

# Illumina MiSeq 高通量测序分析花生仁内生细菌多样性

陈泽斌<sup>1,2</sup>, 高 熹<sup>3\*</sup>, 胡 映<sup>1</sup>, 赵晓红<sup>1</sup>, 郭丽红<sup>2</sup>, 王定康<sup>1</sup>

(1. 昆明学院 农学院, 云南 昆明 650214; 2. 云南省高校特色生物资源开发与利用重点实验室, 云南 昆明 650214;  
3. 云南农业大学 植物保护学院, 云南 昆明 650201)

**摘要:** 为了调查花生仁内生细菌种类多样性, 应用 Illumina MiSeq 高通量测序技术测定花生仁内生细菌的 16S rDNA-V4 变异区序列, 并采用 Mothur 软件整理及统计样品序列数目和操作分类单元 (OTUs) 数量, 分析物种的丰度、分布和  $\alpha$  多样性。研究获得用于分析的有效序列和 OTU 数分别为 69 284 和 25; 稀疏曲线表明测序深度充分, OTU 的数量接近于饱和。花生仁内生细菌主要分布于食酸菌属 (*Acidovorax*) 和希瓦氏菌属 (*Shewanella*), 丰度分别为 83.33% 和 16.67%。且内生细菌的 Chao1 指数为 25.00, Shannon 指数为 1.29。

**关键词:** 高通量测序; 花生仁; 内生细菌; 多样性

**中图分类号:** S565.2 **文献标识码:** A **文章编号:** 1674 - 5639 (2018) 03 - 0086 - 04

**DOI:** 10.14091/j.cnki.kmxyxb.2018.03.017

## Analysis on Diversity of Endophytic Bacteria in Peanut Kernel with Illumina MiSeq High-throughput Sequencing Technology

CHEN Zebin<sup>1,2</sup>, GAO Xi<sup>3\*</sup>, HU Yi<sup>1</sup>, ZHAO Xiaohong<sup>1</sup>, GUO Lihong<sup>2</sup>, WANG Dingkang<sup>1</sup>

(1. College of Agriculture, Kunming University, Kunming, Yunnan, China 650214; 2. Key Laboratory of Special Biological Resource Development and Utilization of Universities in Yunnan Province, Kunming, Yunnan, China 650214;  
3. Plant Protection College, Yunnan Agricultural University, Kunming, Yunnan, China 650201)

**Abstract:** In order to investigate the diversity of endophytic bacterial in peanut kernel, the 16S rDNA-V4 region of peanut kernel was sequenced by Illumina MiSeq high-throughput sequencing technology. Using Mothur softwares, the number of sequences and operational taxonomic units (OTUs) for each sample was sorted and calculated the species abundance and distribution. The research results showed that the number of effective sequences and OTUs for each sample were 69 284/25; the rarefaction curves showed that the sequencing depth was abundant; the number of OTUs was close to saturation. The bacteria in peanut kernel existing in the dominant species are *Acidovorax* and *Shewanella* with a percentage of 83.33% and 16.67%, respectively. The Chao1 index of endophytic bacteria was 25.00, and the Shannon index was 1.29.

**Key words:** high-throughput sequencing; peanut kernel; endophytic bacteria; diversity

目前已从各种健康植物的根、茎、叶、果实和种子中分离出 129 种 (54 个属) 的植物内生细菌, 包括革兰氏阳性和革兰氏阴性细菌<sup>[1]</sup>。与此相关的植物主要有小麦、棉花等<sup>[2]</sup>。优势种主要为芽孢杆菌 (*Bacillus*)、假单胞菌 (*Pseudomonas*)、土壤杆

菌 (*Agrobacterium*)、肠杆菌 (*Enterobacter*) 4 个属<sup>[3]</sup>, 同时还发现所有研究植物中都含有内生细菌。

花生 (*Arachis hypogaea* L.) 含有丰富的蛋白质, 其质量分数为 25% ~ 30%。花生蛋白质中含有 8 种人体必需的氨基酸, 精氨酸质量分数高于其他

收稿日期: 2017 - 08 - 01

**基金项目:** 国家自然科学基金资助项目 (31460491); 昆明学院人才引进项目 (YJL14005); 云南省高校优势特色重点学科 (生态学) 建设资助项目 (05000511311); 云南省特色生物资源开发与利用重点实验室开放基金资助项目 (GXKJ201621)。

**作者简介:** 陈泽斌 (1985—), 男, 云南昆明人, 副教授, 博士, 主要从事植物病理学研究。

\* **通讯作者:** 高熹 (1980—), 女, 福建福州人, 副教授, 博士, 主要从事病虫害综合治理研究, E-mail: chonchon@163.com.

坚果, 生物效能高于大豆<sup>[4]</sup>。花生仁或其他花生制品, 可以为人体提供大量的蛋白质、脂肪和能量, 且可以减少饮食中的饱和脂肪, 促进人体对维生素 E、植物蛋白、纤维素、钙、钾、镁、锌等有益于健康的营养素吸收<sup>[5-6]</sup>。

研究<sup>[7]</sup>发现, 品种、生长期和环境条件是影响植物内生细菌组成的主要影响因素。以往的研究<sup>[8-9]</sup>中, 花生内生细菌主要是通过传统方法培养, 且研究点多集中在花生植株的根、茎和叶。而目前尚未见到关于花生的种子—花生仁内生细菌的报道。因此, 有必要研究花生仁内生细菌的物种组成。本研究采用 Illumina MiSeq 第 2 代测序技术, 对花生仁内生细菌的组成进行全面分析, 与传统培养方法和基于 16S rDNA 的非培养方法相比较, 本文采用的方法可以生成具有更大覆盖率的数据量, 且可以检测到一些含量较低的内生细菌, 而通过传统方法培养的都无法检测出, 因而本文所采用的方法在植物微生物学和遗传领域具有重要的应用价值。

1 材料与方法

1.1 样品来源

“滇引花生 3 号”新鲜花生仁由云南省农科院油料所提供。

1.2 表面消毒

准备一定量的新鲜花生仁, 清水冲洗 1 h 后, 用体积分数为 75% 的乙醇振荡浸泡 1 min, 再用无菌水冲洗 3 次, 然后将其表面水分用无菌滤纸吸干, 表面消毒检验参考文献<sup>[10]</sup>的方法进行。

1.3 总 DNA 提取

下面提取处理后的花生仁的总 DNA, 提取方法主要依据文献<sup>[11]</sup>的相关方法进行。采用 0.8% 琼脂糖凝胶电泳检测处理后的花生仁 DNA 的纯度和质量浓度。将适量的样品置于离心管中离心处理之后, 加入一定的无菌水进行稀释, 一直到其质量浓度为 1 ng/μL。

1.4 16S rDNA-V4 区的 PCR 扩增

以提取的总 DNA 为模板, 引物为 16S rDNA-V4 区特异性引物 (5'-GTTTCGGTGCCAGCGCCGCGG-TAA-3') 和 806R (5'-GTTAGCGGGGACHVGGGT-WTAAT-3'), 接着通过 Fermentas 高保真聚合酶 (EP0502) 进行扩增, 以此来提高检测效率。

1.5 PCR 产物的混样和纯化

对所得产物电泳处理, 选择的电泳液为 2% 的

琼脂糖凝胶, 对所得产物利用凝胶回收试剂盒回收处理, 且保证其纯度满足要求。随后委托北京诺禾致源生物信息科技有限公司进行 MiSeq 测序<sup>[12-15]</sup>。

1.6 生物信息学分析

为了使信息分析结果的准确性满足要求, 对所得原始数据做了拼接和过滤处理, 将其中的嵌合体序列去掉<sup>[16]</sup>, 同时排除宿主叶绿体序列数据, 对其余的有效数据进行 OTU 聚类处理, 选择的相似性水平为 97%, 然后利用相应的 Greengene 数据库注释物种<sup>[17]</sup>。对所得结果进行丰度、α 多样性等方面的分析处理, 以此来确定出微生物群落结构特征信息<sup>[18]</sup>。

1.7 数据分析

在聚类分析时, 选择了 Uparse (版本 7.1) 软件, 同时利用了 RDP classifier 贝叶斯算法对所得 OTU 代表序列分类处理, 并对样本中各分类水平上的菌群进行统计。对菌落进行 α 多样性分析时, 主要依据 Chao1 指数和 Shannon 指数。分析稀释曲线时选择 Mothur 软件, 分别利用 Fasttree 和 SVG 软件绘制热图和对应的分类树图<sup>[19]</sup>。

2 结果与分析

2.1 原始序列的长度和分布

所得结果表明, 花生仁样品 (HS) 的 69 287 条原始序列分布在 198 ~ 348 bp 范围内, 长度为 248 bp 的序列最多, 达 68 885 条 (表 1)。测得的序列长度与 16S rDNA-V4 区的序列长度一致。

表 1 原始序列的长度及分布

| 长度/bp | 数量/条   | 长度/bp | 数量/条 |
|-------|--------|-------|------|
| 198   | 1      | 278   | 1    |
| 208   | 2      | 288   | 32   |
| 218   | 6      | 298   | 28   |
| 228   | 91     | 308   | 11   |
| 238   | 200    | 318   | 3    |
| 248   | 68 885 | 328   | 0    |
| 258   | 26     | 338   | 1    |
| 268   | 0      | 348   | 0    |

2.2 有效序列数及 OUT 丰度聚类

原始序列经过拼接、过滤及去除嵌合体序列, 共得到了 69 284 条有效序列, 可分为 25 个 OTU, 其中 67 984 条为获得分类信息的序列, 1 300 条为低频序列 (图 1)。稀释曲线表明 (图 2), 测序深度充分, OTU 的数量接近于饱和。OTUs 丰度聚类

图表明（图 3），花生内生细菌主要分布在以下 2 个属：食酸菌属（*Acidovorax*）和希瓦氏菌属（*Shewanella*），且食酸菌属丰度（*Acidovorax*）> 希瓦氏菌属丰度（*Shewanella*）。α 多样性分析表明，花生内生细菌 Chao1 指数为 25.00，Shannon 指数为 1.29。

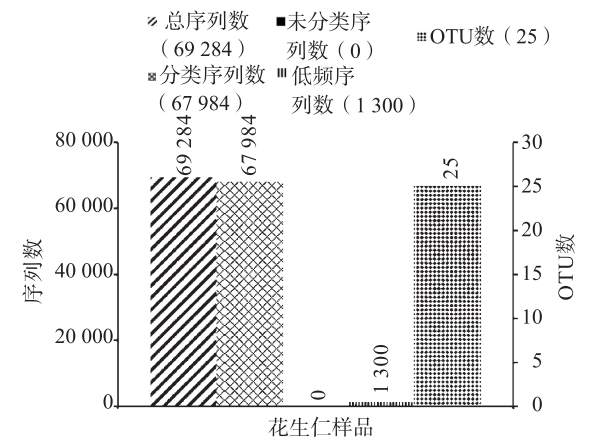


图1 有效序列数及OUT数

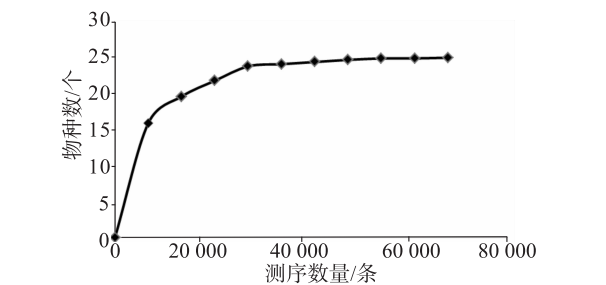


图2 序列的稀释曲线

2.3 菌群在不同分类水平上的丰度

分析所得结果可知（表 2），花生仁内生细菌全部分布在变形菌门，在 γ-变形菌纲中分布的比例为 16.67%，在 β-变形菌纲中分布的比例为 83.33%；而依据目的分类水平结果可知，此种细菌在伯克氏菌目（Burkholderiales, 83.33%）、交替单胞菌目（Alteromonadales, 16.67%）中均有分布；从科的水平来看，花生内生细菌在丛毛单胞菌科（Comamonadaceae, 83.33%）、希瓦氏菌科（Shewanellaceae, 16.67%）中均有分布；而依据属分类结果可知，其在食酸菌属（*Acidovorax*, 83.33%）、希瓦氏菌属（*Shewanella*, 16.67%）中均有分布，由此可见，这 2 个属为花生内生细菌的优势属。

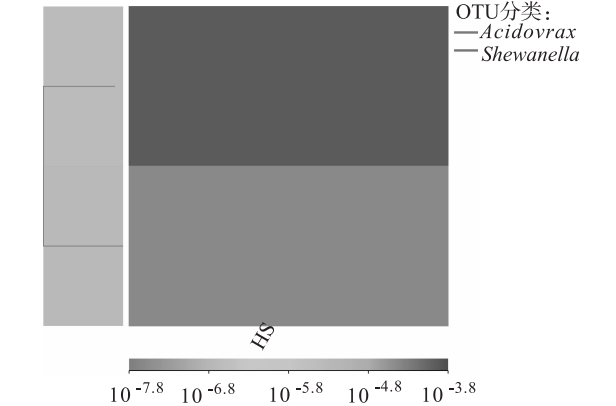


图3 OTUs丰度聚类图

表 2 菌群在不同分类水平上的丰度

| 界                  | 门                        | 纲                               | 目                           | 科                          | 属                             |
|--------------------|--------------------------|---------------------------------|-----------------------------|----------------------------|-------------------------------|
| Bacteria<br>(100%) | Proteobacteria<br>(100%) | Betaprobacteria<br>(83.33%)     | Burkholderiales<br>(83.33%) | Comamonadaceae<br>(83.33%) | <i>Acidovorax</i><br>(83.33%) |
|                    |                          | Gammaproteobacteria<br>(16.67%) | Alteromonadales<br>(16.67%) | Shewanellaceae<br>(16.67%) | <i>Shewanella</i><br>(16.67%) |

注：括号内数字为菌群在该分类水平上的丰度，即所占比例。

3 结论与讨论

16S rRNA 基因长期以来被用于鉴定微生物菌株，直到 2006 年，16S rRNA 基因才首次被高通量测序，此后，Sogin 等<sup>[20]</sup>研究者将高通量测序技术应用于深海微生物群落分析，随之该技术则被广泛应用于肠道微生物研究<sup>[21]</sup>。本文主要是利用此种方法进行了植物内生细菌研究，对比发现其可以有效地避免传统分子生物学方法的缺陷，可以对细菌进行基因组水平的研究，且对厌氧内生微生物也有较高的适用性<sup>[22-23]</sup>，与传统纯培养和 16S rDNA 非

培养方法相比较，第 2 代测序技术可以检测到过去没有发现的，同样发挥重要作用的低丰度内生菌群，该技术可以对全部内生微生物群落进行高效检测，在反映植物内生微生物群落中的细菌比例方面有着较高的应用价值。

大量实验数据说明，叶绿体与细菌的同源性很高<sup>[24]</sup>。在研究中发现，部分序列属于宿主叶绿体和线粒体，为更精确地进行细菌多样性的分析，本研究将其中的叶绿体序列数据排除。为了改进高通量测序技术在植物内生微生物中的应用，下一步应该在不同的测序平台上对 16S rDNA 进行各

种可变区选择处理, 以此来确定出样品的复杂性, 避免叶绿体干扰, 获得更准确的测序结果。

本文所得结果只对属水平上的内生细菌进行了分类, 这是因为 MiSeq 高通量测序技术一般每个样品测定 4~6 万条序列, 由于测序数量大, 充分展示了物种的多样性, 但是由于高通量测序读取长度短, 不能对 16S rRNA 的全部 9 个可变区进行测序, 因此, 通常选择 1~3 个长度为 250~500 bp 的可变区进行测序, 测序片段越短, 越难鉴定到属以下的分类水平。

本研究认为, 花生内生细菌主要分布于以下 2 个属: 食酸菌属 (*Acidovorax*, 83.33%) 和希瓦氏菌属 (*Shewanella*, 16.67%), 目前有关其功能的研究尚未见到相关报道, 仍需要借助传统分离培养方法, 获取菌株实体, 通过试验进一步验证。

#### [参考文献]

- [1] 赵旭, 常思静, 景春娥, 等. 我国植物内生菌研究进展 [J]. 中国沙漠, 2010, 30 (1): 87-91.
- [2] 胡桂萍, 郑雪芳, 尤民生, 等. 植物内生菌的研究进展 [J]. 福建农业学报, 2010, 25 (2): 226-234.
- [3] 徐亚军. 植物内生菌资源多样性研究进展 [J]. 广东农业科学, 2011, 38 (24): 149-152.
- [4] 马勇, 边仁伟. 辽西花生营养成分特性分析 [J]. 中国粮油学报, 2010, 25 (8): 35-37.
- [5] 胡晓辉, 胡东青, 苗华荣, 等. 花生营养品质遗传改良研究进展 [J]. 花生学报, 2012, 41 (1): 36-43, 47.
- [6] 李雄辉, 徐刚. 花生营养奶生产工艺及稳定性的研究 [J]. 食品科学, 1994 (10): 42-45.
- [7] 宋子红, 丁立孝, 马伯军, 等. 花生内生菌的种群及动态分析 [J]. 植物保护学报, 1999, 26 (4): 309-314.
- [8] 王小兵, 骆永明, 刘五星, 等. 花生青枯病内生拮抗细菌的鉴定、抗菌活性及其田间防效 [J]. 中国生物防治学报, 2011, 27 (1): 88-92.
- [9] 方扬, 张小平, 陈露遥, 等. 天府花生内生细菌种群多样性研究 [J]. 西南农业学报, 2008, 21 (2): 353-358.
- [10] 陈泽斌, 代方平, 寸林江, 等. 烟草内生细菌分离方法的优化研究 [J]. 中国烟草学报, 2014, 20 (1): 90-95, 102.
- [11] 陈泽斌, 夏振远, 雷丽萍, 等. 非培养方法解析烟草根部内生细菌的群落结构 [J]. 华北农学报, 2012, 27 (1): 201-209.
- [12] CAPORASO J G, LAUBER C L, WALTERS W A, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample [J]. Proceedings of the National Academy of Sciences, 2011, 108 (1): 4516-4522.
- [13] YOUSSEF N, SHEIK C S, KRUMHOLZ L R, et al. Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys [J]. Applied and Environmental Microbiology, 2009, 75 (16): 5227-5236.
- [14] HESS M, SCZYRBA A, EGAN R, et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen [J]. Science, 2011, 331 (6016): 463-467.
- [15] LUO C, TSEMENTZI D, KYRPIDES N, et al. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample [J]. PloS One, 2012, 7 (2): 1-12.
- [16] EDGAR R C, HAAS B J, CLEMENTE J C, et al. UCHIME improves sensitivity and speed of chimera detection [J]. Bioinformatics, 2011, 27 (16): 2194-2200.
- [17] EDGAR R C. UPARSE: highly accurate OTU sequences from microbial amplicon reads [J]. Nature Methods, 2013, 10 (10): 996-998.
- [18] WANG Q I, GARRITY G M, TIEDJE J M, et al. Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy [J]. Applied and Environmental Microbiology, 2007, 73 (16): 5261-5267.
- [19] DESANTIS T Z, HUGENHOLTZ P, LARSEN N, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB [J]. Applied and Environmental Microbiology, 2006, 72 (7): 5069-5072.
- [20] SOGIN M L, MORRISON H G, HUBER J A, et al. Microbial diversity in the deep sea and the underexplored "rare biosphere" [J]. Proceedings of the National Academy of Sciences, 2006, 103 (32): 12115-12120.
- [21] 南春燕, 马雅军, 徐建农, 等. 中华按蚊幼虫肠道细菌宏基因组的组成研究 [J]. 中国寄生虫学与寄生虫病杂志, 2013, 31 (2): 114-119.
- [22] THOLOZAN J L, CAPPELIER J M, TISSIER J P, et al. Physiological characterization of viable-but-nonculturable campylobacter jejuni cells [J]. Appl Environ Microbiol, 1999, 65 (3): 1110-1116.
- [23] AMMANN R R, LUDWIG W, SCHLEIFER K H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation [J]. Microbiol Rev, 1995, 59 (1): 143-169.
- [24] 胡楷, 吴庆书. 单细胞生物进化研究的进步 [J]. 遗传, 2002, 24 (1): 104-110.