

PCR 反应条件对 16S rRNA 基因扩增子测序准确性的影响

赵枫^{1,2}, 史亚², 王莹², 梁倩², 陈欢^{1,2}, 李樱红³, 窦晓兵^{*1}, 何陆平^{*2}

1 浙江中医药大学生命科学院, 浙江 杭州 310053

2 杭州微数生物科技有限公司, 浙江 杭州 311215

3 浙江省食品药品检验研究院 国家药品监督管理局药品微生物检测与预警重点实验室, 浙江 杭州 310052

赵枫, 史亚, 王莹, 梁倩, 陈欢, 李樱红, 窦晓兵, 何陆平. PCR 反应条件对 16S rRNA 基因扩增子测序准确性的影响[J]. 微生物学通报, 2022, 49(7): 2527-2537

Zhao Feng, Shi Ya, Wang Ying, Liang Qian, Chen Huan, Li Yinghong, Dou Xiaobing, He Luping. Effect of PCR conditions on the accuracy of 16S rRNA gene amplicon sequencing[J]. Microbiology China, 2022, 49(7): 2527-2537

摘要:【背景】16S rRNA 基因扩增子测序技术是一种不依赖培养而获得样本中细菌种群结构、相对丰度等信息的方法。高通量测序技术实验步骤较多, 每一步骤细微的差别都可能在最终的测序结果中放大, 并造成测序结果与实际情况的偏差。【目的】基于 MiSeq 测序平台, 探讨 PCR 反应体系中扩增引物序列、退火温度、模板起始量、扩增循环数和变性时间等 5 个因素对 16S rRNA 基因测序结果的影响。【方法】对 mock DNA 的 16S rRNA 基因扩增子进行测序, 分别分析不同的扩增引物、退火温度、模板起始量、循环数和变性时间对数据准确性的影响。【结果】不同的扩增引物对检测结果有较大的影响, 采用的 4 组引物中, 引物 B (V3-V4, 341F/806R) 的准确性最好, 引物 A (V3-V4, 341F/805R) 次之。比较不同退火温度(52、55 和 60 °C)对检测准确性的影响, 退火温度 60 °C 的结果最接近理论值。模板起始量(2、10 和 50 ng)的检测结果显示, mock DNA 起始量为 2 ng 的结果准确性最高。相较于其他 3 组(15+18、25+8 和 30+8), 循环数为(20+8)的检测结果最接近 mock DNA 的理论值。不同变性时间(3 min 和 5 min)对分析结果无显著性影响。【结论】引物序列、退火温度、模板起始量和循环数是影响 16S rRNA 基因测序结果准确性的重要因素。

关键词: 16S rRNA 基因; 反应条件; 靶向测序

基金项目: 中国科技部科技基础资源调查专项(2021FY100901)

Supported by: Special Project of Science and Technology Basic Resources Survey of China Ministry of Science and Technology (2021FY100901)

*Corresponding authors: E-mail: DOU Xiaobing: xbdou77@163.com; HE Luping: hlpjoke@163.com

Received: 2021-10-15; Accepted: 2021-12-30; Published online: 2022-02-07

Effect of PCR conditions on the accuracy of 16S rRNA gene amplicon sequencing

ZHAO Feng^{1,2}, SHI Ya², WANG Ying², LIANG Qian², CHEN Huan^{1,2}, LI Yinghong³,
DOU Xiaobing^{*1}, HE Luping^{*2}

1 School of Life Sciences, Zhejiang Chinese Medical University, Hangzhou 310053, Zhejiang, China

2 Hangzhou Digital-Micro Biotech Limited Company, Hangzhou 311215, Zhejiang, China

3 NMPA Key Laboratory for Testing and Risk Warning of Pharmaceutical Microbiology, Zhejiang Institute for Food and Drug Control, Hangzhou 310052, Zhejiang, China

Abstract: [Background] 16S rRNA gene amplicon sequencing, a culture-independent approach, is commonly used to identify the bacterial community structure and relative abundance in given samples. The protocol of high-throughput sequencing technology involves many experimental steps, and the nuances of each step may be magnified in the final sequencing results, resulting in unexpected deviations from the actual situation. [Objective] With MiSeq system, we evaluated the impact of five factors in PCR on 16S rRNA gene amplicon sequencing results: primer sequence, annealing temperature, template quantity, cycle number, and denaturation time. [Methods] The 16S rRNA gene amplicon of mock community DNA was sequenced to determine the effects of these five factors on the accuracy of the results achieved. [Results] Primer had great influence on the 16S rRNA sequencing results, and primer B (V3–V4, 341F/806R) showed the highest quantification accuracy among all primers, followed by primer A (V3–V4, 341F/805R). Among the different annealing temperatures (52, 55 and 60 °C), the result of annealing at 60 °C was closest to the theoretical value. As for the DNA template quantity (2, 10 and 50 ng), the deviation was minimal when 2 ng DNA template was used. The result of detection with (20+8) cycles was closest to the theoretical value of mock DNA compared with that of the other three groups (15+18, 25+8, and 30+8). Denaturation time (3 min and 5 min) had no significant effect on the experimental results. [Conclusion] This study revealed that primer sequence, annealing temperature, template quantity, and number of cycles were important factors affecting the accuracy of 16S rRNA gene sequencing results, which laid a foundation for the selection of standardized methods of library construction.

Keywords: 16S rRNA gene; reaction conditions; targeted next-generation sequencing

基于系统进化分子标记的 16S rRNA 基因扩增子测序技术, 通过提取特定环境中的微生物总 DNA 和构建文库, 可以获得大量微生物种类、丰度等相关信息。相较于传统微生物研究方法的分离、培养, 16S rRNA 基因扩增子测序技术更能获得大量的不可培养微生物的信息, 是目前研究微生物群落的重要技术手段^[1]。近年来, 16S rRNA 基因扩增子测序技术在农业、工业、

环境、食品和卫生等领域得到了广泛的运用^[2-4]。

高通量测序技术实验环节步骤较多, 每一步骤细微的差别都可能在最终的测序结果中放大, 并造成测序结果、分析结果与实际情况的偏差。文献报道的误差和偏差来源主要包括采样、提取效率、样本存储条件、扩增参数、模板起始浓度、模板性质(GC 含量和二级结构)、引物错配、聚合酶错误、嵌合体、随机错误及

测序错误等^[5-8]。在 PCR 扩增过程中,存在多种引起 PCR 产物成分变化的因素,如不同引物对 DNA 模板的偏好性不同、退火温度改变引起扩增效率的变化等。因此,运用高通量测序技术分析微生物群落结构特征时,有必要对各个实验环节的条件进行筛选优化。按照微生物组学质量控制计划(Microbiome Quality Control Project, MBQCP)的建议^[9],本研究以模拟细菌群落 DNA(mock DNA)为研究对象,将已知细菌 DNA 进行定量混合,系统全面地评估引物序列、退火温度、模板起始量、循环数及变性时间对 16S rRNA 基因扩增子测序技术一致性和准确性的影响,筛选尽可能真实反映微生物群落的 PCR 反应参数,以期对 16S rRNA 基因扩增子测序技术分析微生物多样性提供有效的参考和建议。

1 材料与方法

1.1 材料

1.1.1 菌种

13 种细菌均采购自美国菌种保藏中心(American Type Culture Collection, ATCC),并通过

16S rRNA 基因进行菌种鉴定,菌种信息见表 1。

1.1.2 主要试剂和仪器

DNA 提取试剂盒, Qiagen 公司; PCR 引物, 上海捷瑞生物工程有限公司; 2×KAPA HiFi HotStart ReadyMix, KAPA Biosystems 公司; 磁珠, Beckman 公司。基因扩增仪、NanoDrop 2000 超微量分光光度计、荧光计, Thermo Fisher 公司; MiSeq PE300 高通量测序仪, Illumina 公司。

1.2 DNA 的提取

细菌在 35 °C 培养 24–48 h 后用 DNA 提取试剂盒提取每种细菌基因组 DNA。荧光计测定各细菌基因组 DNA 浓度,取不等量各菌株基因组 DNA 混合成 mock DNA,菌株的 16S rRNA 基因拷贝数=[DNA 质量/(基因组分子量×摩尔系数)]×每基因组 16S rRNA 基因拷贝数,所占比例=菌株 16S rRNA 基因拷贝数/所有菌株 16S rRNA 基因拷贝数总和(表 1)。

1.3 文库的构建及测序

建库方法采用两步法 PCR。第一步 PCR 对细菌 16S rRNA 基因的 V3–V5 区进行扩增。扩增引物的 3'端与 16S rRNA 基因匹配,5'端的通

表 1 mock DNA 菌种信息

Table 1 Strains information of mock DNA

菌种名称 Strain name	16S rRNA 基因拷贝数 16S rRNA gene copies	dsDNA 分子量 dsDNA molecular weight	DNA 质量 DNA quality (ng)	所占比例(理论值) Proportion (theoretical value, %)
<i>Bacillus cereus</i>	13	3 343 160 354	2.02	35.15
<i>Stenotrophomonas maltophilia</i>	4	2 998 302 813	3.92	23.48
<i>Staphylococcus aureus</i>	6	1 742 835 154	0.87	13.48
<i>Enterococcus faecalis</i>	5	1 988 020 140	0.78	8.85
<i>Streptococcus mutans</i>	6	1 255 876 752	0.22	4.73
<i>Deinococcus radiodurans</i>	4	1 637 006 459	0.39	4.30
<i>Acinetobacter baumannii</i>	6	2 678 611 494	0.39	3.95
<i>Brevundimonas diminuta</i>	2	2 092 323 336	0.39	1.68
<i>Micrococcus luteus</i>	1	1 546 003 127	0.57	1.66
<i>Pseudomonas aeruginosa</i>	4	3 871 812 507	0.20	0.91
<i>Salmonella enterica enterica</i>	7	2 971 604 759	0.08	0.83
<i>Escherichia coli</i>	7	3 031 848 776	0.08	0.81
<i>Curtobacterium citreum</i>	1	2 134 431 388	0.08	0.17

用序列为第二步 PCR 的引物结合位点。第一步 PCR 反应体系(25 μ L): 2 \times KAPA HiFi HotStart ReadyMix 12.5 μ L, 上、下游引物(10 μ mol/L)各 1 μ L, 模板(10 ng) 1 μ L, 无核酸酶水补足至 25 μ L。PCR 反应条件: 95 $^{\circ}$ C 3 min; 95 $^{\circ}$ C 30 s, 55 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 30 s, 25 次循环; 72 $^{\circ}$ C 5 min。反应结束后用磁珠纯化。第二步 PCR 对第一步 PCR 产物进行扩增, 扩增引物的 5'端至 3'端包含测序接头序列、barcode 序列及与第一步 PCR 产物结合序列。第二步 PCR 反应体系(50 μ L): 2 \times KAPA HiFi HotStart ReadyMix 25 μ L, 上、下游引物(10 μ mol/L)各 2 μ L, 纯化产物 5 μ L, 灭菌超纯水补足至 50 μ L。PCR 反应条件: 95 $^{\circ}$ C 3 min; 95 $^{\circ}$ C 30 s, 55 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 30 s, 8 次循环; 72 $^{\circ}$ C 5 min。反应结束后用磁珠纯化。

以上为默认建库方法, 在此基础上调整引物(表 2)、退火温度(52、55 和 60 $^{\circ}$ C)、模板起始量(2、10 和 50 ng)、第一步 PCR 循环数+第二步 PCR 循环数(15+18、20+8、25+8 和 30+8)、变性时间(3 min 和 5 min)等因素, 每组实验设置 5 个重复。荧光计进行文库定量, 等量文库混合, 采用 MiSeq PE300 平台测序。

表 2 本研究所用引物

Table 2 Primers used in this study

引物 Primer	扩增区域 Amplified region	位置 Location	引物序列 Primer sequence (5'→3')	产物长度 PCR product (bp)	熔解温度 T_m ($^{\circ}$ C)
A	V3-V4	341F	F: CCTACGGGNGGCWGCAG	464	59.2
		805R	R: GACTACHVGGGTATCTAATCC		46.3
B	V3-V4	341F	F: CCTAYGGGRBGCASCAG	465	55.8
		806R	R: GGACTACNNGGTATCTAAT		44.0
C	V4-V5	515F	F: GTGYCAGCMGCCGCGGTAA	411	67.7
		926R	R: CCGYCAATTYMTTTRAGTTT		53.1
D	V4	515F	F: GTGYCAGCMGCCGCGGTAA	291	67.7
		806R	R: GGACTACNVGGGTWTCTAAT		44.0

Note: N: A/T/C/G; W: A/T; H: A/T/C; V: G/A/C; Y: C/T; R: A/G; B: G/T/C; S: G/C; M: A/C.

1.4 生物信息学分析

对测序结果的序列质量进行质控和过滤, 去除非靶区域序列及嵌合体, 得到优质序列。利用 UPARSE V8.1.1861 软件对所有样本的全部 effective tag 进行聚类, 默认以 97%的一致性(identity)将序列聚类成为 OTU。用 QIIME V1.9.1 软件计算 UniFrac 距离、构建 UPGMA 样本聚类树。解读样本的多样性和丰度。

2 结果与分析

2.1 PCR 引物对测序结果的影响分析

采用 4 组引物分别进行 PCR 扩增, 扩增产物测序结果与 mock DNA 的理论值在 OTU 水平上进行聚类分析, 引物 A、B 和 C 的测量值与理论值接近, 而引物 D 与理论值偏离最远(图 1)。用测量值与理论值之间的均方根偏差(root mean square deviation, rms.)表示 mock DNA 量化的准确度, 引物 A、B 和 C 的准确性较好, 引物 D 的准确性较差(图 2)。

2.2 PCR 退火温度对测序结果的影响分析

退火温度是影响 PCR 反应的重要因素。4 组引物的熔解温度(T_m)值如表 2 所示, 本研究为每组引物均设置了 3 个退火温度(52、55 和

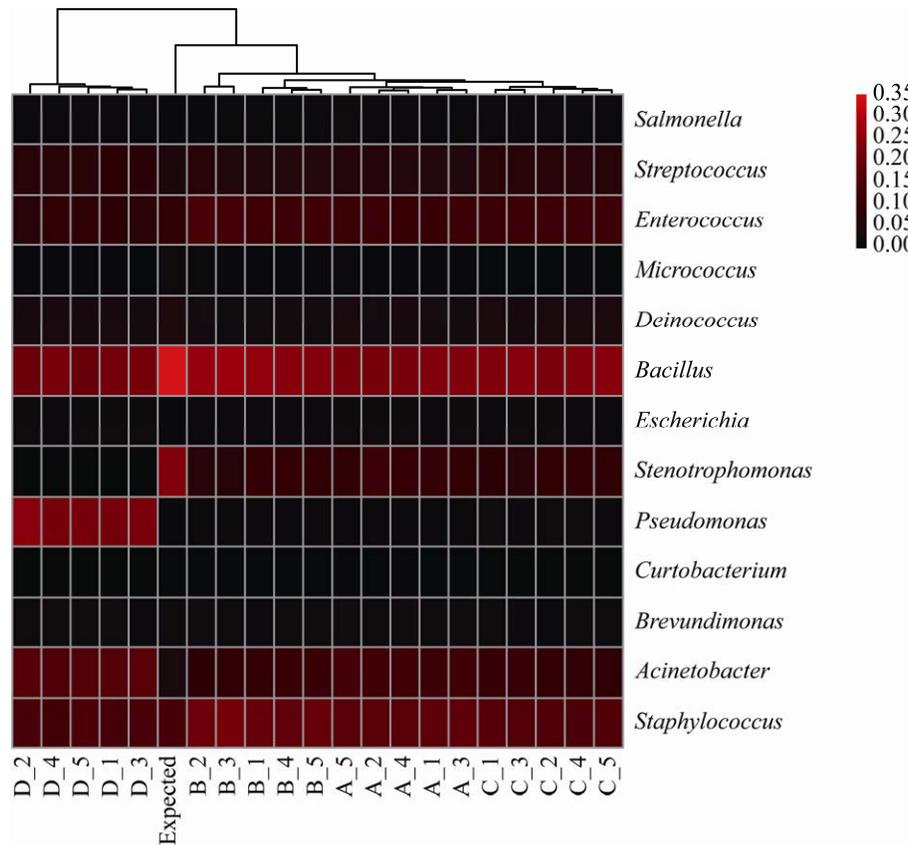


图 1 四组引物在 OTU 水平的聚类分析
Figure 1 Cluster analysis of four pairs of primers at OTU level.

60 °C), 以分析退火温度对测序结果的影响。均方根偏差结果显示, 退火温度对引物 D 的准确值有较大影响, 对其他 3 组引物的影响较小 (图 3)。引物 A 在退火温度为 60 °C 的结果最接

近理论值。因此, 后续优化实验采用引物 A 进行。每个物种丰度结果显示, 4 组引物在不同退火温度下对测量值偏差较大(>2 倍)的物种无明显影响(图 4)。

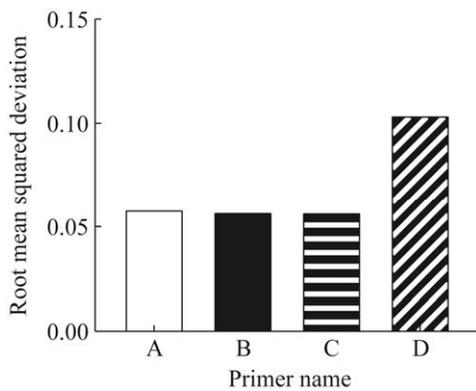


图 2 四组引物的准确性分析
Figure 2 Accuracy analysis of four pairs of primers.

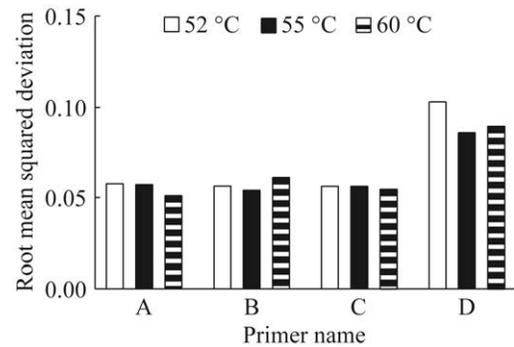


图 3 三种 PCR 退火温度的均方根偏差分析
Figure 3 Root mean square deviation analysis of three PCR annealing temperatures.

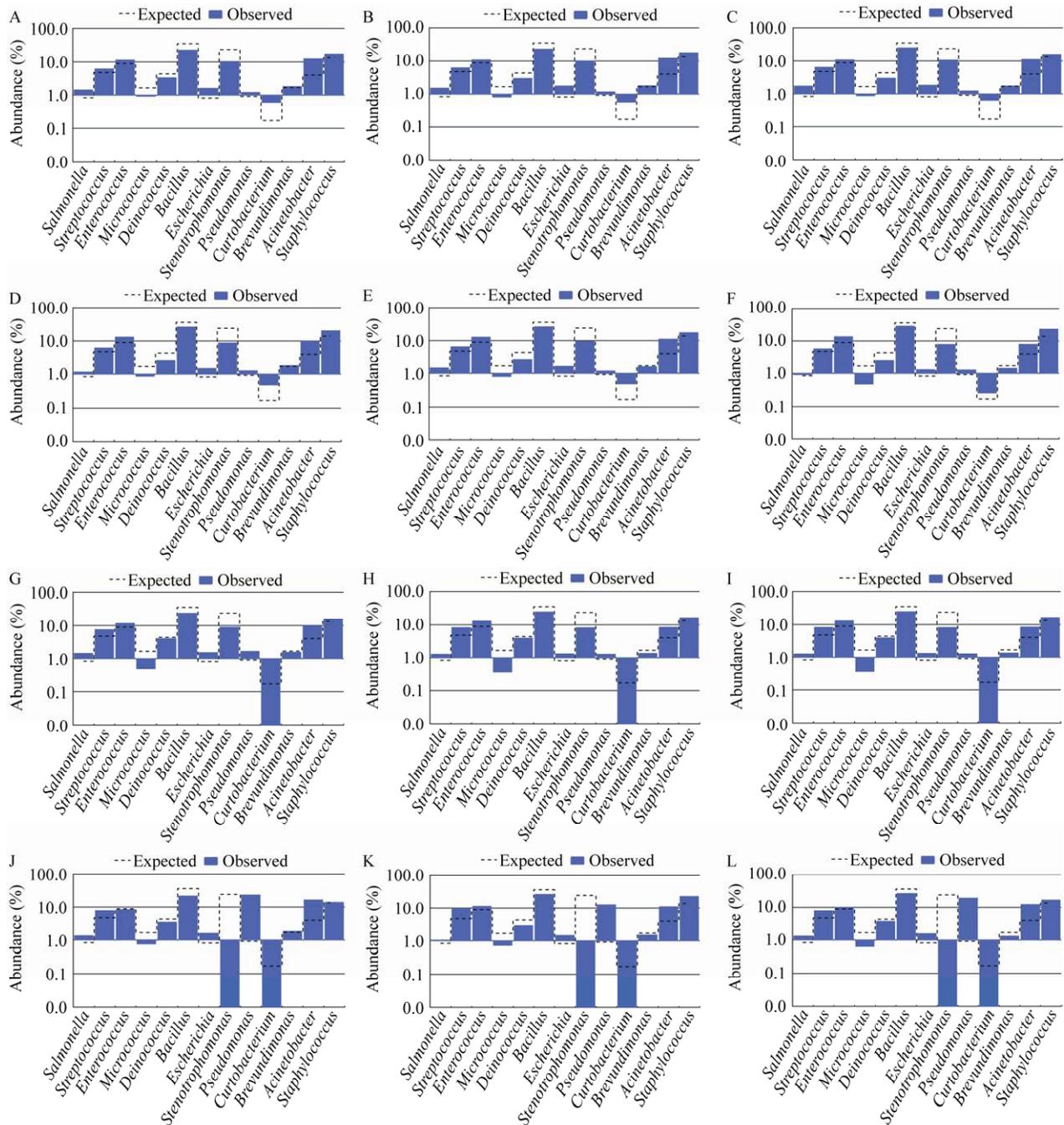


图 4 PCR 退火温度对物种丰度准确性的影响 A–C: 引物 A 退火温度分别为 52、55 和 60 °C 的物种丰度; D–F: 引物 B 退火温度分别为 52、55 和 60 °C 的物种丰度; G–I: 引物 C 退火温度分别为 52、55 和 60 °C 的物种丰度; J–L: 引物 D 退火温度分别为 52、55 和 60 °C 的物种丰度

Figure 4 Effect of annealing temperature on accuracy of species abundance. A–C: The species abundance at annealing temperature of 52, 55 and 60 °C respectively with primer A; D–F: The species abundance at annealing temperature of 52, 55 and 60 °C respectively with primer B; G–I: The species abundance at annealing temperature of 52, 55 and 60 °C respectively with primer C; J–L: The species abundance at annealing temperature of 52, 55 and 60 °C respectively with primer D.

2.3 模板起始量对测序结果的影响分析

用引物 A 分析模板起始量对 mock DNA 测序结果的影响。比较模板起始量(2、10 和 50 ng)的均方根偏差,当模板起始量为 2 ng 时准确性最好(图 5)。不同模板起始量(2、10 和 50 ng)对 mock DNA 各细菌的相对丰度无明显影响,三者均使短小杆菌属和不动杆菌属的测量值与理论值的偏差大于 2 倍(图 6)。

2.4 PCR 循环数对测序结果的影响分析

调整 PCR 循环数,保持 PCR 其他反应参数不变。PCR 进行 20+8、15+18、25+8 和 30+8 次循环(图 7),总循环数分别为 28、33、33 和 38。实验结果显示,总循环数最少的组别(20+8)的测量值最接近 mock DNA 的理论值,准确性最好,25+8、15+18 和 30+8 的准确性较差(图 8)。

2.5 变性时间对测序结果的影响分析

采用 2 个变性时间(3 min 和 5 min)检测其对 mock DNA 测序结果的影响(图 9)。与变性时间为 3 min 的结果相比,5 min 能使寡养单胞菌属的检测值与理论值的偏差减小至 2 倍以下,而对另外两种细菌(不动杆菌属和短小杆菌属)的准确性无明显改善(图 10)。

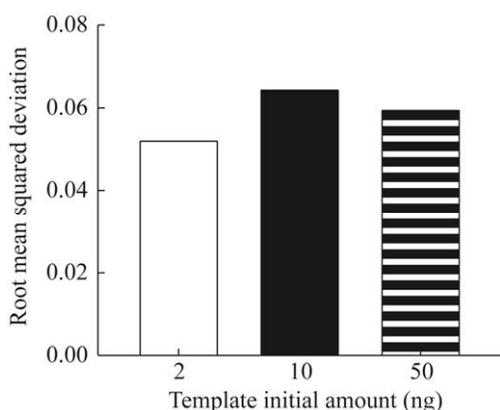


图 5 模板起始量对测序结果的偏差分析
Figure 5 Deviation analysis of template initial amount on sequencing results.

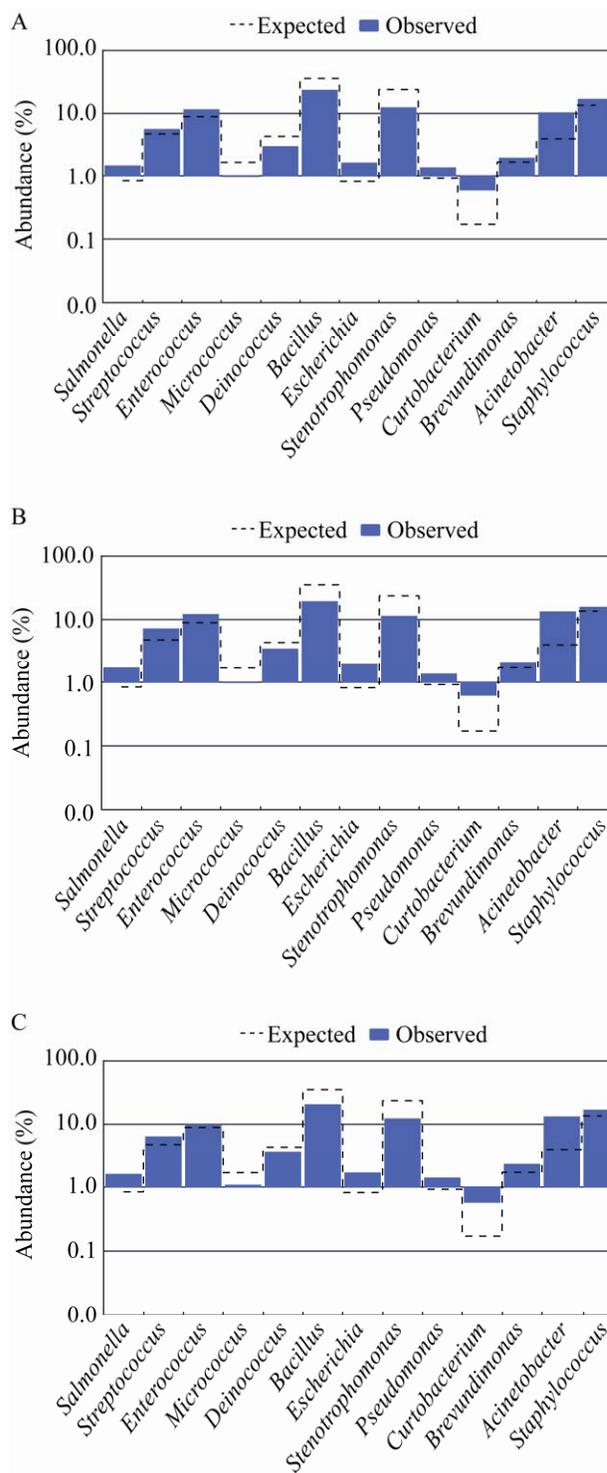


图 6 模板起始量对物种丰度准确性的影响
Figure 6 The effect of template initial amount on the accuracy of species abundance. A: 2 ng; B: 10 ng; C: 50 ng.

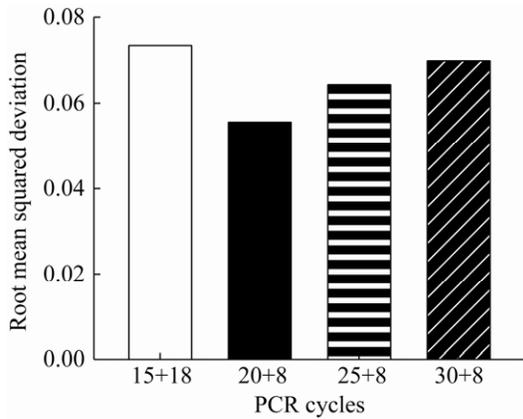


图7 不同 PCR 循环数的测序结果偏差分析
Figure 7 Deviation analysis of sequencing results with different PCR cycles.

3 讨论与结论

随着新一代 DNA 测序技术的出现,人们能同时对多个 DNA 样本的宏基因组进行分析,尤其是以系统进化标记分子的 16S rRNA 基因测序已经成为微生物多样性研究最简捷有效的方法,而 PCR 是基于 16S rRNA 基因测序的基础。在 PCR 扩增过程中,存在多种引起扩增产物成分变化的因素,如引物偏好性、退火温度、酶、变温效率、循环数、模板起始浓度、模板性质(G 含量和二级结构)、引物错配、聚合酶错误、嵌合体及随机错误等^[5,10-12]。

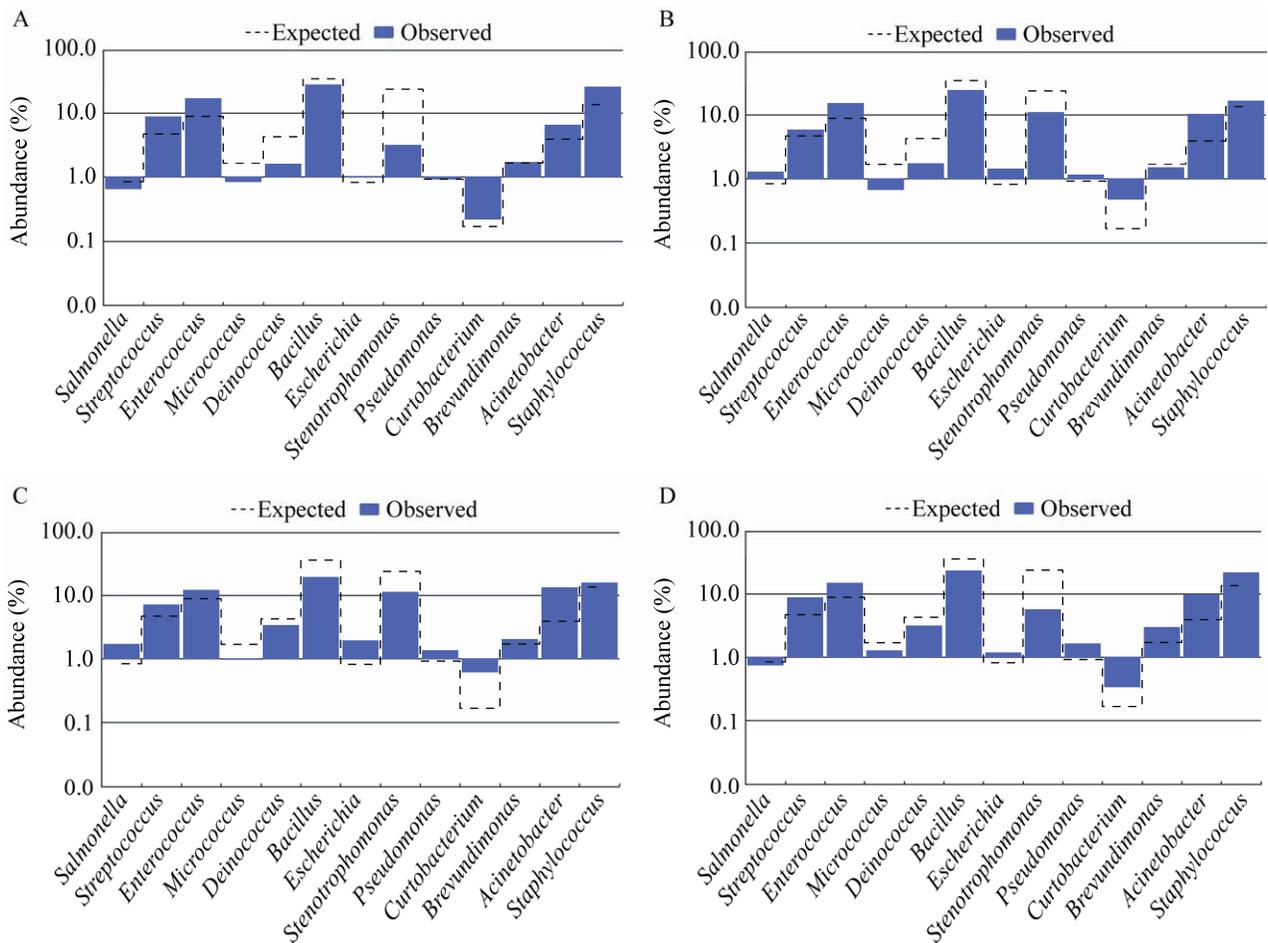


图8 PCR 循环数对物种相对丰度准确性的影响

Figure 8 Effects of PCR cycles on accuracy of species relative abundance. A: 15+18; B: 20+8; C: 25+8; D: 30+8.

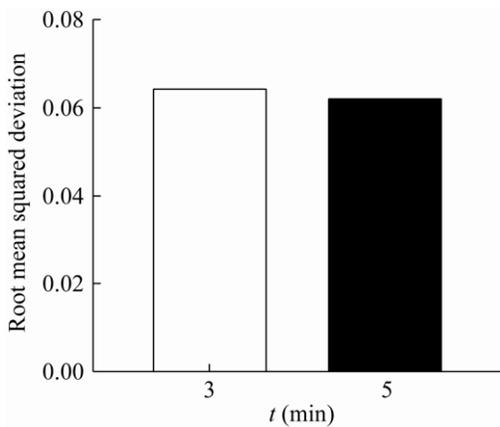


图 9 不同变性时间对测序结果的偏差分析
Figure 9 Deviation analysis of sequencing results with different denaturation time.

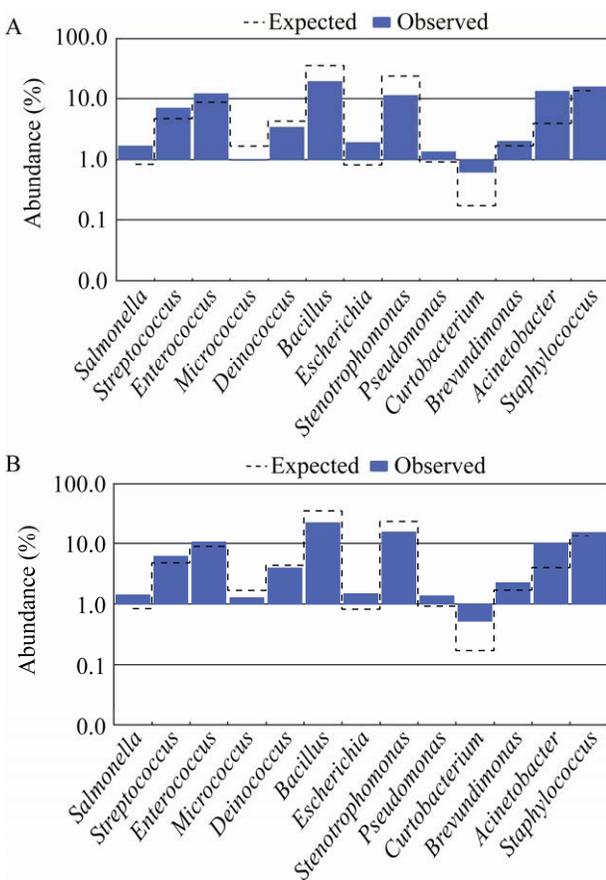


图 10 PCR 变性时间对物种相对丰度准确性的影响

Figure 10 Effects of PCR denaturation time on accuracy of species relative abundance. A: 3 min; B: 5 min.

对这些错误和偏差的认识促进了扩增子文库制备方法的改进。这些方法主要包括是否通过连接添加测序接头、单步 PCR 或多步 PCR, 采用单向还是双向标签, 以及测序引物的选择等^[13]。虽然已经尝试标准化这些方法, 但 PCR 参数对偏差的影响尚未得到系统的研究。本研究考察了基于 PCR 的 16S rRNA 扩增子文库制备方法中的 PCR 参数, 包括引物、退火温度、模板起始量、PCR 循环次数和变性时间对测序结果准确性的影响。

研究表明, 扩增方案和扩增酶的选择对测序结果准确性均有显著影响。两步法 PCR 可以提高 PCR 敏感性, 降低非特异性反应连续放大的可能性, 保证了反应的特异性, 同时保证整个反应的准确性和可行性^[10]。具有校正功能的酶准确度更高, 能显著降低嵌合体的形成, 并且能一定程度上容忍引物和靶标不匹配^[14]。因此, 我们采用校正功能酶及两步法 PCR, 研究 PCR 反应参数对测序结果准确性的影响。

目前二代测序的读长不能覆盖 16S rRNA 基因全长, 需要选择一个有效的高变区(V 区)进行扩增、测序, 但关于 V 区的选择策略尚无公认的准则。16S rRNA 基因不同 V 区会对原核微生物群落结构的分析结果产生明显的影响。一般来说, 可变性和适中的保守性被认为是 V 区选择的标准, V3、V4、V5 和 V6 等区域被广泛应用^[15]。可变区不同则 16S rRNA 基因通用引物不同, 导致微生物群落结构特征分析结果不同。这是因为用于通用引物设计的保守区间在亲缘关系较远的物种之间可能保守性不足, 导致某些种属的细菌扩增效率较差, 从而使微生物群落分析结果出现偏差^[16]。引物偏向性对某些样本影响非常大, 造成对某些种类过低或过高的估计, 甚至有些群体被完全遗漏。V4 区的 515F/806R 在域和门的水平上获得的

多样性最好、得到的测序读长(reads)数量较多、能注释上的比例较高等原因被广泛应用于分析^[17]。通过计算机模拟在 SILVA 数据库中研究了引物的物种覆盖度(taxonomic coverage)和门类覆盖度(phylum spectrum),也推荐 515F/806R 为最好的引物组合^[18]。本研究选取针对 V3-V4、V4-V5 和 V4 区域的 4 组引物,评估了引物选择对 mock DNA 检测结果准确性的影响。然而研究发现,引物 515F/806R 的检测值与 mock DNA 相对丰度的理论值偏差最大,引物 A (V3-V4 区)、B (V3-V4 区)和 C (V4-V5 区)的准确性较好。因此,对不同类型样本的 16S rRNA 基因高变区进行有效性评估,以及研究 V 区和保守引物的选择策略对于原核微生物群落的研究非常必要。

高通量测序过程中,模板的起始量高低会对测序结果造成显著影响^[17],模板量是 PCR 反应取得良好结果的重要保障。一般来说,DNA 模板的适宜用量在一定的范围,在此范围内随着 DNA 量的增加扩增产物越多;但模板过多会导致嵌合体的生成,影响测序结果的准确性。从多种模板中扩增 DNA 片段时,延伸不完整的扩增片段作为引物进行下一个模板的扩增将导致嵌合体的形成。嵌合现象的发生率在 5%–45%之间^[5,9]。本研究结果表明,减少模板起始量能较准确地检测 mock DNA 中各菌的相对丰度。

通过优化起始模板量来最小化 PCR 循环数能有效减少嵌合体的形成,是一种提高结果准确性的有效策略^[2]。本研究对 4 种循环数组合(20+8、15+18、25+8 和 30+8),即总循环数分别为 28、33、33 和 38 的测量值进行比较。总循环数最少的组别(20+8)的测量值最接近 mock DNA 的理论值。随着 PCR 扩增的进行,mock 中各种菌的 DNA 含量较理论值的偏差被逐渐放大,循环数越多被放大得越大。因此,在建

库过程中减少 PCR 循环数有利于测序结果的准确性。

退火温度是影响 PCR 的重要因素,与 PCR 反应的特异性和灵敏度高度相关。退火温度过低可导致非特异性扩增,退火温度过高则影响引物与模板的结合,从而降低 PCR 扩增效率。本研究比较了 4 组引物分别在 3 个退火温度(52、55 和 60 °C)下的测量值,相对于引物种类,不同退火温度的选择对测序结果的准确性影响较小,这与 Gohl 等^[14]的研究结果一致。变性是 PCR 扩增不可或缺的一步,变性时间短则靶序列变性不彻底,易造成扩增失败,极有可能出现假阴性;变性时间过长则损害酶活性。本研究结果显示,变性时间对测序结果无明显影响,其原因可能是 16S rRNA 基因的 GC 含量较一致,变性时间为 3 min 足以打开双链结构。

综上所述,本研究结果认为 16S rRNA 基因扩增子测序中,模板起始量 2 ng、引物 B (V3-V4, 341F/806R)、两步法 PCR (循环数 20+8)及退火温度 60 °C 是测试各组引物和条件中的最优组合。

16S rRNA 扩增子测序技术显著提升了研究复杂环境微生物群组成的能力,但是研究发现其结果准确性受到众多因素的影响,其中 PCR 反应条件是一个重要因素。本研究对不同引物、退火温度、模板起始量、PCR 循环数和变性时间等条件下的扩增子测序分析结果进行了比较,表明引物的选择、模板起始量和 PCR 循环次数是影响微生物群落检测结果的重要因素,我们建议选择合适的引物,并减少模板起始量和 PCR 循环数。本研究为选择标准化方法进行建库奠定了研究基础。

REFERENCES

- [1] Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ,

- Scott KP. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice[J]. *Microbiome*, 2015, 3(1): 26
- [2] 陆晓菊, 官会林, 张正芸, 马永存, 唐蜀昆. 三七连作根际土壤微生物区系的 16S rRNA 系统遗传多样性[J]. *微生物学报*, 2015, 55(2): 205-213
- Lu XJ, Guan HL, Zhang ZY, Ma YC, Tang SK. Microbial distribution and 16S rRNA diversity in the rhizosphere soil of *Panax notoginseng*[J]. *Acta Microbiologica Sinica*, 2015, 55(2): 205-213 (in Chinese)
- [3] 杨娟, 郝志成, 张亚平. 基于 MiSeq 测序分析新疆泥火山土壤细菌群落多样性[J]. *微生物学通报*, 2016, 43(12): 2609-2618
- Yang J, Hao ZC, Zhang YP. Analysis the diversity of soil bacterial community from mud volcano in Xinjiang using MiSeq sequencing[J]. *Microbiology China*, 2016, 43(12): 2609-2618 (in Chinese)
- [4] 王小敏, 徐鹏, 饶传, 王微微, 杨秋丽, 熊久先. 基于 16S rRNA 基因测序技术分析猪链球菌 2 型感染 BALB/c 小鼠粪便样本中微生物的变化[J]. *微生物学报*, 2020, 60(5): 963-971
- Wang XM, Xu P, Rao C, Wang WW, Yang QL, Xiong JX. Effects of *Streptococcus suis* type 2 on the fecal microflora of BALB/c mice based on 16S rRNA high-throughput sequencing[J]. *Acta Microbiologica Sinica*, 2020, 60(5): 963-971 (in Chinese)
- [5] Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform[J]. *Nucleic Acids Research*, 2015, 43(6): e37
- [6] Finotello F, Mastrorilli E, Di Camillo B. Measuring the diversity of the human microbiota with targeted next-generation sequencing[J]. *Briefings in Bioinformatics*, 2018, 19(4): 679-692
- [7] Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM, Consortium UKIG, Satsangi J, Flint HJ, et al. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing[J]. *PLoS One*, 2014, 9(2): e88982
- [8] Nechvatal JM, Ram JL, Basson MD, Namprachan P, Niec SR, Badsha KZ, Matherly LH, Majumdar APN, Kato I. Fecal collection, ambient preservation, and DNA extraction for PCR amplification of bacterial and human markers from human feces[J]. *Journal of Microbiological Methods*, 2008, 72(2): 124-132
- [9] Salipante SJ, Kawashima T, Rosenthal C, Hoogestraat DR, Cummings LA, Sengupta DJ, Harkins TT, Cookson BT, Hoffman NG. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling[J]. *Applied and Environmental Microbiology*, 2014, 80(24): 7583-7591
- [10] D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling[J]. *BMC Genomics*, 2016, 17: 55
- [11] 冯广达, 陈美标, 羊宋贞, 朱红惠. 用于 PCR 扩增的细菌 DNA 提取方法比较[J]. *华南农业大学学报*, 2013, 34(3): 439-442
- Feng GD, Chen MB, Yang SZ, Zhu HH. A comparative study on bacteria DNA extraction methods used for PCR amplification[J]. *Journal of South China Agricultural University*, 2013, 34(3): 439-442 (in Chinese)
- [12] 吴悦妮, 冯凯, 厉舒祯, 王朱珺, 张照婧, 邓晔. 16S/18S/ITS 扩增子高通量测序引物的生物信息学评估和改进[J]. *微生物学通报*, 2020, 47(9): 2897-2912
- Wu YN, Feng K, Li SZ, Wang ZJ, Zhang ZJ, Deng Y. *In-silico* evaluation and improvement on 16S/18S/ITS primers for amplicon high-throughput sequencing[J]. *Microbiology China*, 2020, 47(9): 2897-2912 (in Chinese)
- [13] Zhou HW, Li DF, Tam NFY, Jiang XT, Zhang H, Sheng HF, Qin J, Liu X, Zou F. BIPES, a cost-effective high-throughput method for assessing microbial diversity[J]. *The ISME Journal*, 2011, 5(4): 741-749
- [14] Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies[J]. *Nature Biotechnology*, 2016, 34(9): 942-949
- [15] He Y, Zhou BJ, Deng GH, Jiang XT, Zhang H, Zhou HW. Comparison of microbial diversity determined with the same variable tag sequence extracted from two different PCR amplicons[J]. *BMC Microbiology*, 2013, 13(1): 208
- [16] Ederson CJ, Susilawati E, Smith SL, Wang Q, Chai BL, Farris R, Rodrigues JLM, Thelen KD, Tiedje JM. Bacterial communities in the rhizosphere of biofuel crops grown on marginal lands as evaluated by 16S rRNA gene pyrosequences[J]. *BioEnergy Research*, 2010, 3(1): 20-27
- [17] Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangi JL, Buckler ES, Ley RE. Diversity and heritability of the maize rhizosphere microbiome under field conditions[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(16): 6548-6553
- [18] Klindworth A, Priesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies[J]. *Nucleic Acids Research*, 2013, 41(1): e1