myPhyloDB: A local database storage and retrieval system for the analysis of metagenomic data

Overview:

myPhyloDB is a user-friendly personal database with a browser-interface for accessing and analyzing taxonomic data from multiple projects and/or sequencing runs.

The goal of myPhyloDB is to allow for easy comparisons and statistical analysis of microbial (i.e., fungi or bacteria) taxonomic abundance across projects, soil types, and management scenarios.

For questions/comments or requests for additional features please contact:

Daniel Manter
Soil-Plant-Nutrient Research Unit
USDA-ARS
Fort Collins, CO 80526
phone: (970) 492-7255

This manual is a reference for the use of myPhyloDB.

Table of Contents:

1. Installation

Windows:

Double-click the installer (myPhyloDB_1.0_Win_x64_install.exe) and follow the prompts. If you are upgrading/reinstalling myPhyloDB and would like to keep your current database, make sure the “Default database” is unchecked during installation; otherwise, all components should be selected The program will install a myPhyloDB shortcut to your start menu (Windows 7) or start screen (Windows 8). Clicking the myPhyloDB icon will open a terminal and web-browser that provides the user-interface for running the myPhyloDB program. To exit the program type ctrl-c in the terminal and manually close your browser. Uninstalling myPhyloDB will not remove your database or uploaded files. The default installation folder for myPhyloDB will be: 'C:\Users\<user_name>\AppData\Local\myPhyloDB'.

Linux:

Double-click or run the installer from your terminal (myPhyloDB_1.0_Linux_x64_install.sh). If a previous version of myPhyloDB is detected you will be prompted to either keep your old database or re-install the default database. The program will install a myPhyloDB shortcut to your Desktop. Clicking the myPhyloDB icon will open a terminal and web-browser that provides the user-interface for running the myPhyloDB program. To exit the program type ctrl-c in the terminal and manually close your browser. MyPhyloDB must be manually uninstalled by deleting the appropriate folders. The default installation folder for myPhyloDB will be: 'home/<user_name>/myPhyloDB'.

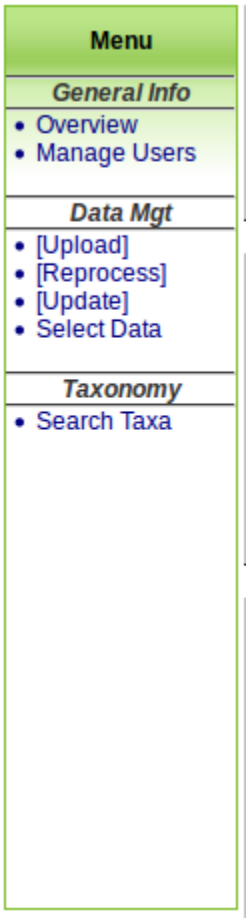
Remote access:

myPhyloDB will run as a local server on your host machine allowing others on your local intranet to access myPhyloDB (unless disabled using your computer's firewall settings) without installing a separate copy. This may be useful for laboratories that want to share data across multiple users. To access myPhyloDB from a remote computer you must first obtain the IP address of the host machine (in a terminal on the host machine, type 'ipconfig' for Windows or 'ifconfig' for Linux), then in the address bar of your remote computer's browser enter the following address 'xxx.xxx.x.xx:8000/myPhyloDB/home/' replacing the x's with the appropriate IP address. Depending upon your local LAN/WAN setup, connection to the host machine may fail using a WiFi connection. If this happens, please try a wired connection to your LAN or contact your local IT support. All data uploads and/or removal of projects by authorized (see Admin section) remote users will be saved to the host computer's installation of myPhyloDB.

2. Home Screen and Sidebar

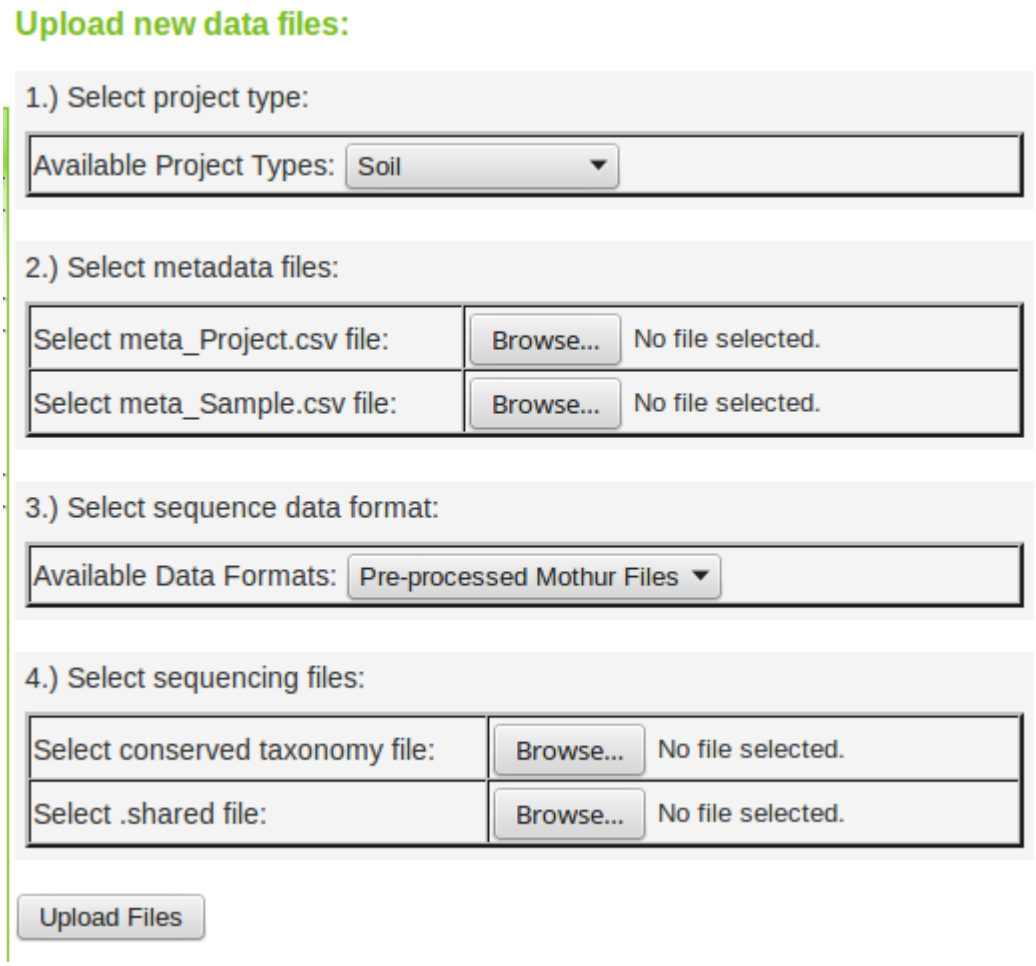
The home screen (http://127.0.0.1:8000/myPhyloDB/home/) provides general information about myPhyloDB as well as links to this instruction manual and example files for uploading new projects into myPhyloDB.

Navigation between the various pages and analyses provided by myPhyloDB is performed using the Menu sidebar at the left of the screen. The first time you launch myPhyloDB, the sidebar should look like the picture to the left. Once you have selected some projects/samples for analysis, the sidebar will show links to the various data analysis pages, as shown in the picture on the right. Selected projects/samples are saved to a 'cookie' and depending upon your browser settings will be stored between sessions.



3. Uploading New Data

To upload data, click “[Upload Data]” on the left hand menu (http://127.0.0.1:8000/myPhyloDB/upload/). For security purposes, this page can only be accessed by an authorized user – to add/remover users see the Admin section of this manual. Uploading new data consists of 4 steps: 1) selecting your project type, 2) selecting your metadata files, 3) selecting your sequence data file format, and 4) selecting your sequencing files.



3.1.1 Project type

myPhyloDB currently supports six different project types (Soil, Air, Water, Microbial, Human-associated, and Human microbiome). Each project type supports a different set of default variables as outlined here (http://www.mothur.org/wiki/MIMarks_Data_Packages). Blank sample files for each project type can be downloaded from myPhyloDB's homepage.

3.1.2 Select metadata files

Each upload requires two files with metadata (project and sample files), example of both can be downloaded from myPhyloDB's homepage. All columns in the Project and Sample files are required and the upload will fail if any changes are made to the header row; however, missing data may be set to 'null'. MyPhyloDB does not perform any unit checking or data conversions, so consistent units should be used for all samples throughout your database.

Only one project can be uploaded at a time (i.e., one row of data in your project file). However, samples (i.e., new sample_name) may be added to an already uploaded project by setting the project_id to the auto-generated UUID found in the DataTable located on the "Select" data page. Similarly, you may add new sequence data to an existing sample by setting both the project_id and sample_id values to the appropriate auto-generated UUIDs.

3.1.3 Select sequence data format

myPhyloDB supports the upload of 1) pre-processed mothur data files, 2) raw 454 pyrosequencing files and 3) raw MiSeq data files. The files required for submission will change depending upon your selection.

3.1.4 Select sequencing files

Example 1: Pre-processed Mothur Files

All of the files necessary to upload a sample pre-processed mothur project can be found on myPhyloDB's homepage (prefix: Example1). This option allows users to upload files that have already been processed using Mothur. To use this option, you will need two mothur-generated files: *.shared and *.cons.taxonomy. The shared file can be generated using the make.shared command but must contain only one OTU level (e.g., label = 1). The taxonomy file can be generated using the classify.otu command using the same OTU level. For example, assuming you have the following three mothur files (final.fasta, final.names, final.groups) run the following commands in mothur to generate the required files.

classify.seqs(fasta=final.fasta, template=gg_13_5_99.fasta, taxonomy=gg_13_5_99.pds.tax)

phylotype(taxonomy=final.pds.wang.taxonomy, name=final.names, label=1)

make.shared(list=final.pds.wang.tx.list, group=final.groups)

classify.otu(taxonomy=final.pds.wang.taxonomy, name=final.names, group=final.groups, list=final.pds.wang.tx.list)

If you follow the above procedure the two files needed for upload will be named: "final.pds.wang.tx.shared" and "final.pds.wang.tx.1.cons.taxonomy". Due to taxa naming differences between the various reference databases (e.g., RDP, GreenGenes, SILVA), it is recommended that a single reference database be used consistently with myPhyloDB. Also, the architecture of myPhyloDB is such that all OTUs must have an entry for all seven main

taxonomic levels (I.e., Kingdom, Phyla, Class, Order, Family, Genus, Species) so to avoid manually editing your taxonomy file we recommend the GreenGenes or SILVA reference databases provided by mothur (www.mothur.org/wiki/Taxonomy_outline). If necessary, 'unclassified' can be used for any taxonomic level without relevant information (e.g.,, species when using RDP).

Example 2: Raw 454 Files

All of the files necessary to upload a sample raw 454 pyrosequencing project can be found on myPhyloDB's homepage (prefix: Example2). Using this option will utilize myPhyloDB's embedded copy of mothur (currently vers 1.35.1) and allow for future reprocessing using myPhyloDB. For this option, you will need to upload three files: sff file (standard 454 flow file), oligo file (file with sample names, barcodes, and primers), and a mothur batch file.

Experienced mothur users may wish to alter the provided batch file to match their current sequencing analysis pipelines; however, please note that the pipeline must create the following 5 files (final.fasta, final.names, final.groups, final.taxonomy, and final.shared); any deviation from the above naming conventions and the upload process will fail.

Example 3: Illumina/MiSeq Files

All of the files necessary to upload a sample raw Illumina/MiSeq project can be found on myPhyloDB's homepage (prefix: Example3). Using this option will utilize myPhyloDB's embedded copy of mothur (currently vers 1.35.1) and allow for future reprocessing using myPhyloDB. For this option, you will need to upload the following files: 3-column config file (file with sample names and fastaq file names), fastaq files (forward and reverse for each sample), and a mothur batch file. Please note, that the current default pipeline only supports the 3-column config file option and processing of fastaq files that have had their barcode/primers removed. Also, please be sure that you select all of the appropriate fastaq files using the available fastaq file chooser, which supports multiple file selection.

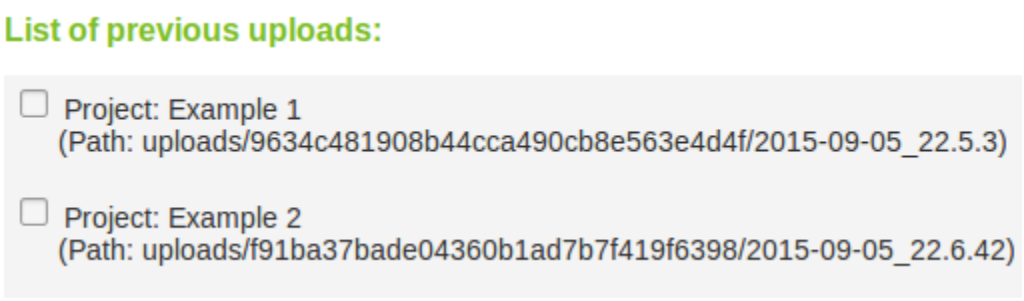
Experienced mothur users may wish to alter the provided batch file to match their current sequencing analysis pipelines; however, please note that the pipeline must create the following 5 files (final.fasta, final.names, final.groups, final.taxonomy, and final.shared); any deviation from the above naming conventions and the upload process will fail.

3.1.5 Upload Files

Once you have completed the four steps above, click “Upload Files” to begin the upload process. Note: the upload process includes multiple steps and may take anywhere from a few minutes to hours depending upon the project size and your computer speed. For your convenience, a progress bar will appear below the “Upload Files” button documenting the status of the upload and parsing steps required to populate the myPhyloDB database.

3.2 Removing Data

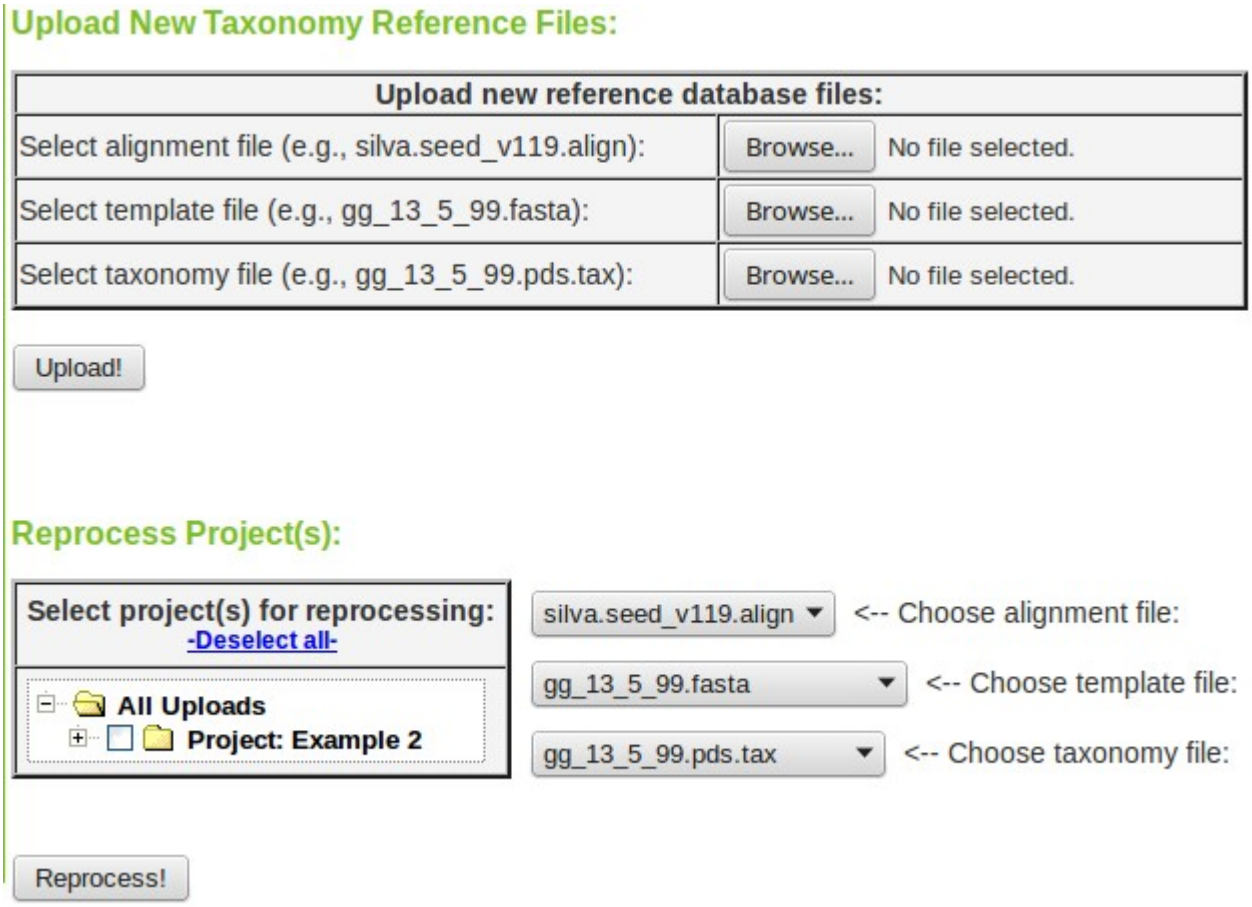
At the bottom of the “[Upload Data]” page is a list of all previous uploads to your myPhyloDB database. Each item in the list is categorized by project name and the upload path, which contains the timestamp when the upload was submitted. If you want to remove any of these uploads simply click the appropriate box and then the “Remove selected” button. Edited projects (i.e., new submission files) can then be uploaded as described above.



4. Reanalyzing Data

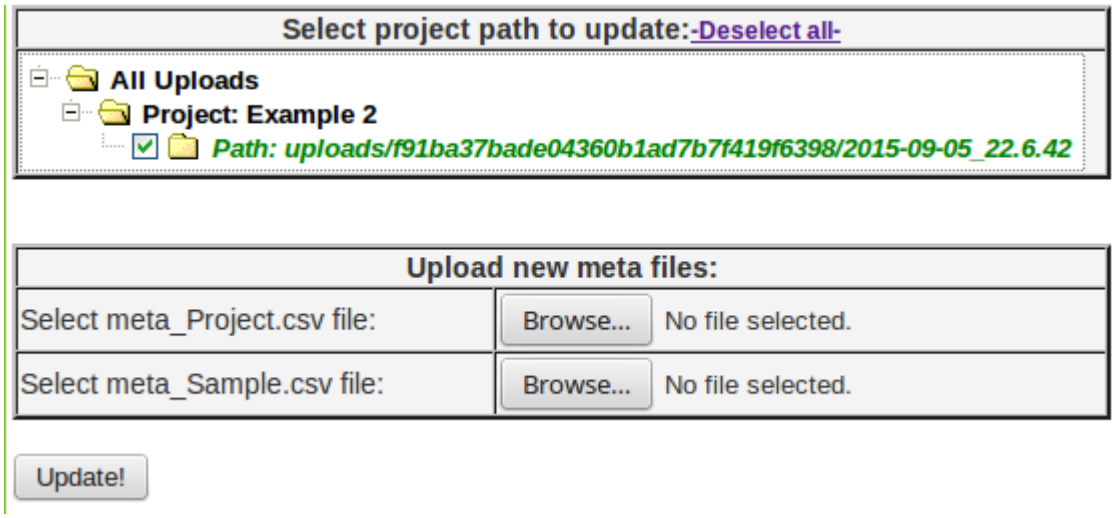
The alignment and classification files (i.e., template and taxonomy) can be conveniently updated for any project(s) contained in myPhyloDB.

To do this, simply upload any new alignment, template, or taxonomic reference files using the [Reprocess] page. Next, select the projects which need to be updated in the project tree, and select the correct (updated) reference files from the drop down menus, then press “Reprocess!”. Note: this will take anywhere from a few minutes to hours depending upon the project size and your computer speed



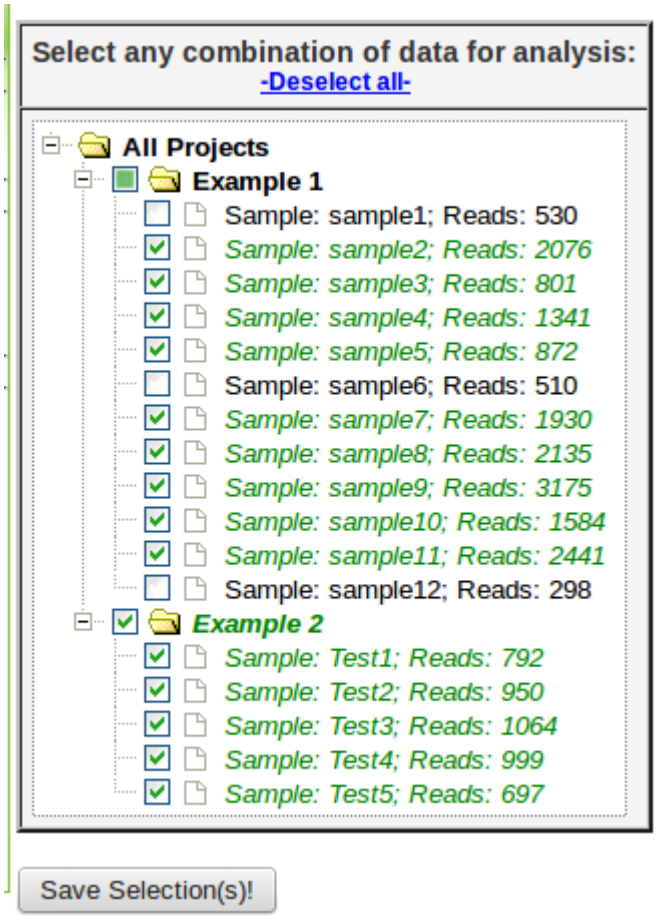
5. Updating Metadata

To update a previously uploaded project with new sample or project metadata, click the [Update] button on the sidebar. Then, select the project and path you wish to have updated. Use the file choosers to select the new project and/or sample file(s) to be used for updating, then press “Update!”. Note: the new project/sample files must contain the correct project and sample UUIDs for the updating procedure to correctly find and update the previously uploaded samples. The correct UUIDs can be obtained from the archived copy of these files (i.e., in the path shown on the project tree) or from the DataTable found on the select data page.



6. Selecting Data for Analysis

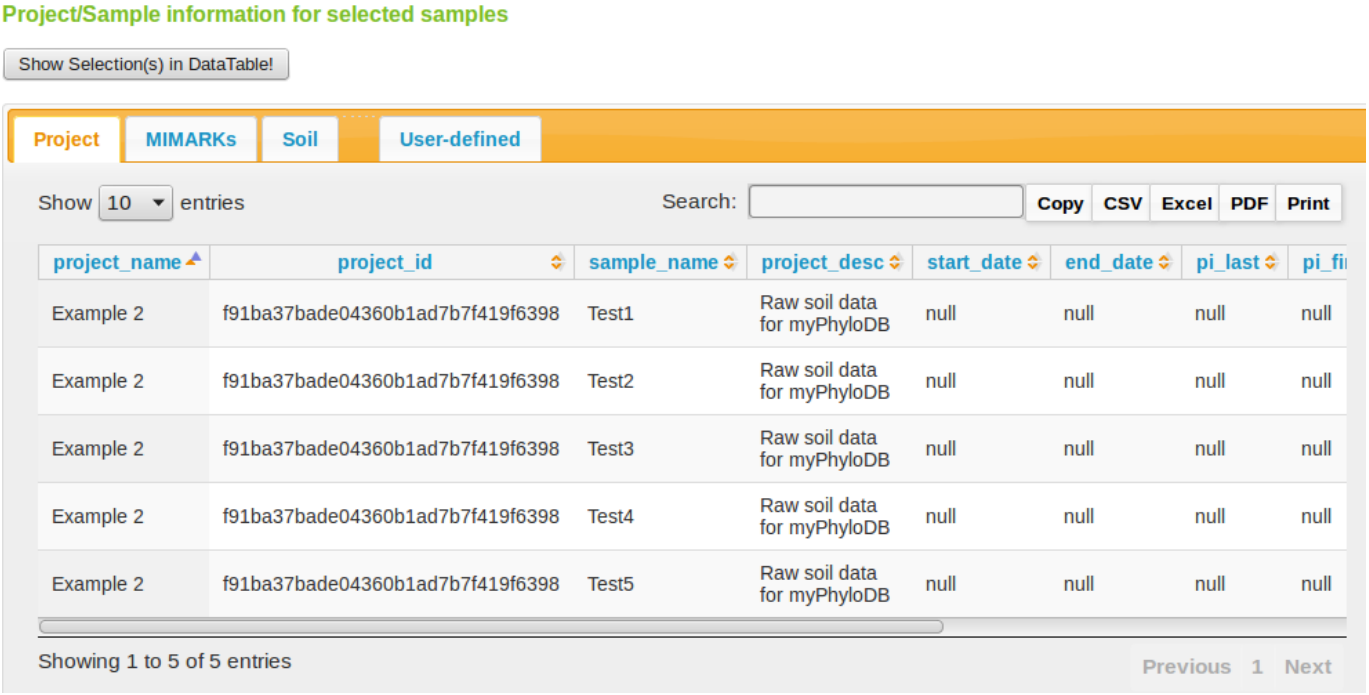
To select data for analysis, click “Select Data” on left hand menu (http://127.0.0.1:8000/myPhyloDB/select/). On the select data use the project/sample tree provided to select any combination of projects or samples desired. By default, if a Project checkbox is selected all samples for that project will also be selected. Each project can be expanded and individual samples can be manually selected/deselected. The project/sample tree is organized by project and sample names; however, the project and sample descriptions can be viewed by hovering the mouse over the appropriate name. In addition, the total number of sequence reads for each sample is shown in parentheses next to the sample name. Hovering the mouse over any sample will also display the sample description.



Completely selected projects will have a green checkmark; whereas, partially selected projects will be filled in with green and the selected samples will have a green checkmark. For your convenience, all selections can be cleared using the -Deselect all- link above the tree.

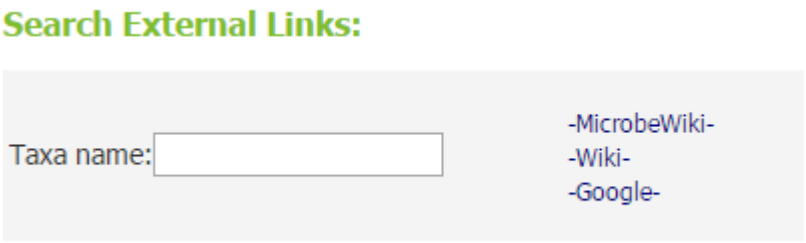
Once you have selected the data you wish to analyze further, click the “Save Selection(s)!” button below the project/sample tree. Note: Upon clicking the button, a pop-up window will appear saying “Selected sample(s) have been recorded!”, press “OK” and proceed to the “Analysis” section of myPhyloDB or explore the selected using the DataTable below. If this window does not appear you may need to restart myPhyloDB with administrator privileges. To do so, right-click on the myPhyloDB icon and select “run as administrator”.

The metadata associated with the each selected project/sample can be displayed in a DataTable by clicking the “Show Selections(s) in DataTable!”. Data is organized into categories (Project, MIMARKS, Soil, Water, etc.). You may switch between these categories using the DataTable tabs. All samples should populate the Project, MIMARKs (minimum information about a marker gene sequence), and User-defined tabs; plus one additional tab (e.g., Soil, Air, Water, etc.) that is dependent upon the project type.



Each DataTable includes a searchbox that can be used to search any field of the displayed table. In addition, each table may be exported to a variety of formats using the button at the top-left of the data table.

7. Search Taxa:



myPhyloDB provides a search Taxa page (http://127.0.0.1:8000/myPhyloDB/taxa/) to allow users to explore the taxonomic data contained in your myPhyloDB database. The "Taxa name" textbox at the top of the page allows users to quickly search various web sites with a user inputted taxa name. The datatable contains the full taxonomic name of each taxa in your database. For each taxonomic level a unique ID was generated by myPhyloDB for internal tracking purposes and to avoid confusion if duplicate taxonomic names exist. All results in myPhyloDB (next section) will include both taxonomic names and IDs which can be used to identify full taxonomic profiles using this data table. You can also export the table data to CSV, Excel, or PDF files or send the data directly to a printer.

All taxa in database:

Show 10 entries Search: Copy CSV Excel PDF Print

	Kingdom Name	Kingdom ID	Phylum Name	Phylum ID	Class Name
0	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Gammaproteobacteria
1	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Deltaproteobacteria
2	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Betaproteobacteria
3	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Gammaproteobacteria
4	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Alphaproteobacteria
5	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Alphaproteobacteria
6	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Gammaproteobacteria
7	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Proteobacteria	22bf9fbd71fd44fca58e573fbd9f987d	Alphaproteobacteria
8	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Acidobacteria	1fa0628580bd4615ad56701c82622145	Acidobacteriia
9	Bacteria	feaf34ed21cf42c68534209baa1f51f0	Bacteroidetes	c03f4a25bef84d3db03394559e08a6c6	Flavobacteriia

Showing 1 to 10 of 481 entries Previous 1 2 3 4 5 .. 49 Next

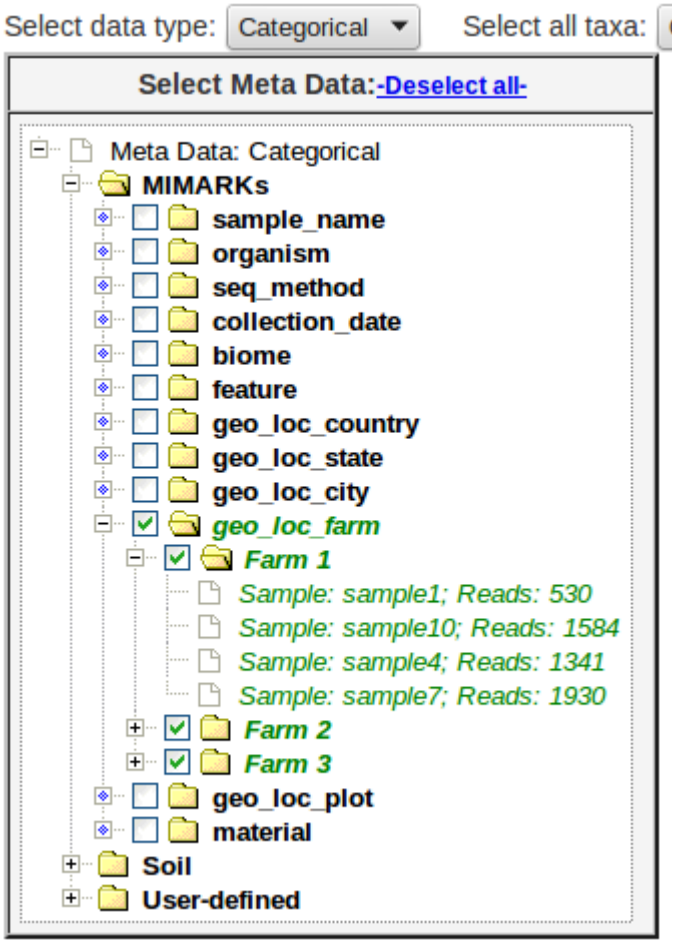
8. Analysis:

Once you have selected the samples you would like to analyze, on the menu sidebar, under the “Analysis” heading, select the type of analysis you would like to perform (Univariate: ANOVA/Regr, DiffAbund, or Multivariate: PCoA).

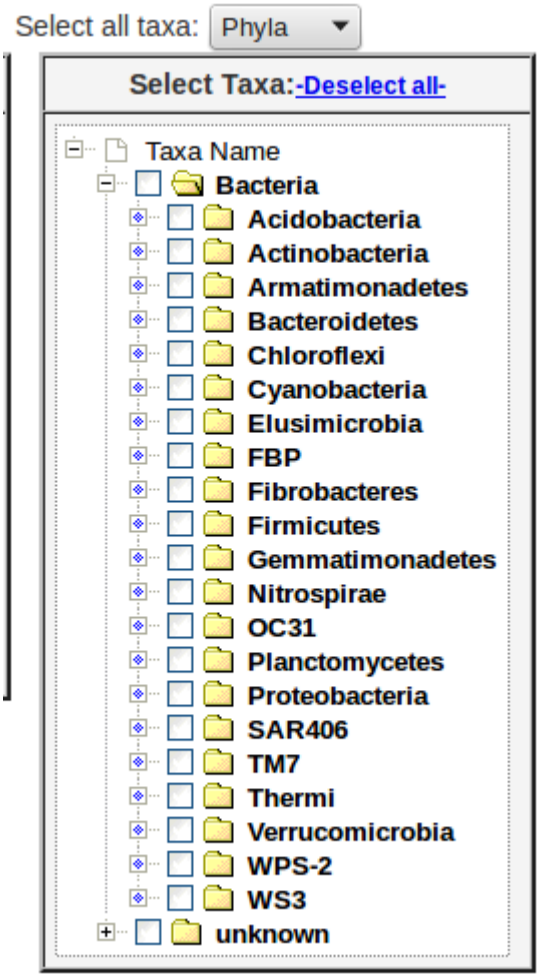
8.1. ANOVA/Regr:

For categorical variables, this analysis (http://127.0.0.1:8000/myPhyloDB/ANOVA/) will produce a bar chart and perform an ANOVA to test if the taxonomic level(s) chosen are significantly different between the meta-variables chosen. For quantitative variables, this analysis will produce a scatter plot and perform a linear regression between the chosen taxonomic level(s) and a single meta-variable.

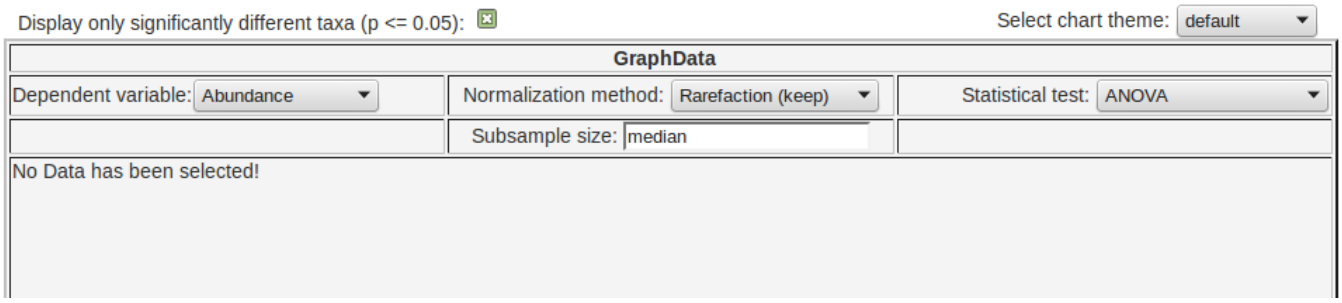
To run this analysis, first select the data type (i.e., categorical or quantitative) from the drop down menu “Select data type”, which will switch the Select Meta Data tree between the variable of each type. Next, select your meta-variable(s) of interest. Any variables where all selected samples are blank (i.e., null values) will generate the following alert “No samples are available for this variable!” upon selection. Also, any samples with null data will not be included in the final analysis. Fully expanding any meta-variable will result in a list of all of the samples that contain non-null values for that variable. The project name for each sample can be found by hovering the mouse over that sample in the tree.



Once you have selected your meta variables, you must select taxonomy data either from the drop down menu (selects all taxa at the select level) or by selecting specific taxonomic name(s) of interest from the taxonomy tree. Any desired combination of taxonomic level(s) and name(s) can be selected using the taxonomic tree by simply selecting the appropriate checkboxes. Please note, that if you use the “Select all taxa” drop down menu, the taxa tree will automatically be emptied of all selections.



The final selections required for analysis are all located within the graph table of the analysis page. Here you can select your dependent variable (abundance (counts), species richness, or Shannon's Diversity Index), normalization method (none, rarefaction, proportion, DESeq2), and statistical test (ANOVA or Welch's two-sided t-test). Optionally you may also choose to display only significantly different taxa ("Display only significantly different taxa" checkbox).

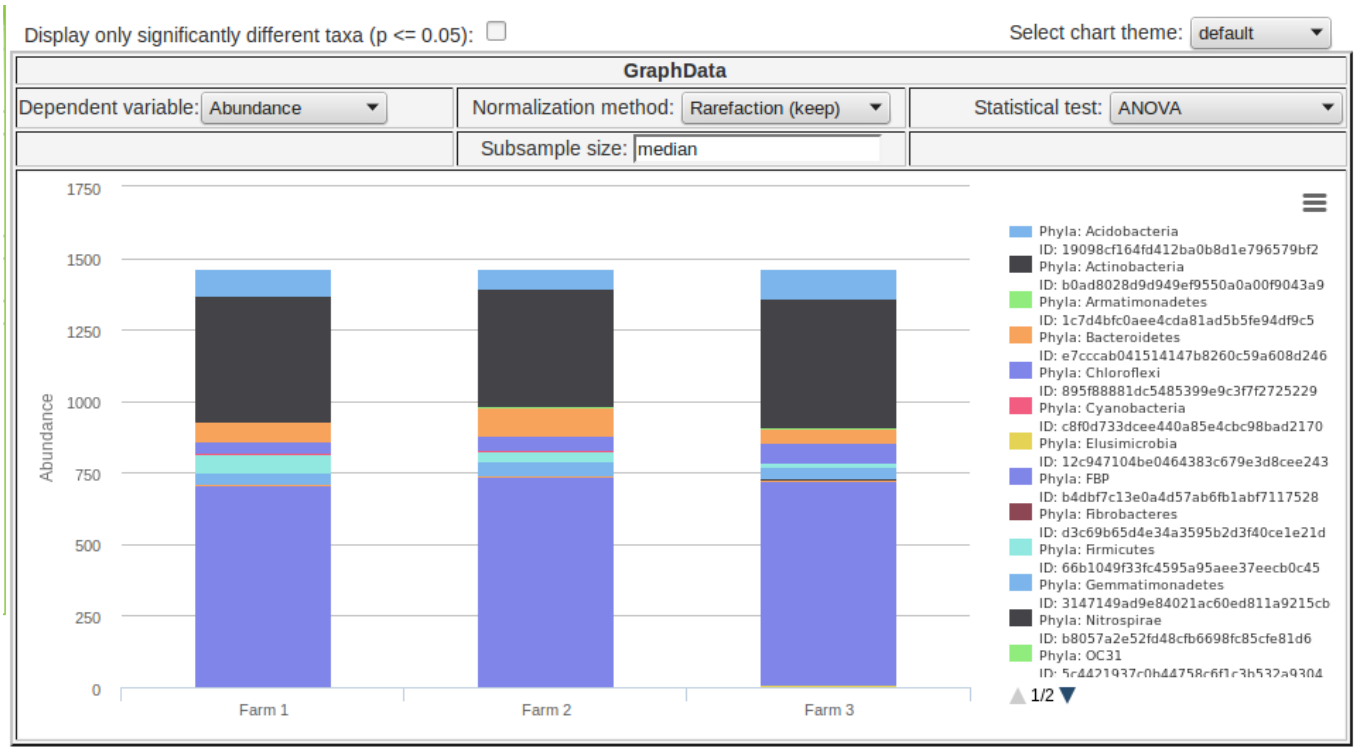


Once you are satisfied with your data choices, click the "Run Analysis!" button on the left menu. The button will change from gray to yellow as analysis is running, then to green when the analysis is complete. If a new combination is selected, the button will change back to gray. If the button turns red check to make sure that both meta-variable(s) and taxonomic name(s) have been selected. Once the analysis is complete (the "Run Analysis" button is green), scroll down to see your results. Any errors (e.g., you forgot to select a taxa level) as well as the current analysis step will be displayed in the graph table.

When you analysis is complete, a new bar graph will be displayed in the graph table along with the statistical results and raw data in the boxes below.

Bar Graph:

The bar graph will display the taxa averages for each meta variable level selected. The chart color theme may be changed (without rerunning the analysis) using the drop down menu above the graph. In addition all myPhyloDB charts can be downloaded by clicking one of the button (3 horizontal bars) just above the figure legend. Specific taxa can also be removed from the displayed graph simply by clicking on the text in the figure legend.



Test Results:

At the top of “Test Results” section is a summary of the data normalization step, which includes the number of samples normalized (or removed) and the number of reads used for rarefaction. This section also displays an ANOVA table and Tukey's HSD posthoc test results (R base package) for each taxa level included in the analysis.

Raw Data (Tabular):

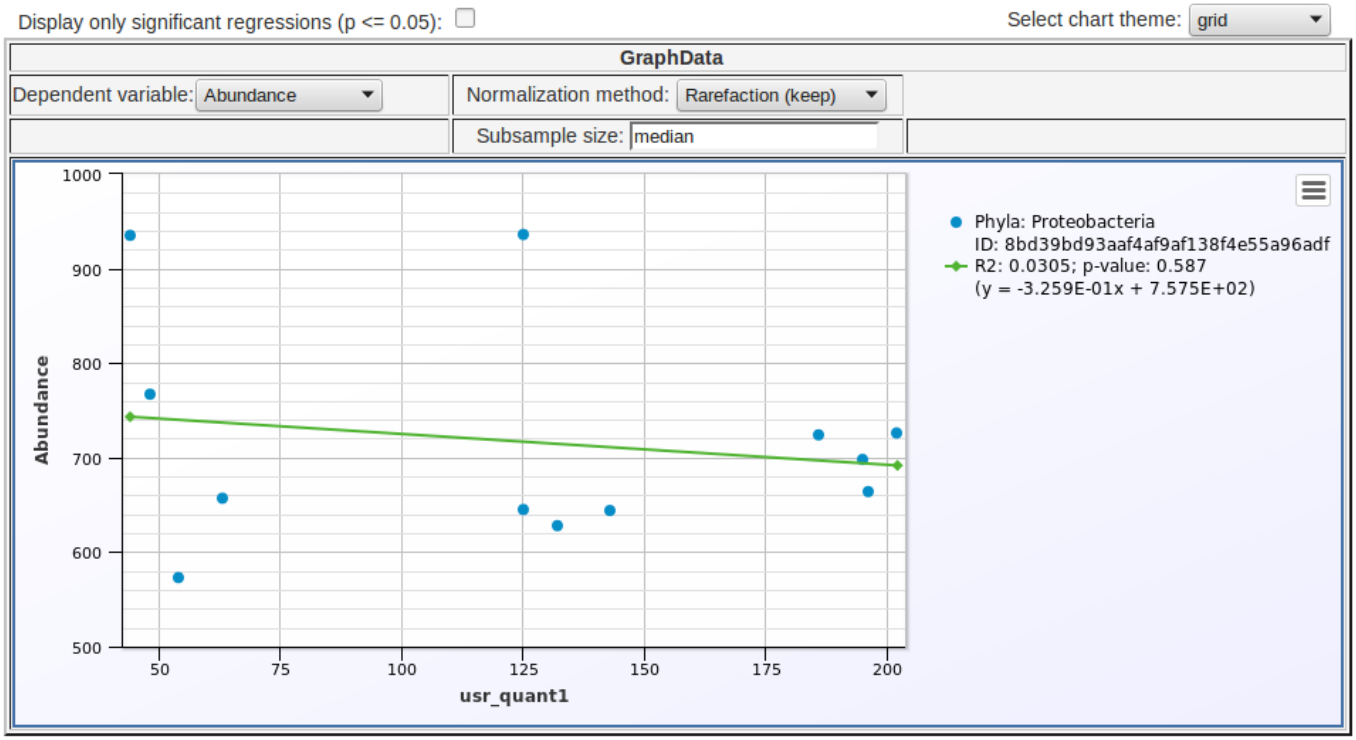
A raw data DataTable is also output to this page, which includes the selected metadata and normalized dependent variable (e.g., abundance counts) for each taxa level included in the analysis. The full taxonomic classification for each taxonomic level in the DataTable can be

obtained by searching the database with the appropriate taxa_id using the “Search Taxa” link on the main menu (page 12).

Raw Data (Biom):

The normalized data is also output to a textbox in biom format to allow for easy export and use with other software packages.

Quantitative variables are handled in a similar manner and will produce a scatter plot instead of a bar graph. However, to avoid unit conflicts only one meta variable may be analyzed at any time; and instead of an ANOVA a simple linear regression analysis is performed.

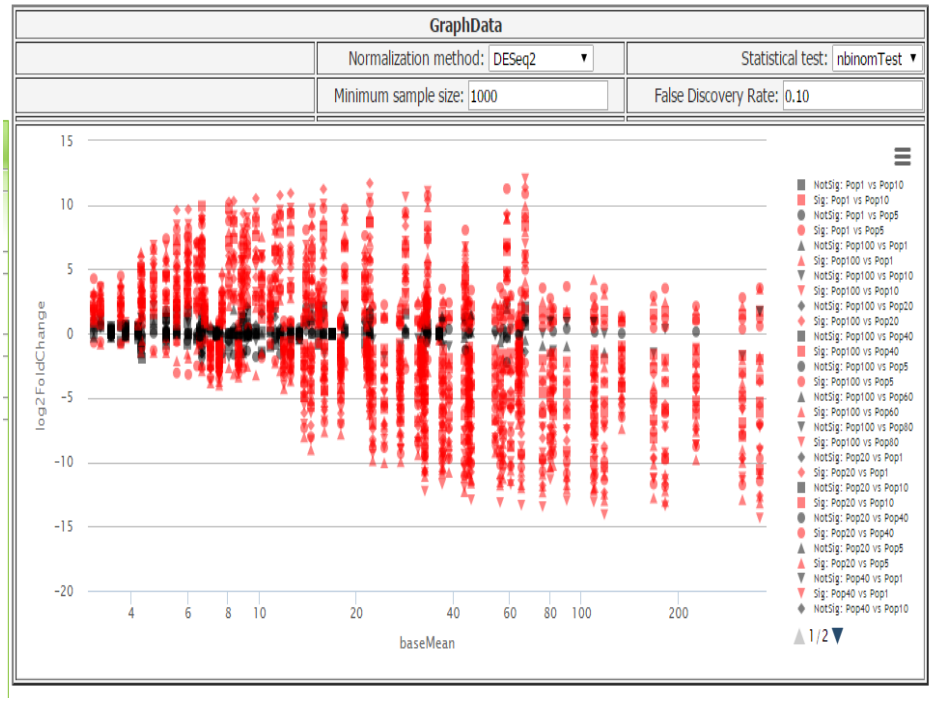


8.2 Diff Abund

The basic layout and data selection of the Diff Abund page is similar to the ANOVA/Regr page of myPhyloDB. The only differences are (1) the removal of the taxonomic tree (i.e., data can only selected by taxa level) and (2) the normalization and statistical test options have been changed. The Diff Abund procedure is part of the DESeq2 R package and more details on the procedure can be found here.

Graph:

The Diff Abund graph has a logarithmic scale for the x-axis (baseMean) and a standard scale for the y-axis (log2FoldChange). All significant data points are plotted in red (significance being determined by a p value of <= 0.05) and non-significant in black. If more than one meta variable is selected, the analysis will compare each independent treatment combination to one another. Main effects (one effect independent of the other) are not allowed; however, the analysis can easily be rerun with only one variable selected. The plot shows a comparison of each sample with every other sample, with respect to your selected meta-variables. As with the other graphs in myPhyloDB, you can toggle individual data points on and off. You can also mouse-over any point on the graph to see a tooltip description.



Normalization Results:

A summary of the samples meeting your normalization criteria.

NbinomTest Resuts:

The nbinom test results are reported in a sortable DataTable. You can search for specific samples to check their results, as well as export the table into CSV, excel spreadsheet, and PDF formats.

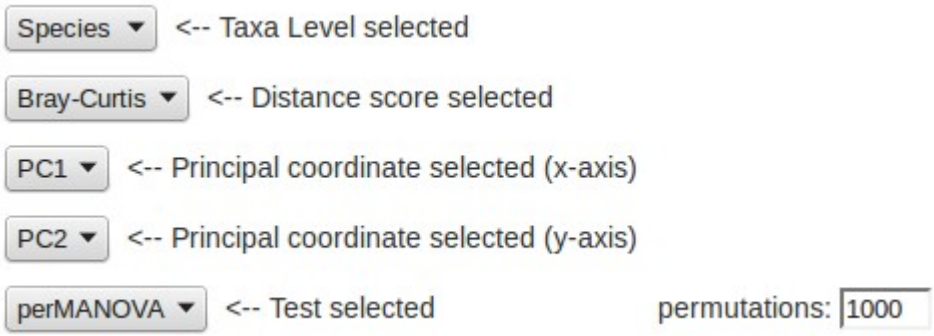
nbinomTest Results:

Show 10 entries Search: Copy CSV Excel PDF Print

	Comparison	Taxa ID	Taxa Name	baseMean	baseMeanA	baseMeanB	log2FoldChange
0	Pop100 vs Pop60	00d167b1ca624880ad60f9ca3781cdb1	Species12	168.370671	0.000000	1.797178	-4.078664e+00
0	Pop100 vs Pop80	00d167b1ca624880ad60f9ca3781cdb1	Species12	168.370671	0.000000	0.234936	-1.521171e+00
0	Pop100 vs Pop20	00d167b1ca624880ad60f9ca3781cdb1	Species12	168.370671	0.000000	77.532797	-9.639622e+00
0	Pop100 vs	00d167b1ca624880ad60f9ca3781cdb1	Species12	168.370671	0.000000	10.348042	-6.729533e+00

8.3 PCoA (principal coordinates analysis)

The PCoA analysis page (http://127.0.0.1:8000/myPhyloDB/PCoA/) layout and operation is similar to the Diff Abund page except for the addition of several drop-down menus.



Meta variables: The selection of meta variables is similar to the ANOVA/Regr page.

Taxa level: Select the taxonomic level (e.g., Phyla, Class, Order, Family, Genus, or Species) you would like to use for analysis.

Distance score: Select the distance score you would like to use for analysis. All scores are calculated using the vegan package in R, except for MorisitaHorn (custom python script based on the calculator in Mothur) and wOdum. The wOdum score can be used to down-weight either rare ($\alpha > 1$) or abundant ($\alpha < 1$) taxa, as discussed here (Manter and Bakker. 2015. BioInformatics. http://dx.doi.org/10.1093/bioinformatics/btv394). When $\alpha = 1$, wOdum is equivalent to Bray-Curtis.

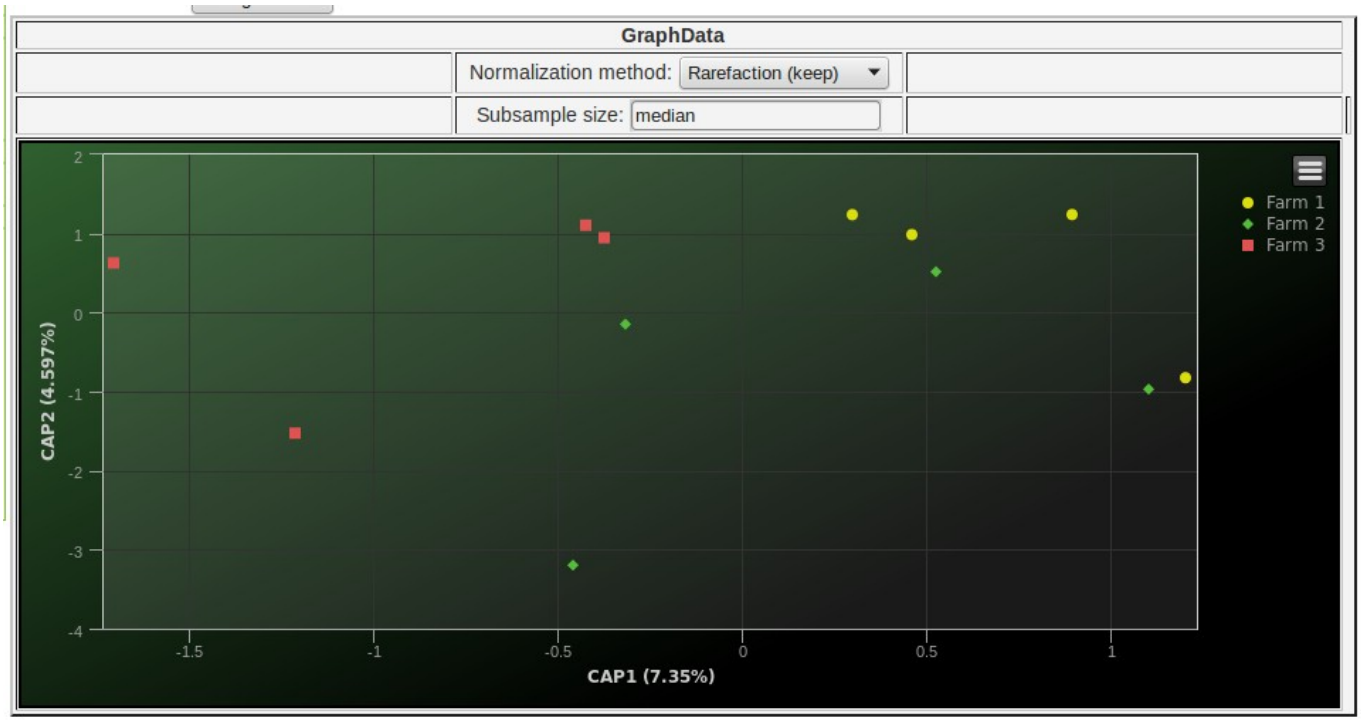
Principal coordinate axis selected (x-axis): This is the axis selected as the x-axis in the displayed graph.

Principal coordinate selected (y-axis):This is the axis selected as the y-axis in the displayed graph.

Test selected: Select whether you would like to perform either an perAMOVA (adonis) or betaDisper analysis of the selected data using the embedded R vegan package.

Graph:

PCoA analysis will produce a two-dimensional ordination plot. Each treatment combination will be displayed with a different symbol. If more than one meta variable was chose, a unique symbol will be used for each treatment combination.



Test Results:

At the top of “Test Results” section is a summary of the data normalization step, which includes the number of samples normalized (or removed) and the number of reads used for rarefaction. This section also displays the perAMOVA or betaDisper test results and the Eigenvalues and proportion of the variance explained for each PCoA axis. All analyses are conducted using the R vegan package.

Test Results:

```
Data Normalization:
All 12 selected samples were included in the analysis...
Data were rarefied to 1462 sequence reads...
================================================


Taxa level: Species
Independent Variable: geo_loc_farm

Distance score: Bray-Curtis
================================================
perMANOVA results:
Permutation: free
Number of permutations: 999

Terms added sequentially (first to last)

             Df SumsOfSqs  MeanSqs F.Model      R2 Pr(>F)
geo_loc_farm  2   0.15032 0.075162 0.74961 0.14279  0.769
Residuals     9   0.90241 0.100268         0.85721
```

Principal Coordinates and Distance Scores:

Also displayed are DataTables of the calculated “Principal Coordinates” and “Distance Scores” in matrix form. These tables can be sorted based on the column of your choice. You can also search for specific samples via id, name, treatment type, etc. As with all tables of this type in myPhyloDB, you can export the table to CSV, Excel, and PDF formats (or just print it directly).

Quantitative variables are handled in a similar manner and will produce a scatter plot between the selected variable (y-axis) and the chosen principal coordinate axis (x-axis). However, to avoid unit conflicts only one meta variable may be analyzed at any time; and instead of a perMANOVA or betaDisper analysis a simple linear regression analysis is performed.

9. Normalization

myPhyloDB provides several options for normalizing your sequence data to a common sampling depth: none, rarefaction (remove), rarefaction (keep), proportion, and DESeq2. A brief description of each procedure follows.

None – no normalization

Rarefaction (remove)

This normalization procedure performs a typical sub-sampling without replacement to the desired subsample size as implemented in Mothur and QIIME. Any sample, with fewer reads than the desired setting will be removed from the analysis. In the text box provided you can enter “min”, “median”, “max”, or any integer desired.

Rarefaction (keep)

This normalization procedure also performs a sub-sampling without replacement to the desired subsample size; however, it will keep all selected samples in the analysis regardless of the initial sample size. Aguirre de Cárcer et al. (Appl Environ Microbiol 2011 77:8795-8798) suggest that sub-sampling to the median number of sequence reads in a dataset can reduce variability and improve analysis. However, for samples with coverage below the subsampling threshold, no normalization procedure was proposed. In order to maintain sampling depths across all samples, myPhyloDB applies a small probability to undetected taxa (i.e., zeros) using a modified additive (Laplace) smoothing technique with $\lambda = 0.1$. The purpose of this small probability is to account for the uncertainty associated with not knowing whether the missing taxa were truly not present, or present but below the detection level, in the observed data. The Lidstone approximated probabilites are then sampled to a user-defined sample size to generate a new taxonomic profile for each sample. In the text box provide you can enter “min”, median, max, or any integer for your desired sample size.

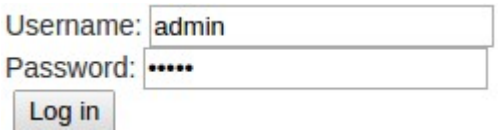
Proportion – all abundances are divided by the total number of sequence reads for that sample

DESeq2 – A discussion can be found here.

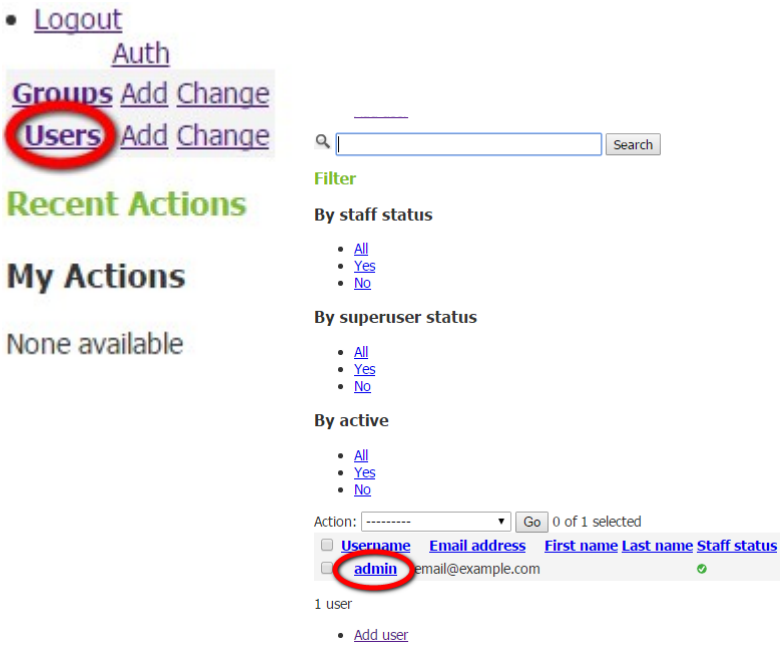
10. Admin

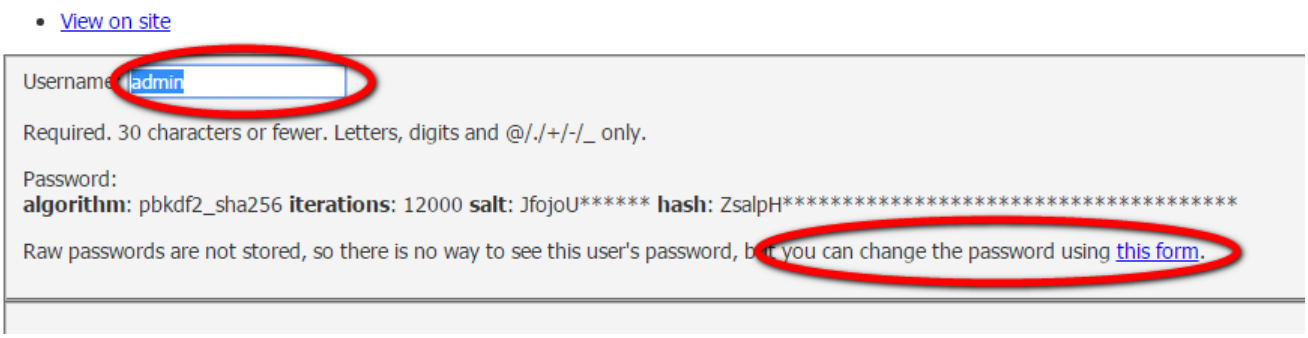
You can access the administrative pages at “127.0.0.1:8000/myPhyloDB/admin”, where you can change the “superuser” or administrator username and password or add/remove authorized users. The default superuser for myPhyloDB is as follows:

username: admin
password: admin
email: admin@example.com

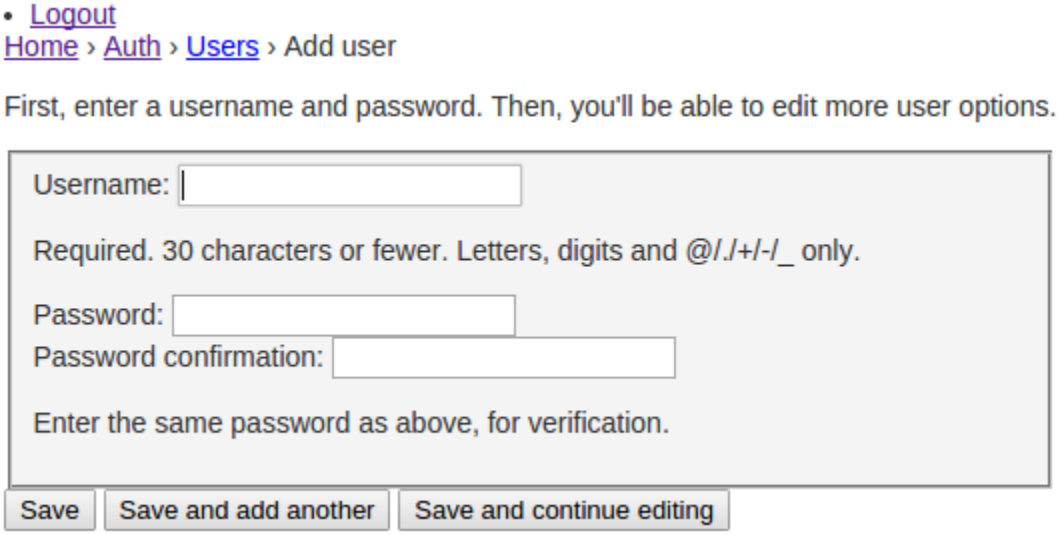


It is highly recommended that you change the default administrative username and password. To change the username click on the 'Users' link in the 'Auth' table. In the table, at the bottom of the next page click on the 'admin' username and change the username on the next page and press the 'Save' button at the bottom of the page. To change the password, click on the 'Change password' at the top-right of the page.





To add new authorized users click on the 'Add' link in the 'Auth' table. Add the desired username and password and press the 'Save' button at the bottom of the page. This user will now have access to the upload page and can add/remove projects from the myPhyloDB database.



There is also support for group permissions from the admin page, although currently these are fairly useless for anything but configuring another admin account.