#Dynamic Assessment of Microbial Ecology (\*DAME\*)

This app is a user-friendly open source platform to mine, interpret and compare taxonomic information from metagenome or 16s rDNA datasets. \*DAME\* uses the R environment to perform microbial ecology data analyses. It is specifically designed to work directly with output files from the QIIME with as minimal file processing as possible at the user end. The software is freely available for use.

The current release (v0.1) assesses α-Diversity, β-Diversity measurements, and differential expression analyses of count data. \*DAME\* requires the BIOM file from QIIME and a CSV file containing the BIOM sample labels and metadata (experimental grouping data) associated with each sample. This app utilizes the Shiny framework to allow for dynamic and real-time interaction with virtually all aspects of the data workflow. Where possible, table and graphic outputs utilize D3 for a fully interactive experience.

##Getting Started:

###Data Import

\*DAME\* requires two files to operate:

1. BIOM file - QIIME output file typically found in the folder (OTU) during otu picking method using either of the programs (pick\_open\_reference\_otus.py, pick\_closed\_reference\_otus.py or pick\_de\_novo\_otus.py).

Note: Use the OTU generated file that has taxonomy details (e.g. otu\_table\_mc3\_w\_tax.biom). \*DAME\* will fail to recognize OTU table without taxonomy details.

Expect upwards of 20 second wait for files >10 MB

2. BIOM Metadata File - .CSV file containing a column with exact sample labels used in QIIMe analysis and experimental groupings. It is recommended to re-purpose the original map file used in QIIME analysis. Make changes in the group names and samples as you wish them to appear in figures and graphs. However please make sure the headers in metadata file and BIOM file matches exactly. This is critical for to work together.

Expect upwards of 15 second wait for files >2 MB.

3. TRE file (optional) - .TRE file, typically found in the same folder (OTU) from QIIME output.

BIOM files that are generated through other pipelines (which are in JSON format) must be converted to HDF5 format before loading into \*DAME\*.

\* Convert from biom to txt:

biom convert –i table.biom –o otu\_table.txt –to-tsv –header-key taxonomy

\* Convert back to biom:

biom convert –i otu\_table.txt –o new\_ otu\_table.biom –to-hdf5 –table-type=”OTU table”–process-obs-metadata taxonomy

###Data Preprocessing

To prepare data suitable for analysis by standard statistical procedures, after loading BIOM and metadata files you will be able to see description of your imported dataset such as number of taxonomic levels, number of OTUs, number of samples, and number of metadata categories. With the help of dropdown menu, data can be visualized at each taxa level. In its raw form you can visualize sample prevalence at each taxa level. Prevalence data can be sorted by all the variable such as percent of total OTUs, total reads, mean sample prevalence etc. Data files generated at this stage are downloadable in excel, pdf and csv format.

You can choose metadata filters, such as groups, variable to exclude and include in final analysis if necessary. Unchecking the boxes ahead of group, treatment names will exclude those particular groups from the final analysis.

\*DAME\* enables the implementation of standardized quality control procedures for microbiome data. You can choose minimum number of reads per measurement where OTUs will be filtered based on the percentage of reads that are above the read count selected. Percent threshold to filter OTUs also can be chosen for minimum numbers of reads that must be in certain percentage range of samples. If desired one can filter out other organisms that are not in the Bacteria domain, e.g., Archaea, etc.

Press the "Finalize Import and Filters" button when the desired filters are selected.

Finalized data set can be visualized at number of taxonomic levels with the use of dropdown menu. Numerical output will summarize the dataset for number of OTUs, number of samples, and number of metadata categories. Dropdown menu allows to visualize the data at each taxa level. You can also visualize sample prevalence at each taxa level. Prevalence data can be sorted by all the factors such as percent of total OTUs, total reads, mean sample prevalence etc. Generated data files are downloadable in excel, pdf and csv format.

###α-Diversity

Diversity of microbial communities can be visualized in the form of bar charts. \*DAME\* provides multiple metrics for measuring microbial a-diversity. In the third tab at the top, α-Diversity of microbiome data can be visualized. You can select any desired taxa level, select the α-diversity parameter such as observed, Chao1, shannan etc. In the final dropdown menu, you can chose variable such as groups, treatment individually or together. Press the finalize the α-diversity tab.

Selecting multiple taxonomic levels will require longer computation time.

Before downloading a final α-diversity plot or data you can chose a file name separately for each piece of data, under the tab ‘show/hide Plotting Data Downloading Options’. Advanced highcharter options you plots can be exported in vertical or horizontal orientation with several options of theme such as available.

α-diversity output in the form of data and plotted graphs can be downloaded in the form of CSV, excel and PDF and graphs can be downloaded in the form of PNG, SVG, JPEG and PDF format. a-diversity anova results are presented as MEAN (SEM). P-value(s) are derived from 1-way ANOVA.

###β-Diversity

After selecting the desired inputs such as taxanomic levels, selecting groups/variables for PERMANOVA analysis and number of desired permutations for PERMANOVA, click the finalize b-diversity tab. Permutational Manova (PERMANOVA) describes if variation in community composition can be attributed to different experimental treatments or control variables.

Next, you can chose the b-diversity parameters on two basis i.e. dissimilarity based indices such as Bray-Curtis, Jaccard, Mountford etc and distance based indices such as Euclidean, Manhattan etc. For ordination method one can chose from a vareity of options such as Principal Cordinate Analysis, Non-Metric Multidimensional scaling etc. In last two drop down menus you can assign color mapping as well as shape mapping variables. Confidence elipses can be toggled for shape and placement, which can also be turned off by unchecking the box. Further there are scatter 3D options are available to change the size of size mapping variable or color opacity of points.

β-diversity PERMANOVA output in the form of data can be downloaded in the form of CSV, excel and PDF. Plotted graphs can be downloaded in the form of PNG, SVG, JPEG and PDF format.

###Differential Abundance

You can determine differential abundance at any taxa level. Once you select the taxonomic level, you can select the differential abundance analysis for variables. For pairwise comparison any two groups can be selected from the dropdown menu. The box plot shows the abundances of the selected taxa in each group/treatment.

Differential abundance is represented in the form of box plots. Raw numerical data used for box plots is also available for download in the form Excel, PDF and CSV format. Boxplots can be downloaded in the form of PNG, SVG, JPEG and PDF format.