HUMAnN2

Back to Table of Contents

All of the code in this page is meant to be run on the command line unless otherwise specified.
Install Anaconda Download and install Anaconda3

Linux: Right-click on the "Linux 64-bit" button under PYTHON 3.5. Mac: use the command-line installer or the graphical installer.

** on the command line:**

```
# install Anaconda with python 3.5
bash Anaconda3-4.0.0-MacOSX-x86_64.sh

# now reopen your terminal window/login again

# now install a "python2" environment also containing anaconda
# This will let you switch between python2 and 3 easily
conda create -n python2 python=2.7 anaconda

# Load the python2 environment
source activate python2
```

Install HUMAnN2

1. Download it
2. Extract
3. Enter the directory, then run the install script.

```
# Choose an installation directory that you have access to
export INSTALL_DIR=/project/flatiron/lib/humann2/

# Run a local install
python setup.py install --home=$INSTALL_DIR
```

4. Add the appropriate paths to the system environment variables

```
# Add these lines to ~/.bash_profile
export PATH=$INSTALL_DIR/bin:$PATH
export PYTHONPATH=$INSTALL_DIR/lib/python:$PYTHONPATH
```

5. Download two databases that it needs

```
humann2_databases --download chocophlan full $INSTALL_DIR/db
humann2_databases --download uniref uniref50_diamond $INSTALL_DIR/db
```

Install Metaphlan2

1. Download it

2. Extract it (to the preferred installation directory, say `/project/flatiron/lib/metaphlan2`.
3. Add the appropriate paths to the system environment variables

```
# Add these lines to ~/.bash_profile
export PATH=/project/flatiron/lib/metaphlan2/biobakery-metaphlan2-f0bfc8620578:$PATH
export mpa_dir=/project/flatiron/lib/metaphlan2/biobakery-metaphlan2-f0bfc8620578
```

Run HUMAnN2

To run on a single fastq (or fasta) file:

```
humann2 -i input.fastq -o output
```

To run on a directory fastq (or fasta) file:

```
# only use R1 reads
for f in dir/*_R1_*.fastq; echo $f; humann2 -i $f -o output --output-format biom --remove-stratified-ou

# note: to run them in parallel (if you have many processors), replace
# the ";" before "done" with a "&"
# But do this with caution

# Then merge the files
merge_otu_tables.py -i $SAMPLE1_genefamilies.biom,$SAMPLE2_genefamilies.biom,$SAMPLE3_genefamilies.biom
merge_otu_tables.py -i $SAMPLE1_pathabundance.biom,$SAMPLE2_pathabundance.biom,$SAMPLE3_pathabundance.bi
merge_otu_tables.py -i $SAMPLE1_pathcoverage.biom,$SAMPLE2_pathcoverage.biom,$SAMPLE3_pathcoverage.biom

# convert to TSV if you want open as a spreadsheet
biom convert --to-tsv -i genefamilies.biom -o genefamilies.txt --table-type "Ortholog table"
biom convert --to-tsv -i pathcoverage.biom -o pathcoverage.txt --table-type "Pathway table"
biom convert --to-tsv -i pathabundance.biom -o pathabundance.txt --table-type "Pathway table"
```