Title: A role for gut microbiota in host niche differentiation

Authors: Greene LK, Williams CV, Junge RE, Mahefarisoa KL, Rajaonarivelo T, Rakotondrainibe H, O'Connell TM, Drea CM

General QIIME2 Pipeline

IMPORTANT NOTES:

comments will begin with a '#'
commands will begin with a '$'
this pipeline was built using QIIME 2-2019.4
Successful commands will always result in green text whereas errors or failures will show in red text
This project had sequences from two runs, which were processed separately until after denoising. This pipeline uses the numbers '1' and '2' to denote files relating to the two sequencing runs. Following merging, a single naming system is used.

1. Downloading Miniconda and QIIME2

```
        # follow instructions online to get miniconda and qiime2 installed (current version 2019.4):
        https://docs.qiime2.org/2019.4/install/native/
```

2. Activate QIIME

```
        $ source activate qiime2-2019.4
```

3. gzip forward reads, reverse reads, and barcode files

```
        $ gzip forward1.fastq
        $ gzip reverse1.fastq
        $ gzip barcoes1.fastq

        $ gzip forward2.fastq
        $ gzip reverse2.fastq
        $ gzip barcodes2.fastq
```

4. Import sequence files (note: The two sets of files must be placed into new directories, here labeled as "Sequences1" and "Sequences2". The files must be renamed within each directory as "forward.fastq.gz" "reverse.fastq.gz" and "barcodes.fastq.qz"

```
        $ qiime tools import
        --type EMPPairedEndSequences
        --input-path Sequences1
        --output-path sequences1.qza

        $ qiime tools import
```

```
        --type EMPPairedEndSequences
        --input-path Sequences2
        --output-path sequences2.qza

### 5. Demultiplex samples:

        $ qiime demux emp-paired
        --i-seqs sequences1.qza
        --m-barcodes-file map_file1.txt
        --m-barcodes-column BarcodeSequence
        --o-per-sample-sequences demux1.qza
        --o-error-correction-details error1.qza
        --p-rev-comp-barcodes
        --p-rev-comp-mapping-barcodes

        $ qiime demux emp-paired
        --i-seqs sequences2.qza
        --m-barcodes-file map_file2.txt
        --m-barcodes-column BarcodeSequence
        --o-per-sample-sequences demux2.qza
        --o-error-correction-details error2.qza
        --p-rev-comp-barcodes
        --p-rev-comp-mapping-barcodes

### 6. visualize demuxed samples

        $ qiime demux summarize
        --i-data demux1.qza
        --o-visualization demux1_visual.qzv

        $ qiime demux summarize
        --i-data demux2.qza
        --o-visualization demux2_visual.qzv

### 7. Denoise samples using DADA2

        $ qiime dada2 denoise-paired
        --i-demultiplexed-seqs demux1.qza
        --p-trim-left-f 0
        --p-trim-left-r 0
        --p-trunc-len-f 150
        --p-trunc-len-r 150
        --o-table table1.qza
        --o-representative-sequences repseqs1.qza
        --o-denoising-stats denoising-stats1.qza

        $ qiime dada2 denoise-paired
        --i-demultiplexed-seqs demux2.qza
        --p-trim-left-f 0
        --p-trim-left-r 0
```

```
        --p-trunc-len-f 150
        --p-trunc-len-r 150
        --o-table table2.qza
        --o-representative-sequences repseqs2.qza
        --o-denoising-stats denoising-stats2.qza

### 8. Visualize denoised stats

        $ qiime metadata tabulate
        --m-input-file denoising-stats1.qza
        --o-visualization denoising-stats1_visual.qzv

        $ qiime metadata tabulate
        --m-input-file denoising-stats2.qza
        --o-visualization denoising-stats2_visual.qzv

### 9. Merge tables and repseqs from the two runs

        $ qiime feature-table merge
        --i-tables table1.qza
        --i-tables table2.qza
        --o-merged-table table.qza

        $ qiime feature-table merge-seqs
        --i-data repseqs1.qza
        --i-data repseqs2.qza
        --o-merged-data repseqs.qza


### 10. Filter any unwanted samples and remove singletons (i.e.,
sequences only contained in 1 sample)

        $ qiime feature-table filter-samples
        --i-table table.qza
        --m-metadata-file samples_to_keep.txt
        --o-filtered-table table_filtered.qza

        $ qiime feature-table filter-features
        --i-table table_filtered.qza
        --p-min-samples 2
        --o-filtered-table table_filtered_no_singletons.qza


### 11. Assign phylogeny

        a. download the SILVA classifier: https://forum.qiime2.org/t/
silva-132-classifiers/3698

        $ qiime feature-classifier classify-sklearn
        --i-classifier silva_classifier.qza
```

```
        --i-reads rep_seqs.qza
        --o-classification assigned_taxa.qza

### 12. Visualize assigned taxa

        $ qiime metadata tabulate
        --m-input-file assigned_taxa.qza
        --o-visualization assigned_taxa_visual.qzv

### 13. filter (remove) chloroplasts and mitochondria

        $ qiime taxa filter-table
        --i-table table_filtered_no_singletons.qza
        --i-taxonomy assigned_taxa.qza
        --p-exclude mitochondria
        --o-filtered-table table_no_mitochondria.qza

        $ qiime taxa filter-table
        --i-table table_no_mitochondria.qza
        --i-taxonomy assigned_taxa.qza
        --p-exclude chloroplast
        --o-filtered-table table_to_use.qza

### 14. Create stacked bar charts

        $ qiime taxa barplot
        --i-table table_to_use.qza
        --i-taxonomy assigned_taxa.qza
        --m-metadata-file map_file.txt
        --o-visualization SBC.qzv

### 15. Summarize feature table

        $ qiime feature-table summarize
        --i-table table_to_use.qza
        --o-visualization table_to_use_visual.qzv
        --m-sample-metadata-file map_file.txt

### 16. Create a phylogenetic tree

        $ qiime phylogeny align-to-tree-mafft-fasttree
        --i-sequences repseqs.qza
        --o-alignment aligned-rep-seqs.qza
        --o-masked-alignment masked-aligned-rep-seqs.qza
        --o-tree unrooted_tree.qza
        --o-rooted-tree rooted_tree.qza

### 17. Create rarefaction plots (note: max depth is depth of sample
with deepest coverage)
```

```
        $ qiime diversity alpha-rarefaction
        --i-table table_to_use.qza
        --i-phylogeny rooted_tree.qza
        --p-max-depth 200000
        --m-metadata-file map_file.txt
        --o-visualization alpha-rarefaction.qzv

### 18. Calculate core alpha and beta diversity metrics

        $ qiime diversity core-metrics-phylogenetic
        --i-phylogeny rooted_tree.qza
        --i-table table_to_use.qza
        --p-sampling-depth 19000
        --m-metadata-file map_file.txt
        --output-dir 19000_core_metrics

### 19. Export diversity metrics

        # note: below is the general formula for exporting any single
file

        $ qiime tools export
        --input-path “file_name”.qza
        --output-path “file_folder_name”

### 20. Create tables with relative abundance of taxa. This project
focused on genus-level resolution, thus p = 6.

        # Collapse taxa
        $ qiime taxa collapse
        --i-table table_to_use.qza
        --i-taxonomy assigned_taxa.qza
        --p-level 6
        --o-collapsed-table table_collapsed.qza

        # Make relative abundances
        $ qiime feature-table relative-frequency
        --i-table table_collapsed.qza
        --o-relative-frequency-table table_relab.qza

        # Export above as a .biom and then convert to a .tsv
        $ qiime tools export
        --input-path table_relab.qza
        --output-path Relative_Abundance/

        $ biom convert
        -i /Relative_Abundance/feature-table.biom
        -o relabund.tsv --to-tsv

### 21. To group samples by a column in a map file prior to making
```

relative abundances, i.e., to calculate average relative abundance across each species, perform the following command before collapsing taxa, and use the outputted table in downstream commands:

```
$ qiime feature-table group
--i-table table_to_use.qza
--m-metadata-file map_file.txt
--m-metadata-column species
--o-grouped-table table_grouped.qza
--p-axis sample
--p-mode mean-ceiling
```