

## · 微生物组测序与分析专题 ·

**朱宝利** 中国科学院微生物研究所研究员，中国科学院大学医学院岗位教授，博士生导师，国家 973 项目首席科学家，病原微生物耐药与耐药基因组学北京市重点实验室主任，中国生物工程学会微生物组学与技术专业委员会主任，《生物工程学报》编委。主要从事病原微生物基因组学、人体肠道微生物组学和免疫基因组学研究。研究成果包括确认细菌耐药谱与农用抗生素使用相关及 *MCR-1* 耐药基因在全球范围内的分布和传播状态等。



## 微生物组测序与分析专家共识

段云峰<sup>1</sup>，王升跃<sup>2</sup>，陈禹保<sup>3</sup>，杨瑞馥<sup>4</sup>，李后开<sup>5</sup>，朱怀球<sup>6</sup>，童贻刚<sup>7</sup>，杜文斌<sup>1</sup>，付钰<sup>1</sup>，胡松年<sup>1</sup>，王军<sup>1</sup>，辛玉华<sup>1</sup>，赵方庆<sup>8</sup>，鲍一明<sup>9</sup>，张雯<sup>10</sup>，李娟<sup>10</sup>，曾明<sup>11</sup>，牛海涛<sup>12</sup>，周欣<sup>13</sup>，李岩<sup>14</sup>，崔生辉<sup>15</sup>，袁静<sup>16</sup>，李俊桦<sup>17</sup>，王加义<sup>14</sup>，刘东来<sup>18</sup>，倪铭<sup>19</sup>，孙青<sup>14</sup>，邓晔<sup>20</sup>，朱宝利<sup>1</sup>

- 1 中国科学院微生物研究所，北京 100101
- 2 上海交通大学医学院附属瑞金医院国家转化医学中心，上海 200025
- 3 北京市计算中心，北京 100094
- 4 中国人民解放军军事医学科学院微生物流行病学研究所，北京 100012
- 5 上海中医药大学，上海 201210
- 6 北京大学，北京 100871
- 7 北京化工大学，北京 100029
- 8 中国科学院北京生命科学研究院，北京 100101
- 9 中国科学院北京基因组研究所（国家生物信息中心），北京 100101
- 10 中国疾病预防控制中心传染病预防控制所，北京 102206
- 11 中国食品药品检定研究院-疫苗室，北京 102629
- 12 北京协和医学院医学实验动物研究所，北京 100021
- 13 中国农业大学，北京 100083
- 14 中国科学院生物物理研究所，北京 100101
- 15 中国食品药品检定研究院食品化妆品检验检测中心，北京 102629
- 16 首都儿科研究所，北京 100020
- 17 深圳华大生命科学研究院，广东 深圳 518083

**Received:** June 26, 2020; **Accepted:** November 19, 2020

**Supported by:** Major State Basic Research Development Program (No. 2018YFC1603903).

**Corresponding author:** Baoli Zhu. Tel/Fax: +86-10-64807362; E-mail: zhubaoli@im.ac.cn  
国家重点研发计划重点专项项目 (No. 2018YFC1603903) 资助。

- 18 中国食品药品检定研究院诊断试剂检定所, 北京 102629  
19 中国人民解放军军事科学院军事医学研究院, 北京 100850  
20 中国科学院生态环境研究中心, 北京 100085

段云峰, 王升跃, 陈禹保, 等. 微生物组测序与分析专家共识. 生物工程学报, 2020, 36(12): 2516–2524.

Duan YF, Wang SY, Chen YB, et al. Expert consensus on microbiome sequencing and analysis. Chin J Biotech, 2020, 36(12): 2516–2524.

**摘要:** 在过去的十几年, 微生物组相关研究和应用持续升温。微生物组逐渐成为生命科学、环境科学和医学等领域的研究焦点。与此同时, 全球多个国家和组织也都积极发起各自的微生物组计划, 进行多方面的布局, 力争在这一具有广阔前景的领域获得战略地位。此外, 无论是科研还是产业应用已经迎来了研究高潮和投融资热潮, 微生物组相关产品和服务也不断出现。然而, 行业在快速发展的同时, 也存在一些不足。由于微生物组测序和分析相关技术和方法发展迅速, 各国研究和应用尚未在技术、方案和数据等标准上达成统一, 国内行业参与者对微生物组也存在认识不足, 对微生物组相关新方法、新技术、新理论等还未能充分掌握和使用。除此之外, 已有的一些标准和指南, 内容过于简单, 实操性也不足, 这不仅给科研数据的整合造成了困难和资源浪费, 还给相关企业进行不良竞争、以次充好提供了机会。更重要的是, 我国尚缺乏微生物组相关的国家标准, 国家微生物组计划仍处于筹备过程。在此背景下, 中国生物工程学会、中国科学院微生物研究所于2019年6月至2020年3月, 共同设立了“微生物组测序与分析专家共识”专项研究课题。中国生物工程学会组织了微生物组相关领域的27位专家以及来自行业内的30多位专业人员, 通过分成4个项目小组、召开4轮研讨会后, 最终形成了涵盖从微生物采集与保存、DNA提取与建库、高通量基因测序和数据分析以及质控标准品等全流程的“微生物组测序与分析专家共识”。本专家共识具有较强可参考性和可操作性, 不仅能指导国内科研和产业机构规范进行微生物组相关产、学、研, 还能为国家相关职能部门提供可参考的技术依据, 保障规模化和规范化的企业利益, 加强行业自律, 避免不规范的企业扰乱市场, 最终促进微生物组相关产业的良性发展。

**关键词:** 微生物组, 高通量基因测序, 专家共识, 国家标准, 微生物组计划

## Expert consensus on microbiome sequencing and analysis

Yunfeng Duan<sup>1</sup>, Shengyue Wang<sup>2</sup>, Yubao Chen<sup>3</sup>, Ruifu Yang<sup>4</sup>, Houkai Li<sup>5</sup>, Huaiqiu Zhu<sup>6</sup>, Yigang Tong<sup>7</sup>, Wenbin Du<sup>1</sup>, Yu Fu<sup>1</sup>, Songnian Hu<sup>1</sup>, Jun Wang<sup>1</sup>, Yuhua Xin<sup>1</sup>, Fangqing Zhao<sup>8</sup>, Yiming Bao<sup>9</sup>, Wen Zhang<sup>10</sup>, Juan Li<sup>10</sup>, Ming Zeng<sup>11</sup>, Haitao Niu<sup>12</sup>, Xin Zhou<sup>13</sup>, Yan Li<sup>14</sup>, Shenghui Cui<sup>15</sup>, Jing Yuan<sup>16</sup>, Junhua Li<sup>17</sup>, Jiayi Wang<sup>14</sup>, Donglai Liu<sup>18</sup>, Ming Ni<sup>19</sup>, Qing Sun<sup>14</sup>, Ye Deng<sup>20</sup>, and Baoli Zhu<sup>1</sup>

1 Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

2 National Center for Translational Medicine, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China

3 Beijing Computing Center, Beijing 100094, China

4 Institute of Microbiology and Epidemiology, Chinese Academy of Military Medical Sciences, Beijing 100012, China

5 Shanghai University of Traditional Chinese Medicine, Shanghai 201210, China

6 Peking University, Beijing 100871, China

7 Beijing University of Chemical Technology, Beijing 100029, China

8 Beijing Institutes of Biological Sciences, Chinese Academy of Sciences, Beijing 100101, China

9 Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation, Beijing 100101, China

10 National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

11 China Institute of Food and Drug Control-Vaccine Room, Beijing 102629, China

12 Institute of Medical Laboratory Animals, Peking Union Medical College, Beijing 100021, China

13 China Agricultural University, Beijing 100083, China

14 *Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China*

15 *Food and Cosmetic Inspection and Testing Center, China Institute of Food and Drug Control, Beijing 102629, China*

16 *Capital Institute of Pediatrics, Beijing 100020, China*

17 *Beijing Genomics Institute at Shenzhen, Shenzhen 518083, Guangdong, China*

18 *Institute of Diagnostic Reagents, China Institute for Food and Drug Control, Beijing 102629, China*

19 *Academy of Military Medicine, Chinese Academy of Military Sciences, Beijing 100850, China*

20 *Research Center for Ecological Environment, Chinese Academy of Sciences, Beijing 100085, China*

**Abstract:** In the past ten years, the research and application of microbiome has continued to increase. The microbiome has gradually become the research focus in the fields of life science, environmental science, and medicine. Meanwhile, many countries and organizations around the world are launching their own microbiome projects and conducting a multi-faceted layout, striving to gain a strategic position in this promising field. In addition, whether it is scientific research or industrial applications, there has been a climax of research and a wave of investment and financing, accordingly, products and services related to the microbiome are constantly emerging. However, due to the rapid development of microbiome sequencing and analysis related technologies and methods, the research and application from various countries have not yet unified on the standards of technology, programs, and data. Domestic industry participants also have insufficient understanding of the microbiome. New methods, technologies, and theories have not yet been fully accepted and used. In addition, some of the existing standards and guidelines are too general with poor practicality. This not only causes obstacles in the integration of scientific research data and waste of resources, but also gives related companies unfair competition opportunity. More importantly, China still lacks national standards related to the microbiome, and the national microbiome project is still in the process of preparation. In this context, the experts and practitioners of the microbiome worked together and developed the consensus of experts. It can not only guide domestic scientific research and industrial institutions to regulate the production, learning and research of the microbiome, the application can also provide reference technical basis for the relevant national functional departments, protect the scale and standardized corporate company's interests, strengthen industry self-discipline, avoid unregulated enterprises from disrupting the market, and ultimately promote the benign development of microbiome-related industries.

**Keywords:** microbiome, high-throughput gene sequencing, expert consensus, national standards, microbiome project

近几年,随着微生物组相关技术和方法的快速发展,人体、动植物、水体、土壤以及特殊环境等领域的微生物组研究也获得了大量突破性成果。特别是在人体微生物组领域,美国分别于2007年和2014年启动了人类微生物计划(Human Microbiome Project, HMP)和人类微生物组整合计划(Integrated Human Microbiome Project, iHMP),不仅揭示了人体微生物组的基本特征,还揭示了微生物组与怀孕、早产、炎症性肠道疾病以及糖尿病等疾病之间的关系<sup>[1-2]</sup>。2016年5月13日,美国科技政策办公室又宣布启动国家微生物组计划(National Microbiome Initiative, NMI),关注不同生态系统的微生物组,推动微生物组研究成果在健康保健、食品生产及环境恢复等领域的应用。

欧盟在人类微生物组领域的布局与美国几乎

同步,2008年,欧盟在FP7计划下启动了人类肠道元基因组(MetaHIT)计划,旨在建立人类肠道微生物基因的参考目录,开发储存、组织和解释这些信息的生物信息学工具,并进一步确定不同人群的菌群基因特征<sup>[3]</sup>。2012年,欧盟又启动了MetaHIT的延续性项目MetaGenoPolis(MGP),关注人类微生物组与心脑血管疾病、能量平衡及脑发育和功能、非传染性慢性疾病的关系研究,旨在利用定量和功能元基因组工具,研究人类肠道微生物组对健康和疾病的影响机制<sup>[4-11]</sup>。

中国也在积极布局微生物组相关研究,中国科学院率先启动了“中国科学院微生物组计划”。2017年,世界微生物数据和中国科学院微生物研究所牵头,联合全球12个国家微生物资源保藏中心共同发起了全球微生物模式菌株基因组和微生物组测序合作计划。

截至目前,除美国、欧盟之外,日本、加拿大、爱尔兰、法国、澳大利亚、韩国等都实施了各自的微生物组计划,中国的国家微生物组计划也在筹备中。

在全球各国加紧布局微生物组科研和产业应用的形势下,微生物组相关的标准显得尤为重要。欧盟曾于2011年启动国际人类微生物组标准(The International Human Microbiome Standards, IHMS)项目,由12个国家、8个合作伙伴和15个参与者组成,总计花费230万欧元,旨在开发标准操作流程,优化人类微生物组领域的的数据质量和可比性<sup>[12]</sup>。2012年,欧盟的MGP项目也建立了4个科学平台以及1个伦理中心,重点关注粪便样本标准化管理、元基因组高通量测序、元基因组数据挖掘及元基因组功能性研究<sup>[13]</sup>。2013年,美国启动了微生物组质量控制计划(The Microbiome Quality Control project, MBQC),目的是推进人类微生物组研究的标准体系建设<sup>[14]</sup>。

在产业应用方面,微生物组相关的应用也在持续跟进,一些检测产品或服务也在逐步上市,相关投融资也很活跃。同时,在市场潜力与应用前景方面,利用微生物组预防、诊断和治疗疾病等领域已经成为企业争夺的技术高地,具有广阔的发展空间和可期的增长趋势。

美国在微生物组领域发展速度最快,投融资频繁、技术先进、企业分布密集。国内也有数十家企业涉及微生物组学上下游的产品研发和应用。产业方向涉及微生物制药、生物合成微生物疗法、基因编辑、抗菌药物研发、粪菌移植、环境治理、生态修复等多个领域。

无论是科研还是产业应用,国际各国都在优先布局微生物组、夺取科技战略高地的情况下,中国并没有落后,微生物组相关的研究论文数量已经比肩欧美,并且在很多方面还取得了令人瞩目的成绩。在微生物组领域,中国有望弯道超车,占据微生物组科研和应用的全球领先地位。

在这样的大背景下,中国迫切需要建立自己

的微生物组相关标准。虽然,在微生物组的标准方面,国内已经有一些地方标准、行业标准和企业标准,但是仍然缺乏统一的国家标准。此外,由于微生物组学与技术发展迅速,技术和方法的更新和迭代都非常快,中国现有的微生物组相关技术标准和指导方案,已经不能很好地满足国内飞速发展的微生物组学相关研究和产业应用需求。

为了帮助全国相关组织、团体以及企业在微生物组研究和应用方面获得全面、权威、实操性强的指导;解决数据格式标准和技术方案不统一;促进微生物组相关产业良性发展,保障规模型和规范企业的利益;加强行业自律,避免不规范的企业扰乱市场,对行业造成不良影响,中国生物工程学会联合其他企事业单位,决定成立“微生物组测序与分析专家共识”专项研究课题(以下简称“专家共识”)。目的是为将来的行业共识和国家标准打下基础,以期指导中国的微生物组科研和临床研究与应用,为国家相关职能部门以及中国的国家微生物组计划提供可参考的技术依据,为形成中国自己的微生物组相关的标准体系奠定基础。本研究课题将重点围绕微生物组测序与分析技术,包括微生物组学相关的样本采集、保存和测序技术、数据分析和生物信息学分析技术的应用,标准品的设计和使用规范等,涉及微生物组研究的各个环节。

“专家共识”采用专家研讨会的形式,邀请微生物组相关领域的专家、技术负责人、科研一线技术人员等共同参与讨论。4个小组经过4轮研讨,最终形成了“专家共识”,具体的形成过程及共识的细则详见《生物工程学报》网络版附录文件。

## 1 微生物组样本采集及处理流程

### 1.1 样本

为了便于来自不同实验室数据的相互交流,有必要对样本按照一定的规范进行统一规范的命名。

### 1.2 样本采集、保存和转运

样本采集、保存和转运方法的选择对样本有

较大影响<sup>[15]</sup>。推荐从新鲜采集的样本中提取微生物基因组 DNA；否则，应将采集的样本立即冷冻和储存，也可使用为特定样本类型设计的保存液及取样套装<sup>[16]</sup>。

### 1.2.1 样本的采集

微生物组样本采集应制定标准化的工作流程。在同一项目研究中需确保采集方法的一致性。避免样本受到外源性污染和获取足够的样本量。

### 1.2.2 样本保存和转运

采集的新鲜样本立即放入液氮中快速冷冻，之后转移至-80℃进行长期保存，或在干冰保存条件下进行转运。如果缺乏低温保存条件，则建议采用微生物样本保存剂进行及时保存<sup>[17]</sup>。样品转运过程应保持低温条件，并最好以 DNA 或 cDNA 状态进行转运。尽管不同保存方法、保存时间所引入的数据变化小于个体间的差异，但因保存方法不一致所产生的数据误差仍应引起重视<sup>[16]</sup>。

## 1.3 微生物样本核酸提取与质控

### 1.3.1 样本前处理

在微生物核酸提取前可移除保护剂，并对样本中残留保存剂进行清洗置换。

### 1.3.2 样本的处理

尽可能确保样本内微生物的组成不发生改变、减少样本在空气中的暴露时间、避免采样过程发生外源性微生物污染。

## 1.4 核酸提取

不同提取方法会影响核酸提取的效率和质量，并对下游基因测序数据产生较大影响。因此，核酸提取需注意考虑不同种类微生物裂解的等效性；确保同一项目使用相同的核酸提取试剂盒。对于低生物量的样本，应格外注意试剂保存、配置和提取过程中的污染问题；核酸纯化过程中尽可能去除抑制剂；避免外源 RNase 污染造成的 RNA 质量不佳；对于含难破壁微生物的样本，推荐使用研磨珠进行研磨<sup>[15]</sup>。

### 1.4.1 核酸质控标准

核酸质控包含浓度质控、纯度质控、片段完整性质控。采用分光光度检测和凝胶电泳可评估核酸浓度、纯度和片段完整性<sup>[15]</sup>。

### 1.4.2 提取后核酸的保存方法

经抽提后的核酸样本需进行分装，单独取用，可在-20℃短暂保存，-80℃长期保存，避免反复冻融。

## 1.5 文库构建

文库构建是高通量测序的重要步骤。文库成败极大依赖于核酸的量及降解情况，质量过低会导致扩增效率、文库得率和多样性降低。

### 1.5.1 建库方法

扩增子测序：细菌及古菌扩增子测序时，需要对 16S rRNA V1-V9 区段进行扩增。目前，通常使用 V3-V4 区段扩增。真菌扩增子测序多采用 18S rRNA 或 ITS 1-2 区段进行扩增。也可以针对细菌、古菌或真菌的目标基因设计特异性引物进行扩增。需注意微量扩增时是否有背景微生物污染引起的假阳性扩增，严格把控实验操作及使用试剂中可能引入的微生物污染是保障测序结果准确性的关键。

元基因组和全基因组建库：尽量采用 PCR-Free 的实验方法，或尽量减少 PCR 循环数。元基因组和全基因组文库片段大小宜在 300-500 bp (进行长读序测序时可更长)，为避免文库交叉污染，可使用双端唯一标签 (Unique Dual Index, UDI) 接头。

RNA 建库：为了在测序过程中确保根据单一链来源的转录本分析结果，应构建链特异性文库。建库前先去 rRNA 后再进行文库构建。

### 1.5.2 文库质量控制

高通量测序常用的质量控制方法包括：Qubit 检测、毛细管电泳、qPCR 和超微量核酸检测仪等。文库质控包括：文库转化率及文库浓度，经 qPCR 检测的合格文库浓度应不低于 2 nmol/L；文库片段分布分析。

### 1.5.3 单分子测序建库

对于单分子测序的建库方法和质控要求,如 PacBio 和纳米孔测序等,请参考网络版“专家共识细则”中的详细内容。

## 2 微生物组测序

微生物基因测序需按照文库分子的长短及丰度来选择合适的测序策略。

### 2.1 测序技术基本要求

微生物组基因测序流程包括测序文库质控及标准制定、测序技术与平台选择、下机数据质量控制及管理 3 个方面。同时,必须保证基因测序所需要的运行环境,保证测序数据的准确性与可重复性。

#### 2.1.1 实验室环境要求

文库制备及质控实验室应达到标准分子生物学实验要求。参考《高通量测序技术临床检测规范化应用北京专家共识(第一版通用部分)》<sup>[18]</sup>。基因测序实验室需严格控制室内环境,包括空间大小、温度、湿度、负压及洁净度等应符合相应要求。

#### 2.1.2 试剂管理要求

可优先采购与测序平台兼容的试剂产品,按照所有质控、测序试剂产品说明书的要求存放,保证文库质控及测序的稳定性及可靠性。临检实验室应优先国家药品监督管理局批准的试剂。具体要求可参考《高通量测序技术临床检测规范化应用北京专家共识(第一版通用部分)》<sup>[18]</sup>。

### 2.2 测序文库处理及质控

在上机测序前,需要对实验预设方案的文库进行质量控制,保证下机数据产出的数量和质量。

#### 2.2.1 高通量测序文库质控

微生物组高通量测序应用广泛,基于所选测序平台及检测项目情况,根据研究方向、样本类型制定合适的质控标准,在此基础上可适当选择合适的质控工具。

#### 2.2.2 单分子测序文库质控

1) 基因组测序文库:单分子基因组测序超长

读序的特点,文库大小的控制要求比较严格,请遵照对应平台的质控标准进行操作。

2) 转录组测序文库:单分子转录组测序是微生物研究的创新方法。质控时应该充分考虑研究方向及测序方法,以便采取合适的质控标准。

3) 扩增子测序文库:单分子扩增子测序可直接对原核 16S rRNA、18S/ITS rRNA 等长目标序列进行测序。

### 2.3 微生物组基因测序技术

#### 2.3.1 测序技术简介

根据原理不同,测序技术共经历了 Sanger 测序、高通量测序、单分子实时测序等阶段。

#### 2.3.2 测序技术及应用

微生物组学研究选择合适的技术手段及平台至关重要。但不同的测序技术对微生物样本、文库特点及质量控制、测序方法及数据质控均不相同,因此,根据组学研究方向和目的,合理选择测序技术及研究方法是行业规范的趋势。

#### 2.3.3 测序平台与应用选择

微生物组测序的技术及平台较多,根据测序目的的不同,可以应用在多个领域。

### 2.4 测序数据的质量控制、管理

#### 2.4.1 质量控制原则

数据的质量检查、生物信息学分析、输出格式、储存等方面都需要标准操作程序(Standard Operating Procedure, SOP)支持。

#### 2.4.2 不同测序平台的质控标准

测序数据从测序仪产出后,注意数据总数量和有效数据量,质控包括接头序列、引物序列、低质量序列和冗余序列评估和去除,参考基因组或目标区域序列比对等步骤。

#### 2.4.3 下机数据的管理规范

测序实验室应规范测序数据命名方式、储存流程及传输规范。数据储存应采用通用的 FASTQ、BAM、BCL 等格式,便于数据交换及实验室间对比评价。

## 2.5 实验流程标准化管理

### 2.5.1 标准操作程序

建立基因测序 SOP 应包括：实验室环境、试剂管理和准备、文库质检、上机测序和数据评估与处理等环节。

### 2.5.2 人员资质要求

实验室人员应具备分子生物学实验背景，或经过相关技能培训获得 PCR 上岗证。建立起系统的培训管理体系，确保在接触真实样本之前对人员进行严格的技能和规范培训。

## 3 微生物组数据分析流程

### 3.1 扩增子测序数据分析

#### 3.1.1 数据质控

在进行正式分析前，要先对下机数据进行质控。高通量测序平台，临床样品测序数据可依据《临床分子病理实验室二代基因测序检测专家共识》<sup>[19]</sup>和《二代测序技术应用于临床肿瘤精准医学诊断的共识》中的质量指标<sup>[20-21]</sup>。

#### 3.1.2 物种注释

在 16S/18S rRNA/ITS 基因扩增子测序分析中，常用的分析策略是对数据过滤后的高质量读序进行聚类生成操作分类单元 (Operational taxonomic units, OTUs)，包括从头聚类 (*De novo*) 和有参聚类 (Closed-reference)。扩增子序列变异 (Amplicon Sequence Variants, ASVs) 在敏感性和特异性方面优于 OTUs 方法<sup>[22]</sup>。获得 OTUs/ASVs 之后的物种注释过程，可以使用不同的参考数据库。目前，3 个常用的细菌和古菌 16S rRNA 基因数据库是：Silva、RDP 和 Greengene。真菌常用的公共数据库有：核糖体内部转录间隔区 (Internal Transcribed Spacer, ITS) 数据库 UNITE、18S rRNA 基因数据库有 Silva。

#### 3.1.3 多样性分析

扩增子测序分析的主要目的之一是识别特定微生物组的多样性，包括  $\alpha$  多样性和  $\beta$  多样性。常用于衡量  $\alpha$  多样性的指数包括 Richness、Chao1、

Shannon、Simpson 和 Faith's PD 指数等。常用于衡量  $\beta$  多样性的指标是距离矩阵，包括 Euclidean metric、Bray Curtis、Jaccard 和 Unweight/weighted UniFrac 距离等。

#### 3.1.4 其他分析

功能分析：可以通过物种信息对微生物组功能进行预测。差异检验分析：样本的组间差异检验主要应用 R 软件包 (如 Vegan) 或 LEfSe 等。相互作用分析：主要涉及环境因子对微生物组的影响。可应用 R CCA/RDA 或 R Vegan 等软件。网络分析：网络分析图形展示可使用 Cytoscape 等软件。

### 3.2 元基因组测序数据分析

元基因组测序可以在全基因组的水平获得微生物组的种群结构、物种多样性、进化关系、功能基因、代谢途径、微生物间相互协作关系及微生物群落与环境之间的关系等多个方面的信息。

#### 3.2.1 序列质控

质控过程主要包括：移除低质量读序、过短读序、接头序列等，还需要去除非微生物的基因组序列。考虑到宿主基因组的干扰和检测的灵敏性，对于人类相关的微生物组样本，可参考《宏基因组分析和诊断技术在急危重症感染应用的专家共识》的意见<sup>[23]</sup>。

#### 3.2.2 元基因组组装

物种注释可选择基于序列组装 (Assembly-based) 或免组装 (Assembly-free) 两种方法。对于人类肠道样本，可采用免组装，而特殊环境样本或新物种，组装后注释的可信度更高。

#### 3.2.3 物种注释

根据不同数据库类型，元基因组的物种注释也被分为基于标记基因 (Marker-based) 的注释和基于微生物全基因组 (Whole genomes) 的注释。

#### 3.2.4 Binning (可选择)

元基因组 Binning 过程是将组装后的众多 Contig 或 Scaffold 按特征进行聚类。可反向获取装配 Contigs 的相应读序进行再组装，从而组装成更长的 Contig/Scaffold 或单一物种基因组草图。

### 3.2.5 基因注释和功能注释

基因注释：元基因组组装后，可应用从头预测的软件包 MetaGeneMark、Prodigal 或 Metagene 等。功能注释：可采用 Diamond，对比蛋白质数据库 eggNOG 和 KEGG 等，其他微生物组的功能数据库还包含抗生素耐药性数据库 (CARD) 和碳水化合物活性酶数据库 (CAZy) 等。

### 3.3 单菌基因组测序数据分析

对分离纯化培养的单菌进行 DNA 提取、文库构建、测序和全基因组组装。主要包括：序列质控、全基因组组装、基因组质量评估、基因注释和功能分析等。

#### 3.3.1 序列质控

常用的高通量基因测序平台的序列和数据过滤策略可参照扩增子测序部分，并移除连续低质量碱基(Q<20)数大于 40%的读序。

#### 3.3.2 全基因组组装

近年来，结合单分子基因测序平台长读序和高通量基因测序高精度的优点，进行“三代加二代”的全基因组组装。

#### 3.3.3 基因组质量参数

组装后的单菌全基因组应进行标准的质量评估，建议遵从美国能源部联合基因组研究所 (DOE JGI) 发布的定义不明微生物基因组的 MISAG 标准。

#### 3.3.4 基因注释和功能分析

细菌编码基因的预测可应用 Glimmer、Prodigal 和 GeneMarkS 等软件。功能预测可应用 HUMAnN2 或 BLAST+。

## 4 检测报告

高通量测序分析结果应提供：测序原始数据和相应的分析报告。报告的主体应包括数据信息、分析方法、分析结果和报告解读。

## 5 参考品与质控品

基于高通量测序平台展开的微生物组研究普

遍缺乏对于减少定量偏差的针对性设计与质控环节，因而，有必要使用标准化的参考品对微生物组测序服务的平台进行评估，并使用标准化的质控品对生产过程及检测过程进行校准<sup>[24]</sup>。

## 6 质量控制

实验室应建立相应的质量管理体系。工作人员均应接受持续培训和能力评估。更新产品内容时，需应用参考品、质控品或标准品进行流程的参数设定和结果验证，判断增加/修改检测内容后结果的准确性和稳定性，以达到检测报告的要求。定期参加室间质评 (或其他行业评估项目)，紧跟行业发展，及时更新行业知识，持续优化流程。

## 7 数据存储

应用结构化的数据库来有效管理原始数据、结果数据和报告。微生物组序列数据的储存应采用通用格式和压缩方式。保存完整的存储、读写和移除日志文件，保证报告的可重复性。

## 8 隐私保护

个人生物样本常常涉及病人的隐私，具有特殊性和敏感性。可采取匿名化或编码处理，保证个人信息不被泄露。同时，从存储、访问和管理等角度设计安全措施。

## 9 知情同意与伦理

临床微生物组样本进行测序分析前，应签署知情同意书。按照我国相关法规和伦理委员会的要求及时递交其他审查文件。

## REFERENCES

- [1] Turnbaugh PJ, Ley RE, Hamady M, et al. The human microbiome project. *Nature*, 2007, 449(7164): 804–810.
- [2] The Integrative HMP (iHMP) Research Network



- Consortium. The integrative human microbiome project. *Nature*, 2019, 569(7758): 641–648.
- [3] Le Chatelier E, Nielsen T, Qin JJ, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature*, 2013, 500(7464): 541–546.
- [4] Routy B, Le Chatelier E, Derosa L, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science*, 2018, 359(6371): 91–97.
- [5] Aron-Wisnewsky J, Prifti E, Belda E, et al. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut*, 2019, 68(1): 70–82.
- [6] Wang XF, Yang ST, Li SH, et al. Aberrant gut microbiota alters host metabolome and impacts renal failure in humans and rodents. *Gut*, 2020, 69(12): 2131–2142.
- [7] Brial F, Le Lay A, Dumas ME, et al. Implication of gut microbiota metabolites in cardiovascular and metabolic diseases. *Cell Mol Life Sci*, 2018, 75(21): 3977–3990.
- [8] Hirschberg S, Gisevius B, Duscha A, et al. Implications of diet and the gut microbiome in neuroinflammatory and neurodegenerative diseases. *Int J Mol Sci*, 2019, 20(12): 3109.
- [9] Kolde R, Franzosa EA, Rahnavard G, et al. Host genetic variation and its microbiome interactions within the human microbiome project. *Genome Med*, 2018, 10(1): 6.
- [10] Blekhman R, Goodrich JK, Huang K, et al. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol*, 2015, 16(1): 191.
- [11] Pennisi E. Meet the psychobiome. *Science*, 2020, 368(6491): 570–573.
- [12] Gan YL, Tao H, Guan JH, et al. iHMS: a database integrating human histone modification data across developmental stages and tissues. *BMC Bioinformatics*, 2017, 18: 103.
- [13] The International Human Microbiome Standards (IHMS)[EB/OL]. (2011). <http://mgps.eu/standard-operating-procedure/>.
- [14] Sinha R, Abnet CC, White O, et al. The microbiome quality control project: baseline study design and future directions. *Genome Biol*, 2015, 16: 276.
- [15] Costea I, Zeller G, Sunagawa S, et al. Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol*, 2017, 35(11): 1069–1076.
- [16] Ma JL, Sheng LL, Hong Y, et al. Variations of gut microbiome profile under different storage conditions and preservation periods: a multi-dimensional evaluation. *Front Microbiol*, 2020, 11: 972.
- [17] Prakash O, Nimonkar Y, Desai D. A recent overview of microbes and microbiome preservation. *Indian J Microbiol*, 2020, 60(3): 297–309.
- [18] 北京市临床检验中心, 北京医学会检验医学分会, 首都医科大学临床检验诊断学系, 等. 高通量测序技术临床检测规范化应用北京专家共识(第一版通用部分). *中华医学杂志*, 2019, 99(43): 3393–3397.
- [19] 《临床分子病理实验室二代基因测序检测专家共识》编写组. 临床分子病理实验室二代基因测序检测专家共识. *中华病理学杂志*, 2017, 46(003): 145–148.
- [20] 中国肿瘤驱动基因分析联盟(CAGC), 中国临床肿瘤学会(CSCO). 二代测序(NGS)技术应用于临床肿瘤精准医学诊断的共识, 第一版(试行). (2016.04.23). <http://v.gcbi.com.cn/?p=166>.
- [21] Gevers D, Pop M, Schloss PD, et al. Bioinformatics for the human microbiome project. *PLoS Comput Biol*, 2012, 8(11): e1002779.
- [22] Xue ZY, Kable ME, Marco ML. Impact of DNA sequencing and analysis methods on 16S rRNA gene bacterial community analysis of dairy products. *mSphere*, 2018, 3(5): e00410–18.
- [23] 宏基因组分析和诊断技术在急危重症感染应用专家共识组. 宏基因组分析和诊断技术在急危重症感染应用的专家共识. *中华急诊医学杂志*, 2019, 28(2): 151–155.
- [24] Hardwick SA, Chen WY, Wong T, et al. Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. *Nat Commun*, 2018, 9(1): 3096.

(本文责编 郝丽芳)