



(21) 申请号 202010103794.X

(22) 申请日 2020.02.20

(65) 同一申请的已公布的文献号
申请公布号 CN 111187813 A

(43) 申请公布日 2020.05.22

(73) 专利权人 予果生物科技(北京)有限公司
地址 100000 北京市东城区和平里东街11
号37号楼213-B25号

专利权人 西咸新区予果微码生物科技有限
公司
予果智造科技(北京)有限公司

(72) 发明人 夏涵

(74) 专利代理机构 北京冠和权律师事务所
11399

代理人 陈国军

(51) Int. Cl.

C12Q 1/6869 (2018.01)

G16B 30/10 (2019.01)

(56) 对比文件

CN 108660199 A, 2018.10.16

CN 110473594 A, 2019.11.19

胡鹏.宏基因组分析和诊断技术概述.《医药
卫生科技》.2020,第1-3页.

Svanella-Dumas et al..Genome
characterization of a divergent isolate
of the mycovirus Botrytis virus F from a
grapevine metagenome.《ARCHIVES OF
VIROLOGY》.2018,第163卷(第11期),第3181-
3183页.

审查员 王震

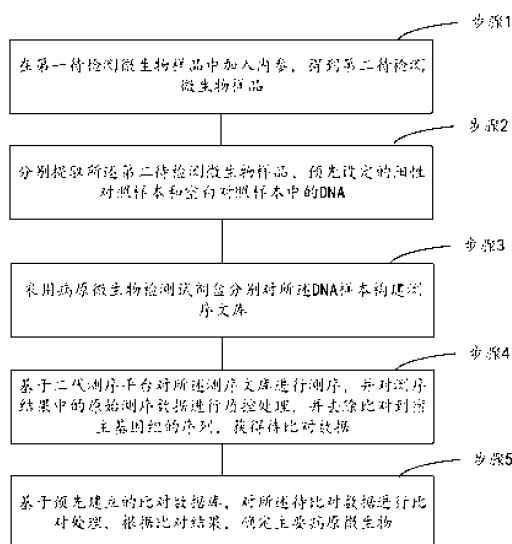
权利要求书3页 说明书9页 附图1页

(54) 发明名称

全流程质控的病原微生物高通量测序检测
方法

(57) 摘要

本发明提供了全流程质控的病原微生物高通量测序检测方法,包括:在第一待检测微生物样品中加入内参,得到第二待检测微生物样品;分别提取第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA;采用病原微生物检测试剂盒分别对提取的DNA样本构建测序文库;基于二代测序平台对测序文库进行测序,并对测序结果中的原始测序数据进行质控处理,并去除比对到宿主基因组的序列,获得待比对数据;基于预先建立的比对数据库,对待比对数据进行比对,根据比对结果,确定主要病原微生物,实现针对感染性疾病的无需预知、无偏好、高通量的病原微生物的快速临床检测,同时实现系统的全流程质控,并降低假阳性。



1.用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,包括:

在第一待检测微生物样品中加入内参,得到第二待检测微生物样品;

分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA;

采用病原微生物检测试剂盒分别对所述DNA样本构建测序文库;

基于二代测序平台对所述测序文库进行测序,并对测序结果中的原始测序数据进行质控处理,并去除比对到宿主基因组的序列,获得待比对数据;

基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物;

基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物时,包括:

步骤A1:利用公式(1)预先建立的比对数据库 S_i ;

$$S_i = \frac{1}{n+m} \sum_{i=1}^n \sum_{j=1}^m \left(\frac{\exp(\frac{f_{ij+1}}{D_i} - \frac{f_{ij}}{D_i})}{\frac{f_{ij}}{D_i}} \right)^{\frac{\sum_{i=1}^n \sum_{j=1}^m \frac{f_{ij+1}}{D_i} \frac{f_{ij}}{D_i}}{n \cdot m}} \quad (1);$$

其中, S_i 表示第i个病原微生物的数据信息,n表示病原微生物的总个数,m表示病原微生物的参数的总个数, f_{ij} 表示病原微生物的第i个病原微生物的第j个参数的参数值,公式(1)中的 D_i 表示病原微生物的第i个病原微生物参数值总和;

步骤A2:将公式(1)得到所述比对数据库与所述待比对数据通过公式(2)得到比对结果值 K_i ;

$$K_i = \ln \left(\frac{\sum_{i=1}^n (e^{\frac{K}{S_i}} + \ln(e^K - e^{S_i}))^n}{n^{(K-S_i)}} \right) \quad (2);$$

其中, K_i 表示所述待比对数据信息与第i个病原微生物的数据信息的比对结果值,K表示所述待比对数据的数据信息;

步骤A3:利用公式(3)对公式(2)所求得的 K_i 进行优化处理得到所述待比对数据与对应的病原微生物之间的相关概率值 D_i ;

$$D_i = \frac{n \times \sum_{i=1}^n K_i \times \left[\frac{K_i}{n} + \frac{\arccos(K_{i+1}) - \arccos(K_i)}{n} \right]}{K_i^2} \times 100\% \quad (3);$$

其中,公式(3)中的 D_i 表示所述待比对数据与第 i 个病原微生物之间的相关概率值;

当公式(3)中的 $D_i \leq 80\%$ 时将其对应的第 i 个病原微生物为非主要病原微生物;

当公式(3)中的 $D_i > 80\%$ 时将其对应的第 i 个病原微生物整理出来即为主要病原微生物;

步骤A4:利用公式(4)对公式(3)中所求得的相关概率值 D_i 进行自动控制处理,将符合要求的主要病原微生物以及非主要病原微生物分拣出来并依据公式(4)求出的比对数量值来自动增减待比对数据;

$$M = \exp \left\{ \frac{\ln(n \times \sum_{i=1}^n (D_i - 80\%)) + \ln((n+l) \times \sum_{i=1}^{n-l} (D_i - 80\%))}{n+l} \right\} \quad (4);$$

其中, M 为比对数量值, l 为自动增减的待比对数据总数且, $l = \frac{M - M_0}{1 + e^{\frac{M - M_0}{M_0}}}$, M_0 为预设

可达到临床检测的主要病原微生物的数量值;

当主要病原微生物的数量小于可达到临床检测的主要病原微生物的数量值的时候系统会自动增减其对比数据总数 l 直至主要病原微生物的数量达到临床检测的主要病原微生物的数量值,从而实现最终的临床检验。

2.如权利要求1所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,

所述比对数据库为预先建立的微生物基因组数据库;

且基于预先建立的微生物基因组数据库,对所述待比对数据进行比对处理,根据比对结果确定主要病原微生物的过程包括:

将所述待比对数据与所述预先建立的微生物基因组数据库进行比对处理,计算比对到所述微生物基因组数据库中的病原微生物基因组的各项参数;

根据计算得到的病原微生物的各项参数,确定所述主要病原微生物;

其中,所述各项参数包括:所述待比对数据比对到病原微生物基因组的序列数、特异序列数、覆盖度和覆盖长度。

3.如权利要求2所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,在计算得到病原微生物的各项参数之后,还包括:

根据所述第二待检测微生物样品的待比对数据比对到内参DNA的特异序列数,判断所述第二待检测微生物样品中的内参检测结果中相应的内参序列数是否为零,若是,判定检测失败。

4.如权利要求2所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,在计算得到病原微生物的各项参数之后,还包括:

根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数,评价整个检测流程成功与否,若相应的微生物特异序列数为零,则判定为实验失败;

根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因

组数据库的序列数,和与所述阳性对照样本相关的参考序列数的比值结果,确定病原微生物的检出效率。

5.如权利要求2所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,在计算得到病原微生物的各项参数之后,还包括:

根据所述预先设定的空白对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数,确定空白对照样本中的微生物检出结果,评价实验的污染情况。

6.如权利要求5所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,在确定空白对照样本中的微生物检出结果之后,还包括:

比较所述预先设定的空白对照样本中的特定微生物的第一检出数量与所述第二待检测微生物样品中的特定微生物的第二检出数量;

当所述第二检出数量与所述第一检出数量的数量比值大于预设比值时,判定所述第二待检测微生物样品中检测出的所述主要病原微生物为阳性结果。

7.如权利要求1所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,所述预先设定的阳性对照样本为已知微生物的混合物;

所述预先设定的空白对照样本为无核酸酶水/缓冲液。

8.如权利要求1所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,在采用病原微生物检测试剂盒分别对所述DNA样本构建测序文库时,包括:

对所述DNA样本进行末端修复、连接接头、DNA纯化、PCR富集和文库纯化。

9.如权利要求1所述的用于非疾病诊断目的的全流程质控的病原微生物高通量测序检测方法,其特征在于,

采用酶解法分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA。

全流程质控的病原微生物高通量测序检测方法

技术领域

[0001] 本发明涉及感染性疾病临床检测技术领域,特别涉及全流程质控的病原微生物高通量测序检测方法。

背景技术

[0002] 感染性疾病病原微生物的传统诊断技术包括镜检、培养、血清学方法等。镜检的敏感性较低,除非样本属高疾病负荷,并且特异性不强。培养的敏感性因病原体而异,尤其是一些所需营养复杂的难培养的微生物,如隐球菌等最常见的3种神经侵袭性真菌可以在标准的、敏感性可变的细菌培养基上培养,至于更多的罕见真菌,则需要特定的培养基;如单纯疱疹病毒仅能从至多5%的单纯疱疹病毒脑炎病例脑脊液中培养出来,且抗生素治疗也会影响培养的敏感性。此外,此方法所需的时间较长,细菌培养至少需要3~5天,用细胞系培养病毒可能需要30天。血清学方法在感染早期的诊断中敏感性不高,会出现体液免疫缺陷的假阴性。聚合酶链式反应(PCR)及其衍生技术是基于特定靶点的有限微生物的检测,且与上述传统诊断技术相似,依赖于对病原微生物的先验认知,且只局限于微生物基因组的一小部分。利用综合的传统诊断技术,仅能在小部分的病例中实现病原微生物检测。

[0003] 宏基因组二代测序(Metagenomic next-generation sequencing,mNGS)为感染性疾病的病原微生物诊断提供了一种无需预知、无偏好、高通量的方法。不需要对病原体进行培养,不依赖于临床预测因素或实验室检测结果,可直接从患者的脑脊液、血液、胸水、腹水、肺泡灌洗液等样本中提取核酸,可以在短时间内一次性检测样本中几乎所有潜在的病原体(病毒、细菌、真菌和寄生虫),包括新发病原体,具有良好的敏感性和特异性。

[0004] 临床宏基因组学的实施,特别强调质量控制和质量保证。由于mNGS检测病原微生物涉及核酸提取、文库构建、上机测序流程的多个实验步骤,因此需要对其整个流程进行严格质控以确保检测结果的可靠性和准确性。mNGS检测病原微生物的临床实施需要设置以下质控:1)如以DNA序列、特定生物体等作为内参,可以质控待检测微生物样品的全流程实验是否成功,若无内参,则很难判断某个样本的阴性结果是否是由于该样本实验失败导致的;2)以多个已知微生物组成的样本作为阳性对照,可用于监控整批待检测微生物样品的核酸提取、文库构建、上机测序、信息分析的过程,若无已知的阳性对照样本与待检测微生物样品同步实施整个实验与分析流程,则无法评估整个检测流程的成功与否;3)空白对照或阴性对照用于监控外源的或试剂的微生物污染和跨样本污染,若无空白对照或阴性对照样本与待检测微生物样品同步实施整个实验与分析流程,则无法排除实验过程中的污染,很难判断待检测微生物样品的阳性结果是否与污染导致的假阳性相关。然而目前尚无针对临床需要的系统的全流程质控性mNGS检测病原微生物的解决方案。

发明内容

[0005] 本发明提供全流程质控的病原微生物高通量测序检测方法,便于实现感染性疾病病原微生物的快速临床诊断,并以加入内参、设置阳性对照样本和空白对照样本,便于实现

全流程质控。

[0006] 本发明提供全流程质控的病原微生物高通量测序检测方法,包括:

[0007] 在第一待检测微生物样品中加入内参,得到第二待检测微生物样品;

[0008] 分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA;

[0009] 采用病原微生物检测试剂盒分别对所述DNA样本构建测序文库;

[0010] 基于二代测序平台对所述测序文库进行测序,并对测序结果中的原始测序数据进行质控处理,并去除比对到宿主基因组的序列,获得待比对数据;

[0011] 基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物。

[0012] 在一种可能实现的方式中,

[0013] 所述比对数据库为预先建立的微生物基因组数据库;

[0014] 且基于预先建立的微生物基因组数据库,对所述待比对数据进行比对处理,根据比对结果确定主要病原微生物的过程包括:

[0015] 将所述待比对数据与所述预先建立的微生物基因组数据库进行比对处理,计算比对到所述微生物基因组数据库中的病原微生物基因组的各项参数;

[0016] 根据计算得到的病原微生物的各项参数,确定所述主要病原微生物;

[0017] 其中,所述各项参数包括:所述待比对数据比对到病原微生物基因组的序列数、特异序列数、覆盖度和覆盖长度。

[0018] 在一种可能实现的方式中,

[0019] 在计算得到病原微生物的各项参数之后,还包括:

[0020] 根据所述第二待检测微生物样品的待比对数据比对到内参DNA的特异序列数,判断所述第二待检测微生物样品中的内参检测结果中相应的内参序列数是否为零,若是,判定检测失败。

[0021] 在一种可能实现的方式中,

[0022] 在计算得到病原微生物的各项参数之后,还包括:

[0023] 根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数,评价整个检测流程成功与否,若相应的微生物特异序列数为零,则判定为实验失败;

[0024] 根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数,和与所述阳性对照样本相关的参考序列数的比值结果,确定病原微生物的检出效率;

[0025] 在一种可能实现的方式中,

[0026] 在计算得到病原微生物的各项参数之后,还包括:

[0027] 根据所述预先设定的空白对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数,确定空白对照样本中的微生物检出结果,评价实验的污染情况。

[0028] 在一种可能实现的方式中,

[0029] 在确定空白对照样本中的微生物检出结果之后,还包括:

[0030] 比较所述预先设定的空白对照样本中的特定微生物的第一检出数量与所述第二

待检测微生物样品中的特定微生物的第二检出数量；

[0031] 当所述第二检出数量与所述第一检出数量的数量比值大于预设比值时,判定所述第二待检测微生物样品中检测出的所述主要病原微生物为阳性结果。

[0032] 在一种可能实现的方式中,

[0033] 所述预先设定的阳性对照样本为已知微生物的混合物;

[0034] 所述预先设定的空白对照样本为无核酸酶水/缓冲液。

[0035] 在一种可能实现的方式中,

[0036] 在采用病原微生物检测试剂盒分别对所述DNA样本构建测序文库时,包括:

[0037] 对所述DNA样本进行末端修复、连接接头、DNA纯化、PCR富集和文库纯化。

[0038] 在一种可能实现的方式中,

[0039] 采用酶解法分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA。

[0040] 在一种可能实现的方式中,

[0041] 基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物时,包括:

[0042] 步骤A1:利用公式(1)预先建立的比对数据库 S_i ;

$$[0043] \quad S_i = \frac{1}{n+m} \sum_{i=1}^n \sum_{j=1}^m \left(\frac{\exp(\frac{f_{ij+1}}{D_i} - \frac{f_{ij}}{D_i})}{\frac{f_{ij}}{D_i}} \right)^{\sum_{j=1}^m \sum_{i=1}^n \frac{f_{ij-1}}{D_i} \cdot \frac{f_{ij}}{D_i}} \quad (1);$$

[0044] 其中, S_i 表示第*i*个病原微生物的数据信息, n 表示病原微生物的总个数, m 表示病原微生物的参数的总个数, f_{ij} 表示病原微生物的第*i*个病原微生物的第*j*个参数的参数值, D_i 表示病原微生物的第*i*个病原微生物参数值总和;

[0045] 步骤A2:将公式(1)得到所述比对数据库与所述待比对数据通过公式(2)得到比对结果值 K_i ;

$$[0046] \quad K_i = \ln \left(\frac{\sum_{i=1}^n (e^{S_i} + \ln(e^K - e^{S_i}))^n}{n^{(K-S_i)}} \right) \quad (2);$$

[0047] 其中, K_i 表示所述待比对数据信息与第*i*个病原微生物的数据信息的比对结果值, K 表示所述待比对数据的数据信息;

[0048] 步骤A3:利用公式(3)对公式(2)所求得的 K_i 进行优化处理得到所述待比对数据与对应的病原微生物之间的相关概率值 D_i ;

$$[0049] \quad D_i = \frac{n \times \sum_{i=1}^n K_i \times \left[\frac{K_i}{n} + \frac{\arccos(K_{i+1}) - \arccos(K_i)}{n} \right]}{K_i^2} \times 100\% \quad (3);$$

[0050] 其中, D_i 表示所述待比对数据与第*i*个病原微生物之间的相关概率值;

[0051] 当 $D_i \leq 80\%$ 时将其对应的第*i*个病原微生物为非主要病原微生物;

[0052] 当 $D_i > 80\%$ 时将其对应的第*i*个病原微生物整理出来即为主要病原微生物;

[0053] 步骤A4:利用公式(4)对公式(3)所求得的相关概率值 D_i 进行自动控制处理,将符合要求的主要病原微生物以及非主要病原微生物分拣出来并依据公式(4)求出的比对数量值来自动增减待比对数据;

$$[0054] \quad M = \exp \left\{ \frac{\ln \left(n \times \sum_{i=1}^n (D_i - 80\%) \right) + \ln \left((n+l) \times \sum_{i=1}^{n+l} (D_i - 80\%) \right)}{n+l} \right\} \quad (4);$$

[0055] 其中, M 为比对数量值, l 为自动增减的待比对数据总数且, $l = \frac{M - M_0}{1 + e^{M - M_0}}$, M_0 为预设

可达到临床检测的主要病原微生物的数量值;

[0056] 当主要病原微生物的数量小于可达到临床检测的主要病原微生物的数量值的时候系统会自动增减其对比数据总数 l 直至主要病原微生物的数量达到临床检测的主要病原微生物的数量值,从而实现最终的临床检验。

[0057] 本发明的有益效果为:实现针对感染性疾病的无需预知、无偏好、高通量的病原微生物的快速临床检测;

[0058] 通过在第二待检测样本中加入内参精确质控全流程,避免出现因实验失败导致的假阴性,同时减少实验环节引入的污染,降低假阳性;

[0059] 通过设置阳性对照样本,质控整批实验,同时评价病原微生物的检出效率,便于评估整个检测流程的成功与否;

[0060] 通过设置空白对照样本,便于有效评估实验污染情况;

[0061] 通过比较第一检出数量与第二检出数量的差异,有效降低假阳性。

[0062] 本发明的其它特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本发明而了解。本发明的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

[0063] 下面通过附图和实施例,对本发明的技术方案做进一步的详细描述。

附图说明

[0064] 附图用来提供对本发明的进一步理解,并且构成说明书的一部分,与本发明的实施例一起用于解释本发明,并不构成对本发明的限制。在附图中:

[0065] 图1为本发明实施例中全流程质控的病原微生物高通量测序检测方法的流程示意图。

具体实施方式

[0066] 以下结合附图对本发明的优选实施例进行说明,应当理解,此处所描述的优选实施例仅用于说明和解释本发明,并不用于限定本发明。

[0067] 相比于传统的病原微生物诊断技术和最相近的现有技术,本方案为感染性疾病诊断病原微生物的临床需要提供了系统的全流程质控性mNGS检测病原微生物的解决方案,如下:

[0068] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法,如图1所示,包括:

- [0069] 步骤1:在第一待检测微生物样品中加入内参,得到第二待检测微生物样品;
- [0070] 步骤2:分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA;
- [0071] 步骤3:采用病原微生物检测试剂盒分别对所述DNA样本构建测序文库;
- [0072] 步骤4:基于二代测序平台对所述测序文库进行测序,并对测序结果中的原始测序数据进行质控处理,并去除比对到宿主基因组的序列,获得待比对数据;
- [0073] 步骤5:基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物。
- [0074] 上述第一待检测微生物样品可以为:人类脑脊液、血液、胸水、腹水、肺泡灌洗液等;
- [0075] 上述内参可以为:合成DNA(脱氧核糖核苷酸);
- [0076] 上述第二代检测微生物样品如:加入合成DNA内参的人类脑脊液;
- [0077] 上述二代测序平台如:illumina平台;
- [0078] 上述病原微生物可以为:细菌、真菌、病毒和寄生虫,如结核分枝杆菌、非结核分枝杆菌、布鲁氏菌、单核细胞增生李斯特菌、铜绿假单胞菌、金黄色葡萄球菌、肠球菌、鲍曼不动杆菌、无乳链球菌、肺炎克雷伯菌、奈瑟氏菌、诺卡氏菌、接合菌、流感嗜血杆菌、肺炎链球菌、大肠埃希菌、隐球菌、曲霉菌、申克氏孢子丝菌、念珠菌、水痘带状疱疹病毒、单纯疱疹病毒、人类疱疹病毒第四型、巨细胞病毒、微小病毒B19、猪肉绦虫、弓形虫、管圆线虫、阿米巴虫等;
- [0079] 上述预先设定的阳性对照样本,可以为已知的微生物的混合物,如可以是包括细菌、真菌、病毒或寄生虫在内的混合物(mix),即微生物mix阳性对照;
- [0080] 上述预先设定的空白对照样本,可以为无核酸酶水/缓冲液,理论上不包含任何微生物,即微生物空白对照;
- [0081] 上述阳性对照样本和空白对照样本是同步实施上述步骤2~5中的整个实验与分析流程的;
- [0082] 上述对原始测序数据进行质控处理,是去除原始测序数据中的低质量和含有接头的序列,且原始测序数据是根据测序文库获得的;
- [0083] 上述去除比对到宿主基因组的序列,是将质控处理后的数据与宿主基因组(人类基因组)比对,去除比对上的数据;
- [0084] 上述待比对数据,是去除原始测序数据中的低质量和含有接头的序列及比对到宿主基因组的序列后剩余的数据;
- [0085] 上述比对数据库是微生物基因组数据库;
- [0086] 上述技术方案的有益效果是:通过基于二代测序平台的宏基因组测序,便于实现无需预知、无偏好、高通量的病原微生物诊断;通过在待检测微生物样品中加入内参(即合成DNA内参),同时设置阳性对照样本和空白对照样本,便于实现病原微生物检测的全流程质控,即全流程QC。
- [0087] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法,
- [0088] 所述比对数据库为:预先建立的微生物基因组数据库;
- [0089] 且基于预先建立的微生物基因组数据库,对所述待比对数据进行比对处理,根据

比对处理结果确定主要病原微生物的过程包括：

[0090] 将所述待比对数据与所述预先建立的微生物基因组数据库进行比对处理，检测比对到所述微生物基因组数据库中的病原微生物基因组的各项参数；

[0091] 根据计算得到的病原微生物的各项参数，确定主要病原微生物；

[0092] 其中，所述各项参数包括：所述待比对数据比对到病原微生物基因组的序列数、特异序列数、覆盖度和覆盖长度。

[0093] 上述序列数是指待比对数据比对到微生物基因组数据库中的特定微生物基因组的所有reads(序列片段)数量；

[0094] 上述特异序列数是指待比对数据中唯一比对到微生物基因组数据库中的特定微生物基因组的reads数量；

[0095] 上述覆盖度是指特定微生物基因组中有reads覆盖的核酸序列长度占该特定微生物整个基因组序列长度的百分比；

[0096] 上述覆盖长度是指特定微生物基因组上有reads覆盖的片段长度之和。

[0097] 上述技术方案的有益效果是：便于根据各项参数，有效的确定主要病原微生物。

[0098] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，在计算得到病原微生物的各项参数之后，还包括：

[0099] 根据所述第二待检测微生物样品的待比对数据比对到内参DNA的特异序列数，判断所述第二待检测微生物样品中的内参检测结果中相应的内参序列数是否为零，若是，判定检测失败。

[0100] 上述技术方案的有益效果是：根据判断内参序列数，便于排除因为检测失败导致的第二待检测微生物样品检测结果假阴性，且内参的存在提高了DNA的有效起始量，减少了实验环节可能引入的污染，降低了假阳性。

[0101] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，在计算得到病原微生物的各项参数之后，还包括：

[0102] 根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数，若相应的微生物特异序列数为零，则判定为实验失败。

[0103] 例如：含有肺炎克雷伯菌的阳性对照样本，若阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库中，得到的肺炎克雷伯菌的特异序列数为0，说明与此阳性对照样本同批次的第二待测微生物样品检测失败。

[0104] 根据所述预先设定的阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数，和与所述阳性对照样本相关的参考序列数的比值结果，确定病原微生物的检出效率。

[0105] 例如：含有肺炎克雷伯菌的阳性对照样本，若阳性对照样本的待比对数据比对到所述预先建立的微生物基因组数据库中，得到的肺炎克雷伯菌的特异序列数为800个，参考序列数为1000个，对应的检出效率为：80%；

[0106] 其中，参考序列数为以阳性对照样本为基础，已知的阳性对照样本中某微生物的特异序列数。

[0107] 上述技术方案的有益效果是：根据与第二待检测微生物样品同批次进行检测的阳性对照样本的检测结果，便于评估整批实验的检测和分析流程成功与否，并且可基于阳性

对照样本的检测结果评价病原微生物的检出效率。

[0108] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0109] 在计算得到病原微生物的各项参数之后，还包括：

[0110] 根据所述预先设定的空白对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数，确定空白对照样本中的微生物检出结果，评价实验的污染情况。

[0111] 例如：空白对照样本理论上是不包含任何微生物，空白对照样本的待比对数据比对到所述预先建立的微生物基因组数据库的序列数，确定空白对照样本中的铜绿假单胞菌特异序列数为500，说明与此空白对照样本同批次的第二待测微生物样品检测流程中存在铜绿假单胞菌的污染。

[0112] 上述技术方案的有益效果是：根据与第二待检测微生物样品同批次进行检测的空白对照样本的检测结果，可评价实验的污染情况。

[0113] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0114] 在计算得到病原微生物的各项参数之后，还包括：

[0115] 比较所述预先设定的空白对照样本中的特定微生物的第一检出数量与所述第二待检测微生物样品中的特定微生物的第二检出数量；

[0116] 当所述第二检出数量与所述第一检出数量的数量比值大于预设比值时，判定所述第二待检测微生物样品中检测出的所述可能的主要致病微生物为阳性结果。

[0117] 例如：预先设定的空白对照样本中的肺炎链球菌的第一检出数量为30，第二待检测微生物样品中的肺炎链球菌的第二检出数量为600，两者比值为20（大于预设比值10），则肺炎链球菌可能是第二待检测微生物样品中的主要致病微生物。

[0118] 上述预设比值10是根据现有的科学实验所确定的数据。

[0119] 上述技术方案的有益效果是：通过比较第二待检测微生物样品与空白对照样本的特定微生物的检出数量，可以显著降低第二待检测微生物样品检出结果的假阳性。

[0120] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0121] 所述预先设定好的阳性对照样本为已知微生物的混合物；

[0122] 所述预先设定好的空白对照样本为无核酸酶水/缓冲液。

[0123] 上述技术方案的有益效果是：通过提前设定好对照样本，便于进行全流程的质控，提高病原微生物检测的准确性和可靠性。

[0124] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0125] 在采用病原微生物检测试剂盒分别对提取的DNA样本构建测序文库时，包括：

[0126] 对所述DNA样本进行末端修复、连接接头、DNA纯化、PCR富集、文库纯化。

[0127] 上述技术方案的有益效果是：采用病原微生物检测试剂盒进行文库构建，便于获得适用于二代测序平台的高质量测序文库。

[0128] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0129] 采用酶解法分别提取所述第二待检测微生物样品、预先设定的阳性对照样本和空白对照样本中的DNA。

[0130] 上述技术方案的有益效果是：采用酶解法进行DNA提取，便于获得适用于二代测序的较全面的DNA样本。

[0131] 本发明实施例提供全流程质控的病原微生物高通量测序检测方法，

[0132] 基于预先建立的比对数据库,对所述待比对数据进行比对处理,根据比对结果,确定主要病原微生物时,包括:

[0133] 步骤A1:利用公式(1)预先建立的比对数据库 S_i :

$$[0134] \quad S_i = \frac{1}{n+m} \sum_{i=1}^n \sum_{j=1}^m \left(\frac{\exp(\frac{f_{ij+1}}{D_i} - \frac{f_{ij}}{D_i})}{\frac{f_{ij}}{D_i}} \right)^{\frac{\sum_{i=1}^n \sum_{j=1}^m f_{ij+1} - f_{ij}}{n \times m}} \quad (1);$$

[0135] 其中, S_i 表示第*i*个病原微生物的数据信息, n 表示病原微生物的总个数, m 表示病原微生物的参数的总个数, f_{ij} 表示病原微生物的第*i*个病原微生物的第*j*个参数的参数值, D_i 表示病原微生物的第*i*个病原微生物参数值总和;

[0136] 步骤A2:将公式(1)得到所述比对数据库与所述待比对数据通过公式(2)得到比对结果值 K_i :

$$[0137] \quad K_i = \ln \left(\frac{\sum_{i=1}^K (e^{\frac{K}{S_i}} + \ln(e^K - e^{\frac{K}{S_i}}))}{n^{(K-S_i)}} \right) \quad (2);$$

[0138] 其中, K_i 表示所述待比对数据信息与第*i*个病原微生物的数据信息的比对结果值, K 表示所述待比对数据的数据信息;

[0139] 步骤A3:利用公式(3)对公式(2)所求得的 K_i 进行优化处理得到所述待比对数据与对应的病原微生物之间的相关概率值 D_i :

$$[0140] \quad D_i = \frac{n \times \sum_{i=1}^n K_i \times \left[\frac{K_i + \arccos(K_{i+1}) - \arccos(K_i)}{n} \right]}{K_i^2} \times 100\% \quad (3);$$

[0141] 其中, D_i 表示所述待比对数据与第*i*个病原微生物之间的相关概率值;

[0142] 当 $D_i \leq 80\%$ 时将其对应的第*i*个病原微生物为非主要病原微生物;

[0143] 当 $D_i > 80\%$ 时将其对应的第*i*个病原微生物整理出来即为主要病原微生物;

[0144] 步骤A4:利用公式(4)对公式(3)所求得的相关概率值 D_i 进行自动控制处理,将符合要求的主要病原微生物以及非主要病原微生物分拣出来并依据公式(4)求出的比对数量值来自动增减待比对数据;

$$[0145] \quad M = \exp \left\{ \frac{\ln(n \times \sum_{i=1}^n (D_i - 80\%)) + \ln((n+l) \times \sum_{i=1}^{n-l} (D_i - 80\%))}{n+l} \right\} \quad (4);$$

[0146] 其中, M 为比对数量值, l 为自动增减的待比对数据总数且, $l = \frac{M - M_0}{1 + e^{M - M_0}}$, M_0 为预设可达到临床检测的主要病原微生物的数量值;

[0147] 当主要病原微生物的数量小于可达到临床检测的主要病原微生物的数量值的时候系统会自动增减其对比数据总数 l 直至主要病原微生物的数量达到临床检测的主要病原微生物的数量值,从而实现最终的临床检验。

[0148] 上述技术方案的有益效果是:通过对病原微生物的参数值分析得到比对数据库,

是为了保证得到的比对数据库的全面真实并且数据化,并且利用对比公式对所述比对数据库与所述待比对数据进行比对,保证了所述待比对数据与每一个病原微生物都进行对比,从而保证对比数据的可靠性,并且利用优化公式的方法得到相关概率值,从而使得最终得到的结果准确保证准确性,利用自动控制增减对比数据来达到可以临床检测极大的缩减了检测和比对的时间。

[0149] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

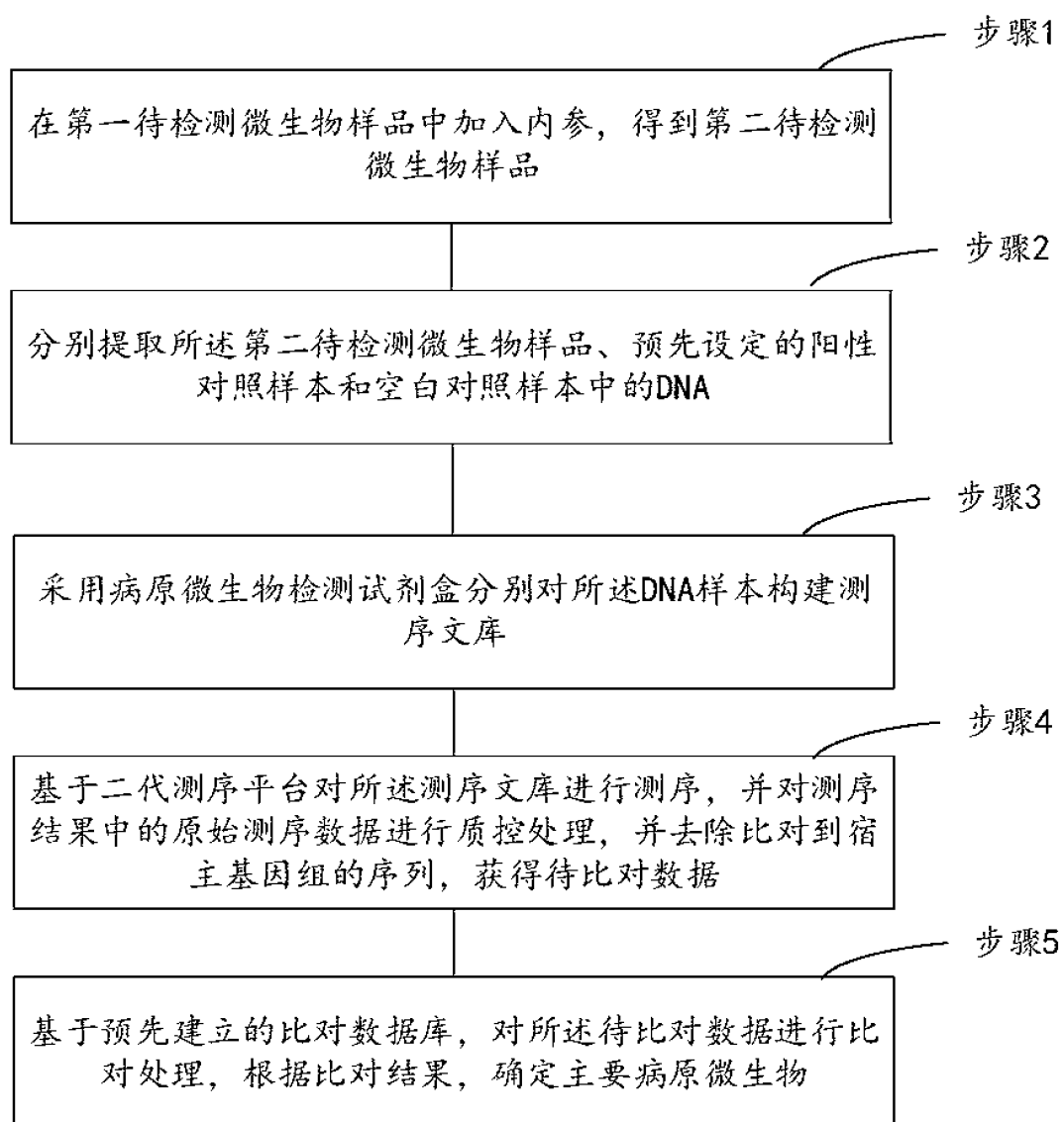


图1