MMDE-Wiki Documentation

Release latest

October 06, 2015

Contents

1 Amplicon Sequencing **1**
- 1.1 Experimental Design . . . 1
- 1.2 Sample Preparation . . . 1
- 1.3 Sequencing Library Preparation . . . 1
- 1.4 OTU Clustering . . . 2
- 1.5 PICRUSt . . . 3

2 Shotgun Sequencing **5**

CHAPTER 1

Amplicon Sequencing

1.1 Experimental Design

TODO: document experimental design issues

1.1.1 High-diversity samples

1.1.2 Low-diversity samples

1.2 Sample Preparation

TODO: document sample preparation

1.2.1 Sample Collection

1.2.2 DNA extraction

1.3 Sequencing Library Preparation

TODO: document library preparation

1.3.1 Library Types

These preparation methods are currently supported by the UMGC

Earth Microbiome Protocol (EMP)

- V4
- V4 (version 2)
- ITS1

Dual-indexed

- V1-V3
- V4

- V4 + PNA blockers
- V3-V5
- V4-V6
- V5-V6
- V3 (in development)
- V3-V4 (in development)

IIS (in development)

- V4
- V4-V6

1.4 OTU Clustering

TODO: OTU clustering

1.4.1 Process

The following steps are recommended for performing OTU picking using Qiime. Two packages are available to automate the process of running these steps on a dataset. If you have an MSI account you can use the "metagenomicsQC" package to automatically process a dataset using MSI computing resources and generate a html report summarizing the Qiime results. If you do not have access to MSI we provide a toolkit for generating a bash script to run all of these steps.

1. **Configuring Qiime**: FIX
2. **Create Mapping File**: FIX
3. **Trimming/Filtering/Converting**
   (a) Overlapping paired-end reads: Read pairs are stitched together and amplicon primers are removed using PandaSeq. Sequence IDs are converted to Qiime format and fastq files are converted to fasta format.
   (b) Non-overlapping paired-end reads: Samples with paired end reads that don't overlap are treated like single-end reads; the second (R2) read is ignored
   (c) Single-end reads: 3' ends are quality trimmed and the amplicon primer is removed. Sequence IDs are converted to Qiime format and fastq files are converted to fasta format. (Qiime scripts convert_fasta_qual_fastq.py and split_libraries.py used)
4. **Chimera Detection**: Chimeras are detected using ChimeraSlayer's usearch61 method.
5. **OTU Picking**: Qiime's pick_open_reference_otus.py script is used to pick OTUs using usearch61.
6. **Qiime Plots**: A series of plots based on the OTU table are generated using Qiime
7. **Alpha Diversity**:
8. **Normalization**:
9. **Beta Diversity**:

1.4.2 Running at MSI

Log in

Submit job

1.4.3 Running on a Mac or Linux computer

Qiime is a large software package with many dependencies that can be difficult to install. We recommend using MacQIIME is a precompiled installation of QIIME, with most of its dependecies, placed in one easy-to-install package. The MMDE-toolkit

Installing Qiime

1. Install MacQIIME or refer to the QIIME installation guide for other options.
2. Install USEARCH, versions 5.2.236(really?) and 6.1 are both required and not included in Qiime due to license issues
3. Install Pandaseq
4. Download MMDE-toolkit

Running Qiime

1. Create mappingfile
2. Start MacQIIME
3. Generate QIIME bash script:

```
qiime-1.9.1.pl > qiime.sh
```

4. Execute QIIME bash script:

```
bash qiime.sh
```

1.5 PICRUSt

TODO: document PICRUSt

1.5.1 Installing PICRUSt

1.5.2 Running PICRUSt

CHAPTER 2

Shotgun Sequencing

Guidelines for shotgun metagenomic sequencing will be developed in the future.

The Minnesota Microbiome Data Engine Wiki is developed and maintained by

- Dan Knights Lab
- University of Minnesota Genomics Center
- University of Minnesota Informatics Institute