

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-526314
(P2024-526314A)

(43)公表日 令和6年7月17日(2024.7.17)

(51)国際特許分類

F I

G 1 6 B 40/30 (2019.01)

G 1 6 B 40/30

G 0 6 N 99/00 (2019.01)

G 0 6 N 99/00

1 8 0

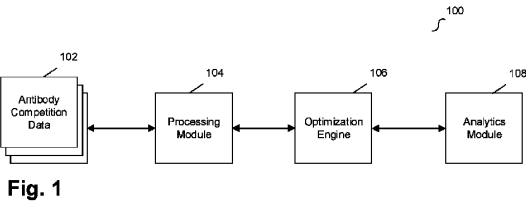
審査請求 未請求 予備審査請求 未請求 (全24頁)

(21)出願番号	特願2024-500487(P2024-500487)	(71)出願人	522359523
(86)(22)出願日	令和4年7月8日(2022.7.8)		アブセラ バイオロジクス インコーポ レイテッド
(85)翻訳文提出日	令和6年2月22日(2024.2.22)		カナダ国 ブイ 5 ワイ 0 エー 1 プリテ ィッシュ コロンビア , パンクーバー , ユーコン ストリート 2 2 1 5
(86)国際出願番号	PCT/US2022/036517	(74)代理人	100078282
(87)国際公開番号	WO2023/009293		弁理士 山本 秀策
(87)国際公開日	令和5年2月2日(2023.2.2)	(74)代理人	100113413
(31)優先権主張番号	63/219,578		弁理士 森下 夏樹
(32)優先日	令和3年7月8日(2021.7.8)	(74)代理人	100181674
(33)優先権主張国・地域又は機関	米国(US)		弁理士 飯田 貴敏
(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,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く	(74)代理人	100181641
			弁理士 石川 大輔
		(74)代理人	230113332
			最終頁に続く

(54)【発明の名称】 隠れた変数の親和性を使用する抗体競合モデル

(57)【要約】

実施形態は、結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出する。例えば、複数の抗体および抗原に関する抗体競合データを受け取ることができ、抗体競合データは、抗体間のペアワイズ競合を示すデータ値を含む。訓練データを生成するために、抗体競合データをプロセッシングすることができる。訓練データおよび最適化エンジンを使用して、複数の隠れた変数および隠れた変数に関する親和性スコアを導出することができ、隠れた変数に関する親和性スコアは、それぞれの抗体について導出され、隠れた変数は、抗体間の競合を引き起こす抗原に関する競合因子を表す。



【特許請求の範囲】**【請求項 1】**

結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するための方法であって、

抗体間のペアワイズ競合を示すデータ値を含む、複数の抗体および抗原に関する抗体競合データを受け取ること；

訓練データを生成するために、前記抗体競合データをプロセッシングすること；ならびに前記訓練データおよび最適化エンジンを使用して、複数の隠れた変数および前記隠れた変数に関する親和性スコアを導出することであって、前記隠れた変数に関する親和性スコアが、それぞれの抗体について導出され、前記隠れた変数が、前記抗体間の競合を引き起こす前記抗原に関する競合因子を表すことを含む、方法。

10

【請求項 2】

第 1 の隠れた変数が、前記抗原に関する第 1 の競合因子を表し、所与の抗体に関連する前記第 1 の隠れた変数に関する導出された親和性スコアが、前記第 1 の競合因子に関する前記所与の抗体の競合の程度を示す、請求項 1 に記載の方法。

【請求項 3】

前記第 1 の競合因子が、前記抗体間の競合を引き起こす前記抗原のエピトープに対応する、請求項 2 に記載の方法。

【請求項 4】

受け取った抗体競合データが、複数の実験ランに由来するデータを含み、それぞれの実験ランが、抗体のセット間のペアワイズ競合を示すデータ値を生成し、前記複数の実験ランが、抗体の異なるセットに関する抗体競合データを生成する、請求項 2 に記載の方法。

20

【請求項 5】

前記抗体競合データをプロセッシングすることが、前記複数の実験ランに由来する前記抗体競合データを組み合わせることを含む、請求項 4 に記載の方法。

【請求項 6】

前記複数の隠れた変数および前記隠れた変数に関する前記親和性スコアを導出することが、前記異なるセットの抗体から、前記抗体に関する親和性スコアを導出することを含む、請求項 5 に記載の方法。

30

【請求項 7】

前記隠れた変数が、前記訓練データに由来するペアワイズ競合データ値、前記隠れた変数に関する前記抗体の親和性スコアを表す隠れたロジット値を使用して、前記抗体に関する前記隠れたロジット値を最適化することによって導出される、請求項 1 に記載の方法。

【請求項 8】

前記抗体の隠れたロジット値が、損失関数、前記訓練データに由来する前記ペアワイズ競合データ値、および前記損失関数を最適化するために前記隠れたロジット値を調整する勾配技術を使用して最適化される、請求項 7 に記載の方法。

【請求項 9】

前記隠れた変数および前記隠れた変数に関する前記親和性スコアが、第 1 の隠れた変数に関する前記抗体の隠れたロジット値を最初に最適化すること；ならびに

40

前記第 1 の隠れた変数の最初の最適化後にさらなる隠れた変数を連続的に加えること、および前記第 1 の隠れた変数およびそれぞれの連続的に加えられたさらなる隠れた変数に関する抗体の隠れたロジット値を連結的に最適化することによって導出される、請求項 8 に記載の方法。

【請求項 10】

2 つの抗体に関して最適化された前記隠れたロジット値を使用して、前記 2 つの抗体に関するペアワイズ競合スコア予測を生成することをさらに含む、請求項 7 に記載の方法。

【請求項 11】

50

前記受け取った抗体競合データが、前記２つの抗体に関するペアワイズ競合データを含まない、請求項１０に記載の方法。

【請求項１２】

結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するためのシステムであって、

プロセッサ；ならびに

前記プロセッサによる実行のためのインストラクションを保存するメモリであって、前記インストラクションが、前記プロセッサを、

抗体間のペアワイズ競合を示すデータ値を含む、複数の抗体および抗原に関する抗体競合データを受け取る；

10

訓練データを生成するために、前記抗体競合データをプロセッシングする；ならびに

前記訓練データおよび最適化エンジンを使用して、複数の隠れた変数および前記隠れた変数に関する親和性スコアを導出し、前記隠れた変数に関する親和性スコアが、それぞれの抗体について導出され、前記隠れた変数が、前記抗体間の競合を引き起こす前記抗原に関する競合因子を表す

ように構成する、メモリを含む、システム。

【請求項１３】

第１の隠れた変数が、前記抗原に関する第１の競合因子を表し、所与の抗体に関連する前記第１の隠れた変数に関する導出された親和性スコアが、前記第１の競合因子に関する前記所与の抗体の競合の程度を示す、請求項１２に記載のシステム。

20

【請求項１４】

前記第１の競合因子が、前記抗体間の競合を引き起こす前記抗原のエピトープに対応する、請求項１３に記載のシステム。

【請求項１５】

プロセッサによって実行された場合、前記プロセッサに、結合パターンを探索するための抗体競合データに基づいて隠れた変数を導出するインストラクションが保存された非一時的コンピュータ可読媒体であって、実行された場合、前記インストラクションが、前記プロセッサに、

抗体間のペアワイズ競合を示すデータ値を含む、複数の抗体および抗原に関する抗体競合データを受け取らせる；

30

訓練データを生成するために、前記抗体競合データをプロセッシングさせる；ならびに

前記訓練データおよび最適化エンジンを使用して、複数の隠れた変数および前記隠れた変数に関する親和性スコアを導出させ、前記隠れた変数に関する親和性スコアが、それぞれの抗体について導出され、前記隠れた変数が、前記抗体間の競合を引き起こす前記抗原に関する競合因子を表す、

非一時的コンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【０００１】

連邦政府により支援された研究または開発に関する陳述

40

本発明は、国防高等研究計画局（Defense Advanced Research Projects Agency）によって与えられたD18AC00002の下、米国政府の支援によって為された。米国政府は、本発明において、ある特定の権利を有する。

【０００２】

分野

本開示の実施形態は一般に、結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出することに関する。

【背景技術】

【０００３】

50

背景

モノクローナル抗体（「m A B」）発見は複雑であり、時間がかかり、資源集約的な技術的課題である。m A B 発見の 1 つの構成要素は、抗原に結合するときに抗体がどのように競合するかを理解することを含む。エピトープビニングは、これをさらに理解することができる有益な技術である。しかしながら、従来の手法は、これらの抗体がどのように競合するかに関する見解が限られたモデルしか達成できない。

【発明の概要】

【課題を解決するための手段】

【0004】

要旨

10

本開示の実施形態は、結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するためのシステムおよび方法を対象とする。複数の抗体および抗原に関する抗体競合データを受け取ることができ、抗体競合データは、抗体間のペアワイズ競合を示すデータ値を含む。訓練データを生成するために、抗体競合データをプロセッシングすることができる。訓練データおよび最適化エンジンを使用して、複数の隠れた変数および隠れた変数に関する親和性スコアを導出することができ、隠れた変数に関する親和性スコアは、それぞれの抗体について導出され、隠れた変数は、抗体間の競合を引き起こす抗原に関する競合因子を表す。

【0005】

実施形態の特徴および利点は、以下の説明に記載されるか、または説明から明らかであるか、または本開示の実施によって学習され得る。

20

【0006】

さらなる実施形態、詳細、利点、および改変は、添付の図面と併用される、好ましい実施形態の以下の詳細な説明から明らかとなるであろう。

【図面の簡単な説明】

【0007】

【図1】図1は、例示的实施形態に従って結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するためのシステムを示す。

【0008】

【図2】図2は、例示的实施形態による計算システムの略図を示す。

30

【0009】

【図3】図3は、モノクローナル抗体に関する競合データを示す従来のヒートマップを示す。

【0010】

【図4】図4は、競合データに基づいてモノクローナル抗体をビニングするための以前のネットワーク手法を示す。

【0011】

【図5】図5は、モノクローナル抗体に関する競合力学を示す。

【0012】

【図6】図6は、例示的实施形態に従って結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するための流れ図を示す。

40

【発明を実施するための形態】

【0013】

詳細な説明

実施形態は、ペアワイズ抗体競合データに基づいてモノクローナル抗体間の競合パターンを示す隠れた変数の情報を導出する。例えば、抗体 - 抗原結合の予測数学モデルを、最適化エンジンによって発見することができる。一部の実施形態では、最適化エンジンは、抗体のペアが互いに競合するかどうかに関する予測を生成するための基礎を形成する隠れた変数のセットを導出することができる。これらの隠れた変数は、結合時に抗体によって使用されるエピトープ結合資源または「抗原不動産 (antigen real estate)」とし

50

て大まかに考えることができる。

【0014】

一部の実施形態では、モデルが、これらの資源が実際に抗原表面上のどこに存在するかに関して不可知であるため、変数は、「隠れている」。例えば、それぞれの隠れた変数は、抗体が結合するために使用する抗原上の一部のエピトープ資源のためのプレースホルダーであってもよい。一部のインプリメンテーションでは、一部の隠れた変数は、一部の他の競合因子（例えば、エピトープ / 位置競合以外の）を表してもよい。

【0015】

一部の実施形態では、最適化エンジンは、隠れた変数のロジット値を生成し、これらのロジット値を、抗体に関する訓練データ中に存在する観察された競合データ値（例えば、ペアワイズ抗体競合）と比較することができる。一部の実施形態では、損失関数が最適化される、および / または測定基準が達成される（例えば、収束が達成される）まで、抗体の隠れた変数のロジット値を調整する勾配を実装することによって、損失関数を最適化することができる。例えば、抗体に関する隠れた変数のロジット値の最適化は、隠れた変数によって表される競合因子（例えば、隠れた変数によって表される抗原上のエピトープに関して）に関する抗体の競合レベルを示す / 予測する隠れた変数の親和性スコアを達成することができる。一部の実施形態では、抗体間のペアワイズ競合予測スコアを、抗体に関するロジット値を使用して生成することができる。

10

【0016】

実施形態はまた、異なる抗体競合データ上で訓練された複数の隠れた変数のモデルに由来する予測（例えば、競合スコア）を組み合わせることによって集合学習技術を実装することもできる。例えば、それぞれの隠れた変数のモデルを、異なるセットの抗体に関する競合データを使用して訓練することができる。2つの抗体が競合するかどうかに関する予測を、いくつかの訓練された隠れた変数のモデルに由来する競合スコアを組み合わせることによって生成することができる。

20

【0017】

一部の実施形態では、ランドマーク抗体相関モデルは、測定されていない抗体間のペアワイズ競合（例えば、特定の抗原に対する）を予測するための所定のセットのランドマーク抗体に関する競合測定値を使用することができる。例えば、競合予測が望まれる抗体のペアを考慮して、ランドマーク抗体を用いてそれぞれの抗体の競合測定値間の相関を算出することができる。相関値に基づいて、競合尤度を予測することができる。

30

【0018】

従来のエピトープビニングは、同じ結合領域（例えば、エピトープ）について競合する抗体が、ピンの中に一緒にグループ化されるように分析される競合データを導出するための、コンビナトリアル様式（例えば、ペアワイズ）での抗体の試験（例えば、「実験ラン（experimental run）」を実行するデバイスを使用する）を含む。エピトープビニング実験は、大量のデータを生成する。例えば、一部のビニング実験では、それぞれの実験ランに関して、関与する抗体のペア毎に、データポイント（例えば、数値）が生成される。一部のランは、実験あたり最大で384の抗体を含んでもよく、これは、異なる抗体間のペアワイズ競合に関する最大で $384 * 384 = 147,456$ の観測値が存在することを意味する。さらに、時には、1回の実験ランにおける現在のデバイス支援よりもさらに大きい抗体群にわたってエピトープビニングを実行することが有利であるか、またはこれらの抗体の全てのペアに関して競合実験を実行することなく、新しく発見された抗体を用いて以前のエピトープビニングランを拡張することが有利である。

40

【0019】

実施形態は、単一または複数のエピトープビニングランの結果を分析および理解するための改善されたモデルを達成する。さらに、改善されたモデルは、実験結果を、それぞれの抗体を単一のピンに単に割り当てるよりも有益な方法でそれらを一緒にグループ化することができるように、個々の抗体の特性に起因すると考えることができる。実施形態は、現在のデバイス制限によって、一度に384の抗体に制限される、複数のエピトープビニ

50

ング実験からの結果を組み合わせる技術を支援する。実施形態はまた、実験全体を繰り返すことなく、新しい抗体を用いて現行のエピトープビニングランを拡張することもできる。さらに、異なるエピトープビニングランに関与した抗体について（例えば、それらについて、それらが競合するかどうかに関する直接的な実験情報がない）、モデルの実施形態は、これらの抗体が競合するかどうかに関する予測を支援する。

【0020】

実施形態は、特定の抗原に対するペアワイズ抗体競合測定値を収集および組織化するための技術を、それらを実際に測定する実験ランを実行する前に（またはそれを行うことなく）、これらのペアワイズ抗体競合測定値を予測することができるモデルを使用することによって最適化する。したがって、実施形態は、従来のエピトープビニング手法と比較した場合、所望の抗体競合データを生成する（および資源効率を有意に改善する）のに必要な実験ランの回数を有意に減少させることができる。

10

【0021】

ここで、本開示の実施形態を詳細に参照し、その例は添付の図面に示されている。以下の詳細な説明において、いくつかの特定の詳細は、本開示の完全な理解を提供するために記載される。しかしながら、本開示はこれらの特定の詳細がなくても実施することができることが当業者には明らかであろう。他の例では、周知の方法、手順、構成要素、および回路は、実施形態の態様を不必要にわかりにくくしないように、詳細に説明されていない。可能な限り、同様の参照番号が同様の要素について使用されるであろう。

【0022】

20

図1は、例示の実施形態に従って結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するためのシステムを示す。システム100は、抗体競合データ102、プロセッシングモジュール104、最適化エンジン106、および分析モジュール108を含む。例えば、抗体競合データ102は、抗体および抗原とのその相互作用（例えば、ペアワイズ競合）を特徴付ける数値結果を生成する表面プラズモン共鳴（「SPR」）実験技術から生成されたデータを含んでもよい。一部のインプリメンテーションでは、抗体競合データは、Carterra（登録商標）LSA（商標）機器を使用して生成される。

【0023】

一部の実施形態では、抗体競合データ102は、いくつかの実験ランに由来するデータを含んでもよい。例えば、実験ランは、所与の抗原に関する2つの抗体間のペアワイズ競合を示す数値を生成することができ、全抗体競合データ102において、いくつかの（例えば、数十、数百、または数千）の抗体間の相互作用に関するデータを含んでもよい。一部の実施形態では、競合データ102は、バイナリー値（例えば、1または0、真または偽など）を使用して2つの抗体が競合するかどうかを示すバイナリーペアワイズ競合データを含む。

30

【0024】

プロセッシングモジュール104は、訓練データが最適化エンジン106のために生成されるように、抗体競合データ102をプロセッシングすることができる。例えば、抗体競合データ102は、複数の実験ランに由来するデータを含んでもよく、プロセッシングモジュール104は、最適化エンジン106によるプロセッシングにとって好適な様式でこのデータを組み合わせることができる。プロセッシングモジュール104の実施形態は、関数（例えば、バイナリー値を割り当てる関数）を使用して競合データ102に由来する数値を変換するか、または他の好適なデータ変換を実行することもできる。

40

【0025】

最適化エンジン106は、プロセッシングモジュール104によって生成された訓練データに基づいて隠れた変数および関与する抗体に関する隠れた変数の親和性スコアを導出することができる。例えば、最適化エンジン106は、隠れた変数のロジット値（例えば、抗体の隠れた変数の親和性スコアを表すロジット値）を生成し、これらのロジット値を、抗体に関する訓練データ中に存在する観察された競合データ値（例えば、ペアワイズ抗体

50

競合)と比較することができる。一部の実施形態では、損失関数が最適化される、および/または測定基準が達成される(例えば、収束が達成される)まで、抗体の隠れた変数のロジット値を調整する勾配を実装することによって、損失関数を最適化することができる。

【0026】

例えば、抗体に関する隠れた変数のロジット値の最適化は、隠れた変数によって表される競合因子(例えば、隠れた変数によって表される抗原上のエピトープ)に関する抗体の競合レベルを示す/予測する隠れた変数の親和性スコアを達成することができる。一部のインプリメンテーションでは、隠れた変数は、エピトープ位置を越えた抗原結合のための競合因子(例えば、同じ結合位置に関する競合を越えた干渉/競合因子)と相関してもよい。

10

【0027】

分析モジュール108は、最適化エンジン106からの出力に基づいて抗体に関する競合情報を生成することができる。例えば、最適化エンジン106は、複数の抗体間の競合を予測/発見するためのモデルを出力することができる。特に、最適化エンジン106によって生成されたモデルは、異なる競合因子(例えば、異なるエピトープまたは他の競合因子)について競合する抗体を発見することができる。したがって、分析モジュール108を使用して、異なる隠れた変数の親和性値を有する抗体(例えば、異なる方法で抗原について競合する抗体)のパネルを生成することができる。そのようなパネルは、正の処置転帰に対する多様な経路を提供することができ、かくして、正の健康転帰をもたらすモノクローナル抗体の製造/発見に対する改善を意味する。

20

【0028】

図2は、実施形態による計算システム200の略図である。図2に示されるように、システム200は、プロセッサ212、データ214、メモリー216、および/またはシステム200の他のコンポーネント間で情報を通信するように構成された、バス210、ならびに他の要素を含んでもよい。プロセッサ212は、コマンドを実行し、計算を実施し、および/またはシステム210の機能を制御するように構成された1つまたは複数の一般的な、または特定の目的のプロセッサを含んでもよい。プロセッサ212は、マイクロプロセシングデバイスなどの単一の集積回路を含んでもよいが、または複数の集積回路デバイスおよび/もしくはは組み合わせて動作する回路基板を含んでもよい。プロセッサ212は、メモリー216に保存されたオペレーティングシステム218、最適化エンジン220、および/または他のアプリケーションなどのソフトウェアを実行してもよい。

30

【0029】

通信コンポーネント222は、ネットワーク(示されない)上でシステム200の1つまたは複数のコンポーネントから別のデバイスに送信されるデータをプロセッシング(例えば、エンコード)すること、およびネットワーク上の別のシステムから受信したデータを、システム200の1つまたは複数のコンポーネントのためにプロセッシング(例えば、デコード)することなどによって、システム200および他のデバイスのコンポーネント間の接続を可能にしてもよい。例えば、通信コンポーネント222は、無線ネットワーク通信を提供するように構成されたネットワークインタフェースカードを含んでもよい。任意の好適な無線通信プロトコルまたは技術を、Wi-Fi、Bluetooth(登録商標)、Zigbee(登録商標)、ラジオ、赤外線、および/または移動体通信技術およびプロトコルなどの通信コンポーネント222によって実装することができる。一部の実施形態では、通信コンポーネント222は、Ethernet(登録商標)などの、有線ネットワーク接続、技術、およびプロトコルを提供してもよい。

40

【0030】

システム200は、プロセッサ212に関する情報およびインストラクションを保存することができるメモリー216を含む。メモリー216の実施形態は、データを回収し、読み取り、書き込み、改変し、保存するためのコンポーネントを含有する。メモリー2

50

１６は、プロセッサ２１２によって実行された時に機能を実施するソフトウェアを保存してもよい。例えば、オペレーティングシステム２１８（およびプロセッサ２１２）は、システム２００のためのオペレーティングシステム機能を提供することができる。最適化エンジン２２０（およびプロセッサ２１２）は、実施形態による抗体競合を予測／発見するためのモデルを生成することができる。最適化エンジン２２０の実施形態を、インメモリ構成として実装することができる。メモリ２１６のソフトウェアモジュールは、オペレーティングシステム２１８、最適化エンジン２２０、ならびに他のアプリケーションモジュール（描写されない）を含んでもよい。

【００３１】

メモリ２１６は、システム２００のコンポーネントによってアクセス可能な非一時的コンピュータ可読媒体を含む。例えば、メモリ２１６は、ランダムアクセスメモリ（「RAM」）、ダイナミックRAM（「DRAM」）、スタティックRAM（「SRAM」）、読み出し専用メモリ（「ROM」）、フラッシュメモリ、キャッシュメモリ、および／または任意の他の型の非一時的コンピュータ可読媒体の任意の組合せを含んでもよい。データベース２１４は、システム２００のコンポーネントのための保管を提供するためにシステム２００の他のコンポーネントに通信接続される（バス２１０を介してなど）。データベース２１４の実施形態は、論理的に関連する記録またはファイルの統合された収集物中にデータを保存することができる。

【００３２】

データベース２１４は、データウェアハウス、分散型データベース、クラウド型データベース、安全なデータベース、分析データベース、生産データベース、非生産データベース、エンドユーザーデータベース、遠隔データベース、インメモリーデータベース、リアルタイムデータベース、リレーショナルデータベース、オブジェクト指向型データベース、階層型データベース、多次元データベース、Hadoop分散ファイルシステム（「HDFS」）、NoSQLデータベース、または当業界で公知の任意の他のデータベースであってもよい。システム２００のコンポーネントは、プロセッサ２１２が、ユーザーに対して情報、データ、および任意の他の好適な表示物を表示することができるディスプレイ２２４、キーボードなどのＩ／Ｏデバイス２２６、ならびにコンピュータマウスなどのＩ／Ｏデバイス２２８または任意の他の好適なＩ／Ｏデバイスにさらに連結される（例えば、バス２１０を介して）。

【００３３】

一部の実施形態では、システム２００は、システムアーキテクチャ、分散システム、または他の好適なシステムの要素であってもよい。例えば、システム２００は、Carterra（登録商標）LSA（商標）機器の種々のモジュール、抗体競合データを生成するための任意の他の好適なデバイス、または任意の他の好適なモジュールを含んでもよい、１つまたは複数の追加の機能モジュールを含んでもよい。

【００３４】

システム２００の実施形態は、別々のデバイスのための関連機能を遠隔で提供することができる。一部の実施形態では、システム２００の１つまたは複数のコンポーネントは、実装されなくてもよい。例えば、システム２００は、ディスプレイ、１つまたは複数のプロセッサ、およびメモリを含むが、図２に示されるシステム２００の１つまたは複数の他のコンポーネントを含まない、タブレット、スマートフォン、または他の無線デバイスであってもよい。一部の実施形態では、システム２００のインプリメンテーションは、図２に示されていない追加のコンポーネントを含んでもよい。図２は、単一のシステムとしてシステム２００を描写するが、システム２００の機能を、クラウドインフラストラクチャー内で、または任意の他の好適な様式で、分散システムとして、異なる位置で実装することができる。一部の実施形態では、メモリ２１６、プロセッサ２１２、および／またはデータベース２１４は、分散型である（システム２００を示す複数のデバイスまたはコンピュータにわたる）。一実施形態では、システム２００は、計算デバイス（例えば、スマートフォン、タブレット、コンピュータなど）の一部であってもよい。

【 0 0 3 5 】

モノクローナル抗体（「m A B」）発見は複雑であり、時間消費的であり、資源集約的な技術的課題である。m A B 発見の 1 つの構成要素は、抗原に結合する時に抗体がどのように競合するかを理解することを含む。エピトープビニングは、これをさらに理解する有益な手法である。特に、従来のエピトープビニングは、同じ結合領域（例えば、エピトープ）について競合する抗体が、ピンの中に一緒にグループ化されるように分析される競合データを導出するための、コンビナトリアル様式（例えば、ペアワイズ）での抗体の試験（例えば、「実験ラン」を実行するデバイスを使用する）を含む。実験ランによって生成される例示的な競合データは、図 3 のヒートマップ 3 0 0 に描写されている。ヒートマップ 3 0 0 は、行および列にわたって異なる抗体を含み、2 つの抗体の交点の数値は、それらの間のペアワイズ競合を示す。

10

【 0 0 3 6 】

エピトープビニングに対する一部の以前の手法は、グラフクラスタリングアルゴリズムに基づく。これらの手法では、抗体は、競合グラフ内の接続エッジの近接性または数に基づいて単一のクラスターに割り当てられる。図 4 は、競合データに基づいてモノクローナル抗体をビニングするための従来のネットワーク手法を示す。ネットワークグラフ 4 0 0 は、以前のグラフクラスタリング手法を使用するピン 4 0 2 を含む。図 4 に描写されるように、それぞれの抗体は、抗体の競合プロファイルに基づいて、単一のピン 4 0 2、またはクラスターを割り当てられる。

20

【 0 0 3 7 】

抗体を単一のピンに割り当てるよりもむしろ、実施形態は、隠れた変数のセットに基づいてそれぞれの抗体の親和性スコア数を割り当てる。例えば、これらの隠れた変数を使用して、観察されなかった抗体のペア間の競合を予測することができ、それらは、抗体が抗原に結合するために使用する結合パターンの理解において本質的に有用でもある。

【 0 0 3 8 】

いくつかの利点があるが、実施形態によって実装される隠れた変数の手法によって達成される。例えば、実施形態は、それぞれの抗体を単一のクラスターに割り当てるクラスターに基づくモデルよりも高い忠実度で観察された実験データをモデル化することができる。具体的には、実験データが抗体競合の非推移的パターンを示す場合、これは、抗体を単一のクラスターに割り当てるモデルを使用してよく表すことができない。図 5 は、以前の手法においてこの欠点を示すモノクローナル抗体に関する競合力学を示す。具体的には、略図 5 0 0 は、

30

グループ A はグループ B と相当量競合する
 グループ B はグループ C と相当量競合する
 グループ A はグループ C と相当量競合しない
 ことを描写する。

【 0 0 3 9 】

クラスターに基づくモデルは、これらのグループを割り当てるべき単一のクラスターを決定することができない。しかしながら、「隠れた変数」モデルは、それぞれのグループの異なる親和性の抗体を 2 つの異なる隠れた変数に割り当てることによってこのパターンの競合を説明することができる：

40

【 表 1 】

	隠れた変数 0	隠れた変数 1
グループ A	高親和性	低親和性
グループ B	高親和性	高親和性
グループ C	低親和性	高親和性

50

【 0 0 4 0 】

したがって、以前の手法は限られた洞察を有していたが、より高い忠実度の分析を、隠れた変数の手法の実施形態を使用して導出することができる。単一の抗体が、同時に、および特定のグループへのメンバーシップに関するバイナリー判断よりもむしろ親和性数値を用いて、1より多いクラスターに属するのを可能にするものとして隠れた変数を考えるのが有用であり得る。同時に、これらの特性は、モデルに、抗体競合に関する予測を行わせるのを可能にする。

【 0 0 4 1 】

クラスターに基づく手法の別の制限は、得られるモデルが、競合データを組み合わせて、実験装備の異なるランにわたってより大きい競合行列を生成することなどによって、抗体が競合するかどうかに関するロバストな予測を行うことができないということである。実施形態は、それぞれの抗体を単一のクラスターに単に割り当てるよりもむしろ、それぞれの抗体に対する親和性数値を異なる隠れた変数に割り当てるモデルを生成する。この手法は、異なるエピトープビニングランに由来する抗体が互いに競合するかどうかに関する数値予測を支援する。

【 0 0 4 2 】

複数のエピトープビニングランに由来するデータを一緒に結合させる利点は、一般に公知の行列競合問題に対する新規手法と考えることができる。例えば、複数のエピトープビニングランは、完全に交差しないことが多いため、抗体競合の全てのペアを記述する行列は不完全である（例えば、複数のランに及ぶ抗体が表形式で行および列上に列挙される場合、一部の交点に関するデータは失われているであろう）。行列がペアワイズ抗体競合を表す場合、実施形態によって取られる隠れた変数の親和性手法は、利用可能な競合データに基づく最適化を手段として不完全な行列を「完成させる」ことができる。

【 0 0 4 3 】

一部の実施形態では、モデルに由来する数値予測を、モデルが、ノイズのある、および/または矛盾する実験的証拠を含み、かくして、より高い、またはより低い信頼性で予測を行うことを可能にする（例えば、証拠の強度に応じて）、信頼スコアと解釈することができる。隠れた変数の親和性スコアのさらなる利益は、モデルが、t分布型確率的近傍埋め込み法（「t - S N E」）または一様多様体の近似と投影（「U M A P」）などの次元削減技術を支援して、複数のエピトープビニングランにわたる抗体のグループ間の関係を示す2次元クラスタリングプロットを生成することができるということである。例えば、Euclidean、Manhattanなどのような任意の好適な距離測定基準を使用して、それぞれの抗体に関する隠れた変数の親和性の間でペアワイズ距離行列を計算し、その距離行列を、UMAPなどの次元削減システムによって実行することができる。一部の技術は、エピトープビニングランのセットに関する完全な競合行列を帰属させ、その列と行との間の距離を使用して抗体に関するペアワイズ距離行列を計算し、次元削減システムによってそのペアワイズ距離行列を送信することもできる。

【 0 0 4 4 】

実施形態における隠れた変数の親和性スコアを、それぞれの隠れた変数を有するそれぞれの抗体の親和性を表す「隠れたロジット」の表として保存することができる。例えば、隠れたロジットは、任意の有限数であってもよい。一部の実施形態では、正の値は、より高い親和性を表し、負の値は、より低い親和性を表す。以下は、抗体および3つの隠れた変数のセットに関する、それぞれの抗体の隠れたロジット値を示す表の例である。

10

20

30

40

【表 2】

	H0 ロジット	H1 ロジット	H2 ロジット
Ab1	2.20	-2.20	-0.85
Ab2	1.39	-2.20	0.00
Ab3	-2.20	2.20	-1.39
Ab4	0.85	1.39	-2.20

10

【0 0 4 5】

モデルにおいて隠れたロジットを使用する前に、一部の実施形態では、シグモイド関数によって数値を送信することができる。例えば、シグモイド関数は、それらを、(0 . . . 1) の範囲に変換することができる。ここで、0 より大きい隠れたロジット値は、0 . 5 より大きい隠れた変数になる。以下は、隠れたロジットの例である。

【表 3】

	H0	H1	H2
Ab1	0.9	0.1	0.3
Ab2	0.8	0.1	0.5
Ab3	0.1	0.9	0.2
Ab4	0.7	0.8	0.1

20

【0 0 4 6】

「隠れたロジット」の値が、それぞれの抗体とそれぞれの隠れた変数との間の正規化された親和性スコアを表すため、「隠れたロジット」の標識が使用される。一部の実施形態では、隠れた変数の親和性を、抗体が結合するために特定の隠れた変数を必要とするかどうかに関するバイナリー判断と解釈することができるように、できるだけ0 または1 に近いものとして隠れた変数の親和性を飽和させるために、モデルフィッティングが使用されるが、それは矛盾する証拠および他の要因のため、いつも可能というわけではない。

30

【0 0 4 7】

一部の実施形態では、2 つの抗体が競合するかどうかに関する予測は、これらの2 つの抗体が、重複する隠れた変数の資源をどれくらい必要とするかの評価基準に基づく。この評価基準を達成するための1 つの手法は、問題の行内の対応する列における値を一緒に乗算する、ドット積演算である。最後に、予測される競合スコアは、一部の実施形態では、シグモイド演算によって送信され、その値は、(0 . . . 1) の範囲内にある。

40

【0 0 4 8】

以下は、一部の実施形態により、2 つの抗体が競合するかどうかをモデルがどのように予測するかを示すアルゴリズムの例である：

【数 1】

競合(Ab₁, Ab₂) = シグモイド $\left(-\alpha + 2\alpha \sum_{h=0}^{N-1} HV[Ab_1, H_h] HV[Ab_2, H_h]\right)$

【0 0 4 9】

上記アルゴリズムにおいて、HV は、隠れた変数の表中での検索を示す（例えば、シグモイド関数を使用して範囲(0 . . . 1)に変換された後の隠れたロジット）。値 α は、

50

外側シグモイド上の温度パラメーターを表してもよく、それは任意の好適な値（例えば、5、または任意の他の好適な値）を取ってもよい。アルゴリズムのこの実施形態は、シグモイド関数の2つの適用を実装することに留意されたい：1）HVを創出する場合、隠れた変数の表、および2）競合関数を計算する時の最も外側の演算。

【0050】

実施形態はまた、異なる抗体競合データ上で訓練された複数の隠れた変数のモデルに由来する予測（例えば、競合スコア）を組み合わせることによって集合学習技術を実装することもできる。例えば、それぞれの隠れた変数のモデルを、異なるセットの抗体に関する競合データ（例えば、無作為化された訓練セット）を使用して訓練することができる。2つの抗体が競合するかどうかに関する予測を、いくつかの訓練された隠れた変数のモデルに由来する競合スコア（例えば、上で開示されたような、ドット積演算によって算出される）を組み合わせることによって生成することができる。組み合わせたスコアは、任意の他の好適な数学演算によって算出された、平均、加重平均、または組合せであってもよい。

10

【0051】

一部の実施形態では、複数のバージョンの隠れた変数のモデルは、異なるサブセットの抗体競合訓練データを使用して訓練される。例えば、所与のサブセットの訓練データ内の、抗体のグループに関する大部分のペアワイズ競合測定値は完全に除去される。換言すれば、訓練データのサブセットを生成するための無作為なペアワイズ競合測定値を単に除去するよりもむしろ、実施形態は、抗体のグループに関するペアワイズ競合データの大部分を選択的に除去する。異なるサブセットの訓練データ内の抗体のグループに関する競合データのこの選択的除去は、非相関型の訓練された隠れた変数のモデルを達成する。非相関型モデルは、それらを集合手法で組み合わせる場合に、より良好な結果を達成する。

20

【0052】

一部の実施形態では、訓練データのサブセットを生成するために抗体のグループに関するペアワイズ競合データの大部分が除去されるが、このグループに関する一部の競合データを維持することができる。例えば、訓練データの全セットに由来する所定のセットの抗体を、持続的抗体として指定し、これらの持続的抗体に関する競合データを、訓練データのサブセットにわたって維持することができる。これらの実施形態では、訓練データのサブセットを生成するために抗体のグループに関するペアワイズ競合データを選択的に除去する場合、抗体のグループと持続的抗体との間のペアワイズ競合データは、訓練データのサブセット中に維持される。これらの実施形態は、持続的抗体に関する競合データによって支援される訓練からそれぞれ恩恵を受ける非相関型モデルを訓練することができる。

30

【0053】

この集合技術の1つの利点は、集合中の個々のモデルに由来する予測における変動（例えば、算出された分散測定基準）が、その全体的な予測において集合モデルの信頼性または確実性の判定基準を提供するということである。さらに、これらの変動はまた、基礎データ分布が変化する場合、予測がどれくらい変化し得るかの指標も提供する。一部の実施形態では、集合技術は、1つまたは複数の隠れた変数のモデルおよび任意の他の好適なモデル（例えば、ランドマーク相関モデル）に由来するペアワイズ抗体競合予測を組み合わせることができる。

40

【0054】

一部の実施形態は、予測スコアを生成するためにドット積演算を利用する。一部の単純なドット積モデルは機械学習の分野においてある特定の問題を構築し、最適化するために存在するが、最適化モデルの実施形態と一部の現存するモデルの実施形態との間には違いがある：

- ・ `word2vec` などのある特定の自然言語プロセッシングモデルは、近隣の単語にわたって確率分布をモデル化するが、実施形態は、バイナリーイベントとして抗体競合をモデル化する。例えば、実施形態は、可能な場合、外側シグモイドを飽和させることを目的とするが、これは `word2vec` においてはあまり望ましくないため、この相違に一部

50

起因して、外側シグモイドの前のシフトが実装される；

・一部の現行モデルは、大きい単語の埋め込みなどの、実施形態よりもより隠れた多くの変数を使用することができる。例えば、一部の実施形態による隠れた変数のモデルが自然言語プロセッシングモデルと比較される場合、一部の相違は、隠れた変数のモデルの実施形態における訓練データがより小さいこと；および隠れた変数のモデルの実施形態が、パターンを解明するために隠れた変数を調査するために使用されることである。

【 0 0 5 5 】

一部の実施形態では、隠れた変数の数およびドット積後のシフトの程度は、モデルの調整可能なハイパーパラメーターである。例えば、一部のデータセットに関するデータ内の多くの競合パターンを表すには、実験的には、5 ～ 1 0 個の隠れた変数で十分であるが、任意の他の好適な数の隠れた変数を実装することができる。解釈を容易にするために、隠れた変数の値は、一部の実施形態では、(0 . . . 1) の範囲内に制約される。

【 0 0 5 6 】

一部の実施形態では、モデル訓練を使用して、隠れた変数の値を算出する。例えば、隠れた変数の値は、エピトープビニングランに由来する実験データに基づいて導出される。ランによって生成される実験データの例は、以下の通りである：

【表 4 】

分析物	リガンド	競合
Ab1	Ab2	あり
Ab1	Ab3	なし
Ab1	Ab4	あり
Ab1	Ab5	なし
Ab3	Ab4	あり
Ab3	Ab5	あり
Ab,	Ab5	あり

【 0 0 5 7 】

表形式の実験データは、異なるエピトープビニングランに由来する実験結果の、連結、または積み重ねを支援することに留意されたい。この連結は、複数の異なるエピトープビニングランに由来するデータを使用して隠れた変数の値の連結最適化を達成するために一部の実施形態によって実装される数値最適化手順と共に使用される。同じ隠れた変数が、異なるラン中に存在していた抗体について同じ意味を有するようにするために、両方のエピトープビニングランに関与したクロスラン (cross-run) 抗体の十分なセットが維持されることに留意されたい。クロスラン抗体を最適に選択するために、1 つまたは複数の予測モデルは、その競合挙動において互いの相関が最も低い 1 回目のランに由来する抗体の小さいセットを同定することができ、これらのものを、その後のランにおいて使用されるクロスラン抗体として選択することができる。

【 0 0 5 8 】

実施形態は、モデル予測が損失関数に従って実際の実験データに一致するようにその値を最適化するための、勾配降下の形態などの数値最適化技術を使用して、隠れた変数および隠れた変数の値 (例えば、隠れた変数の親和性スコア) を導出する。いくつかの潜在的なアルゴリズムまたは技術を使用して、このタスクを達成することができる。以下は、一部の実施形態による最適化手順の例である。

・ 0) 0 . 5 の不定の隠れた変数の値を意味する、全ての抗体に関する 0 . 0 の隠れたロジット値で初期化された、単一の隠れた変数 H_0 から開始する。

・ これは勾配が最も高いシグモイド曲線上の点である、すなわち、シグモイドが最小飽和であることに留意されたい。

・ 1) Broyden - Fletcher - Goldfarb - Shanno アルゴリズム (「BFGS」) または L - BFGS などの最適化手順を使用して、訓練セット上のモデルの競合予測の交差エントロピー損失を最適化する。

・ 単一の隠れた変数を用いるだけで、この最適化問題は凸状であり得る。

・ 解は、データセット中の単一の最大の競合グループを表し、予測するためにこの単一の隠れた変数を使用する傾向がある。

・ 必要に応じて、交差エントロピーに小さい L_2 ノルムの項を加える。

・ 2) 最適化が収束した後、それぞれの抗体に関する 0 . 5 の隠れた変数の値を意味する、0 . 0 の中間のロジット値に再初期化された、それぞれの抗体に関する新しい隠れた変数の列を加える。

・ 3) ステップ 1 に由来する同じ損失関数および最適化手順を使用してペアワイズ競合イベントを予測するために、全ての抗体に関する全ての隠れた変数を連結的に最適化する。

・ この最適化問題は、最早凸状ではないが、第 1 の隠れた変数が、一部の抗体について 0 . 5 からおそらく移動したため、対称性は破壊され、オブティマイザー (optimizer) が前進することができる。

・ この点で、オブティマイザーは、データセット中の第 2 の最も一般的な競合グループを拾うことが多い。

・ 提供された検証セット上でのモデルの予測性能が収束するまで、ステップ 2 および 3 を繰り返して、新しい隠れた変数を増加的に加えることができる。

・ この最適化手順は全体として決定性であることに留意されたい。

【 0 0 5 9 】

隠れた変数を増加的に加えるこの技術は、弱学習器の反復的連続が訓練される場合、それぞれの 1 つが、以前の反復においてモデルによって作られた誤差を補正する、機械学習からブーストする着想を連想させる。上記の最適化手順は、ブーストとは異なるが、古いおよび新しい隠れた変数が一緒に連結的に最適化される点で、より早期の反復に由来する隠れた変数が、新しいものが加えられるたびに精緻化されることを可能にする。

【 0 0 6 0 】

隠れた変数を増加的に加える手法はまた、主成分分析 (「PCA」) において行われるように、より低いランクの近似を使用する行列の表示と漠然と類似している。しかしながら、上記の最適化手順は、行列再構築関数が、非線形性を含有し、標準的な切り捨てられた特異値分解の構造を追跡せず、損失関数が、複数のエピトープビニングランを一緒に連結することを可能とするものである、欠測値を有する不完全な行列の再構築を可能にする点で PCA とは異なる。

【 0 0 6 1 】

隠れた変数のモデルの実施形態はまた、例えば、訓練の相違に少なくとも起因して、word2vec などの自然言語プロセッシングモデルとも異なる。一部のさらなる相違は、以下の通りである。

・ 隠れた変数のモデルの実施形態が、バッチよりもむしろ全訓練セットを使用して最適化する。例えば、実施形態における訓練セットは、多くの自然言語プロセッシングアプリケーションのものよりも小さい規模である。したがって、全訓練セットを、それぞれのラウンドの勾配計算のために使用することができ、LBFGS のような二次最適化法を使用して、速度収束を助け、ハイパーパラメーター調整を回避することができる。

・ 二次オブティマイザーは、2 層のシグモイドのため、隠れた変数のモデルのトポロジーの実施形態上でより良好に機能し得る。任意のシグモイドが飽和する場合、それから生じる勾配シグナルは非常に弱くあり得、実施形態はシグモイドを飽和に向かって押し込む

10

20

30

40

50

。

・隠れた変数のモデルの実施形態は、多数の隠れた変数も、それらのための無作為な初期化も必要としない。

・隠れた変数のモデルの実施形態は、固定サイズの埋め込みから開始するよりもむしろ、隠れた変数を増加的に加えることができる。

【0062】

一部の実施形態では、他の最適化技術を実装することもできる。例えば、最適化のための別の選択肢は、固定数の隠れた変数から開示することであってよい。これらの実施形態では、隠れた変数の無作為の初期化を実装して、問題の対称性を破壊することができ、オプティマイザーに進展させることを可能とするが、しかしながら、このインプリメンテーションは、隠れた変数を増加的に加える必要性を回避する。任意の他の好適な最適化技術を実装することができる。

10

【0063】

一部の実施形態では、モデルが訓練されたら、隠れた変数の親和性スコアを使用して、抗体間の競合の傾向を理解することができる。例えば、実施形態の新規特性は、以前に実装されたクラスタリング技術と比較した場合、実施形態が、抗体を複数のグループと関連させるということである。例えば、0.5などの一部のカットオフ値で隠れた変数に関するそれぞれの抗体の親和性スコアを閾値化することによって、これを達成することができる。一部の実施形態では、抗体は、複数の隠れた変数について0.5より高い親和性を有し、かくして、1より多いグループに属してもよい。この新規グループのメンバーシップは、抗体のこれらのグループがどのように交差するかに関するさらなる疑問を可能にする。

20

【0064】

一部の実施形態を使用して、「競合サンドイッチ」と本明細書で称される競合のパターンを検出することができる。競合サンドイッチは、図5に示される競合プロファイルをする抗体の3つのグループを含む。すなわち、
グループAはグループBと相当量競合する
グループBはグループCと相当量競合する
グループAはグループCと相当量競合しない。

【0065】

この場合、グループAおよびCは、「パンのスライス」であり、グループBは、サンドイッチの「具」である。図5は、競合サンドイッチのグラフに基づく概観を描写する。

30

【0066】

一部の環境では、競合サンドイッチは、1つが他の2つの間に「サンドイッチ」された、エピトープの部分的な隣接性または順序を示し得る。この着想を考える一般的な方法は、それが、移行性が保持されない連結性グラフ内のグループを同定するということである。この効果は、多くの状況において興味深いものであり得る。例として、競合サンドイッチが人々の集団間の関連を示し得る方法を考慮すると、

・グループAにおけるほぼ全ての人が互いに知っており、グループBにおけるほぼ全ての人も知っている

40

・グループBにおけるほぼ全ての人が互いに知っており、グループCにおけるほぼ全ての人も知っている

・グループCにおけるほぼ全ての人が互いに知っているが、グループAにおける誰もほとんど知らない。

【0067】

これは、これらの人々の集団が互いに知っている2つの異なる / 別個の基礎となる理由があり、一部の人々（例えば、グループB）については、両方の理由が適用されることを示し得る。同様に、競合サンドイッチは、一部の状況では、抗体間の異なる / 別個の競合を示すことができる。

【0068】

50

一部の実施形態では、導出された隠れた変数および隠れた変数に関する親和性スコアは、競合を観察するための明確な実験ランを必要とすることなく、抗体間の競合を予測するモデルを示すことができる。換言すると、実験的に試験および観察されていない抗体のペアに関して、導出されたモデルを使用して、競合を予測することができる。導出されたモデルの実施形態を、抗体競合に関する予想またはシミュレーションツールと考えることができる。したがって、実施形態は、資源および時間効率を改善することによって、抗体間の競合試験を改善する。

【 0 0 6 9 】

一部の実施形態では、導出されたモデルは、抗体間の高忠実度競合力学を予測することもできる。例えば、異なる隠れた変数の親和性スコアを示す抗体は、異なる結合機構を示し得るが、類似する隠れた変数の親和性スコアを示す抗体は、類似する結合機構を示し得る。隠れた変数の親和性スコアの記述子は、抗体を個々のピンに単に関連付ける以前の手法と比較した場合、競合力学のより洗練された見解を示す。導出されたモデルの実施形態は、よりロバスタなモノクローナル抗体発見および製造プロセスを達成するための多様な隠れた変数の親和性スコアを示す抗体の選択を可能にする。

【 0 0 7 0 】

一部のインプリメンテーションでは、抗体競合を、ランドマーク抗体相関モデルを使用して予測することができる。例えば、ランドマーク抗体相関モデルは、所定のランドマーク抗体のセットを用いてその競合に関してそれぞれの抗体を特徴付けることができる。所定のランドマーク抗体を用いてこれらの競合測定値を考慮する1つの方法は、本明細書に開示される隠れた変数のモデルにおける隠れた変数の代わりとなるものとしてであるが、しかしながら、所定のランドマーク抗体を用いた競合測定値は隠れていない。

【 0 0 7 1 】

以下は、ランドマーク抗体競合測定値のセットを参照するランドマーク抗体競合モデルを示す例である。

【表 5】

ランドマーク抗体競合	D	E	F	G
リガンドとして A と競合する	あり	なし	あり	なし
分析物として A と競合する	あり	なし	あり	なし
リガンドとして B と競合する	なし	あり	なし	なし
分析物として B と競合する	なし	なし	なし	なし
リガンドとして C と競合する	あり	なし	あり	なし
分析物として C と競合する	あり	なし	あり	なし

【 0 0 7 2 】

この例では、抗体 A、B、および C はランドマーク抗体であり、他の 4 つの抗体、D、E、F、および G のそれぞれとの競合を示す測定値が取られた。ペアワイズ予測が望まし

い抗体のセット（例えば、抗体 D、E、F、および G）について、ランドマーク抗体相関モデルの実施形態は、所定のランドマーク抗体（例えば、A、B、C）の同じセットに対する以前に取られた競合測定値を利用する。しかしながら、これらの他の非ランドマーク抗体間のペアワイズ競合測定値は取られておらず、したがって、これらの値は未知である（例えば、抗体 D および E が互いに競合するかどうかは未知である）。

【0073】

抗体 D および E の間の競合の尤度を予測するために、これらの 2 つの抗体の列の間の相関が算出される。上の例によると、列 D および F の間の相関は 1.0 であり、それらがランドマーク抗体に対する正確に同じ競合プロファイルを有するため、抗体 D および F が互いに競合する高い尤度が予測される。他方で、抗体 D および E の間の相関は 0 に近く、これは、それらが互いに競合する可能性が低いことを示している。抗体 G の列の標準偏差は 0 であり、それぞれの列の標準偏差は相関係数の分母に現れるため、抗体 G と、他の全ての抗体との相関は定義されていないことに留意されたい。

10

【0074】

図 6 は、例示的实施形態に従って結合パターンを発見するための抗体競合データに基づいて隠れた変数を導出するための流れ図を示す。一実施形態では、図 6 の機能は、メモリーまたは他のコンピュータ可読媒体または有形媒体に保存されたソフトウェアによって実装され、プロセッサによって実行される。他の実施形態では、それぞれの機能は、ハードウェアによって（例えば、特定用途向け集積回路（「ASIC」）、プログラマブルゲートアレイ（「PGA」）、フィールドプログラマブルゲートアレイ（「FPGA」）などの使用によって）、またはハードウェアとソフトウェアとの任意の組合せによって実行され得る。

20

【0075】

602 では、複数の抗体および抗原に関する抗体競合データを受け取ることができ、抗体競合データは、抗体間のペアワイズ競合を示すデータ値を含む。例えば、実験ラン（例えば、表面プラズモン共鳴（「SPR」）実験技術から生成されたデータ）は、抗体のセットに関するコンビナトリアル（例えば、ペアワイズ）競合データを生成することができる。

【0076】

一部の実施形態では、受け取った抗体競合データは、複数の実験ランに由来するデータを含み、それぞれの実験ランは、抗体のセット間のペアワイズ競合を示すデータ値を生成し、複数の実験ランは、抗体の異なるセットに関する抗体競合データを生成する。

30

【0077】

604 では、訓練データを生成するために、抗体競合データをプロセッシングすることができる。例えば、競合データのプロセッシングは、競合に関するバイナリー表示への数学的変換などの、データ変換を含んでもよい。一部の実施形態では、抗体競合データのプロセッシングは、複数の実験ランに由来する抗体競合データを組み合わせることを含む。

【0078】

606 では、訓練データおよび最適化エンジンを使用して、複数の隠れた変数および隠れた変数に関する親和性スコアを導出することができ、隠れた変数に関する親和性スコアは、それぞれの抗体について導出され、隠れた変数は、抗体間の競合を引き起こす抗原に関する競合因子を表す。例えば、第 1 の隠れた変数は、抗原に関する第 1 の競合因子を表してもよく、所与の抗体に関連する第 1 の隠れた変数に関する導出された親和性スコアは、第 1 の競合因子に関する所与の抗体の競合の程度を示す。一部の実施形態では、第 1 の競合因子は、抗体間の競合を引き起こす抗原のエピトープに対応する。一部の実施形態では、複数の隠れた変数および隠れた変数に関する親和性スコアを導出することは、異なるセットの抗体（例えば、異なる実験ランに関連する異なるセットの抗体）から抗体に関する親和性スコアを導出することを含む。

40

【0079】

一部の実施形態では、隠れた変数は、訓練データに由来するペアワイズ競合データ値、

50

隠れた変数に関する抗体の親和性スコアを表す隠れたロジット値を使用して、抗体に関する隠れたロジット値を最適化することによって導出される。例えば、抗体の隠れたロジット値は、損失関数、訓練データに由来するペアワイズ競合データ値、および損失関数を最適化するために隠れたロジット値を調整する勾配技術を使用して最適化され得る。

【0080】

一部の実施形態では、隠れた変数および隠れた変数に関する親和性スコアは、第1の隠れた変数に関する抗体の隠れたロジット値を最初に最適化すること、ならびに第1の隠れた変数の最初の最適化後にさらなる隠れた変数を連続的に加えること、および第1の隠れた変数およびそれぞれの連続的に加えられたさらなる隠れた変数に関する抗体の隠れたロジット値を連続的に最適化することによって導出される。

10

【0081】

一部の実施形態では、2つの抗体に関するペアワイズ競合スコア予測を、2つの抗体に関して最適化された隠れたロジット値を使用して生成することができる。例えば、受け取った抗体競合データ（例えば、訓練を生成するためにプロセッシングされた）は、2つの抗体に関するペアワイズ競合データを含まなくてもよい。一部の実施形態では、ペアワイズ競合スコア予測は、部分的には、2つの抗体に関する隠れたロジット値に対してドット積演算を実施することによって生成される。

【0082】

一部の実施形態では、導出された隠れた変数および隠れた変数に関する親和性スコアは、競合を観察するための明確な実験ランを必要とすることなく、抗体間の競合を予測するモデルを示すことができる。換言すると、実験的に試験および観察されていない抗体のペアに関して、導出されたモデルを使用して、競合を予測することができる。導出されたモデルの実施形態を、抗体競合に関する予想またはシミュレーションツールと考えることができる。したがって、実施形態は、資源および時間効率を改善することによって、抗体間の競合試験を改善する。

20

【0083】

一部の実施形態では、導出されたモデルは、抗体間の高忠実度競合力学を予測することもできる。例えば、異なる隠れた変数の親和性スコアを示す抗体は、異なる結合機構を示し得るが、類似する隠れた変数の親和性スコアを示す抗体は、類似する結合機構を示し得る。隠れた変数の親和性スコアの記述子は、抗体を個々のピンに単に関連付ける以前の手法と比較した場合、競合力学のより洗練された見解を示す。導出されたモデルの実施形態は、よりロバストなモノクローナル抗体発見および製造プロセスを達成するための多様な隠れた変数の親和性スコアを示す抗体の選択を可能にする。

30

【0084】

本明細書を通して記載された本開示の特性、構造、または特徴を、1つまたは複数の実施形態において、任意の好適な様式で組み合わせることができる。例えば、本明細書を通じた、「一実施形態」、「一部の実施形態」、「ある特定の実施形態」、「ある特定の複数の実施形態」または他の類似する用語の使用は、実施形態と関連して記載される特定の特性、構造、または特徴が本開示の少なくとも1つの実施形態に含まれ得るという事実を指す。かくして、本明細書を通じた、語句「一実施形態」、「一部の実施形態」、「ある特定の実施形態」、「ある特定の複数の実施形態」または他の類似する用語の出現は、必ずしも全て同じ群の実施形態を指すわけではなく、記載される特性、構造、または特徴を、1つまたは複数の実施形態において任意の好適な様式で組み合わせることができる。

40

【0085】

当業者であれば、上で考察された実施形態を、異なる順序のステップ、および/または開示されるものとは異なる構成の要素を用いて実施することができることを容易に理解するであろう。したがって、本開示は概略された実施形態を考慮するが、本開示の趣旨および範囲内にありながら、ある特定の改変、変化、および代替的な構築が明らかとなるであろうことが当業者には明らかであろう。したがって、本開示の境界を決定するためには、添付の特許請求の範囲を参照するべきである。

50

【図面】

【 図 1 】

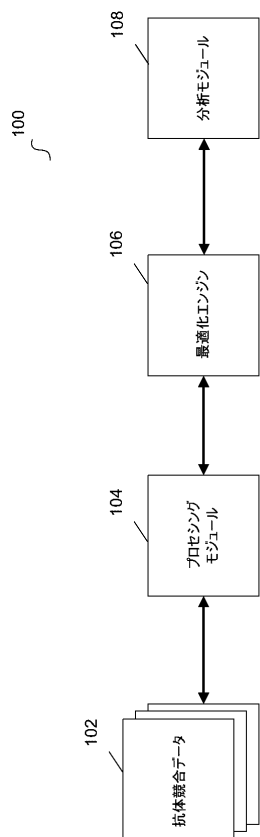


Fig. 1

【 図 2 】

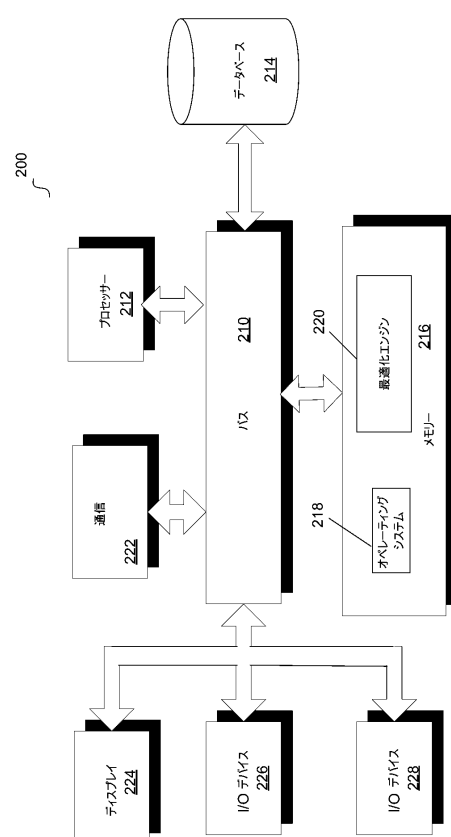


Fig. 2

【 図 3 】

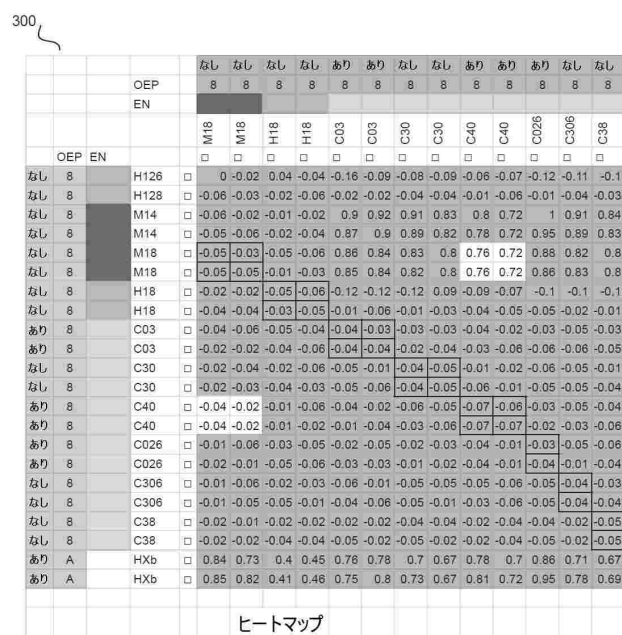


FIG. 3

【圖 4】

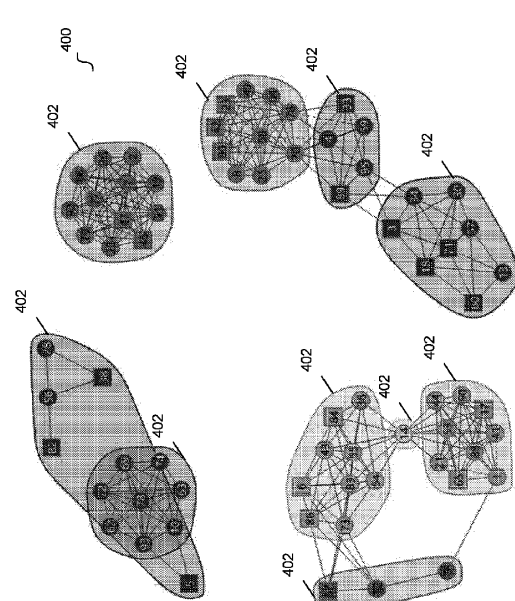
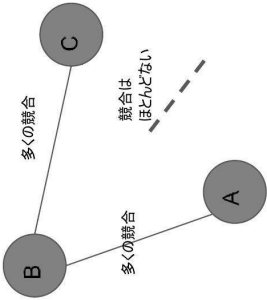


Fig. 4

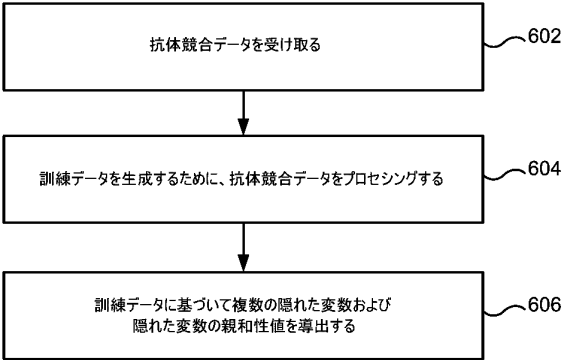
【 図 5 】

500



【 図 6 】

Fig. 5



10

20

Fig. 6

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/036517

A. CLASSIFICATION OF SUBJECT MATTER
INV. G16B40/30
ADD.

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)
G16B C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/192039 A1 (JOJIC NEBOJSA [US] ET AL) 16 August 2007 (2007-08-16) paragraph [0007] - paragraph [0031] paragraph [0045] - paragraph [0052] figures 1-5	1-15
A	Mariette Awad ET AL: "Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers" In: "Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers", 30 April 2015 (2015-04-30), Apress, XP055500724, ISBN: 978-1-4302-5989-3 page 12 - page 13	1-15

☒ Further documents are listed in the continuation of Box C.

☒ 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

28 March 2023

Date of mailing of the international search report

06/04/2023

Name and mailing address of the ISA/
European Patent Office, P.B. 5618 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Schechner-Resom, G

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/036517

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Pittala Srivamshi ET AL: "Learning Context-aware Structural Representations to Predict Antigen and Antibody Binding Interfaces", bioRxiv, 3 June 2019 (2019-06-03), XP055863145, DOI: 10.1101/658054 Retrieved from the Internet: URL:https://www.biorxiv.org/content/10.1101/658054v1.full.pdf [retrieved on 2021-11-18] the whole document -----	1-15

1

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

page 2 of 2

10

20

30

40

50

INTERNATIONAL SEARCH REPORT				International application No PCT/US2022/036517	
Information on patent family members					
Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 2007192039 A1		16-08-2007	EP 2149120 A1		03-02-2010
			US 2007192039 A1		16-08-2007
			WO 2008131311 A1		30-10-2008

10

20

30

40

50

フロントページの続き

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,N
E,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,
CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,J
M,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,ST,SV,SY,T
H,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

弁護士 山本 健策

- (72)発明者 ヒューズ, クリストファー タデウス
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ベルトラン ドゥ ピュイレモン, ヴァレンティヌ ジュリー レイラ
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ドッキング, トーマス ロデリック
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 クラフト, ルーカス
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ハニー, ステファン エドワード
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ジェブソン, ケビン リチャード
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ゴゴルサ, トマス
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 ヤップ, ジョーダン ジョン
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付
- (72)発明者 フォード, アレクサンダー スウォール
カナダ国 ブイ 5 ワイ 0 エー 1 ブリティッシュ コロンビア, バンクーバー, ユーコン ストリ
ート 2 2 1 5, アブセラ バイオロジクス インコーポレイテッド 気付