

1. 一种微生物的种特异共有序列的获得方法，其特征在于，所述方法至少包括以下步骤：

S100，寻找候选共有序列：基于聚类算法对属于同一菌种的各个目标菌株的特异性序列进行聚类，获得多个候选种特异共有序列；

S200，验证和获得初次筛选种特异共有序列：

判断候选种特异共有序列是否满足以下条件：

1) 株种覆盖度满足预设值；

2) 有效拷贝数满足预设值；

若候选种特异共有序列满足以上所有条件，则为种特异共有序列；

其中，

株种覆盖度=（出现该候选种特异共有序列的目标菌株数量/目标菌株的总数量）*100%；

所述有效拷贝数采用公式（I）进行计算：

$$\sum_{i=0}^n Ci * \left(\frac{Si}{Sall} \right) \quad (I) ;$$

其中，

n 为候选种特异共有序列拷贝数梯度的总数；

Ci 为第 i 个候选种特异共有序列对应的拷贝数；

Si 为出现第 i 个候选种特异共有序列的菌株数量；

Sall 为目标菌株总数量。

2. 如权利要求 1 所述的微生物的种特异共有序列的获得方法，其特征在于，所述特异性序列是指，同属于同一目标菌株的目标片段，所述目标片段所在区域是所述目标菌株的特异性区域。
3. 如权利要求 2 所述的微生物的种特异共有序列的获得方法，其特征在于，所述特异性区域为特异性多拷贝区域。
4. 如权利要求 2 所述的微生物的种特异共有序列的获得方法，其特征在于，所述特异性区域的获得方法包括如下步骤：

S110，将微生物目标片段与一个或多个对比菌株全基因组序列分别进行两两比对，去除相似性超过预设值的片段，获得若干剩余片段，作为第一轮切割片段 T₁~T_n，其中，n 为 ≥1 的整数；

S120, 将所述第一轮切割片段 $T_1 \sim T_n$ 分别与其他对比菌株的全基因组序列比对, 去除相似性超过预设值的片段, 获得剩余切割片段的集合作为微生物目标片段的候选特异性区域;

S130, 验证和获得特异性区域: 判断候选特异性区域是否满足以下条件:

1) 搜寻公共数据库, 寻找是否存在与候选特异性区域相似性大于预设值的其他物种;

2) 将候选特异性区域分别与各对比菌株的全基因组序列, 以及微生物目标片段来源菌株的寄主的全基因组序列进行比对, 寻找是否存在相似性大于预设值的片段;

若候选特异性区域均不满足以上条件, 则为微生物目标片段的特异性区域。

5. 如权利要求 4 所述的微生物的种特异共有序列的获得方法, 其特征在于, 还包括以下特征中的一项或多项:

a. 所述方法能够区别所述微生物目标片段来源菌株与对比菌株是否为同一种或亚种;

b. 所述相似性是指微生物目标片段的覆盖率与匹配率的乘积, 所述覆盖率=(相似序列片段长度/(微生物目标片段末端值-微生物目标片段起始值+1))%;

c. 步骤 S120 中, 所述第一轮切割片段 $T_1 \sim T_n$ 分别与其他对比菌株的全基因组序列分组迭代比对;

d. 所述相似性预设值超过 80%;

e. 两个进行对比的序列之间的各个碱基的位置不发生交叉;

f. 所述方法还包括以下步骤: S111, 将选定的相邻微生物目标片段进行两两比对, 若出现相似性低于预设值的比对结果, 则发出警报并显示目标株种对应的筛选条件。

6. 如权利要求 5 所述的微生物的种特异共有序列的获得方法, 其特征在于, 将第一轮切割片段 T_n 与其他对比菌株的全基因组序列进行分组迭代比对时包括下列步骤:

S121, 将其余对比菌株分为 P 组, 每组若干个对比菌株;

S122, 将第一轮切割片段 T_n 同时与第 1 组中各对比菌株的全基因组序列进行一对一的比对, 去除相似性超过预设值的片段, 获得若干剩余片段作为第一轮切割片段 T_n 的第一轮候选序列库;

S123, 将第一轮切割片段 T_n 的上一轮候选序列库同时与下一组中各对比菌株的全基因组序列进行一对一的比对, 去除相似性超过预设值的片段, 获得若干剩余片段作为第一轮切

割片段 T_n 的下一轮候选序列库；从第一轮候选序列库开始重复步骤 S122 直至获得第 P 轮候选序列库作为第一轮切割片段 T_n 的候选特异性序列库；

所有第一轮切割片段的候选特异性序列库的集合即为候选特异性区域。

7. 如权利要求 3 所述的微生物的种特异共有序列的获得方法，其特征在于，多拷贝区域的获得方法包括如下步骤：

S140，寻找候选多拷贝区域：对微生物目标片段进行内部比对，寻找相似性满足预设值的待测序列对应的区域作为候选多拷贝区域，所述相似性是指待测序列的覆盖率与匹配率的乘积；

S150，验证获得多拷贝区域：获得候选多拷贝区域拷贝数的中值；若候选多拷贝区域拷贝数的中值大于 1，则记为多拷贝区域。

8. 如权利要求 7 所述的微生物的种特异共有序列的获得方法，其特征在于，还包括以下特征中的一项或多项：

- a. 覆盖率=(相似序列的长度/(待测序列末端值-待测序列起始值+1))%；
- b. 当微生物目标片段为不完整的多条基序时，将各基序连接起来再寻找候选多拷贝区域；
- c. 候选多拷贝区域拷贝数的中值的获得方法为：确定各候选多拷贝区域在微生物目标片段上的位置，获得待验证候选多拷贝区域的每个碱基位置上覆盖的其他候选多拷贝区域的个数，计算该待验证候选多拷贝区域的拷贝数的中值；
- d. 在步骤 S150 中，还可计算候选多拷贝区域拷贝数的 95%置信区间；优选的，在计算候选多拷贝区域拷贝数的 95%置信区间时，以候选多拷贝区域的碱基数为样本数，候选多拷贝区域中各碱基对应的拷贝数值为样本值计算。

9. 如权利要求 1 所述的微生物的种特异共有序列的获得方法，其特征在于，还包括如下步骤中的一项或多项：

S300，针对初次筛选种特异共有序列，根据探针引物设计规则，进行探针、引物设计，获得候选探针引物；将候选探针引物序列比对到所有目标菌株全基因组上，计算每条探针引物序列对应的菌株覆盖度，筛选出菌株覆盖度都满足预设值的候选探针引物，将筛选出的候选探针引物对应的初次筛选种特异共有序列作为最终的种特异共有序列；

S400，若步骤 S200 中候选共有序列的株种覆盖度均未达到预设值，则将候选共有序列进行组合，从中筛选出株种覆盖度能够达到预设值并且共有序列最少的组合，以筛选出的组合作为候选共有序列，采用步骤 S200 验证和获得初次筛选种特异共有序列。

10. 如权利要求 9 所述的微生物的种特异共有序列的获得方法，其特征在于，还包括如下步骤：

S500，若步骤 S300 中所述候选探针引物的菌株覆盖度均未达到预设值，则将初次筛选种特异共有序列进行组合，从中筛选出株种覆盖度能够达到预设值并且共有序列最少的组合，以筛选出的组合作为候选共有序列，采用步骤 S200 验证和获得初次筛选种特异共有序列。

11. 如权利要求 9 或 10 所述的微生物的种特异共有序列的获得方法，其特征在于，在步骤 S400 和 S500 中，将共有序列数目由低到高进行组合，来进行选择。

12. 如权利要求 9 或 10 所述的微生物的种特异共有序列的获得方法，其特征在于，当所述目标菌株的数量进行更新时，将原有的候选探针引物比对到更新的目标菌株全基因组上，计算覆盖度，验证原有的候选探针引物是否可以覆盖更新的目标菌株。

13. 一种微生物的种特异共有序列的获得装置，其特征在于，所述装置至少包括以下模块：

候选共有序列寻找模块，用于基于聚类算法对属于同一菌种的各个目标菌株的特异性序列进行聚类，获得多个候选种特异共有序列；

初次筛选种特异共有序列验证和获得模块，用于判断候选种特异共有序列是否满足以下条件：

1) 株种覆盖度满足预设值；

2) 有效拷贝数满足预设值；

若候选种特异共有序列满足以上所有条件，则为种特异共有序列；

其中，

株种覆盖度 = (出现该候选种特异共有序列的目标菌株数量 / 目标菌株的总数量) * 100%；

所述有效拷贝数采用公式 (I) 进行计算：

$$\sum_{i=0}^n C_i * \left(\frac{S_i}{S_{all}} \right) \quad (I) ;$$

其中，

n 为候选种特异共有序列拷贝数梯度的总数；

C_i 为第 i 个候选种特异共有序列对应的拷贝数；

S_i 为出现第 i 个候选种特异共有序列的菌株数量；

S_{all} 为目标菌株总数量。

14. 如权利要求 13 所述的微生物的种特异共有序列的获得装置，其特征在于，所述特异性序列是指，同属于同一目标菌株的目标片段，所述目标片段所在区域是所述目标菌株的特异性区域。
15. 如权利要求 14 所述的微生物的种特异共有序列的获得装置，其特征在于，所述特异性区域为特异性多拷贝区域。
16. 如权利要求 13 所述的微生物的种特异共有序列的获得装置，其特征在于，所述装置还包括以下模块，用于获得特异性区域：

第一轮切割片段获得模块，用于将微生物目标片段与一个或多个对比菌株全基因组序列分别进行两两比对，去除相似性超过预设值的片段，获得若干剩余片段，作为第一轮切割片段 $T_1 \sim T_n$ ，其中， n 为 ≥ 1 的整数；

候选特异性区域获得模块，用于将所述第一轮切割片段 $T_1 \sim T_n$ 分别与其余对比菌株的全基因组序列比对，去除相似性超过预设值的片段，获得剩余切割片段的集合作为微生物目标片段的候选特异性区域；

特异性区域验证和获得模块，用于判断候选特异性区域是否满足以下条件：

- 1) 搜寻公共数据库，寻找是否存在与候选特异性区域相似性大于预设值的其他物种；
 - 2) 将候选特异性区域分别与各对比菌株的全基因组序列，以及微生物目标片段来源菌株的寄主的全基因组序列进行比对，寻找是否存在相似性大于预设值的片段；
- 若候选特异性区域均不满足以上条件，则为微生物目标片段的特异性区域。
17. 如权利要求 16 所述的微生物的种特异共有序列的获得装置，其特征在于，还包括以下特征中的一项或多项：
 - a. 所述方装置能够区别所述微生物目标片段来源菌株与对比菌株是否为同一种或亚种；
 - b. 所述相似性是指微生物目标片段的覆盖率与匹配率的乘积，所述覆盖率=(相似序列片段长度/(微生物目标片段末端值-微生物目标片段起始值+1))%；
 - c. 所述候选特异性区域获得模块中，所述第一轮切割片段 $T_1 \sim T_n$ 分别与其余对比菌株的全基因组序列分组迭代比对；
 - d. 所述相似性预设值超过 80%；
 - e. 两个进行对比的序列之间的各个碱基的位置不发生交叉；
 - f. 第一轮切割片段获得模块还包括以下子模块，原始数据相似性比较子模块，用于

将选定的相邻微生物目标片段进行两两比对，若出现相似性低于预设值的比对结果，则发出警报并显示目标株种对应的筛选条件。

18. 如权利要求 17 所述的微生物的种特异共有序列的获得装置，其特征在于，将第一轮切割片段 T_n 与其余对比菌株的全基因组序列进行分组迭代比对时，所述候选特异性区域获得模块包括下列子模块：

对比菌株分组子模块，用于将其余对比菌株分为 P 组，每组若干个对比菌株；

第一轮候选序列库获得子模块，用于将第一轮切割片段 T_n 同时与第 1 组中各对比菌株的全基因组序列进行一对一的比对，去除相似性超过预设值的片段，获得若干剩余片段作为第一轮切割片段 T_n 的第一轮候选序列库；

候选特异性区域获得子模块，用于将第一轮切割片段 T_n 的上一轮候选序列库同时与下一组中各对比菌株的全基因组序列进行一对一的比对，去除相似性超过预设值的片段，获得若干剩余片段作为第一轮切割片段 T_n 的下一轮候选序列库；从第一轮候选序列库开始重复候选特异性区域获得子模块直至获得第 P 轮候选序列库作为第一轮切割片段 T_n 的候选特异性序列库；

所有第一轮切割片段的候选特异性序列库的集合即为候选特异性区域。

19. 如权利要求 15 所述的微生物的种特异共有序列的获得装置，其特征在于，所述装置还包括以下模块，用于获得多拷贝区域：

候选多拷贝区域寻找模块，用于对微生物目标片段进行内部比对，寻找相似性满足预设值的待测序列对应的区域作为候选多拷贝区域，所述相似性是指待测序列的覆盖率与匹配率的乘积；

多拷贝区域验证获得模块，用于获得候选多拷贝区域拷贝数的中值；若候选多拷贝区域拷贝数的中值大于 1，则记为多拷贝区域。

20. 如权利要求 19 所述的微生物的种特异共有序列的获得装置，其特征在于，还包括以下特征中的一项或多项：

a. 覆盖率=(相似序列的长度/(待测序列末端值-待测序列起始值+1))%；

b. 当微生物目标片段为不完整的多条基序时，将各基序连接起来再寻找候选多拷贝区域；

c. 所述多拷贝区域验证获得模块中还包括候选多拷贝区域拷贝数的中值的获得子模块，用于确定各候选多拷贝区域在微生物目标片段上的位置，获得待验证候选多拷贝区域的每个碱基位置上覆盖的其他候选多拷贝区域的个数，计算该待验证候选多拷贝区域的拷

贝数的中值；

d. 多拷贝区域验证获得模块还用于计算候选多拷贝区域拷贝数的 95%置信区间；优选的，在计算候选多拷贝区域拷贝数的 95%置信区间时，以候选多拷贝区域的碱基数为样本数，候选多拷贝区域中各碱基对应的拷贝数值为样本值计算。

21. 如权利要求 13 所述的微生物的种特异共有序列的获得装置，其特征在于，所述装置还包括以下模块的一个或多个：

最终的种特异共有序列筛选模块，用于针对初次筛选种特异共有序列，根据探针引物设计规则，进行探针、引物设计，获得候选探针引物；将候选探针引物序列比对到所有目标菌株全基因组上，计算每条探针引物序列对应的菌株覆盖度，筛选出菌株覆盖度都满足预设值的候选探针引物，将筛选出的候选探针引物对应的初次筛选种特异共有序列作为最终的种特异共有序列；

第一共有序列组合筛选模块，用于当初次筛选种特异共有序列验证和获得模块中，候选共有序列的株种覆盖度均未达到预设值时，将候选共有序列进行组合，从中筛选出株种覆盖度能够达到预设值并且共有序列最少的组合，以筛选出的组合作为候选共有序列，利用初次筛选种特异共有序列验证和获得模块，验证和获得初次筛选种特异共有序列。

22. 如权利要求 21 所述的微生物的种特异共有序列的获得装置，其特征在于，所述装置还包括以下模块：

第二共有序列组合筛选模块，若最终的种特异共有序列筛选模块中，所述候选探针引物的菌株覆盖度均未达到预设值，则将初次筛选种特异共有序列进行组合，从中筛选出株种覆盖度能够达到预设值并且共有序列最少的组合，以筛选出的组合作为候选共有序列，利用初次筛选种特异共有序列验证和获得模块，验证和获得初次筛选种特异共有序列。

23. 如权利要求 21 或 22 所述的微生物的种特异共有序列的获得装置，其特征在于，第一共有序列组合筛选模块和第二共有序列组合筛选模块中，将共有序列数目由低到高进行组合，来进行选择。

24. 如权利要求 21 或 22 所述的微生物的种特异共有序列的获得装置，其特征在于，所述装置还包括以下模块：

序列更新覆盖度模块，用于当所述目标菌株的数量进行更新时，将原有的候选探针引物比对到更新的目标菌株全基因组上，计算覆盖度，验证原有的候选探针引物是否可以覆

盖更新的目标菌株。

25. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现权利要求 1-12 任一所述的微生物的种特异共有序列的获得方法。
26. 一种计算机处理设备，包括处理器及权利要求 25 所述的计算机可读存储介质，其特征在于，所述处理器执行所述计算机可读存储介质上的计算机程序，实现权利要求 1-12 任一所述的微生物的种特异共有序列的获得方法的步骤。
27. 一种电子终端，其特征在于，包括：处理器、存储器、及通信器；所述存储器用于存储计算机程序，所述通信器用于与外部设备进行通信连接，所述处理器用于执行所述存储器存储的计算机程序，以使所述终端执行权利要求 1-12 任一所述的微生物的种特异共有序列的获得方法。
28. 如权利要求 1-12 所述的微生物的种特异共有序列的获得方法，权利要求 13-24 所述的微生物的种特异共有序列的获得装置，权利要求 25 所述的计算机可读存储介质，权利要求 26 所述的计算机处理设备或权利要求 27 所述的电子终端，用于核苷酸扩增中模板序列的筛选。
29. 一种微生物菌种鉴定方法，包括：通过扩增的方法鉴定目标菌株中是否含有采用种特异共有序列，所述种特异共有序列采用如权利要求 1-12 所述的微生物的种特异共有序列的获得方法，权利要求 13-24 所述的微生物的种特异共有序列的获得装置，权利要求 25 所述的计算机可读存储介质，权利要求 26 所述的计算机处理设备或权利要求 27 所述的电子终端获得。
30. 如权利要求 29 所述的微生物菌种鉴定方法，其特征在于，还包括以下特征中的一项或多项：
 - a. 所述方法能够区别所述微生物目标片段来源菌株与对比菌株是否为同一种或亚种；
 - b. 微生物选自细菌、病毒、真菌、变形虫、隐孢子虫、鞭毛虫、微孢子虫、梨形虫、疟原虫、弓形虫、毛滴虫或动质体中的一种或多种。