

SZTT/SZGIA

深圳市团体标准

SZTT /SZGIA 1-2016

基于高通量测序的环境微生物检测 第1部分：基本规程

Detection of environmental microorganisms based on high-throughput sequencing

Part 1 : Basic code

2016-10-18 发布

2016-10-19 实施

深圳基因产学研资联盟 发布

前 言

本标准编写格式遵循了 GB/T 1.1-2009 给出的规则编写。

本标准的附录A-B都是资料性附录。

本标准由深圳基因产学研资联盟提出并归口。

本标准负责起草单位：深圳华大基因研究院、深圳市谱元基因研究院、深圳华大基因科技有限公司、蓝色彩虹（深圳）科技有限公司、华大精准营养（深圳）科技有限公司、深圳弘睿康生物科技有限公司、深圳微健康基因科技有限公司、深圳市恒创基因科技有限公司。

本标准主要起草人：肖亮、覃俊杰、谢强、杜佳婷、李倩一、刘靓、李陶莎、陈蕾、胡鹏、戴文魁、尹三军。

本标准为首次发布。

基于高通量测序的微生物检测

第1部分：基本规程

1 范围

本标准规定了采用高通量测序技术进行环境微生物检测的方法，包括样本的采集、保存，实验室工作条件，样本处理程序、质量控制及关键分析步骤。

本标准适用于利用宏基因组学方法进行环境样品中微生物检测的机构。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 6682 分析实验室用水规格和实验方法

GB 19489 实验室生物安全通用要求

GB/T 29859 生物信息学术语

GB/T 30989-2014 高通量基因测序技术规程

SN/T 1193 基因检验实验室技术要求

3 术语和定义

下列定义适用于本文件。

3.1

环境 environment

包括人体(动物)环境与自然环境，前者包括但不限于口腔、皮肤、生殖道，肠道等，后者包括但不限于水体、空气，土壤等。

3.2

微生物 microorganism

本标准中定义为古细菌、细菌、真菌。

3.3

宏基因组学 metagenomics

又称环境微生物基因组学、元基因组学。以直接从环境样品中提取全部微生物的DNA为研究对象,揭示环境样品所包含的全部微生物的遗传组成及其群落功能。

3.4

相对丰度 relative abundance

某一特定种类微生物在环境总微生物群落中所占的相对比例，通常以百分比表示。

3.5

稳定剂 stabilizer

保持微生物核酸物质稳定的化学试剂。

3.6

错误率 error rate

基于高通量测序的特征，其输出的结果可能存在错误。错误率即表示测序结果中某一碱基可能被测错的概率。

3.7

检测限 detection limit

本标准中定义为在给定的可靠程度内，可从样品中检测到目标物质（物种或序列）的最小测序数据量。

4 缩略语

下列缩略语适合于本文件。

DNA——脱氧核糖核酸（deoxyribonucleic acid）

5 检测工作程序

5.1 原理

运用高通量测序技术对环境样品中微生物的总体遗传物质进行测序，然后通过测序结果与现有数据库中序列的比对，来检测样品中所含的微生物的种类和相对丰度。

5.2 检测对象

存在于环境样本中的微生物。

5.3 样品采集

——应避免样品被其它物质污染。不同环境，不同类型的样品，应采取不同的采样方法以保证样品中微生物群落的代表性和准确性。

——样品采集应经过伦理审查和生物安全评价审查。

5.4 样品保存和运输

宜在干冰条件下运输，并于-80℃保存，不应在室温下放置超过2小时。保存过程中应避免反复冻融。如需常温保存或运输，宜将样品浸润于稳定剂中。

5.5 实验室工作程序

5.5.1 实验室条件及所需仪器设备

- 应具备符合 GB19489-2008 实验室生物安全通用要求和 SN/T 1193 基因检验实验室技术要求的实验室。
- 实验室用水规格应符合 GB/T 6682 分析实验室用水规格和实验方法的规定。
- 实验室洁净等级至少应达到 30 万级。
- 实验室应按照不同工作内容划分独立区域。如试剂储存和准备区, 样本制备区, 扩增区等, 各区域间应避免交叉污染。
- 实验室温度范围应控制在 20~25 摄氏度。

5.5.2 样本接收

核对并录入编号, 确认样本信息, 检查样本保温状态 (填充的保温物质是否仍有残留), 检查样品本身状态 (有无泄漏, 污染)。记录相关信息, 并保证样品的可追溯性。

5.5.3 样本中微生物总 DNA 提取

- 使用稳定可重复的方法提取微生物总 DNA, 宜使用相关试剂盒。
- 提取过程中应避免核酸污染及降解。
- 提取的总 DNA 质量, 如 DNA 总量、浓度、片段大小、纯度, 应达到相应高通量测序平台的最低要求。其中人体 (动物) 样品所包含的人体 (动物) DNA 含量占总 DNA 的比例, 应符合相应部位的正常范围。(各部位的标准有待具体定)
- 提取后的 DNA, 如保存 7 天以内, 可置于-20℃条件保存; 如超过 7 天应保存于-80℃条件下; 宜避免反复冻融。

5.5.4 建库 (可选) 与高通量测序

按照不同测序仪配套的待测序核酸样本质量要求、标准文库制备方法和质控方法进行核酸样本质控并按照规定的使用量取样, 以及进行文库制备和质控 (如有)。将质控完成的待测序文库进行测序。过程参见附录A。

5.6 数据分析

5.6.1 数据处理

- 质量值: 应达到相应高通量测序平台的要求, 并由此计算有效的测序数据量, 错误率高于 1% 的碱基比例应低于 20%。
- 数据量: 有效的测序数据量应达到声明目标物种检测限。

5.6.2 微生物物种注释

各机构应对所使用的注释方法和数据库版本号进行描述, 以满足可重复性的要求。若无特殊方法, 对于已知物种而言, 宜使用数据库中序列相似度应达到95%以上, 覆盖度达到90%以上的序列物种信息进行注释。

5.6.3 序列拼接与组装 (如有)

各机构应对所使用的组装方法 (包括但不限于软件, 参数, 流程) 进行描述, 以满足可重复性要求, 并保证组装结果序列的质量满足与数据库进行比对和注释的要求。参考数据库参见附录B。

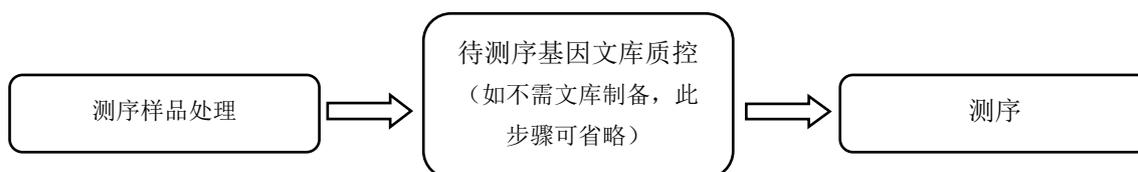
5.6.4 微生物物种相对丰度计算（如有）

各机构应对所使用的相对丰度计算方法进行描述，满足可重复性要求且所得计算结果应尽可能反映目标物种在环境样本中实际的相对丰度。如无特殊方法，宜将有效的测序数据比对到组装好的基因集上（相似度95%以上，覆盖度90%以上），得到基因水平上的相对丰度分布情况。然后将注释到同一物种的基因序列相对丰度进行叠加，得到该物种的相对丰度结果。

附 录 A
(资料性附录)
高通量基因测序过程

A.1 高通量基因测序过程

完成高通量基因测序，按照如下流程进行：



图A.1 高通量基因测序流程图

A.2 测序样本处理

对待测序的生物样本进行核酸提取，并对提取的核酸进行质控检测。

A.3 高通量基因测序文库制备

A.3.1 DNA样品打断

对预处理好的DNA样品，采用物理或者化学的方法将待测序的核酸DNA打断到符合测序策略要求的片段大小。

A.3.2 接头连接

将打断后的DNA片段制备成两端连有接头（含有与扩增及测序引物互补的序列），中间有短序列标签结构的文库或者标签文库。

A.3.3 PCR（可选）

将接头连接后的文库进行PCR扩增放大得到待测序文库或者标签文库。

A.4 质量控制

将待测序文库或者标签文库进行质量控制检测。宜按照琼脂糖凝胶电泳方法或者Q-PCR方法进行文库质控，建库过程中宜设置阴性对照与阳性对照。

附 录 B
(资料性附录)
参考数据库

B 1 美国能源部联合基因组研究所 (DOE-JGI) 的细菌核糖体RNA基因序列数据库

<http://greengenes.secondgenome.com>

B 2 美国国立生物技术信息中心 (NCBI) 的核酸序列数据库

<ftp://ftp.ncbi.nlm.nih.gov/blast/db/>

B 3 美国国立生物技术信息中心 (NCBI) 的细菌基因组数据库

<http://www.ncbi.nlm.nih.gov/genome>

B 4 美国国立生物技术信息中心 (NCBI) 的物种分类数据库

<http://www.ncbi.nlm.nih.gov/taxonomy>

B 5 美国能源部联合基因组研究中心 (DOE-JGI) 的微生物基因组数据库

<https://img.jgi.doe.gov/>

参 考 文 献

Marco, Diana, ed. *Metagenomics: current innovations and future trends*. Horizon Scientific Press, 2011.
