**1 Environment**

The environment is available from: <https://github.com/SMUJYYXB/GGMP-Regional-variations>. See details in Environment part.

**2 Data**

The sequences were deposited in the European Nucleotide Archive (ENA) under accession numbers PRJEB27507.

**3 Scripts**

3.1 Preparing raw data:

16S rRNA V4 raw data preparing: <Ion_torrent_preprocessing_forward.pl>

ITS1 DNA raw data preparing: <Ion_torrent_preprocessing_paired_ITS.pl>

3.2 Perform quality control of the raw data and delete the chimera:

3.2.1 <uchime_processing_b.pl>

3.2.2 <get_nonchimera_seq.pl>

3.3 Normalize the sequence data: <Sequences_Normalization.py>

3.4 Get the biom data: pick\_de\_novo\_otus.py

**4 Direction for use**

4.1 Configuring the system environment files and variables based on the (1) Environment.

4.2 Location of the files: put the scripts, metadata folder and supplementary files in the same path.

4.3 Run pipeline

4.3.1 Preparing raw data:

nohup perl Ion\_torrent\_preprocessing\_forward.pl <$PWD/fastq> <$PWD/primer.txt> <$PWD/output\_dir> &

nohup perl Ion\_torrent\_preprocessing\_paired\_ITS.pl <$PWD/fastq> <$PWD/primer.txt> <$PWD/output\_dir> &

4.3.2 Perform quality control of the raw data and delete the chimera:

nohup perl uchime\_processing\_b.pl <number> <tag.list> &

number: divide the tag.list file into n parts

tag.list: list the path of the files you need to detect chimera

nohup perl get\_nonchimera\_seq.pl <newname.fa.list> <output directory> &

newname.fa.list: list the absolute path of newname.fa flies of the former step into a new file

nohup python Sequences\_Normalization.py -p <Path> -e <suffix> -n <normalized numbers> &

cat \*normalized.fa > nochimera.normalized.fasta

4.3.3 Get the biom and tre data:

nohup pick\_de\_novo\_otus.py -i nochimera.normalized.fasta -o output -p denovo.params.txt &