# Method

The method is only used as a reference for publication purpose.

Customers are responsible for the related risks of duplicate checking.

# 1. Experimental Procedure

## 1.1 Extraction of genome DNA

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

## 1.2 Amplicon Generation

16S rRNA/18SrRNA/ITS genes of distinct regions(16S V4/16S V3/16S V3-V4/16S V4-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 15 μL of Phusion® High -Fidelity PCR Master Mix (New England Biolabs); 2 μM of forward and reverse primers, and about 10 ng template DNA. Thermal cycling consisted of initial denaturation at 98℃ for 1 min, followed by 30 cycles of denaturation at 98℃ for 10 s, annealing at 50℃ for 30 s, and elongation at 72℃ for 30 s. Finally 72℃ for 5 min.

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

## 1.4 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 Illumina NovaSeq platform and 250 bp paired-end reads were generated.

# 2. Bioinformatics Analysis Pipeline

## 2.1 Paired-end reads assembly and quality c ontrol

### 2.1.1 Data split

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

### 2.1.2 Sequence assembly

Paired-end reads were merged using FLASH (V1.2.11, http://ccb.jhu.edu/software/FLASH/) (Magoc T et al.,2011), 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.

### 2.1.3 Data Filtration

Quality filtering on the raw tags were performed using the fastp (Version 0.23.1) software to obtain high-quality Clean Tags( Bokulich NA et al.,2012).

### 2.1.4 Chimera removal

The tags were compared with the reference database(Silva database (16S/18S), https://www.arb-silva.de/; Unite Database(ITS), https://unite.ut.ee/) using UCHIME Algorithm (http://www.drive5.com/usearch/manual/uchime_algo.html)to detect chimera sequences, and then the chimera sequences were removed (Edgar RC et al.,2011). Then the Effective Tags finally obtained.

# 3. Reference

Bokulich NA, Subramanian S, Faith JJ, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing.Nature Methods. 2012;10(1):57-59. doi:10.1038/nmeth.2276.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection.Bioinformatics. 2011;27(16):2194-2200. doi:10.1093/bioinformatics/btr381.

Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies.Bioinformatics. 2011;27(21):2957-2963. doi:10.1093/bioinformatics/btr507.