

微生物学报 *Acta Microbiologica Sinica*
54 (12) :1489 – 1499; 4 December 2014
ISSN 0001 – 6209; CN 11 – 1995/Q
http://journals.im.ac.cn/actamicrocn
doi: 10.13343/j.cnki.wsxb.2014.12.012

高通量测序和 DGGE 分析土壤微生物群落的技术评价

夏围围^{1,2}, 贾仲君^{1*}

¹中国科学院南京土壤研究所, 土壤与农业可持续发展国家重点实验室, 江苏 南京 210008

²中国科学院大学, 北京 100049

摘要: 【目的】比较新一代高通量测序与传统的变性梯度凝胶电泳 (Denaturing Gradient Gel Electrophoresis, DGGE) 指纹图谱技术, 评价两种技术研究土壤微生物群落结构的优缺点。【方法】针对新西兰典型草地和森林土壤, 以 16S rRNA 基因为靶, 通过高通量测序和 DGGE 技术分析土壤微生物群落的组成、丰度和多样性, 比较两种方法在土壤微生物研究中的适用性。【结果】在不同的微生物分类水平, 高通量测序草地土壤检测到 22 门, 54 纲, 60 目, 131 科, 350 属; 而 DGGE 仅检测到 6 门, 9 纲, 8 目, 10 科, 10 属, 表明 DGGE 显著低估了土壤微生物的群落组成。森林土壤也得到了类似规律, 高通量测序的检测灵敏度是 DGGE 的 3.8、6.7、6.4、19.2 及 39.4 倍。进一步分析土壤中主要微生物类群的相对丰度, 发现分类水平越低, 高通量测序与 DGGE 的结果差异越大, 尤其在科和属的水平上差异最大。以高通量测序结果为标准, DGGE 明显高估了土壤中大多数微生物类群的相对丰度, 最高可达 2000 倍。两种方法都表明草地土壤的多样性指数高于森林土壤, 但 DGGE 多样性指数的绝对值远低于高通量测序结果。【结论】高通量测序能够较为全面和准确的反映土壤微生物群落结构, 而 DGGE 仅能够反映有限的优势微生物类群, 在很大程度上极可能低估土壤微生物的物种组成并高估其丰度。

关键词: 高通量测序, 变性梯度凝胶电泳 (Denaturing Gradient Gel Electrophoresis, DGGE), 土壤微生物群落
中图分类号: Q938 **文章编号:** 0001-6209 (2014) 12-1489-11

微生物是土壤最活跃的组成。从定植于土壤母质的蓝绿藻开始, 到土壤肥力的形成, 土壤微生物参与了土壤发生、发展、发育的全过程, 推动了地球表层系统关键生态过程, 维系了陆地生态系统物质与能量的良性循环, 被称为地球元素生物地球化学循环的引擎^[1]。然而, 由于微生物肉眼不可见, 从分类学的角度准确刻画和定义土壤中数量巨大的微生物类群, 一直是学术界的重大理论和技术问题^[2]。例如, 19 世纪末建立的平板分离培养技术被认为是

土壤微生物研究的重要里程碑, 但由于微生物体积小、形态简单, 基于纯培养技术的表型特征分类具有较大的局限性。因此, 长期以来由于理论和技术的限制, 从土壤学的角度很难定量研究微生物区系在复杂土壤环境中动态变化规律, 从而原位表征微生物区系的演变特征及其在农业生产实践中的重要作用^[3]。20 世纪 90 年代, 以 rRNA 为基础的生物三域分类理论逐渐得到学术界认可^[4], 通过对环境样品中大量 rRNA 核苷酸序列比对分析, 发现土壤中最

基金项目: 中国科学院战略性先导科技专项 (B 类) (XDB15040000)

* 通信作者。Tel: +86-25-86881311; Fax: +86-25-86881000; E-mail: jia@issas.ac.cn

作者简介: 夏围围 (1986 -), 女, 安徽淮南人, 博士研究生, 主要从事土壤微生物研究。E-mail: wwxia@issas.ac.cn

收稿日期: 2014-02-17; 修回日期: 2014-07-03

多可达 99% 的微生物尚未被培养,其功能尚未可知^[5]。同时,以 rRNA 序列比对为基础的分子指纹图谱技术极大地改变了传统的研究理念和方法,将土壤微生物学的研究对象从单个菌种资源发展到整体的微生物群落演替及功能意义。

20 世纪 90 年代,DNA 测序技术快速发展并在分子生态学中得到了广泛应用。克隆文库和聚丙烯酰胺凝胶电泳 (Denaturing Gradient Gel Electrophoresis, DGGE) 等为代表分子指纹图谱技术极大地推动了微生物生态学的研究^[6]。这些技术以复杂环境样品如土壤为研究对象,直接提取其中微生物的基因组 DNA,进一步采用 16S rRNA 基因的通用引物获得所有微生物的 PCR 扩增产物,通过构建 16S rRNA 基因的克隆文库^[7],或者开展 DGGE 凝胶电泳分析^[8],比较克隆子或者 DGGE 条带的序列,推测复杂环境微生物群落的分类学地位及其对环境条件的响应与适应机理。与克隆文库技术相比,DGGE 更加简单高效,能够同时最多分析 32 个样品,其条带的数量和亮度可以较好地指征土壤主要微生物类群的变化。然而,每克土壤中微生物据估算高达 1 - 10 亿,传统的分子指纹图谱如 DGGE 和克隆文库获得的微生物 DNA 序列通常低于 100 条,检测限低、工作量大,特别对于土壤中数量少,但具有重要功能的土壤微生物类群区系,传统分子指纹图谱技术具有较大的局限性^[6]。

21 世纪以来,新一代高通量测序技术的发展日新月异,可直接测序 16S rRNA 基因的 PCR 产物,每次分析获得的基因序列数以百万甚至亿万计,不仅通量高,而且能够同时分析上百个不同的样品,是解析复杂环境中微生物群落物种组成和相对丰度的重要工具^[9]。2007 年焦磷酸高通量测序技术首次应用于土壤微生物 16S rRNA 基因遗传多样性研究^[10],迄今在 ISI 科学引文数据库已有 500 余篇土壤微生物相关论文报道。然而,相比于传统分子指纹图谱技术,新一代高通量测序的技术优势仍处于一种定性的描述,尚未见系统的技术评价报道。据此,本研究针对新西兰两种典型草地土壤和森林土壤,以 16S rRNA 基因为靶,通过 454 高通量测序和 DGGE 技术研究土壤微生物群落,分析比较土壤微生物的组成、丰度和多样性,明确两种方法在评价土壤微生物群落结构方面的优缺点,为微生物分子生态学研究提供一定的参考。

1 材料和方法

1.1 土壤样品概况

供试土壤采自新西兰两种典型生态系统,包括 Palmerston North 奶牛牧场(40°38'S,175°61'E)的草地土壤(GS, Grassland Soil)、新西兰北岛中部的 Puruki 森林野外试验站(38°36'S,176°13'E)的森林土壤(FS, Forest Soil)。两种土壤样品采集时间分别为 2010 年 2 月和 2008 年 6 月。每块试验地随机选取 6 个取样点,采集 0 cm - 10 cm 的表层新鲜土壤,除去杂物、细根,研磨并过 2 mm 筛,然后将 6 个样点的土壤混合均匀,保存于 4℃ 待用。草地和森林土壤的基本理化性质分别如下:pH, 5.41 和 5.18;含水量,31.2% 和 54.3%;硝态氮,143.8 和 275.6 g/kg;铵态氮,16.7 和 19.7 g/kg;有机碳,45.09 和 91.74 g/kg。

1.2 土壤微生物 DNA 提取

针对每种土壤,分别称取 0.5 g,各 3 个重复。用 FastDNA® Spin Kit for Soil 试剂盒(MP Biomedicals)提取土壤微生物基因组总 DNA,将提取得到的土壤 DNA 溶解于 70 μL 无菌 TE 缓冲液(10 mmol/L Tris-HCl, 1 mmol/L EDTA, pH 8.0),详细操作步骤请参考试剂盒说明书。通过微量紫外分光光度计(NanoDrop® ND-1000 UV-Vis)测定 DNA 浓度和纯度(OD_{260}/OD_{280} 和 OD_{260}/OD_{230})。利用 TE 缓冲液将土壤总 DNA 进行 10 倍梯度稀释后,进行 PCR 扩增,通过凝胶电泳分析土壤 DNA 中可能存在的腐殖质及其对 PCR 扩增的影响。土壤 DNA 保存于 -20℃ 冰箱待用。

1.3 焦磷酸高通量测序分析土壤微生物群落

1.3.1 PCR 扩增土壤微生物 16S rRNA 基因与焦磷酸 454 测序:针对森林和草地土壤的 6 个土壤 DNA,首先利用通用引物扩增其中的土壤微生物 16S rRNA 基因,随后纯化 PCR 产物后将其等摩尔混合上机测序,具体步骤如下。通用引物分别为 515F(5'-GTGCCAGCMGCCGCGG-3')和 907R(5'-CCGTCAATTCMTTTRAGTTT-3')^[11]。如表 1 所示,用于焦磷酸测序的通用引物含有 11 个碱基的特异 Tag 标签,用以区分不同的土壤样品。此外,每对引物中同时包含 adaptor 和 4 个碱基的 key 序列以提高下游测序效率,具体信息参考已有报道^[12]。PCR

扩增反应为: 0.25 μL 的 TaKaRa Ex Taq HS (5 U/ μL), 5.0 μL 的 $10 \times$ Buffer (Mg^{2+} Plus), 4.0 μL 的 dNTP Mixture (各 2.5 mmol/L), 1.0 μL 的 20 $\mu\text{mol/L}$ 引物, 加入 2.0 μL 稀释 10 倍的 DNA 模板至 50 μL 反应体系, 每次 PCR 反应均设置无菌水的阴性对照。PCR 扩增条件为: 94 $^{\circ}\text{C}$, 5 min; (94 $^{\circ}\text{C}$, 30 s; 55 $^{\circ}\text{C}$, 30 s; 72 $^{\circ}\text{C}$, 45 s), 33 个循环; 72 $^{\circ}\text{C}$ 10 min。获得扩增产物后, 利用 Agarose Gel DNA Fragment Recovery Kit Ver. 2.0 试剂盒 (TaKaRa) 进行切胶纯化, 并将其溶于 30 μL DNase-free H_2O 。进一步通过 2.0% 琼脂糖凝胶电泳检测 PCR 产物纯化效果, 测定纯化后 PCR 产物的浓度。将 6 个土壤样品 16S rRNA 基因的 PCR 纯化产物等摩尔数混合, 利用中国科学院南京土壤研究所测试中心的罗氏公司 454 FLX Titanium 测序仪上机分析。

1.3.2 焦磷酸高通量数据统计分析: 如表 1 所示, 草地 (GS) 和森林土壤 (FS) 高通量测序分别获得了 59733 和 56235 条原始序列, 平均每个样品的序列属分别为 19911 和 18745。利用 mothur 软件包对焦磷酸序列进行分析, 通过设置如下参数去除低质量序列, 包括最小长度 (minlength) 150 bp、最大长度 (maxlength) 600 bp, 最大模糊碱基数 (maxambig) 0, 最大同聚物长度 (maxhomop) 10, 平均质量分数

(qaverage) 20。处理后的草地 (GS) 和森林土壤 (FS) 分别得到 51208 和 47945 条高质量序列, 平均每个样品分别获得 17069 和 15982 条序列。进一步通过 RDP Classifier 软件包进行系统发育分析, 在 80% 置信度下将每个样品中的所有 16S rRNA 基因序列进行分类, 获得门、纲、目、科和属各水平下的分类单元。以草地土壤的一个重复样品为例, N 为所有序列之和, 划分到每个分类单元 (Taxon) 的序列数为 n_i 。根据公式 $P_i = n_i/N$, 计算每个分类单元的相对丰度 P_i 开展后续分析。基于高通量测序的微生物相对丰度 P_i , 简称高通量丰度。

不同分类水平下的土壤微生物多样性的计算策略如下。根据 mothur 软件包的设置, 默认 80%、90% 和 95% 的序列相似度可近似作为微生物不同门、目、属的分类标准。进一步根据公式 $H_{shannon} = -\sum_{i=1}^{S_{obs}} \frac{n_i}{N} \ln \frac{n_i}{N}$ 和 $D_{simpson} = \frac{\sum_{i=1}^{S_{obs}} n_i (n_i - 1)}{N(N - 1)}$ 计算每个样品的香农 (Shannon) 指数和辛普森 (Simpson) 指数。 $H_{shannon}$ 代表香农多样性指数, $D_{simpson}$ 代表辛普森多样性指数。 n_i 为每个分类单元中包含的序列数, N 为所有序列之和, S_{obs} 代表高通量测序所有检测到的分类单元总数。基于高通量测序的微生物群落多样性, 简称高通量多样性。

表 1. 高通量测序特异引物标签及测序结果概述

Table 1. Pyrosequencing summary of 16S rRNA genes in grassland soil (GS) and forest soil (FS).

Sample	Replicate	Tag sequence (5' \rightarrow 3')	Total reads	High quality reads	High quality reads average size
GS	1	CGAGCAGCGCT	19418	16453	389.4
	2	CGAGGACTACT	21571	18552	389.1
	3	CGATACTACGT	18744	16203	389.4
FS	1	CGTACGTCGAT	19151	16322	389.2
	2	CTACTCGTAGT	17481	14847	390.0
	3	GTACAGTACGT	19603	16776	389.3

1.4 变性梯度凝胶电泳 (DGGE) 指纹图谱分析土壤微生物群落

1.4.1 PCR 扩增 16S rRNA 基因及电泳: 如前所述, 利用通用引物 515F 和 907R 扩增土壤样品 16S rRNA 基因, 但该引物对不含焦磷酸测序的 adaptor 接头、Tag 标签和 key 序列等, 仅在 515F 引物前段添加 GC 夹 (GC-clamp: 5' -CGCCCGCCGCGCGCGGC GGGCGGGCGGGGGCACGGGGG-3')^[8], 以便更

好地分离不同 DGGE 条带^[13]。PCR 反应条件如前所述, 取 3.0 μL 的 PCR 产物, 通过 1.2% 琼脂糖凝胶电泳 PCR 扩增特异性。进一步利用 NanoDrop 定量 PCR 产物, 每个样品采用约 150 ng 的 16S rRNA 基因 PCR 产物进行 DGGE 电泳分析。DGGE 聚丙烯酰胺凝胶浓度为 8%, 变性梯度范围为 45% - 70%, 电泳条件为: 0.5 \times TAE 缓冲液, 80V 电压、60 $^{\circ}\text{C}$ 电泳 16h。电泳分析通过 Bio-Rad D-Code

System system 完成后,利用 SYBR 染料染色 30 min 进行后续分析。

1.4.2 DGGE 割胶、克隆和测序: DGGE 变性梯度凝胶利用 Bio-Rad 公司的凝胶成像系统 (Quantity One Bio-Rad, USA) 在紫外光下拍摄电泳图片。将 DGGE 凝胶上肉眼可辨的优势条带仔细切下并置于 0.5 mL 离心试管。利用 50 μ L 的无菌水反复冲洗 3 次,并通过移液器枪头将其尽量捣碎后加入 20 μ L 无菌水 4 $^{\circ}$ C 浸泡过夜。取 2.0 μ L 上清液作为模板,进一步利用引物 GC clamp-515F 和 907R 进行 PCR 扩增。PCR 扩增体系和反应条件如前所述。将 PCR 产物链接到 *Peasy-T3* 载体 (Trans) 并转化到 *E. coli* DH5 α 感受态细胞中,在含有氨苄青霉素的 LB 培养基上选择具有 Amp⁺ 抗性的白色转化子。通过 M13F 和 M13R 载体引物进行 PCR 扩增以验证阳性克隆。每个 DGGE 条带选择 2-3 个阳性克隆送到 Invitrogen 公司 (上海) 进行测序。DNA 序列分析如前所述,在 80% 的置信水平通过 RDP Classifier 进行分类,鉴定到门、纲、目、科和属不同水平。

1.4.3 DGGE 指纹图谱的统计分析: DGGE 电泳图谱,运用 Quantity One 软件,建立泳道,排除背景干扰 (Rolling Disk Size = 0-20),建立并校正条带,最后自动生成定量报告。报告中包括每个条带的轨迹定量值 Trace (Int \times mm),由条带的光密度和峰面积自动计算得到。根据 $P_i = n_i/N$, 计算相对丰度 P_i 。其中, n_i 为每个条带的轨迹定量值, N 为所有条带的轨迹定量值 n_i 之和。基于 DGGE 图谱的相对丰度 P_i , 简称 DGGE 丰度。

再根据公式 $H_{shannon} = - \sum_{i=1}^{S_{obs}} \frac{n_i}{N} \ln \frac{n_i}{N}$ 和 $D_{simpson} = \frac{\sum_{i=1}^{S_{obs}} n_i (n_i - 1)}{N(N-1)}$, 计算 DGGE 法推导的微生物群落多样性。 $H_{shannon}$ 代表香农多样性指数, $D_{simpson}$ 代表辛普森多样性指数, n_i 为 DGGE 图谱中条带的轨迹定量值, N 为所有条带的轨迹定量值 n_i 之和, S_{obs} 代表 DGGE 所有检测出来的条带总数。基于 DGGE 图谱的微生物群落多样性, 简称 DGGE 多样性。

1.5 显著性分析

利用 SPSS 16.0 软件,采用单因素方差 (One-way ANOVA) 分析高通量测序和 DGGE 两种方法得到的微生物丰度、多样性指数之间的差异, $P < 0.05$ 表示差异显著。

2 结果和分析

2.1 草地和森林土壤微生物群落结构

如图 1-A 所示,在微生物分类门的水平,高通量测序发现草地土壤和森林土壤拥有 9 个相同的优势微生物类群 (丰度 $> 0.7\%$ 为优势类群), 它们分别为变形菌门 (Proteobacteria)、放线菌门 (Actinobacteria)、酸杆菌门 (Acidobacteria)、拟杆菌门 (Bacteroidetes)、厚壁菌门 (Firmicutes)、疣微菌门 (Verrucomicrobia)、浮霉菌门 (Planctomycetes)、绿弯菌门 (Chloroflexi) 和芽单胞菌门 (Gemmatimonadetes)。这些微生物分别占草地和森林土壤微生物总量的 92.8% 和 94.3%。对于相对丰度 $< 0.7\%$ 的弱势微生物类群,草地和森林土壤中共同拥有 11 个门,但这些数量上占弱势的微生物类群仅占草地和森林土壤微生物总量的 1.52% 和 0.69%。此外,草地土壤含有 4 个独特的门 (丰度 $< 0.05\%$), 森林土壤含有 1 个特异的门,其丰度远远低于 0.01%。

如图 1-B 所示,草地和森林土壤的 DGGE 指纹图谱分别检测到 15 和 12 个优势条带并进行了克隆测序,代表了 DGGE 优势微生物。在微生物分类门水平,草地和森林土壤共有 5 个相同的微生物类群,包括 Proteobacteria, Acidobacteria, Bacteroidetes, Firmicutes 和 Planctomycetes。此外, Chloroflexi 仅存于草地土壤。

如表 2 所示,在微生物分类门、纲、目、科和属水平上,将高通量测序所有序列划分到已知的微生物单元,对微生物单元进行统计。以草地土壤为例,高通量测序发现 22 门、54 纲、60 目、131 科和 350 属,而 DGGE 技术仅检测到 6 门、9 纲、8 目、10 科和 10 属,前者检测灵敏度是后者的 3.7、6.0、7.5、13.1 和 35.0 倍。这一结果表明, DGGE 严重低估了土壤微生物的组成,分类水平越低,高通量测序检测到的微生物越多,与 DGGE 相比的优势愈加明显。森林土壤的分析也得到了类似的规律,在微生物门、纲、目、科和属分类水平上,高通量测序的检测灵敏度分别是 DGGE 的 3.8、6.7、6.4、19.2 及 39.4 倍。两种土壤的高通量测序和 DGGE 结果均表明在科和属的水平上,高通量测序与 DGGE 差异最大。

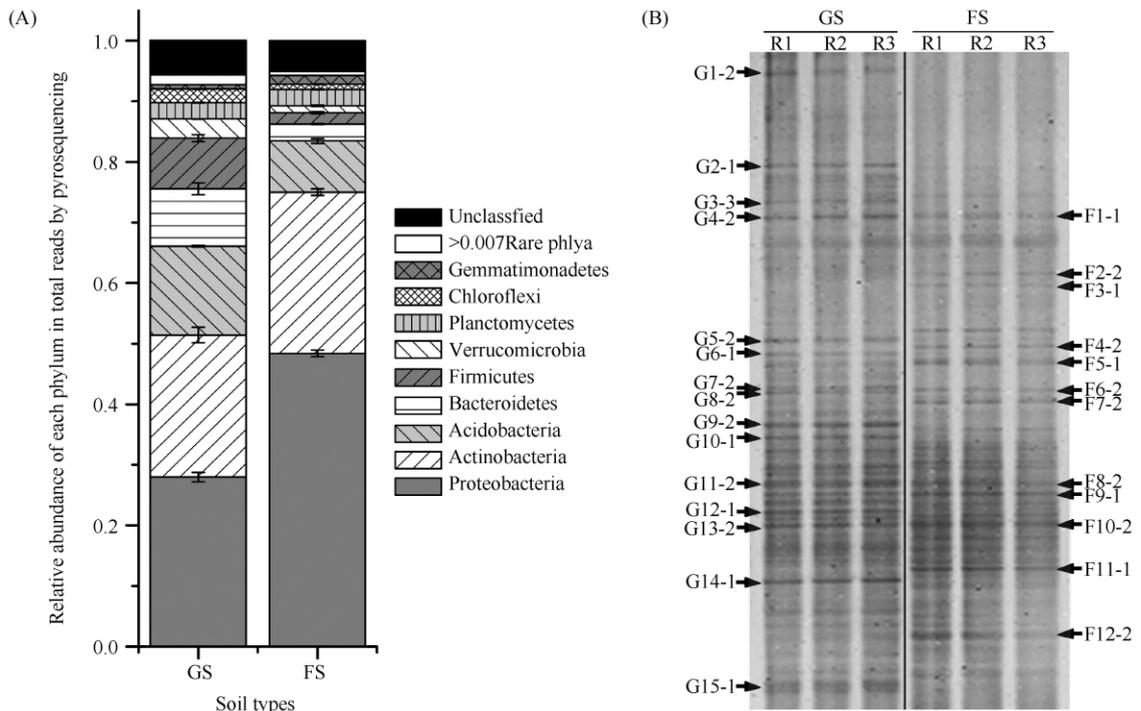


图 1. 草地 (GS) 和森林 (FS) 土壤微生物在门水平的高通量相对丰度 (A) 及 DGGE 指纹图谱 (B)

Figure 1. Microbial community structure at phylum level by pyrosequencing (A) and DGGE fingerprints (B) of 16S rRNA genes in grassland soil (GS) and forest soil (FS).

表 2. 高通量测序及 DGGE 技术检测到的微生物单元总数

Table 2. Microbial taxon number of grassland (GS) and forest (FS) soils by DGGE and pyrosequencing

Sample	Taxonomic level	Number of microbial taxon		Ratio of pyrosequencing to DGGE
		Pyrosequencing	DGGE	
GS	Phylum	22 ± 1	6	3.7
	Class	54 ± 2	9	6.0
	Order	60 ± 1	8	7.5
	Family	131 ± 4	10	13.1
	Genus	350 ± 12	10	35.0
FS	Phylum	19 ± 1	5	3.8
	Class	47 ± 1	7	6.7
	Order	51 ± 2	8	6.4
	Genus	276 ± 10	7	39.4

2.2 高通量测序土壤微生物的群落组成与 DGGE 的相关比较

如表 3 所示,分类水平越低、越多的高通量序列不能被归属到已知的微生物单元。尽管如此, DGGE 检测水平仍然远远低于高通量测序。如图 2 所示,高通量测序发现草地土壤 (GS) 微生物在分类门、纲、目、科和属上分别占微生物总量的 94.34%、91.57%、64.08%、56.02% 和 64.48%, 而 DGGE 主要微生物只占微生物总量的 65.42%、47.44%、31.46%、11.98% 和 10.04%。换言之, DGGE 在上

述各分类水平分别有 28.78%、43.74%、32.20%、43.12% 和 50.88% 的微生物无法检测到。分类水平越低, DGGE 的检测灵敏度越低, 尤其是在科和属的水平上, 高通量的检测灵敏度是 DGGE 的 4-6 倍。森林土壤微生物的高通量测序和 DGGE 比较分析结果与草地土壤规律一致。

此外, 高通量测序检测到的一些优势微生物类群也无法被 DGGE 检测 (图 2), 尤其是在科和属的水平上。以森林土壤为例, 在微生物分类门水平, DGGE 反映的微生物均为高通量优势类群, 但未检出全部 9

个优势类群,如 Actinobacteria (26.6%)、Verrucomicrobia (1.14%) 和 Gemmatimonadetes (1.45%),以及所有的高通量测序中弱势类群。在分类科的水平上,高通量检测出了21个优势类群,占微生物总量的48.91%;DGGE仅检测出了其中3个,占微生物总量的10.46%,慢生根瘤菌科(Bradyrhizobiaceae: 4.89%),生丝微生物科(Hyphomicrobiaceae: 4.60%),康奈斯氏杆菌科(Conexibacteraceae: 3.10%),假诺卡氏菌科

(Pseudonocardiaceae: 2.54%)等优势类群均未被DGGE检出。在分类属水平上,高通量检测出了22个优势类群,占微生物总量的42.11%;DGGE只检测出了其中4个,占微生物总量的13.25%,慢生根瘤菌属(Bradyrhizobium: 3.69%),康奈斯氏杆菌属(Conexibacter: 3.10%),酸杆菌属 Gp1 属(2.20%),假诺卡氏菌属(Pseudonocardia: 2.18%)等均未被DGGE检出。

表3. 高通量序列被归属到已知微生物分类单元的比例

Table 3. Ratio of classified taxa in GS and FS soils by pyrosequencing

Taxonomic level	Ratio of classified taxa in GS, /%		Ratio of classified taxa in FS, /%	
	Taxon > 0	Taxon > 0.05%	Taxon > 0	Taxon > 0.05%
Phylum	94.34 ± 0.09	94.20 ± 0.14	94.97 ± 0.20	94.81 ± 0.24
Class	91.57 ± 0.29	91.18 ± 0.32	93.57 ± 0.30	93.19 ± 0.27
Order	64.08 ± 0.69	63.66 ± 0.71	75.78 ± 0.54	75.41 ± 0.50
Family	56.02 ± 1.27	55.10 ± 1.24	59.43 ± 0.56	58.53 ± 0.52
Genus	64.48 ± 0.96	60.92 ± 0.94	63.59 ± 0.71	60.64 ± 0.49

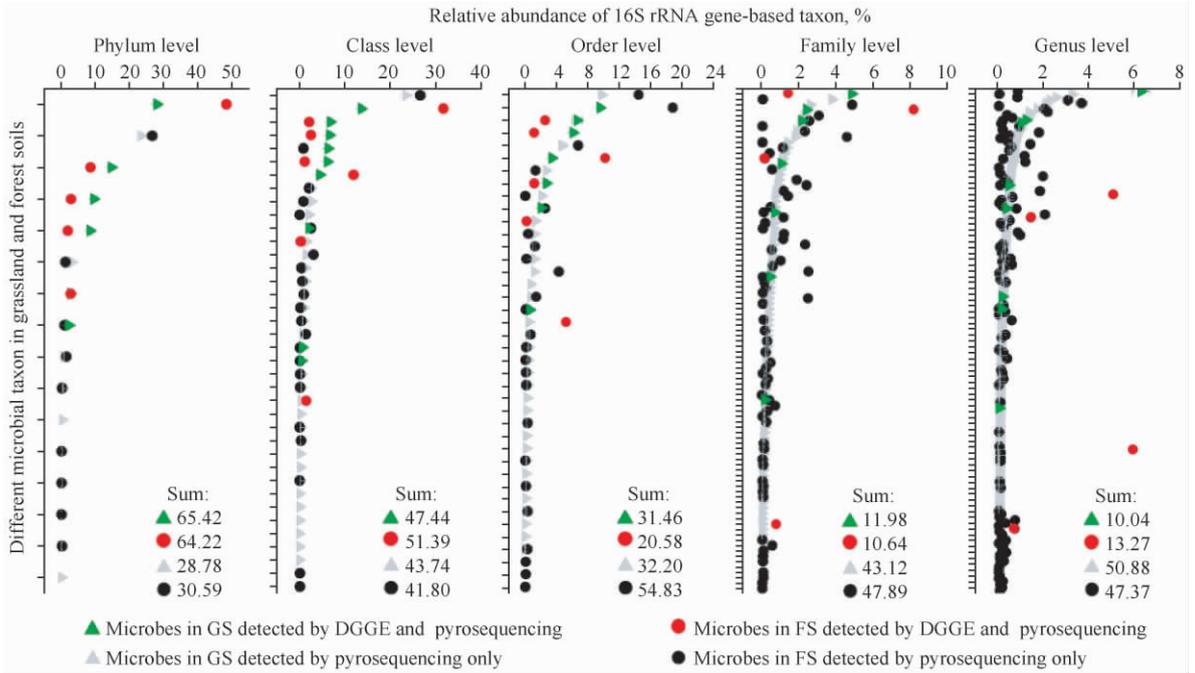


图2. 高通量测序和 DGGE 指纹图谱比较分析草地和森林土壤土壤微生物在不同分类水平上的群落组成

Figure 2. Comparative analysis of microbial community structure by DGGE and pyrosequencing of 16S rRNA genes in grassland soil (GS) and Forest soil (FS) at different taxonomic levels. The solid triangles in green and gray represent pyrosequencing and DGGE data from GS soil, respectively. The solid cycles in red and black represent pyrosequencing and DGGE data from FS soil, respectively. Taxon with the relative abundance more than 0.05% was included only.

2.3 DGGE 分析土壤微生物群落结构与高通量测序的相关比较

比较 DGGE 指纹图谱和高通量测序量化的两种

土壤微生物丰度(分别简称 DGGE 丰度和高通量丰度),发现草地土壤在分类门水平有6个类群,其中4个 DGGE 丰度显著高于高通量丰度($P < 0.05$),如

Proteobacteria (DGGE: 35.0% ; Pyrosequencing: 28.0%)、 Bacteroidetes (DGGE: 21.3% ; Pyrosequencing: 9.50%)、 Chloroflexi (DGGE: 7.12% ; Pyrosequencing: 2.22%)、 Planctomycetes (DGGE: 4.03% ; Pyrosequencing: 2.72%)。换言之, 假设以高通量丰度为标准, 在分类门水平, 67% (4/6) 的微生物的丰度有被 DGGE 过高估计的可能。类似地, 在纲、目、科和属水平, 分别有 56%、63%、90% 和 80% 的微生物存在高估的可能 ($P < 0.05$)。森林土壤与草地土壤的分析结果类似, 依次按上述分类水平, DGGE 高估微生物的概率分别为 60%、86%、88%、100% 和 86%。也就是说, 在各分类水平上, DGGE 均高估了大多数土壤微生物的相对丰度, 而且分类水平越低, DGGE 高估微生物丰度的可能性就越大, 特别是在科和属的水平上。DGGE 丰度与高通量丰度相符程度不超过 22%。

如图 3 所示, 以电泳图谱量化丰度表示, DGGE 优势微生物在微生物总量中的相对分度普遍高于其在高通量总体中的比例, 尤其在科和属的水平上。以草地土壤为例, 在分类门水平, DGGE 图谱 6 个优势微生物门, 占总丰度的 84%, 其在高通量结果中的反映为 63%, DGGE 是高通量的 1.3 倍。也就是说 DGGE 对优势微生物整体丰度的估计是高通量测序的 1.3 倍。同理, 在纲、目、科和属水平上, DGGE 分别是高通量测序结果的 1.8、2.3、4.7 和 3.1 倍。森林土壤与草地土壤的规律类似, 按分类水平由高到低, DGGE 分析结果分别偏高 1.0、1.5、2.9、4.7 和 3.8 倍。

就单个微生物组分的丰度而言, DGGE 和高通量测序两种方法的差异更加显著 (表 4)。以草地土壤为例, 在科水平上, 产碱菌科 (Alcaligenaceae) 的 DGGE 丰度为 4.80%, 高通量丰度只有 0.004%, 差异约为 1200 倍; 丛毛单胞菌科 (Comamonadaceae) 的 DGGE 丰度为 10.0%, 高通量丰度只有 0.48%, 差异大于 20 倍; 肠杆菌科 (Enterobacteriaceae) 的 DGGE 丰度为 2.63%, 高通量测序未检测到。在属水平上, 食酸菌属 (*Acidovorax*) 的 DGGE 丰度为 3.45%, 高通量丰度 0.08%, 两者差异大于 40 倍; 酸杆菌 *Gp5* 属的 DGGE 丰度为 6.94%, 高通量丰度 0.41%, 差异大于 13 倍; 沙雷氏菌属 (*Serratia*: 2.63%) 和苍白杆菌属 (*Ochrobactrum*: 1.14%) 未被高通量测序方法检测到。与高通量测序方法相比,

DGGE 法高估了某些微生物组分。森林土壤中也有类似情况, 如 Enterobacteriaceae, 被高估 2000 倍以上; 新鞘氨醇杆菌属 (*Novosphingobium*), 被高估 200 倍以上; 拉乌尔菌属 (*Raoultella*: 4.15%) 和牻牛瘤胃菌属 (*Proteiniclasticum*: 2.14%) 同样未被高通量测序方法检测到。

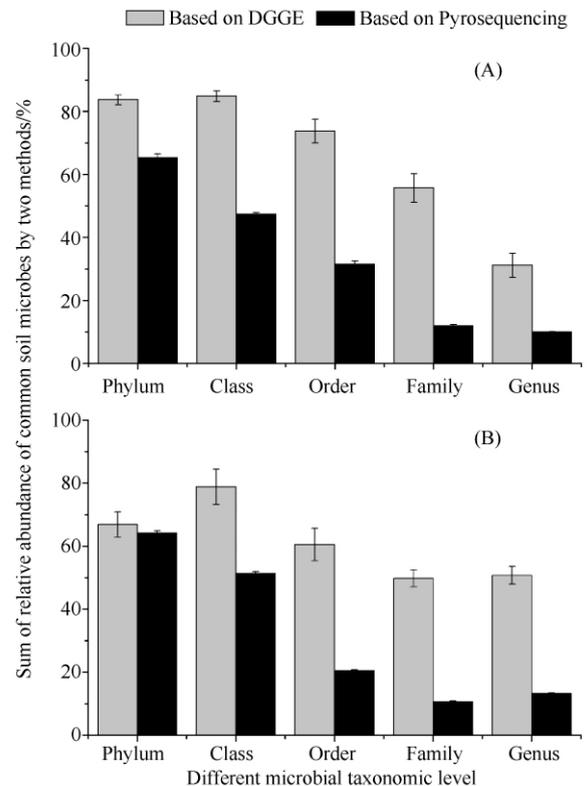


图 3. 以 DGGE 为基础的草地 (A) 和森林 (B) 土壤优势微生物在各分类水平的总丰度以及其在高通量测序结果中的反映

Figure 3. The total abundance of common microbes calculated from both DGGE band intensity and pyrosequencing reads. GS (A) and FS (B) represent grassland soil and forest soil, respectively.

2.4 高通量测序和 DGGE 比较分析土壤微生物多样性

DGGE 条带的数量和亮度可在一定程度反映微生物群落结构, 如表 5 所示, 基于 DGGE 指纹图谱的草地和森林土壤微生物 Shannon 指数分别为 3.43 和 3.26, Simpson 指数分别为 0.9603 和 0.9551。基于高通量测序的群落多样性分析表明, 草地土壤和森林在 80% 序列相似度 (门水平) 下, Shannon 指数分别为 5.07 和 4.75, Simpson 指数分别为 0.9878 和 0.9829, 显著高于 DGGE 多样性指数 ($P < 0.05$); 在 95% 序列相似度 (属水平) 下, Shannon 指数分别为

7.12 和 6.93, Simpson 指数分别为 0.9981 和 0.9976, 远高于 DGGE 多样性指数。两种土壤之间多样性的比较, 在整体趋势上, DGGE 与高通量测序结果一致, 均表明草地较森林土壤的多样性偏高。

但在各个分类水平下, DGGE 为基础的群落多样性均显著低于高通量测序为基础的群落多样性 ($P < 0.05$)。在分类门水平下两种方法表示的多样性指数之间相差最小, 在属水平下差异最大。

表 4. DGGE 为基础的微生物丰度及其在高通量测序结果中的反映

Table 4. Estimated relative abundance of common microbes by both DGGE and pyrosequencing methods

Sample	Taxonomic level	Microbial components	DGGE bands	Microbial relative abundance, %	
				DGGE	Pyrosequencing
GS	Family	Chitinophagaceae	G2-G4	11.9 ± 2.36	4.88 ± 0.43
		Comamonadaceae	G1, G4, G8	10.0 ± 1.66	0.48 ± 0.09
		Anaerolineaceae	G11	7.12 ± 0.40	0.72 ± 0.09
		Burkholderiales_incertae_sedis	G14	5.85 ± 1.14	1.08 ± 0.17
		Phyllobacteriaceae	G5	4.80 ± 1.67	0.20 ± 0.01
		Alcaligenaceae	G5	4.80 ± 1.67	0.004 ± 0.003
		Planctomycetaceae	G10	4.03 ± 0.37	2.16 ± 0.15
		Xanthomonadaceae	G1	3.45 ± 1.37	2.46 ± 0.13
		Enterobacteriaceae	G6	2.63 ± 0.92	—
		Bruceaceae	G7	1.14 ± 0.56	0.002 ± 0.004
	Genus	<i>Gp5</i>	G9	6.94 ± 0.91	0.51 ± 0.04
		<i>Terrimonas</i>	G2	3.86 ± 0.06	1.08 ± 0.09
		<i>Dokdonella</i>	G1	3.45 ± 1.37	0.41 ± 0.02
		<i>Acidovorax</i>	G1	3.45 ± 1.37	0.08 ± 0.02
		<i>Gp6</i>	G3	2.80 ± 1.98	6.31 ± 0.39
		<i>Ferruginibacter</i>	G3	2.80 ± 1.98	1.27 ± 0.22
		<i>Chitinophaga</i>	G3	2.80 ± 1.98	0.19 ± 0.04
		<i>Serratia</i>	G6	2.63 ± 0.92	—
		<i>Variovorax</i>	G8	1.32 ± 0.37	0.18 ± 0.03
		<i>Ochrobactrum</i>	G7	1.14 ± 0.56	—
FS	Family	Xanthomonadaceae	F3, F4, F6, F10-F12	26.5 ± 1.83	8.20 ± 0.18
		Burkholderiaceae	F2, F5	8.39 ± 0.62	0.78 ± 0.03
		Chitinophagaceae	F1	4.41 ± 0.97	1.48 ± 0.09
		Sphingomonadaceae	F8	4.24 ± 1.78	0.18 ± 0.00
		Enterobacteriaceae	F9	4.15 ± 0.46	0.002 ± 0.004
	Genus	Clostridiaceae 1	F2	2.14 ± 0.70	0.03 ± 0.01
		<i>Rhodanobacter</i>	F3, F4, F6, F10, F11	16.9 ± 0.91	5.96 ± 0.16
		<i>Gp2</i>	F12	9.54 ± 2.73	1.47 ± 0.07
		<i>Burkholderia</i>	F2, F5	8.39 ± 0.62	0.74 ± 0.03
		<i>Rhizomicrobium</i>	F7	5.36 ± 1.79	5.08 ± 0.17
<i>Novosphingobium</i>	F8	4.24 ± 1.78	0.02 ± 0.02		
<i>Raoultella</i>	F9	4.15 ± 0.46	—		
<i>Proteiniclasticum</i>	F2	2.14 ± 0.70	—		

表 5. DGGE 及高通量测序估算土壤微生物的多样性

Table 5. Estimated microbial diversity indexes of GS and FS soils by DGGE and pyrosequencing

Index	Sample	DGGE	Pyrosequencing		
			80% similarity	90% similarity	95% similarity
$H_{shannon}$	GS	3.43 ± 0.04	5.07 ± 0.05	6.31 ± 0.05	7.12 ± 0.05
	FS	3.26 ± 0.03	4.75 ± 0.01	5.99 ± 0.04	6.93 ± 0.03
$D_{simpson}$	GS	0.9630 ± 0.0018	0.9878 ± 0.0010	0.9962 ± 0.0003	0.9981 ± 0.0001
	FS	0.9551 ± 0.0034	0.9829 ± 0.0009	0.9942 ± 0.0005	0.9976 ± 0.0001

3 讨论

土壤中微生物数量巨大、种类繁多,土壤微生物的研究很大程度上依赖于方法和技术的进步。假设本研究土壤中每个微生物细胞平均拥有 5 个 16S rRNA 基因拷贝^[14],实时荧光定量 PCR 结果表明,草地和森林土壤中微生物细胞数量分别为 3800 万和 2400 万。但 DGGE 图谱中每个样品的可视条带数量仅为几十条,似乎很难全面反映土壤微生物的数量和物种组成。而高通量测序每个样品则至少能够获得上千条序列,本研究中单个样品的微生物 16S rRNA 基因序列不少于 14800 条,在微生物门、纲、目、科、属的分类水平上,其检测灵敏度是 DGGE 的 3.8、6.7、6.4、19.2 及 39.4 倍,高通量测序具有明显的检测优势。

21 世纪以来,DNA 测序技术取得革命性突破,新一代高通量测序 16S rRNA 基因极有可能成为一种常规手段分析土壤微生物的物种多样性变化规律,成为土壤质量评价的常规方法^[15]。事实上,高通量测序土壤微生物多样性的基本原理与传统的分子指纹图谱基本一致。两者都以土壤样品的 DNA 作为模板,获得土壤微生物群落 16S rRNA 基因的 PCR 扩增产物后,开展下游分析。高通量测序法具有分析样本量多、单个样品序列通量高的优点。DGGE 不仅需要进行繁琐的凝胶制作和电泳分离等下游工作,而且具有较大的随机性,很难检测到低丰度甚至是个别高丰度的土壤微生物类群。例如,本研究两种土壤中,高通量测序发现 Actinobacteria 均占总体微生物的 20% 以上,但 DGGE 技术则未能检测到丰度上占优势的 Actinobacteria。值得注意的是,本研究中高通量测序和 DGGE 分析的 PCR 扩增引物相同,表明 DGGE 技术分析土壤微生物主要类群的技术条件仍待进一步优化。Rettedal 等研究发现,采用含 40 bp GC 夹的引物进行 PCR 扩增,可能导致相同 16S rRNA 基因序列的扩增片段迁移到 DGGE 的不同位置产生多个条带,造成对物种丰度的过高或过低估计^[16]。此外,G/C 含量等因素也可能影响 PCR 产物的 DGGE 凝胶电泳分离效果,甚至产生基因片段共迁移的现象^[17]。例如,本研究中同一个条带检测出了归属于不同分类单元的微生物 16S rRNA 序列,也导致此类微生物的丰度难以准确

估计。克服这些技术问题,需要多次尝试聚丙烯酰胺凝胶的浓度和变性剂的梯度,费时费力,且不易得到理想的结果。事实上,由于受到仪器灵敏度等主观客观因素的影响,实际操作过程中很难将 DGGE 图谱上所有条带全部割胶测序,也很难对同一个条带进行构建大量克隆文库,排除不同 16S rRNA 片段的共迁移现象。因此,传统基于条带亮度的微生物丰度分析,基于条带序列的微生物组成研究仍需其它方法的补充验证。

然而,与高通量测序技术研究微生物多样性相比,DGGE 也具有一些不可替代的优点,能够更加直观地比较和分析微生物群落结构的变化规律。例如,我们最近的研究表明,尿素培养酸性森林土壤 8 周后,土壤中硝态氮浓度显著增加,表明土壤中发生了强烈的硝化作用。同时,利用 DGGE 技术分析氨氧化古菌 *amoA* 功能基因的遗传多样性表明,与未培养的零时刻土壤相比,尿素培养土壤中氨氧化古菌发生了明显的变化,表明古菌极可能主导了酸性土壤氨氧化^[12]。类似采用硝化抑制剂的研究也表明,泉古菌特异 *amoA* 和 16S rRNA 基因的 DGGE 指纹图谱与土壤氨氧化通量变化规律高度吻合,这些结果表明古菌在酸性土壤氨氧化过程中发挥了重要作用^[18]。由于 DGGE 可检测 DNA 片段中 1 bp 碱基的变异^[13],该技术研究土壤中特定功能微生物的变化规律如氨氧化古菌和甲烷氧化菌,则具有一定的技术优势。

高通量测序能够较为全面和准确的反映土壤微生物群落结构,而 DGGE 仅能分析有限的优势微生物类群,存在高估物种丰度以及低估微生物群落大小和多样性的可能。DGGE 技术能够直观、定性地比较土壤微生物群落的差异;高通量测序技术则能够在整体微生物群落水平分析物种遗传多样性,并能较为客观地反映其中低丰度的重要功能微生物。随着高通量测序成本不断降低,海量数据分析方法日趋成熟,高通量测序将成为一项常规的实验手段,在土壤微生物学研究中发挥重要作用。

致谢:感谢新西兰农业科学院 (AgResearch Ltd) Bowatte Saman 博士和新西兰土地保护研究中心 (Landcare Research) Kevin Tate 博士提供土壤样品。感谢中国科学院南京土壤研究所黄蓉在高通量测序过程中提供的技术帮助。

参考文献

- [1] Song C, Wu J, Lu Y, Shen Q, He J, Huang Q, Jia Z, Leng S, Zhu Y. Advances of soil microbiology in the last decade in China. *Advance in Earth Science*, 2013, 28(10): 1087-1105. (in Chinese)
宋长青, 吴金水, 陆雅海, 沈其荣, 贺纪正, 黄巧云, 贾仲君, 冷疏影, 朱永官. 中国土壤微生物学研究 10 年回顾. *地球科学进展*, 2013, 28(10): 1087-1105.
- [2] Pace NR. Problems with "Prokaryote". *Journal of Bacteriology*, 2009, 191(7): 2008-2010.
- [3] van der Heijden MGA, Bardgett RD, and van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters*, 2008, 11(3): 296-310.
- [4] Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, 1977, 74(11): 5088-5090.
- [5] Pace NR. A molecular view of microbial diversity and the biosphere. *Science*, 1997, 276(5313): 734-740.
- [6] Wintzingerode FV, Göbel UB, Stackebrandt E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews*, 1997, 21(3): 213-229.
- [7] Stahl DA, Lane DJ, Olsen GJ, Pace NR. Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. *Science*, 1984, 224(4647): 409-411.
- [8] Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*, 1993, 59(3): 695-700.
- [9] Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF. Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 2013, 499(7459): 431-437.
- [10] Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD. Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal*, 2007, 1(4): 283-290.
- [11] Xia WW, Zhang CX, Zeng XW, Feng YZ, Weng JH, Lin XG, Zhu JG, Xiong ZQ, Xu J, Cai ZC, Jia ZJ. Autotrophic growth of nitrifying community in an agricultural soil. *The ISME Journal*, 2011, 5: 1226-1236.
- [12] Lu L, Han WY, Zhang JB, Wu YC, Wang BZ, Lin XG, Zhu JG, Cai ZC, Jia ZJ. Nitrification of archaeal ammonia oxidizers in acid soils is supported by hydrolysis of urea. *The ISME Journal*, 2012, 6(10): 1978-1984.
- [13] Sheffield VC, Cox DR, Lerman LS, Myers RM. Attachment of a 40-base-pair (G + C)-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National Academy of Sciences of the United States of America*, 1989, 86(1): 232-236.
- [14] Shrestha PM, Noll M, Liesack W. Phylogenetic identity, growth-response time and rRNA operon copy number of soil bacteria indicate different stages of community succession. *Environmental Microbiology*, 2007, 9(10): 2464-2474.
- [15] Zheng Y, Jia Z. Next generation sequencing and stable isotope probing of active microorganisms responsible for aerobic methane oxidation in red paddy soils. *Acta Microbiologica Sinica*, 2013, 53(2): 173-184. (in Chinese)
郑燕, 贾仲君. 新一代高通量测序与稳定性同位素示踪 DNA/RNA 技术研究稻田红壤甲烷氧化的微生物过程. *微生物学报*, 2013, 53(2): 173-184.
- [16] Rettedal EA, Clay S, Brözel VS. GC-clamp primer batches yield 16S rRNA gene amplicon pools with variable GC clamps, affecting denaturing gradient gel electrophoresis profiles. *FEMS Microbiology Letters*, 2010, 312(1): 55-62.
- [17] Sekiguchi H, Tomioka N, Nakahara T, Uchiyama H. A single band does not always represent single bacterial strains in denaturing gradient gel electrophoresis analysis. *Biotechnology Letters*, 2001, 23(15): 1205-1208.
- [18] Offre P, Prosser JI, Nicol GW. Growth of ammonia-oxidizing archaea in soil microcosms is inhibited by acetylene. *FEMS Microbiology Ecology*, 2009, 70(1): 99-108.

Comparative analysis of soil microbial communities by pyrosequencing and DGGE

Weiwei Xia^{1,2}, Zhongjun Jia^{1*}

¹ State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, Jiangsu Province, China

² University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: [Objective] We aimed to assess the advantage and disadvantage of next-generation pyrosequencing and traditional Denaturing Gradient Gel Electrophoresis (DGGE) in fingerprinting analysis of soil microbial communities.

[Methods] We analyzed microbial compositions, abundance and diversity of typical grassland and forest soils by 16S rRNA gene-based pyrosequencing and DGGE to compare the accuracy and reproducibility of the two techniques on soil microbial communities. [Results] For grassland soils, pyrosequencing technique revealed 22 phyla, 54 classes, 60 orders, 131 families and 350 genera; DGGE only detected 6 phyla, 9 classes, 8 orders, 10 families and 10 genera. The results show that DGGE greatly underestimated soil community compositions. Similar results were obtained for forest soils, and the detection sensitivity of pyrosequencing of forest soils was 3.8, 6.7, 6.4, 19.2 and 39.4 times higher than that of DGGE at the taxonomic levels of phylum, class, order, family and genera respectively. Furthermore, DGGE overwhelmingly overestimated the relative abundance of dominant microorganisms represented by the high-intensity bands, leading up to a 2000-fold difference. Both DGGE and pyrosequencing showed consistent results of microbial diversity changing patterns, although the DGGE-based diversity index was much lower than pyrosequencing. [Conclusion] Pyrosequencing thus provides more comprehensive and accurate fingerprints of soil microbial community structure than DGGE. DGGE only can represent a few numerically dominant phylotypes with apparent overestimation of their relative abundance in soil microbial communities.

Keywords: pyrosequencing, denaturing gradient gel electrophoresis (DGGE), soil microbial community

(本文责编:王晋芳)

Supported by the "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDB15040000)

* Corresponding author. Tel: +86-25-86881311; Fax: +86-25-86881000; E-mail: jia@issas.ac.cn

Received: 17 February 2014 / Revised: 3 July 2014