



(12) 发明专利

(10) 授权公告号 CN 106663148 B

(45) 授权公告日 2021.03.26

(21) 申请号 201580044401.X

(22) 申请日 2015.08.11

(65) 同一申请的已公布的文献号
申请公布号 CN 106663148 A

(43) 申请公布日 2017.05.10

(30) 优先权数据
62/037,413 2014.08.14 US

(85) PCT国际申请进入国家阶段日
2017.02.14

(86) PCT国际申请的申请数据
PCT/IB2015/056093 2015.08.11

(87) PCT国际申请的公布数据
W02016/024213 EN 2016.02.18

(73) 专利权人 皇家飞利浦有限公司
地址 荷兰艾恩德霍芬

(72) 发明人 S·卡玛拉卡兰 J·N·保尔森

(74) 专利代理机构 永新专利商标代理有限公司
72002

代理人 李光颖 王英

(51) Int.Cl.
G16B 20/00 (2019.01)
G16B 30/00 (2019.01)

(56) 对比文件
US 2002120408 A1, 2002.08.29
CN 1479220 A, 2004.03.03
US 2011246084 A1, 2011.10.06
US 2010169122 A1, 2010.07.01
CN 1304042 A, 2001.07.18
CN 103571972 A, 2014.02.12
CN 103858009 A, 2014.06.11

审查员 王青

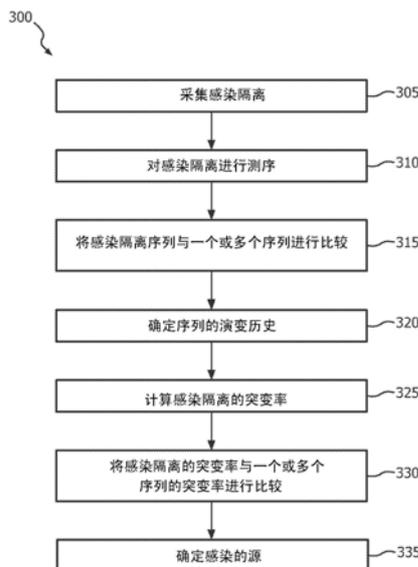
权利要求书2页 说明书8页 附图7页

(54) 发明名称

用于跟踪和识别感染传送的系统和方法

(57) 摘要

本公开描述了用于确定感染传送的源的系统和方法。种系发生学方法被用于确定感染隔离的演变历史和复制速率。感染隔离的演变距离和/或复制速率可以与其他隔离进行比较。基于对演变距离和/或复制速率的比较,进行对感染传送的源的确定的。



1. 一种用于识别感染隔离的方法,包括:
 - 接收表示感染隔离序列的电子数据;
 - 将表示所述感染隔离序列的所述电子数据与被存储在数据库中的至少一个参考序列进行比较,以确定所述感染隔离序列与所述至少一个参考序列之间的变异;
 - 至少部分地基于所述变异来确定感染隔离的演变历史的至少部分;
 - 至少部分地基于所述演变历史的所述部分来计算所述感染隔离的突变率;
 - 将所述感染隔离的所述突变率与所述至少一个参考序列的突变率进行比较;
 - 至少部分地基于所述感染隔离的所述突变率与所述至少一个参考序列的所述突变率之间的差异来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的;并且
 - 提供对所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的指示。
2. 根据权利要求1所述的方法,其中,提供的所述指示用于至少部分地基于对所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定来对所述第一存储库或所述第二存储库进行隔离检疫。
3. 根据权利要求1所述的方法,其中,所述指示是在电子显示器上的文本。
4. 根据权利要求1所述的方法,其中,确定所述演变历史的至少部分包括计算演变距离。
5. 根据权利要求4所述的方法,其中,所述演变距离是通过Jukes-Cantor方法计算的。
6. 根据权利要求1所述的方法,其中,所述突变率是通过平均路径长度方法计算的。
7. 一种用于识别感染隔离的方法,包括:
 - 利用至少一个处理单元将被存储在能由所述至少一个处理单元访问的存储器中的感染隔离的序列与被存储在能由所述至少一个存储单元访问的数据库中的至少一个参考序列进行比较,以确定感染隔离序列与所述至少一个参考序列之间的变异;
 - 利用所述至少一个处理单元至少部分地基于所述变异来确定所述感染隔离的复制速率;
 - 利用所述至少一个处理单元将所述感染隔离的所述复制速率与所述至少一个参考序列的复制速率进行比较,以至少部分地基于所述感染隔离的所述复制速率与所述至少一个参考序列的所述复制速率的差异来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的;并且
 - 利用显示器向用户提供对所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定。
8. 根据权利要求7所述的方法,其中,所述第一存储库是活体有机体并且所述第二存储库是非活体域。
9. 根据权利要求8所述的方法,还包括向所述用户提供通知,以在所述确定是所述感染隔离为由所述非活体域传送的情况下对装备的件进行消毒。
10. 根据权利要求8所述的方法,还包括向所述用户提供通知,以在所述确定是所述感染隔离为由所述非活体域传送的情况下针对所述感染隔离来测试装备的件。
11. 根据权利要求8所述的方法,还包括向所述用户提供通知,以在所述确定是所述感染隔离为由所述活体有机体传送的情况下针对所述感染隔离来测试患者。
12. 根据权利要求8所述的方法,还包括向所述用户提供通知,以在所述确定是所述感

染隔离为由所述活体有机体传送的情况下隔离检疫患者。

13. 根据权利要求7所述的方法,还包括生成所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定的报告,以用于传送到感染跟踪代理。

14. 一种用于识别感染隔离的系统,包括:

处理单元;

存储器,其能由所述处理单元访问;

数据库,其能由所述处理单元访问;以及

显示器,其被耦合到所述处理单元;

其中,所述处理单元被配置为:

将被存储在所述存储器中的感染隔离的序列与被存储在所述数据库中的至少一个参考序列的进行比较,以确定感染隔离序列与所述至少一个参考序列之间的变异;

至少部分地基于所述变异来确定所述感染隔离的复制速率;

将所述感染隔离的所述复制速率与所述至少一个参考序列的复制速率进行比较,以至少部分地基于所述感染隔离的所述复制速率与所述至少一个参考序列的所述复制速率的差异来确定所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的;并且

向所述显示器提供对所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定。

15. 根据权利要求14所述的系统,还包括能由所述处理单元访问的计算机系统,其中,所述处理单元被配置为提供所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定。

16. 根据权利要求14所述的系统,还包括测序单元,所述测序单元被配置为将所述感染隔离的所述序列提供到所述存储器。

17. 根据权利要求14所述的系统,其中,所述处理单元包括多个处理单元。

用于跟踪和识别感染传送的系统和方法

背景技术

[0001] 医疗保健相关联的感染 (HAI) 是在针对与感染不相关的状况的医疗保健处置期间接收到的患者受到的感染。入院进行手术流程以修复疝气而随后在康复病房中在手术位点处出现葡萄糖球菌感染的健康患者是患者受到HAI的范例。HAI在医学文献中常常被称为医院内感染。根据在2011年由CDC实行的手术,住院的每25位患者中大约有1位将染上HAI。研究估计存在大约721000例HAI。HAI引起或者促成每年大约75000例死亡。

[0002] 医院内感染能够引起重度肺炎以及尿路、血流和身体的其他部分的感染。许多类型难以利用抗生素治疗,并且抗生素耐药性正扩散至会感染医院外部的人的革兰氏阴性菌。在USA,最常见类型的医院范围感染是肺炎(21.8%),紧接着是手术位点感染(21.8%)以及胃肠道感染(17.1%)。(Magill SS、Edwards JR、Bamberg W等人的“Multistate Point-Prevalence Survey of Health Care-Associated Infections”,N Engl J Med 2014;370:1198-208。)

[0003] 根据CDC的2009年度报告,HAI每年花费美国医院大约350亿美元。许多花费与更长的患者停留、隔离医院的部分、以及对感染源的发现和根除有关。HAI中的大约25.6%被认为是由诸如导管和呼吸机的医学设备所引起的。其余的感染被认为与医院内的手术流程和其他源相关联。存在这样的需求:能够跟踪医院相关联的感染并且确定传送源,以降低额外感染的数量。(Scott II、R.D.的“The Direct Medical Costs of Healthcare-Associated Infections in U.S.Hospitals and the Benefits of Prevention”,CDC,2009年3月。)

发明内容

[0004] 根据本发明的说明性实施例,一种方法,可以包括:接收表示感染隔离序列的电子数据;将表示所述感染隔离序列的所述电子数据与被存储在数据库中的至少一个参考序列进行比较,以确定在所述感染隔离序列与所述至少一个参考序列之间的变异;可以至少部分地基于所述变异来确定所述感染隔离的演变历史的至少部分;可以至少部分地基于所述演变历史的所述部分来计算所述感染隔离的种系发生学 (phylogenetic) 度量;将所述感染隔离的所述种系发生学度量与所述至少一个参考序列的种系发生学度量进行比较;可以至少部分地基于在所述感染隔离的所述种系发生学度量与所述至少一个参考序列的所述种系发生学度量之间的差异,来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的;并且提供所述感染隔离是为由第一存储库或第二存储库传送的指示。所述演变距离可以由Jukes-Cantor方法来计算。

[0005] 根据本发明的说明性实施例,一种方法,可以包括:利用至少一个处理单元将可以被存储在能由所述至少一个处理单元访问的存储器中的感染隔离的序列与被存储在能由所述至少一个处理单元访问的数据库中的至少一个参考序列进行比较,以确定感染隔离序列与所述至少一个参考序列之间的变异;可以至少部分地基于所述变异利用所述至少一个处理单元来确定所述感染隔离的复制速率;利用所述至少一个处理单元将所述感染隔离的所述复制速率与所述至少一个参考序列的复制速率进行比较,以可以至少部分地基于所述

感染隔离的复制速率与所述至少一个参考序列的复制速率的差异来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的；并且利用显示器向用户提供对所述感染隔离是否是由第一存储库还是由第二存储库传送的所述确定。所述第一存储库可以是活体有机体并且第二存储库可以是非活体域。

[0006] 根据本发明的说明性实施例，一种方法，可以包括：利用至少一个处理单元将可以被存储在能由所述至少一个处理单元访问的存储器中的感染隔离的序列与被存储在能由所述至少一个处理单元访问的数据库中的至少一个参考序列进行比较，以确定在所述感染隔离序列与所述至少一个参考序列之间的变异；可以至少部分地基于所述变异利用所述至少一个处理单元来确定所述感染隔离距所述至少一个参考序列的演变距离；利用所述至少一个处理单元将所述感染隔离的所述演变距离与被存储在所述数据库中的多个序列距所述至少一个参考序列的演变距离的分布进行比较，以可以至少部分地基于所述感染隔离的所述演变距离与所述多个序列的演变距离的所述分布之间的差异来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的；并且利用显示器向用户提供所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的确定。对所述感染隔离是否是由第一存储库还是由第二存储库传送的所述确定可以至少部分地基于所述感染隔离的所述演变距离是否落在多个序列的演变距离的分布的期望置信区间内。

[0007] 根据本发明的说明性实施例，一种系统，可以包括：处理单元、能由所述处理单元访问的存储器、能由所述处理单元访问的数据库、以及被耦合到所述处理单元的显示器，其中，所述处理单元可以被配置为：将被存储在所述存储器中的感染隔离的序列与被存储在所述数据库中的至少一个参考序列进行比较，以确定在所述感染隔离与所述至少一个参考序列之间的变异；至少部分地基于所述变异来确定所述感染隔离的复制速率；将所述感染隔离的所述复制速率与所述至少一个参考序列的复制速率进行比较，以至少部分地基于所述感染隔离的所述复制速率与所述至少一个参考序列的所述复制速率的差异来确定所述感染隔离是否是由第一存储库还是由第二存储库传送的；并且向所述显示器提供所述感染隔离是否是由所述第一存储库还是由所述第二存储库传送的所述确定。所述系统还可以包括能由所述处理单元访问的计算机系统，其中，所述处理单元可以被配置为提供所述感染隔离是否是由第一存储库还是由第二存储库传送的所述确定。所述系统还可以包括测序单元，所述测序单元可以被配置为将所述感染隔离的所述序列提供到所述存储器。

附图说明

- [0008] 图1是根据本公开的实施例的传送向量的框图。
- [0009] 图2A是根据本公开的实施例的种系发生学树。
- [0010] 图2B是根据本公开的实施例的定日期的种系发生学树。
- [0011] 图3是根据本公开的实施例的方法的流程图。
- [0012] 图4是根据系统的实施例的系统的框图。
- [0013] 图5是根据本公开的实施例的生长速率曲线的图。
- [0014] 图6是根据本公开的实施例的经验性累积分布函数图。
- [0015] 图7是根据本公开的实施例的密度图。
- [0016] 图8是根据本公开的实施例的过程的流程图。

具体实施方式

[0017] 下文对特定示范性实施例的描述在本质上仅仅是示范性的,并且根本不旨在限制本发明或者其应用或使用。在下文中对本系统和方法的实施例的详细描述中对附图进行参考,所述附图形成说明书的部分,并且其中,通过在其中可以实践所描述的系统和方法的图示特定实施例的方式示出。足够详细地描述这些实施例,以使得本领域技术人员能够实践当前所公开的系统和方法,并且应当理解,可以利用其他实施例,并且可以做出结构和逻辑变化而不偏离本系统的主旨和范围。

[0018] 因此,下文的详细描述没有限制性的含义,并且仅由权利要求来定义本系统的范围。在本文中在附图中的附图标记的(一个或多个)首数字通常对应于附图编号,其中,例外情况是,由相同的附图标记来识别在多个附图中出现的相同部件。此外,出于简洁的目的,当特定特征对于本领域技术人员而言将是显而易见的时,将不讨论对这些特定特征的详细描述,从而不混淆对本系统的描述。

[0019] 可能由诸如细菌、病毒、真菌、寄生虫或其他有机体的病原体引起感染。一些感染可能由同时存在的多种类型的有机体引起。感染可能由各种不同的传送向量来扩散。在图1中图示了范例传送向量。在一些实例中,可能在两个活体有机体之间传送感染。例如,医生可能处置具有感染的第一患者。如果医生不对其手彻底地消毒,他可能将所述感染从第一患者传送到第二患者。该范例可以对应于传送向量100。受感染的有机体还可能将感染直接传送到另一有机体。例如,受感染的人可能在另一人的附近打喷嚏或咳嗽,所述另一人吸入病原体并且随后变得被感染。该范例可以对应于传送向量105。在其他实例中,感染可能从非活体域被传送到活体有机体。非活体域可能是诸如桌台或台面的表面。非活体域也可以是诸如透析机器或内窥镜的医学装备的件。例如,当患者身上的开放性伤口接触非无菌检查桌台时,患者可能接触感染。该范例被图示为传送向量110。

[0020] 确定传送向量(例如,患者如何被感染)可以允许将感染追溯回其源。然后,所述源可以被消除,以防止感染的进一步传送。例如,如果源是患者,则所述患者可以被隔离以避免感染其他患者,直到所述感染能够利用处置(例如,抗生素)来清除。在一个范例中,如果源是装备的件,则所述装备可以被消毒或者利用新的装备来替换。在医院设置中,将感染追溯回其源的能力可以降低HAI的发生率以及其相关联的成本。其还可以减少利用并非链接到HAI的消毒剂不必要地处置区的时间和成本。

[0021] 当检测到诸如HAI的感染时,可以由医学工作人员从患者、表面、食物、装备或其他怀疑的源收集样本。样本可以包括组织、血液、水、表面的分泌物。然后,可以处理所述样本以将引起所述感染的病原体与样本中的其他材料隔离。然后,可以通过各种方法来分析感染隔离。所述分析可以确定病原体类型、物类、抗药性和/或其他性质。尽管该分析可以利用确定有效的处置来辅助,但是其可能不足以识别传送向量。

[0022] 通过收集来自受感染有机体或非活体域的感染隔离并且分析其基因序列,可以确定通过使用种系发生学方法的感染的传送向量。感染隔离是样本的包括来自病原体的基因信息的分量。种系发生学是对有机体之间的演变关系的研究。这样的关系常常被表示为诸如树的加权图。在图2A-B中示出了种系发生学树的范例。种系发生学方法分析有机体的基因序列的所有或部分。通过确定感染的演变历史,感染可以被追溯回其源。所述演变历史可以提供对感染的不同事件如何相关或不相关的理解。例如,可以将来自多位受感染患者的

感染隔离的序列进行比较。可以确定患者被感染时的时间。还可以确定,所述患者中的一位或多位患者被不同的源感染。

[0023] 存在多种种系发生学方法,包括基于演变距离、简约和最大似然的方法。基于距离的方法是其中在每个有机体之间计算演变距离。基于在有机体的基因序列之间的相似性的程度来计算所述演变距离。用于确定演变距离的一种这样的方法被称为Jukes-Cantor (由M.N.Munro编辑的T.H.Jukes,C.R.Cantor的Evolution of protein molecules In Mammalian protein metabolism,第III卷、第21-132页(1969))方法,其中,从基因中的任意特定核苷酸到另一核苷酸的转变,即转换或颠换,能够以相同的概率发生:

$$[0024] \quad Q = \begin{bmatrix} -\frac{3\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & -\frac{3\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & -\frac{3\mu}{4} & \frac{\mu}{4} \\ \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & -\frac{3\mu}{4} \end{bmatrix} \quad \text{等式 1}$$

[0025] 在以上等式1中,瞬时率矩阵Q表示每时刻在核苷酸对之间的变化率。P—概率转换矩阵被给出为:

$$[0026] \quad p(t) = e^{Qt} \quad \text{等式 2}$$

[0027] 结果,在该模型下的在任意两个有机体之间的演变距离是简单的:

$$[0028] \quad d_{ab} = -\frac{3}{4} \ln \left(1 - \frac{4}{3} p \right) \quad \text{等式 3}$$

[0029] 其中,p是在所述序列之间不同的沿着单个核苷酸多态性(SNP)/DNA的位点的数量。所述距离转到无限,因为p接近平衡值(位点的75%不同)。然而,该简单模型不考虑生物学因素:转换(嘌呤到嘌呤(a-g)或者嘧啶到嘧啶(t-c))或颠换(嘌呤到嘧啶或者反之亦然)以不同的比例发生。另一距离模型,Kimura 2参数模型(Kimura,Motoo.的“A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences”,Journal of molecular evolution 16.2(1980):111-120)试图对此进行校正。在这种情况下:

$$[0030] \quad d = -\frac{1}{2} \ln \left[(1 - 2p - q) (\sqrt{1 - 2q}) \right] \quad \text{等式 4}$$

[0031] 针对p(转换的比例)和q(颠换的比例)。

[0032] 一旦感染隔离序列已经被比较以确定其演变距离,则可以确定演变的比率。然后,在来自样本的感染隔离之间的演变距离和关系能够以诸如树图的图形形式来进行绘制。邻接法(Saitou N,Nei M.的“The neighbor-joining method:a new method for reconstructing phylogenetic trees”,Molecular Biology and Evolution,第4卷、第4期、第406-425页(1987年7月))是建立无根树的一种方法。所述方法通过首先找到具有相同父节点k的相邻叶i和j的对来校正序列之间的不相等的演变比率。亦即,叶i和j可以从公共病原体k演变而来的病原体。叶i和j然后可以被从叶节点的列表中移除,并且k被添加到节点的当前列表,并且重新计算节点距离。该算法是贪婪“最小演变”算法的范例。

[0033] 使用许多感染隔离序列样本的树形表示,可以估计一种感染对于另一种感染的相对定时。不失一般性地,被称为平均路径长度(MPL)的方法可以被使用(Britton,Tom等人的“Phylogenetic dating with confidence intervals using mean path lengths”,

Molecular phylogenetics and evolution 24.1 (2002) :58-65)。MPL方法利用从该节点到其下游的所有叶的距离的均值来估计节点的年龄(age)。在假设相似分子钟亦即,演变的比率的情况下,能够计算估计的节点年龄的标准差。使用该方法,可以针对不同的感染隔离来计算突变率。

[0034] 感染的源可以基于复制速率来确定。病原体可以在不同的环境中以不同的速率复制。可能具有以不同速率复制的病原体的存储库的不同环境可以包括,但不限于:血液、唾液、食物、手术台、水槽、厕所和床上用品。复制速率可以与突变率相关。复制速率、总复制、突变率、总核苷酸改变以及表示种系发生学关系、种系发生学状态的其他度量,样本的再现或复制历史可以被用于在本文中所描述的度量的附加或者作为对在本文中所描述的度量的替换。这样一来,任何种系发生学度量、量度、指示器或预测器处在本发明的范围之内,并且在下文被称为种系发生学度量。通常,突变率针对以较低速率复制的病原体而言能够是较低的。如果病原体的复制速率对于一环境是已知的,则其可以基于所计算的突变率来确定源自该环境的感染隔离。例如,病原体在诸如血液的流体环境中比在诸如家具表面的干燥环境中可以更快速地复制。如果从患者获取的感染隔离与参考感染隔离相比呈现高突变率,则可以确定该患者从另一活体有机体被传送感染。该确定可以允许医院工作人员将其对感染牵制的努力聚焦于其他患者而非针对污染来测试装备。这可以允许医院工作人员减少确定感染源的时间,并且消除对装备不必要消毒的成本。

[0035] 图3图示了根据本公开的实施例的确定感染的源的方法300的流程图。医学工作人员可以在步骤305处首先收集来自不同环境的样本,以采集感染隔离或多个感染隔离。样本可以被例行地收集或者可以响应于识别的HAI来收集。所述感染隔离可以在步骤310处根据在本领域中已知的测序技术来处理和测序。提供测序技术的公司的范例包括454生命科学,罗氏公司和太平洋生物科学公司。在本领域中已知的测序技术可以允许测定感染隔离的整个基因组。医院可能在现场具有测序技术,或者医院可以将样本发送到分离的测序公司。序列的数字表示可以被生成并且被存储在一个或多个处理单元能访问的存储器中,以允许对序列进行分析。除非另外指明,将假设基于基因序列的任何比较、分析或确定是由一个或多个处理单元执行的。然后,所述感染隔离序列可以由处理单元在步骤315处与一个或多个序列进行比较。其他序列可以来自其他收集的感染隔离,来自公共或私人数据库的已知病原体的参考序列,和/或来自其他源的序列。所述比较可以包括确定在感染隔离序列与一个或多个其他序列之间的变异。在下文中,变异或基因变异涉及单核苷酸差异或以基因水平差异。变异可以使用诸如BWA-samtools和Golden Helix的现有软件工具找到。

[0036] 这些变异可以被用在步骤320处以确定与所述一个或多个序列有关的感染隔离序列的演变历史。例如,可以假设给定感染因子跨所有传送域、活体或非活体以相似的比率呈现单核苷酸突变。结果,单核苷酸改变的累积数量可以被用于估计序列复制发生的总数量,即,世代的总数量。这样一来,能够做出进一步的推理:在演变历史中的变异或者世代的总数量可以归于传送域的类型。所述演变历史可以通过上文所描述的方法中的一种方法或另一种方法来确定。基于所述演变历史,在步骤325处计算感染隔离的突变率。所述突变率可以通过上文所描述的方法中的一种方法或另一种方法来计算。

[0037] 根据一个说明性实施例,所述感染隔离的所述突变率在步骤330处与一个或多个其他序列的突变率进行比较。基于对突变率的比较,在步骤335中进行对感染的源的确定。

可以通过突变率来对感染隔离序列和其他序列进行聚类,并且每个聚类被分配给不同感染源。

[0038] 根据另一说明性实施例,所述感染隔离的所述复制速率在步骤330处与一个或多个其他序列的复制速率进行比较。基于对复制速率的比较,在步骤335中进行对感染的源的确定。可以通过复制速率来对感染隔离序列和其他序列进行聚类,并且每个聚类被分配给不同感染源。所述感染隔离序列可以与来自不同已知环境的参考序列复制速率进行比较并且被分配源环境,如将在下文中更为详细地描述的。可以执行确定感染源或者感染隔离的感染源的类别的其他方法。

[0039] 根据本公开的实施例的被用于确定感染的传送的源的系统400的范例被示为在图4中的框图。数字形式的感染隔离序列可以被包括在存储405中。存储器405可以是处理单元415能访问的。处理单元415可以包括一个或多个处理单元。处理单元415可以访问包括一个或多个序列的数据库410。处理单元415可以将其确定的结果提供给显示器820和/或数据库410。显示器420可以是用户可见的电子显示器。任选地,处理单元415还可以访问计算机系统425。计算机系统425可以包括额外的数据库、存储器和/或处理单元。计算机系统425可以是系统400的部分或者由系统400远程地访问。在一些实施例中,系统400还可以包括测序单元430。测序单元430可以处理所述感染隔离以生成序列并产生感染隔离序列的数字形式。

[0040] 图2A-B图示了根据本公开的实施例的范例。在该范例中,可以从处在相似或不同环境中的各种存储库收集10个HAI样本,并且可以对针对感染隔离的整个基因组进行测序。可以确定变异并且可以计算演变距离。基于这些演变距离,可以绘制原始树200,如图2A中所示的。在该范例中,原始树200是基于Kimura 2参数距离估计使用邻近连接树方法来生成的。树200图示了在10个HAI之间的演变距离和关系。分支的长度与这些计算的距离成比例。例如,对应于两个不同HAI的节点11和节点12具有0.72的演变距离。如先前所描述的,这是在两个HAI之间的相似度的量度。演变距离越大,在两个HAI的序列之间的相似度越低。针对每个节点/HAI样本在时间上的相对年龄可以通过使用定日期(dating)方法(诸如上文所描述的MPL方法)基于演变距离的原始树200来计算。在图2B中示出了定日期的树205。在该范例中,在原始树100中,从第11节点到第12节点的边缘长度为0.72,并且从第12节点到节点/叶t10的边缘距离为0.99。如果特定节点的年龄是已知的,则可以估计其余的节点的年龄。在该范例中,第11节点的年龄是已知的,并且是针对所述树的所有其他节点的公共祖先。针对第11节点的MPL统计结果为1.56。假设节点11偏离两天前的最小值,则其余的节点可以通过除以 $0.78 = 1.56/2$ 被变换为在天中的最小年龄。例如,使用针对年龄的该方程,节点12偏离节点11,1.479天之前。

[0041] 使用该定日期信息,可以预期样本可以具有针对特定环境的小的突变数量。如果演变距离与其他样本或参考序列相比是高的,和/或所计算的突变率或复制速率与其他样本或参考序列相比更高,则对应于该样本的感染可以被确定为从以下环境传送得到:其允许病原体的高的复制速率。高复制速率可以允许针对突变的更大的概率和发生率,增加在不同时间点处收集的感染隔离的序列之间的差异。例如,与另一细菌序列具有大的演变距离的细菌的序列可以被确定为已经通过活体有机体传送。活体有机体可以比非活体域提供更多的营养并且支持针对病原体的更好的复制速率。因此,呈现在期望阈值之上或之下的演变距离、累积突变发生率、突变率或者其组合的感染隔离可以被确定为由特定的源传送。

[0042] 备选地或者除了上文所使用的方法之外,可以获知针对一种或多种环境的病原体的复制速率或生长速率。在图5A中示出了根据本公开的实施例的生长速率的范例绘图500。基于已知的生长速率,可以预期每世代的突变的具体数量。

[0043] 根据一个说明性实施例,参考图5B,通过将样本的序列与被存储在数据库中的过往样本的序列进行比较,可以针对在研究中的样本520确定种系发生地最接近的相邻者。一旦识别了最接近的相邻者,就可以推测差异的数量(被归一化到核苷酸基的总数量比较)。在假设针对该有机体的每复制周期的公共突变率的情况下,可以针对多个存储库(例如,在图5B中所示的存储库1、2和3)来绘制特征突变曲线。突变曲线可以基于在不同存储库(即,传送域)中的该有机体的复制速率来选取。基于在研究中的样本的隔离与其确定的最近的相邻者之间的按时间顺序排列的时间差异,针对不同的生长速率来估计累积的突变的预期数量。最接近实际的、观察到的突变的生长速率可以被识别为样本的生长速率,并且已知与该生长速率相关联的存储库可以被识别为传送的最可能的源(在图5B中针对样本520的存储库2)。

[0044] 感染隔离可以基于其估计的生长速率被拟合到生长速率曲线。例如,如果感染隔离被发现具有低的生长速率(例如,0.2)并且拟合到曲线505,则其可以被确定通过表面传送感染。如果感染隔离被发现具有中等生长速率(例如,0.5)并且拟合到曲线510,则其可以确定通过诸如厕所的湿环境传送感染。如果感染隔离被发现具有高的生长速率(例如,1.0)并且拟合到曲线515,则所述感染隔离可以被确定为已经通过活体有机体传送。

[0045] 除了上文所使用的方法之外或者对于上文所使用的方法而言备选地,可以估计感染隔离样本和/或参考序列的演变距离分布。新的感染隔离样本的演变距离可以与已知参考序列或先前分析的感染隔离样本的演变距离进行比较。经验累积分布函数(ECDF)可以使用已知的统计学方法来生成。在图6中示出了根据本公开的实施例的示范性ECDF绘图600。在该范例中,如果感染隔离被确定为落在期望的置信区间(例如,95%)之外,则所述感染隔离可以被确定已经从与参考序列的不同的源传送。在图6中所示的绘图中,可以确定具有 $x > .8$ 的任何感染隔离序列具有与参考序列相比高的突变率,并且因此,所述感染通过活体有机体传送。

[0046] 图7图示了基于演变距离根据聚类序列的本公开的实施例的不同的方法。尽管以不同的图形形式来表示,如在核密度绘图700中所示的,所述方法类似于参考图6所描述的方法。如果感染隔离序列落在参考序列的演变距离的期望置信区间之外,则所述感染隔离序列被确定为已经通过与参考序列的不同的源而传送。在绘图700中,具有0.2或更大的演变距离的感染隔离可以被确定为已经在相对年龄较低时通过活体有机体传送。

[0047] 尽管参考序列被分组为在图6和图7中所示的样本的单个聚类,但是其可以具有参考序列的多个聚类。每个聚类可以表示针对特定环境中的病原体的演变距离。针对属于每个聚类的新的感染隔离的置信区间然后可以使用已知的统计方法来计算。然后,感染的传送的源可以基于所述感染隔离序列对于哪一个聚类具有包括的最高概率来确定。

[0048] 图8图示了概述根据确定感染的传送的源的根据本公开的实施例的范例过程800A和800B的流程图。过程800A和/或800B可以被包括在图3中的方法300的步骤335中。方法300的步骤335可以包括用于确定感染的传送的源的一个或多个过程。过程800A基于感染隔离的突变率是否比期望阈值 X 更大或更小来确定感染源是否是第一存储库还是第二存储库。

过程800B基于感染隔离的演变距离是否落在置信区间X之内来确定所述感染源是否是第一存储库还是第二存储库。所述置信区间可以基于一个或多个样本序列的分布。还可以执行其他过程。尽管仅示出了两个存储库,但是本公开的实施例可以包括在任意数量的存储库之间做出确定。例如,在过程800A中,值X可以由定义突变率的范围的多个值来替换。每个范围可以与不同的存储库相关联。

[0049] 一旦通过一个或多个处理单元做出了对感染的传送的源的确定的,基于被存储在存储器中的感染隔离序列,可以通过至少一个处理单元将所述确定提供给用户、数据库和/或计算机系统。用户可以是在医院的感染控制工作人员的成员。用户可以在被耦合到处理单元的电子显示器上接收视觉指示器。然后,所述用户可以使用对感染源的所述确定来协调感染控制工作人员的努力,以采取措施来隔离和移除感染的源。例如,如果所述感染源被确定为来自活体有机体,诸如特定的患者,所述感染控制工作人员可以隔离患者,并且确保护士和医生在处置所述患者之后在接触其他患者之前使用提高的消毒方法。如果所述感染源被确定为来自非活体域,诸如支气管镜,则所述装备可以在被用于额外的患者之前被消毒并且再针对病原体进行测试。

[0050] 多个感染隔离可以被确定为全部来自不同的不相关的源。在这些实例中,感染控制工作人员可以确定消毒流程和/或感染控制协议已经在医疗保健设施处出故障。可以执行对恰当的流程的医学工作人员的进一步的分析的/或再训练,以减少感染的传送。

[0051] 所述确定和感染隔离序列可以被存储在数据库中,以用于将来的感染源确定。所述数据还可以被转移到能由额外的用户、医院或代理访问的计算机系统。这可以允许更容易的信息共享并且改进感染控制。其还可以允许对感染的更容易的报告,其可能是规定所要求的。例如,所述处理单元可以自动地生成具有所述确定的报告,并且将所述报告提供到感染疾病跟踪代理,诸如美国疾病控制中心(CDC)。

[0052] 当然,应当意识到,根据本系统、设备和方法,以上实施例或过程中的任意一个可以与一个或多个其他实施例和/或过程进行组合,或者分离,和/或在分离的设备和设备部分之中执行。

[0053] 最终,以上讨论旨在仅仅图示说明本系统,而不应当被解释为将权利要求限定到任何特定的实施例或实施例的组。因此,尽管已经参考示范性实施例具体详细地描述了本系统,但是还应当意识到,本领域技术人员可以在不背离如在下文的权利要求书中所阐述的本发明的宽和意图的泛的主旨和范围的情况下设计许多修改和备选实施例。因此,说明书和附图被认为是采取说明性的方式,而不旨在限制权利要求的范围。

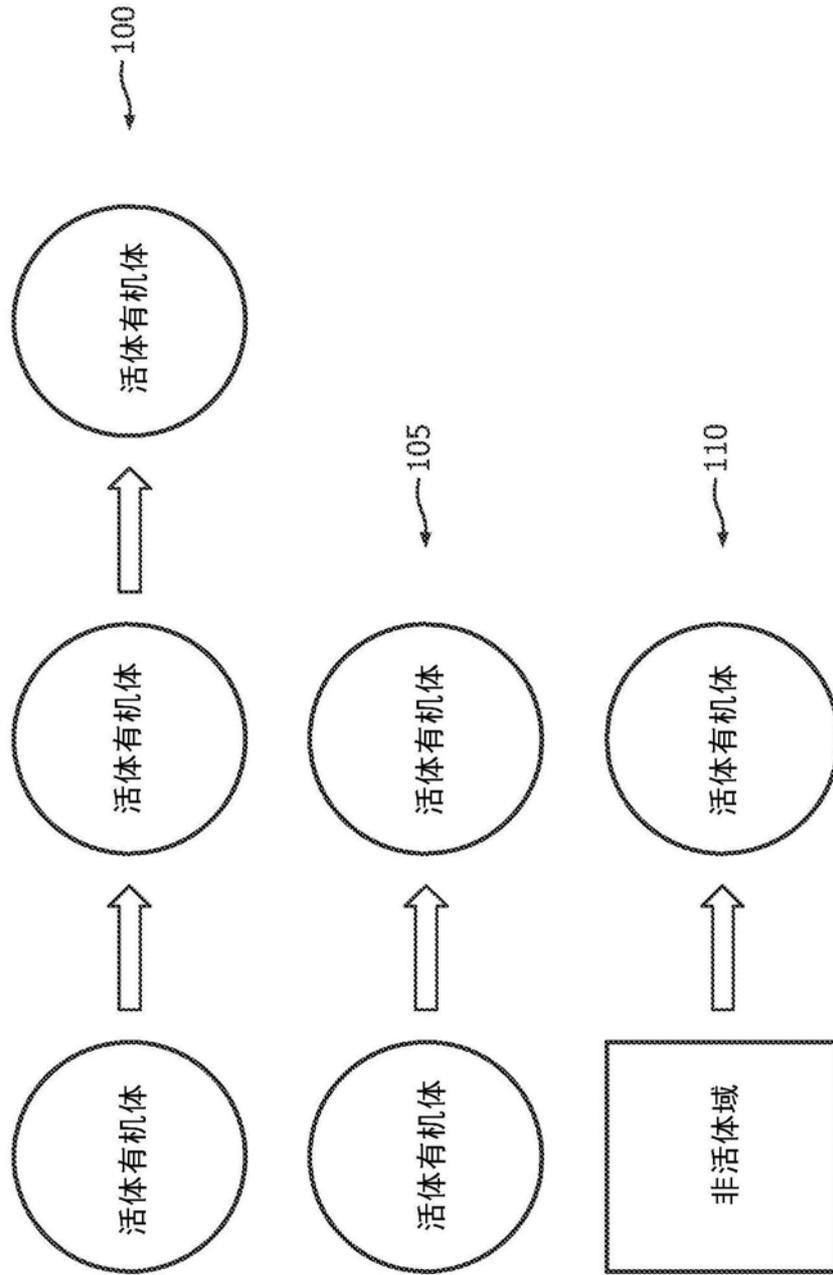


图1

原始树 - 200

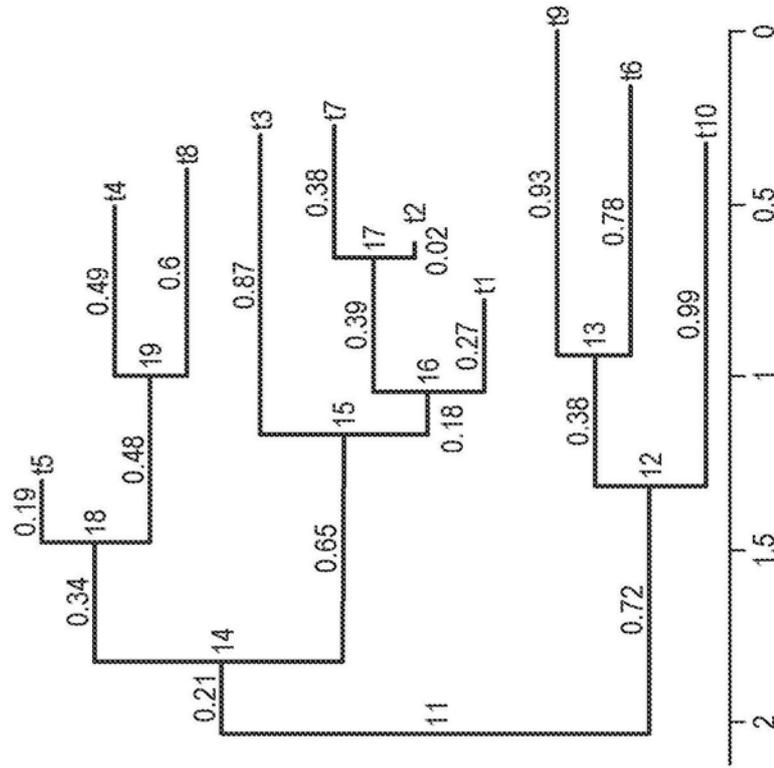


图2A

定日期的MPL树 - 205

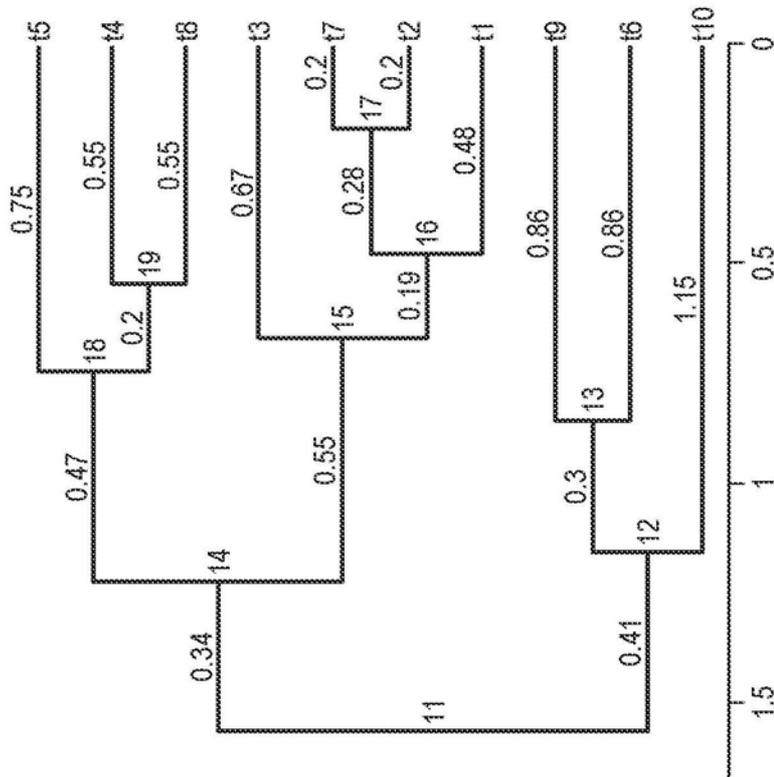


图2B

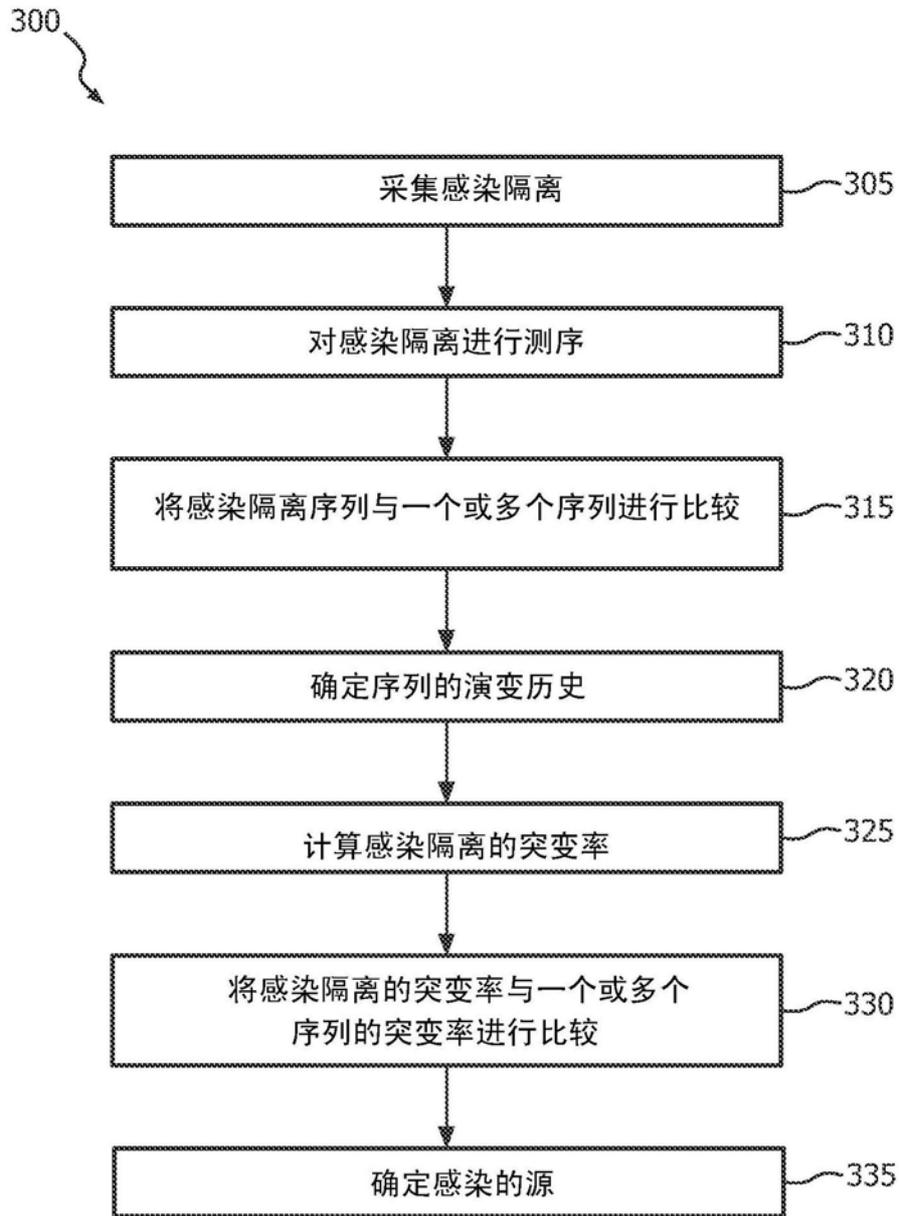


图3

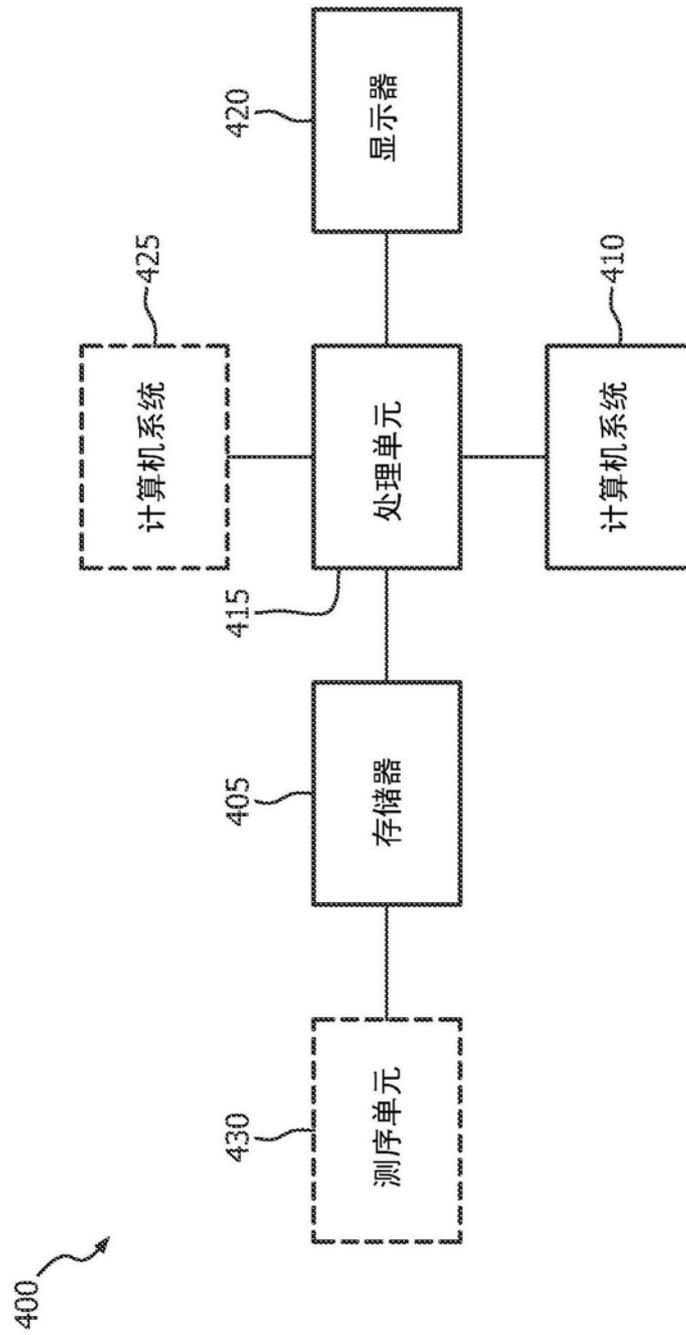


图4

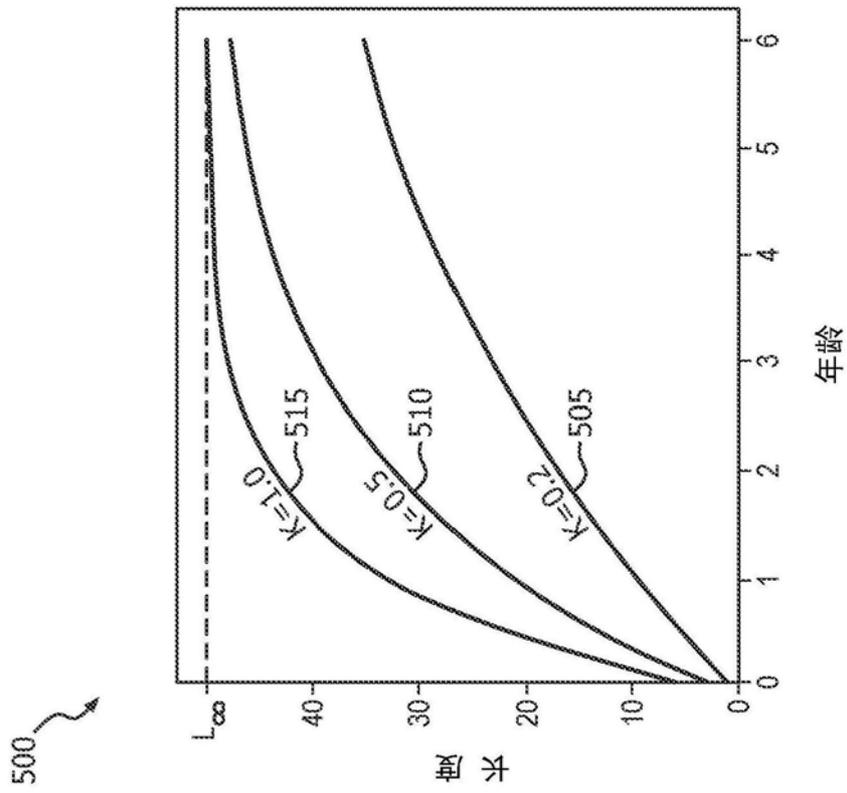


图5A

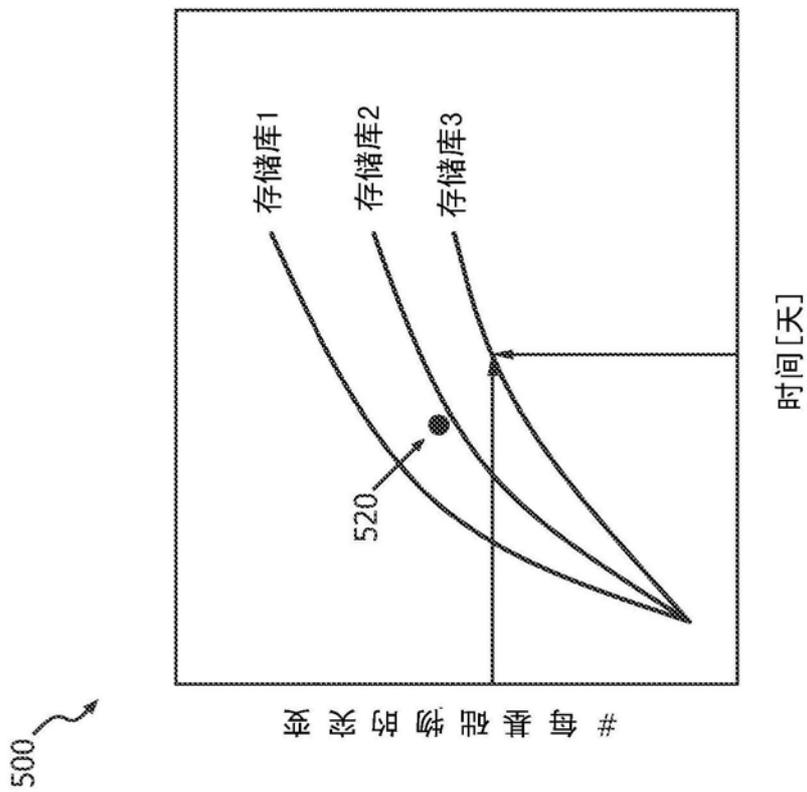


图5B

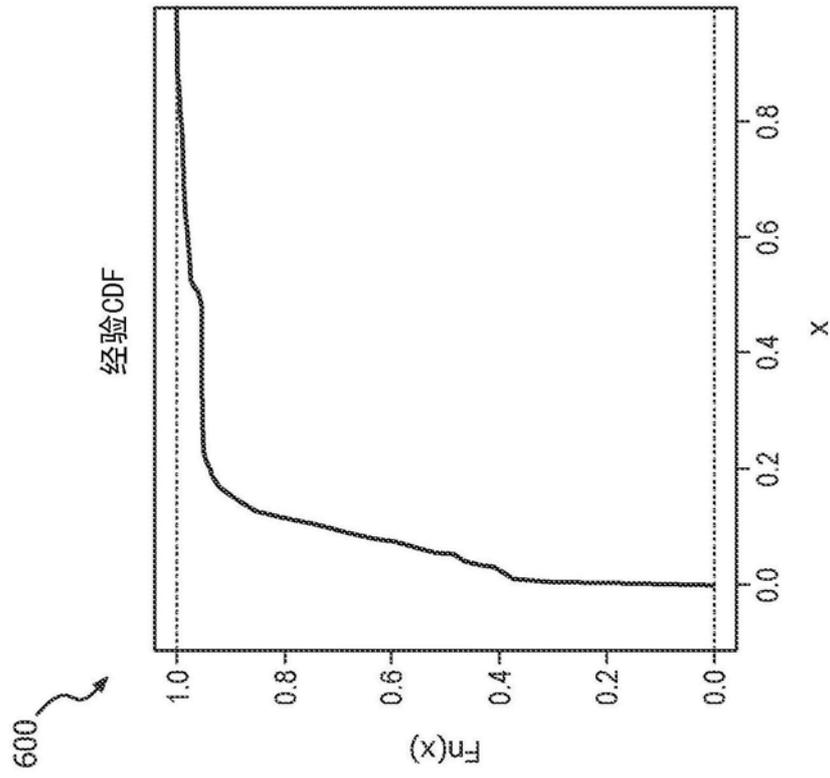


图6

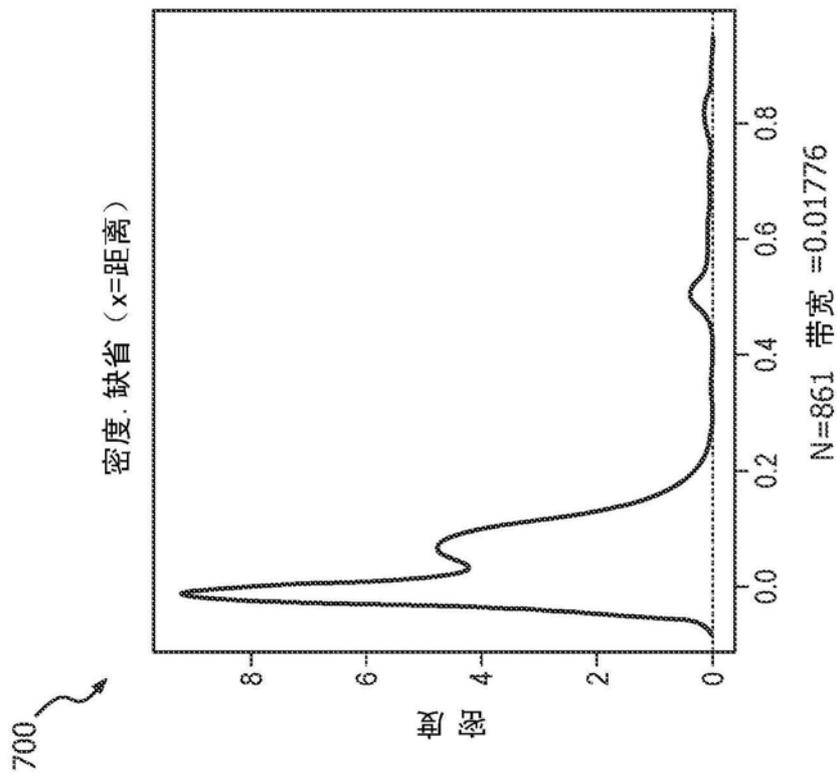


图7

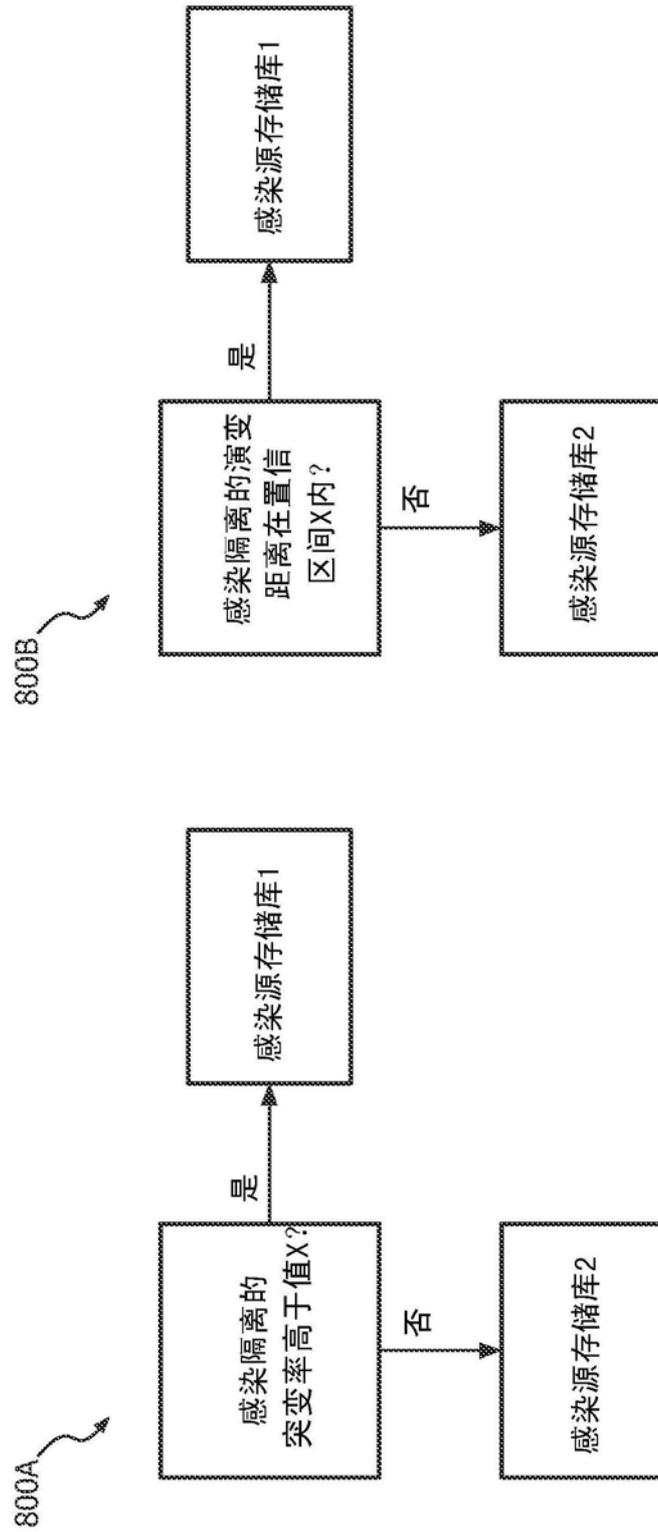


图8