Illumina sequencing and statistical analysis

Sequencing

1. Extraction of genome DNA

Total genome DNA from samples was extracted using CTAB/SDS method. DNA concentration and purity was monitored on 1% agarose gels. According to the concentration, DNA was diluted to 1ng/μL using sterile water.

2. Amplicon Generation

16S rRNA/18SrRNA/ITS genes of distinct regions(16SV4/16SV3/16SV3-V4/16SV4-V5, 18S V4/18S V9, ITS1/ITS2, Arc V4) were amplified used specific primer(e.g. 16S V4: 515F-806R, 18S V4: 528F-706R, 18S V9: 1380F-1510R, et. al) with the barcode. All PCR reactions were carried out with Phusion® High-Fidelity PCR Master Mix (New England Biolabs).

3. PCR Products quantification and qualification

Mix same volume of 1X loading buffer (contained SYB green) with PCR products and operate electrophoresis on 2% agarose gel for detection. Samples with bright main strip between 400-450bp were chosen for further experiments.

4. PCR Products Mixing and Purification

PCR products was mixed in equidensity ratios. Then, mixture PCR products was purified with Qiagen Gel Extraction Kit(Qiagen, Germany).

5. Library preparation and sequencing

Sequencing libraries were generated usingTruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA) following manufacturer's recommendations and

index codes were added. The library quality was assessed on the Qubit@ 2.0 Fluorometer (Thermo Scientific) and Agilent Bioanalyzer 2100 system. At last, the library was sequenced on an IlluminaHiSeq2500 platform and 250 bp paired-end reads were generated.

Data analysis

1. Paired-end reads assembly and quality control

1.1 Data split

Paired-end reads was assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence.

1.2 Sequence assembly

Paired-end reads were merged using FLASH (V1.2.7,http://ccb.jhu.edu/software/FLASH/) [1],a very fast and accurate analysis tool, which was designed to merge paired-end reads when at least some of the reads overlap the read generated from the opposite end of the same DNA fragment, and the splicing sequences were called raw tags.

1.3 Data Filtration

Quality filtering on the raw tags were performed under specific filtering conditions to obtain the high-quality clean tags[2]according to the QIIME(V1.7.0, http://qiime.org/index.html)[3]quality controlled process.

1.4 Chimera removal

The tags were compared with the reference database(Gold database, http://drive5.com/uchime/uchime_download.html)using UCHIME

algorithm(UCHIME Algorithm, http://www.drive5.com/usearch/manual/uchime_algo.html)[4]to detect chimera sequences, and then the chimera sequences were removed[5]. Then the Effective Tags finally obtained.

2. OTU cluster and Species annotation

2.1 OTU Production

Sequences analysis were performed by Uparse software (Uparse v7.0.1001，http://drive5.com/uparse/)[6]. Sequences with ≥97% similarity were assigned to the same OTUs. Representative sequence for each OTU was screened for further annotation.

2.2 Species annotation

For each representative sequence, the GreenGene Database (http://greengenes.lbl.gov/cgi-bin/nph-index.cgi)[7]was used based on RDP classifier(Version 2.2, http://sourceforge.net/projects/rdp-classifier/)[8]algorithmto annotate taxonomic information.

2.3 Phylogenetic relationship Construction

In order to study phylogenetic relationship of different OTUs, and the difference of the dominant species in different samples(groups), multiple sequence alignment were conducted using the MUSCLE software (Version 3.8.31，http://www.drive5.com/muscle/)[9].

2.4 Data Normalization

OTUs abundance information were normalized using a standard of sequence number

corresponding to the sample with the least sequences. Subsequent analysis of alpha diversity and beta diversity were all performed basing on this output normalized data.

3. Alpha Diversity

Alpha diversity is applied in analyzing complexity of species diversity for a sample through6 indices, including Observed-species, Chao1, Shannon, Simpson, ACE, Good-coverage. All this indices in our samples were calculated with QIIME(Version 1.7.0) and displayed with R software(Version 2.15.3).

Two indices were selected to identify Community richness:

Chao - the Chao1 estimator (http://www.mothur.org/wiki/Chao);

ACE - the ACE estimator (http://www.mothur.org/wiki/Ace);

Two indices were used to identify Community diversity:

Shannon - the Shannon index (http://www.mothur.org/wiki/Shannon);

Simpson - the Simpson index (http://www.mothur.org/wiki/Simpson);

One indice to characterized Sequencing depth:

Coverage - the Good's coverage (http://www.mothur.org/wiki/Coverage)

Reference

[1] Magoč, T.; Salzberg, S.L. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011, 27(21), 2957–2963.
[2] Bokulich, N.A.; Subramanian, S.; Faith, J.J.; Gevers, D.; Gordon, J.I.; Knight, R.; Mills, D.A.; Caporaso, J.G. Quality-filtering vastly improves diversity estimates from Illuminaamplicon sequencing. Nat. Methods 2013, 10(1), 57–59.
[3] Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 2010, 7(5), 335–336.
[4] Robert, C.E.; Brian, J.H.; Jose, C.C. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011, 27(16), 2194–2200.
[5] Haas, B.J.; Gevers, D.; Earl, A.M.; Feldgarden, M.; Ward, D.V.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Highlander, S.K.; Sodergren, E.; Methé, B.; DeSantis, T.Z.

Human Microbiome Consortium; Petrosino, J.F.; Knight, R.; Birren, B.W. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res. 2011, 21(3), 494–504.
[6] Edgar, R.C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 2013, 10(10), 996–998.
[7] DeSantis, T.Z.; Hugenholtz, P.; Larsen N. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 2006, 72(7), 5069–5072.
[8] Wang, Q.; Garrity, G.; Tiedje, J. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 2007, 73(16), 5261–5267.
[9] Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004, 32(5), 1792–1797.

Enzyme assay by double antibody sandwich method

The enzymatic activity was determined by dehydrogenase (DHA), catalase (CAT), and ammonia monooxygenase (AMO) activity using double antibody sandwich method. ELISA reagent kits were provided by Shanghai Enzyme-linked Biotechnology Co., Ltd., China. Samples were freeze-dried first, and then ground and screened through a 16 mesh sieve for enzyme assay. First, the HRP-labeled DHA, CAT, and AMO antibodies were coated on microplates to prepare solid-phase antibodies, then the corresponding enzymes were sequentially added to the microplates, forming antibody-antigen-ELISA antibody complexes. After thorough washing, substrate TMB was added for color development. TMB was converted to blue under the catalysis of HRP enzyme, and finally to yellow under the action of acid. The color depth was positively correlated with the enzyme content in the sample. Finally, measure the absorbance value using a microplate reader at a wavelength of 450nm, and calculate the enzyme content in the sample using a standard curve. The amount of enzyme that catalyzes 1 μmol of substrate in 1 min was defined as 1 International Unit (IU).

Determination of nitrification and denitrification potential

The nitrification potential were determined by constant temperature shaking culture with nitrification culture solution. The nitrification culture solution was prepared as follows, mix KH_2PO_4 solution (0.2 mg/L), K_2HPO_4 solution (0.2 mg/L), and $(NH_4)_2SO_4$ solution (0.05 mg/L) in a volume ratio of 3:7:30, and then adjust the pH to 7.2 with a dilute solution of H_2SO_4 or NaOH. The specific operation was as follows, (1) dry samples at 40 ℃, weigh an appropriate amount of sample and put it into a volumetric flask, then add nitrification culture solution in the ratio of 1.0 g sample : 10 mL solution, and cover the flask with absorbent cotton (group A); (2) take nitrification culture solution without sample added as blank control (group B), shake group A and group B at 20 ℃ for 24 h in a constant temperature water oscillator; (3) measure NO_3^--N concentration in the solution, and the difference in NO_3^--N concentration between group A and group B is the nitrification potential of the sample (10^{-2} mg/(kg·h)), which is expressed as the amount of NO_3^--N (mg) produced within 1 hour by 1.0 kg dried sample.

The denitrification potential were determined by constant temperature culture with denitrification culture solution. The denitrification culture solution was prepared as follows, mix KH_2PO_4 solution (0.2 mg/L), K_2HPO_4 solution (0.2 mg/L), KNO_3 solution (0.03 mg/L), and $C_6H_{12}O_6$ solution (0.07 mg/L) in a volume ratio of 3:7:30:10, and then adjust the pH to 7.2 with a dilute solution of H_2SO_4 or NaOH. The C: N ratio of the mixed solution is approximately 4:1. The specific operation was as follows, (1) dry samples at 40 ℃, weigh an appropriate amount of sample and put it

into a volumetric flask, then add denitrification culture solution in the ratio of 1 g sample : 10 mL solution, and cover the flask with rubber stopper or sealing film (group A); (2) take denitrification culture solution without sample added as blank control (group B), put group A and group B in a constant temperature incubator at 20 ℃ for 72 h; (3) filter 10 mL of suspension every 24 hours and measure NO_3^--N concentration, and supplement 10mL of denitrification culture solution after each suspension extraction. The difference in NO_3^--N concentration between pre-cultivation and post-cultivation is the denitrification potential of the sample (mg/(kg·h)), which is expressed as the amount of NO_3^--N (mg) consumed within 1 hour by 1 kg dried sample.