

Qiime 2 on HPC cluster:

What is Qiime 2?

QIIME 2 is a powerful, extensible, and decentralized microbiome analysis package with a focus on data and analysis transparency. QIIME 2 enables researchers to start an analysis with raw DNA sequence data and finish with publication-quality figures and statistical results.

QIIME 2 is a complete redesign and rewrite of the QIIME 1 microbiome analysis pipeline. QIIME 2 will address many of the limitations of QIIME 1, while retaining the features that make QIIME 1 a powerful and widely-used analysis pipeline.

The full documentation for the Qiime 2 is found in the following links:

[Documentation](#)

Versions Available:

- Qiime 2 v 2020.6.0

How to load a version of Qiime 2?

To load a version of Qiime 2 on the HPC, use the following command:

```
module avail bio/qiime_2
```

Verify by using this command:

```
module list
```

Qiime 2 has a dependency on miniconda. So, the output should be two modules- miniconda and Qiime 2.

How to use Qiime 2 on the cluster?

Since the dataset used for Qiime is relatively big, it is preferable to use script rather than a bash terminal on compute node.

Input files:

This example and input files are directly taken from the [docs](#).

```
# Sample metadata
wget -O "sample-
metadata.tsv" "https://data.qiime2.org/2022.2/tutorials/moving-
pictures/sample_metadata.tsv"

# Make a directory for sequence reads
mkdir emp-single-end-sequences

# Download sequences reads
wget \
  -O "emp-single-end-sequences/barcodes.fastq.gz" \
  "https://data.qiime2.org/2022.2/tutorials/moving-pictures/emp-single-
end-sequences/barcodes.fastq.gz"

wget \
  -O "emp-single-end-sequences/sequences.fastq.gz" \
  "https://data.qiime2.org/2022.2/tutorials/moving-pictures/emp-
single-end-sequences/sequences.fastq.gz"
```

The Script:

To run a slurm job, the user must prepare input files. Users can use other software to do so. However, it is advised to purge all the modules before starting to use slurm script.

```
module purge
```

Use the following template for the script,

```
#!/bin/bash
#SBATCH -J test
#SBATCH -n 1
#SBATCH -c 16
#SBATCH -p main
#SBATCH -q main
#SBATCH --mem=10G

#Load the module
module load bio/qiime 2

#Go to the test directory
cd $SLURM_SUBMIT_DIR

#Convert input into Qiime artifact
qiime tools import --type EMPSingleEndSequences --input-path emp-
single-end-sequences \
--output-path emp-single-end-sequences.qza

# Demultiplex the sequence
qiime demux emp-single --i-seqs emp-single-end-sequences.qza --m-
barcodes-file sample-metadata.tsv \
--m-barcodes-column barcode-sequence \
--o-per-sample-sequences demux.qza \
--o-error-correction-details demux-details.qza
```

Schedule the job with the following sbatch command.

```
sbatch script.sbatch
```

All the processed files will be generated in the same directory as the sbatch script.

Where to find help?

If you are stuck on some part or need help at any point, please contact OIT at the following address.

<https://ua-app01.ua.edu/researchComputingPortal/public/oitHelp>