

痕量样本高通量测序文库构建

李博星¹, 陈楠^{1,2}, 陶许诺¹, 陈璐萍¹, 朱宝利¹, 律娜^{1*}

(1. 中国科学院微生物研究所, 北京 100101; 2. 中国科学院遗传与发育生物学研究所, 北京 100101)

摘要: 目的 通过调整建库过程中的测序接头加入量、片段选择的执行与否、文库扩增循环数等条件, 建立一种基于 Illumina 测序平台的高效稳定的痕量样本高通量文库构建方法。**方法** 采用不同类型的样本, 设计不同起始 DNA 量(最低至 0.1 ng)的文库构建策略, 主要对接头使用量、扩增循环数等进行优化, 并通过数据分析评估文库质量。**结果** 对于起始量低至 0.1 ng 的 DNA 样本, 基因组文库构建成功率达 100%。对于基因组较小的细菌, 0.1 ng 的起始 DNA 量的数据组装效果与 100 ng 起始 DNA 量相近; 对于基因组较大、结构复杂的细菌, 0.1 ng 的起始 DNA 量的数据组装效果显著差于 100 ng 起始 DNA 量, 但可通过增加测序数据量的方法提高组装效果。对于元基因组测序, 起始 DNA 量为 1 ng 时, 制备的测序文库与常规起始 DNA 量制备的文库得到的微生物群落结构相近。**结论** 本研究建立的 0.1 ng 起始 DNA 量文库构建方法稳定, 可用于更多微生物基因组测序、元基因组测序的应用领域。

关键词: 高通量测序; 文库构建; 微生物; 痕量

Construction of high throughput sequencing library for trace samples

LI Bo-Xing¹, CHEN Nan^{1,2}, TAO Xu-Nuo¹, CHEN Lu-Ping¹, ZHU Bao-Li¹, LV Na^{1*}

(1. Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;

2. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China)

ABSTRACT: Objective To establish an efficient and stable high-throughput library construction method of trace samples based on Illumina sequencing platform by adjusting the amount of sequencing adapter, the execution of fragment selection, the number of library amplification cycles and so on. **Methods** Different types of samples were used to design the construction strategy of library with different initial DNA amount (as low as 0.1 ng). The usage of adapter and the number of amplification cycles were optimized, and the quality of library was evaluated by data analysis. **Results** The success rate of genomic library construction was 100% for DNA samples with the initial amount as low as 0.1 ng. For bacteria with small genome, the data assembly effect of 0.1 ng initial DNA was similar to that of 100 ng initial DNA. For bacteria with large genome and complex structure, the data assembly effect of 0.1 ng initial DNA was significantly worse than that of 100 ng initial DNA, but the assembly effect could be improved by increasing the amount of sequencing data. For metagenomic sequencing, when the initial DNA quantity was 1 ng, the microbial community structure of the library was similar to that of the library with conventional initial DNA quantity. **Conclusion** The construction method of 0.1 ng initial DNA library established in this study is stable and can be used in more application fields of microbial genome sequencing and metagenome sequencing.

基金项目: 国家重点研发计划项目(2018YFC1603900、2017YFA0505901)

Fund: Supported by the National Key Research and Development Program of China (2018YFC1603900, 2017YFA0505901)

*通信作者: 律娜, 助理研究员, 主要研究方向为分子生物技术应用。E-mail: lvna@im.ac.cn

*Corresponding author: LV Na, Assistant Professor, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. E-mail: lvna@im.ac.cn

KEY WORDS: high throughput sequencing; library construction; microorganism; trace

0 引言

微生物主要包括细菌、真菌、古菌、病毒等微小生物,是自然界分布最广泛、生物多样性最为丰富的生命形式,在农业、工业、食品、医疗卫生等方面发挥着重要作用。近年来,高通量测序技术(next generation sequencing, NGS)由于其通量高、成本低等特点,已被广泛应用于微生物研究的各个领域,如基因组学研究、元基因组学研究等,并取得了广泛的研究成果^[1-4]。除此之外,高通量测序在食品发酵、食品微生物检测、益生菌类保健食品的研发中也发挥了重要的作用,例如借助高通量测序技术手段,研究传统发酵食品泡菜、食醋、奶酪等发酵过程中的质量稳定性与微生物多样性之间的密切关系^[5]。

Illumina 测序平台是目前应用最广泛的 NGS 技术平台^[6],将双链 DNA 经过打断、末端补平处理后,再与特定的 Illumina 测序接头(adapter)段按照正确的方向相连接的过程即为文库制备,得到的 DNA 片段即为测序文库^[7]。测序文库质量高低是能否获得有效数据的关键影响因素之一,是保证 NGS 结果的基础。影响文库质量的因素有很多,如 DNA 质量、文库构建的 DNA 投入量、接头的加入量、文库扩增循环数等^[8-10]。目前,常规 DNA 文库构建的起始量一般要求大于 100 ng^[11]。然而,某些样本非常稀少、取样困难、保存不当等原因会导致无法得到足够的 DNA 用于文库构建。

虽然已有研究对低起始量的文库构建进行了研究,如通过多重置换扩增(multiple displacement amplification, MDA)的方法将 DNA 量增加 9 个数量级,以达到建库的要求,但这样会显著改变微生物菌落结构,影响结果的可靠性^[12-13]。鉴于此,本研究针对低起始量 DNA 样本(最低至 0.1 ng),基于现有基因组测序、元基因组测序文库构建方法进行优化,并通过数据分析评估文库质量,建立一种高效稳定的文库构建方法,以期拓展微生物基因组测序、元基因组测序的应用领域提供参考。

1 材料与方法

1.1 仪器与试剂

Bioruptor UCD-30 非接触式超声波破碎仪(比利时 Diagenode 公司); Invitrogen™Qubit2.0 荧光计、Applied Biosystems PCR 热循环仪、2720Invitrogen™DynaMag™-96 侧面磁力架(美国 Thermo Fisher Scientific 公司); BioTek Epoch 全波长酶标仪(美国伯腾仪器有限公司); VORTEX-5 涡旋混合器(江苏海门市其林贝尔仪器制造有限公司)。

QIAGEN DNeasy® PowerSoil® Pro Kit (德国 QIAGEN 公司); 试剂盒 A (Rapid DNA Library Prep Kit for Illumina)、试剂盒 B (Hyper Prep Kit for Illumina)、Qubit assay kit (美国 Thermo Fisher Scientific 公司); 酵母提取物(英国 OXOID 公司); TPY 液体培养基(250 g, 青岛海博生物技术有限公司); 氯化钠(分析纯, 国药集团化学试剂有限公司)。

用于该实验的大肠杆菌由本实验室从粪便中分离培养获得; 沙门氏菌由本实验室从鸡肉中分离培养获得; 金黄色葡萄球菌株由外实验室赠予; 双歧杆菌购买于中国科学院菌种保藏中心, 菌株编号 CGMCC1.5082。

1.2 实验方法

1.2.1 基因组 DNA 提取与稀释

本研究选取平均鸟嘌呤和胞嘧啶(guanine and cytosine, GC)含量不同的革兰氏阴性菌(大肠杆菌、沙门氏菌)和革兰氏阳性菌(金黄色葡萄球菌和双歧杆菌)每种各 1 株。粪便样本取自 1 名志愿者。

使用 QIAGEN DNeasy® PowerSoil® Pro Kit, 按照说明书提取细菌和粪便基因组 DNA。Nanodrop 测定浓度, 根据光密度(optical density, OP) OD₂₆₀/OD₂₈₀ 值判断 DNA 的纯度。琼脂糖凝胶电泳用于鉴定基因组 DNA 的完整程度。Qubit 2.0 用于准确测定双链 DNA 的浓度。

根据 Qubit 测得的基因组 DNA 浓度, 对每个样本进行梯度稀释, 确保起始基因组 DNA 的量为 100.0、10.0、1.0、0.5 和 0.1 ng; 元基因组测序文库 1000、500、100、50、1 ng 的起始量梯度。每个梯度设置 3~4 个重复。

1.2.2 基因组文库构建

采用 2 种不同的试剂盒(A、B)对不同起始量的样本进行基因组测序文库制备, 其中所有基因组文库采用试剂盒 A 进行构建, 大肠杆菌和沙门氏菌的基因组文库同时还采用试剂盒 B 进行构建, 所有元基因组文库均用试剂盒 B 进行文库制备。此外, 根据 DNA 的起始量, 为达到成功制备测序文库的目的, 对不同 DNA 起始量的建库流程进行优化。

主要建库流程为超声破碎仪对基因组 DNA 进行片段化, 磁珠纯化 DNA 片段后进行末端修复和 3'加测序接头、片段选择以及扩增富集等步骤。根据基因组 DNA 起始量的不同, 选择不同的测序接头浓度及加入量(表 1), 在接头连接反应过程中对起始 DNA 量小于 1 ng 的样本, 增加 4 °C 孵育步骤, 时间不小于 1 h, 以增加接头的连接效率。文库制备过程中的片段选择可以提高文库质量和基因组拼接结果, 但是此过程中会损失 DNA, 为了保证能够获得可测序的基因组文库, 当起始 DNA 量小于 10 ng 时, 将不进行片段选择。最后用 PCR 法富集正确连接测序接头的 DNA 片段, PCR 的循环数随着起始 DNA 浓度的降低增加 2~5 个循环(表 1)。

表 1 DNA 起始量和 adapter 使用量
Table 1 DNA initiations and adapter usages

起始 DNA 量/ng	adapter 储备液浓度 ($\mu\text{mol/L}$)	稀释倍数	adapter 加入体积 / μL	是否片段选择	PCR 循环数
1000	15	1 \times	5	是	10
500	15	1 \times	5	是	10
100	15	1 \times	5	是	10
50	15	1 \times	5	否	10
10	15	5 \times	5	否	12
1	15	20 \times	5	否	15
0.5	15	50 \times	5	否	15
0.1	15	50 \times	3	否	15

1.2.3 文库质检与测序

对构建好的文库进行初步质检, 采用 Qubit 检测文库浓度, 琼脂糖凝胶电泳检测基因组测序文库片段大小。初步质检合格的文库, 根据预计的测序数据量成比例混合后, 使用荧光定量 PCR 和 LabChip Touch 进行进一步质检, 计算上机量。

使用 Illumina Novaseq 6000 进行双端 150 bp (PE150) 模式进行测序。基因组测序每个样本按照 1 G 的预计数据量进行测序; 元基因组测序每个样本按照 2 G 的数据量进行测序。

1.2.4 数据分析

(1) 数据质控

对测序得到的原始测序数据(raw data)进行质控, 获得有效数据(clean data)。数据质控包括: 去除接头污染的碱基数大于 5 bp 的序列, 去除序列中质量值 $Q \leq 19$ 的碱基占总碱基的 50% 以上的低质量序列, 以及去除含 N 比例大于 5% 的序列。

(2) 基因组组装

对于质控后得到的高质量基因组测序数据, 使用基于 SPAdes^[14]的 Unicycler 短序列组装软件对有效数据进行初步组装, 然后根据序列与组装得到的 contigs 的 paired-end 和 overlap 关系, 对组装结果进行优化, 再使用 GapCloser 1.12 软件对组装结果进行修补, 最后去除冗余序列得到最终的组装结果。

(3) 元基因组物种注释

对于质控后得到的高质量元基因组测序数据, 先使用 SOAP denovo^[15]进行元基因组组装, 组装主要分为 3 步: 单样本组装, 多样本组装, 结果合并和丰度过滤。然后使用 metaphlan2^[16]软件分析, 获得不同分类层级(门、纲、目、科、属、种)的微生物群落结构和丰度。

2 结果与分析

2.1 文库制备结果

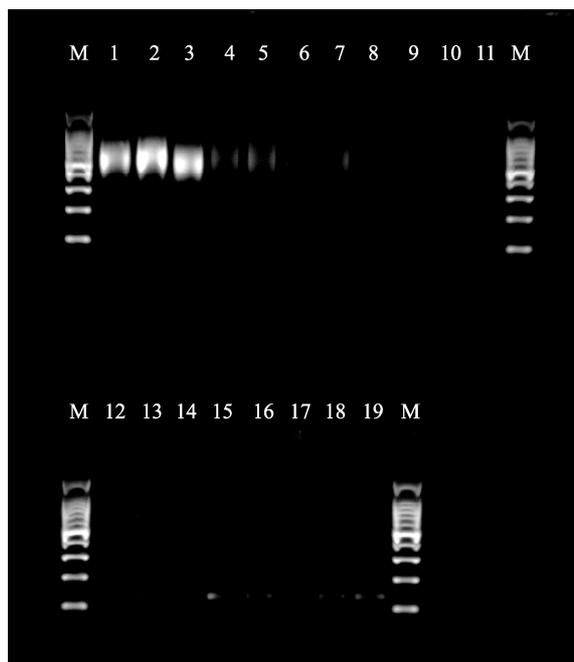
通过调整接头的浓度和 PCR 循环数, 起始 DNA 量从

0.1 ng 到 100 ng 的单菌基因组样本, 用试剂盒 A 均成功制备出测序文库。文库的质量浓度在 4.508~89.026 ng/ μL 之间。琼脂糖凝胶电泳结果显示(图 1), 经片段选择后的测序文库, 文库长度主要在 400~700 bp, 而没有经片段选择的文库, 文库长度成较大范围的弥散状。通过凝胶电泳结果还发现, 起始 DNA 的量从 1 ng 开始, 随着起始 DNA 量的减少, 接头和引物二聚体的比例逐步增多。虽然低起始量 DNA 所制备的文库浓度偏低, 但为了验证此种文库是否可以测序以及数据的质量和拼接结果, 将所有文库等量混在一起进行测序。

为了验证不同试剂盒的差异, 大肠杆菌和沙门氏菌各取 3 个不同的菌株, 用试剂盒 B 制备不同起始 DNA 量的测序文库。从图 1 可以看出, 没有经过片段选择的文库接头/引物二聚体的污染严重, 该污染会影响测序的数据质量, 浪费测序数据。因此, 试剂盒 B 文库制备不论为何种起始 DNA 量, 全部样本需进行片段选择, 该方法有效地去除了接头和引物二聚体的污染, 文库琼脂糖凝胶电泳质检结果与图 2 相似。由于片段选择的加入, 在文库制备过程中损失部分 DNA 片段, 导致起始 DNA 量在 0.5 ng 时有 1 个沙门氏菌样本建库失败。通过比较图 1 和图 2 发现, 虽然用同样的片段选择流程, 试剂盒 B 制备的基因组测序文库片段更集中, 且一致性更好, 由此得到的数据更利于后续的基因组拼接等生物信息分析。因此, 在起始 DNA 的量能够达到 1 ng 时, 建议用试剂盒 B 结合片段选择的方法制备基因组文库。

根据环境样本的特性, 某些样本无法获得足够量的原始样本(如拭子、齿垢样本等)以获得 1 μg 以上的基因组 DNA。而起始 DNA 量的多少不仅影响基因组文库能否制备成功, 还可能影响微生物群落结构的分析结果。以平均每个细菌的基因组为 5 Mb 计算, 1 ng 的元基因组 DNA 约含有 10 万个细菌的基因组, 而 1 μg 元基因组 DNA 则含有 1 亿个细菌。此外, 本研究所用样本为粪便样本, 人类肠道含有约 10^{12} 个细菌^[17-19], 起始 DNA 量过低会导致细菌种类过少, 故将元基因组的最低起始量提高到 1 ng。本研究通过分析一

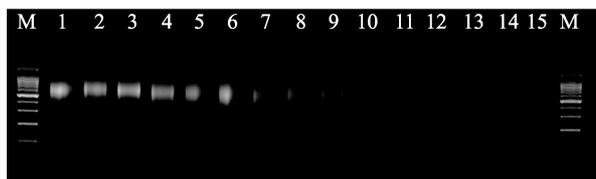
个环境样本不同起始量 DNA 制备的元基因组文库所产生的数据, 研究不同起始 DNA 量对元基因组微生物群落结构分析结果的影响。为了能够获得更多的有效数据, 元基因组文库均使用试剂盒 B 进行构建, 起始 DNA 的量从 1 μg 到 1 ng 都成功制备了基因组文库, 除 1 ng 起始量制备的测序文库质量浓度(2.176~9.05 $\text{ng}/\mu\text{L}$)较低外, 其他各起始量样本均获得质量浓度 20 $\text{ng}/\mu\text{L}$ 以上的基因组文库, 且片段大小集中、统一(图 2)。



注: M: 100 bp DNA marker, 条带从上往下依次为 1500、1000、900、800、700、600、500、400、300、200、100 bp; 1~3: 100 ng 大肠杆菌 DNA 起始量; 4~7: 10 ng 大肠杆菌 DNA 起始量; 8~11: 1 ng 大肠杆菌 DNA 起始量; 12~15: 0.5 ng 大肠杆菌 DNA 起始量; 16~19: 0.1 ng 大肠杆菌 DNA 起始量。

图 1 基因组测序不同起始量文库琼脂糖电泳

Fig.1 Agarose electrophoresis of genomic sequencing with different starting amounts libraries



注: M: 100 bp DNA marker, 条带从上往下依次为 1500、1000、900、800、700、600、500、400、300、200、100 bp; 1~3: 1000 ng 粪便 DNA 起始量; 4~6: 500 ng 粪便 DNA 起始量; 7~9: 100 ng 粪便 DNA 起始量; 10~12: 50 ng 粪便 DNA 起始量; 13~15: 1 ng 粪便 DNA 起始量。

图 2 元基因组测序不同起始量文库琼脂糖电泳

Fig.2 Agarose electrophoresis of metagenomic sequencing with different initial libraries

2.2 测序数据质量

所有测序文库在 Illumina Novaseq 6000 测序平台, 进行双端 150 bp (PE150)模式测序。对测序所得的原始测序数据(raw data)进行质控, 鉴定不同测序文库下机数据质量, 过滤后的高质量测序数据(clean data)进行后续的基因组拼接或微生物群落结构分析。每个细菌基因组测序获得 raw data 在 0.8~2.5 Gb 之间, 质控后获得 0.3~2.4 Gb 的 clean data。每个元基因组文库获得 1.0~3.0 Gb 的 raw data, 数据过滤后获得 0.9~2.9 Gb 的 clean data。

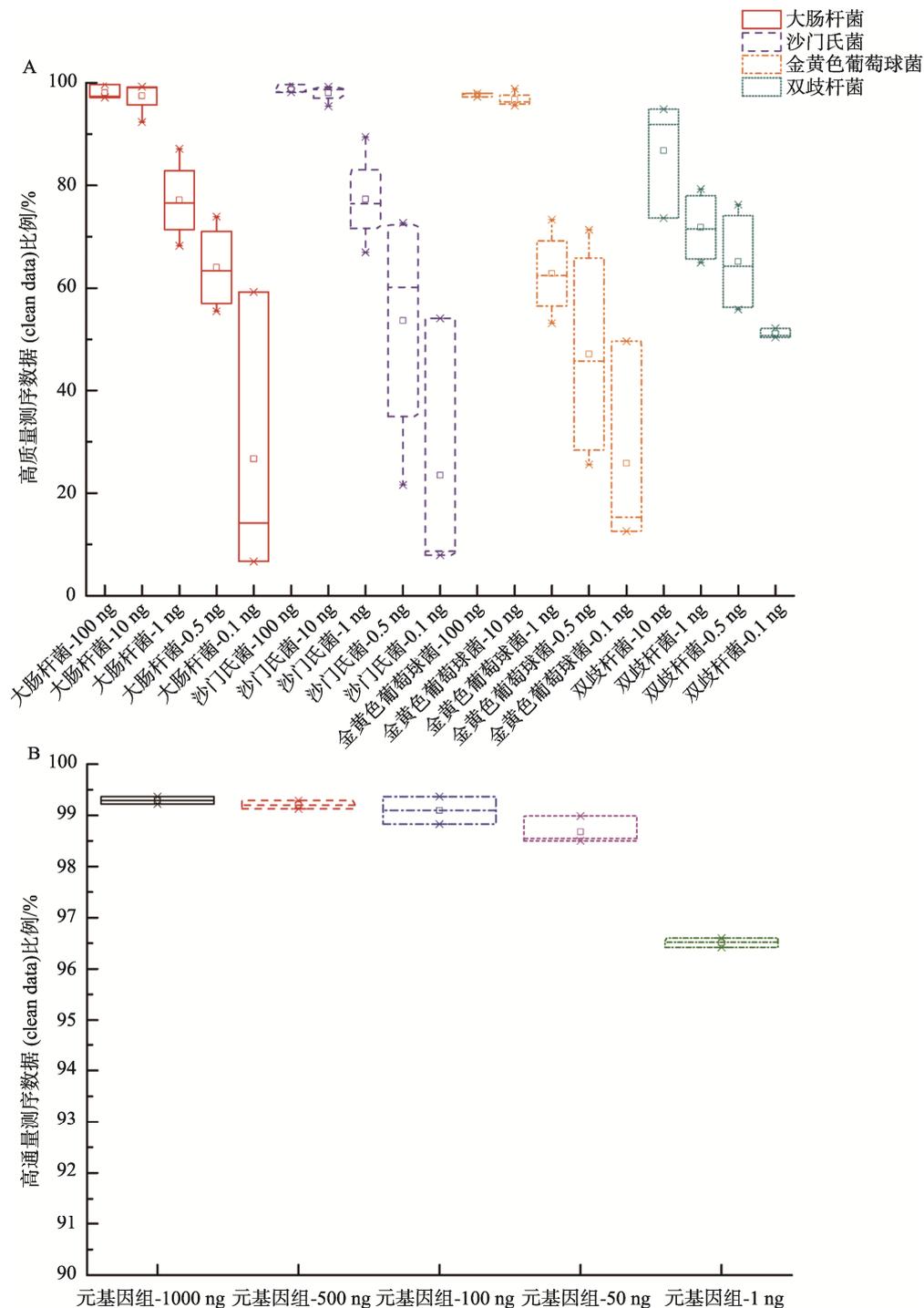
通过分析每个样本 clean data 的比例, 发现试剂盒 A 制备的基因组文库, 随着起始量的降低, 基因组测序文库的 clean data 比例较大程度下降。起始 DNA 量为 100 和 10 ng 时的 clean data 比例大于 90%, 起始 DNA 量为 1 ng 时的 clean data 平均比例降至 62%~77%, 至起始 DNA 量为 0.1 ng 时的 clean data 比例降低至 50%以下(图 3A)。而试剂盒 B 结合片段选择制备的基因组文库的数据质量优于试剂盒 A 制备的基因组文库, 尤其当起始 DNA 量为 1.0 和 0.5 ng 时, 文库测序所得数据的 clean data 的比例依然高于 90%(结果未展示)。使用相同方法制备的元基因组文库所得数据的 clean data 的比例同样很高, 1 ng 的起始 DNA 量制备的测序文库所得数据的 clean data 的比例均达到 96%, 数据质量整体良好(图 3B)。

2.3 不同起始量基因组测序文库对基因组组装的影响

选择了 4 种细菌进行不同起始 DNA 量基因组文库构建, 包括 2 种革兰氏阴性菌(大肠杆菌和沙门氏菌)和 2 种革兰氏阳性菌(金黄色葡萄球菌和双歧杆菌), 以 scaffold 数量和 N50 的长度标评估不同起始 DNA 量文库对基因组组装的影响。一般情况下, 测序文库均一性越好、测序数据质量越高, 相同物种的基因组组装后所得 scaffold 数量越少、N50 越长, 组装所得的基因组越完整。scaffold 数量结果显示(图 4), 所有样本的起始 DNA 量在 1 ng 以上时, 组装所得 scaffold 数量几乎无差异; 随着起始 DNA 量的降低, 不同细菌表现出的 scaffold 数量变化有所不同。大肠杆菌在起始 DNA 量降到 0.5 ng 时, scaffold 的数量较大幅度增高; 沙门氏菌的起始 DNA 量降低到 0.1 ng 时, scaffold 数才较大幅度增高; 而金黄色葡萄球菌和双歧杆菌在起始 DNA 量降到 0.1 ng 时, scaffold 数仅略有升高。除了测序文库质量和测序数据质量外, 基因组的大小和复杂程度也会影响 scaffold 数量, 基因组越大越复杂, scaffold 数量越多。本研究结果显示, 大肠杆菌的 scaffold 数量在 80~106 个, 沙门氏菌在 48~87 个, 金黄色葡萄球菌的在 26~52 个, 而双歧杆菌的 scaffold 数量仅为 11~19 个。大肠杆菌和沙门氏菌的基因组主要分布在 4~6 Mb 范围内, GC 含量在 50%左右; 金黄色葡萄球菌和双歧杆菌的基因组主要分布在 2~3 Mb 范围内, GC 含量分别在 32%和

62%左右。综上, 本研究发现基因组越大、结构越复杂, scaffold 数量对文库制备起始量越敏感; 而文库制备起始量不影响高 GC 含量细菌的基因组组装结果。图 5 为不同 DNA 起始量各细菌基因组拼接后 N50 的平均长度, 发现不同细菌随着起始 DNA 量的变化, N50 变化趋势同样呈现一定的

差异。大肠杆菌的 N50 随起始 DNA 量的降低仅略有降低; 双歧杆菌和沙门氏菌的 N50 则在 DNA 起始量在 0.1 ng 时较大程度降低; 金黄色葡萄球菌在起始 DNA 量小于 1 ng 时, N50 的值出现增长的情况, 但是查看每个样本的 N50 值时发现个体间差异很大。



注: A: 细菌基因组文库 clean data 比例; B: 元基因组文库 clean data 比例。

图 3 不同起始量文库测序数据质控分析

Fig.3 Data QC analysis of libraries constructed from different input DNA

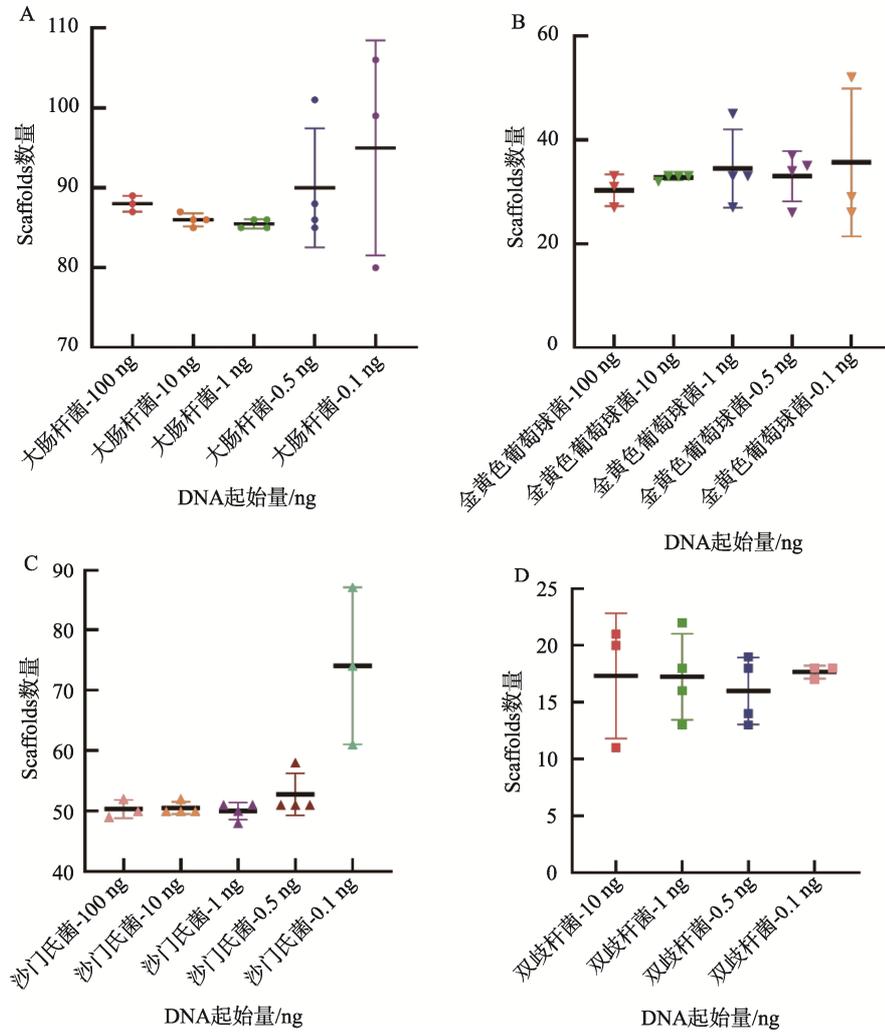


图 4 基因组测序不同起始量基因组组装 scaffold 数量
Fig.4 Scaffold counts of libraries constructed from different input DNA

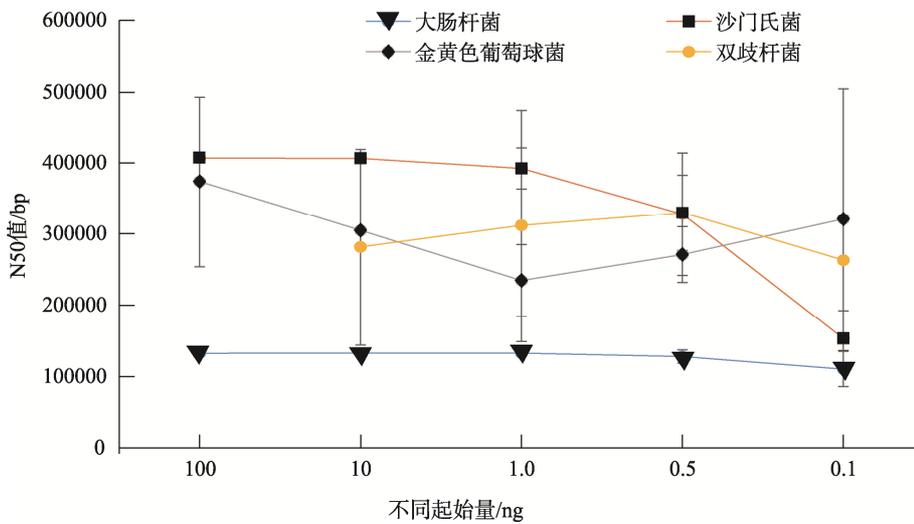


图 5 基因组测序不同起始量基因组组装平均 N50 值(n=3)
Fig.5 N50 lengths of libraries constructed from different input DNA (n=3)

2.4 元基因组测序不同起始量文库微生物组成

为了评估低起始量元基因组 DNA 文库对元基因组测序的影响, 对测序所得的不同 DNA 起始量的元基因组序列使用 metaphlan2 对细菌门、纲、目、科、属、种 6 个分类水平进行了物种注释。结果显示, 不同起始量文库的物种结构和丰度在各分类级别均无较大差异, 具有一致性(图 6)。本研究所用粪便样本在不同分类水平的细菌种类分别为门 4 种、纲 9 种、目 11 种、科 22 种、属 32 种、种 76 种, 且细菌种类相对比较集中, 如在门分类水平上, 样本中丰度最高的 2 个细菌类群占到 97% 以上, 分别为拟杆菌门(*Bacteroidetes*, 平均约占 45.4%)和厚壁菌门(*Firmicutes*, 平均约占 51.9%), 与文献报道的粪便样本主要细菌类群一致^[20-21]; 在属分类水平上, 拟杆菌属(*Bacteroides*, 平均约占 34.8%)、*Blautia* 属(平均约占 19.3%)和巨单胞菌属(*Megamonas*, 平均约占 18.0%) 3 个属的丰度占到了 70% 以上。结合数据质量及

物种注释结果, 用较低的 DAN 起始量制备基因组文库所得测序结果不会对元基因组测序的结果造成较大影响, 因此, 对于难获取或保存不当导致 DNA 量少不满足常规建库的样本, 仍可通过元基因组测序进行研究并得到相对可靠的结果。

3 结论与讨论

本研究通过优化文库构建流程, 对于起始量低至 0.1 ng 的样本, 可稳定完成基因组文库构建, 且对于基因组较大的细菌, 可通过增加测序数据量的方法提高组装效果, 达到和常规起始量建库一致的数据分析结果。不同的文库制备试剂盒以及是否加入片段选择步骤对数据质量有一定的影响。对于元基因组测序, 起始量为 1 ng DNA 制备的测序文库与常规起始量制备的基因组文库, 微生物群落结构相近。研究结果可为拓展微生物基因组测序、元基因组测序的应用领域提供参考。

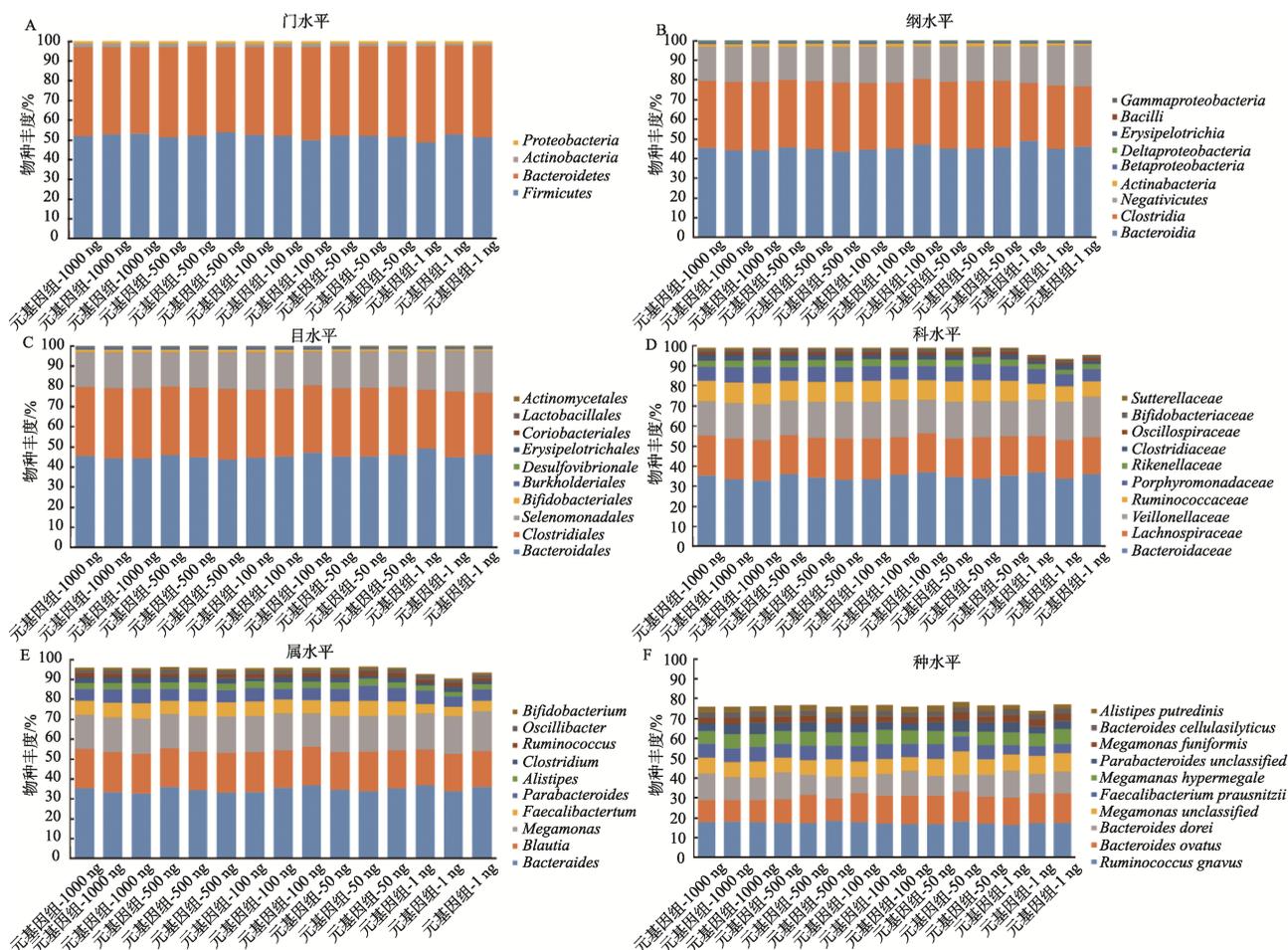


图 6 元基因组测序不同起始量文库细菌结构和丰度(前 10)

Fig.6 Bacterial composition and abundance of libraries constructed from different input DNA (top10)

参考文献

- [1] LIU L, LI Y, LI S, *et al.* Comparison of next-generation sequencing systems [J]. *J Biomed Biotechnol*, 2012, 2012: 251364.
- [2] SHENDURE J, BALASUBRAMANIAN S, CHURCH GM, *et al.* DNA sequencing at 40: Past, present and future [J]. *Nature*, 2017, 550(7676): 345–353.
- [3] HEATHER JM, CHAIN B. The sequence of sequencers: The history of sequencing DNA [J]. *Genomics*, 2016, 107(1): 1–8.
- [4] GARRIDO-CARDENAS JA, GARCIA-MAROTO F, ALVAREZ-BERMEJO JA, *et al.* DNA Sequencing sensors: An overview [J]. *Sensors*, 2017, 17(3): 588.
- [5] 夏梦雷, 杨帆, 陆错, 等. 高通量测序技术在传统发酵食品微生物群落中的应用研究[J]. *中国酿造*, 2021, 40(5): 7.
XIA ML, YANG F, LU K, *et al.* Application of high throughput sequencing technology in microbial community of traditional fermented food [J]. *China Brew*, 2021, 40(5): 7.
- [6] METZKER ML. Sequencing technologies-The next generation [J]. *Nat Rev Genet*, 2010, 11(1): 31–46.
- [7] HEAD SR, KOMORI HK, LAMERE SA, *et al.* Library construction for next-generation sequencing: Overviews and challenges [J]. *Biotechniques*, 2014, 56(2): 61.
- [8] HESS JF, KOHL TA, KOTROVÁ M, *et al.* Library preparation for next generation sequencing: A review of automation strategies [J]. *Biotechnol Adv*, 2020, 41: 107537.
- [9] SIDSTEDT M, RÅDSTRÖM P, HEDMAN J. PCR inhibition in qPCR, dPCR and MPS-Mechanisms and solutions [J]. *Anal Bioanal Chem*, 2020, 412(3): 2009–2023.
- [10] 王霞, 欧阳艳艳, 王晶, 等. 高通量测序 DNA 文库定量质控技术研究 [J]. *计量学报*, 2020, 41(10): 1308–1312.
WANG X, OUYANG YY, WANG J, *et al.* Research on characterization method for library DNA reference material for next generation sequencing [J]. *Acta Metrol Sin*, 2020, 41(10): 1308–1312.
- [11] 杨祖玉, 王开乐, 吴大飞, 等. 以痕量 DNA 为基础的二代测序文库构建方法: 中国, CN201410158326.7[P]. 2014-07-23.
YANG ZY, WANG KL, WU DF, *et al.* Construction of second generation sequencing library based on trace DNA: China, CN201410158326.7 [P]. 2014-07-23.
- [12] YILMAZ S, ALLGAIER M, HUGENHOLTZ P. Multiple displacement amplification compromises quantitative analysis of metagenomes [J]. *Nat Method*, 2010, 7(12): 943–944.
- [13] PROBST AJ, WEINMAIER T, DESANTIS TZ, *et al.* New perspectives on microbial community distortion after whole-genome amplification [J]. *PLoS One*, 2015, 10(5): e0124158.
- [14] BANKEVICH A, NURK S, ANTIPOV D, *et al.* SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing [J]. *J Comput Biol*, 2012, 19(5): 455–477.
- [15] LUO R, LIU B, XIE Y, *et al.* SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler [J]. *Gigascience*, 2012, 1(1): 18.
- [16] TRUONG DT, FRANZOSA EA, TICKLE TL, *et al.* MetaPhlan2 for enhanced metagenomic taxonomic profiling [J]. *Nat Methods*, 2015, 12(10): 902–903.
- [17] LI D, FENG Y, TIAN M, *et al.* Gut microbiota-derived inosine from dietary barley leaf supplementation attenuates colitis through PPAR γ signaling activation [J]. *Microbiome*, 2021, 9(1): 83.
- [18] DAVID LA, MAURICE CF, CARMODY RN, *et al.* Diet rapidly and reproducibly alters the human gut microbiome [J]. *Nature*, 2014, 505(7484): 559–563.
- [19] SENDER R, FUCHS S, MILO R. Are we really vastly outnumbered revisiting the ratio of bacterial to host cells in humans [J]. *Cell*, 2016, 164: 337–340.
- [20] HUTTENHOWER C, GEVERS D, KNIGHT R, *et al.* Structure, function and diversity of the healthy human microbiome [J]. *Nature*, 2012, 486(7402): 207–214.
- [21] WEXLER AG, GOODMAN AL. An insider's perspective: Bacteroides as a window into the microbiome [J]. *Nat Microbiol*, 2017, 2: 17026.

(责任编辑: 于梦娇 张晓寒)

作者简介



李博星, 硕士, 主要研究方向为高通量测序方向的研究。
E-mail: 18700943237@163.com

律娜, 助理研究员, 主要研究方向为分子生物技术应用。
E-mail: lvna@im.ac.cn