

附件1：微生物全基因组测序技术指导原则草案公示稿（第一次）

微生物全基因组测序技术指导原则

本指导原则对全基因组测序技术用于药品微生物控制给予通用性技术规定，为药用原料、辅料、制药用水、中间产品、终产品、包装材料、环境、设备和人员等药品全生命周期质量控制中微生物精准鉴定、溯源分析和风险识别等提供指导。

微生物全基因组测序（Microbial whole-genome sequencing）是指利用高通量测序技术对微生物个体的整个基因组序列进行测定，获取遗传信息的过程。高通量测序技术主要包括：边合成边测序、半导体测序、DNA（Deoxyribonucleic acid, DNA）纳米球测序、连接酶测序等第二代测序技术（又称下一代测序，Next Generation Sequencing）和基于单分子测序（Single Molecule Sequencing）的第三代测序技术。

第二代测序技术的基本原理主要是利用物理或酶切的方法将待测样本的基因组打断到1kb以内的DNA片段，在其两端连接特定接头序列后，固定于测序介质中，通过核酸扩增技术，如聚合酶链式反应、等温扩增技术等将待测样本放大收集成库，然后进行平行循环测序。

当需要获得微生物样本基因组精细图、完成图时，可采用能够实现大片段测序读长的第三代测序技术。第三代测序技术的基本原理主要有：采用荧光标记脱氧核糖核苷酸，用光学镜头实时记录DNA合成过程中新引入脱氧核糖核苷酸的荧光变化，通过不断地重复合成、成像、淬灭等过程进行单分子荧光测序；或采用电泳技术驱动单个分子逐一通过纳米孔，通过检测不同碱基的电信号，进行单分子纳米孔测序。

本指导原则以目前发展成熟、应用较为广泛的第二代测序技术为主要技术手段，对实验室的一般要求、全基因组测序的主要技术指标、技术流程、影响测序结果的主要因素、方法学考察和应用指导等方面进行通用性技术规定。

一、实验室的一般要求

1. 实验场地及人员

27 开展微生物全基因组测序的实验环境应具备分子生物学实验室的基本条
28 件，并符合相应级别的生物安全等级要求。实验区域一般应设置：试剂储存和
29 准备区、样本制备区、扩增区、核酸测序及分析区，各个区域在物理空间上相
30 互独立，并标识明确；另外，根据使用仪器的功能，相关区域可适当合并。应
31 单向流进入各工作区域，按照试剂储存和准备区、样本制备区、扩增区、核酸
32 测序及分析区的先后顺序进行实验操作。实验区域应定期进行清洁消毒。实
33 验人员应具备分子生物学和微生物学专业背景，或经专业培训。

34 2. 实验仪器

35 实验室一般应具备高通量核酸测序仪、核酸扩增仪、片段分析仪、核酸定
36 量仪、生物安全柜、混匀器、高速离心机、水浴或加热模块、冰箱、微量加样
37 器等分子生物学检验常用仪器设备。影响测序质量的仪器设备应定期进行性
38 能确认和维护，以保证仪器处于良好的运行状态。

39 3. 实验试剂

40 除另有规定外，所有实验使用的试剂均应不含DNA和DNA降解酶，宜大体
41 积配制、小体积分装，并保证试剂的无菌性，必要时可采用高压灭菌或0.22
42 μm 孔径滤膜过滤除菌。用于核酸扩增的相关试剂应避免反复冻融。关键试剂
43 应制定质量控制程序，以确保试剂质量。采用适宜的商品化试剂或试剂盒进
44 行核酸提取、文库构建和核酸测序时，应按照说明书操作，并符合说明书中的
45 质量控制要求。

46 二、全基因组测序的主要技术指标

47 1. 测序通量

48 测序通量是指单次测序可获得序列信息的基因片段数量或可测定的DNA
49 （以碱基表示）数量。核酸测序仪器的测序通量直接关系到测序输出的数据
50 量。微生物的基因组DNA较小，但不同种属之间变化幅度较大，如：葡萄球菌
51 属、埃希菌属、假单胞菌属、沙门菌属等常见细菌的基因组DNA大小约3~6 Mbp；
52 酵母菌的基因组DNA大小约12~16 Mbp；典型致病霉菌的基因组DNA通常大于
53 30 Mbp。

54 在进行微生物全基因组测序时，应根据待测样本基因组大小、样本数量
55 等实际需求，选择适宜测序通量的测序仪器和配套试剂，保证测序结果的准

56 确性。

57 2. 碱基识别质量

58 碱基识别质量是衡量碱基正确识别的概率（通常以数字值直接表示）。碱
59 基识别质量与碱基识别错误率之间的关系为： $Q = -10 \lg P$ （ Q 为碱基识别质量，
60 P 为碱基识别错误率）。 $Q=20$ 代表碱基识别正确率 $\geq 99\%$ ； $Q=30$ 代表碱基识别正
61 确率 $\geq 99.9\%$ 。高通量测序仪器应能自动判读碱基识别质量。

62 三、技术流程

63 全基因组测序的一般流程包括：测序样本的获得、测序文库的构建、全基
64 因组测序和数据分析等。

65 1. 测序样本的获得

66 全基因组测序主要用于待测微生物的核酸序列测定。待测微生物应进行
67 分离纯化，以获得生长状态稳定的纯培养物，可参考“微生物鉴定指导原则”
68 （通则9204）。分离纯化后的纯培养物应采用适宜的方法，可参考“细菌DNA
69 特征序列鉴定法”（通则1021），获得浓度、纯度和完整性良好的基因组测序
70 样本。

71 2. 测序文库的构建

72 测序文库是指将基因组样本随机打断后，在其两端加入特定接头序列
73 （adapters），并经过大规模平行扩增，形成的DNA片段集合。测序文库中样本
74 的核酸浓度、纯度、片段的大小分布等因素，都会影响测序输出的数据量和碱
75 基识别质量。应对构建的测序文库进行纯化、定量、均一化处理，使文库中各
76 待测样本的浓度保持均等；必要时，采用凝胶电泳或毛细管电泳等方法检测
77 文库的质量。

78 3. 全基因组测序

79 将测序文库中的待测样本固定在测序介质中，通过特定接头序列，将测
80 序引物与待测核酸序列进行结合。加入底物脱氧核糖核苷酸，在DNA聚合酶作
81 用下，使结合在待测核酸序列上的测序引物进行延伸，并利用信号收集器采
82 集信号，包括但不限于光信号、电信号或离子信号等，通过信号分析软件对采
83 集到的信号进行分析，获得待测样本的碱基序列信息，以及物理通量、有效通
84 量、测序读长、测序深度、碱基识别质量等参数。

85 4. 数据分析

86 采用适宜的序列分析方法和软件，对得到的核酸测序下机数据进行序列
87 拼接，最终获得待测微生物样本的全基因组序列信息。

88 四、影响测序结果的主要因素

89 1. 待测样本核酸质量

90 应采用适宜的方法提取待测样本的基因组DNA，并保证提取的基因组DNA
91 在适宜的浓度和纯度范围内，无蛋白、多糖等污染。一般情况下，核酸浓度宜
92 不低于10 ng/ μ l， A_{260}/A_{280} 比值宜在1.8~2.0之间。核酸浓度较低，或发生降
93 解等导致质量不佳的情况，可导致基因组DNA片段化不完全，影响文库质量，
94 进而影响测序深度和测序结果。

95 2. 测序文库质量

96 应对测序文库进行质量控制。当测序文库中包含多个待测样本时，不同
97 样本的核酸浓度应基本一致，保证测序后的输出数据量均匀稳定。推荐采用
98 荧光分析法定量检测不同样本的基因组DNA浓度，测序文库制备完成后，采用
99 适宜的稀释倍数，确定上机测序文库的浓度。

100 3. 测序深度

101 测序深度是指待测样本中某个指定核苷酸被检测的次数。一般高通量测
102 序仪器输出的测序深度指待测样本基因组序列中核苷酸被检测次数的平均值。
103 测序深度与基因组覆盖率之间是正相关，测序深度越大，重复测序次数越多，
104 待测样本基因组覆盖率越大，测序带来的错误率也会随着测序深度的提高而
105 降低。一般而言，基因组测序深度应不少于50倍；建立全基因组序列参考数据
106 库时，测序深度应不少于100倍。

107 4. 碱基识别质量

108 碱基识别质量是评价测序结果准确率的重要因素。根据核酸测序仪器的
109 正常运行参数，单个样本的核酸测序的结果应保证 $Q_{20} \geq 80\%$ 或 $Q_{30} \geq 70\%$ ；也
110 即测序数据中80%及以上的碱基正确率大于99%，或者70%及以上的碱基正确率
111 大于99.9%。

112 五、方法学考察

113 除考察影响测序结果的主要因素，包括：待测样本核酸质量、测序文库质

114 量、测序深度、碱基识别质量等，还应进行相应的分析方法学考察；可在测序
115 过程中增加已知序列的参考品，评估测序仪器性能，以保证全基因组测序结
116 果的准确性和重现性。

117 六、应用指导

118 微生物全基因组序列能够提供全面丰富的遗传信息，通过全基因组序列
119 的比对分析，可以实现待测微生物，包括：标准菌株、模式菌株、质控菌株、
120 生产检定用菌（毒）种、益生菌等，以及从药用原料、辅料、制药用水、中间
121 产品、终产品、包装材料和环境等中检出污染微生物等的精准鉴定、溯源分析
122 以及风险评估等。

123 精准鉴定

124 当基于常规生化筛选、表型和基因型鉴定方法无法获得待测微生物样本
125 准确的鉴定信息时，可利用全基因组测序技术获得更加精准的鉴定结果或遗
126 传变异信息等。全基因组序列分析还对研究微生物的系统进化具有重要价值，
127 有助于新种或亚种的发现和遗传分类单元的系统发育解析，提高对新种或亚
128 种的生物学认识。

129 溯源分析

130 当出现无菌试验结果阳性、培养基灌装等模拟工艺失败、生产过程严重
131 异常事件时，如常规基因型鉴定方法无法提供足够的分辨力，可在获得菌种
132 鉴定信息的基础上，采用全基因组测序技术对目标微生物以及相关环节中分
133 离的同种微生物进行全基因组序列的同源性分析，结合污染调查信息，实现
134 目标微生物的溯源分析

135 风险评估

136 全基因组序列包含了微生物菌株全部的遗传信息，基于全基因组数据分
137 析还能够用于毒力、耐药以及其他基因的功能分析与表型预测，为开展微生
138 物的风险评估分析提供参考依据。

起草单位：上海市食品药品检验研究院 联系电话：1800677839

复核单位：中国食品药品检定研究院、天津市药品检定研究院、辽宁省药品检验检测院

参与单位：浙江现代生物技术发展中心、中国工业微生物菌种保藏中心