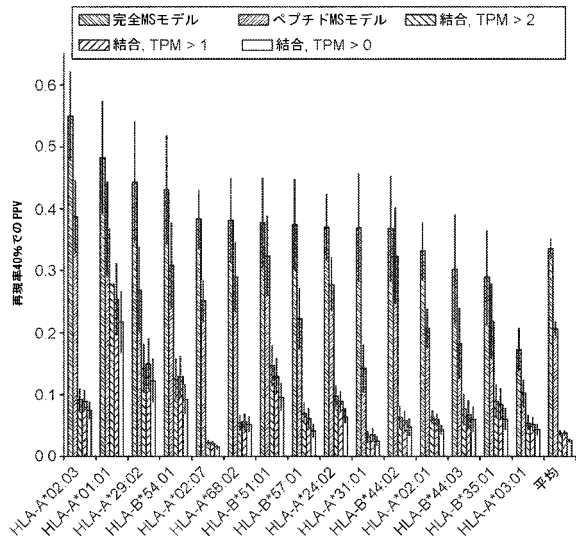
【図16】



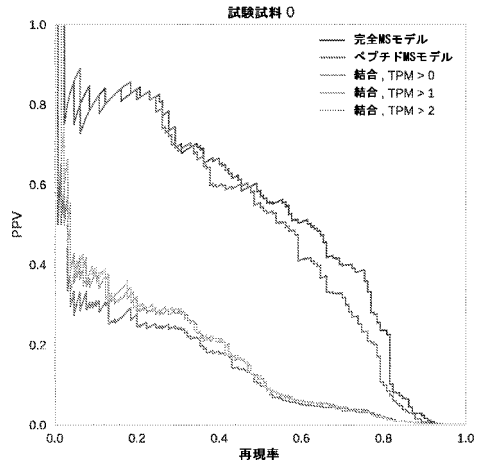
【図17A】



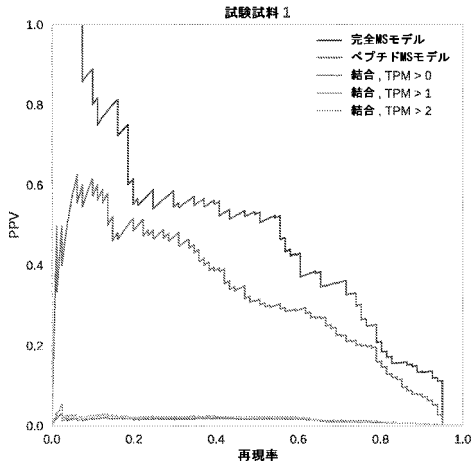
【図17B】



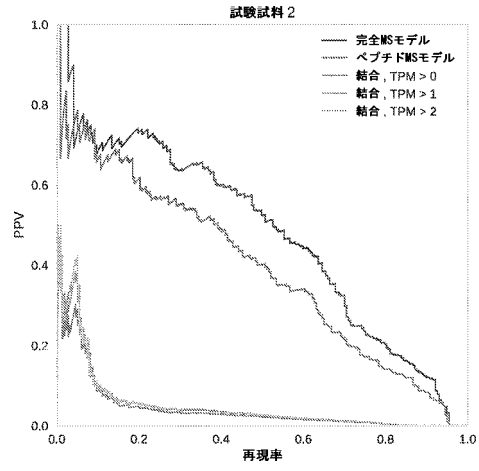
【図17C】



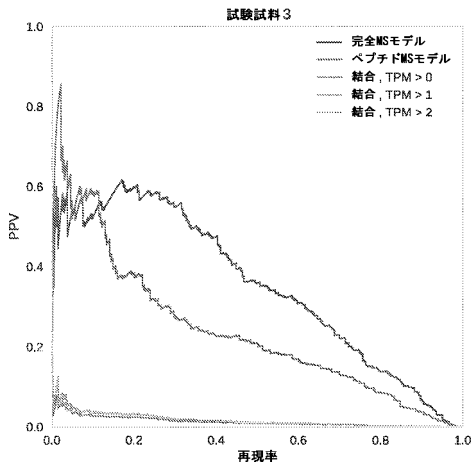
【図17D】



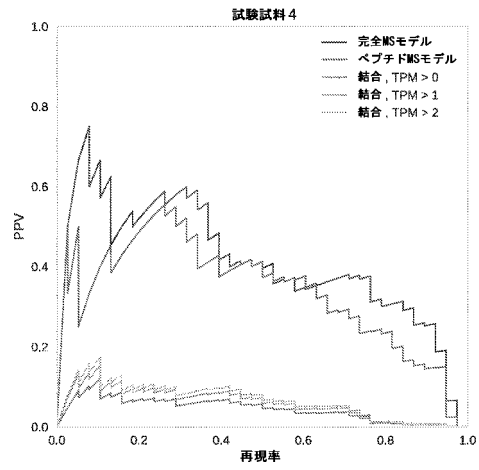
【図17E】



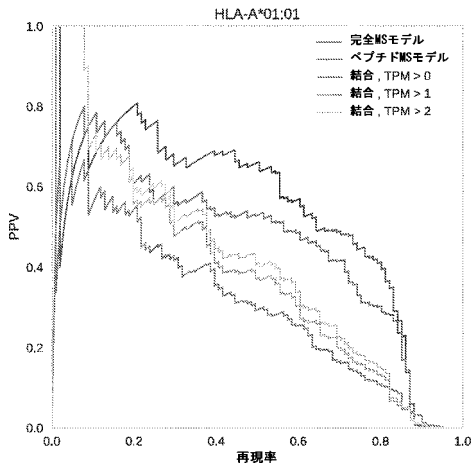
【図17F】



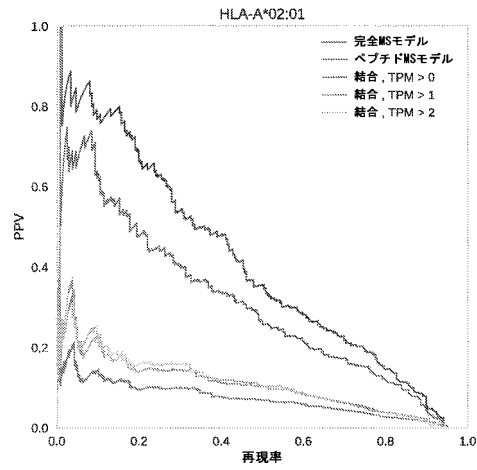
【図17G】



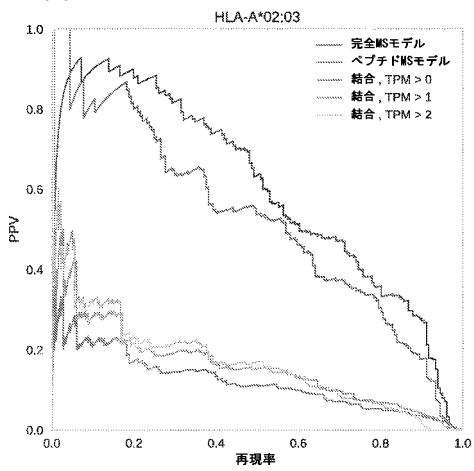
【図 17 H】



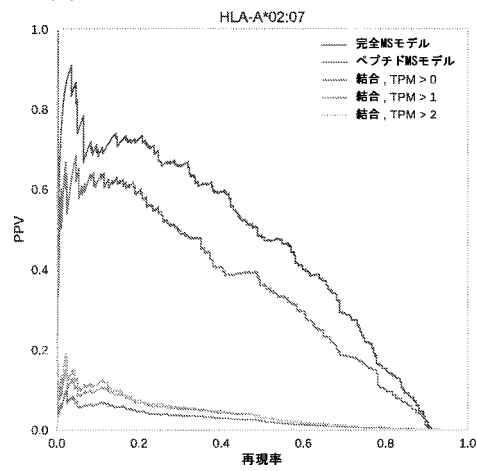
【図 17 I】



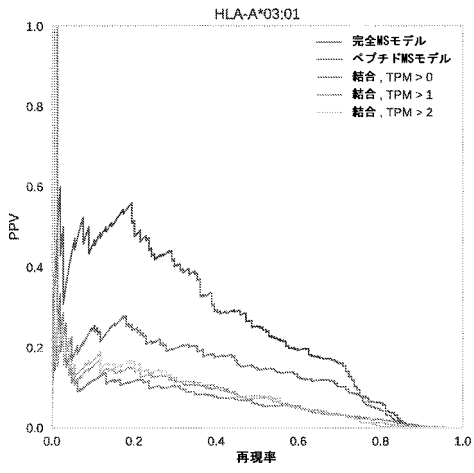
【図 17 J】



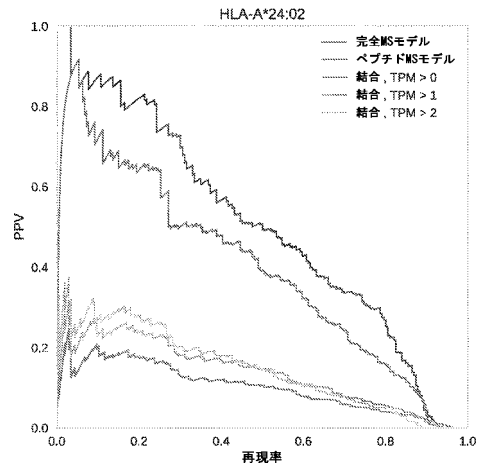
【図 17 K】



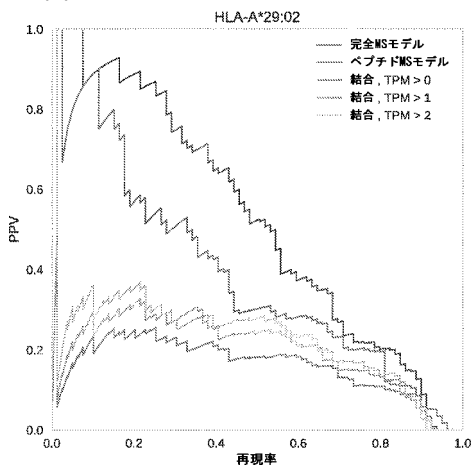
【 図 17 L 】



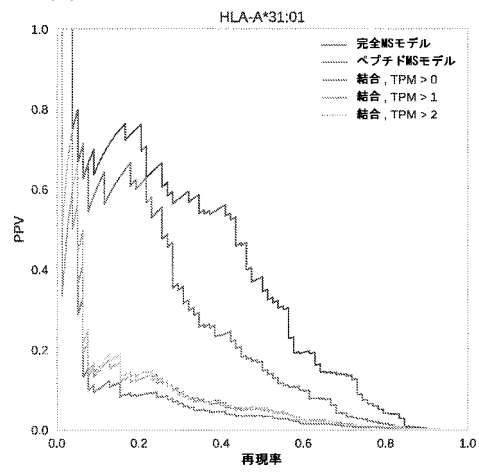
【 図 17 M 】



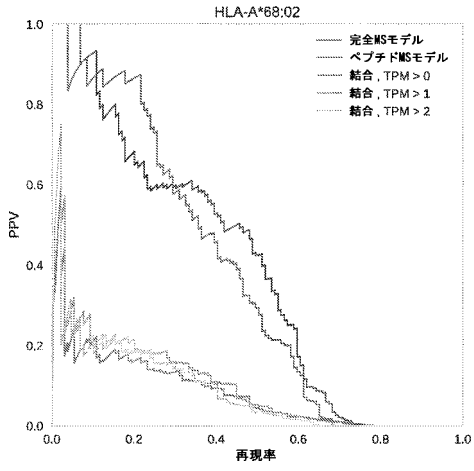
【 図 17 N 】



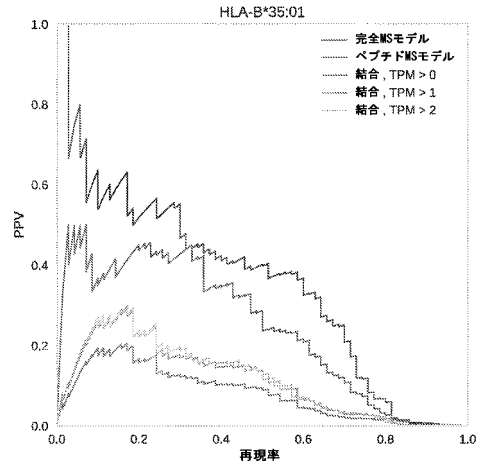
【 図 17 O 】



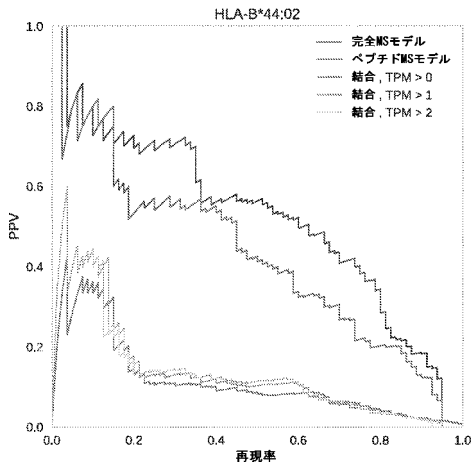
【図 17 P】



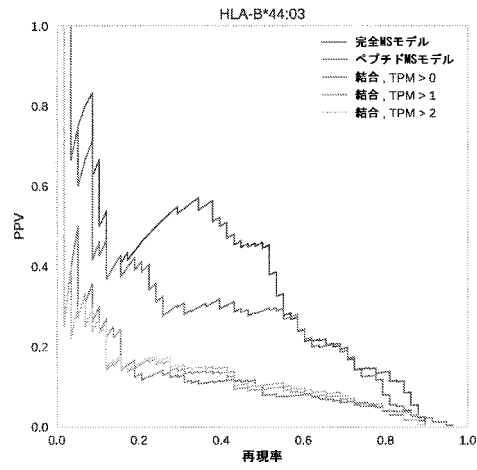
【図 17 Q】



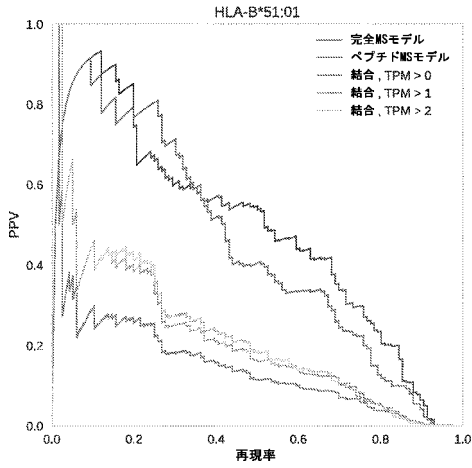
【図 17 R】



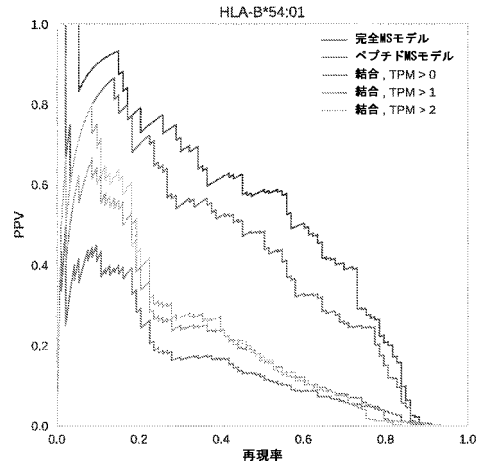
【図 17 S】



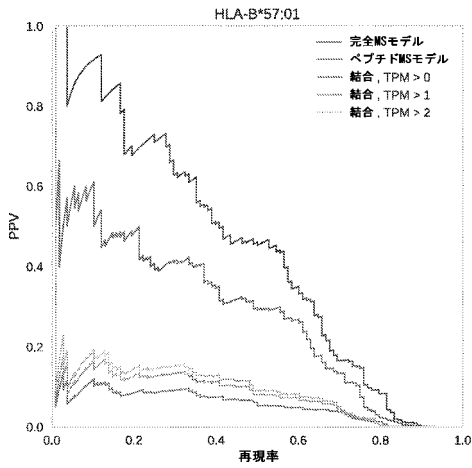
【 図 17 T 】



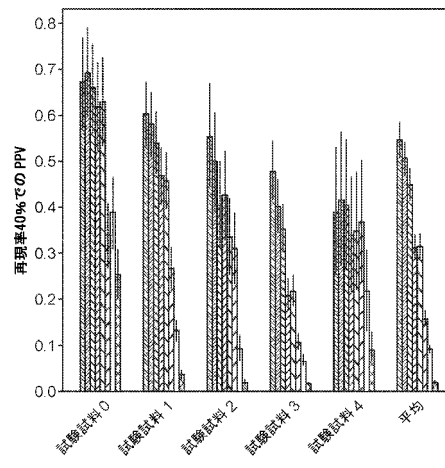
【 図 17 U 】



【 図 17 V 】

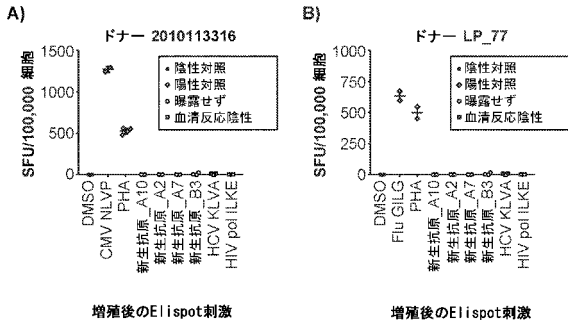


【 図 18 】

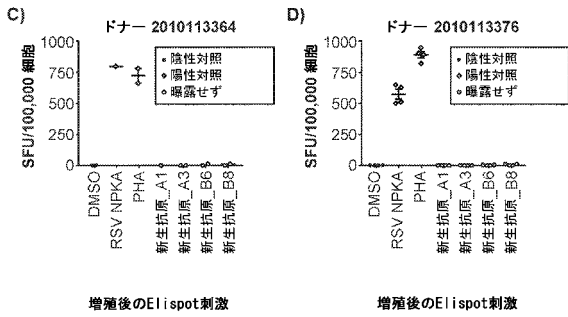
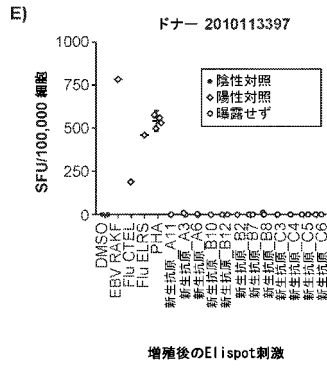


- 完全MSモデル
- MSモデル、フランキング配列なし
- MSモデル、フランキング配列または遺伝子毎係数なし
- すべての長さを一緒に訓練したペプチドのみのMSモデル
- それぞれの長さを別々に訓練したペプチドのみのMSモデル
- 直鎖ペプチドのみのMSモデル
- MixMHCpred 1.1
- 結合親和性

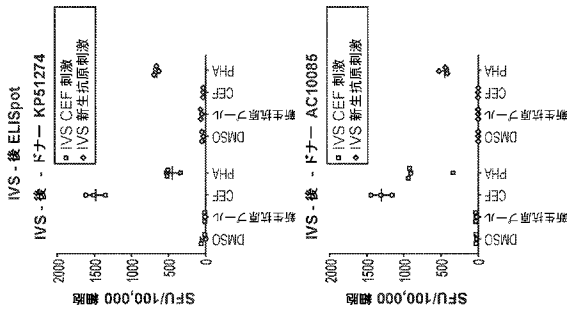
【 図 19 A - 1 】



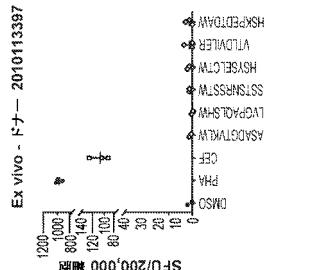
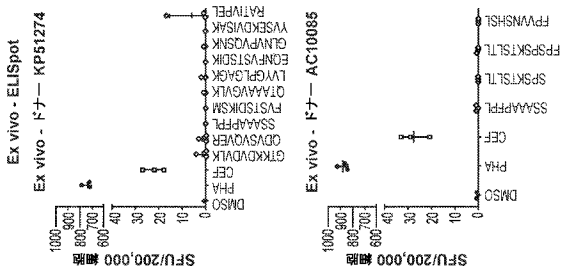
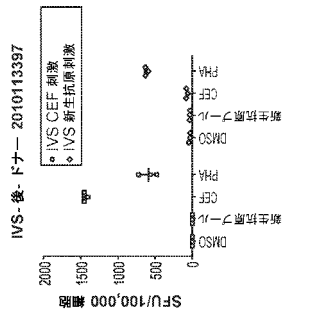
【 図 19 A - 2 】



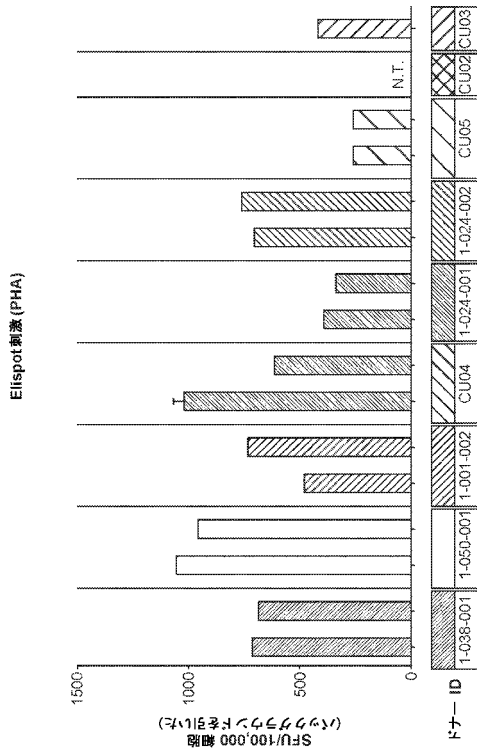
【 図 19 B - 1 】



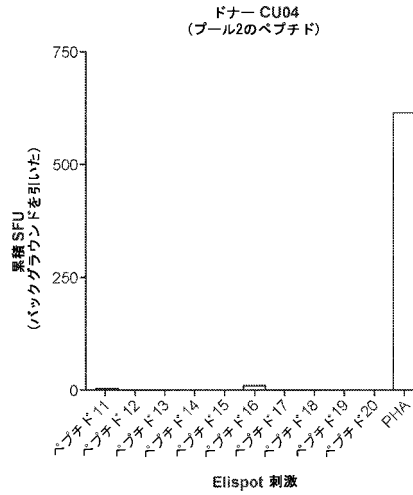
【 図 19 B - 2 】



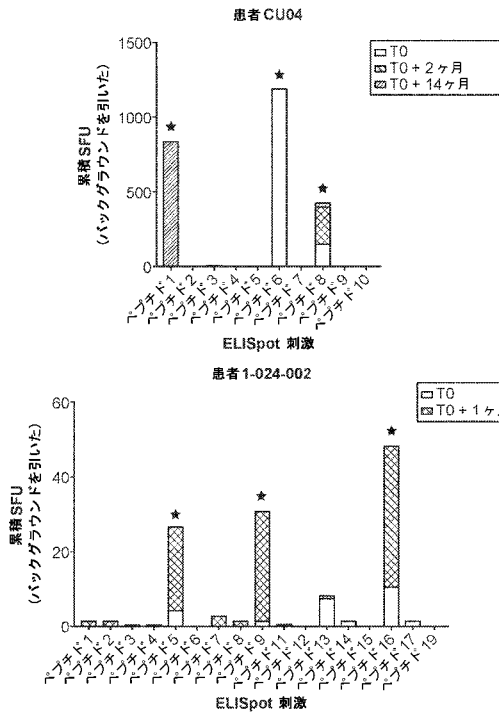
【図20】



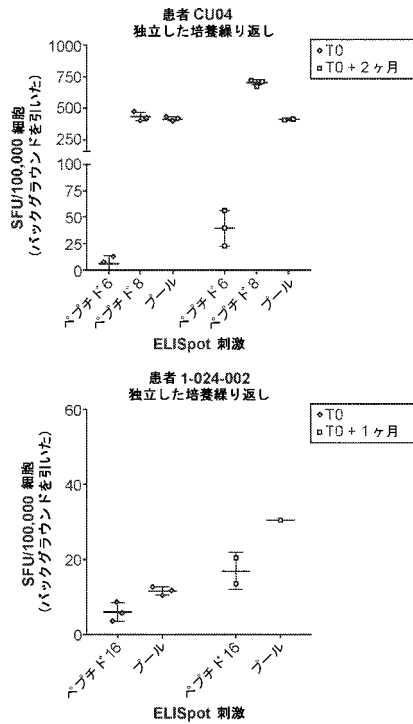
【図21A】



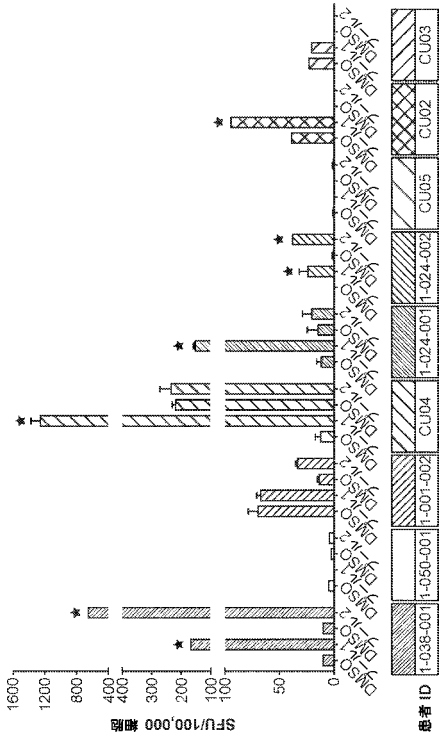
【図21B】



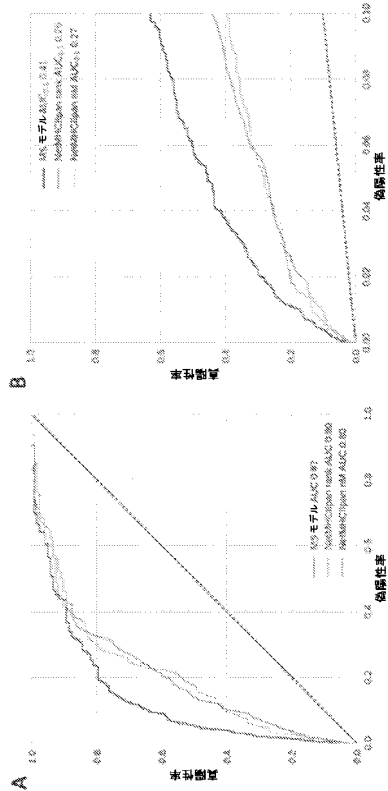
【図21C】



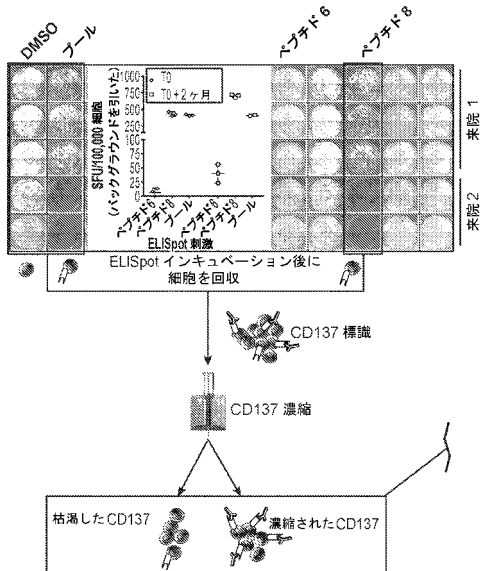
【図22】



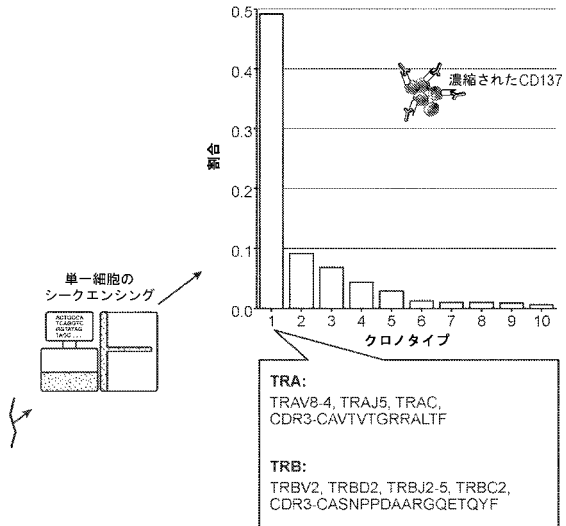
【図23】



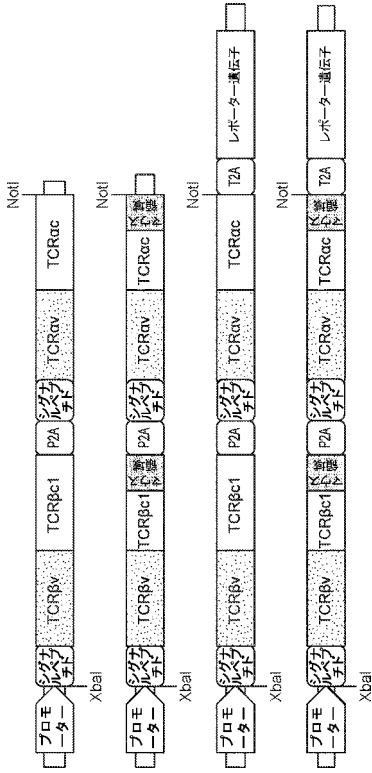
【図24-1】



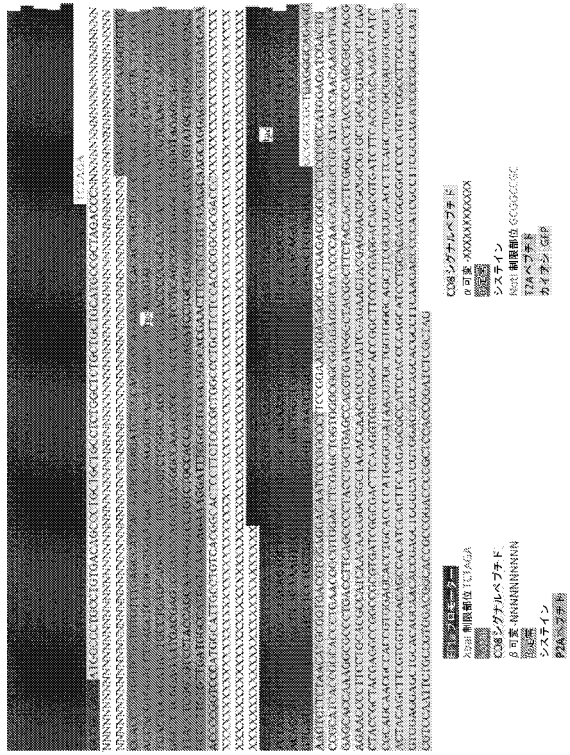
【図24-2】



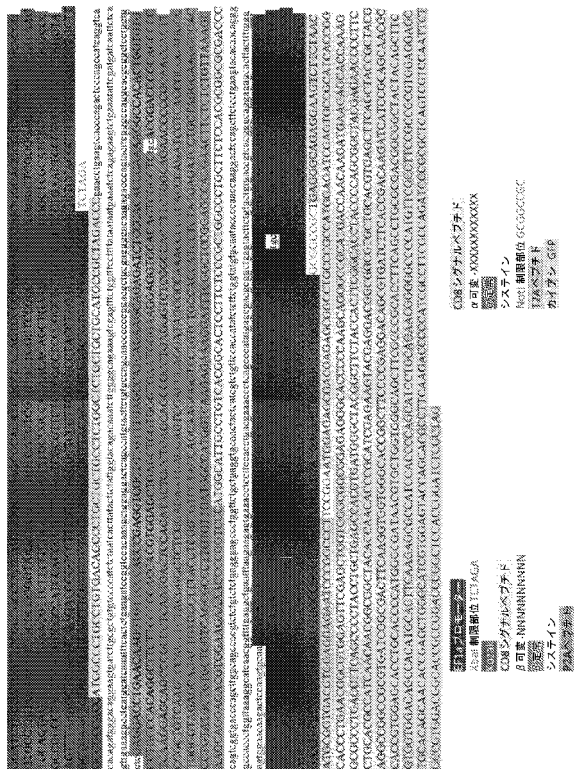
【 図 2 5 】



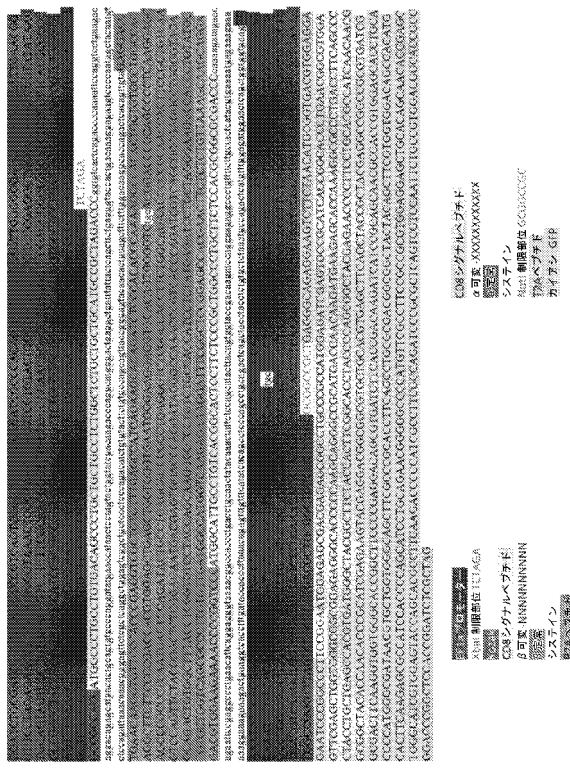
【 図 2 6 】



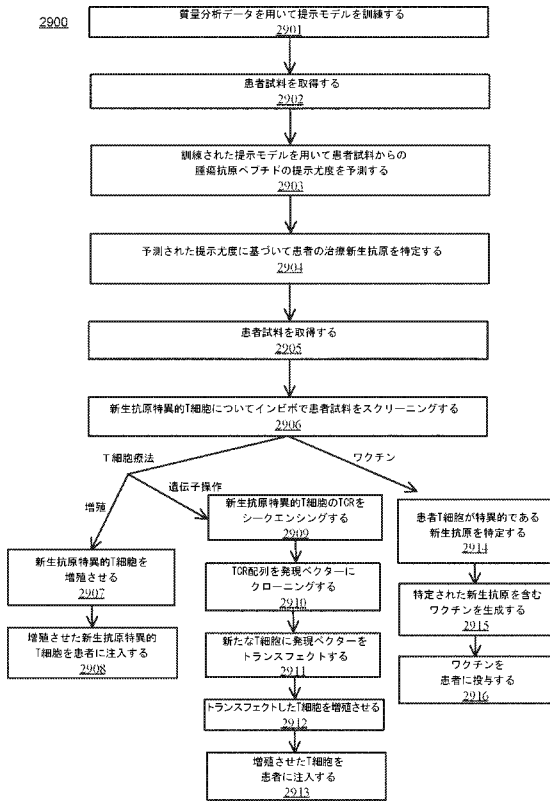
【 図 2 7 】



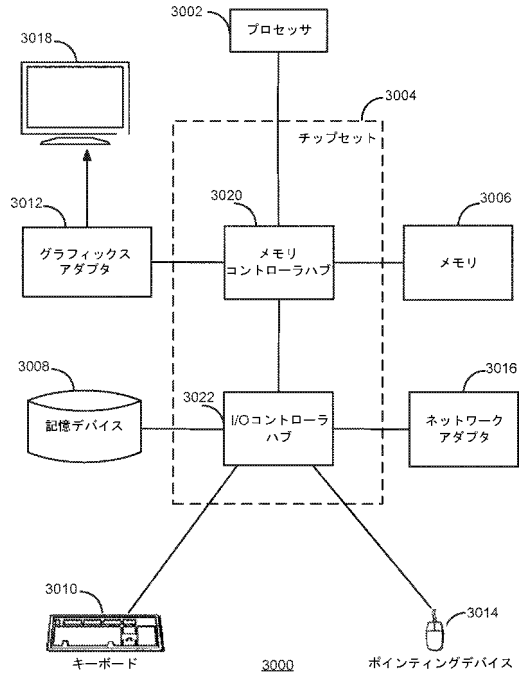
【 図 2 8 】



【図 29】



【図 30】



【手続補正書】

【提出日】令和2年5月20日(2020.5.20)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0010

【補正方法】変更

【補正の内容】

【0010】

TIL 新生エピトープデータセット及び予想される新生抗原反応性T細胞の特定タスクにおける本明細書に開示されるモデルの予測性能は、HLAのプロセッシング及び提示をモデル化することによって、治療に有用な新生エピトープの予測を得ることが今や可能であることを示す。要約すると、この作業は、抗原標的化免疫療法のための実用的なインシリコの抗原特定を可能とすることにより、患者の治療に向かう進行を加速するものである。

【本発明1001】

対象の1つ以上の腫瘍細胞に由来し腫瘍細胞の表面上に提示される可能性の高い少なくとも1種類の新生抗原に対して抗原特異的な1つ以上のT細胞を特定するための方法であって、

前記対象の前記腫瘍細胞及び正常細胞からエクソーム、トランスクリプトーム、または全ゲノムのヌクレオチドシーケンシングデータのうちの少なくとも1つを取得する工程であって、前記ヌクレオチドシーケンシングデータが、前記腫瘍細胞由来の前記ヌクレオチドシーケンシングデータと前記正常細胞由来の前記ヌクレオチドシーケンシングデータとを比較することにより特定された新生抗原のセットのそれぞれのペプチド配列を表すデータを取得するために用いられ、各新生抗原のペプチド配列が、前記ペプチド配列を前記対象の前記正常細胞から特定された対応する野生型のペプチド配列とは異なるものとする少なくとも1つの変化を含む、前記工程；

前記新生抗原のそれぞれの前記ペプチド配列を対応する数値ベクトルにコード化する工程であって、各数値ベクトルが、前記ペプチド配列を構成する複数のアミノ酸、及び前記ペプチド配列における前記アミノ酸の位置のセットに関する情報を含む、前記工程；

前記新生抗原のセットについて提示尤度のセットを生成するために、機械学習させた提示モデルに前記数値ベクトルをコンピュータプロセッサを使用して入力する工程であって、前記セット内の各提示尤度が、対応する新生抗原が1つ以上のMHCアレルによって前記対象の前記腫瘍細胞の前記表面上に提示される尤度を表し、前記機械学習させた提示モデルが、

複数の試料における各試料について、前記試料中に存在するものとして特定されたMHCアレルのセット内の少なくとも1つのMHCアレルに結合したペプチドの存在を測定する質量分析によって得られた標識、ならびに

前記試料のそれぞれについて、前記ペプチドを構成する複数のアミノ酸、及び前記ペプチドにおける前記アミノ酸の位置のセットに関する情報を含む数値ベクトルとしてコードされた訓練ペプチド配列

を含む訓練データセットに少なくとも基づいて特定される複数のパラメータと、入力として受け取られた前記数値ベクトルと、前記数値ベクトル及び前記パラメータに基づいた出力として生成された前記提示尤度との間の関係を表す関数とを含む、前記工程；

選択された新生抗原のセットを生成するために、前記新生抗原のセットのサブセットを、前記提示尤度のセットに基づいて選択する工程；

前記サブセット中の前記新生抗原のうちの少なくとも1つに対して抗原特異的である1つ以上のT細胞を特定する工程；ならびに

前記1つ以上の特定されたT細胞を返す工程を含む、前記方法。

[本発明1002]

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、前記1つ以上のMHCアレルのそれぞれについて、前記ペプチド配列の特定の位置の特定のアミノ酸に基づいて前記MHCアレルが前記新生抗原を提示するかどうかを示す依存性スコアを生成するために、前記機械学習させた提示モデルを前記新生抗原の前記ペプチド配列に適用することを含む、本発明1001の方法。

[本発明1003]

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、各MHCアレルについて、対応するMHCアレルが対応する新生抗原を提示する尤度を示す対応するアレル毎尤度を生成するために、前記依存性スコアを変換すること、及び前記新生抗原の前記提示尤度を生成するために、前記アレル毎尤度を組み合わせることをさらに含む、本発明1002の方法。

[本発明1004]

前記依存性スコアを変換することが、前記新生抗原の提示を前記1つ以上のMHCアレルにわたって相互排他的なものとしてモデル化する、本発明1003の方法。

[本発明1005]

前記機械学習させた提示モデルに前記数値ベクトルを入力する工程が、前記提示尤度を生成するために、前記依存性スコアの組み合わせを変換することであって、前記依存性スコアの組み合わせを変換することが、前記新生抗原の提示を前記1つ以上のMHCアレル間で干渉するものとしてモデル化する、前記変換することをさらに含む、本発明1002の方法。

[本発明1006]

前記提示尤度のセットが、少なくとも1つ以上のアレル非相互作用特性によってさらに特定され、かつ、

前記アレル非相互作用特性に基づいて、前記対応する新生抗原のペプチド配列が提示さ

れるかどうかを示す、前記アレル非相互作用特性についての依存性スコアを生成するために、前記機械学習させた提示モデルを前記アレル非相互作用特性に適用することをさらに含む、本発明1002～1005のいずれかの方法。

[本発明1007]

前記1つ以上のMHCアレルにおける各MHCアレルについての前記依存性スコアを、前記アレル非相互作用特性についての前記依存性スコアと組み合わせること、
対応するMHCアレルが対応する新生抗原を提示する尤度を示す、各MHCアレルについてのアレル毎尤度を生成するために、各MHCアレルについての前記組み合わせられた依存性スコアを変換すること、及び
前記提示尤度を生成するために、前記アレル毎尤度を組み合わせること
をさらに含む、本発明1006の方法。

[本発明1008]

前記MHCアレルのそれぞれについての前記依存性スコアと、前記アレル非相互作用特性についての前記依存性スコアとを組み合わせること、及び
前記提示尤度を生成するために、前記組み合わせられた依存性スコアを変換すること
をさらに含む、本発明1006の方法。

[本発明1009]

前記1つ以上のMHCアレルが2つ以上の異なるMHCアレルを含む、本発明1001～1008のいずれかの方法。

[本発明1010]

前記ペプチド配列が、アミノ酸9個以外の長さを有するペプチド配列を含む、本発明1001～1009のいずれかの方法。

[本発明1011]

前記ペプチド配列をコード化する工程が、ワン・ホットエンコーディングスキームを用いて前記ペプチド配列をコード化することを含む、本発明1001～1010のいずれかの方法。

[本発明1012]

前記複数の試料が、
(a) 単一のMHCアレルを発現するように操作された1つ以上の細胞株、
(b) 複数のMHCアレルを発現するように操作された1つ以上の細胞株、
(c) 複数の患者から得られた、または複数の患者に由来する1つ以上のヒト細胞株、
(d) 複数の患者から得られた新鮮なまたは凍結された腫瘍試料、及び
(e) 複数の患者から得られた新鮮なまたは凍結された組織試料
のうちの少なくとも1つを含む、本発明1001～1011のいずれかの方法。

[本発明1013]

前記訓練データセットが、
(a) 前記ペプチドのうちの少なくとも1つについてのペプチド-MHC結合親和性の測定値に関連するデータ、及び
(b) 前記ペプチドのうちの少なくとも1つについてのペプチド-MHC結合安定性の測定値に関連するデータ
のうちの少なくとも1つをさらに含む、本発明1001～1012のいずれかの方法。

[本発明1014]

前記提示尤度のセットが、RNA-seqまたは質量分析により測定される、前記対象における前記1つ以上のMHCアレルの少なくとも発現レベルによってさらに特定される、本発明1001～1013のいずれかの方法。

[本発明1015]

前記提示尤度のセットが、
(a) 前記新生抗原のセット内の新生抗原と前記1つ以上のMHCアレルとの間の予測される親和性、及び
(b) 前記新生抗原コード化ペプチド-MHC複合体の予測される安定性
のうちの少なくとも1つを含む特性によってさらに特定される、本発明1001～1014のい

れかの方法。

[本発明1016]

前記数値的尤度のセットが、

(a) そのソースタンパク質配列内の、前記新生抗原コード化ペプチド配列に隣接する C 末端配列、及び

(b) そのソースタンパク質配列内の、前記新生抗原コード化ペプチド配列に隣接する N 末端配列

のうちの少なくとも1つを含む特性によってさらに特定される、本発明1001～1015のいずれかの方法。

[本発明1017]

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記腫瘍細胞表面上に提示される尤度が増大している新生抗原を選択することを含む、本発明1001～1016のいずれかの方法。

[本発明1018]

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記対象において腫瘍特異的な免疫応答を誘導することができる尤度が増大している新生抗原を選択することを含む、本発明1001～1017のいずれかの方法。

[本発明1019]

前記選択された新生抗原のセットを選択する工程が、前記提示モデルに基づいて、選択されない新生抗原と比較してプロフェッショナル抗原提示細胞 (A P C) によってナイーブ T 細胞に提示されることができる尤度が増大している新生抗原を選択することを含み、任意で、前記 A P C が樹状細胞 (D C) である、本発明1001～1018のいずれかの方法。

[本発明1020]

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して中枢性寛容または末梢性寛容によって阻害される尤度が減少している新生抗原を選択することを含む、本発明1001～1019のいずれかの方法。

[本発明1021]

前記選択された新生抗原のセットを選択する工程が、前記機械学習させた提示モデルに基づいて、選択されない新生抗原と比較して前記対象において正常組織に対する自己免疫応答を誘導することができる尤度が減少している新生抗原を選択することを含む、本発明1001～1020のいずれかの方法。

[本発明1022]

前記1つ以上の腫瘍細胞が、肺癌、メラノーマ、乳癌、卵巣癌、前立腺癌、腎臓癌、胃癌、結腸癌、精巣癌、頭頸部癌、膵臓癌、脳癌、B細胞リンパ腫、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、T細胞リンパ球性白血病、非小細胞肺癌、及び小細胞肺癌からなる群から選択される、本発明1001～1021のいずれかの方法。

[本発明1023]

前記選択された新生抗原のセットから個別化がんワクチンを構築するための出力を生成することをさらに含む、本発明1001～1022のいずれかの方法。

[本発明1024]

前記個別化がんワクチン用の出力が、前記選択された新生抗原のセットをコードした少なくとも1つのペプチド配列または少なくとも1つのヌクレオチド配列を含む、本発明1023の方法。

[本発明1025]

前記機械学習させた提示モデルが、ニューラルネットワークモデルである、本発明1001～1024のいずれかの方法。

[本発明1026]

前記ニューラルネットワークモデルが M H C アレルについての複数のネットワークモデ

ルを含み、それぞれのネットワークモデルが複数のMHCアレルのうちの対応するMHCアレルに割り当てられ、かつ1つ以上の層に配置された一連のノードを含む、本発明1025の方法。

[本発明1027]

前記ニューラルネットワークモデルが、前記ニューラルネットワークモデルのパラメータを更新することによって訓練され、少なくとも2つのネットワークモデルの前記パラメータが、少なくとも1回の訓練イテレーションについて一緒に更新される、本発明1026の方法。

[本発明1028]

前記機械学習させた提示モデルが、ノードの1つ以上の層を含むディープラーニングモデルである、本発明1025～1027のいずれかの方法。

[本発明1029]

前記1つ以上のT細胞を特定する工程が、前記1つ以上のT細胞を増殖させる条件下で前記1つ以上のT細胞を前記サブセット中の前記新生抗原のうちの1つ以上と共培養することを含む、本発明1001～1028のいずれかの方法。

[本発明1030]

前記1つ以上のT細胞を特定する工程が、前記T細胞とMHCマルチマーとの結合を可能とする条件下で前記1つ以上のT細胞を前記サブセット中の前記新生抗原のうちの1つ以上を含む前記MHCマルチマーと接触させることを含む、本発明1001～1029のいずれかの方法。

[本発明1031]

前記1つ以上の特定されたT細胞の1つ以上のT細胞受容体(TCR)を特定することをさらに含む、本発明1001～1030のいずれかの方法。

[本発明1032]

前記1つ以上のT細胞受容体を特定することが、前記1つ以上の特定されたT細胞のT細胞受容体配列をシーケンシングすることを含む、本発明1031の方法。

[本発明1033]

本発明1001～1032のいずれかの前記サブセット中の少なくとも1つの選択された新生抗原に対して抗原特異的な単離T細胞。

[本発明1034]

前記1つ以上の特定されたT細胞受容体のうちの少なくとも1つを発現するように複数のT細胞を遺伝子操作する工程、

前記複数のT細胞を増殖させる条件下で前記複数のT細胞を培養する工程、及び前記増殖させたT細胞を前記対象に注入する工程
をさらに含む、本発明1032の方法。

[本発明1035]

前記1つ以上の特定されたT細胞受容体のうちの少なくとも1つを発現するように前記複数のT細胞を遺伝子操作する工程が、

前記1つ以上の特定されたT細胞のT細胞受容体配列を発現ベクターにクローニングすること、及び

前記複数のT細胞のそれぞれに前記発現ベクターをトランスフェクトすること
を含む、本発明1034の方法。

[本発明1036]

前記1つ以上の特定されたT細胞を増殖させる条件下で前記1つ以上の特定されたT細胞を培養する工程、及び

前記増殖させたT細胞を前記対象に注入する工程
をさらに含む、本発明1001～1035のいずれかの方法。

[本発明1037]

前記サブセット中の前記新生抗原のうちの少なくとも1つに対して抗原特異的である前記1つ以上のT細胞が、前記対象からの5～30mLの全血を使用して特定される、本発明10

01～1036のいずれかの方法。

[本発明1038]

前記新生抗原のサブセットが最大で20種類の新生抗原を含み、前記1つ以上の特定されたT細胞が、前記新生抗原のサブセット中の少なくとも2種類の新生抗原を認識する、本発明1001～1037のいずれかの方法。

[本発明1039]

前記1つ以上のMHCアレルがクラスI MHCアレルである、本発明1001～1038のいずれかの方法。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0012

【補正方法】変更

【補正の内容】

【0012】

【図1A】新生抗原の特定に対する現在の臨床的アプローチを示す。

【図1B】予測された結合ペプチドのうち、腫瘍細胞上に提示されるものは5%未満であることを示す。

【図1C】新生抗原予測の特異性の問題の影響を示す。

【図1D】結合予測が、新生抗原の特定に充分ではないことを示す。

【図1E】ペプチド長の関数としてのMHC-I提示の確率を示す。

【図1F】Promega社のダイナミックレンジ標準から生成された、例示的なペプチドスペクトルを示す。図1Fは、SEQ ID NO:1を開示する。

【図1G】特性の追加が、いかにモデルの陽性適中率を増大させるかを示す。

【図2A】一実施形態による、患者におけるペプチド提示の尤度を特定するための環境の概略である。

【図2B】一実施形態による、提示情報を取得する方法を説明する。図2Bは、SEQ ID NO:28を開示する。

【図2C】一実施形態による、提示情報を取得する方法を説明する。図2Cは、出現順にそれぞれSEQ ID NO:3～8を開示する。

【図3】一実施形態による、提示特定システムのコンピュータ論理構成要素を説明する、ハイレベルブロック図である。

【図4】一実施形態による、訓練データの例示的なセットを説明する。図4は、出現順にそれぞれSEQ ID NO:10～13として「ペプチド配列」、SEQ ID NO:15、29～30、及び30として「C末端フランキング配列を開示する。

【図5】MHCアレルに関連した例示的なネットワークモデルを説明する。

【図6A】一実施形態による、MHCアレルによって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。

【図6B】別の実施形態による、MHCアレルによって共有される例示的なネットワークモデル $NN_H(\cdot)$ を説明する。

【図7】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

【図8】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

【図9】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

【図10】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

【図11】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提示尤度の生成を説明する。

【図12】例示的なネットワークモデルを用いた、MHCアレルに関連したペプチドの提

示尤度の生成を説明する。

【図13A】NSCLC患者における変異負荷の標本度数分布を示す。

【図13B】一実施形態による、患者が最小変異負荷を満たすかどうかの選択基準に基づいて選択された患者に対してシミュレートしたワクチン中の提示新生抗原の数を示す。

【図13C】一実施形態による、提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、従来技術のモデルによって特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。

【図13D】HLA-A*02:01についての単一アレル毎提示モデルに基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者と、HLA-A*02:01及びHLA-B*07:02についてのアレル毎提示モデルの両方に基づいて特定された治療サブセットを含むワクチンに関連付けられた選択された患者との間のシミュレートしたワクチン中の提示新生抗原の数を比較したものである。一実施形態によれば、ワクチン容量は、 $v = 20$ 個のエピトープに設定される。

【図13E】一実施形態による、変異負荷に基づいて選択された患者と、期待効用値スコアにより選択された患者との間でシミュレートしたワクチン中の提示新生抗原の数を比較したものである。

【図14A】各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及び $TPM > 0, 1$ 、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの再現率40%における陽性適中率(PPV)を比較したものである。

【図14B】各試験試料が1:10,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む15種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及び $TPM > 0, 1$ 、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの再現率40%におけるPPVを比較したものである。

【図14C】それぞれの試験試料が少なくとも1つの既存のT細胞応答を示す患者から採取された12種類の異なる試験試料からなる試験セットについて、「完全MSモデル」、「ペプチドMSモデル」、及び $TPM > 0, 1$ 、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルによって特定された、上位5、10、及び20位にランクされた体細胞変異について、T細胞(例えば、既存のT細胞応答)によって認識された体細胞変異の割合を比較したものである。

【図14D】それぞれの試験試料が少なくとも1つの既存のT細胞応答を示す患者から採取された12種類の異なる試験試料からなる試験セットについて、「完全MSモデル」、「ペプチドMSモデル」、及び $TPM > 0, 1$ 、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルによって特定された、上位5、10、及び20位にランクされた最小新生エピトープについて、T細胞(例えば、既存のT細胞応答)によって認識された最小新生エピトープの割合を比較したものである。

【図15A】9人の患者における患者特異的新生抗原ペプチドプールに対するT細胞応答の検出を示す。

【図15B】4人の患者における個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図15C】患者CU04についてELISpotウェルの例示的画像を示す。

【図16】各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」及び「アンカー残基のみのMSモデル」の再現率40%における陽性適中率(PPV)を比較したものである。

【図17A】各モデルを図14Aからの試験試料0で試験する場合の「完全MSモデル」

、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17B】各試験試料が1:5,000の提示ペプチドと非提示ペプチドとの比を有する単一アレル細胞株試験データセットからの除外ペプチドを含む15種類の異なる試験試料からなる試験セットで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの再現率40%におけるPPVを比較したものである。

【図17C】各モデルを図14Aからの試験試料0で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17D】各モデルを図14Aからの試験試料1で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17E】各モデルを図14Aからの試験試料2で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17F】各モデルを図14Aからの試験試料3で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17G】各モデルを図14Aからの試験試料4で試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17H】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*01:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17I】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17J】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:03細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17K】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*02:07細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM > 0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry 1.2.0結合親和性モデルの完全な適合率 - 再現率曲線を示す。

【図17L】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14B

のHLA-A*03:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17M】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*24:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17N】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*29:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17O】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*31:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17P】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*68:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17Q】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*35:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17R】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*44:02細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17S】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*44:03細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17T】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*51:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17U】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14BのHLA-A*54:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図17V】1:10,000の提示ペプチドと非提示ペプチドとの比を有する図14B

のHLA-A*57:01細胞株の試験データセットからの除外ペプチドで各モデルを試験する場合の「完全MSモデル」、「ペプチドMSモデル」、及びTPM>0、1、及び2の3つの異なる遺伝子発現閾値を有するMHCFlurry1.2.0結合親和性モデルの完全な適合率-再現率曲線を示す。

【図18】各試験試料が1:2500の提示ペプチドと非提示ペプチドとの比を有する除外腫瘍試料を含む5種類の異なる試験試料からなる図14Aの試験セットで各モデルを試験する場合のヒト腫瘍におけるHLA提示ペプチドをモデル化するためのMSモデルの異なるバージョン及び初期のアプローチ²⁹の再現率40%における陽性適中率(PPV)を比較したものである。

【図19A-1】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図19A-2】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。

【図19B-1】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。図19Bは、出現順にそれぞれ27、24、21~22、31~36、21、37~45を開示する。

【図19B-2】HLAを一致させた健康なドナーにおいて新生抗原を用いた対照実験の結果を示す。図19Bは、出現順にそれぞれ27、24、21~22、31~36、21、37~45を開示する。

【図20】図15Aに示される各ドナー及び各インビトロ増殖についてPHA陽性対照に対するT細胞応答の検出を示す。

【図21A】患者CU04におけるプール#2中のそれぞれの個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図21B】患者CU04の3回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的新生抗原ペプチドに対するT細胞応答の検出を示す。

【図21C】患者CU04の2回の来院のそれぞれにおける、また、患者1-024-002の2回の来院のそれぞれにおける(各来院は異なる時点で行われる)、個々の患者特異的新生抗原ペプチドに対する、及び患者特異的新生抗原ペプチドプールに対するT細胞応答の検出を示す。

【図22】図15Aの患者について2つの患者特異的新生抗原ペプチドプール及びDMSO陰性対照に対するT細胞応答の検出を示す。

【図23】「MSモデル」、HLA-DRB1*15:01及びHLA-DRB5*01:01にわたって最も低いNetMHCIIpanパーセントイルランクを採用する、「NetMHCIIpanrank」:NetMHCIIpan3.1⁷⁷、及びHLA-DRB1*15:01及びHLA-DRB5*01:01にわたってnM単位の最も強い親和性を採用する、「NetMHCIIpannM」:NetMHCIIpan3.1の、HLA-DRB1*15:01/HLA-DRB5*01:01試験データセット内のペプチドのランク付けにおける予測性能を比較したものである。

【図24-1】NSCLC患者の末梢血由来の新生抗原特異的メモリーT細胞のTCRをシーケンシングするための方法を示す。図24は、出現順にそれぞれ46~48を開示する。

【図24-2】NSCLC患者の末梢血由来の新生抗原特異的メモリーT細胞のTCRをシーケンシングするための方法を示す。図24は、出現順にそれぞれ46~48を開示する。

【図25】TCRをレシピエント細胞に導入するためのTCRコンストラクトの例示的な実施形態を示す。

【図26】治療法の開発のためにTCRを発現系にクローニングするための例示的なP526コンストラクト骨格のヌクレオチド配列を示す。図26はSEQ ID NO:49を開示する。

【図 27】治療法の開発のために患者の新生抗原特異的 TCR のクロナタイプ 1 を発現系にクローニングするための例示的なコンストラクト配列を示す。図 27 は SEQ ID NO: 50 を開示する。

【図 28】治療法の開発のために患者の新生抗原特異的 TCR のクロナタイプ 3 を発現系にクローニングするための例示的なコンストラクト配列を示す。図 28 は SEQ ID NO: 51 を開示する。

【図 29】一実施形態により、患者に個別化された新生抗原特異的治療を行うための方法のフローチャートである。

【図 30】図 1 及び 3 に示した実体を実施するための例示的なコンピュータを説明する。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0173

【補正方法】変更

【補正の内容】

【0173】

V I . B . 1 . 総合的 H L A ペプチドシーケンシングのための M S 検出限界の研究

ペプチド Y V Y V A D V A A K (SEQ ID NO: 1) を用いて、何が検出の限界かを、LC カラム上にロードした様々な量のペプチドを用いて決定した。試験したペプチドの量は、1 p m o l、1 0 0 f m o l、1 0 f m o l、1 f m o l、及び 1 0 0 a m o l であった。(表 1) 結果を図 1 F に示す。これらの結果は、検出の最低限界 (L o D) がアトモルの範囲 ($1 0^{-18}$) にあること、ダイナミックレンジが 5 桁に及ぶこと、及び、シグナル対ノイズが、低いフェムトモル範囲 ($1 0^{-15}$) でシーケンシングに十分であるように見えることを示す。

【手続補正 4】

【補正対象書類名】明細書

【補正対象項目名】0177

【補正方法】変更

【補正の内容】

【0177】

提示特定システム 160 は、1 つ以上の提示モデルを通して提示尤度を決定する。具体的には、提示モデルは、所定のペプチド配列が、関連する M H C アレルのセットについて提示されるかどうかの尤度を生成し、尤度は、記憶装置 165 に保存された提示情報に基づいて生成される。例えば、提示モデルは、ペプチド配列「Y V Y V A D V A A K (SEQ ID NO: 1)」が、試料の細胞表面上のアレルのセット H L A - A * 0 2 : 0 1、H L A - A * 0 3 : 0 1、H L A - B * 0 7 : 0 2、H L A - B * 0 8 : 0 3、H L A - C * 0 1 : 0 4 について提示されるかどうかの尤度を生成し得る。提示情報 165 は、M H C アレルによってペプチドが提示されるようにこれらのペプチドが様々なタイプの M H C アレルに結合するかどうかについての情報を含み、これは、モデルにおいて、ペプチド配列中のアミノ酸の位置に応じて決定される。提示モデルは、提示情報 165 に基づいて、認識されていないペプチド配列が、M H C アレルの関連するセットと結合して提示されるかどうかを予測することができる。上記に述べたように、提示モデルはクラス I 及びクラス II M H C アレルの両方に適用することができる。

【手続補正 5】

【補正対象書類名】明細書

【補正対象項目名】0179

【補正方法】変更

【補正の内容】

【0179】

V I I . B . 1 . アレル相互作用情報

アレル相互作用情報は、主として、ヒト、マウスなど由来の 1 つ以上の特定された M H

C分子によって提示されていることが公知である、特定されたペプチド配列を含む。注目すべきことに、これは、腫瘍試料から取得されたデータを含んでもよく、または含まなくてもよい。提示されたペプチド配列は、単一のMHCアレルを発現する細胞から特定されてもよい。この例において、提示されたペプチド配列は、概して、あらかじめ決定されたMHCアレルを発現するように操作されてその後合成タンパク質に曝露された単一アレル細胞株から収集される。MHCアレル上に提示されたペプチドは、酸溶出などの技法によって単離され、質量分析により特定される。図2Bは、あらかじめ決定されたMHCアレルHLA-DRB1*12:01上に提示された例示的なペプチド

YEMFNDSQRAPDDKMF (SEQ ID NO: 2)

が単離され、質量分析により特定される、この例を示す。この状況においては、ペプチドが、単一のあらかじめ決定されたMHCタンパク質を発現するように操作された細胞を通して特定されるため、提示されたペプチドとそれが結合したMHCタンパク質との間の直接の関連が、決定的に既知である。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0180

【補正方法】変更

【補正の内容】

【0180】

提示されたペプチド配列はまた、複数のMHCアレルを発現する細胞から収集されてもよい。典型的にヒトにおいては、6種類の異なるタイプのMHC I分子及び最大で12種類の異なるタイプのMHC II分子が細胞で発現している。そのような提示されたペプチド配列は、複数のあらかじめ決定されたMHCアレルを発現するように操作されている複数アレル細胞株から特定されてもよい。そのような提示されたペプチド配列はまた、正常組織試料または腫瘍組織試料のいずれかの、組織試料から特定されてもよい。この例において特に、MHC分子は、正常組織または腫瘍組織から免疫沈降させることができる。複数のMHCアレル上に提示されたペプチドは、同様に、酸溶出などの技法によって単離され、質量分析により特定されることができる。図2Cは、6種類の例示的なペプチド

YEMFNDSKF (SEQ ID NO: 3), HROEIFSHDFJ (SEQ ID NO: 4),

FJIEJFOESS (SEQ ID NO: 5), NEIOREIREI (SEQ ID NO: 6),

JFKSIFEMMSJDSSUIFLKSJFIEIFJ (SEQ ID NO: 7), 及び

KNFLENFIESOFI (SEQ ID NO: 8)

が、特定されたクラスI MHCアレルHLA-A*01:01、HLA-A*02:01、HLA-B*07:02、HLA-B*08:01、及びクラスII MHCアレルHLA-DRB1*10:01、HLA-DRB1:11:01上に提示されており、単離され、質量分析により特定される、この例を示す。単一アレル細胞株とは対照的に、結合したペプチドが、特定される前のMHC分子から単離されるため、提示されたペプチドとそれが結合したMHCタンパク質との間の直接の関連は、未知である可能性がある。

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0182

【補正方法】変更

【補正の内容】

【0182】

アレル相互作用情報はまた、所定のMHCアレルと所定のペプチドとの間の結合親和性の測定値または予測値も含むことができる。1つ以上の親和性モデルが、そのような予測値を生成することができる(72, 73, 74)。例えば、図1Dに示した例に戻ると、提示情報165は、ペプチドYEMFNDSKF (SEQ ID NO: 3)とクラスIアレルHLA-A*01:01との間の1000nMの結合親和性予測値を含み得る。I

C50 > 1000 nmであるペプチドはわずかしが、MHCによって提示されず、より低いIC50値が、提示の確率を増大させる。提示情報165は、ペプチドKNFLFNFI E S O F I (SEQ ID NO: 8)とクラスIIアレルHLA - DRB1 : 11 : 01との間の結合親和性予測値を含み得る。

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0194

【補正方法】変更

【補正の内容】

【0194】

VII. B. 2. アレル非相互作用情報

アレル非相互作用情報は、そのソースタンパク質配列内の、新生抗原コード化ペプチドに隣接するC末端配列を含むことができる。MHC - Iでは、C末端フランキンゲン配列は、ペプチドのプロテアソームプロセッシングに影響を及ぼし得る。しかし、C末端フランキンゲン配列は、ペプチドが小胞体に輸送され、細胞の表面上のMHCアレルと遭遇する前に、プロテアソームによってペプチドから切断される。その結果、MHC分子は、C末端フランキンゲン配列についてのいかなる情報も受け取らず、したがって、C末端フランキンゲン配列の効果は、MHCアレルタイプに応じて変動することができない。例えば、図2Cに示した例に戻ると、提示情報165は、ペプチドのソースタンパク質から特定された、提示されたペプチドFJIEJFOESS (SEQ ID NO: 5)のC末端フランキンゲン配列FOEIFNDKSLDKFJI (SEQ ID NO: 9)を含み得る。

【手続補正9】

【補正対象書類名】明細書

【補正対象項目名】0233

【補正方法】変更

【補正の内容】

【0233】

図4は、1つの実施形態による、訓練データ170Aの例示的なセットを説明する。具体的には、訓練データ170A中の最初の3つのデータ例は、アレルHLA - C * 01 : 03を含む単一アレル細胞株、ならびに3種類のペプチド配列QCEIOWAREFLKEIGJ (SEQ ID NO: 10), FIEUHFWI (SEQ ID NO: 11), 及びFEWRHRJTRUJR (SEQ ID NO: 12)

からのペプチド提示情報を示す。訓練データ170A中の4番目のデータ例は、アレルHLA - B * 07 : 02、HLA - C * 01 : 03、HLA - A * 01 : 01を含む複数アレル細胞株、及びペプチド配列QIEJQIEIJE (SEQ ID NO: 13)からのペプチド情報を示す。最初のデータ例は、ペプチド配列QCEIOWARE (SEQ ID NO: 14)が、アレルHLA - DRB3 : 01 : 01によって提示されなかったことを示す。前の2つの段落において議論したように、ネガティブ標識されたペプチド配列は、データ管理モジュール312によってランダムに生成されてもよく、または提示されるペプチドのソースタンパク質から特定されてもよい。訓練データ170Aはまた、ペプチド配列 - アレルペアについて、1000 nMの結合親和性予測値及び1時間の半減期の安定性予測値も含む。訓練データ170Aはまた、ペプチドFJELFISBOSJFIE (SEQ ID NO: 15)のC末端フランキンゲン配列、及び10² TPMのmRNA定量測定値などの、アレル非相互作用変数も含む。4番目のデータ例は、ペプチド配列QIEJQIEIJE (SEQ ID NO: 13)が、アレルHLA - B * 07 : 02、HLA - C * 01 : 03、またはHLA - A * 01 : 01のうちの1つによって提示されたことを示す。訓練データ170Aはまた、アレルの各々についての結合親和性予測値及び安定性予測値、ならびに、ペプチドのC末端フランキンゲン配列及びペプチドについてのmRNA定量測定値も含む。

【手続補正10】

【補正対象書類名】明細書

【補正対象項目名】0343

【補正方法】変更

【補正の内容】

【0343】

IX. 実施例5：予測モジュール

予測モジュール320は、配列データを受け取って、提示モデルを用いて配列データ中の候補新生抗原を選択する。具体的には、配列データは、患者の腫瘍組織細胞から抽出されたDNA配列、RNA配列、及び/またはタンパク質配列であってよい。予測モジュール320は、配列データを、MHC-Iについては8~15個のアミノ酸を有する、またはMHC-IIについては6~30個のアミノ酸を有する複数のペプチド配列 p^k に処理する。例えば、予測モジュール320は、所定の配列

IEFROEIFJEF (SEQ ID NO: 16)

を、9個のアミノ酸を有する3種類のペプチド配列

“IEFROEIFJ (SEQ ID NO: 17),” “EFROEIFJE (SEQ ID NO: 18),” 及び

“FROEIFJEF (SEQ ID NO: 19)”

に処理することができる。一実施形態では、予測モジュール320は、患者の正常組織細胞から抽出された配列データをその患者の腫瘍組織細胞から抽出された配列データと比較して1つ以上の変異を有する部分を特定することによって、変異したペプチド配列である候補新生抗原を特定することができる。

【手続補正11】

【補正対象書類名】明細書

【補正対象項目名】0428

【補正方法】変更

【補正の内容】

【0428】

これらのペプチドのうち3つでは、これはモデルとMHCflurry1.2.0との間のHLAカバレッジの差による。ペプチドYEHEDVKEA (SEQ ID NO: 20)は、MHCflurry1.2.0によってカバーされていないHLA-B*49:01によって提示されると予測される。同様に、ペプチドSSAAAPFPL (SEQ ID NO: 21)及びFVSTSDIKSM (SEQ ID NO: 22)は、やはりMHCflurry1.2.0によってカバーされていないHLA-C*03:04により提示されると予測される。原理的にはすべてのアレルをカバーするpan特異的結合親和性予測ツールであるオンラインNetMHCpan4.0 (BA) 予測ツールは、SSAAAPFPL (SEQ ID NO: 21)をHLA-C*03:04に対する強い結合物質としてランクし(23.2 nM、患者1-024-002で2位にランク)、FVSTSDIKSM (SEQ ID NO: 22)のHLA-C*03:04に対する弱い結合(943.4 nM、患者1-024-002で39位にランク)及びYEHEDVKEA (SEQ ID NO: 20)のHLA-B*49:01に対する弱い結合(3387.8 nM)を、また、やはりこの患者に存在するがモデルによってカバーされていないHLA-B*41:01に対するより強い結合(208.9 nM、患者1-038-001で11位にランク)を予測している。したがって、これら3つのペプチドのうち、FVSTSDIKSM (SEQ ID NO: 22)は結合親和性予測によれば漏れていたであろうし、SSAAAPFPL (SEQ ID NO: 21)は捕捉されていたであろうし、YEHEDVKEA (SEQ ID NO: 20)のHLA拘束性は不明である。

【手続補正12】

【補正対象書類名】明細書

【補正対象項目名】0429

【補正方法】変更

【補正の内容】

【0429】

ペプチド特異的T細胞応答が逆置み込みされた残りの5つのペプチドは、モデルによって判定された最も可能性の高い提示アレルがやはりMHCflurry1.2.0によってカバーされていた患者に由来するものであった。これら5つのペプチドのうち4つ(4/5)は標準的な500nMの閾値よりも強い予測結合親和性を有し、上位20位にランクされたが、モデルによるランクよりもいくぶん低いランクであった(ペプチド DENITTIQF (SEQ ID NO: 23), QDVSVQVER (SEQ ID NO: 24), EVADAATLTM (SEQ ID NO: 25), DTVEYPYTSF (SEQ ID NO: 26))

は、MHCflurryによって2、14、7、及び9位にランクされたのに対して、モデルによってそれぞれ0、4、5、7位にランクされた)。ペプチド GTKKDVLK (SEQ ID NO: 27) はCD8T細胞により認識され、モデルによって1位にランクされたが、MHCflurryによるランクは70位であり、予測結合親和性は2169nMであった。

【手続補正13】

【補正対象書類名】明細書

【補正対象項目名】0552

【補正方法】変更

【補正の内容】

【0552】

補足の表3

NSCLC患者においてT細胞認識について試験したペプチド

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールの反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCflurry MHCflurry のランク	MHCflurry (nM)	MHCflurryによりカバーされる最も高確率の制限
1-001-002	HSPFATSL	52	N		1-001-002_pool_1	N	snp	HERC2	A2060S	41.9	HLA-C*01:02	0	95	5169.68205	偽
1-001-002	DPEEVIVT	53	N		1-001-002_pool_1	N	snp	CLTC	S989L	272.1	HLA-B*51:01	1	61	3455.25069	真
1-001-002	ELDPDQLQEV	54	N		1-001-002_pool_1	N	snp	KATNAL1	D407Y	12.81	HLA-A*01:01	2	1	24.2177849	真
1-001-002	TPLTKDVTL	55	N		1-001-002_pool_1	N	snp	AP3B1	S817T	44.4	HLA-B*08:01	3	2	48.9740194	真
1-001-002	DGVGKSAL	56	N		1-001-002_pool_1	N	snp	KRAS	G12D	40.75	HLA-B*08:01	4	89	4714.29522	真
1-001-002	YTTVRALTL	57	N		1-001-002_pool_1	N	snp	TNFAIP1	R48L	45.62	HLA-B*08:01	5	26	973.417701	真
1-001-002	TPSAAVKLI	58	N		1-001-002_pool_1	N	snp	STARD5	M108V	1.95	HLA-B*51:01	6	39	2030.48603	真
1-001-002	WPVLLINV	59	N		1-001-002_pool_1	N	del_fs	ZMAT3	V240fs	14.99	HLA-B*51:01	7	16	600.564752	真
1-001-002	ELNARRCSF	60	N		1-001-002_pool_1	N	snp	PQLC1	R109C	33.89	HLA-B*08:01	8	5	62.0439997	真
1-001-002	QMKNPILFL	61	N		1-001-002_pool_1	N	snp	STXBP1	R171L	38.76	HLA-B*08:01	9	20	674.64733	真
1-001-002	LTEKVSLLK	62	N		1-001-002_pool_2	N	snp	BICD2	E489K	42.66	HLA-A*01:01	10	10	428.744925	真
1-001-002	SPETATSL	63	N		1-001-002_pool_2	N	snp	HERC2	A2060S	41.9	HLA-B*08:01	11	4	59.1155419	真
1-001-002	NVDMRTISF	64	N		1-001-002_pool_2	N	snp	STOM	K93N	360.6	HLA-B*08:01	12	30	1490.72261	真
1-001-002	TSIVVSQTL	65	N		1-001-002_pool_2	N	snp	WDR19	A282V	18.12	HLA-B*08:01	13	176	9862.33009	真
1-001-002	HIKIEPVAL	66	N		1-001-002_pool_2	N	snp	KIF5	T163I	25.77	HLA-B*08:01	14	27	1122.27455	真
1-001-002	DSPDGSNGL	67	N		1-001-002_pool_2	N	snp	OSER1	S119N	20.7	HLA-C*01:02	15	471	21598.414	偽
1-001-002	YTAVHYAASY	68	N		1-001-002_pool_2	N	snp	ANKRD52	A559S	18.32	HLA-A*01:01	16	0	11.5906737	真
1-001-002	VGADGVGKSAL	69	N		1-001-002_pool_2	N	snp	KRAS	G12D	40.75	HLA-C*01:02	17	370	17985.3612	偽
1-001-002	MMPPPIPGI	70	N		1-001-002_pool_2	N	snp	ZNF207	Q409L	186.04	HLA-B*51:01	18	136	7609.76602	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点) 変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurryのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
1-001-002	FPVPGMTNQ	71	N		1-001-002_pool_2	N	snp	FER	C759F	67.36	HLA-B*51:01	19	38	1999.07208	真
1-024-001	VTHHAPLSW	72	N		1-024-001_pool_1	Y	snp	OS8PL11	G489W	24.12	HLA-A*32:01	0	7	77.009026	真
1-024-001	GTKKDDVLIK	27	Y		1-024-001_pool_1	Y	snp	RTFDC1	E177K	61.32	HLA-A*03:01	1	70	2168.51668	真
1-024-001	GLNVPVQSNK	34	N		1-024-001_pool_1	Y	snp	HERC6	R218L	8.7	HLA-A*03:01	2	4	59.675168	真
1-024-001	VVVGACGVGK	73	N		1-024-001_pool_1	Y	snp	KRAS	G12C	40.05	HLA-A*03:01	3	11	133.648023	真
1-024-001	AQFAGKQIY	74	N		1-024-001_pool_1	Y	snp	SHC3	E376D	8.88	HLA-A*32:01	4	91	3715.42819	真
1-024-001	KVVLPSDVTSY	75	N		1-024-001_pool_1	Y	snp	COL7A1	R468S	25.42	HLA-A*32:01	6	85	3234.15772	真
1-024-001	MIMKNISTK	76	N		1-024-001_pool_1	Y	snp	PTPN6	E471K	105.39	HLA-A*03:01	7	0	12.2301919	真
1-024-001	DLAGGTFDV	77	N		1-024-001_pool_1	Y	snp	HSPA8	G201A	736.61	HLA-B*27:05	9	353	18290.7955	真
1-024-001	LIFDLAGGTF	78	N		1-024-001_pool_1	Y	snp	HSPA8	G201A	736.61	HLA-C*02:02	11	57	1716.74204	偽
1-024-001	INVLFDLA	79	N		1-024-001_pool_1	Y	snp	HSPA8	G201A	736.61	HLA-A*32:01	17	621	27984.1357	真
1-024-001	VVGACGVGK	80	N		1-024-001_pool_2	N	snp	KRAS	G12C	40.05	HLA-A*03:01	5	19	197.846108	真
1-024-001	VIMLNGTKK	81	N		1-024-001_pool_2	N	snp	RTFDC1	E177K	61.32	HLA-A*03:01	8	10	122.750322	真
1-024-001	LAGGTFDV	82	N		1-024-001_pool_2	N	snp	HSPA8	G201A	736.61	HLA-C*02:02	10	632	28384.8834	偽
1-024-001	LRNSGGEVF	83	N		1-024-001_pool_2	N	del_fs	CEP128	R102fs	11.31	HLA-B*27:05	12	46	1020.95087	真
1-024-001	VVLPSDVTSY	84	N		1-024-001_pool_2	N	snp	COL7A1	R468S	25.42	HLA-A*32:01	13	62	1925.29397	真
1-024-001	IFDLAGGTF	85	N		1-024-001_pool_2	N	snp	HSPA8	G201A	736.61	HLA-C*02:02	14	427	21255.2074	偽
1-024-001	GILLDEAKRLLY	86	N		1-024-001_pool_2	N	snp	ZNF416	Q49K	11.89	HLA-A*03:01	15	24	354.82068	真
1-024-001	SVLLPENVTIK	87	N		1-024-001_pool_2	N	snp	UBASH3B	G307V	12.11	HLA-A*03:01	16	23	228.127132	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバールされる最も高確率の制限	完全MSモデルのランク	MHCFlurry (nm)	MHCFlurryのランク	MHCFlurryによりカバールされる最も高確率の制限
1-024-001	DLAGGTFDVS	88	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-A*32:01	18	487	23357.3292	真
1-024-001	IFDLAGGTFDV	89	N		1-024-001_pool_2	N	chr11_123059991_C_G	snp	HSPA8	G201A	736.61	HLA-C*02:02	19	563	25887.4267	偽
1-024-002	AEWRNGSTSSL	90	N		1-024-002_pool_1	Y	chr3_122703943_C_G	snp	PARP14	P1095A	129.52	HLA-A*68:01	0	8	126.397714	真
1-024-002	YWSEKDVISAK	35	N		1-024-002_pool_1	Y	chr2_43889858_G_A	snp	LRPPRC	T1335I	79.08	HLA-A*68:01	1	9	136.482978	真
1-024-002	EGSLGISHTR	91	N		1-024-002_pool_1	Y	chr18_62157782_C_A	snp	PIGN	W83L	20.74	HLA-A*68:01	2	6	88.2623459	真
1-024-002	IPASVSAPK	92	N		1-024-002_pool_1	Y	chr13_109784018_C_A	snp	IRS2	S679I	63.55	HLA-A*68:01	3	16	224.278982	真
1-024-002	QDVSQVER	24	Y		1-024-002_pool_1	Y	chr9_64411223_T_G	snp	ANKRD20A 4	M646R	8.92	HLA-A*68:01	4	14	193.974327	真
1-024-002	LVVGASGVGK	93	N		1-024-002_pool_1	Y	chr12_25245351_C_T	snp	KRAS	G12S	72.77	HLA-A*68:01	6	41	1238.56407	真
1-024-002	RATVPEL	36	N		1-024-002_pool_1	Y	chr7_131463253_A_T	snp	MKLN1	D521V	84.08	HLA-C*03:04	7	266	16010.7063	偽
1-024-002	SSAAAPFPL	21	Y		1-024-002_pool_1	Y	chr6_13711102_T_A	snp	RANBP9	H135L	43.5	HLA-C*03:04	8	103	4565.97417	偽
1-024-002	GVSKIIIGGNPK	94	N		1-024-002_pool_1	Y	chr4_10116175_C_T	snp	WDR1	D26N	134.53	HLA-A*68:01	9	125	6797.60699	真
1-024-002	EQNFVSTSDIK	33	個々に試験せず		1-024-002_pool_1	Y	chr3_25791346_A_C	snp	OXS5M	K109T	12.82	HLA-A*68:01	17	156	9099.70986	真
1-024-002	RTQDVSQVER	95	N		1-024-002_pool_2	Y	chr9_64411223_T_G	snp	ANKRD20A 4	M646R	8.92	HLA-A*68:01	5	53	1847.42359	真
1-024-002	EAGNNSRVPR	96	N		1-024-002_pool_2	Y	chr2_74046630_G_T	snp	TET3	G238V	56.35	HLA-A*68:01	10	13	161.242762	真
1-024-002	RYVLHVAA	97	N		1-024-002_pool_2	Y	chr3_122703943_C_G	snp	PARP14	P1095A	129.52	HLA-A*68:01	11	176	10453.627	真
1-024-002	VSKIIIGGNPK	98	N		1-024-002_pool_2	Y	chr4_10116175_C_T	snp	WDR1	D26N	134.53	HLA-A*68:01	12	38	954.724495	真
1-024-002	QPSGVPISL	99	N		1-024-002_pool_2	Y	chr12_14478436_GG_T	mnp	ATF7IP	G1021L	123.21	HLA-A*68:01	13	139	7795.97025	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurryのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
1-024-002	DVSVQVER	100	N		1-024-002_pool_2	Y	chr9_64411223_I_G	snp	ANKRD20A 4	M646R	8.92	HLA-A*68:01	14	7	123.489687	真
1-024-002	FVSTSDIKSM	22	Y		1-024-002_pool_2	Y	chr3_25791346_A_C	snp	OXSM	K109T	12.82	HLA-C*03:04	15	128	7025.56581	偽
1-024-002	FPWVSHSL	39	N		1-024-002_pool_2	Y	chr1_116062776_G_C	snp	SI22A15	A396P	8.57	HLA-C*03:04	16	155	9082.40652	偽
1-024-002	APPRLGDSAL	101	N		1-024-002_pool_2	Y	chr6_13711102_I_A	snp	RANBP9	H135L	43.5	HLA-A*68:01	18	196	11590.601	真
1-024-002	ATVPELNEI	102	N		1-024-002_pool_2	Y	chr7_131463253_A_T	snp	MKLN1	D521V	84.08	HLA-A*68:01	19	365	19785.1419	真
1-038-001	QEFAPLGTV	103	N	プールの結果を参照	1-038-001_pool_1	Y	chr2_219501883_G_T	snp	GMPPA	G92V	21.6	HLA-B*49:01	0	31	3481.07375	偽
1-038-001	MNQVLHAY	104	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr14_100354547_C_G	snp	WARS	D148H	757.21	HLA-C*07:01	12	422	27180.1513	偽
1-038-001	HEDVKEAI	105	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	16	300	24830.2411	偽
1-038-001	GPYFVQAV	106	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr1_111242326_C_T	snp	CHI3L2	L379F	122.33	HLA-B*49:01	1	19	1176.97782	偽
1-038-001	YEHEDVKEAI	107	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	2	212	22559.0306	偽
1-038-001	EESVMLITV	108	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr1_155833354_CC_AG	mnp	AGMAT	G105L	1.03	HLA-B*49:01	3	109	17185.8013	偽
1-038-001	IEEDSAEKI	109	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr6_84215849_C_A	snp	CEP162	E82D	15.62	HLA-B*49:01	4	171	20568.515	偽
1-038-001	TEEDVKIKF	110	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr7_95105459_C_A	snp	SAMD9	M213I	68.23	HLA-B*49:01	5	226	22894.2742	偽
1-038-001	NEQSKILKV	111	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chrX_70375298_C_G	snp	KIF4A	L625V	19.51	HLA-B*49:01	6	141	19054.8385	偽
1-038-001	VDNIIQSI	112	個々に試験せず	プールの結果を参照	1-038-001_pool_1	Y	chr20_2654879_G_T	snp	NOP56	M167I	89.39	HLA-B*49:01	7	119	17928.6022	偽

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurry MHCFlurry のランク	MHCFlurryによりカバーされる最も高確率の制限	
1-038-001	YEHEDVKEA	20	Y		1-038-001_pool_2	Y	chr8_96231911_C_G	snp	UQCRB	D41H	174.81	HLA-B*49:01	9	250	23419.567	偽
1-038-001	YWSEVPVSV	113	個々に試験せず		1-038-001_pool_2	Y	chr17_2330604_G_A	snp	TSR1	H561Y	48.21	HLA-C*17:01	10	0	6.07874308	偽
1-038-001	SELTVHQRI	114	個々に試験せず		1-038-001_pool_2	Y	chr19_37564705_G_C	snp	ZNF571	L575V	19.07	HLA-B*49:01	11	159	19886.0407	偽
1-038-001	VGVGKSAL	115	個々に試験せず		1-038-001_pool_2	Y	chr12_25245350_C_A	snp	KRAS	G12V	91.89	HLA-C*17:01	13	388	26432.7668	偽
1-038-001	DMNQVLHAY	116	個々に試験せず		1-038-001_pool_2	Y	chr14_100354547_C_G	snp	WARS	D148H	757.21	HLA-C*07:01	14	64	10286.4383	偽
1-038-001	NEKGKALY	117	個々に試験せず		1-038-001_pool_2	Y	chr17_51294040_G_T	snp	UTP18	M547I	63.21	HLA-C*07:01	15	339	25564.2874	偽
1-038-001	TEYKLVVGVAV	118	個々に試験せず		1-038-001_pool_2	Y	chr12_25245350_C_A	snp	KRAS	G12V	91.89	HLA-B*49:01	17	233	23113.572	偽
1-038-001	QEFAPLGTVG	119	個々に試験せず		1-038-001_pool_2	Y	chr2_219501883_G_T	snp	GMPPA	G92V	21.6	HLA-B*49:01	18	338	25558.5468	偽
1-038-001	QEVNRTLLNV	120	個々に試験せず		1-038-001_pool_2	Y	chr17_4085728_C_A	snp	ZZEF1	G863V	63	HLA-B*49:01	19	124	18359.7482	偽
1-038-001	VEMLGLJSC	121	個々に試験せず		1-038-001_pool_2	Y	chr4_168427109_C_A	snp	DDX60L	A631S	44.71	HLA-B*49:01	8	267	23949.2398	偽
1-050-001	LFDMMVSV	122	N		1-050-001_pool_1	N	chr1_193097666_T_C	snp	GLRX2	N94S	17.92	HLA-A*29:02	0	1	44.54051	真
1-050-001	ISTFRQCAL	123	個々に試験せず		1-050-001_pool_1	N	chr17_80346815_G_T	snp	RNF213	R2827L	330.56	HLA-C*16:01	10	322	22721.4424	偽
1-050-001	YNTDIEFY	124	個々に試験せず		1-050-001_pool_1	N	chr15_26580447_G_T	snp	GABRB3	T185N	2.2	HLA-A*29:02	16	20	447.152559	真
1-050-001	EETPPFSNY	125	N		1-050-001_pool_1	N	chr21_31266125_T_A	snp	TIAM1	Y283F	13.99	HLA-B*44:03	1	26	537.02592	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurry MHCFlurryのランク	MHCFlurryによりカバーされる最も高確率の制限	
1-050-001	QASGNHHVW	126	個々に試験せず		1-050-001_pool_1	N	chr22_30893501_T_C	snp	OS8P2	Y677H	7.86	HLA-B*44:03	19	109	7506.81856	真
1-050-001	EEVPIIAI	127	個々に試験せず		1-050-001_pool_1	N	chr18_5419733_G_A	snp	EP841L3	S495L	51.69	HLA-B*44:03	2	17	390.306194	真
1-050-001	IEHNIIRNAKY	128	個々に試験せず		1-050-001_pool_1	N	chr3_52617347_T_G	snp	PBRM1	D578A	65.68	HLA-B*44:03	3	10	186.953378	真
1-050-001	AERLDVKAI	129	個々に試験せず		1-050-001_pool_1	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-B*44:03	5	34	1075.19965	真
1-050-001	LFQGGKDLQY	130	個々に試験せず		1-050-001_pool_1	N	chr17_80346815_G_T	snp	RNF213	R2827L	330.56	HLA-A*29:02	6	54	2855.46701	真
1-050-001	DTSPVAVAL	131	個々に試験せず		1-050-001_pool_1	N	chr5_73074790_T_C	snp	FCHO2	L543S	43.6	HLA-A*26:01	8	91	5750.39585	真
1-050-001	AEETPPFSNY	132	N		1-050-001_pool_2	N	chr21_31266125_T_A	snp	TIAM1	Y283F	13.99	HLA-B*44:03	9	16	364.187996	真
1-050-001	AAKAALEDF	133	個々に試験せず		1-050-001_pool_2	N	chr3_47661451_C_G	snp	SMARCC1	E721D	39.53	HLA-C*16:01	11	307	22125.437	偽
1-050-001	EVTPIAIR	134	個々に試験せず		1-050-001_pool_2	N	chr18_5419733_G_A	snp	EP841L3	S495L	51.69	HLA-A*26:01	12	125	9269.11767	真
1-050-001	DVKAIGPLV	135	個々に試験せず		1-050-001_pool_2	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-A*26:01	13	90	5692.75283	真
1-050-001	NETPVAVLI	136	個々に試験せず		1-050-001_pool_2	N	chr7_79453094_C_A	snp	MAGI2	G76V	2.29	HLA-B*44:03	14	13	253.431553	真
1-050-001	LFVVFQTVY	137	個々に試験せず		1-050-001_pool_2	N	chr1_159535913_A_T	snp	OR10I5	L32Q	0.9	HLA-A*29:02	15	9	139.510048	真
1-050-001	AEERLDVKAI	138	個々に試験せず		1-050-001_pool_2	N	chr14_103339252_G_T	snp	EIF5	M275I	89.97	HLA-B*44:03	17	38	1465.22509	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点) 試験せず	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点) 変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	IMHCFIurryのランク	IMHCFIurry (nM)	MHCFIurryによりカバーされる最も高確率の制限
1-050-001	ASGNHHVW	139		個々に試験せず		1-050-001_pool_2	chr22_30893501_I_C	snp	OSBP2	Y677H	7.86	HLA-C*16:01	18	173	13216.9384	偽
1-050-001	KLFHDMNVS	140		個々に試験せず		1-050-001_pool_2	chr1_193097666_I_C	snp	GIRX2	N94S	17.92	HLA-A*29:02	4	21	453.621334	真
1-050-001	ETPPFSNVNTL	141		個々に試験せず		1-050-001_pool_2	chr21_31266125_I_A	snp	TIAM1	Y283F	13.99	HLA-A*26:01	7	172	13162.6216	真
CU04	DENITTIQF	23		Y		CU04_pool_1	chr4_22413213_C_A	snp	ADGRA3	C734F	20.67	HLA-B*18:01	0	2	8.27203164	真
CU04	MELKVESF	142		N		CU04_pool_1	chr1_37874128_G_C	snp	INPP5B	Q606E	36.85	HLA-B*18:01	1	5	13.0510076	真
CU04	EHIPESAGF	143		N		CU04_pool_1	chr3_9943508_G_C	snp	CRELD1	Q347H	29.9	HLA-B*38:01	2	103	4218.0095	真
CU04	YHGDPMPCL	144		N		CU04_pool_1	chr12_7066530_C_T	snp	CLS	P295L	157.54	HLA-B*38:01	3	12	76.7416543	真
CU04	DEERIPVL	145		N		CU04_pool_1	chr7_5752914_I_C	snp	RNF216	M45V	49.2	HLA-B*18:01	4	29	387.328968	真
CU04	EVADAATLTM	25		Y		CU04_pool_1	chr1_52288541_A_C	snp	ZFYVE9	K845T	70.08	HLA-A*26:01	5	7	38.7340629	真
CU04	IEVEVEI	146		N		CU04_pool_1	chr7_135598004_C_G	snp	NUP205	L691V	42.37	HLA-B*18:01	6	21	209.301169	真
CU04	DTVEPYTSF	26		Y		CU04_pool_1	chr14_34713369_C_A	snp	CFI2	D66Y	16.65	HLA-A*26:01	7	9	42.7267485	真
CU04	VEIQLT	147		N		CU04_pool_1	chr11_62827178_C_G	snp	STX5	E134Q	83.43	HLA-B*18:01	8	3	11.6727539	真
CU04	LEKAVHAY	148		N		CU04_pool_1	chr7_138762364_G_T	snp	ATP6V04	P163H	47.21	HLA-B*18:01	9	0	3.63590379	真
CU04	EEADFLAY	149		N		CU04_pool_2	chr6_10556704_C_T	snp	GCNT2	P94L	25.19	HLA-B*18:01	10	1	6.48490966	真
CU04	ENITTIQFY	150		N		CU04_pool_2	chr4_22413213_C_A	snp	ADGRA3	C734F	20.67	HLA-A*26:01	11	16	135.44155	真
CU04	FHATNPLNL	151		N		CU04_pool_2	chr14_75117203_C_G	snp	NEK9	D252H	20.29	HLA-B*38:01	12	8	39.1165673	真
CU04	VFKDLSVTL	152		N		CU04_pool_2	chrX_40597563_G_A	snp	ATP6AP2	E145K	88.26	HLA-B*38:01	13	45	1080.8332	真
CU04	QAVAAVKL	153		N		CU04_pool_2	chr17_42104792_I_A	snp	DHX58	M513L	35.87	HLA-C*12:03	14	136	6872.44	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurryのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
CU04	IQQIQNCI	154	N		CU04_pool_2	N	chr2_67404159_G_C	snp	ETAA1	E493Q	38.47	HLA-B*38:01	15	59	1665.0162	真
CU04	VAKGFISM	155	N		CU04_pool_2	N	chr2_85395579_C_T	snp	CAPG	E314K	151.69	HLA-C*12:03	16	107	5236.61406	真
CU04	QTKPASILY	156	N		CU04_pool_2	N	chr2_32487684_AG_A	del_fs	BIRC6	G2619fs	111.74	HLA-A*26:01	17	47	1143.73481	真
CU04	DHFETIKY	157	N		CU04_pool_2	N	chr1_220024376_C_G	snp	EPRS	M277I	76.64	HLA-B*18:01	18	6	29.8996386	真
CU04	VEPYTFS	158	N		CU04_pool_2	N	chr14_34713369_C_A	snp	CFL2	D66Y	16.65	HLA-B*18:01	19	4	12.3783994	真
CU05	SVSDISEYR	159	N		CU05_pool_1	N	chr12_15670870_G_C	snp	EPS8	Q64E	52.56	HLA-A*68:02	0	1	6.0399624	真
CU05	YTFEIQGVNGV	160	N		CU05_pool_1	N	chr1_22865138_C_G	snp	EPHB2	A410G	74.99	HLA-A*68:02	1	22	132.877429	真
CU05	IYTSSGQLQLF	161	N		CU05_pool_1	N	chr10_73293336_I_C	snp	CFAP70	E636G	30.45	HLA-A*24:02	2	17	46.3526841	真
CU05	FATPSLHTSV	162	N		CU05_pool_1	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	4	16	43.8761927	真
CU05	AVSKPGLDYEL	163	N		CU05_pool_1	N	chr14_77026556_I_A	snp	IRF2BPL	M413L	58.51	HLA-A*68:02	5	274	13566.6012	真
CU05	KYINKTIRV	164	N		CU05_pool_1	N	chr19_2328426_C_T	snp	LSM7	D20N	76.01	HLA-A*24:02	8	32	318.671051	真
CU05	ETTEEMIKYV	165	N		CU05_pool_1	N	chr6_80040624_G_A	snp	TTK	G804E	17.14	HLA-A*68:02	9	37	398.324158	真
CU05	VVSHPLVYW	166	N		CU05_pool_1	N	chr4_106232956_C_G	snp	TBCK	D478H	71.17	HLA-A*68:02	11	235	10875.8686	真
CU05	DIFQVVKAI	167	N		CU05_pool_1	N	chr1_198754369_C_A	snp	PTPRC	L1204I	104.6	HLA-A*68:02	13	36	394.198029	真
CU05	FAFDAVSKPGI	168	N		CU05_pool_1	N	chr14_77026556_I_A	snp	IRF2BPL	M413L	58.51	HLA-A*68:02	18	65	1067.11951	真
CU05	SVSDISEYR	169	N		CU05_pool_2	N	chr12_15670870_G_C	snp	EPS8	Q64E	52.56	HLA-A*68:02	3	94	2050.45825	真
CU05	YTFEIQGV	170	N		CU05_pool_2	N	chr1_22865138_C_G	snp	EPHB2	A410G	74.99	HLA-A*68:02	6	11	26.6362167	真
CU05	ATPSLHTSV	171	N		CU05_pool_2	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	7	25	177.027506	真
CU05	DFATPSLHTSV	172	N		CU05_pool_2	N	chr17_80345147_A_T	snp	RNF213	D2271V	735.31	HLA-A*68:02	10	185	7619.02631	真
CU05	KYINKTIRVKF	173	N		CU05_pool_2	N	chr19_2328426_C_T	snp	LSM7	D20N	76.01	HLA-A*24:02	12	42	538.209517	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点)	個々のペプチド反応 注記	プールID	プール反応 (任意の時点)	変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモザイクによりカバーされる最も高確率の制限	完全MSモザイクのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限	
CU05	SVKPHLCSL	174	N		CU05_pool_2	N	chr17_35363437_C_T	snp	SIFN11	R124H	91.5	HLA-A*68:02	14	88	1897.58723	真
CU05	DISEYRVEHL	175	N		CU05_pool_2	N	chr12_15670870_G_C	snp	EP58	Q64E	52.56	HLA-A*68:02	15	59	885.161001	真
CU05	WVWVSHPHLV	176	N		CU05_pool_2	N	chr4_106232956_C_G	snp	TBCK	D478H	71.17	HLA-A*68:02	16	15	40.725305	真
CU05	KVFKLGKRV	177	N		CU05_pool_2	N	chrX_24810777_G_A	snp	POLA1	E1017K	19.31	HLA-A*68:02	17	61	954.869111	真
CU05	VSRPGLDYEL	178	N		CU05_pool_2	N	chr14_77026556_T_A	snp	IRF2BP1	M413L	58.51	HLA-A*68:02	19	258	12457.5646	真
CU02	SPSKTSLTL	37	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr12_132750694_G_T	snp	ANKLE2	P266T	43.78	HLA-B*07:02	0	7	20.5140939	真
CU02	ASADGTVKLW	40	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr16_1977246_A_G	snp	TBL3	I545V	26.23	HLA-B*57:01	1	20	77.5504026	真
CU02	LVGPAQLSHW	41	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr8_143930249_G_A	snp	PLEC	P863L	528.48	HLA-B*57:01	4	42	287.473059	真
CU02	QTAANAAGVLIK	31	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr7_77773271_A_G	snp	RSBN1L	T584A	25.89	HLA-A*03:01	5	19	76.1012011	真
CU02	FPSKTSLSL	38	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr12_132750694_G_T	snp	ANKLE2	P266T	43.78	HLA-B*07:02	6	26	131.765585	真
CU02	SSTSNRSSTW	42	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr10_96604023_G_A	snp	PIK3AP1	R733W	9.84	HLA-B*57:01	7	30	162.029882	真
CU02	LVYVPLGAGK	32	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr13_33821175_C_T	snp	RFC3	S44L	9.76	HLA-A*03:01	8	2	8.21211585	真
CU02	HSYSELCVW	43	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr8_119802006_C_G	snp	TAF2	D194H	29.74	HLA-B*57:01	9	3	10.120376	真
CU02	VTLDVILER	44	個々に試験せず	プールの結果を参照	CU02_pool_1	Y	chr9_108979413_T_G	snp	CTNNA1L	E323D	32.44	HLA-B*57:01	10	136	2107.24068	真

患者	ペプチド	SEQ ID NO:	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点) 注記	個々のペプチド反応 (任意の時点) 注記	プールID	プール反応 (任意の時点) 変異	変異タイプ	遺伝子	タンパク質への影響	TPM	完全MSモデルによりカバーされる最も高確率の制限	完全MSモデルのランク	MHCFlurryのランク	MHCFlurry (nM)	MHCFlurryによりカバーされる最も高確率の制限
CU02	HSKPEDTDAW	45	個々に試験せず	個々のペプチド反応結果を参照	CU02_pool_1	chr12_133057238_A_G	snp	ZNF84	TI175A		29.84	HLA-B*57:01	11	23	90.7546185	真
CU03	IAASRSVVM	179	個々に試験せず		CU03_pool_1	chr1_230868472_G_A	snp	C1orf198	A14V		36.47	HLA-C*12:03	0	19	146.699014	真
CU03	AAIAASRSV	180	個々に試験せず		CU03_pool_1	chr1_230868472_G_A	snp	C1orf198	A14V		36.47	HLA-C*12:03	2	42	492.404622	真
CU03	AASRSVVM	181	個々に試験せず		CU03_pool_1	chr1_230868472_G_A	snp	C1orf198	A14V		36.47	HLA-C*12:03	6	116	3437.73836	真
CU03	EMDMHLSDY	182	個々に試験せず		CU03_pool_1	chr5_37180032_T_A	snp	C5orf42	I1908L		14.78	HLA-A*01:01	8	7	35.7275148	真
CU03	VENQKHSL	183	個々に試験せず		CU03_pool_1	chr12_30728769_C_T	snp	CAPRIN2	S554N		6.69	HLA-B*08:01	10	124	3970.47602	真
CU03	QYMDSSLVKI	184	個々に試験せず		CU03_pool_1	chr10_60788061_G_T	snp	CDK1	S107I		26.84	HLA-A*23:01	7	8	50.3301427	真
CU03	SASLHPATV	185	個々に試験せず		CU03_pool_1	chr2_25929006_C_T	snp	KIF3C	R785H		17.29	HLA-C*12:03	9	30	260.370195	真
CU03	VPDQSKQL	186	個々に試験せず		CU03_pool_1	chr6_63685063_T_G	snp	PHF3	N447K		47.53	HLA-B*08:01	13	130	4071.14261	真
CU03	IVFIATSEF	187	個々に試験せず		CU03_pool_1	chr11_65976483_A_T	snp	SART1	N554I		70.53	HLA-B*15:03	5	3	17.4168253	真
CU03	YPAQPPVL	188	個々に試験せず		CU03_pool_1	chr20_44066022_C_A	snp	TOX2	S382Y		11.56	HLA-B*08:01	11	101	2455.95947	真

【補正対象書類名】明細書
 【補正対象項目名】0554
 【補正方法】変更
 【補正の内容】
 【0554】

補足の表5

I V S 対 照 実 験 に お け る T S N A 及 び 感 染 症 エ ピ ト ャ ー プ

ペプチド名	配列	SEQ ID NO:	由来(細胞株, 遺伝子)	予測されたHLA拘束性	予測された結合親和性(mM)	変異位置	変異
新生抗原_A1	APKKSIKL	189	H2009 PPFIA3	B*07:02	125	chr19-49140014	C → T
新生抗原_A2	LLLEVVVHHL	190	H128 FANCA	A*02:01	6	chr16-89808348	C → T
新生抗原_A3	FTDEKVKAY	191	H2122 PDE10A	A*01:01	41	chr6-165543564	G → T
新生抗原_A6	RTAKQNPLTK	192	H2122 GPR183	A*03:01	138	chr13-99295446	G → A
新生抗原_A7	FLAPTGPV	193	H128 NTM	A*02:01	8	chr11-131911555	T → C
新生抗原_A10	RLADAEKLFQL	194	H128 PLEKHG4	A*02:01	201	chr16-67284435	G → A
新生抗原_A11	RTAKQNPLTKK	195	H2122 GPR183	A*03:01	131	chr13-99295446	G → A
新生抗原_B2	IMYLTGMVNK	196	H2009 GSPT1	A*03:01	33	chr16-11891120	G → A
新生抗原_B3	TLQELSHAL	197	H128 PRPF19	A*02:01	106	chr11-60902829	G → T
新生抗原_B6	VSQPVAPSY	198	Colo829 KIAA0319L	A*01:01	948	chr1-35479047	C → T
新生抗原_B7	RLFTPIAGY	199	H2126 CYP26B1	A*03:01	157	chr2-72133060	G → C
新生抗原_B8	ITEEPILMTY	200	H2122 RP111	A*01:01	308	chr8-10611205	C → A
新生抗原_B10	KVTGHRWLK	201	H2009 BSG	A*03:01	51	chr19-579577	G → A
新生抗原_B12	KLSEQILKK	202	H2009 TLR5	A*03:01	39	chr1-223110532	C → G
新生抗原_C3	GTKPNPHVY	203	H2126 OAS3	A*03:01	7336	chr12-112961105	G → T
新生抗原_C4	QQQQVVTNK	204	H2126 LRP1	A*03:01	2361	chr12-57162861	G → T
新生抗原_C5	KVLGKGSFAK	205	H2126 PLK2	A*03:01	40	chr5-58459089	G → A
新生抗原_C6	SVQAPVPPK	206	H2009 ENGASE	A*03:01	279	chr17-79084548	C → G
EBV RAKF	RAKFKQLL	207	EBV BZLF-1	B*08:01	457	Nan	Nan
Flu CTCL	CTELKLSDY	208	Influenza NP	A*01:01	39	Nan	Nan
Flu ELRS	ELRSRYWAI	209	Influenza A	B*08:01	12	Nan	Nan
CMV NLVP	NLVPMTATV	210	CMV pp65	A*02:01	45	Nan	Nan
Flu GILG	GILGFVFTL	211	Influenza MP	A*02:01	20	Nan	Nan
HCV KLVA	KLVALGINAV	212	HCV NS3	A*02:01	49	Nan	Nan
HIV ILKE	ILKEPVHGV	213	HIV pol	A*02:01	144	Nan	Nan
RSV NPKA	NPKASLLSL	214	RSV NP	B*07:02	60	Nan	Nan
*新生抗原配列中の 変異ペプチドを 下線で示す。 **腫瘍細胞株: Colo829, H128, H2009, H2122, H2126							

10247349

【手続補正15】
 【補正対象書類名】明細書
 【補正対象項目名】0555

【補正方法】変更

【補正の内容】

【0555】

補足の表6

T C R 核酸配列

【補正対象書類名】明細書

【補正対象項目名】0556

【補正方法】変更

【補正の内容】

【0556】

補足の表7

頻度及びCDR3a/b配列

クロナタイプID	頻度	割合	cdr3s_aa	cdr3s_nt
クロナタイプ1	386	0.491719745	TRA:CAVTVTGRRLTF (SEQ ID NO:47); TRB:CASNPPDAARGQETQYF (SEQ ID NO:48)	TRA:TGTGCTGTGACCGTACGGGGCAGGAGGAGCAGCTACTTTTT (SEQ ID NO:238); TRB:TGTGCCAGCA/ CCCCCGGACGCTGCAGGGGCAAGAGACCCAGTACTTC (SEQ ID NO:239)
クロナタイプ3	53	0.067515924	TRA:CALNARLMF (SEQ ID NO:229); TRB:CASSYREYNTEAFF (SEQ ID NO:230)	TRA:TGTGCTCTAAATGCCAGACTCATGTTT (SEQ ID NO:240); TRB:TGTGCCAGCAGTTACCGGGAGT/CAACACTGAAGCTTTCTTT (SEQ ID NO:241)
クロナタイプ4	34	0.043312102	TRA:CATGLDLDKLI (SEQ ID NO:231); TRB:CAVVRWGGNQFYF (SEQ ID NO:232); TRB:CASSLAGTGGNYEQFF (SEQ ID NO:233)	TRA:TGTGCTACGGGGCCTAGATTGGACAAGCTCATCTTT (SEQ ID NO:242); TRB:TGTGCTGTCAGATE GGGCGGTAAACCAGTTCTATTTT (SEQ ID NO:243); TRB:TGTGCCAGCAGCTTAGCCGGGACAGGGGGT AATTATGAGCAGTTCTTC (SEQ ID NO:244)
クロナタイプ6	10	0.012738854	TRA:CAVTVTGRRLTF (SEQ ID NO:47); TRB:CASNPPDAARGQETQYF (SEQ ID NO:48); TRB:CASSYREYNTEAFF (SEQ ID NO:230)	TRA:TGTGCTGTGACCGTACGGGGCAGGAGGAGCAGCTACTTTTT (SEQ ID NO:238); TRB:TGTGCCAGCA/ CCCCCGGACGCTGCAGGGGCAAGAGACCCAGTACTTC (SEQ ID NO:239); TRB:TGTGCCAGCAGT TACCCGGAGTACAACACTGAAGCTTTCTTT (SEQ ID NO:241)
クロナタイプ9	7	0.008917197	TRA:CAVVLDSNYQLIW (SEQ ID NO:234); TRB:CSATRGHLSNPQHF (SEQ ID NO:235)	TRA:TGTGCTGTGGTGGATAGCAACTATCAGTTAATCTGG (SEQ ID NO:245); TRB:TGCAGTGCAA (CAGGGGGCACTTGAGCAATCAGCCCCAGCATT (SEQ ID NO:246)
クロナタイプ10	5	0.006369427	TRA:CATASRQGGSEKLVF (SEQ ID NO:236); TRB:CASSRGGGTDYQYF (SEQ ID NO:237)	TRA:TGTGCTACGGGTAGCCGTCAGGGGGATCTGAAAAGCTGGTCTTT (SEQ ID NO:247); TRB:TGTGTC CCAGCAGCCGAGGGGGGCGACAGATACGAGTATTTT (SEQ ID NO:248)

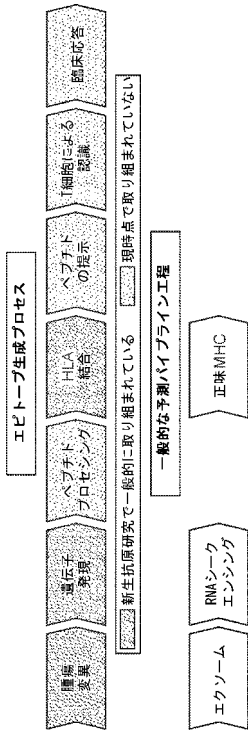
【半 統 補 正 1 7】

【補正対象書類名】明細書
【補正対象項目名】0557
【補正方法】変更
【補正の内容】
【0557】
補足の表8
配列及びV、D、J遺伝子

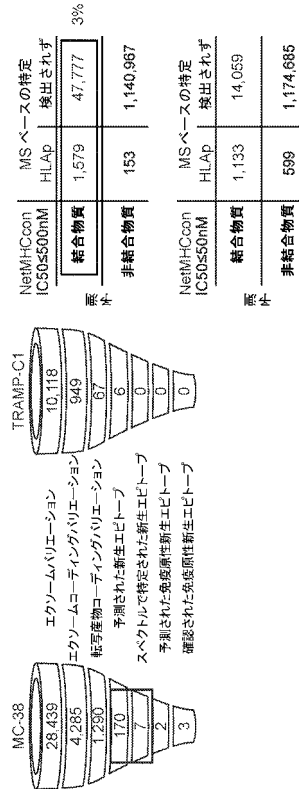
クロノタイプID	コンセンサスID	長さ	種	V 遺伝子	D 遺伝子	J 遺伝子	C 遺伝子	完全株	プロダクティブ	cdi3	SEQ ID NO:	cdi3_nt	SEQ ID NO:	リード	Umis
0	クロノタイプ1_1	1263	TRA	TRAV8-4	None	TRAJ5	TRAC	真	真	CAVTVIGRRALIF	47	TGTGCTGTGACCGTCAACGGGAGGAGGAGCAGTACTTTT	238	2717775	10562
1	クロノタイプ1_2	1332	TRB	TRBV2	TRBD2	TRBJ2-5	TRBC2	真	真	CASNPPDARGQEQYF	48	TGTGCCCAAGACCCCGGACGCTGCGAGGGGACAAGAGACCCAGTACTTC	239	1796338	7242
2	クロノタイプ10_1	604	TRA	TRAV17	None	TRAJ57	TRAC	真	真	CATASRGGSEKLVF	236	TGTGCTACGGCTAGCCGTCAGGCGGATCTGAAAAGGCTGTCTT	247	4006	13
3	クロノタイプ10_2	774	TRB	TRBV7-6	TRBD1	TRBJ2-3	TRBC2	真	真	CASSRGGTDIQYF	237	TGTGCCAGCAGCCGAGGGGGGACAGATAGCGAGTATTT	248	9401	35
103	クロノタイプ3_1	1463	TRB	TRBV6-1	TRBD2	TRBJ1-1	TRBC1	真	真	CASSYREYNTEAFF	230	TGTGCCAGCAGTTACCCGGGAGTACAACACTGGAAGCTTCTTT	241	243324	894
104	クロノタイプ3_2	608	TRA	TRAV6	None	TRAJ31	TRAC	真	真	CALNARLMF	229	TGTGCTCTAAATGCCAGACTCATGTT	240	87200	247
125	クロノタイプ4_1	683	TRA	TRAV17	None	TRAJ34	TRAC	真	真	CATGLDLKLVF	231	TGTGCTACGGGCTAGATTGGCAAGCTCATCTT	242	131803	385
126	クロノタイプ4_2	659	TRA	TRAV41	None	TRAJ49	TRAC	真	真	CAVRWGGNQIPYF	232	TGTGCTGTGAGATGGGGGGTAAACAGTTCATTTT	243	88320	307
127	クロノタイプ4_3	1315	TRB	TRBV7-9	TRBD1	TRBJ2-1	TRBC2	真	真	CASSIAGTGGVEQFF	233	TGTGCCAGCAGCTTAGCCGGGACAGGGGTAAATATGAGCAGTCTTC	244	237200	830
171	クロノタイプ6_1	1007	TRB	TRBV6-1	TRBD2	TRBJ1-1	TRBC1	真	真	CASSYREYNTEAFF	230	TGTGCCAGCAGTTACCCGGGAGTACAACACTGGAAGCTTCTTT	241	23803	105
172	クロノタイプ6_2	713	TRA	TRAV8-4	None	TRAJ5	TRAC	真	真	CAVTVIGRRALIF	47	TGTGCTGTGACCGTCAACGGGAGGAGAGCAGTACTTTT	238	24293	106
173	クロノタイプ6_3	1147	TRB	TRBV2	TRBD2	TRBJ2-5	TRBC2	真	真	CASNPPDARGQEQYF	48	TGTGCCCAAGACCCCGGACGCTGCGAGGGGACAAGAGCAGTACTTC	239	34437	132
242	クロノタイプ9_1	568	TRA	TRAV72	None	TRAJ33	TRAC	真	真	CAVLDLSYQLIW	234	TGTGCTGTGTGTGGATAGCAACTATCAGTTAATCTGG	245	9883	29
243	クロノタイプ9_2	1102	TRB	TRBV2D-1	TRBD1	TRBJ1-5	TRBC1	真	真	CSATRGHLSNQPHF	235	TGCAGTGTCAACAGGGGGGCACTTGAGCAATCAAGCCCGCAGCATTTT	246	38619	126

- 【補正対象書類名】 図面
- 【補正対象項目名】 全図
- 【補正方法】 変更
- 【補正の内容】

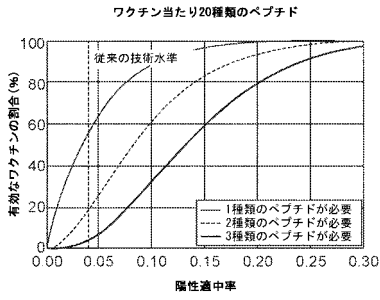
【図 1 A】



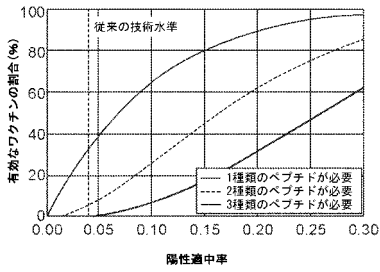
【図 1 B】



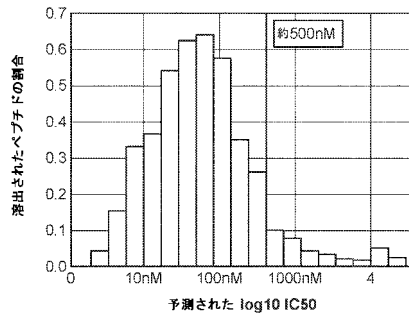
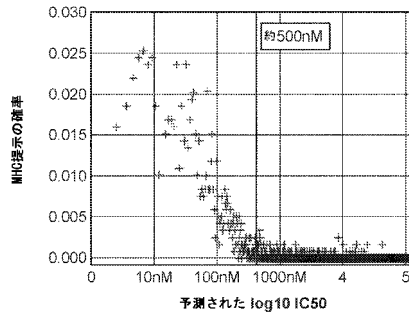
【図 1 C】



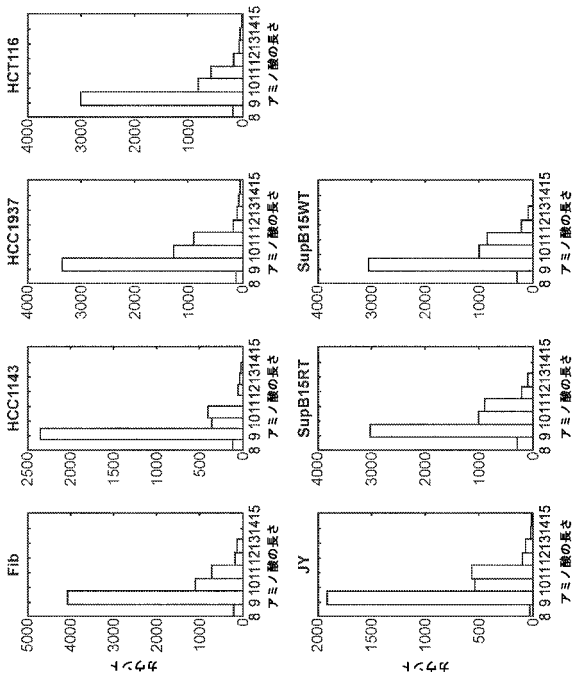
ワクチン当たり10種類のペプチド



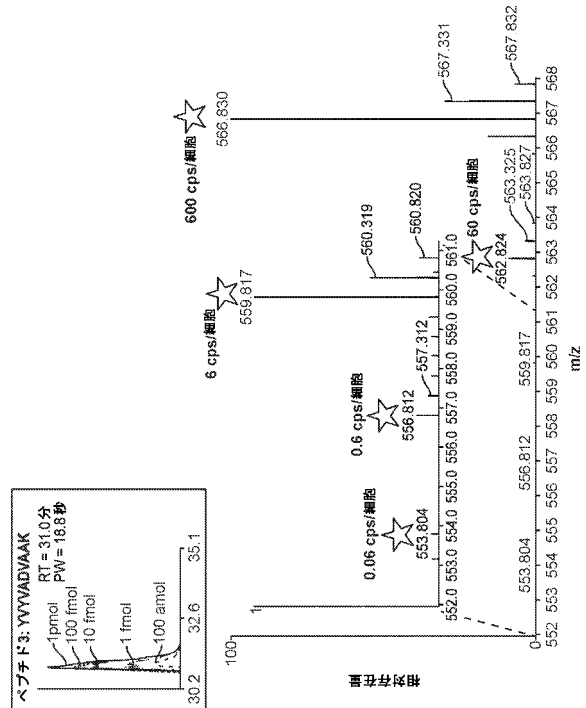
【図 1 D】



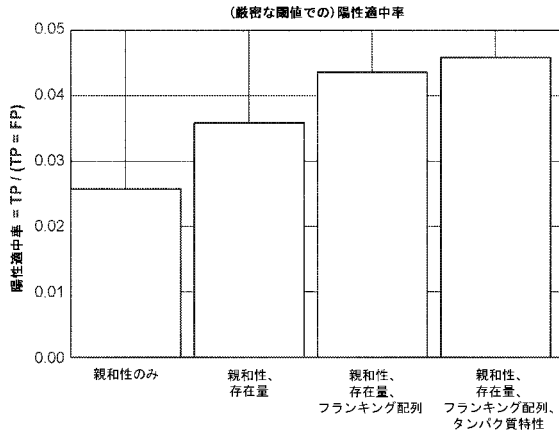
【図 1 E】



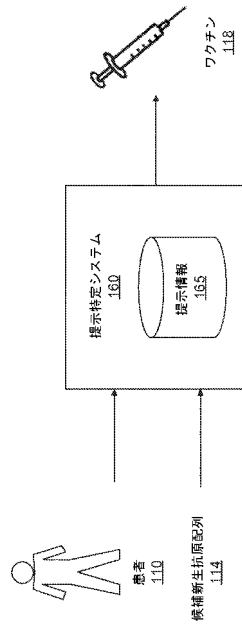
【図 1 F】



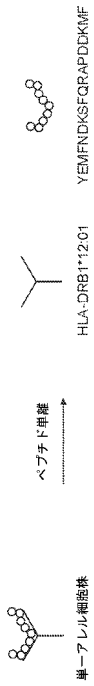
【図 1 G】



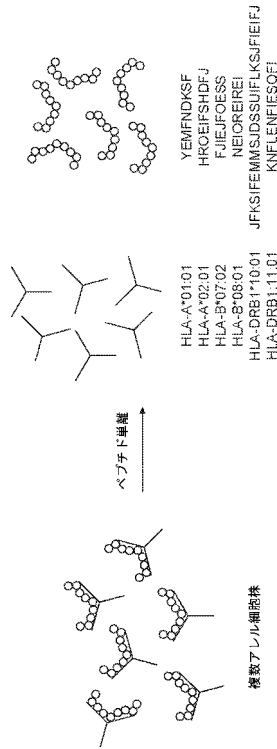
【図 2 A】



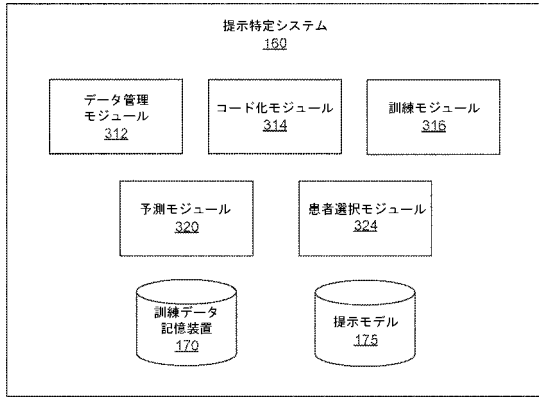
【図 2 B】



【図 2 C】



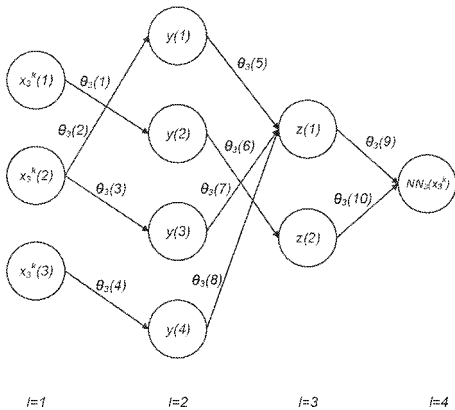
【 図 3 】



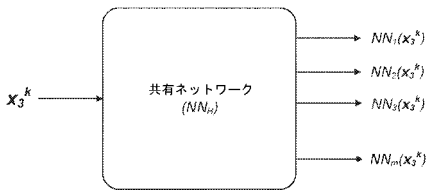
【 図 4 】

アレル依存性 (x)		アレル非依存性 (w)		mRNA Q. (m・TPM)	ラベル (y)
親和性 (p-nm)	安定性 (s-n)	アレル (p)	アレル (q)		
GCEIOWAREFLKEIGJ	1000	HLA-DRB3*01:01	FJELFISBOSJFIE	10 ²	提示されない
FIEUHFVI	1500	HLA-C*01:03	FEGRKUCCI	10 ³	提示される
FEWRHRJTRUUR	650	HLA-C*01:03	PJFIOEJOUJGEIO	10 ¹	提示される
QIEJQIEJIE	500	HLA-B*07:02	PJFIOEJOUJGEIO	1	提示される
	600	HLA-C*01:03			
	1200	HLA-A*01:01			

【 図 5 】



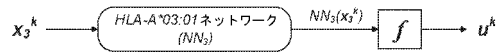
【 図 6 A 】



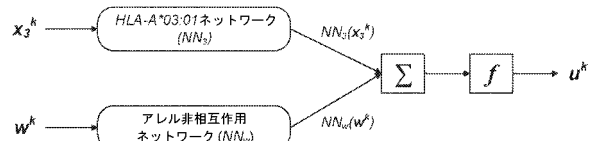
【 図 6 B 】



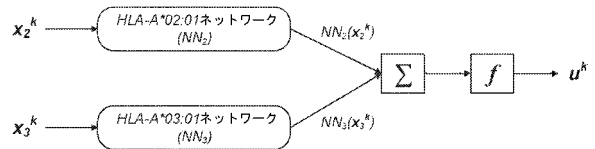
【 図 7 】



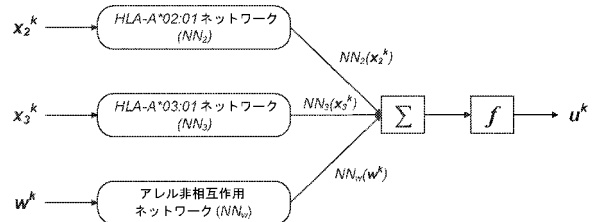
【 図 8 】



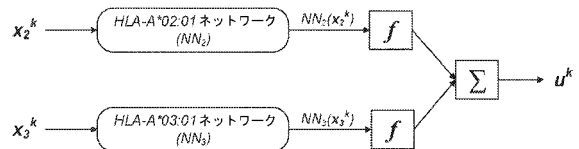
【 図 9 】



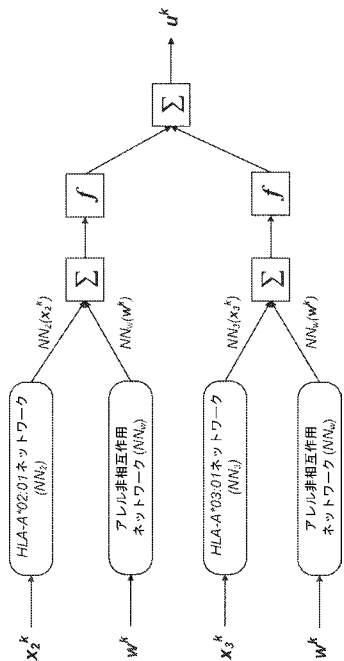
【 図 10 】



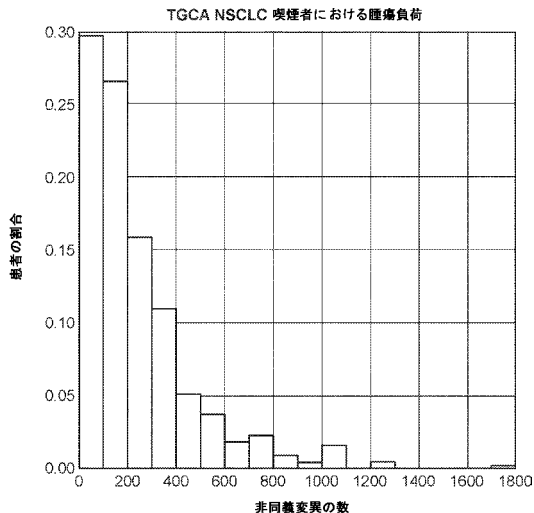
【 図 11 】



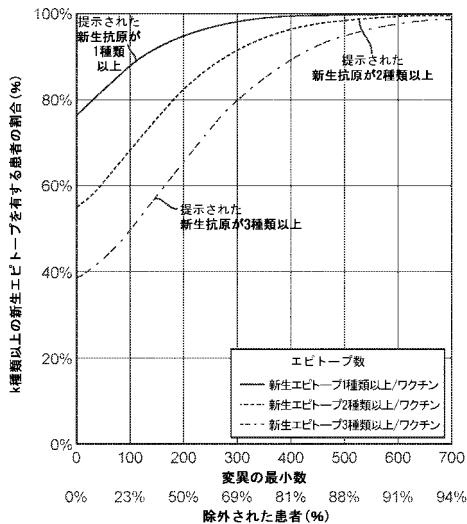
【図 1 2】



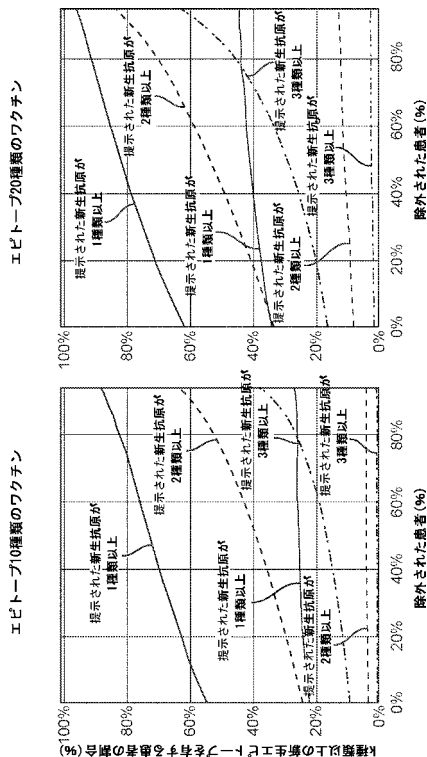
【図 1 3 A】



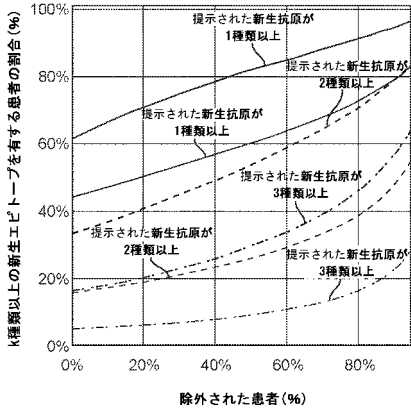
【図 1 3 B】



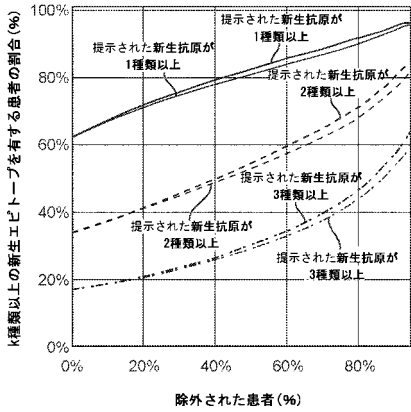
【図 1 3 C】



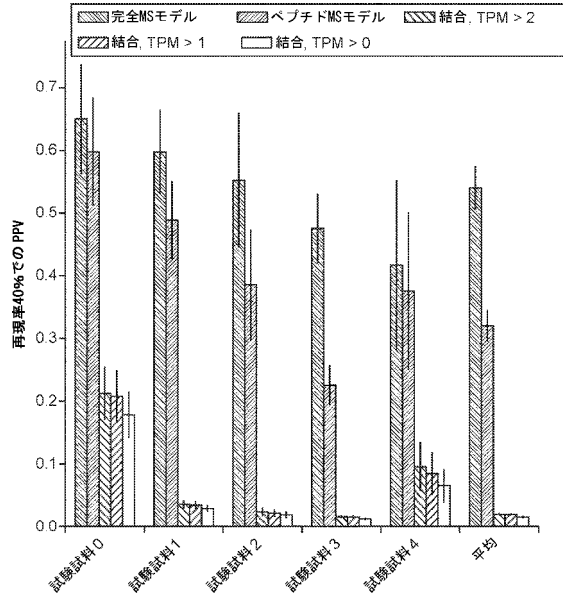
【図 1 3 D】



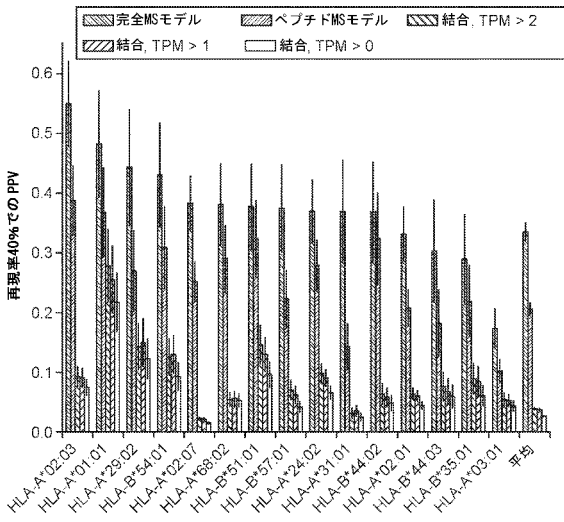
【図 1 3 E】



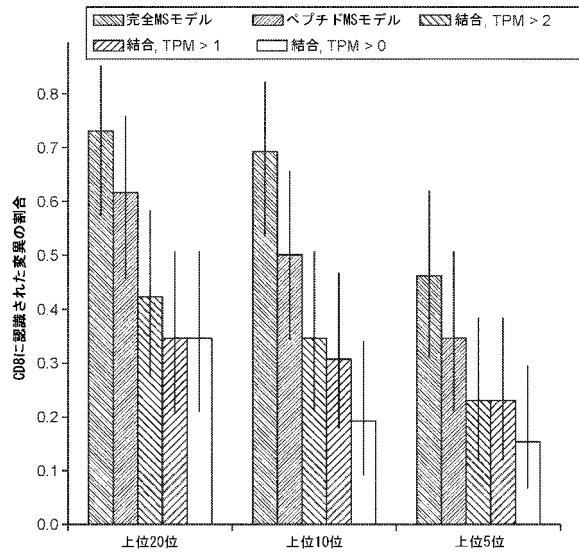
【図 1 4 A】



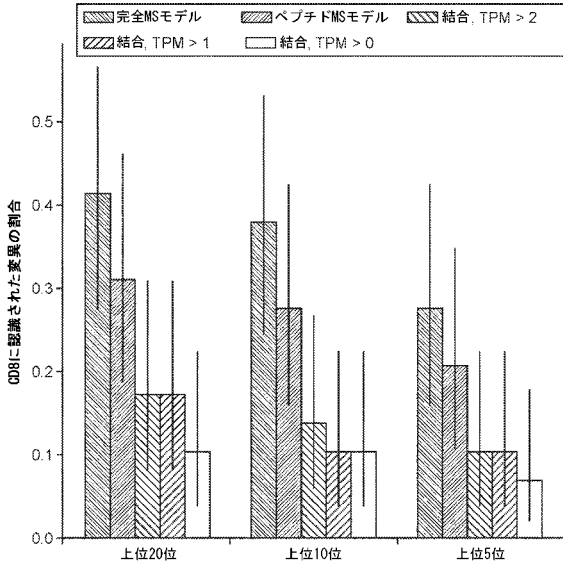
【図 1 4 B】



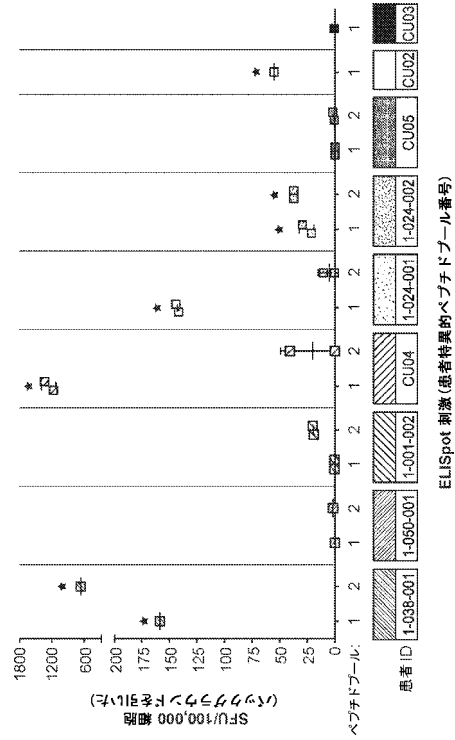
【図 1 4 C】



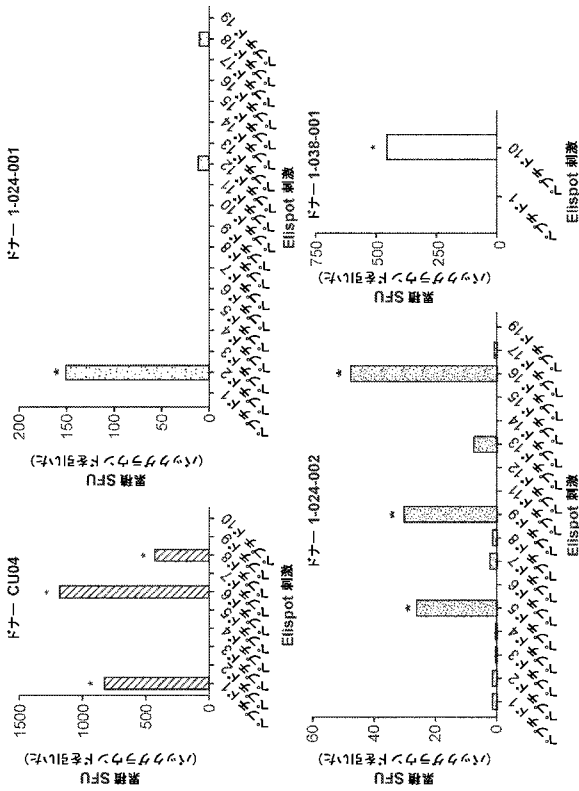
【図 14 D】



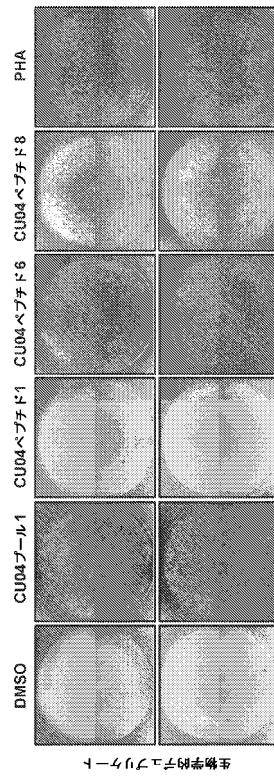
【図 15 A】



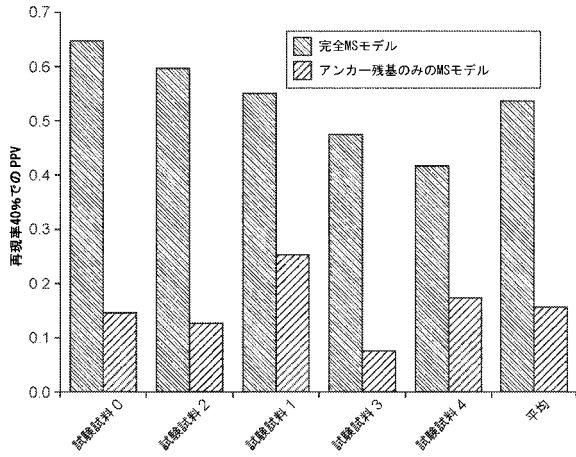
【図 15 B】



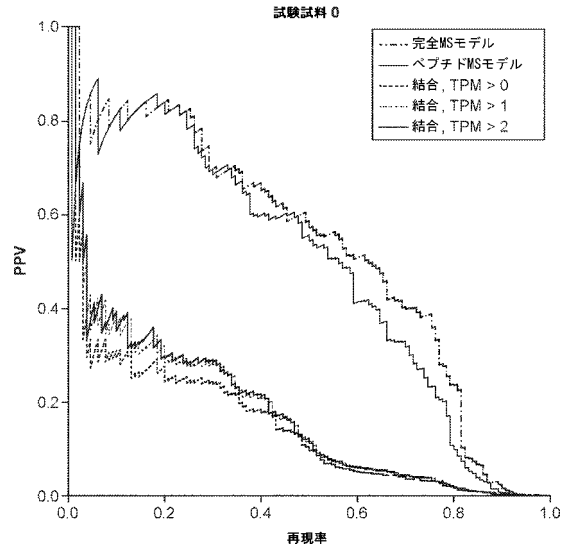
【図 15 C】



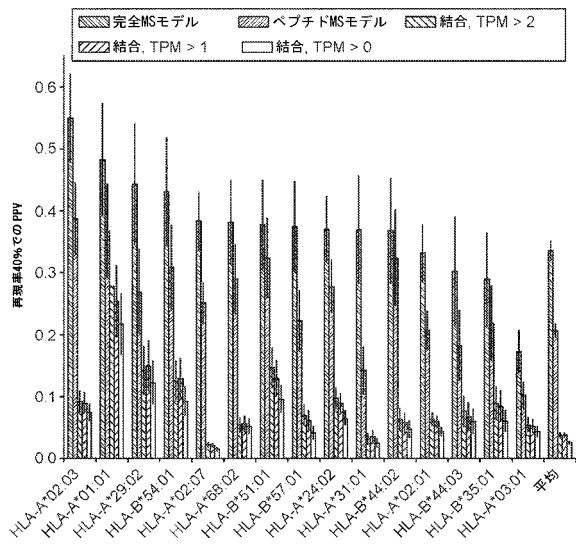
【図16】



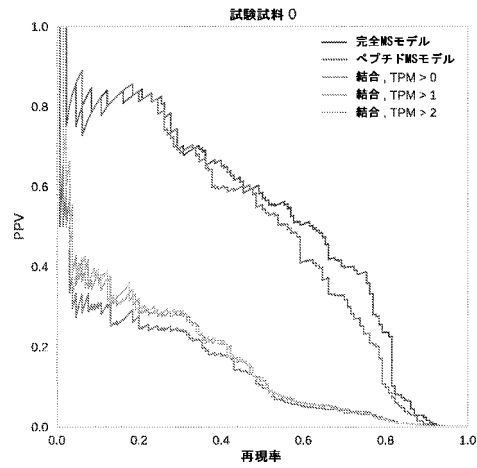
【図17A】



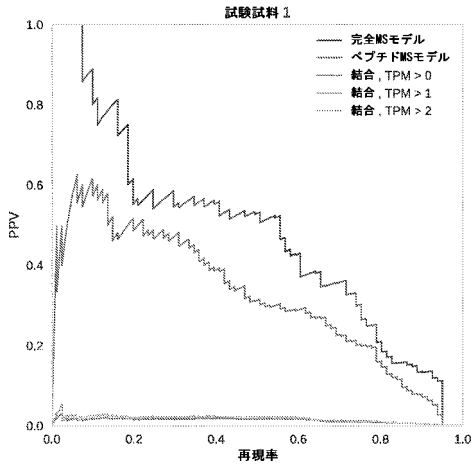
【図17B】



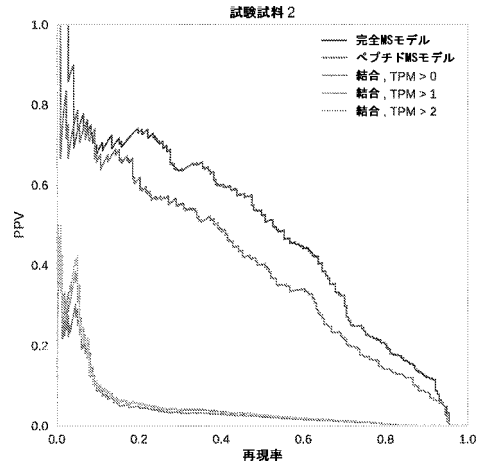
【図17C】



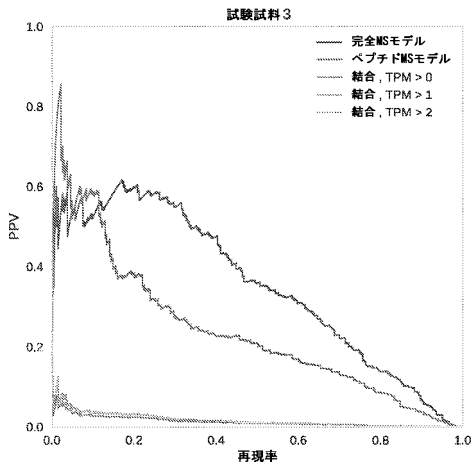
【図 17 D】



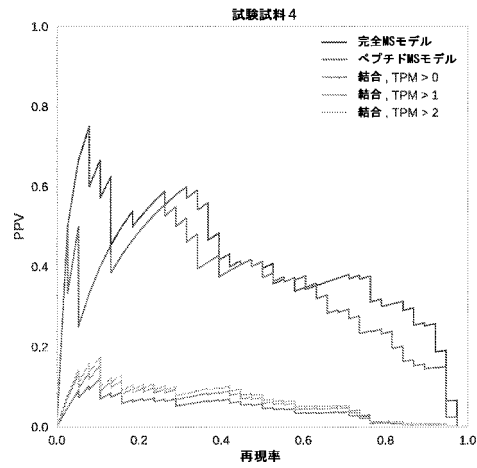
【図 17 E】



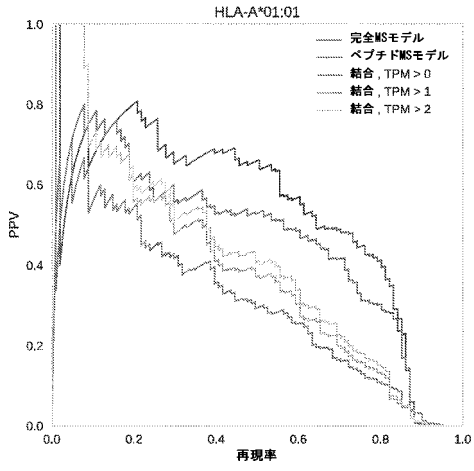
【図 17 F】



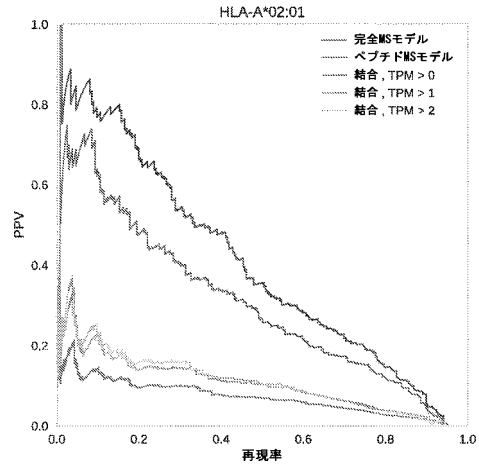
【図 17 G】



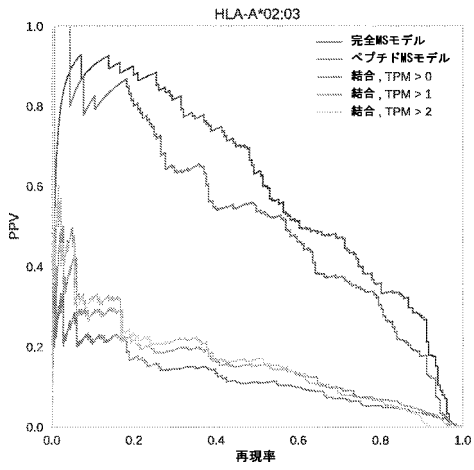
【図 17 H】



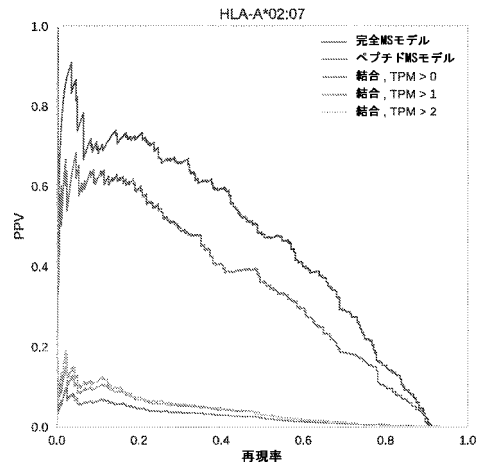
【図 17 I】



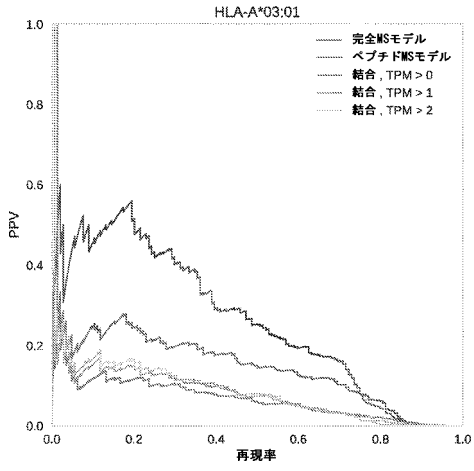
【図 17 J】



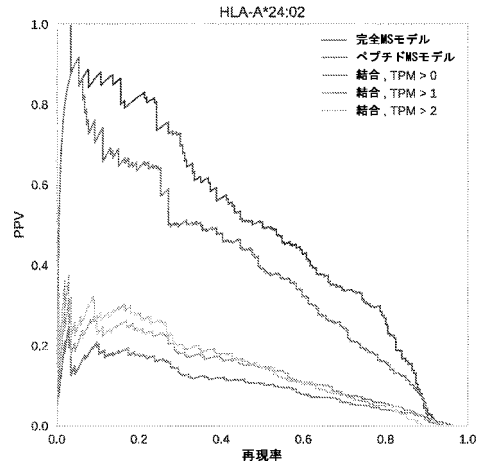
【図 17 K】



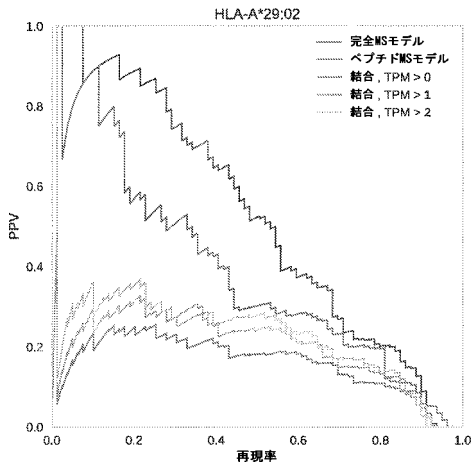
【図 17 L】



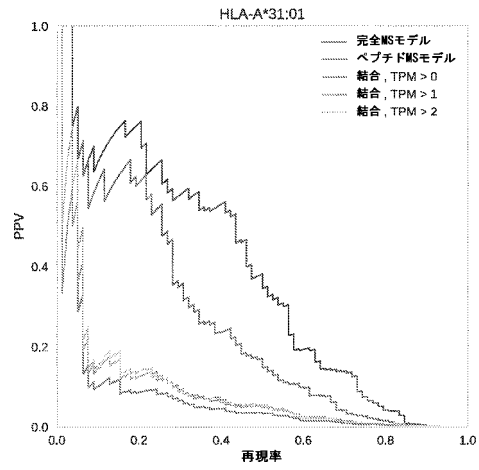
【図 17 M】



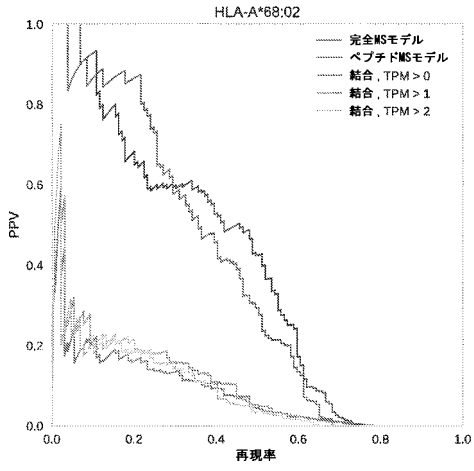
【図 17 N】



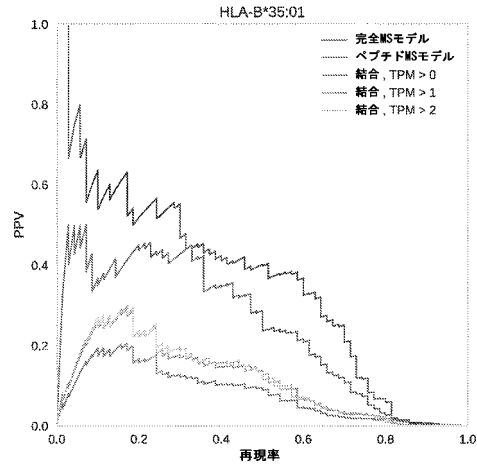
【図 17 O】



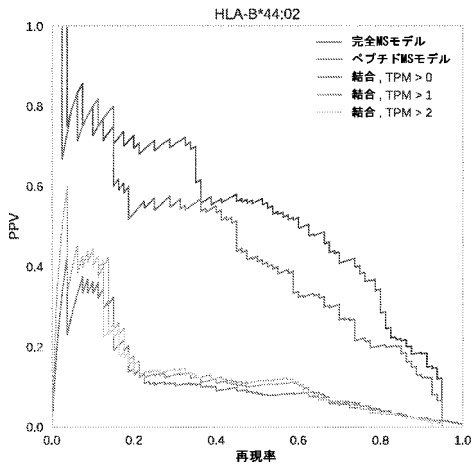
【図 17 P】



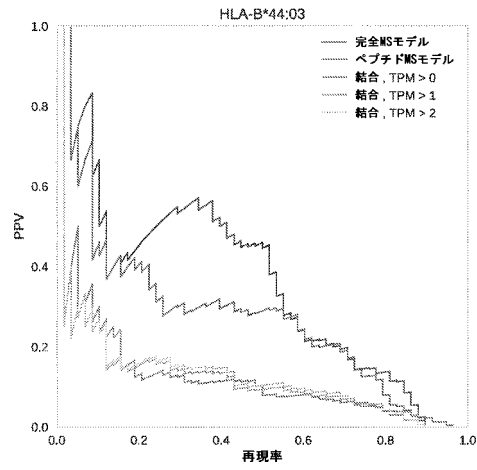
【図 17 Q】



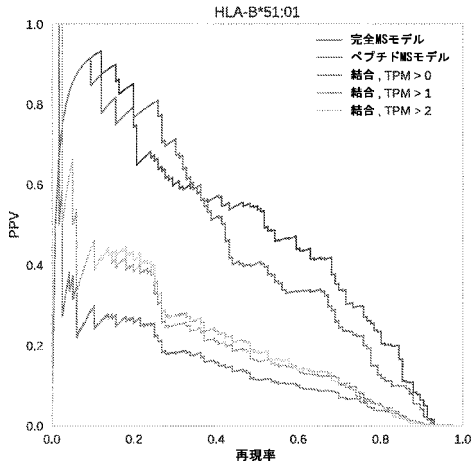
【図 17 R】



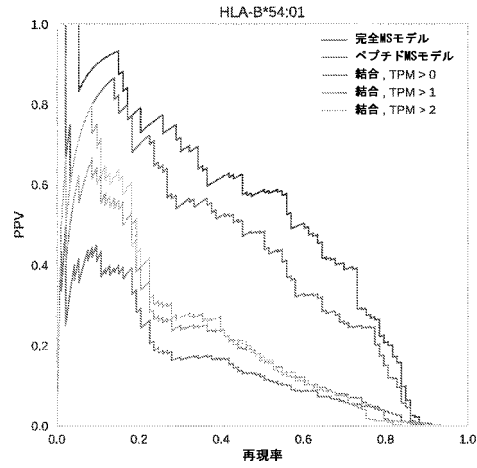
【図 17 S】



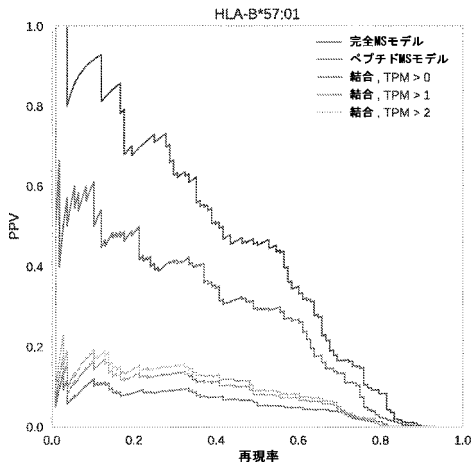
【図17T】



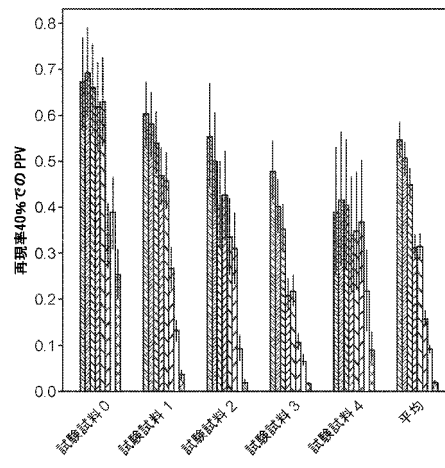
【図17U】



【図17V】

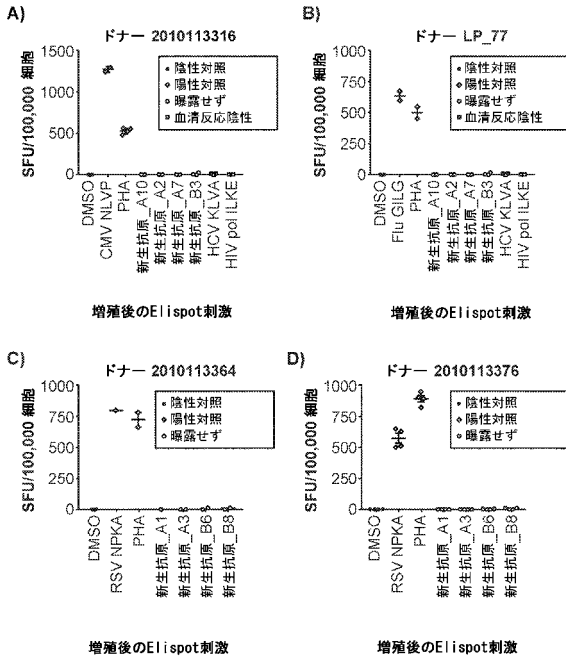


【図18】

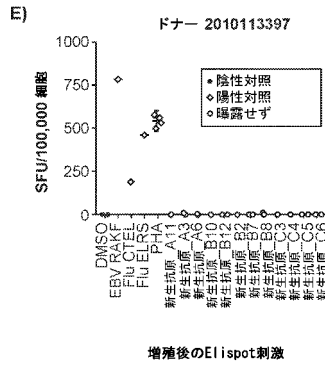


- 完全MSモデル
- MSモデル、フランキング配列なし
- MSモデル、フランキング配列または遺伝子毎係数なし
- すべての長さを一緒に訓練したペプチドのみのMSモデル
- それぞれの長さを別々に訓練したペプチドのみのMSモデル
- 直鎖ペプチドのみのMSモデル
- MixMHCpred 1.1
- 結合親和性

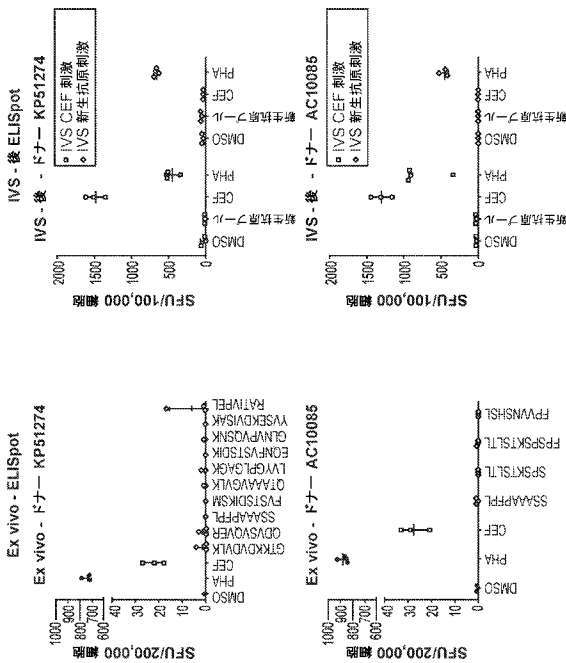
【 図 19 A - 1 】



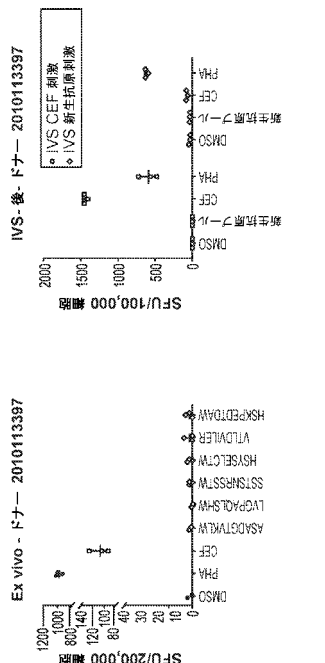
【 図 19 A - 2 】



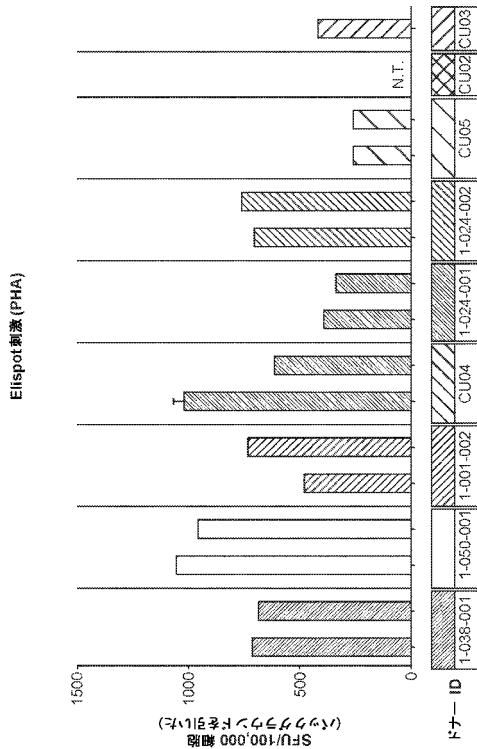
【 図 19 B - 1 】



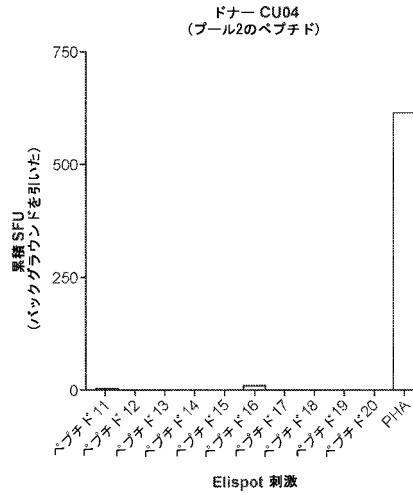
【 図 19 B - 2 】



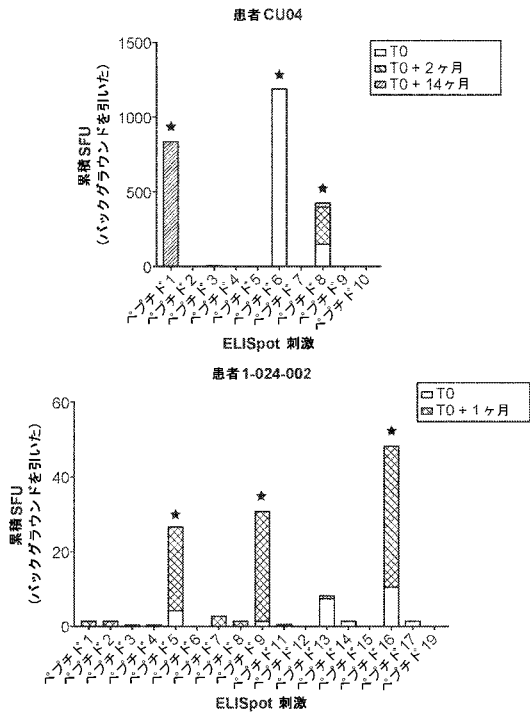
【図 2 0】



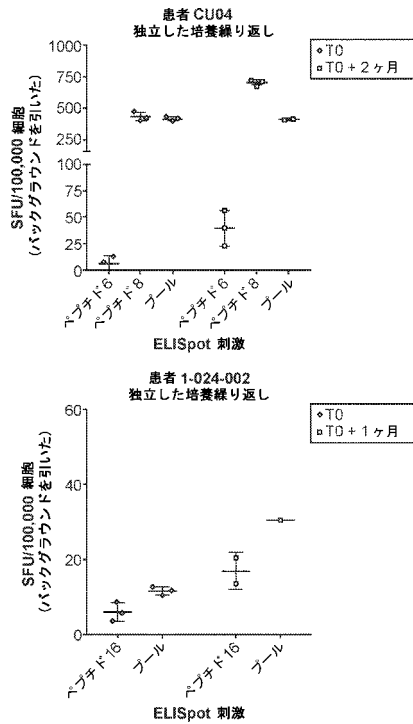
【図 2 1 A】



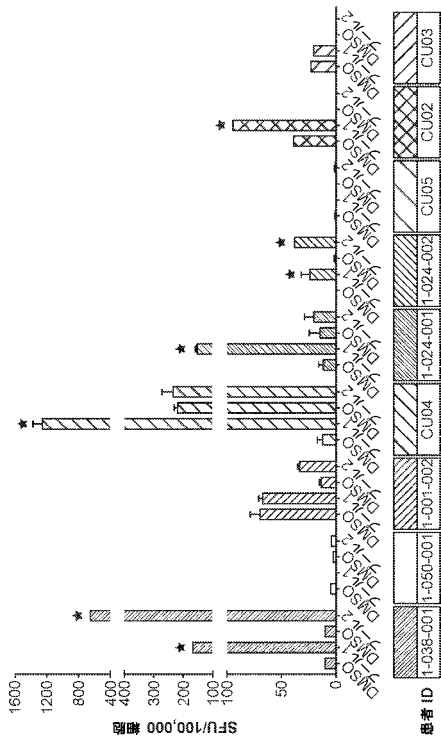
【図 2 1 B】



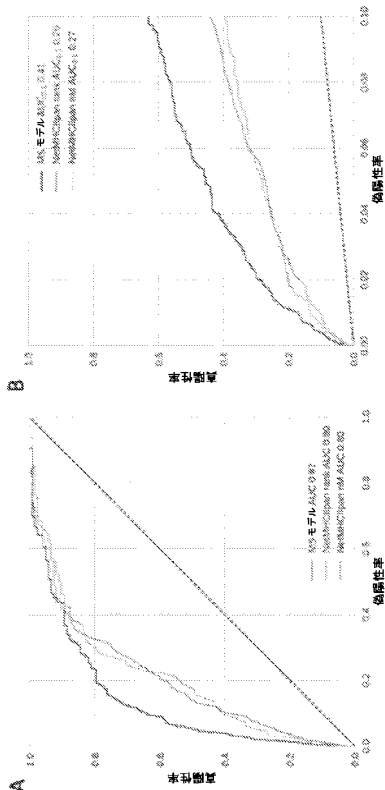
【図 2 1 C】



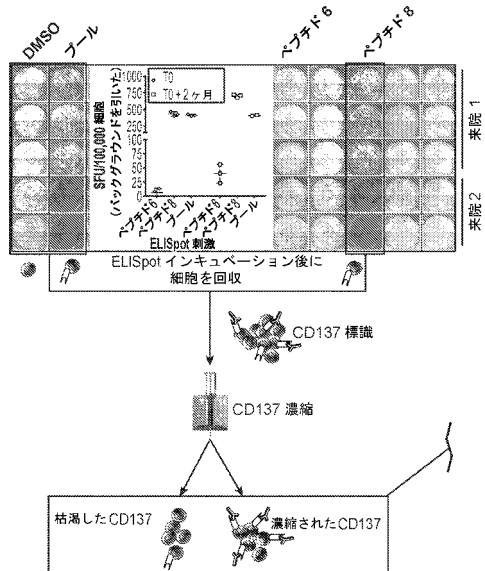
【図22】



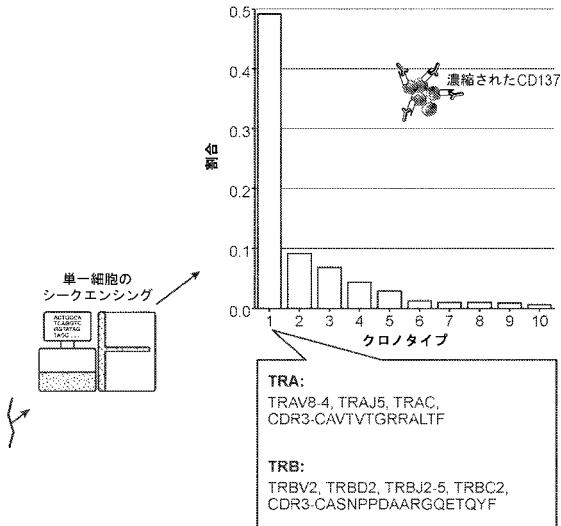
【図23】



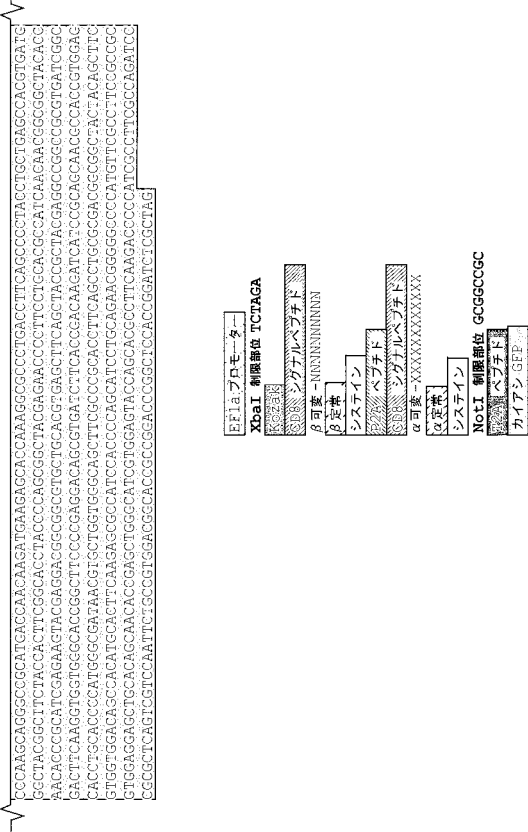
【図24-1】



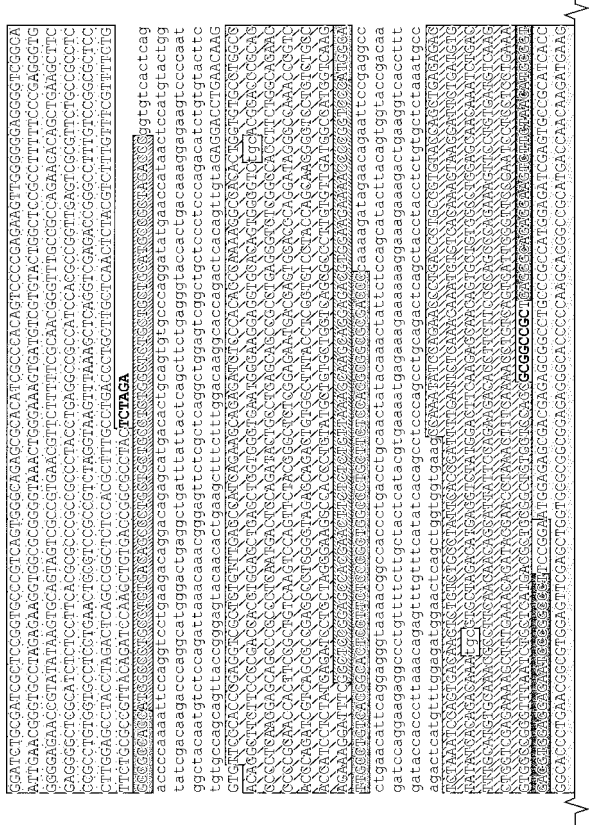
【図24-2】



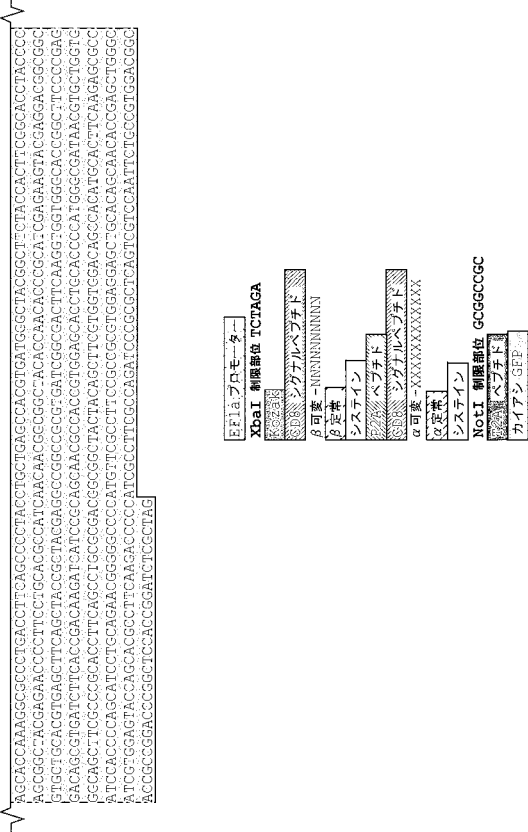
【 図 27 - 2 】



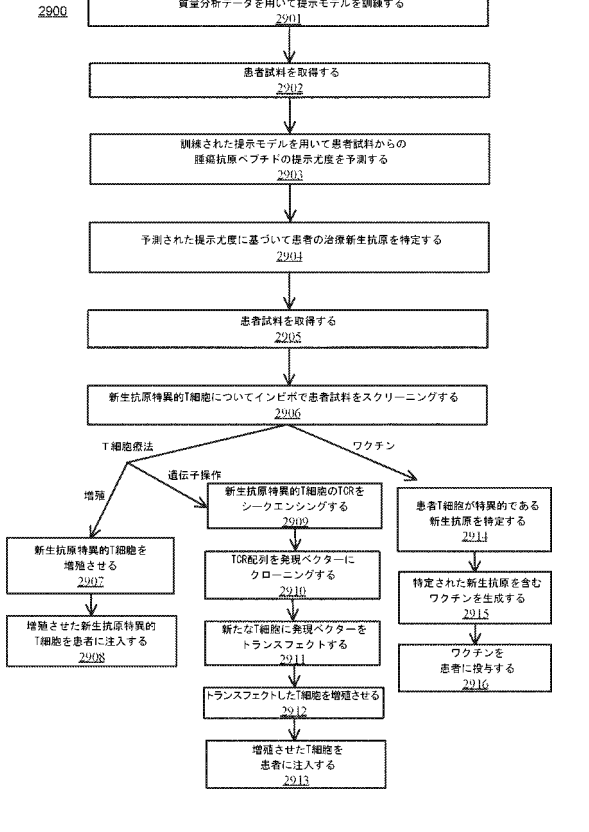
【 図 28 - 1 】



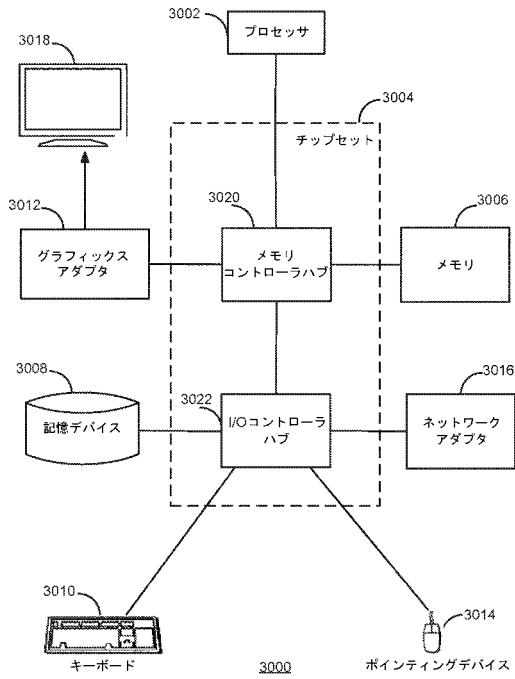
【 図 28 - 2 】



【 図 29 】



【図 30】



【手続補正 19】

【補正対象書類名】明細書

【補正対象項目名】配列表

【補正方法】追加

【補正の内容】

【配列表】

2020532323000001.app

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No. PCT/US2018/049614
--

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - A61K 35/17; A61K 39/00; C12Q 1/68; G01N 33/68; G06F 19/22 (2018.01) CPC - A61K 35/17; A61K 39/0011; C12Q 1/6886; G01N 33/6848; G06F 19/22 (2018.08)		
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) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC - 424/93.71; 424/185.1 (keyword delimited)		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2017/0212984 A1 (GRITSTONE ONCOLOGY, INC.) 27 July 2017 (27.07.2017) entire document	1-8
Y	WO 2016/044530 A1 (THE JOHNS HOPKINS UNIVERSITY) 24 March 2016 (24.03.2016) entire document	1-8
A	US 2016/0125129 A1 (BIONTECH AG et al) 05 May 2016 (05.05.2016) entire document	1-8
A	WO 2016/100975 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY et al) 23 June 2016 (23.06.2016) entire document	1-8
A	WO 2016/146035 A1 (SYZ CELL THERAPY CO.) 22 September 2016 (22.09.2016) entire document	1-8
A	WO 2016/187508 A2 (THE BROAD INSTITUTE INC. et al) 24 November 2016 (24.11.2016) entire document	1-8
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 27 November 2018		Date of mailing of the international search report 04 JAN 2019
Name and mailing address of the ISA/US Mall Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300		Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/049614

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

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

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

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

3. Claims Nos.: 9-39
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

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

- Remark on Protest**
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

フロントページの続き

(51) Int.Cl.	F I		テーマコード(参考)
C 1 2 N 15/10 (2006.01)	C 1 2 N	15/10	1 0 0 Z
C 1 2 N 5/10 (2006.01)	C 1 2 N	5/10	

(31)優先権主張番号 62/554,286

(32)優先日 平成29年9月5日(2017.9.5)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/703,197

(32)優先日 平成30年7月25日(2018.7.25)

(33)優先権主張国・地域又は機関

米国(US)

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

(74)代理人 100142929

弁理士 井上 隆一

(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 ヤレンスカイ ローマン

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル ホートン ストリート 5 9 5
9 スイート 3 0 0 グリットストーン オンコロジー インコーポレイテッド

(72)発明者 ブリク - スリバン ブレンダン

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル ホートン ストリート 5 9 5
9 スイート 3 0 0 グリットストーン オンコロジー インコーポレイテッド

(72)発明者 パスビー ジェニファー

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル ホートン ストリート 5 9 5
9 スイート 3 0 0 グリットストーン オンコロジー インコーポレイテッド

(72)発明者 デイビス マシュー ジョセフ

アメリカ合衆国 9 4 6 0 8 カリフォルニア州 エメリービル ホートン ストリート 5 9 5

- 9 스위트 300 그리ットストーン オンコロジー インコーポレイテッド
(72)発明者 ヤング ローレン エリザベス
アメリカ合衆国 94608 カリフォルニア州 エメリービル ホートン ストリート 595
9 스위트 300 그리ットストーン オンコロジー インコーポレイテッド
(72)発明者 フランシス ジョシュア マイケル
アメリカ合衆国 94608 カリフォルニア州 エメリービル ホートン ストリート 595
9 스위트 300 그리ットストーン オンコロジー インコーポレイテッド
(72)発明者 パーマー クリスティーン
アメリカ合衆国 94608 カリフォルニア州 エメリービル ホートン ストリート 595
9 스위트 300 그리ットストーン オンコロジー インコーポレイテッド
(72)発明者 スコベルン モイカ
アメリカ合衆国 94608 カリフォルニア州 エメリービル ホートン ストリート 595
9 스위트 300 그리ットストーン オンコロジー インコーポレイテッド
F ターム(参考) 4B063 QA01 QA13 QA18 QQ02 QQ42 QR08 QR55 QR62 QS25 QS34
QS36 QS38 QX02
4B065 AA94X AB01 BA02 BA25 CA45