

基于高通量测序技术的丛枝菌根真菌多样性研究方法

High-throughput Sequencing-based Method for Characterizing Biodiversity of Arbuscular Mycorrhizal Fungi

付伟¹, 武慧¹, 陈保冬^{1,*}

¹城市与区域生态国家重点实验室, 中国科学院生态环境研究中心, 北京

*通讯作者邮箱: bdchen@rcees.ac.cn

摘要: 丛枝菌根 (Arbuscular mycorrhizal, AM) 真菌是一类可以与大多数陆地植物形成专性共生体系的土壤真菌类群。在过去的十几年中, 高通量测序技术已经彻底改变了菌根生态学的研究方法, 使得我们有能力深入探究不同环境中的 AM 真菌群落。相较于传统的形态学鉴定和 Sanger 测序方法, 高通量测序可以快速、高通量、较准确的分析大量样品的 AM 真菌群落组成, 使得区域和全球尺度上的 AM 真菌多样性和生物地理学研究成为可能。测序方法相较于传统方法具有更高的检测灵敏度, 目前通过形态学鉴定大约定义了 300 种 AM 真菌, 而通过分子手段定义的物种数超过 1000 种。虽然高通量测序技术已经广泛应用于微生物的群落研究中, 但是由于 AM 真菌在生物学和遗传学上的特异性, 使得该方法在 AM 真菌群落研究中的应用存在一些认识误区。本方法以 AM 真菌的 rRNA 小亚基 (Small subunit rRNA, SSU) 测序为例, 基于二代测序平台 (Illumina MiSeq) 和最新的去噪数据分析方法 DADA2, 介绍了 AM 真菌多样性的研究方法。本方法适用于土壤和植物根系 AM 真菌多样性的研究。

关键词: 丛枝菌根真菌, 多样性, rRNA 小亚基 (Small subunit rRNA, SSU), DADA2

材料与试剂

1. 常规试剂与材料

1.1 剪刀

1.2 镊子

1.3 酒精灯

1.4 液氮

1.5 2 mm 土壤筛

1.6 塑封袋

1.7 无菌培养皿

1.8 无菌手套

1.9 ddH₂O (双蒸水, DNase-Free)

2. FastDNATM SPIN Kit (MP Biomedicals, 116560200) , 保存温度: 15-30 °C

3. 1.5 ml/2 ml 无菌离心管 (Axygen, MCT-150-C-S-1)

4. 高保真 DNA 聚合酶 (Thermo scientific, F530L) , -20 °C 保存

5. 引物 (生工生物, WANDA/AML2) , 溶解后-20 °C 保存

仪器设备

1. 样本裂解仪 (MP Biomedicals, Fastprep-24 5G)

2. PCR 仪 (Eppendorff, Mastercycler pro)

3. NanoDrop 1000 (Thermo scientific, ND-3300)

4. 冷冻干燥机 (LABCONCO, FreeZone 4.5)

5. 离心机 (Thermo scientific, Heraeus Fresco 21)

6. 涡旋混合器 (MoBio Laboratories, G560E)

7. 电子天平 (Mettler Toledo, ML54)

软件和数据库

1. R statistics (v3.6.0, <https://www.r-project.org/>)

2. R package: DADA2 (v1.14.0, <https://benjjneb.github.io/dada2/index.html>)

3. R package: vegan (v2.5.6, <https://cran.r-project.org/web/packages/vegan/index.html>)

4. MaarjAM database (下载日期: 2020.01.16, <https://maarjam.botany.ut.ee/>)

实验步骤

1. 土壤和根系样品的前处理

1.1 根系样品

将根系样品低温保存运送到实验室, 用水将根系表面黏附的土壤和有机质清洗干净

(必要时可使用超声清洗)，并迅速用吸水纸将根系擦干，置于无菌培养皿中；继而用无菌剪刀将根系剪成 0.5~1 cm 根段，随机称取~250 mg 根系样品于 1.5 ml 离心管中，然后使用液氮冷冻后将样品冻干，冻干样品于-20 °C 保存。

1.2 土壤样品

将土壤样品低温保存运送到实验室，过 2 mm 筛后，充分混匀，去除所有植物根系和残体，然后称取~500 mg 土壤样品置于 2 ml 离心管中，使用液氮冷冻后将样品冻干，冻干样品于-20 °C 保存。

注：在样品前处理过程中使用的剪刀和镊子等与样品直接接触的实验器具须使用酒精灯灼烧，消除 DNA 污染，且处理间器具不可混用。

2. DNA 的提取

推荐使用商业 DNA 提取试剂盒，以提高 DNA 提取的稳定性，如使用 FastDNA™ SPIN Kit (MP Biomedicals, CA, USA) 配合 MP Fastprep-24™ 5G 样本裂解仪提取土壤 DNA。根系样品经液氮研磨后也可使用此方法对其 DNA 进行提取，DNA 浓度使用 NanoDrop 1000 超微量分光光度计进行检测。

DNA 提取步骤参见：<https://www.mpbio.com/media/productattachment/LS082019-EN-FastDNA-SPIN-Kit-for-Soil-116560200-Manual.pdf>

3. 扩增引物的选择

目前，在使用靶向扩增子测序研究 AM 真菌群落时主要考虑三个基因片段：rRNA 小亚基 (Small Subunit rRNA, SSU), 转录间隔区 (Internal Transcribed Spacer, ITS) 和 rRNA 大亚基 (Large Subunit rRNA, LSU) (图 1)。SSU rRNA 是 AM 真菌多样性研究中使用最广泛的区域，此区域的引物可以在科的水平上扩增绝大多数 AM 真菌类群 (Egan *et al.*, 2018, Lekberg *et al.*, 2018)，覆盖度较好，但是因为此区域相对较为保守，因此在种水平的区分度并不高 (Bruns and Taylor, 2016)，更适合于较高分类水平下的相关研究。相较于其它研究区域，SSU rRNA 拥有较为完善的数据库 MaarjAM (<https://maarjam.botany.ut.ee/>)，可以用于划分虚拟分类单元 (Virtual taxa, VT)，将形态学与分子生物学鉴定相结合，极大的促进了不同研究之间的横向比较 (Öpik *et al.*, 2010)。ITS 是真菌多样性研究的通用区域，其在 AM 真菌多样性研究中的应用次之，但有逐渐增多的趋势。ITS 区在 AM 真菌多样性研究中的应用主要受限于两点：(1) 通用真菌 ITS 引物对 AM 真菌的扩增效率较低；(2)

ITS 序列在 AM 真菌种内的变异度较大，因此难以进行进化关系的分析。然而，ITS 区域相较于 SSU rRNA 变异性更高，在低分类水平下的分辨率较高，因此其检测 AM 真菌群落响应的敏感性可能更高 (Gao *et al.*, 2018)。此外，ITS 作为真菌多样性研究的通用区域，其拥有较为完善的数据库 (如 UNITE (Kõljalg *et al.*, 2013))，以及可以在真菌群落水平揭示 AM 真菌的群落响应和相互作用关系。LSU rRNA 在 AM 真菌多样性的研究中应用最少，此区域拥有较为合适的变异度，但是其引物的偏好性较强 (Kohout *et al.*, 2014)，且多数扩增区域较长，不适用于二代测序。此外，还可以联合 SSU、ITS 和 LSU 三个区域，使用 PacBio 三代测序平台开展 AM 真菌多样性研究，此方法极大的提高了相关研究在 AM 真菌低分类水平下的区分度，可以在种水平上考察 AM 真菌的群落变化 (Krüger *et al.*, 2009; Schlaeppi *et al.*, 2016)，但是此方法的成本相对较高，对实验操作和数据分析要求较高。综上所述，每一个扩增区域都有自己的优缺点和适用范围 (表 1)，因此在选择扩增区域和引物时应明确要解决的科学问题和研究目标，以便采用更合适的实验方法。结合以上分析，为了更广泛的适用范围，本方法以最常用的 SSU rRNA 测序为例，使用 WANDA/AML2 引物组合 (表 1)，介绍与其相关的 AM 真菌多样性研究方法。

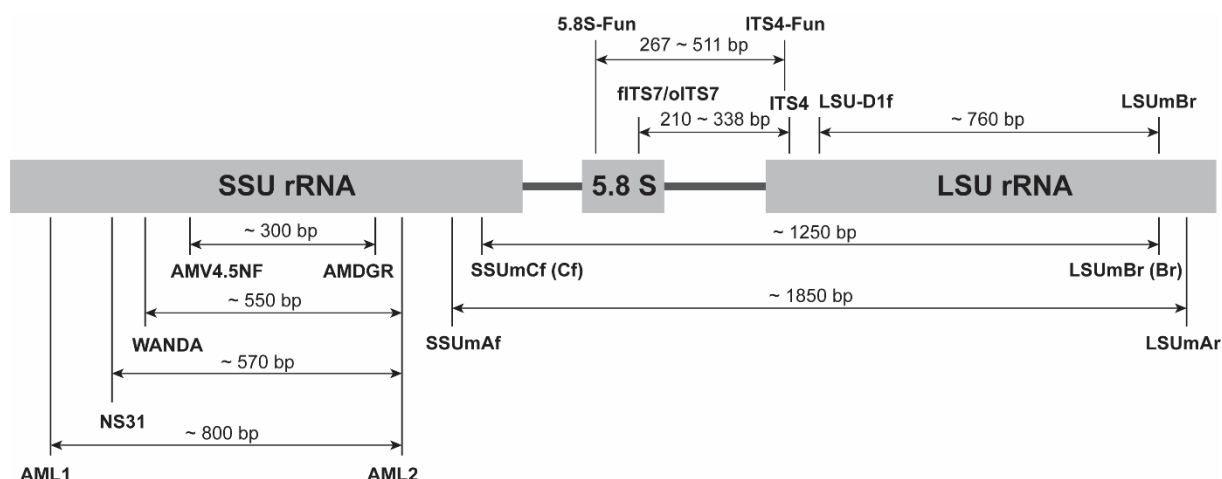


图 1. AM 真菌多样性研究可用引物

表 1 AM 真菌多样性研究可用引物列表及其优缺点

引物组合	优点	缺点	参考文献
AML1/AML2 AMV4.5F/AMDGR	可以使用 MaarjAM 数据库。	长度较短无法在低分类水平下有较好的区分度。	Sato <i>et al.</i> , 2005; Lee <i>et al.</i> , 2008; Veresoglou <i>et al.</i> , 2019
WANDA/AML2	1. MaarjAM 数据库; 2. 扩增效率较高, 覆盖度较好, 且偏好性较低。 3. 片段长度有所提高。	长度有所提高, 但是在低分类水平下的区分度仍然不高。	Dumbrell <i>et al.</i> , 2011; Egan <i>et al.</i> , 2018; Lekberg <i>et al.</i> , 2018
NS31/AML2	1. MaarjAM 数据库; 2. 传统的测序区域。	此区域较长仅能使用 454 和 PacBio, 但区分度远不及 Cf/Br。	Hiiesalu <i>et al.</i> , 2014
SSUmAf/LSUmAr SSUmCf/LSUmBr fITS7/ITS4	1. 解决了 ITS 引物对 AM 真菌扩增效率低的问题; 2. 相较于 SSU 在低分类水平下有较好的区分度。	1. 种间变异度较大; 2. 三次 PCR 扩增可能会影响 AM 真菌群落结构。	Krüger <i>et al.</i> , 2009; Deveautour <i>et al.</i> , 2018, 2020
ITS7/ITS4 或 oITS7/ITS4	1. 相较于 SSU 在低分类水平下有较好的区分度; 2. oITS7 针对 AM 真菌序列设计, 理论上覆盖度较好。	1. 种间变异度较大; 2. 引物的扩增效率可能不高, 影响下游分析。	Ihrmark <i>et al.</i> , 2012; Kohout <i>et al.</i> , 2014; Lekberg <i>et al.</i> , 2018
5.8S-Fun/ITS4-Fun	1. 相较于 SSU 在低分类水平下有较好的区分度; 2. 设计时同时考虑 AM 真菌, 理论上覆盖度较好; 3. 可以在真菌群落水平上研究 AM 真菌的多样性。	1. 种间变异度较大; 2. 引物的扩增效率可能不高, 影响下游分析。	Taylor <i>et al.</i> , 2016; Gao <i>et al.</i> , 2018

Cf/Br	片段由 SSU, ITS 和 LSU 组成，在 AM 真菌种水平上拥有较高的区分度。	只能使用 PacBio 三代测序平台，成本较高，数据分析要求较高。	Schlaeppli <i>et al.</i> , 2016
SSUmAf/LSUmAr LSU-D1f/LSUmBr	包含 LSU 的 D1 和 D2 区拥有较好的低分类水平区分度。	此区域较长仅能使用 454 和 PacBio，但区分度不及 Cf/Br。	Krüger <i>et al.</i> , 2009; Senés- Guerrero and Schüßler 2015

110

111

4. PCR 扩增 (以引物 WANDA/AML2 为例)

4.1 20 µl PCR 反应体系

5 × Phusion buffer	1 ×	4
Phusion High-Fidelity DNA Polymerase	0.02 U/µl	0.2
dNTP	200 µM	1.6
WANDA	0.4 µM	0.4
AML2	0.4 µM	0.4
BSA	1 µg/µl	1
DNA 模板	30-70 ng	1
ddH ₂ O		11.4
Total		20

注：本反应体系仅供参考，具体应根据所使用聚合酶和 DNA 模板的性质进行调整。

4.2 PCR 反应条件

温度 Temperature	时间 Time	循环 Cycle
99 °C	5 min	
99 °C	20 s	35 ×
63 °C	20 s	35 ×
72 °C	20 s	35 ×
72 °C	10 min	
4 °C	Hold	

注：本反应条件仅供参考，具体反应条件应根据聚合酶和 DNA 模板的性质进行调整，此处建议使用高保真聚合酶，PCR 产物保存于-20°C。

5. 扩增与测序

使用两步 PCR 法对目的片段进行扩增 (Egan *et al.*, 2018, Lekberg *et al.*, 2018)。首先使用 WANDA/AML2 按照 4.1 的反应体系和 4.2 的 PCR 反应条件对目的片段进行第一次扩增，获得 PCR 产物稀释 10 倍后用于第二次 PCR 扩增。第二次 PCR

扩增使用 10 次循环，添加相应的 barcodes 和 Illumina adapters 等。PCR 产物经纯化、混样 (等摩尔浓度)、建库后，使用 Illumina MiSeq PE300 进行双端测序，测序深度建议大于 10,000 条 reads。扩增与测序也可交由测序服务商进行。

结果与分析

扩增子测序数据的处理主要有聚类 (Clustering algorithm) 和去噪 (Denoising algorithm) 两种分析方法，去噪分析方法因为其具有较高的准确性和区分度，越来越受到研究者的青睐，正在逐渐的取代传统的聚类分析方法 (Callahan *et al.*, 2017)。本方法我们采用 DADA2 (Divisive amplicon denoising algorithm) 去噪分析方法来分析 AM 真菌的 SSU rRNA 扩增子测序数据。

1. DADA2 是一种运行于 R 语言中的软件包 (Callahan 等, 2016)，数据分析前请先安装 R，并将 DADA2 软件包安装于 R 中，加载到当前运行环境中，并设定当前工作路径

```
rm (list=ls () )
```

```
## 通过 Bioconductor 安装 DADA2
```

```
if (!requireNamespace ("BiocManager", quietly = TRUE) )
```

```
  install.packages ("BiocManager")
```

```
BiocManager::install ("dada2", version = "3.6.0")
```

```
library (dada2) ; packageVersion ("dada2")
```

```
## 修改工作路径到当前位置，即数据的存放目录
```

```
setwd ("~/data")
```

```
path <- "~/data"
```

```
list.files (path)
```

2. 生成文件路径列表，提取样品名称 (本例中使用三个样品 RA1, RA2 和 RA3)

```
## 生成正向和反向序列文件的路径列表，数据命名格式: samplename.R1.fq.gz
```

```
fnFs <- sort (list.files (path, pattern = ".R1.fq.gz", full.names = TRUE) )
```

```
fnRs <- sort (list.files (path, pattern = ".R2.fq.gz", full.names = TRUE) )
```

```
## 提取样品名称
```

```
sample.names <- sapply (strsplit (basename (fnRs) , "[.]") , '[', 1)
```

3. 数据质量控制，去除低质量序列


```

155 plotQualityProfile (fnRs[1:2]) ## 查看序列质量信息（图2a）
156 filtFs <- file.path (path, "filtered", paste0 (sample.names, "_F_filt.fq.gz") )
157 filtRs <- file.path (path, "filtered", paste0 (sample.names, "_R_filt.fq.gz") )
158 names (filtFs) <- sample.names
159 names (filtRs) <- sample.names
160 ## 根据数据的质量信息对数据进行过滤，因为正向和反向序列的重合区域只有 25-
161 50 bp，所以此步骤不可以直接将序列右侧低质量序列切除，只能通过收紧质量过滤
162 参数过滤低质量序列。在此代码中我们用 trimleft 切除了引物序列，并收紧了质量过
163 滤参数，不允许序列中有 N 出现，正向和反向序列中仅允许有一个错误，最小序列
164 长度为 200 bp，此处质量参数设置较为严格，样品 RA1 的 40,471 条序列中仅有
165 16,742 条通过（图 2. a, b），如序列质量较低导致序列去除率过高，可适当的放松
166 质量参数，增加序列的通过率。若数据本身为非压缩数据，compress 设置为 FALSE;
167 Windows 操作系统下不支持多核计算，multithread 设置为 FALSE。
168 out <- filterAndTrim (fnFs, filtFs, fnRs, filtRs, trimleft = c (20,22) , maxN = 0, maxEE
169 = c (1, 1) , minLen = 200, truncQ = 2, rm.phix = TRUE, compress = TRUE,
170 multithread = FALSE)
171 head (out)
172 ## 查看质量控制后的序列质量信息，如序列质量没有得到改善，可以在上一步骤中
173 收紧质量过滤参数，提高质控后序列质量。在本例中样品 RA1 的序列在质量控制后，
174 序列质量值的平均数显著升高，说明序列质量得到了明显的改善（图 2）。
175 plotQualityProfile (c (filtFs[1], filtRs[1]) )
176

```

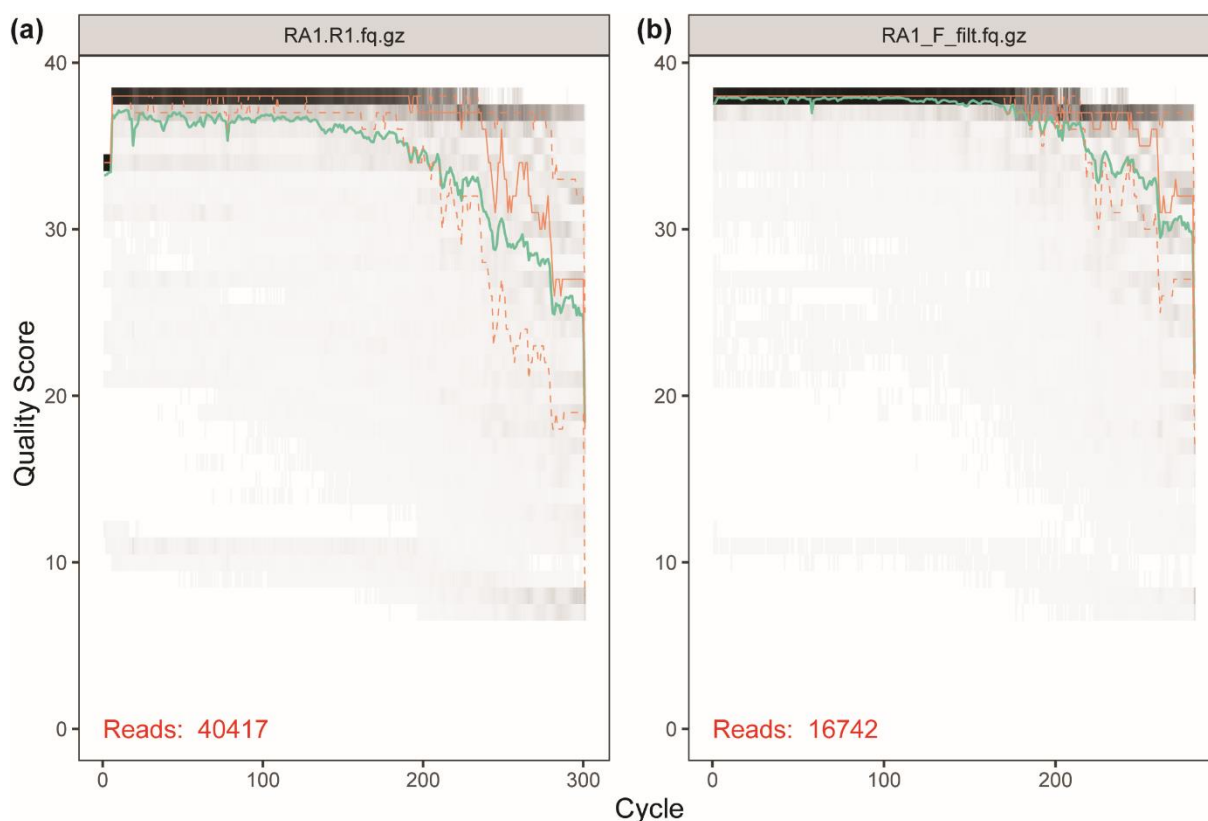


图 2. 序列质量控制. (a) 质量控制前序列质量信息; (b) 质量控制后序列质量信息
本图为序列每个位置质量分布的灰度热图，绿线表示质量值的平均数，橙色实线为质量值的中位数，橙色虚线分别表示上、下四分位数

4. DADA2 去噪算法的核心，使用机器学习算法对序列进行去噪

使用机器学习算法学习测序错误发生的概率

```
errF <- learnErrors (filtFs, multithread = FALSE)
```

```
errR <- learnErrors (filtRs, multithread = FALSE)
```

```
plotErrors (errR, nominalQ = TRUE)
```

DADA2 去噪算法的核心，样品推断，根据错误发生率对序列进行去噪，去重复，列出每一个样品中的独特序列 (Unique sequences) (图 3)

```
dadaFs <- dada (filtFs, err = errF, multithread = FALSE)
```

```
dadaRs <- dada (filtRs, err = errR, multithread = FALSE)
```

```
dadaRs[[1]]
```

```
> dadaFs <- dada(filtFs, err = errF, multithread = FALSE)
Sample 1 - 16742 reads in 10545 unique sequences.
Sample 2 - 14492 reads in 9493 unique sequences.
Sample 3 - 12163 reads in 7911 unique sequences.
> dadaRs <- dada(filtRs, err = errR, multithread = FALSE)
Sample 1 - 16742 reads in 5338 unique sequences.
Sample 2 - 14492 reads in 4282 unique sequences.
Sample 3 - 12163 reads in 3766 unique sequences.
> dadaRs[[1]]
dada-class: object describing DADA2 denoising results
104 sequence variants were inferred from 5338 input unique sequences.
Key parameters: OMEGA_A = 1e-40, OMEGA_C = 1e-40, BAND_SIZE = 16
```

图 3. 样品推断并列独特序列

5. 合并正反向序列，构建独特序列表格

```
## 正向序列和反向序列的合并，此处可以设置最小合并区域，默认为 10 bp
mergers <- mergePairs (dadaFs, filtFs, dadaRs, filtRs, verbose = TRUE,
minOverlap = 20)
head (mergers[[1]])
## 构建序列表格
seqtab <- makeSequenceTable (mergers)
## 查看序列表格的基本信息
dim (seqtab)
table (colSums (seqtab > 0) )
table (rowSums (seqtab > 0) )
```

6. 去除序列中的嵌合体

```
seqtab.nochim <- removeBimeraDenovo (seqtab, method = "consensus",
multithread = FALSE, verbose = TRUE)
dim (seqtab.nochim)
table (colSums (seqtab.nochim > 0) )
table (rowSums (seqtab.nochim > 0) )
sum (seqtab.nochim) /sum (seqtab) ## 查看嵌合体序列在总序列数中的比例
table (nchar (getSequences (seqtab.nochim) ) ) ## 查看序列长度分布
```

7. 追踪各个步骤中的序列数 (图 4)

```
getN <- function (x) sum (getUniques (x) )
```

```
track <- cbind (out, sapply (dadaFs, getN) , sapply (dadaRs, getN) , sapply
(mergers, getN) ,rowSums (seqtab.nochim) )
colnames (track) <- c ("input", "filtered", "denoisedF", "denoisedR", "merged",
"nochim")
rownames (track) <- sample.names
head (track)
```

```
> head(track)
```

	input	filtered	denoisedF	denoisedR	merged	nochim
RA1	40417	16742	15104	16514	8547	5332
RA2	33306	14492	13092	14262	7521	5075
RA3	28807	12163	11034	11988	6184	4199

图 4. 追踪各个步骤中的序列数

8. 进一步通过序列长度过滤非目标序列 (可选)

根据计算获得的目标序列长度过滤序列表格, 此步骤与引物、*barcode* 序列等密切相关

```
seqtab.nochim2 <- seqtab.nochim[ ,nchar (colnames (seqtab.nochim) ) %in%
495:510]
dim (seqtab.nochim2)
```

9. 输出 ASV (Amplicon sequence variants) 表格用于后续分析

```
write.csv (t (seqtab.nochim2), file = "seqtab.nochim.filtered.csv")
```

10. 使用 ASV 表格中的序列与 MaarjAM 数据库进行比对, 序列一致性 (Sequence identity) 大于 97%, 覆盖度 (Sequence coverage) 大于 95% 的 ASV 即可认为与数据库中的虚拟分类单元 (Virtual taxa, VT) 相对应, 继而用于下游分析 (Lekberg *et al.*, 2018) 。如使用 DADA2 中的朴素贝叶斯分类算法 (Naive Bayesian classifier) , 可使用 Sliva 18S 数据库, 但是本数据库并不是 AM 真菌的专一数据库, 比对效果不及 MaarjAM 数据库, 因此须自行根据 MaarjAM 建库, 建库格式请参考 DADA2 官网。如需要在不同的相似水平上对 ASVs 进行聚类, 可以使用 VSEARCH 软件进行 (Rognes *et al.*, 2016)

11. 下游的 AM 真菌多样性分析可以从 ASV 水平和 VT 水平同时进行，分析方法与其他微生物类似，例如使用 **Vegan** 软件包绘制 AM 真菌 ASV 水平的稀释曲线并计算其 α 多样性

```
library (vegan) ## 使用前须安装
```

```
asv.table <- seqtab.nochim ## 此处以嵌合体过滤后的 ASV 表格为例进行计算
```

```
head (asv.table)
```

```
## 以所有样品中的最小序列数抽平 ASV 表格，以稀释后的 ASV 表格做稀释曲线  
如图 5
```

```
r.asv.table <- rrarefy (asv.table, 4199) ## 本例中样本最小序列数为 4199 (图 4)
```

```
rarecurve (r.asv.table, step = 1, xlab = "Sample size", ylab = "ASV numbers",  
label = TRUE, col = "red")
```

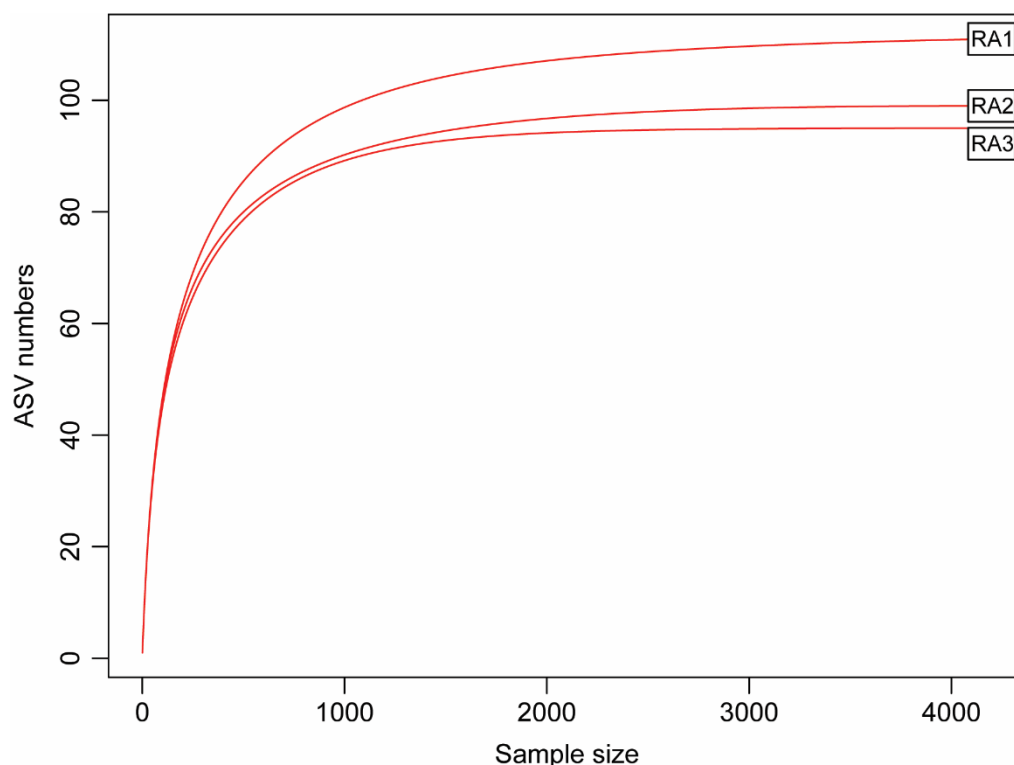


图 5. 样本稀释曲线

```
## 计算各样本的  $\alpha$  多样性指数
```

```
richness <- specnumber (r.asv.table) ## Species richness
```

```
shannon <- diversity (r.asv.table) ## Shannon entropy (base e)
```

```
simpson <- diversity (r.asv.table, "simpson") ## Simpson diversity
```

```
inverse.simpson <- diversity (r.asv.table, "invsimpson")    ## Inverse Simpson
pielou.evenness <- shannon/log (richness)    ## Pielou evenness
shannon.evenness <- shannon/richness    ## Shannon evenness
simpson.evenness <- simpson/richness    ## Simpson evenness
## 合并各样本的  $\alpha$  多样性指数，输出结果如图 6 所示
asv.alpha.div <- data.frame (richness, shannon, simpson, inverse.simpson,
pielou.evenness,
shannon.evenness, simpson.evenness)
head (asv.alpha.div)

> head(asv.alpha.div)
      richness shannon  simpson  inverse.simpson  pielou.evenness  shannon.evenness  simpson.evenness
RA1    111  4.067838  0.9753012   40.48783      0.8637460      0.03664719      0.008786497
RA2     99  4.035747  0.9753187   40.51649      0.8782681      0.04076512      0.009851704
RA3     95  3.992145  0.9739278   38.35500      0.8766476      0.04202258      0.010251871
```

图 6. AM 真菌 α 多样性计算结果

参考文献

1. Bruns, T. D., and Taylor, J. W. (2016) . [Comment on “Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism”](#). *Science* 351 (6257) : 826-826.
2. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. and Holmes, S.P. (2016) . [DADA2: high-resolution sample inference from Illumina amplicon data](#). *Nat methods* 13 (7) : 581-583.
3. Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017) . [Exact sequence variants should replace operational taxonomic units in marker-gene data analysis](#). *ISME J* 11 (12) : 2639-2643.
4. Deveautour, C., Donn, S., Power, S. A., Bennett, A. E., and Powell, J. R.(2018) . [Experimentally altered rainfall regimes and host root traits affect grassland arbuscular mycorrhizal fungal communities](#). *Mol Ecol* 27 (8) : 2152-2163.
5. Deveautour, C., Power, S. A., Barnett, K. L., Ochoa-Hueso, R., Donn, S., Bennett, A. E., and Powell, J. R. (2020) . [Temporal dynamics of mycorrhizal fungal](#)

[communities and co-associations with grassland plant communities following experimental manipulation of rainfall.](#) *J Ecol* 108 (2) : 515-527.

6. Dumbrell, A. J., Ashton, P. D., Aziz, N., Feng, G., Nelson, M., Dytham, C., Fitter, A. H., and Helgason, T. (2011) . [Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing.](#) *New Phytol* 190 (3) : 794-804.

7. Egan, C. P., Rummel, A., Kokkoris, V., Klironomos, J., Lekberg, Y., and Hart, M. (2018) . [Using mock communities of arbuscular mycorrhizal fungi to evaluate fidelity associated with Illumina sequencing.](#) *Fungal Ecol* 33: 52-64.

8. Gao, C., Montoya, L., Xu L., Madera, M., Hollingsworth, J., Purdom, E., Hutmacher, R. B., Dahlberg, J. A., Coleman-Derr, D., Lemaux, P. G., and Taylor, J. W. (2018) . [Strong succession in arbuscular mycorrhizal fungal communities.](#) *ISME J* 13 (1) : 214-226.

9. Hiiesalu, I., Pärtel, M., Davison, J., Gerhold, P., Metsis, M., Moora, M., Öpik, M., Vasar, M., Zobel, M., and Wilson, S. D. (2014) . [Species richness of arbuscular mycorrhizal fungi: associations with grassland plant richness and biomass.](#) *New Phytol* 203 (1) : 233-244.

10. Ihrmark, K., Bodeker, I.T., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandstrom-Durling, M., Clemmensen, K. E. and Lindahl, B. D. (2012) . [New primers to amplify the fungal ITS2 region evaluation by 454-sequencing of artificial and natural communities.](#) *FEMS Microbiol Ecol* 82 (3) : 666-677.

11. Kohout, P., Sudová, R., Janoušková, M., Čtvrtlíková, M., Hejda, M., Pánková, H., Slavíková, R., Štajerová, K., Vosátka, M., and Sýkorová, Z. (2014) . [Comparison of commonly used primer sets for evaluating arbuscular mycorrhizal fungal communities: is there a universal solution?](#) *Soil Biol Biochem* 68: 482-493.

12. Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F., Bahram, M., Bates, S. T., Bruns, T. D., Bengtsson, - Palme J., and Callaghan, T. M. (2013) .

- [Towards a unified paradigm for sequence based identification of fungi.](#) *Mol Ecol* 22 (21) : 5271-5277.
13. Krüger, M., Stockinger, H., Krüger, C., and Schussler, A. (2009) . [DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi.](#) *New Phytol* 183 (1) : 212-223.
14. Lee, J., Lee, S., and Young, J. P. (2008) . [Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi.](#) *FEMS Microbiol Ecol* 65 (2) : 339-349.
15. Lekberg, Y., Vasar, M., Bullington, L. S., Sepp, S. K., Antunes, P. M., Bunn, R., Larkin, B. G., and Öpik, M. (2018) . [More bang for the buck? Can arbuscular mycorrhizal fungal communities be characterized adequately alongside other fungi using general fungal primers?](#) *New Phytol* 220 (4) : 971-976.
16. Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J. M., Reier, U., and Zobel, M. (2010) . [The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi \(Glomeromycota\).](#) *New Phytol* 188 (1) : 223-241.
17. Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016) . [VSEARCH: a versatile open source tool for metagenomics.](#) *PeerJ*, 4:e2584.
18. Sato, K., Suyama, Y., Saito, M., and Sugawara, K. (2005) . [A new primer for discrimination of arbuscular mycorrhizal fungi with polymerase chain reaction-denature gradient gel electrophoresis.](#) *Grassl Sci* 51 (2) : 179-181.
19. Schlaeppi, K., Bender, S. F., Mascher, F., Russo, G., Patrignani, A., Camenzind, T., Hempel, S., Rillig, M. C., and van der Heijden, M. G. 2016. [High-resolution community profiling of arbuscular mycorrhizal fungi.](#) *New Phytol* 212 (3) : 780-791.
20. Senés-Guerrero, C., and Schüßler, A. (2015) . [A conserved arbuscular mycorrhizal fungal core-species community colonizes potato roots in the Andes.](#) *Fungal Divers* 77 (1) : 317-333.
21. Taylor, D. L., Walters, W. A., Lennon, N. J., Bochicchio, J., Krohn, A., Caporaso, J. G., and Pennanen, T. (2016) . [Accurate estimation of fungal diversity and](#)

[abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing.](#) *Appl Environ Microbiol* 82 (24) : 7217-7226.

22. Veresoglou, S. D., Liu, L., Xu, T. L., Rillig, M. C., Wang, M. E., Wang, J. T., Chen, Y. L., Hu, Y. J., Hao, Z. P., and Chen, B. D. (2019) . [Biogeographical constraints in Glomeromycotinan distribution across forest habitats in China.](#) *J Ecol* 107 (2) :684-695.

致谢

感谢 DADA2 的开发者 Callahan 博士在 GitHub 上的解疑答惑，以及由他开发的数据分析软件。感谢国家自然科学基金面上项目 (41877050) 和国家重点研发计划项目 (2016YFC0500702) 的资助。