**Extraction and analysis of microbiomes associated with rice roots**

Joseph Edwards1,a,b, Christian Santos-Medellín1,a, and Venkatesan Sundaresan1,2,\*

1 Department of Plant Biology, University of California, Davis

2 Department of Plant Sciences, University of California, Davis

a These authors contributed equally

b Current address: Department of Integrative Biology, University of Texas, Austin

\* To whom correspondence should be addressed: sundar@ucdavis.edu

Abstract

Plant roots associate with a wide diversity of prokaryotic microbiota comprised of bacteria and archaea. Rhizospheric microbiota, the communities of microbes in the soil adjacent to the root, can contain over x million number of cells per gram of soil and can play important roles for the fitness of the host plant. Subsets of the rhizospheric microbiota can colonize the root surface (rhizoplane) and the root interior (endosphere), forming an intimate relationship with the host plant. Compositional analysis of these communities is important to develop tools in order to manipulate root-associated microbiota for increased crop productivity. Due to the reduced cost and increasing throughput of next-generation sequencing, major advances in deciphering these communities have recently been achieved, mainly through the use of amplicon sequencing of the 16S rRNA gene. Here we first present a protocol for dissecting the microbiota from various root compartments, developed using rice as a model. We next present a method for amplifying fragments of the 16S rRNA gene using a dual index approach. Finally, we present a simple workflow for analyzing the resulting sequencing data to make ecological inferences.

Keywords

Root microbiome, amplicon sequencing, rhizosphere

Background

Various plant root niches host different microbial communities (microbiota) originating from the soil [1–5]. Distinct microbiota acquired by each root niche likely have varying metabolic potential and may therefore impact the health of the host plant in different ways [6]. Prokaryotic community composition in root-associated microbiota can be inferred through the use of 16S rRNA gene sequencing [7]. The relatively low cost of sequencing now allows for comparative studies across plant species using datasets gathered by different research groups; however, small aberrations in specimen collection, sequencing, and analysis protocols may lead to large differences in the inferred microbial communities [8]. We present this protocol detailing how to collect and analyze plant root microbiota in an attempt to promote reproducibility across the plant microbiome field.

Materials and Reagents

1. Falcon 50 mL conical centrifuge tubes (Corning , catalog number: 352070)

2. Autoclaved phosphate buffered saline (PBS) solution (~100 mL / plant)

3.Filtered pipette tips (10, 200, 1000 µL)

4.DNeasy PowerSoil kit (Qiagen, catalog number 1288-100)

5. HotStar High Fidelity DNA polymerase kit (Qiagen, catalog number: 202602)

6. Agencourt Ampure XP beads (Beckman Coulter, catalog number: A63880)

7. Qubit dsDNA HS assay kit (ThermoFisher Scientific, catalog number: Q32851)

8. Qubit assay tubes (ThermoFisher Scientific, catalog number: Q32856)

9. NucleoSpin gel and PCR clean-up kit (Macherey-Nagel, catalog number: 740609.250)

10. Ethanol 200 proof

11. Agarose

12. Gloves

13. 0.2 mL PCR tubes

13. DNA gel loading dye

14. Nuclease-free water

15. 1.5 mL microfuge tubes

16. 1.5 mL non-stick microfuge tubes

Equipment

1. Pipettes (2.5, 10, 200, 1000 µL)

2. Ultrasonic cleaning bath, 40 kHz (Branson 1800, CPX-952-116R)

3. Dissection tools (scissors and forceps)

4. Microcentrifuge

5. Mini-Beadbeater-96 high-throughput cell disrupter (Biospec, catalog number: 1001)

6. Electrophoresis gel unit

7. Qubit fluorometer (catalog number: Q33226)

9. Blade

10. 96 well magnetic plate

11. 1.5 mL tube magnetic rack

Procedure

1. Compartment separation of root-associated microbiota

The following protocol uses a combination of washing and sonicating steps to separate the rhizosphere, rhizoplane, and endosphere fractions of the root-associated microbiota. This approach has been successfully employed to harvest compositionally distinct communities that harbor microorganisms enriched in each of these spatial compartments [3]. The protocol for separation of the rhizoplane is based on the method for endospheric bacteria isolation developed by Lundberg et al. [1]. The method utilizes a bath sonicator to remove the microbiota in the rhizoplane, and avoids hypochlorite treatment for the reasons detailed in Lundberg et al. (2012). Because the DNA yields from rhizoplane samples are low, the sequences can exhibit higher variability following PCR amplification, and may require additional replicates to draw statistically significant conclusions. It should be noted that the compartment dissection protocol cannot ensure complete purity of samples free of contamination from adjacent compartments, especially where they overlap spatially, but as noted above, it has been proven to be efficient and reproducible for studies of overall compositional profiles [3,9,10].



**Figure 1. Root harvesting and processing for microbiome studies.** (A) Rice roots pulled out of the soil (B) Rice roots after being vigorously shaken to remove loose soil. The red box indicates the ~5 cm of root cut with flame-sterilized scissors. (C) Rice root section collected into 50 mL Falcon tubes for compartment separation. The soil layer firmly attached to the roots constitutes the rhizosphere. (D) Rice root section after being thoroughly washed with sterile PBS solution.

1. Using gloves, harvest the rice plant by firmly holding the shoot and slowly pulling the root system out of the ground (Fig 1A). In the case of seedlings, carefully scoop the roots to avoid ripping the tissue.
2. Vigorously shake the roots to remove loose soil, leaving only the soil layer firmly attached to the root. This layer constitutes the rhizosphere compartment (Fig 1B).
3. Using flame-sterilized scissors, cut ~5 cm of root immediately below the root-shoot junction (red box Fig 1B, Fig 1C) and place the tissue in a sterile 50 mL Falcon tube with 15 mL of autoclaved PBS solution. For potted plants, avoid collecting roots immediately adjacent to the inner walls.
4. Separate the rhizosphere fraction by vortexing the roots until the attached soil is suspended in the PBS solution (Sup Fig 1).
5. Using flame-sterilized forceps, transfer the roots to a new 50 mL falcon tube and wash thoroughly with PBS solution to remove any soil still adhered to the roots (Fig 1D). Add 10 mL of fresh PBS solution to the clean roots. The roots should be completely free of soil before proceeding to Step 1f (Fig 1D).
6. Separate the rhizoplane compartment by sonicating the roots for 30 seconds at 50-60 Hz and transferring the 10 mL of PBS with the sonicated microbes (Sup Fig 1).
7. Add enough fresh PBS to fully cover the roots and sonicate for 30 seconds at 50-60 Hz. Discard the PBS and repeat this step once more. The thrice-sonicated roots constitute the endosphere compartment (Sup Fig 1).

2. DNA Extraction

Use the DNeasy PowerSoil kit to isolate the genomic DNA from the root-associated communities. The input for each of the compartments is as follows:

1. For the rhizosphere compartment, add 500 µL of the soil suspension generated in Step 1d to a PowerBead tube.
2. For the rhizoplane compartment, add 1.5 mL of the microbial suspension generated in Step 1f to a microfuge tube. Spin down at 10,000 x g for 1 minute and discard 1 mL of supernatant. Add 1 mL of microbial suspension, spin down (10,000 x g for 1 minute), and discard 1 mL. Repeat these steps once more for a total of three centrifugations. Resuspend the pellet, and transfer the 500 µL of concentrated suspension to a PowerBead tube.
3. For the endosphere compartment, use fire-sterilized forceps to transfer 0.25 g of the thrice-sonicated roots from Step 1g to a PowerBead tube. Pre-homogenize the root tissue by beadbeating the sample for 1 minute.

After adding the samples to the PowerBead tubes, follow the PowerSoil kit protocol with these adjustments:

1. After adding Solution C1, the PowerBead tubes can be homogenized using a beadbeater for 2 minutes instead of vortexing them for 10 minutes.
2. Elute the final product in 30 µL of Solution C6 instead of 100 µL.

3. 16S rRNA amplification

For library construction, this protocol uses the 515F/806R universal primer set to amplify the V4 region of the 16S rRNA gene. Primer design for Illumina sequencing follows the one described in Caporaso et al, 2012 [7], except both forward and reverse primers are barcoded (Fig 2). By using a unique combination of barcodes for each sample, this dual-indexing strategy allows us to multiplex a large number of libraries with a limited amount of primers. Full primer sequences are provided in Supplementary Table 1. We recommend working with sets of 24 samples, in which all reactions share the same barcoded 515F primer but have a unique 806R barcode. Additionally, it is important to run a negative control for each individual reaction to detect any potential contamination.



**Figure 2. Schematic of 16S V4 region amplicon**. (A) Genomic regions before amplification. The primer binding sites are blue and the number corresponds to the position within the 16S rRNA gene where the primers bind. (B) Amplicon after PCR amplification. FBC stands for forward barcode and RBC stands for reverse barcode. (C) The sequencing strategy for the amplicons. Note that three custom primers are used in the sequencing: a primer for the forward read starting at position 515, a primer for the reverse sequencing read starting at position 806, and a primer for the reverse barcode.

1. For a set of 24 reactions and 24 negative controls, prepare a master mix using the following recipe:

|  |  |  |
| --- | --- | --- |
|  | 1 reaction | 50 reactions |
| RNase-Free Water | 6.25 µL | 312.5 µL |
| HotStar HiFidelity PCR Buffer 5X | 2.5 µL | 125 µL |
| 10 µM barcoded 515F Primer | 1.25 µL | 62.5 µL |
| HotStar HiFidelity DNA Polymerase | 0.25 µL | 12.5 µL |

1. Aliquot 20.5 µL of the master mix into each PCR tube
2. Add 2.5 µL of the corresponding 10 µM barcoded 806R primer into each tube and mix well by pipetting.
3. Aliquot 11.5 µL of each reaction into new PCR tubes to run as negative controls.
4. Add 1 µL of the corresponding template to the remaining 11.5 µL.
5. Cap the tubes and spin down.
6. Run the following touchdown PCR program:

Initial denaturation:

* 95ºC, 5 minutes

7 cycles, decreasing the annealing temperature 2ºC each cycle:

* 95ºC, 45 seconds
* 65ºC, 1 minute (- 2ºC / cycle)
* 72ºC, 1:30

30 cycles:

* 95ºC, 45 seconds
* 50ºC, 30 seconds
* 72ºC, 1:30

Final extension:

* 72ºC, 10 minutes

Pause:

* 4ºC, ∞

4. Gel

1. Add 1 µL of PCR product to 5 µL of gel loading dye (1X)
2. Run the samples on a 1% agarose gel at 120 V for 20 minutes.
3. Verify proper amplification (expected band size is ~400 bp long) and absence of contamination in the negative controls.

5. PCR Cleanup

1. Remove beads from 4ºC and allow them to reach room temperature.
2. Prepare a fresh batch of 70% ethanol solution (500 µL / reaction).
3. Aliquot 9 µL of the PCR product into a tube.
4. Add 5.4 µL (0.6 volume) of Ampure XP beads, mix by pipetting, and let incubate at room temperature for 5 minutes.
5. Transfer the tubes to a magnet plate and let stand for 2 minutes.
6. Carefully remove the cleared solution without disturbing the beads.
7. Keeping the tubes on the magnet plate, add 200 µL of 70% ethanol, incubate for 30 seconds, and remove with a pipette. Repeat this step once more for a total of two ethanol washes. For the final wash, remove all ethanol from the bottom.
8. With the tubes still on the magnet plate, air-dry the beads for 2 minutes. Take care not to overdry the beads as this will prevent the beads from being resuspended in Step 5i.
9. Take the plates off the magnet plate, add 50 µL of nuclease-free water, and resuspend the beads by pipetting up and down.
10. Place the tube back to the magnet plate and incubate for 1 minute.
11. Transfer 35 µL of the eluant to a new tube.

6. Quantification

Use the Qubit dsDNA HS assay to measure the concentrations of the purified libraries:

1. Set up two Assay Tubes for the standards and one for each sample to be quantified.
2. For each tube, prepare 200 µL of Qubit Working Solution by mixing 199 µL of Qubit buffer and 1 µL of Qubit Working Solution
3. For the standards, aliquot 190 µL of working solution to 0.5 Assay Tubes, add 10 µL of the corresponding standard, and mix by vortexing.
4. For the samples, aliquot 195 µL of working solution to 0.5 Assay Tubes, add 5 µL of the corresponding sample, and mix by vortexing.
5. Incubate tubes for 2 minutes at room temperature.
6. Select the dsDNA High Sensitivity Assay on the Qubit Fluorometer, read the standards, and run each sample.

7. Pooling

1. Multiplex ~150 libraries per MiSeq sequencing run. If multiple sequencing runs are needed, randomize the libraries across sequencing runs to avoid batch effects.
2. Based on the range of concentrations obtained, determine a target amount of DNA to be pooled. Aim for at least 5 ng of DNA per sample.
3. For each library, calculate the volume need by dividing the target amount of DNA by the sample concentration.
4. Pool the libraries in a non-stick RNase-free 1.5 mL microfuge tube.
5. Avoid pipetting volumes lower than 1 µL. If some samples are too concentrated, predilute them before pooling.

8. Concentration

1. Remove AmpureXP beads from 4ºC and allow them to reach room temperature.
2. Prepare a fresh batch of 70% ethanol solution.
3. Add 1.8 volumes of Ampure XP beads to the pooled libraries, mix by pipetting, and let incubate at room temperature for 5 minutes.
4. Transfer the tube to a magnet and let stand for 2 minutes.
5. Carefully remove the cleared solution without disturbing the beads.
6. Keeping the tubes on the magnet, add 1.5 mL of 70% ethanol, incubate for 30 seconds, and remove with a pipette. Repeat this step once more for a total of two ethanol washes. For the final wash, remove all ethanol from the bottom.
7. With the tube still on the magnet, air-dry the beads for 2 minutes
8. Take the tube off the magnet plate, add 50 µL of nuclease-free water, and resuspend the beads by pipetting.
9. Place the tube back to the magnet and incubate for 1 minute.
10. Transfer the cleared eluant to a new tube.

8. Gel cleanup

1. Run the concentrated pool of libraries on a 1.8% agarose gel at 120 V for 40 minutes.
2. Using a new blade, excise the ~400 bp band.
3. Purify the libraries using the NucleoSpin Gel and PCR Clean-up kit.

9. Sequencing

1. Submit the pooled libraries for 2 x 250 MiSeq sequencing. Use the custom sequencing primers XXX (Table X).
2. Include a PhiX control for low diversity samples.

Software

1. Python2

2. R

Sequence Analysis

A more detailed version of the sequence analysis pipeline including code for running the analysis can be found on [GitHub](https://github.com/bulksoil/BioProtocol/blob/master/ANALYSIS/SeqProcessing/seq_processing.md).

1. Compile metadata in a spreadsheet

1. Assign each sample a unique identifier. This identifier should be unique not just across a single experiment, but across all previous experiments as well. The identifier should only contain alphanumeric and period (“.”) characters.
2. Create column(s) for barcodes. Each sample should have a unique barcode for each run.
3. Create additional columns for the experimental variables associated with each sample. Some typical examples of experimental variables are root compartment, plant developmental stage, plant genotype, plot location, date, year, and collector.

Note: *A detailed protocol for designing a metadata table can be found at the* [*Earth Microbiome Project's website*](http://www.earthmicrobiome.org/protocols-and-standards/metadata-guide/)*.*

2. Download sequencing files from sequencing facility

1. Downloading of individual files can be accomplished either through an FTP client or by using command line tools (such as wget).

3. Demultiplex sequeces

1. There should be 4 files provided by the sequencing facility.

4. Construct full length contiguous sequences.

1. Full length sequences can be assembled using PANDAseq [11]. Note that full length contigs are not necessary for clustering OTUs using DADA2 (see below).

5. Cluster Sequences and build OTU table

1. If clustering with QIIME [12], several options and algorithms are available. The user may want to conduct reference based clustering against a database of 16S rRNA genes (closed reference clustering) or a user may want to perform *de novo* OTU clustering. This method is known as closed reference clustering. Alternatively a user may prefer to perform a hybrid between these two methods where sequences are first referenced against a database. Reads that do not have a match within the database are then clustered *de novo*. This method is known as open reference clustering. The user can define the similarity threshold for one read to be considered a match with an entry in the database. Historically, > 97% sequence identity has been used as the standard for clustering sequences into operational taxonomic units (OTUs). One relatively new and extremely fast method for closed reference clustering of sequences into OTUs is through using the NINJA-OPS pipeline [13]. NINJA-OPS leverages the speed and memory efficiency of Bowtie [14], mapping reads back to a synthetic genome of concatenated 16S genes. This method can be performed on a laptop computer.
2. Alternatively, users may prefer to bin sequences based upon exact matches using DADA2 [15]. If using this method, there is no need to construct full length contiguous sequences before clustering.

6. Assign taxonomies to OTUs

1. Multiple algorithms exist to assign taxonomies to the OTU sequences. If using closed reference OTU clustering, there is no need to perform this step because the database sequences have already been classified [16]. QIIME defaults to using the uclust [17] algorithm for taxonomic assignment, while DADA2 uses the RDP naive bayes method [18] for assigning taxonomies to sequences.

Data analysis

A detailed description of how to perform the data analysis including R code can be found on [GitHub](https://github.com/bulksoil/BioProtocol/blob/master/ANALYSIS/Stats/MicrobiomeAnalysis.rmd). In this tutorial we use data from Santos-Medellín et. al, 2017 [9] to illustrate analytical techniques.



**Figure 3. Example analysis of amplicon microbiome data.** (A) Principal coordinates analysis showing microbial community structure between root compartments using Bray dissimilarities. Each point represents the microbial community in one particular sample. (B) Alpha diversity within each community using two commonly used metrics. Richness measures how many unique OTUs were detected in each sample while Shannon-entropy measures the randomness or uncertainty in a community. (C) The distribution of the 10 most abundant phyla in the dataset. We show the similar compartments between sites host similar distributions of microbes when analyzing at the phylum level. (D) Differentially abundant microbes in the rhizosphere and endosphere compartments compared to bulk soil. Each point represents a single microbial OTU. The colored points represent OTUs that were significantly differentially abundant in one of the comparisons. The color of the point represents the direction of enrichment. Differential abundance analyses were carried out using DESeq2.

1. Remove plastidial and mitochondrial sequences from the dataset.

1. Mitochondria and plastids are a result of an ancient endosymbiosis event. The mitochondria and plastids have retained their own ribosomal machinery, therefore a fraction of the resulting sequences will belong to these organelles. These reads are not part of the root microbiota and should be removed from the dataset before further analysis. This is not to say that organellar reads are not useful - these reads can be used for quantification purposes [3], but they should not be considered part of the microbiota.
2. Mitochondrial and plastidial OTUs can be identified via their associated taxonomies by searching for ‘mitochondria’ in the Family column and ‘Chloroplast’ under the Class column.

2. Normalize the sequencing depth for each library.

Although the libraries were pooled in a equimolar concentration, sequencing depth can vary a few orders of magnitude between each library. It is therefore necessary to normalize data to ensure the each sample is equally represented in the analysis.

1. A few methods exist for normalization. Rarefaction is the process of randomly sampling from the pool of OTUs until a desired depth is achieved. It is of note that this method removes much of the data the user has acquired. For example, if the user has two libraries *a* and *b* and the depth, *d*, of each library is *da* = 100,000 and *db* = 5000, the user may choose to rarefy to 5000 sequences. This does not discard any sequences from *db*, but it removes 95% of the data from *da*. OTUs with low representation may be discarded using this method. Relative abundance is a method which divides the count of each OTU by the sequencing depth such that the user is left with proportional representation of each OTU in each library. This method makes full use of all the data the user has acquired. Depending on the particular analysis, the user may prefer to use alternative methods implemented in high throughput sequencing statistical libraries such as edgeR [19] or DESeq [20].

3. Ensure that the order of samples in the metadata file (also known as a mapping file) matches the order of samples in the OTU count table.

4. Remove low prevalence OTUs from the data.

1. Low prevalence or non-reproducible OTUs may add unnecessary noise to the dataset. There is no specific rule of thumb for removing low abundance OTUs, but one metric that has been previously used is to remove OTUs that are not present in at least 5% of the samples [10,21].

5. Beta diversity plots (Fig. 3A)

Beta diversity measures the differences in microbiota composition between the samples.

1. Calculate pairwise dissimilarities between each sample. There are ecologically appropriate metrics for this task such as Bray-Curtis, Jaccard, and UniFrac [22] dissimilarity metrics.
2. Using the calculated dissimilarities, perform principal coordinate analysis (PCoA).
3. Plot the resulting axes and color the points based upon the factor of interest.

6. Alpha diversity plots (Fig. 3B)

Alpha diversity measures the diversity within each sample.

1. Calculate alpha diversity metric for each sample. Popular metrics are the Shannon index, the Simpson Index, species richness, and Faith’s phylogenetic distance.
2. Plot resulting calculations, comparing the factors of interest.

7. Phylum level analysis (Fig. 3C)

1. Summarize the mean representation of each phylum in each sample type.
2. Plot the results. There are many ways to plot these results. Here we have chosen to display the data using a stacked bar plot. We have also only retained the 10 most highly represented phyla.

8. OTU differential abundance (Fig. 3D).

Note: The statistical distributions of bacterial and archaeal OTU abundances do not follow a gaussian distribution and typically cannot be log-transformed to fit a normal distribution. Therefore, methods assuming normal distributions are not recommended for performing differential abundance tests. Available statistical packages (such as edgeR or DESeq) are recommended in order to properly model OTU distributions [23].

1. Load non-normalized count data into the statistical package of choice.
2. Normalize for sequencing depth.
3. Model sample level and OTU level dispersions.
4. Fit the model using a design matrix.
5. Perform differential abundance tests.
6. Plot results with the average abundance on the x-axis and the fold change between the sample types on the y-axis.

Recipes

Phosphate buffered saline solution (1 L)

* 8 g NaCl
* 0.2 g KCl
* 1.44 g Na2HPO4
* 0.24 g KH2PO4

pH 7.4

**Acknowledgements**

We thank Kelsey Galimba, Cassandra Ramos, Paul Tisher, John Jaeger, Eugene Lurie, Bao Nguyen, and Zach Liechty for development of this protocol as well as helpful discussions to streamline the processes. We thank Derek Lundberg and Jeff Dangl (University of North Carolina, Chapel Hill) for kindly providing their sonication protocol for separating the rhizoplane from endosphere fractions. We thank Srijak Bhatnagar and Jonathan Eisen for their help in establishing the analysis protocol. V.S. acknowledge the support of National Science Foundation Awards DBI- 0923806 and IOS-1444974 and USDA Agricultural Experiment Station grant number CAD-XXX-6973-H. JE and CS acknowledge support from the Elsie Taylor Stocking Memorial Research Fellowship. JE acknowledges support from the Henry Jastro Research Fellowhip and CMS-M acknowledges support from the University of California Institute for Mexico (UCMEXUS)/Consejo Nacional de Ciencia y Tecnología (CONACYT) and Secretaría de Educación Pública (Mexico).

Supplemental Figures



**Supplementary Figure 1. Compartment separation of root-associated microbiota**

**References**

1. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature. Nature Research; 2012;488: 86–90. doi:10.1038/nature11237

2. Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, et al. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature. Nature Research; 2012;488: 91–5. doi:10.1038/nature11336

3. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, et al. Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl Acad Sci U S A. 2015;112: E911-20. Available: http://www.pnas.org/content/112/8/E911.full

4. Wagner MR, Lundberg DS, del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun. Nature Publishing Group; 2016;7: 1–15. Available: http://dx.doi.org/10.1038/ncomms12151%5Cnhttp://10.1038/ncomms12151

5. Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, et al. The soil microbiome influences grapevine-associated microbiota. MBio. American Society for Microbiology; 2015;6.

6. Finkel OM, Castrillo G, Herrera Paredes S, Salas González I, Dangl JL. Understanding and exploiting plant beneficial microbes. Curr Opin Plant Biol. 2017;38: 155–163. doi:10.1016/j.pbi.2017.04.018

7. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. Nature Publishing Group; 2012;6: 1621–1624. doi:10.1038/ismej.2012.8

8. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat Commun. Nature Publishing Group; 2017;8: 1784. doi:10.1038/s41467-017-01973-8

9. Santos-Medellín C, Edwards J, Liechty Z, Nguyen B, Sundaresan V. Drought Stress Results in a Compartment-Specific Restructuring of the Rice Root-Associated Microbiomes. MBio. American Society for Microbiology; 2017;8: e00764-17. doi:10.1128/mbio.00764-17

10. Edwards JA, Santos-Medellín CM, Liechty ZS, Nguyen B, Lurie E, Eason S, et al. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. Gore J, editor. PLOS Biol. Public Library of Science; 2018;16: e2003862. doi:10.1371/journal.pbio.2003862

11. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics. BioMed Central; 2012;13: 31. doi:10.1186/1471-2105-13-31

12. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. Nature Publishing Group; 2010;7: 335–336. doi:10.1038/nmeth.f.303

13. Al-Ghalith GA, Montassier E, Ward HN, Knights D. NINJA-OPS: Fast Accurate Marker Gene Alignment Using Concatenated Ribosomes. Eisen JA, editor. PLoS Comput Biol. Public Library of Science; 2016;12: e1004658. doi:10.1371/journal.pcbi.1004658

14. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. Nature Research; 2012;9: 357–359. doi:10.1038/nmeth.1923

15. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. Nature Publishing Group; 2016;13: 581–583. doi:10.1038/nmeth.3869

16. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. American Society for Microbiology; 2006;72: 5069–72. doi:10.1128/AEM.03006-05

17. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26: 2460–2461. doi:10.1093/bioinformatics/btq461

18. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. American Society for Microbiology (ASM); 2007;73: 5261–7. doi:10.1128/AEM.00062-07

19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. Oxford University Press; 2010;26: 139–140. doi:10.1093/bioinformatics/btp616

20. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. BioMed Central Ltd; 2010;11: R106. doi:10.1186/gb-2010-11-10-r106

21. Callahan BJ, Sankaran K, Fukuyama JA, McMurdie PJ, Holmes SP. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. F1000Research. 2016;5: 1492. doi:10.12688/f1000research.8986.1

22. Lozupone C, Knight R. UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. Society. 2005;71: 8228–8235. doi:10.1128/AEM.71.12.8228

23. McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Comput Biol. Public Library of Science; 2014;10: e1003531. Available: http://dx.doi.org/10.1371%2Fjournal.pcbi.1003531