

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6867959号  
(P6867959)

(45) 発行日 令和3年5月12日 (2021.5.12)

(24) 登録日 令和3年4月13日 (2021.4.13)

(51) Int. Cl.

F I

G 1 6 B 10/00 (2019.01)

G 1 6 B 10/00

G 1 6 H 10/00 (2018.01)

G 1 6 H 10/00

G 1 6 H 20/00 (2018.01)

G 1 6 H 20/00

C 1 2 Q 1/6869 (2018.01)

C 1 2 Q 1/6869

Z

請求項の数 13 (全 11 頁)

(21) 出願番号 特願2017-566789 (P2017-566789)  
 (86) (22) 出願日 平成28年7月11日 (2016.7.11)  
 (65) 公表番号 特表2018-528503 (P2018-528503A)  
 (43) 公表日 平成30年9月27日 (2018.9.27)  
 (86) 国際出願番号 PCT/IB2016/054139  
 (87) 国際公開番号 WO2017/009770  
 (87) 国際公開日 平成29年1月19日 (2017.1.19)  
 審査請求日 令和1年5月9日 (2019.5.9)  
 (31) 優先権主張番号 62/191,684  
 (32) 優先日 平成27年7月13日 (2015.7.13)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)

(73) 特許権者 590000248  
 コーニンクレッカ フィリップス エヌ  
 ヴェ  
 KONINKLIJKE PHILIPS  
 N. V.  
 オランダ国 5656 アーヘー アイン  
 ドーフェン ハイテック キャンパス 5  
 2  
 (74) 代理人 100122769  
 弁理士 笛田 秀仙  
 (74) 代理人 100163809  
 弁理士 五十嵐 貴裕

最終頁に続く

(54) 【発明の名称】 病院環境における感染の追跡

(57) 【特許請求の範囲】

【請求項 1】

感染の広がりを追跡する装置のプロセッサの作動方法において、  
 複数の感染サンプルのそれぞれに関して、配列データを取得するステップと、  
 病原性生物の経時的な変化の予想範囲を決定するステップであって、異なるサンプルに  
 関する前記配列データに線形回帰を使用し、前記変化の予想範囲の境界を定める予測区間  
 を計算する、ステップと、

少なくとも1つの感染サンプルについて、該サンプルと前記複数の感染サンプルからの  
 少なくとも1つの追加的な感染サンプルとの間の経時的な突然変異変化の数を計算するス  
 テップと、

前記経時的な突然変異変化の計算された数のそれぞれについて、前記少なくとも1つの  
 サンプルと前記少なくとも1つの追加的なサンプルとの間の前記経時的な突然変異変化の  
 数が前記予測区間内にあるかどうかを決定するステップと、

前記予測区間内にある前記経時的な突然変異変化の計算された数のそれぞれについて、  
 関連付けられるサンプル対を潜在的な伝染として識別するステップとを有する、作動方法  
 。

【請求項 2】

密接に関連付けられるサンプルを識別するため配列決定されたサンプルの系統発生を構  
 築すること、及び

前記複数の感染サンプルの少なくとも1つについて、前記サンプルと、前記系統発生に

より決定される少なくとも1つの密接に関連付けられるサンプルとの間の経時変化の数を計算することにより、前記変化の予想範囲が決定される、請求項1に記載の作動方法。

【請求項3】

前記変化の予想範囲が、時間にわたり同じ患者から採取される、又は伝染した可能性が高い2人の患者から採取される複数のサンプルから決定される、請求項1に記載の作動方法。

【請求項4】

前記系統発生を構築するステップが、配列決定されたサンプルの各対の間のペアワイズ距離を計算するステップと、距離行列を使用して系統樹を構築するステップとを有する、請求項1に記載の作動方法。

【請求項5】

前記感染サンプル配列データが、完全ゲノム配列決定又は標的配列決定を通して得られる、請求項1に記載の作動方法。

【請求項6】

前記サンプル間の突然変異変化の計算された数が、両方のサンプルがベースコールを持つゲノム位置の総数で割ることにより正規化される、請求項1に記載の作動方法。

【請求項7】

前記変化の予想範囲内の複数の変化を伴うエッジを横断するサンプルの系統樹において幅優先探索を行うことにより、前記感染サンプルのどれが同じアウトブレイクに由来するかを決定するステップを更に有する、請求項4に記載の作動方法。

【請求項8】

配列決定された感染サンプルの表示された系統樹において、潜在的な伝染であると識別された感染サンプルの対をマークするステップを更に有する、請求項7に記載の作動方法。

【請求項9】

複数の系統樹を作成するステップを更に有し、各樹が、同じアウトブレイクに由来するものとしてマークされたサンプルから構築される、請求項8に記載の作動方法。

【請求項10】

感染の広がりを追跡するシステムであって、  
複数の感染サンプルのそれぞれに関して、配列データを受信する受信機と、  
 前記受信された配列データを格納するメモリと、  
 プロセッサとを有し、前記プロセッサが、  
異なるサンプルに関する前記配列データに線形回帰を使用し、予測区間を計算して、病原性生物の経時的な変化の予想範囲を決定し、

少なくとも1つの感染サンプルについて、該サンプルと前記複数の感染サンプルからの少なくとも1つの追加的な感染サンプルとの間の経時的な突然変異変化の数を計算し、

前記経時的な突然変異変化の計算された数の各々について、サンプル間の前記経時的な突然変異変化の数が、前記予測区間内にあるかどうかを決定し、及び

前記予測区間内にある前記経時的な突然変異変化の計算された数の各々について、前記関連付けられるサンプル対を潜在的な伝染として識別する、システム。

【請求項11】

前記プロセッサが、  
 密接に関連付けられるサンプルを識別するため、配列決定されたサンプルの系統発生を構築すること、及び

前記複数の感染サンプルの少なくとも1つについて、前記サンプルと前記系統発生により決定された前記少なくとも1つの密接に関連付けられるサンプルとの間の経時変化の数を計算することにより、前記変化の予想範囲を決定する、請求項10に記載のシステム。

【請求項12】

前記プロセッサが、配列決定されたサンプルの各対の間のペアワイズ距離を計算し、距離行列を使用して系統樹を構築することにより、前記系統発生を構築する、請求項11に

10

20

30

40

50

記載のシステム。

【請求項 13】

前記プロセッサが更に、前記予期される変化範囲における複数の変化を伴うエッジを横切るサンプルの系統樹における幅優先探索を行うことにより、前記感染サンプルのどれが同じアウトブレイクに由来するかを決定する、請求項 12 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本書に記載される様々な実施形態は、医療関連感染、より詳細には、しかし排他的ではないが、院内感染の広がりを追跡するための方法及び装置に関する。

10

【背景技術】

【0002】

ヘルスケア関連感染症（HAI）は、別の状態のヘルスケア治療中に患者により獲得された感染症である。医学文献におけるHAIはしばしば、院内感染と呼ばれる。HAIは致命的であり、病院で頻繁に発生する。それらはしばしば、細菌又は真菌の原因を持つ。入院した患者20人のうち約1人がHAIに罹患する。2002年に実施された最新の調査では、約170万のHAIが存在すると推定されている。これらの170万のHAIは、毎年約99,000人の死亡を引き起こしたか、又はこれに寄与している。ヨーロッパでも同様の傾向が確認されており、そこでは、HAI感染が、年間25,000人の死者の3分の2に寄与しているように見える。

20

【発明の概要】

【発明が解決しようとする課題】

【0003】

院内感染は、重度の肺炎並びに尿路、血流及び体の他の部分の感染を引き起こし得る。多くのタイプは、抗生物質で治療するのが難しく、抗生物質抵抗性が、グラム陰性菌に広がり、病院外の人々に感染する可能性がある。米国では、病院全体で最も頻繁に感染するタイプは、尿路感染症（36%）、外科手術部位感染（20%）、血流感染症及び肺炎（共に11%）である。

【0004】

HAIの重要な経済的影響は、1970年代半ばに実施された院内感染管理（SENIC）の有効性に関する調査に基づき1992年に確認された。公表時点では、医療における直接費は66.5億ドルと推定された（2007年にCPIを用いてインフレ調整済み）。しかしながら、最近公表された証拠によれば、現在の直接費は284億ドルから338億ドルと推定される。その費用の大部分は、患者の長期滞在、病院の隔離、及び感染源の発見と根絶とに関連付けられる。

30

【0005】

上記によれば、HAIの制御及び根絶を容易にするため、HAIの識別及び追跡のための方法及び装置を提供することが望ましい。

【課題を解決するための手段】

【0006】

本書で使用される表現及び用語は、説明のためのものであって限定するものではないことを理解されたい。

40

【0007】

本書に開示される様々な実施形態は概して、医療関連感染に関連し、より詳細には、病院環境における院内感染の広がりを追跡する方法、装置及びプロセッサにより実行される命令を格納する非一時的機械読み取り可能媒体に関する。この目的のため、いくつかの実施形態は、細菌及び真菌から得られる病院関連感染を追跡し、それらを感染源の特定、関連症例の識別などを支援するため、進化的態様で他のサンプルと比較する。

【0008】

種々の実施形態は更に、感染制御手順を支援する目的のため、（別個の無関係の感染で

50

はなく)単一のアウトブレイクにおいて広がっているかもしれない感染サンプルを識別する。同じタイプの感染を持つ複数の患者がいる場合、院内感染対コミュニティが獲得した感染(即ち、病院環境外で伝染した感染)においてどの感染が広がっているのかを特定することが重要である。

#### 【0009】

一般に、感染の広がりを追跡する方法及びシステムが開示される。より一般的には、存在する主要な病原体(例えば、細菌、ウイルス又は真菌)を決定するのに、感染のサンプル由来の配列決定データが用いられる。次いで、変異体のリスト(SNP及びインデルなど)を得るため、配列決定読み取り値が病原体の参照配列にマッピングされる。配列サンプルのコレクションで、各サンプル間のペアワイズ距離が計算され、距離行列が作成される。この距離行列は、感染アウトブレイクがどのように広がったかを示すサンプルの系統樹を構築するために使用される。どの感染が関連付けられるかを決定するため、過去に病原体がどのように進化したかの測定値を用いて、患者-患者間伝染の場合に見られると予想される経時的な期待変化率が決定される。

10

#### 【0010】

予想される変化率は、異なるサンプルについて経時的に見られる変化の数を含むサンプルデータに線形回帰を使用して計算されることができる。HAIが伝染されるときに見られると予想される変化の予想数の境界となる区間である予測区間も計算される。95%のような特定の閾値の予測区間を使用すると、密接に関連付けられるサンプルが、患者間で伝染されるか、又は無関係であるかのいずれかに分類されることができる。関連し得るサンプルの対を決定した後、それらのサンプルは、関連するサンプルの周りにボックスを置くなど、様々な方法により系統樹において視覚的にマークされることができる。これらのサンプルは、典型的には系統樹のサブツリーに入るべきである。

20

#### 【0011】

本書に開示される様々な実施形態は、感染の広がりを追跡する方法を含む。この方法は、密接に関連付けられるサンプルを識別するため、感染サンプル配列データに基づき感染配列決定サンプルの系統発生を構築するステップと、病原性生物の経時的な変化の予想範囲を決定するステップと、少なくとも1つの感染サンプルについて、該サンプルと少なくとも1つの追加的なサンプルとの間の経時的な突然変異変化の数を計算するステップと、上記経時的な突然変異変化の計算された数のそれぞれについて、サンプル間の上記経時的な突然変異変化の数が変化の予想範囲内にあるかどうかを決定するステップと、変化の予想範囲内にある上記経時的な突然変異変化の計算された数のそれぞれについて、関連付けられるサンプル対を潜在的な伝染としてマーキングするステップとを有する。

30

#### 【0012】

いくつかの実施形態では、上記のような感染の広がりを追跡する方法が開示され、上記複数の感染サンプルの少なくとも1つについて、上記サンプルと、上記系統により決定される少なくとも1つの密接に関連付けられるサンプルとの間の経時変化の数を計算することにより、上記変化の予想範囲が決定される。

#### 【0013】

様々な実施形態では、上記のような感染の広がりを追跡する方法が開示され、上記変化の予想範囲が、時間にわたり同じ患者から採取される、又は伝染した可能性が高い2人の患者から採取される複数のサンプルから決定される。

40

#### 【0014】

様々な実施形態では、上記のような感染の広がりを追跡する方法が開示され、上記系統樹を構築するステップが、配列されたサンプルの各対の間のペアワイズ距離を計算するステップと、距離行列を使用して例えばFastTree2又はKimura2パラメータモデルのような手法で系統樹を構築するステップとを有する。

#### 【0015】

様々な実施形態では、上記のような感染の広がりを追跡する方法が開示され、病原性生物の経時的な変化の予想範囲を決定するステップは、異なるサンプルに関する経時変化を

50

含むサンプルデータに線形回帰を使用するステップと、予測区間を計算するステップを有する。

【0016】

様々な実施形態では、上記のような感染の広がりを追跡する方法が開示され、上記感染サンプル配列データが、完全ゲノム配列決定又は標的配列決定を通して得られる。

【0017】

様々な実施形態では、上記のような感染の広がりを追跡する方法が開示され、上記サンプル間の突然変異変化の計算された数が、両方のサンプルがベースコールを持つゲノム位置の総数で割ることにより正規化される。

【0018】

様々な実施形態では、上記のような感染の広がりを追跡する方法が、上記変化の予想範囲内の複数の変化を伴うエッジを横断するサンプルの系統樹において幅優先探索を行うことにより、上記感染サンプルのどれが同じアウトブレイクに由来するかを決定するステップを更に有する。

【0019】

様々な実施形態では、上記のような感染の広がりを追跡する方法が、配列決定された感染サンプルの表示された系統樹において、潜在的な伝染であると識別された感染サンプルの対をマーキングするステップを更に有する。

【0020】

様々な実施形態では、上記のような感染の広がりを追跡する方法が、複数の系統樹を作成するステップを更に有し、各樹が、同じアウトブレイクに由来するものとしてマークされたサンプルから構築される。

【0021】

様々な実施形態が、感染の広がりを追跡するシステムに関する。このシステムが、感染サンプル配列データを受信する受信機と、上記受信された感染サンプル配列データを格納するメモリと、プロセッサとを有し、上記プロセッサが、密接に関連付けられるサンプルを識別するため、配列決定されたサンプルの系統発生を構築し、病原性生物の経時的な変化の予想範囲を決定し、少なくとも1つの感染サンプルについて、該サンプルと少なくとも1つの追加的な感染サンプルとの間の経時的な突然変異変化の数を計算するステップと、上記経時的な突然変異変化の計算された数の各々について、サンプル間の上記経時的な突然変異変化の数が、上記変化の予想範囲内にあるかどうかを決定し、上記変化の予想範囲内にある上記経時的な突然変異変化の計算された数の各々について、上記関連付けられるサンプル対を潜在的な伝染としてマークする。

【0022】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、上記プロセッサが、上記複数の感染サンプルの少なくとも1つについて、上記サンプルと上記系統により決定される上記少なくとも1つの密接に関連付けられるサンプルとの間の経時変化の数を計算することにより、上記変化の予想範囲を決定する。

【0023】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、構成されたプロセッサが、時間にわたり同じ患者から、又は伝染の可能性のある2人の患者から採取される複数のサンプルから、変化の予想範囲を決定する。

【0024】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、上記プロセッサが、配列されたサンプルの各対の間のペアワイズ距離を計算し、距離行列を使用して系統樹を構築することにより、上記系統発生を構築する。

【0025】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、上記プロセッサが、異なるサンプルについての経時変化を含むサンプルデータに線形回帰を使用し、予測区間を計算して病原性生物についての経時的な変化の予想範囲を決定する

10

20

30

40

50

。

【 0 0 2 6 】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、感染サンプル配列データが、完全ゲノム配列データ又は標的配列データのソースから受け取られる。

【 0 0 2 7 】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、サンプル間の突然変異変化の計算された数が、両方のサンプルがベースコールを持つゲノム位置の総数で割ることにより正規化される。

【 0 0 2 8 】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、プロセッサが、上記変化の予想範囲において複数の変化を伴うエッジを横切る上記サンプルの系統樹における幅優先探索を行うことにより、上記感染サンプルのどれが同じアウトブレイクに由来するかを決定するよう更に構成される。

【 0 0 2 9 】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、プロセッサが、配列決定された感染サンプルの表示された系統樹において、潜在的な感染であると識別された上記感染サンプルの対をマーキングするよう更に構成される。

【 0 0 3 0 】

様々な実施形態において、上記のような感染の広がりを追跡するシステムが開示され、そこでは、上記プロセッサが、複数の系統樹を作成するよう更に構成され、各樹は、同じアウトブレイクからのものとしてマークされたサンプルから構築される。

【 0 0 3 1 】

上記の説明を読むと、様々な代替の実施形態が当業者に明らかになるであろう。これらの実施形態は、本開示の主題の範囲及び趣旨内にあるとみなされるべきである。

【図面の簡単な説明】

【 0 0 3 2 】

【図 1 A】感染を追跡する方法の実施形態の概略図である。

【図 1 B】感染を追跡する方法の実施形態の概略図である。

【図 2】感染を追跡するシステムの実施形態の概略図である。

【発明を実施するための形態】

【 0 0 3 3 】

図面において、同様の参照文字は一般に、異なる図を通して同じ部分を指す。以下の説明では、様々な実施形態が、以下の図面を参照して説明される。

【 0 0 3 4 】

本書に提示される説明及び図面は、様々な原理を説明する。本書に明示的に記載又は図示していないが、これらの原理を具現化し、本開示の範囲内に含まれる様々な構成を当業者が考案できる点を理解されたい。本書で使用する「又は」という用語は、別段の指示がない限り（例えば、「又は他の」又は「又はその代わりに」）、非排他的な又は（即ち及び/又は）を意味する。更に、本書に記載される様々な実施形態は、必ずしも相互に排他的ではなく、本書に記載される原理を組み込む追加の実施形態を生成するために組み合わせられることができる。

【 0 0 3 5 】

様々な実施形態の本開示は、上述の必要性を満たす。実施形態は一般に、医療関連感染症、より詳細には、院内感染症を追跡する方法及び装置に関する。

【 0 0 3 6 】

図 1 を参照すると、様々な実施形態による院内感染を追跡する方法の概略図 1 0 0 が示される。ステップ 1 1 0 において、複数の感染サンプルにおける各感染サンプルが配列決定される。配列決定は、例えば I l l u m i n a H i S e q 又は M i S e q 又は P a c i

10

20

30

40

50

f i c B i o s c i e n c e s などの次世代シーケンシング技術といった様々な配列決定技術を用いて行われることができる。実施形態は、生物体の完全ゲノム配列決定又は標的配列決定からのデータを利用することができる。

#### 【 0 0 3 7 】

配列決定データは処理され、配列決定された生物に関する1つ又は複数の参照配列と比較される。適切な参照配列は例えば、事前知識又はMLSTタイピングを用いて識別されることができる。参照配列が選択されると、配列決定されたデータは、BWA、Bowtieなどの既知の方法を使用して参照配列に整列され得、配列決定されたデータと参照配列との間の相違は例えば、samtools又はGATKといった変種を呼び出すための様々な方法により識別されることができる。ゲノムにおける変異は、重要な遺伝子において又はゲノムのあまり重要でない領域においてどの突然変異が生じるかを決定するために注釈を付けられることができる。

10

#### 【 0 0 3 8 】

ステップ120では、配列決定されたサンプルの各対の間のペアワイズ距離が、例えばR統計的モデリング言語及びAPEのようなRパッケージ(<http://cran.r-project.org/web/packages/ape/index.html>で入手可能)を使用して計算され、距離行列を形成するために使用される。距離行列は、差の絶対数に基づかれることができ、又はJukes-Cantorのような確率モデルに基づかれることができる。次いで、この距離マトリクスは、配列決定されたサンプルの系統樹を構築し、配列決定された各サンプルに関する任意の密接に関連付けられるサンプルを識別するために使用され得る。

20

#### 【 0 0 3 9 】

ステップ130において、伝染される感染の経時的な変化の予想範囲が決定される。経時的な変化の予想範囲を決定するための1つの手順は、系統樹が密接に関連していることを示す近隣のものの、即ちそれらの間に限られた数の変化(例えば10個のSNP)しかなく、進化論的に関連付けられると考えられるものを調べ、それらのゲノム間の変化の範囲を測定し、それらの起源の感染の隔離の間の時間で割ることを含む。単位時間あたりの平均変化範囲を決定するため、線形回帰がサンプルデータに使用されることもできる。経時的な変化の予想範囲を決定するための代替的な手順は、例えば、既知の時間にわたり同じ患者からサンプルを取るといった制御された方法を利用し、第三者ソース若しくは出版物から事前に決定された情報を使用して観察された変化を測定すること、又は既知の院内感染から採取される複数のサンプルから変化の予想範囲を決定することを含むことができる。時間にわたり見られると変化の予想範囲に関する予測区間を計算するのに、他の方法が使用されることができ、これらの区間は、予想される変化の典型的な範囲の境界を示すために使用されることができる。

30

#### 【 0 0 4 0 】

ステップ140において、少なくとも1つの感染サンプルについて、そのサンプルと少なくとも1つの追加の感染サンプルとの間の経時的な突然変異変化の数が計算され、これらのサンプル対の間の経時的な変化の数が、変化の予想範囲の区間内にあるかないかが決定される。

40

#### 【 0 0 4 1 】

例えば、サンプルの各対の間のペアワイズ距離は、サンプル間のSNP差の数に基づき計算されることができる。突然変異変化の数は例えば、インデル(挿入及び欠失)、ゲノム再編成(転位及び転座)、コピー数の変化、遺伝子の不在又は存在、又はこれらの特徴のいくつかの組み合わせを計数することにより決定されることもできる。これらの変化は、生物体の完全ゲノム、又は生物体の染色体若しくはプラスミドといったゲノムの一部において測定されることができる。

#### 【 0 0 4 2 】

ペアワイズ距離が計算されると、サンプル間の時間差を考慮して、サンプル対間の差の数が、伝染された感染から予想される差の範囲内にあるかどうかを決定するため、各サン

50

ブルが、他のすべてのサンプルと比較される。サンプル間の突然変異変化の計算された数は、両方のサンプルがベースコールを持つゲノム位置の総数を用いて正規化されることができ、これにより、サンプル配列決定プロセスの物理的形質導入において生じる誤差及びノイズが低減される。

#### 【0043】

ステップ150において、経時的な変化の計算された数が変化の予想範囲の区間内にあるサンプルの各対について、関連付けられるサンプル対が潜在的に伝染した感染としてマークされる。例えば、変化の数が、以前のデータに基づき予測される変化の数の何らかの予測区間、例えば95%以内である場合、サンプルは、潜在的に伝染した感染としてマークされる。

10

#### 【0044】

様々な実施形態では、上記の伝染した感染を識別する方法が、感染サンプルのどれが同じアウトブレイクに由来するかを決定するステップ160と共に開示される。これは、あるグラフ上で幅優先探索を行うことにより達成されることができる。そのグラフでは、各患者がノードであり、前述のステップが患者間で伝染した可能性があるとして決定する場合、2つのノード間にエッジが存在する。斯かる探索では、1人の患者から始めて、その患者から感染した可能性が高いとマークされる他の患者に注目して、同じアウトブレイク内の患者を本質的に特定する。このプロセスは、元の患者からの感染を受けた患者からの感染を受けた可能性の高い患者を新しい患者が識別されなくなるまで識別することにより、繰り返される。

20

#### 【0045】

様々な実施形態では、上記のような伝染した感染を識別する方法が、感染したと識別される感染サンプルを、配列決定された感染サンプルの系統樹においてマーキングするステップ170と共に開示される。系統樹は、上記の伝染決定に関連して計算される元の系統樹であってもよいし、又は上記のような伝染した感染を説明する再計算された系統樹であってもよい。伝染した感染の実際の表示は例えば、表示されるツリー上のサンプルのクラスタ又はサブツリーの強調表示とすることができる。

#### 【0046】

様々な実施形態では、上述したように伝染した感染を識別する方法が、複数の系統樹を作成するステップ180とともに開示され、各樹は、別々のアウトブレイクに関連付けられる。

30

#### 【0047】

図2を参照すると、様々な実施形態による伝染した感染を識別するシステムの概略図200が示される。システム200は、感染サンプル配列データ204を受信する受信機202と、受信された感染サンプル配列データ204を格納するメモリ206と、プロセッサ208とを含む。

#### 【0048】

プロセッサ208は、上述したように、配列決定されたサンプルの系統発生を構築し、配列決定された各サンプルについて密接に関連付けられる任意のサンプルを識別し、伝染した感染の経時的な変化の予想範囲を決定し、上記複数の感染サンプルのそれぞれについて、そのサンプルと少なくとも1つの他の感染サンプルとの間の経時的な変化の数を計算し、経時的な変化の計算された数のそれぞれについて、時間にわたる変化の数が変化の予想範囲の区間内にあるかどうかを決定し、及び変化の予想範囲の区間内にある経時的な変化の計算された数のそれぞれに対して、関連付けられるサンプルのペアを伝染した感染としてマークする。

40

#### 【0049】

プロセッサ208は、1つ又は複数のマイクロプロセッサ、フィールドプログラマブルゲートアレイ(FPGA)、特定用途向け集積回路(ASIC)、他の同様のデバイス、又はそれらの組み合わせを含むことができる。メモリ206は、キャッシュ(L1/L2/L3キャッシュ)、システムメモリ、又はストレージなどの様々なメモリデバイスを含む

50



ことができる。本書に使用される「非一時的機械可読媒体」という用語は、揮発性（例えばDRAM及びSRAM）及び不揮発性（例えば、フラッシュ、磁気及び光学）メモリの両方を含むが、一時的信号を排除するものとして理解されたい。

【0050】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、構成されたプロセッサが、複数の感染サンプルの少なくとも1つに対して、サンプルと系統発生により決定される密接に関連した少なくとも1つのサンプルとの間の経時的な変化の数を計算することにより、変化の予想範囲を決定する。

【0051】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、構成されたプロセッサが、時間にわたり同じ患者から、又は伝染の可能性のある2人の患者から採取される複数のサンプルから、変化の予想範囲を決定する。

10

【0052】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、構成されたプロセッサが、配列決定されたサンプルの各対の間のペアワイズ距離を計算し、距離行列を使用して系統樹を構築することにより、系統発生を構築する。

【0053】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、構成されたプロセッサが、異なるサンプルについての経時的変化を含むサンプルデータに線形回帰を使用し、予測区間を計算することにより、病原性生物についての経時的な変化の予想範囲を決定する。

20

【0054】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、感染サンプル配列データが、完全ゲノム配列データ又は標的配列データのソースから受け取られる。

【0055】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、サンプル間の突然変異変化の計算された数が、両方のサンプルがベースコールを持つゲノム位置の総数で割ることにより正規化される。

【0056】

30

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、プロセッサが、上記変化の予想範囲において複数の変化を伴うエッジを横切る上記サンプルの系統樹において幅優先探索を行うことにより、上記感染サンプルのどれが同じアウトブレイクに由来するかを決定するよう更に構成される。

【0057】

様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、プロセッサが、配列された感染サンプルの表示された系統樹において、潜在的な感染であると識別された上記感染サンプルの対をマークするよう更に構成される。

【0058】

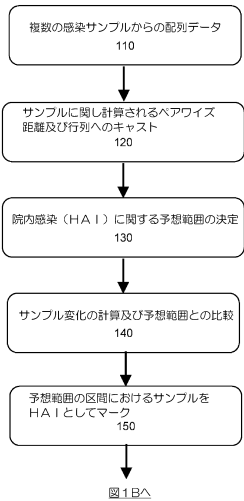
様々な実施形態では、上記のような伝染した感染を識別するシステムが開示され、そこでは、上記プロセッサが、複数の系統樹を作成するよう更に構成され、各樹は、同じアウトブレイクからのものとしてマークされたサンプルから構築される。

40

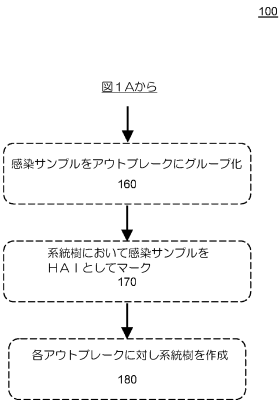
【0059】

実施形態が特定の程度で説明されてきたが、本開示の主旨及び範囲から逸脱することなく要素の構成及び配置の詳細に多くの変更が加えられることができることは明らかである。本発明は、例示のために本書に記載された実施形態に限定されない点を理解されたい。

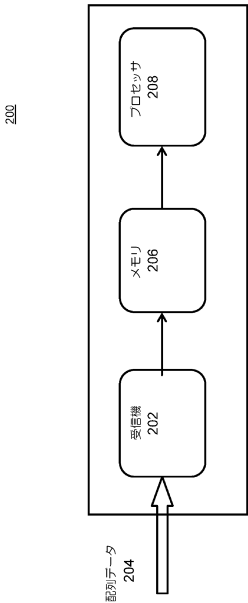
【図 1 A】



【図 1 B】



【図 2】



---

フロントページの続き

(72)発明者 リン ヘンリー

オランダ国 5 6 5 6 アーエー アインドーフェン ハイ テック キャンパス ビルディング  
5

(72)発明者 カマラカラン シサーサン

オランダ国 5 6 5 6 アーエー アインドーフェン ハイ テック キャンパス ビルディング  
5

審査官 梅岡 信幸

(56)参考文献 米国特許出願公開第 2 0 0 8 / 0 2 0 6 7 6 7 ( U S , A 1 )

(58)調査した分野(Int.Cl. , D B 名)

G 1 6 B 5 / 0 0 - 9 9 / 0 0

G 1 6 H 1 0 / 0 0 - 8 0 / 0 0

C 1 2 Q 1 / 0 0 - 3 / 0 0