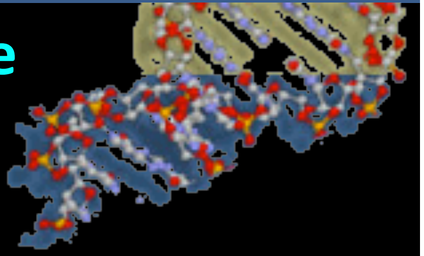
VAMPS

Visualization and Analysis of Microbial Population Structures

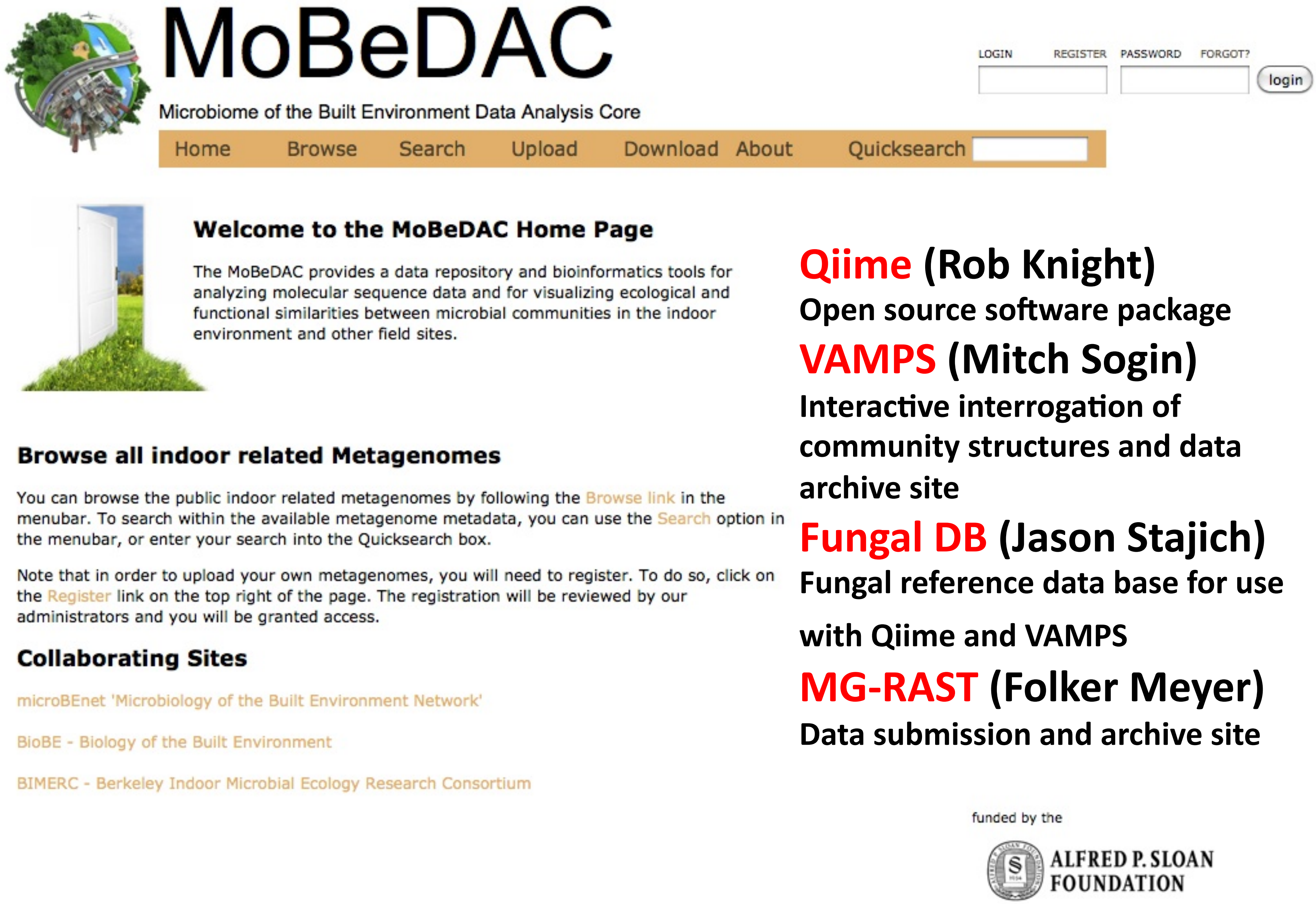


The role of VAMPs in the MoBEDAC initiative

The 13th Workshop of the Genomic Standards Consortium
Shenzhen, China,
March 7, 2012

Mitchell L. Sogin (sogin@mbl.edu)
Andy Voorhis (avoorhis@mbl.edu)
John Hufnagle (jhufnagle@mbl.edu)
Anna Shipunova (ashipunova@mbl.edu)
Susan Huse (shuse@mbl.edu)

http://mobedac.org/



MoBeDAC

Microbiome of the Built Environment Data Analysis Core

Welcome to the MoBeDAC Home Page

The MoBeDAC provides a data repository and bioinformatics tools for analyzing molecular sequence data and for visualizing ecological and functional similarities between microbial communities in the indoor environment and other field sites.

Browse all indoor related Metagenomes

You can browse the public indoor related metagenomes by following the Browse link in the menubar. To search within the available metagenome metadata, you can use the Search option in the menubar, or enter your search into the Quicksearch box.

Note that in order to upload your own metagenomes, you will need to register. To do so, click on the Register link on the top right of the page. The registration will be reviewed by our administrators and you will be granted access.

Collaborating Sites

microBEnet 'Microbiology of the Built Environment Network'

BioBE - Biology of the Built Environment

BIMERC - Berkeley Indoor Microbial Ecology Research Consortium

funded by the

ALFRED P. SLOAN FOUNDATION

Qiime (Rob Knight)

Open source software package

VAMPS (Mitch Sogin)

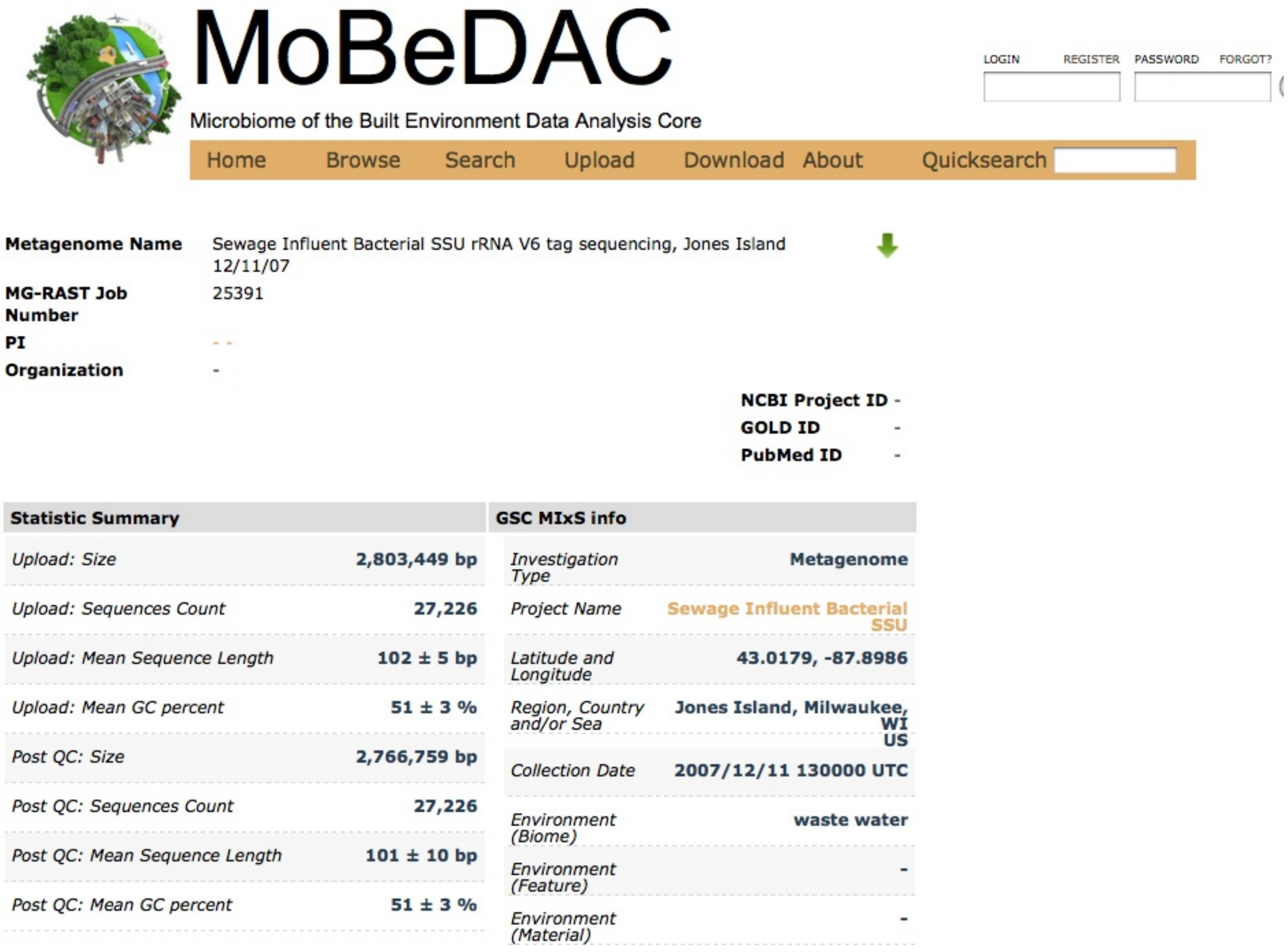
Interactive interrogation of community structures and data archive site

Fungal DB (Jason Stajich)

Fungal reference data base for use with Qiime and VAMPS

MG-RAST (Folker Meyer)

Data submission and archive site

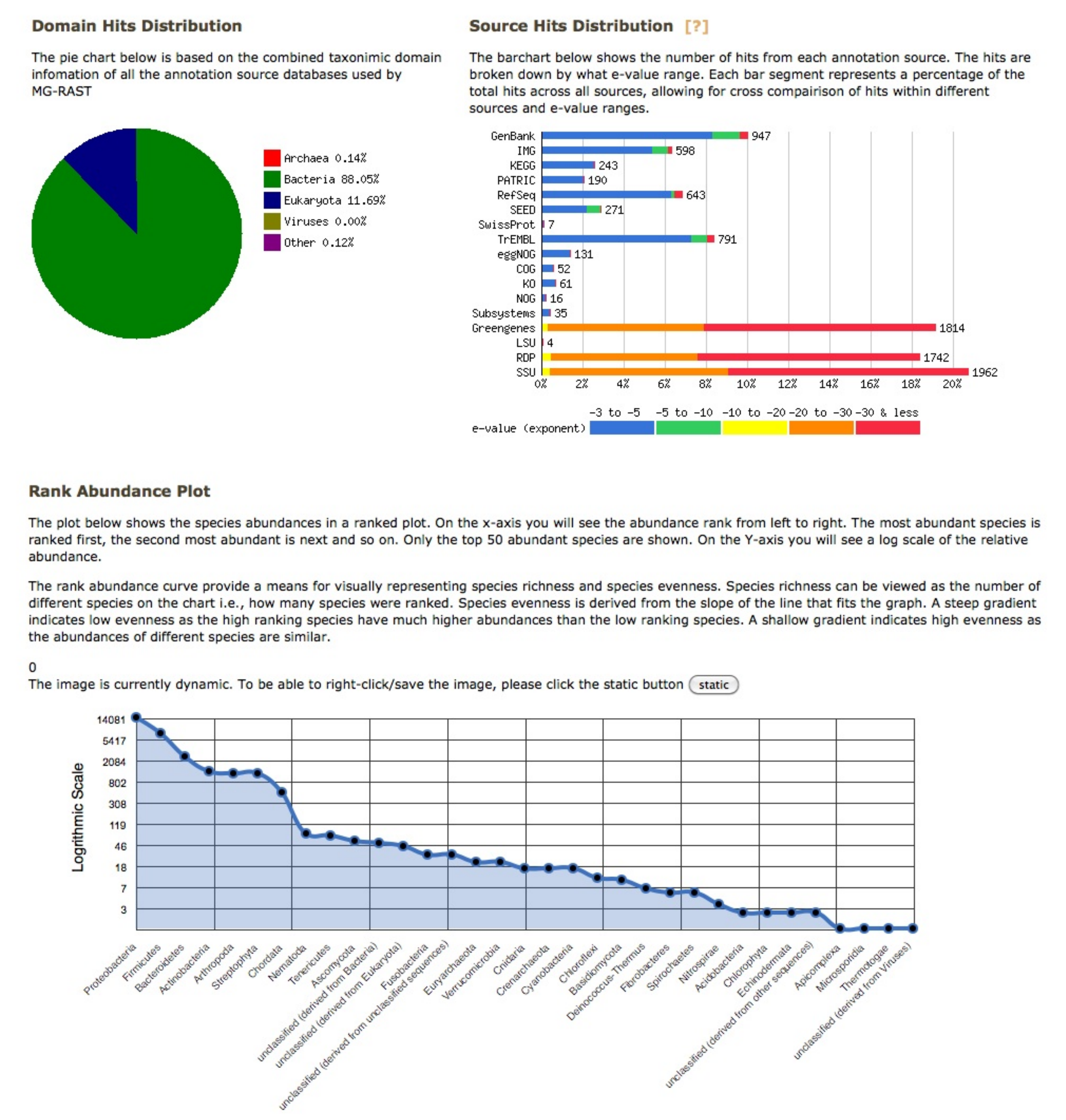


Metagenome Name	Sewage Influent Bacterial SSU rRNA V6 tag sequencing, Jones Island 12/11/07
MG-RAST Job Number	25391
PI	- -
Organization	-

NCBI Project ID	-
GOLD ID	-
PubMed ID	-

Statistic Summary	
Upload: Size	**2,803,449 bp**
Upload: Sequences Count	**27,226**
Upload: Mean Sequence Length	**102 ± 5 bp**
Upload: Mean GC percent	**51 ± 3 %**
Post QC: Size	**2,766,759 bp**
Post QC: Sequences Count	**27,226**
Post QC: Mean Sequence Length	**101 ± 10 bp**
Post QC: Mean GC percent	**51 ± 3 %**

GSC MIxS info	
Investigation Type	**Metagenome**
Project Name	**Sewage Influent Bacterial SSU**
Latitude and Longitude	**43.0179, -87.8986**
Region, Country and/or Sea	**Jones Island, Milwaukee, WI US**
Collection Date	**2007/12/11 130000 UTC**
Environment (Biome)	**waste water**
Environment (Feature)	**-**
Environment (Material)	**-**



Domain Hits Distribution

The pie chart below is based on the combined taxonimic domain infomation of all the annotation source databases used by MG-RAST

Source Hits Distribution [?]

The barchart below shows the number of hits from each annotation source. The hits are broken down by what e-value range. Each bar segment represents a percentage of the total hits across all sources, allowing for cross compairison of hits within different sources and e-value ranges.

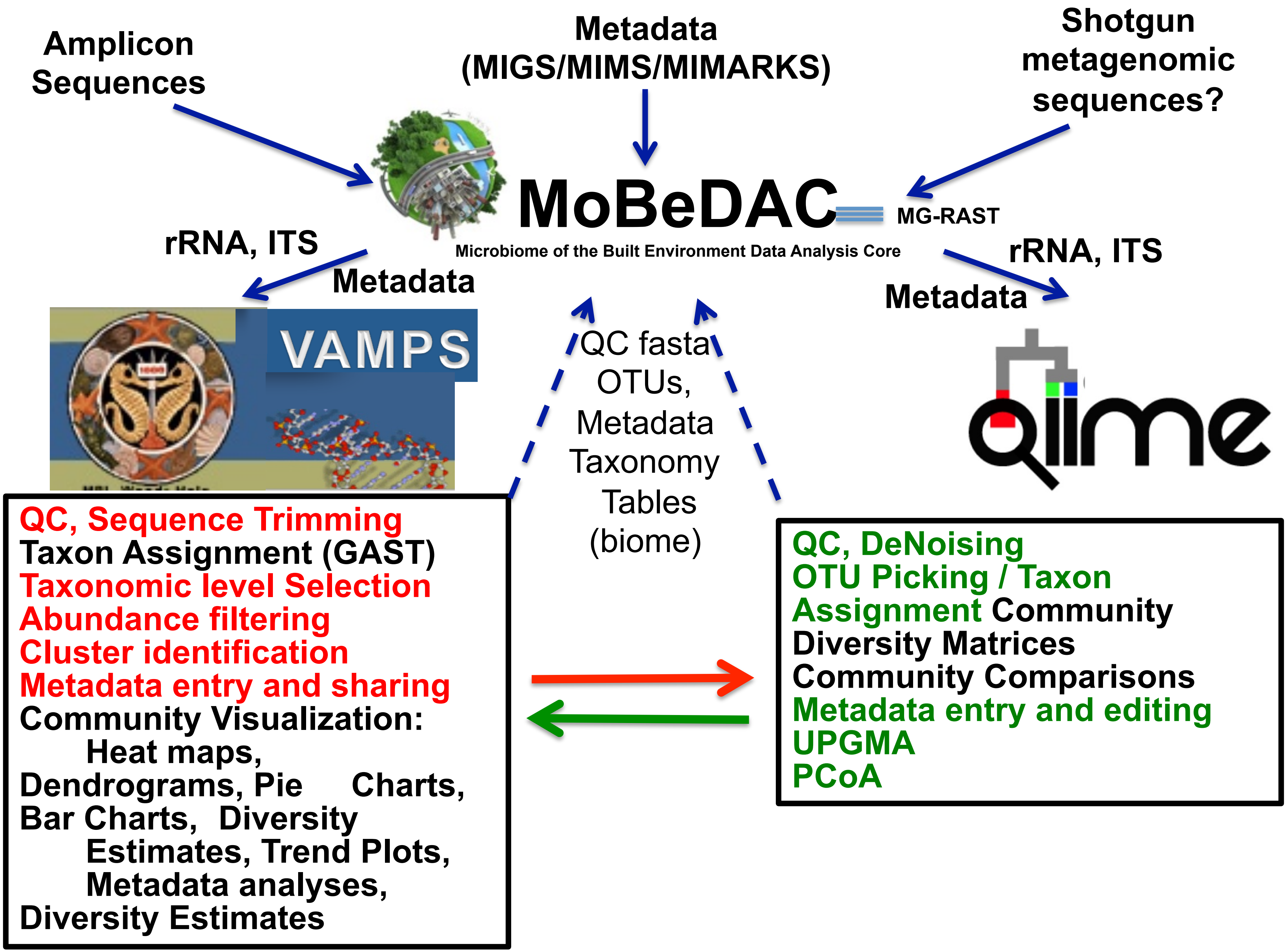
Rank Abundance Plot

The plot below shows the species abundances in a ranked plot. On the x-axis you will see the abundance rank from left to right. The most abundant species is ranked first, the second most abundant is next and so on. Only the top 50 abundant species are shown. On the Y-axis you will see a log scale of the relative abundance.

The rank abundance curve provide a means for visually representing species richness and species evenness. Species richness can be viewed as the number of different species on the chart i.e., how many species were ranked. Species evenness is derived from the slope of the line that fits the graph. A steep gradient indicates low evenness as the high ranking species have much higher abundances than the low ranking species. A shallow gradient indicates high evenness as the abundances of different species are similar.

0

The image is currently dynamic. To be able to right-click/save the image, please click the static button



Amplicon Sequences

Metadata (MIGS/MIMS/MIMARKS)

Shotgun metagenomic sequences?

MoBeDAC

MG-RAST

Microbiome of the Built Environment Data Analysis Core

rRNA, ITS

Metadata

VAMPS

rRNA, ITS

Metadata

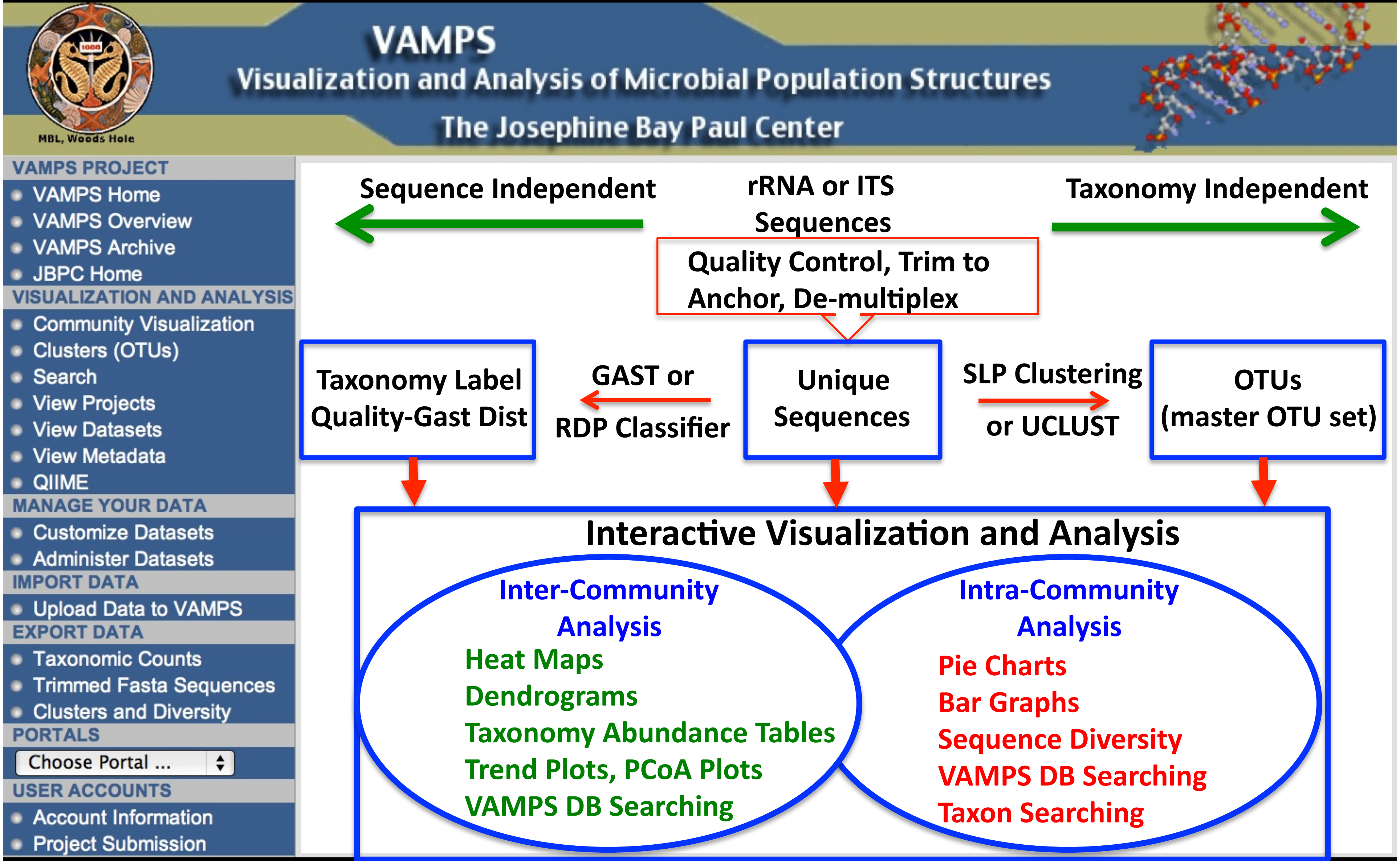
qiime

QC fasta
OTUs,
Metadata
Taxonomy
Tables
(biome)

QC, Sequence Trimming
Taxon Assignment (GAST)
Taxonomic level Selection
Abundance filtering
Cluster identification
Metadata entry and sharing
Community Visualization: Heat maps, Dendrograms, Pie Charts, Bar Charts, Diversity Estimates, Trend Plots, Metadata analyses, Diversity Estimates

QC, DeNoising
OTU Picking / Taxon Assignment
Community Diversity Matrices
Community Comparisons
Metadata entry and editing
UPGMA
PCoA

753 users/523 projects/6546 samples/323,000,000 reads **http://vamps.mbl.edu**



VAMPS

Visualization and Analysis of Microbial Population Structures

The Josephine Bay Paul Center

Sequence Independent ← rRNA or ITS Sequences → Taxonomy Independent

Quality Control, Trim to Anchor, De-multiplex

Unique Sequences

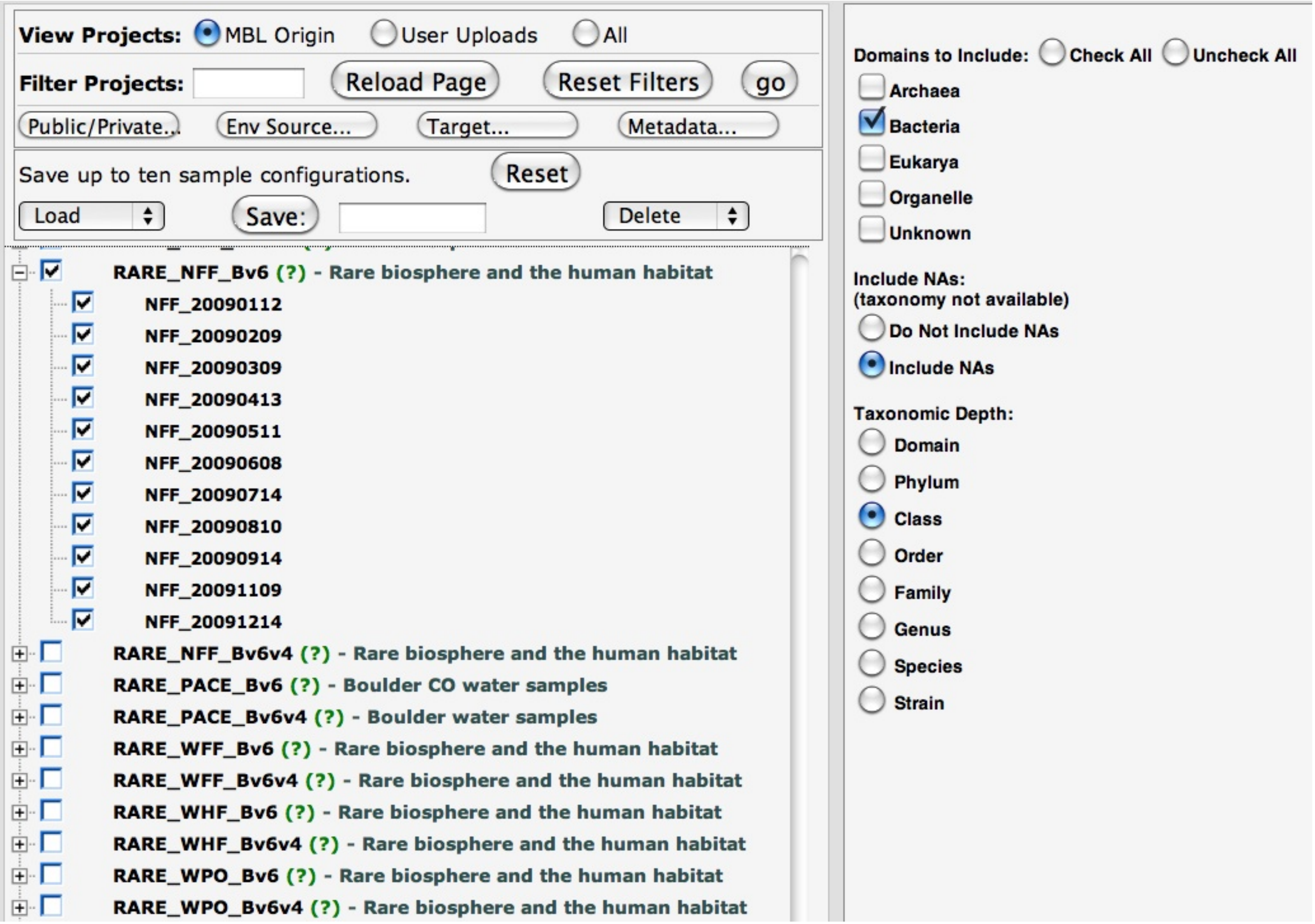
GAST or RDP Classifier → Taxonomy Label Quality-Gast Dist

SLP Clustering or UCLUST → OTUs (master OTU set)

Interactive Visualization and Analysis

Inter-Community Analysis
- Heat Maps
- Dendrograms
- Taxonomy Abundance Tables
- Trend Plots, PCoA Plots
- VAMPS DB Searching

Intra-Community Analysis
- Pie Charts
- Bar Graphs
- Sequence Diversity
- VAMPS DB Searching
- Taxon Searching



View Projects: ● MBL Origin ○ User Uploads ○ All

Filter Projects:

Save up to ten sample configurations.

- ☑ **RARE_NFF_Bv6** (?) - **Rare biosphere and the human habitat**
  - ☑ **NFF_20090112**
  - ☑ **NFF_20090209**
  - ☑ **NFF_20090309**
  - ☑ **NFF_20090413**
  - ☑ **NFF_20090511**
  - ☑ **NFF_20090608**
  - ☑ **NFF_20090714**
  - ☑ **NFF_20090810**
  - ☑ **NFF_20090914**
  - ☑ **NFF_20091109**
  - ☑ **NFF_20091214**
- ☐ **RARE_NFF_Bv6v4** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_PACE_Bv6** (?) - **Boulder CO water samples**
- ☐ **RARE_PACE_Bv6v4** (?) - **Boulder water samples**
- ☐ **RARE_WFF_Bv6** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_WFF_Bv6v4** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_WHF_Bv6** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_WHF_Bv6v4** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_WPO_Bv6** (?) - **Rare biosphere and the human habitat**
- ☐ **RARE_WPO_Bv6v4** (?) - **Rare biosphere and the human habitat**

Domains to Include: ○ **Check All** ○ **Uncheck All**

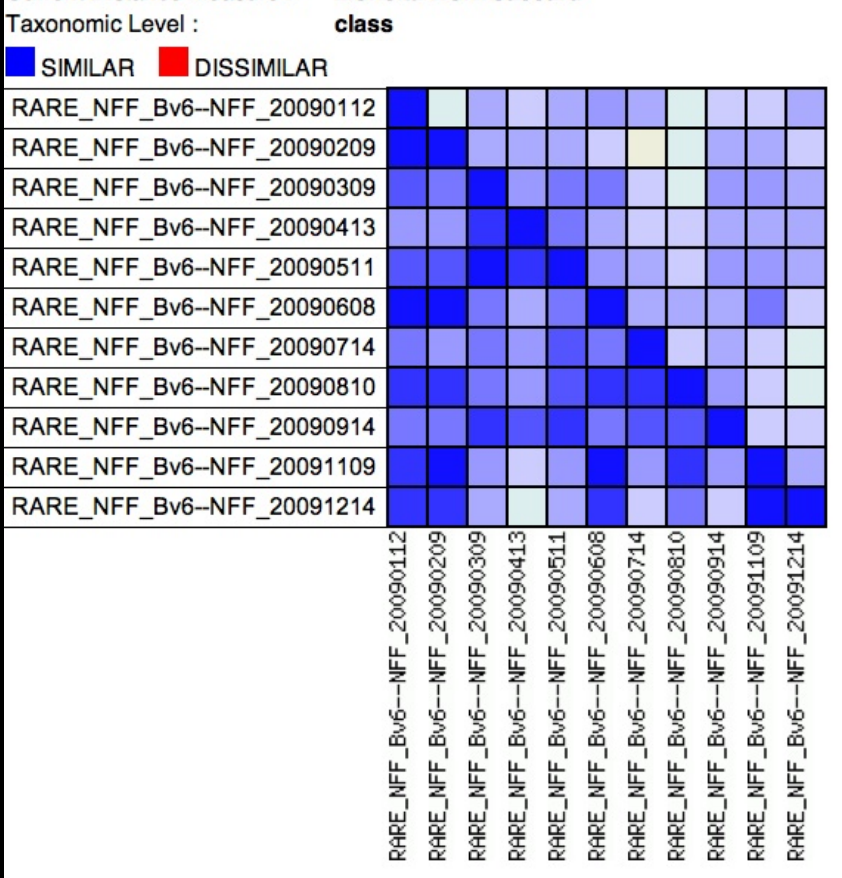
☐ **Archaea**
☑ **Bacteria**
☐ **Eukarya**
☐ **Organelle**
☐ **Unknown**

Include NAs:
(taxonomy not available)

○ **Do Not Include NAs**
● **Include NAs**

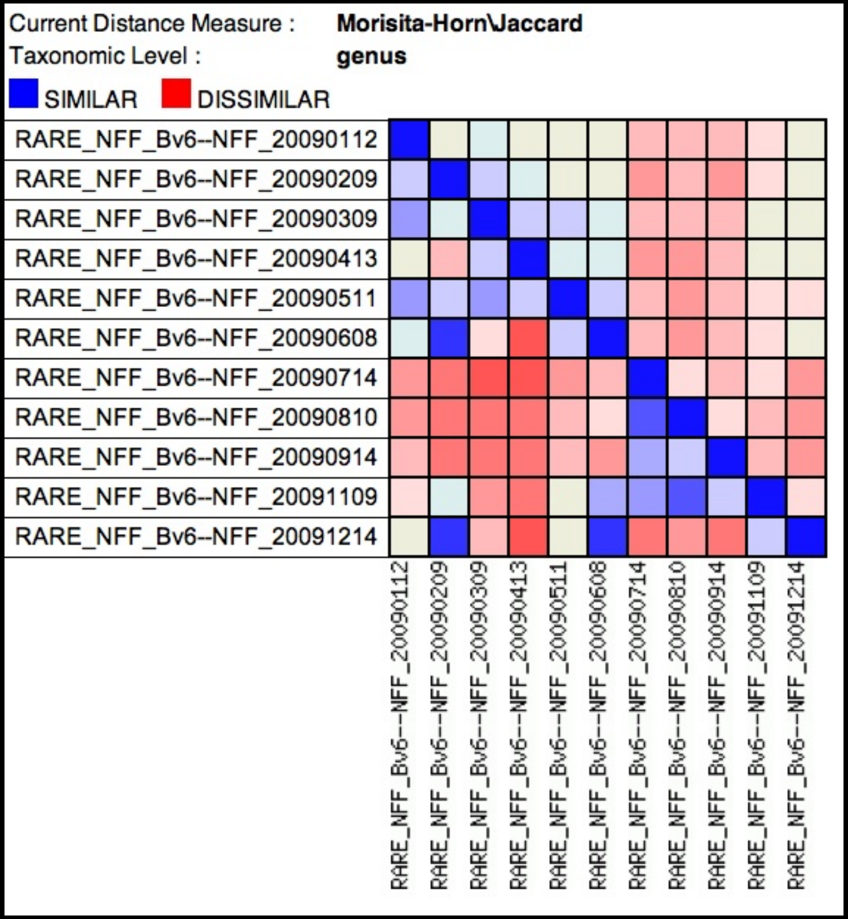
Taxonomic Depth:

○ **Domain**
○ **Phylum**
● **Class**
○ **Order**
○ **Family**
○ **Genus**
○ **Species**
○ **Strain**



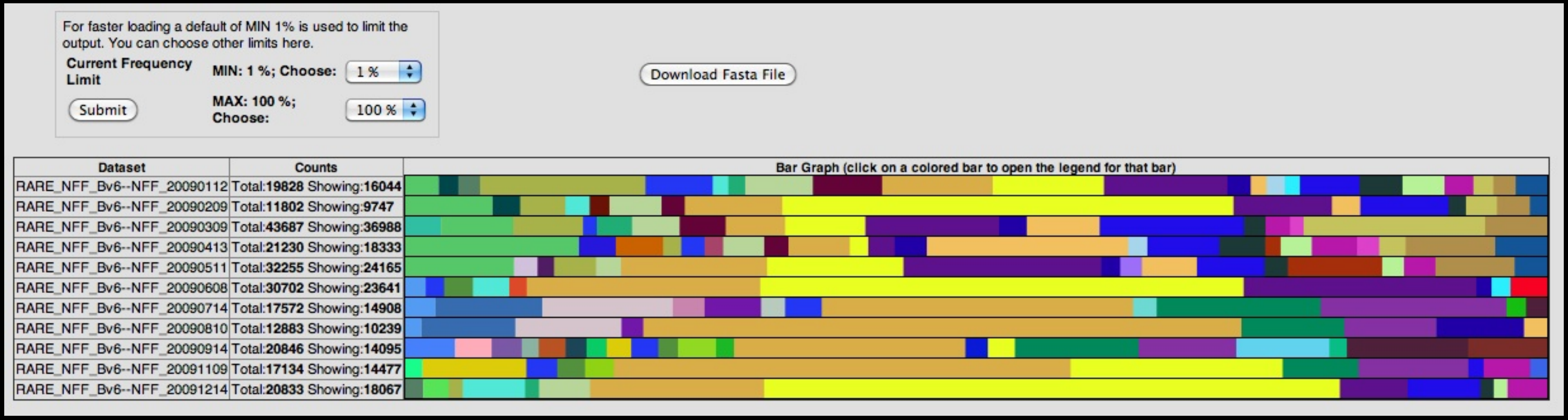
Current Distance Measure : **Morisita-Horn\Jaccard**
Taxonomic Level : **class**

🟦 SIMILAR 🟥 DISSIMILAR



Current Distance Measure : **Morisita-Horn\Jaccard**
Taxonomic Level : **genus**

🟦 SIMILAR 🟥 DISSIMILAR



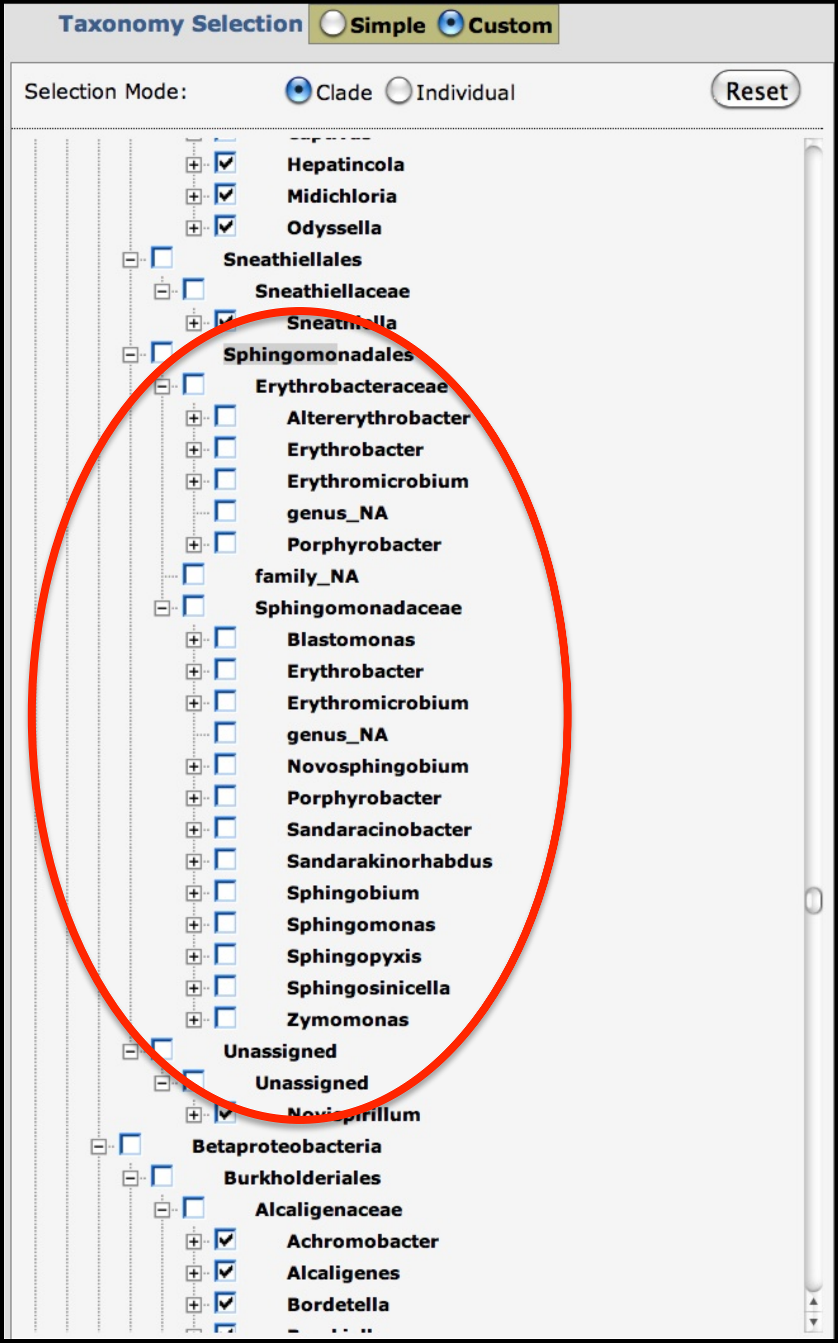
For faster loading a default of MIN 1% is used to limit the output. You can choose other limits here.

Current Frequency Limit

MIN: 1 %; Choose: 1 %

MAX: 100 %; Choose: 100 %

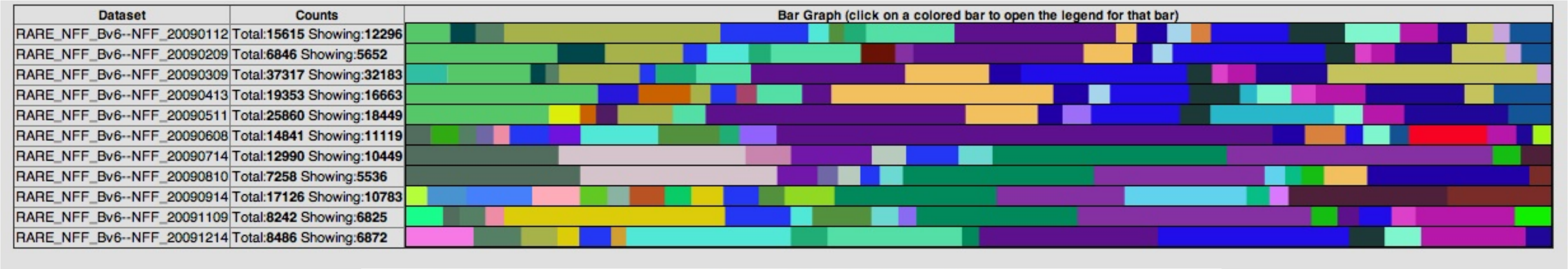
Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**19828** Showing:**16044**	
RARE_NFF_Bv6--NFF_20090209	Total:**11802** Showing:**9747**	
RARE_NFF_Bv6--NFF_20090309	Total:**43687** Showing:**36988**	
RARE_NFF_Bv6--NFF_20090413	Total:**21230** Showing:**18333**	
RARE_NFF_Bv6--NFF_20090511	Total:**32255** Showing:**24165**	
RARE_NFF_Bv6--NFF_20090608	Total:**30702** Showing:**23641**	
RARE_NFF_Bv6--NFF_20090714	Total:**17572** Showing:**14908**	
RARE_NFF_Bv6--NFF_20090810	Total:**12883** Showing:**10239**	
RARE_NFF_Bv6--NFF_20090914	Total:**20846** Showing:**14095**	
RARE_NFF_Bv6--NFF_20091109	Total:**17134** Showing:**14477**	
RARE_NFF_Bv6--NFF_20091214	Total:**20833** Showing:**18067**	



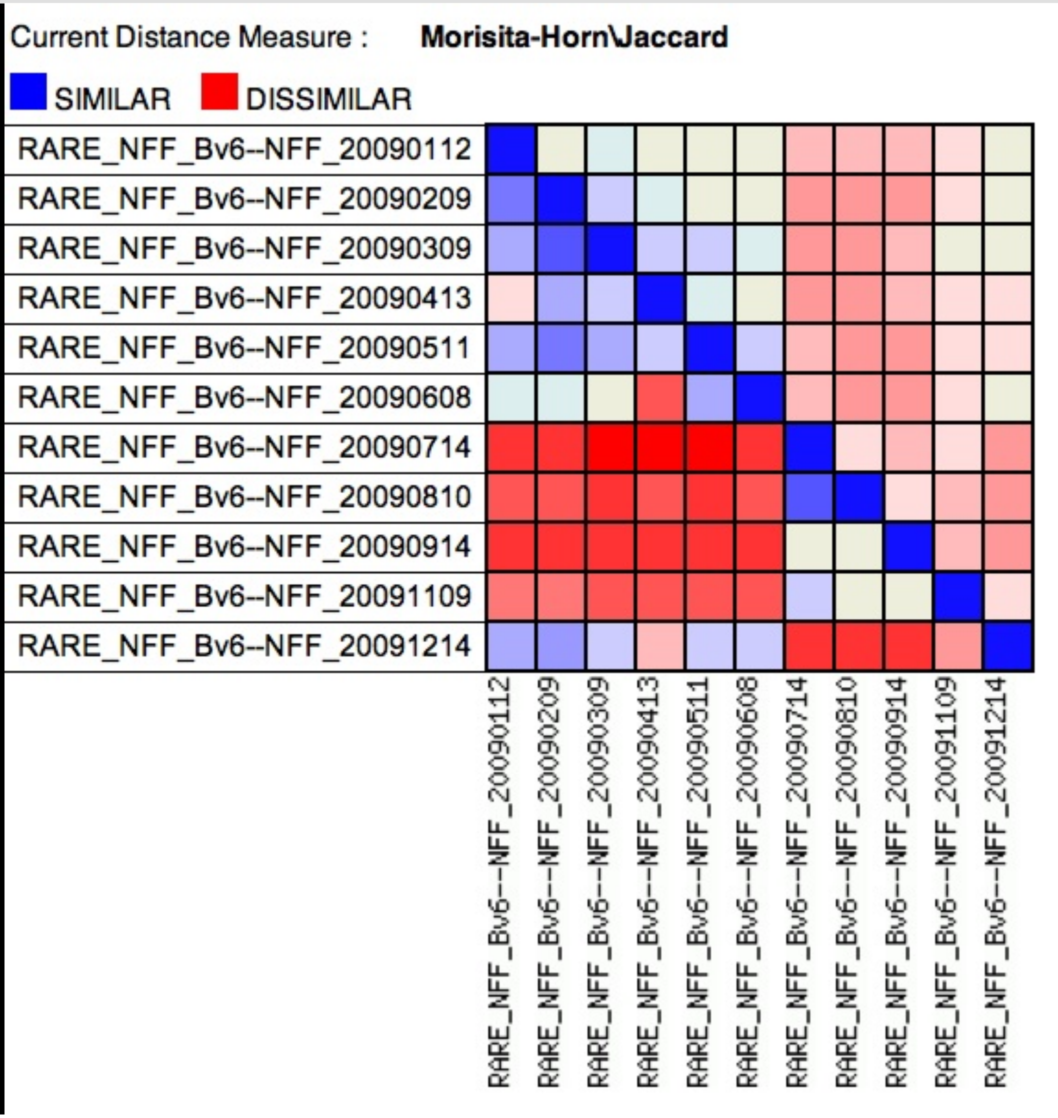
Taxonomy Selection ○ Simple ● Custom

Selection Mode: ● Clade ○ Individual Reset

- ☑ Hepatincola
- ☑ Midichloria
- ☑ Odyssella
- ☐ Sneathiellales
  - ☐ Sneathiellaceae
    - ☑ Sneathiella
- ☐ Sphingomonadales
  - ☐ Erythrobacteraceae
    - ☐ Altererythrobacter
    - ☐ Erythrobacter
    - ☐ Erythromicrobium
    - ☐ genus_NA
    - ☐ Porphyrobacter
  - ☐ family_NA
  - ☐ Sphingomonadaceae
    - ☐ Blastomonas
    - ☐ Erythrobacter
    - ☐ Erythromicrobium
    - ☐ genus_NA
    - ☐ Novosphingobium
    - ☐ Porphyrobacter
    - ☐ Sandaracinobacter
    - ☐ Sandarakinorhabdus
    - ☐ Sphingobium
    - ☐ Sphingomonas
    - ☐ Sphingopyxis
    - ☐ Sphingosinicella
    - ☐ Zymomonas
- ☐ Unassigned
  - ☐ Unassigned
    - ☑ Novispirillum
- ☐ Betaproteobacteria
  - ☐ Burkholderiales
    - ☐ Alcaligenaceae
      - ☑ Achromobacter
      - ☑ Alcaligenes
      - ☑ Bordetella

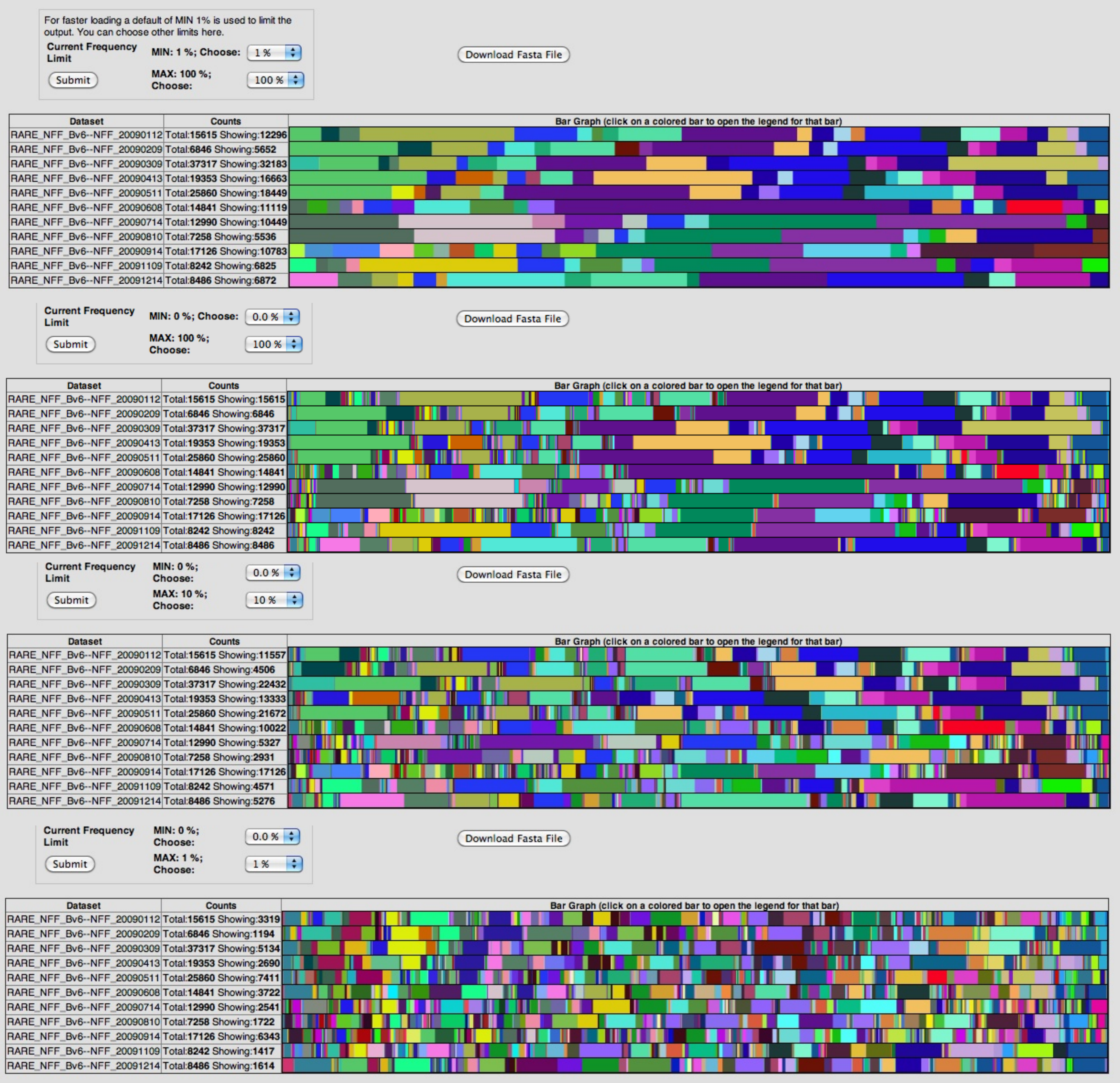


Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**15615** Showing:**12296**	
RARE_NFF_Bv6--NFF_20090209	Total:**6846** Showing:**5652**	
RARE_NFF_Bv6--NFF_20090309	Total:**37317** Showing:**32183**	
RARE_NFF_Bv6--NFF_20090413	Total:**19353** Showing:**16663**	
RARE_NFF_Bv6--NFF_20090511	Total:**25860** Showing:**18449**	
RARE_NFF_Bv6--NFF_20090608	Total:**14841** Showing:**11119**	
RARE_NFF_Bv6--NFF_20090714	Total:**12990** Showing:**10449**	
RARE_NFF_Bv6--NFF_20090810	Total:**7258** Showing:**5536**	
RARE_NFF_Bv6--NFF_20090914	Total:**17126** Showing:**10783**	
RARE_NFF_Bv6--NFF_20091109	Total:**8242** Showing:**6825**	
RARE_NFF_Bv6--NFF_20091214	Total:**8486** Showing:**6872**	



Current Distance Measure : **Morisita-Horn\Jaccard**

🟦 SIMILAR 🟥 DISSIMILAR



For faster loading a default of MIN 1% is used to limit the output. You can choose other limits here.

Current Frequency Limit
MIN: 1 %; Choose: 1 %
MAX: 100 %; Choose: 100 %

Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**15615** Showing:**12296**	
RARE_NFF_Bv6--NFF_20090209	Total:**6846** Showing:**5652**	
RARE_NFF_Bv6--NFF_20090309	Total:**37317** Showing:**32183**	
RARE_NFF_Bv6--NFF_20090413	Total:**19353** Showing:**16663**	
RARE_NFF_Bv6--NFF_20090511	Total:**25860** Showing:**18449**	
RARE_NFF_Bv6--NFF_20090608	Total:**14841** Showing:**11119**	
RARE_NFF_Bv6--NFF_20090714	Total:**12990** Showing:**10449**	
RARE_NFF_Bv6--NFF_20090810	Total:**7258** Showing:**5536**	
RARE_NFF_Bv6--NFF_20090914	Total:**17126** Showing:**10783**	
RARE_NFF_Bv6--NFF_20091109	Total:**8242** Showing:**6825**	
RARE_NFF_Bv6--NFF_20091214	Total:**8486** Showing:**6872**	

Current Frequency Limit
MIN: 0 %; Choose: 0.0 %
MAX: 100 %; Choose: 100 %

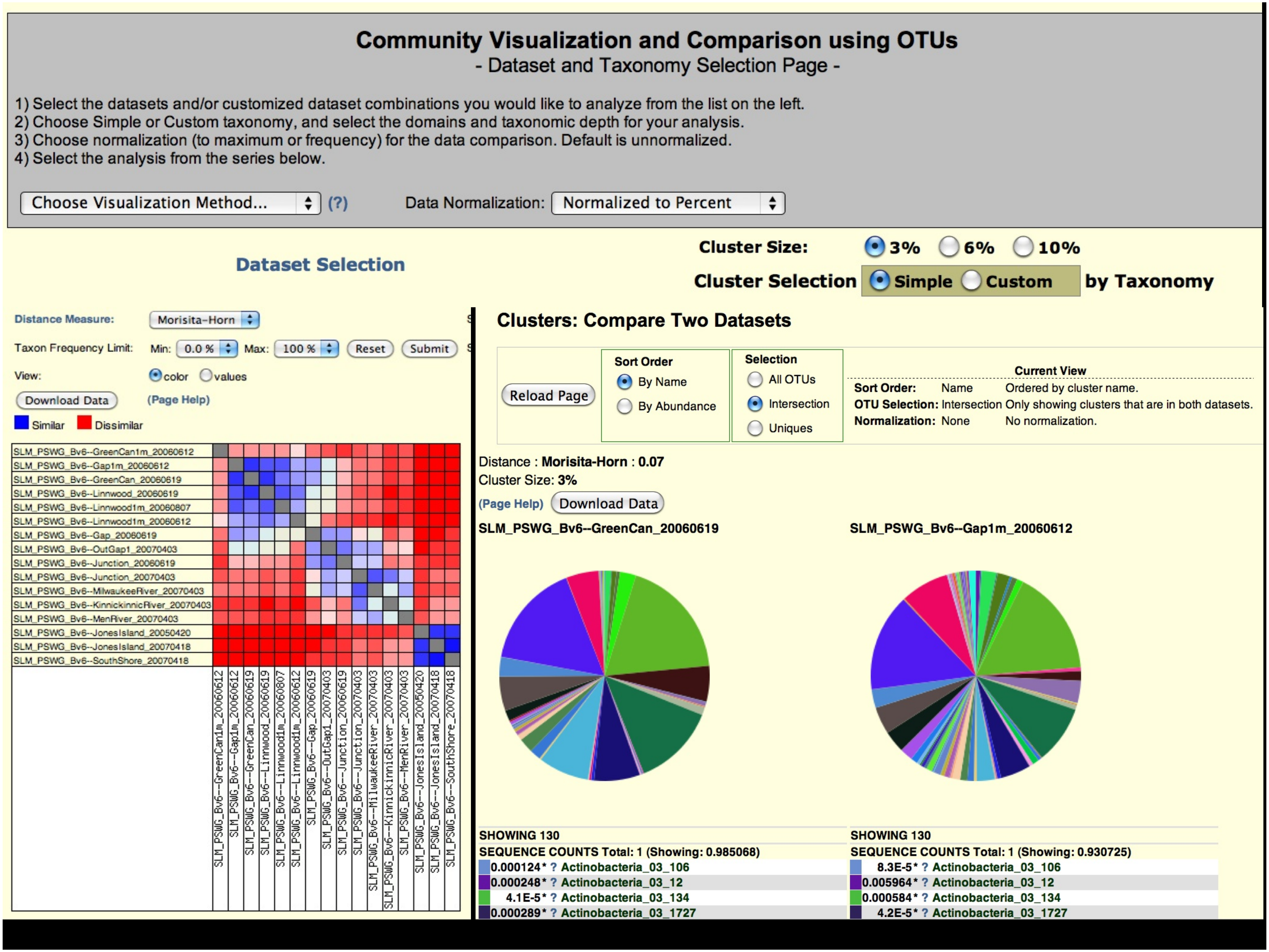
Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**15615** Showing:**15615**	
RARE_NFF_Bv6--NFF_20090209	Total:**6846** Showing:**6846**	
RARE_NFF_Bv6--NFF_20090309	Total:**37317** Showing:**37317**	
RARE_NFF_Bv6--NFF_20090413	Total:**19353** Showing:**19353**	
RARE_NFF_Bv6--NFF_20090511	Total:**25860** Showing:**25860**	
RARE_NFF_Bv6--NFF_20090608	Total:**14841** Showing:**14841**	
RARE_NFF_Bv6--NFF_20090714	Total:**12990** Showing:**12990**	
RARE_NFF_Bv6--NFF_20090810	Total:**7258** Showing:**7258**	
RARE_NFF_Bv6--NFF_20090914	Total:**17126** Showing:**17126**	
RARE_NFF_Bv6--NFF_20091109	Total:**8242** Showing:**8242**	
RARE_NFF_Bv6--NFF_20091214	Total:**8486** Showing:**8486**	

Current Frequency Limit
MIN: 0 %; Choose: 0.0 %
MAX: 10 %; Choose: 10 %

Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**15615** Showing:**11557**	
RARE_NFF_Bv6--NFF_20090209	Total:**6846** Showing:**4506**	
RARE_NFF_Bv6--NFF_20090309	Total:**37317** Showing:**22432**	
RARE_NFF_Bv6--NFF_20090413	Total:**19353** Showing:**13333**	
RARE_NFF_Bv6--NFF_20090511	Total:**25860** Showing:**21672**	
RARE_NFF_Bv6--NFF_20090608	Total:**14841** Showing:**10022**	
RARE_NFF_Bv6--NFF_20090714	Total:**12990** Showing:**5327**	
RARE_NFF_Bv6--NFF_20090810	Total:**7258** Showing:**2931**	
RARE_NFF_Bv6--NFF_20090914	Total:**17126** Showing:**17126**	
RARE_NFF_Bv6--NFF_20091109	Total:**8242** Showing:**4571**	
RARE_NFF_Bv6--NFF_20091214	Total:**8486** Showing:**5276**	

Current Frequency Limit
MIN: 0 %; Choose: 0.0 %
MAX: 1 %; Choose: 1 %

Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
RARE_NFF_Bv6--NFF_20090112	Total:**15615** Showing:**3319**	
RARE_NFF_Bv6--NFF_20090209	Total:**6846** Showing:**1194**	
RARE_NFF_Bv6--NFF_20090309	Total:**37317** Showing:**5134**	
RARE_NFF_Bv6--NFF_20090413	Total:**19353** Showing:**2690**	
RARE_NFF_Bv6--NFF_20090511	Total:**25860** Showing:**7411**	
RARE_NFF_Bv6--NFF_20090608	Total:**14841** Showing:**3722**	
RARE_NFF_Bv6--NFF_20090714	Total:**12990** Showing:**2541**	
RARE_NFF_Bv6--NFF_20090810	Total:**7258** Showing:**1722**	
RARE_NFF_Bv6--NFF_20090914	Total:**17126** Showing:**6343**	
RARE_NFF_Bv6--NFF_20091109	Total:**8242** Showing:**1417**	
RARE_NFF_Bv6--NFF_20091214	Total:**8486** Showing:**1614**	



Community Visualization and Comparison using OTUs

- Dataset and Taxonomy Selection Page -

1) Select the datasets and/or customized dataset combinations you would like to analyze from the list on the left.
2) Choose Simple or Custom taxonomy, and select the domains and taxonomic depth for your analysis.
3) Choose normalization (to maximum or frequency) for the data comparison. Default is unnormalized.
4) Select the analysis from the series below.

Data Normalization: Normalized to Percent

Cluster Size: ● 3% ○ 6% ○ 10%

Cluster Selection ● Simple ○ Custom **by Taxonomy**

Dataset Selection

Distance Measure: Morisita-Horn

Taxonomy Frequency Limit: Min: 0.0 % Max: 100 %

View: ● color ○ values

🟦 Similar 🟥 Dissimilar

SLM_PSWG_Bv6--GreenCan1m_20060612
SLM_PSWG_Bv6--Gap1m_20060612
SLM_PSWG_Bv6--GreenCan_20060619
SLM_PSWG_Bv6--Linnwood_20060619
SLM_PSWG_Bv6--Linnwood1m_20060807
SLM_PSWG_Bv6--Linnwood1m_20060612
SLM_PSWG_Bv6--Gap_20060619
SLM_PSWG_Bv6--OutGap1_20070403
SLM_PSWG_Bv6--Junction_20060619
SLM_PSWG_Bv6--Junction_20070403
SLM_PSWG_Bv6--MilwaukeeRiver_20070403
SLM_PSWG_Bv6--KinnickinnicRiver_20070403
SLM_PSWG_Bv6--MenRiver_20070403
SLM_PSWG_Bv6--JonesIsland_20050420
SLM_PSWG_Bv6--JonesIsland_20070418
SLM_PSWG_Bv6--SouthShore_20070418

Clusters: Compare Two Datasets

Sort Order
● By Name
○ By Abundance

Selection
○ All OTUs
● Intersection
○ Uniques

Current View

Sort Order:	Name	Ordered by cluster name.
OTU Selection:	Intersection	Only showing clusters that are in both datasets.
Normalization:	None	No normalization.

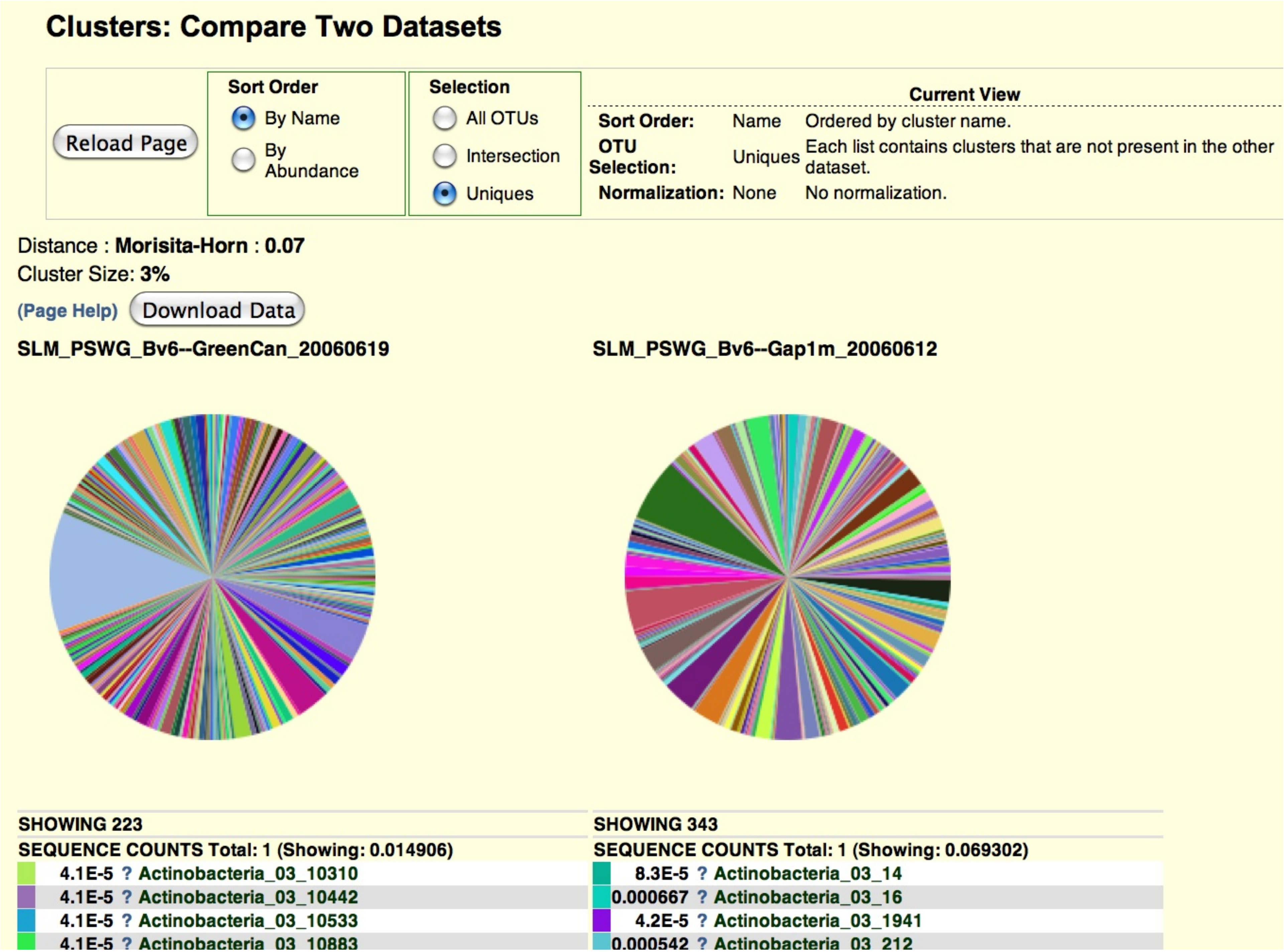
Distance : **Morisita-Horn** : **0.07**
Cluster Size: **3%**

SLM_PSWG_Bv6--GreenCan_20060619

SHOWING 130
SEQUENCE COUNTS Total: 1 (Showing: 0.985068)
0.000124* ? Actinobacteria_03_106
0.000248* ? Actinobacteria_03_12
4.1E-5* ? Actinobacteria_03_134
0.000289* ? Actinobacteria_03_1727

SLM_PSWG_Bv6--Gap1m_20060612

SHOWING 130
SEQUENCE COUNTS Total: 1 (Showing: 0.930725)
8.3E-5* ? Actinobacteria_03_106
0.005964* ? Actinobacteria_03_12
0.000584* ? Actinobacteria_03_134
4.2E-5* ? Actinobacteria_03_1727



Clusters: Compare Two Datasets

Sort Order
● By Name
○ By Abundance

Selection
○ All OTUs
○ Intersection
● Uniques

Current View

Sort Order:	Name	Ordered by cluster name.
OTU Selection:	Uniques	Each list contains clusters that are not present in the other dataset.
Normalization:	None	No normalization.

Distance : **Morisita-Horn** : **0.07**

Cluster Size: **3%**

SLM_PSWG_Bv6--GreenCan_20060619

SHOWING 223

SEQUENCE COUNTS Total: 1 (Showing: 0.014906)

🟩	4.1E-5	? Actinobacteria_03_10310
🟪	4.1E-5	? Actinobacteria_03_10442
🟦	4.1E-5	? Actinobacteria_03_10533
🟩	4.1E-5	? Actinobacteria_03_10883

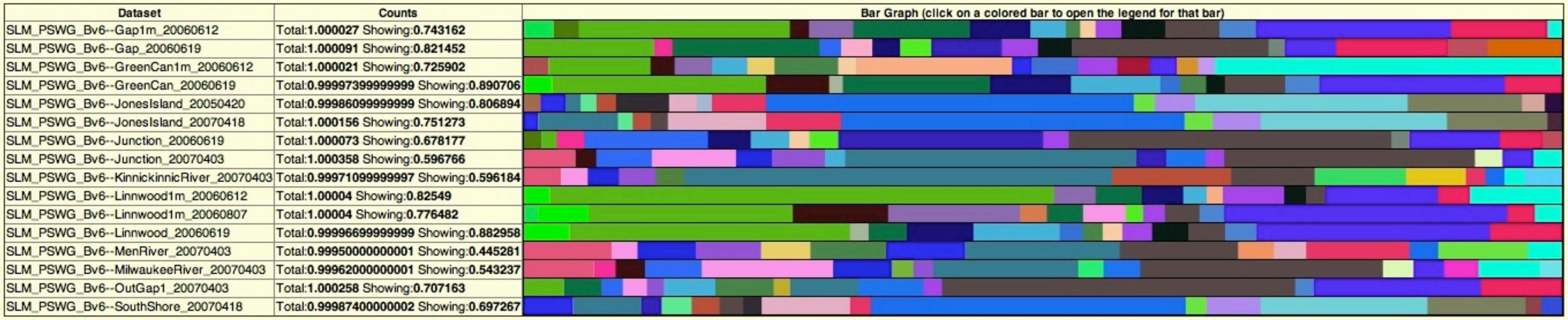
SLM_PSWG_Bv6--Gap1m_20060612

SHOWING 343

SEQUENCE COUNTS Total: 1 (Showing: 0.069302)

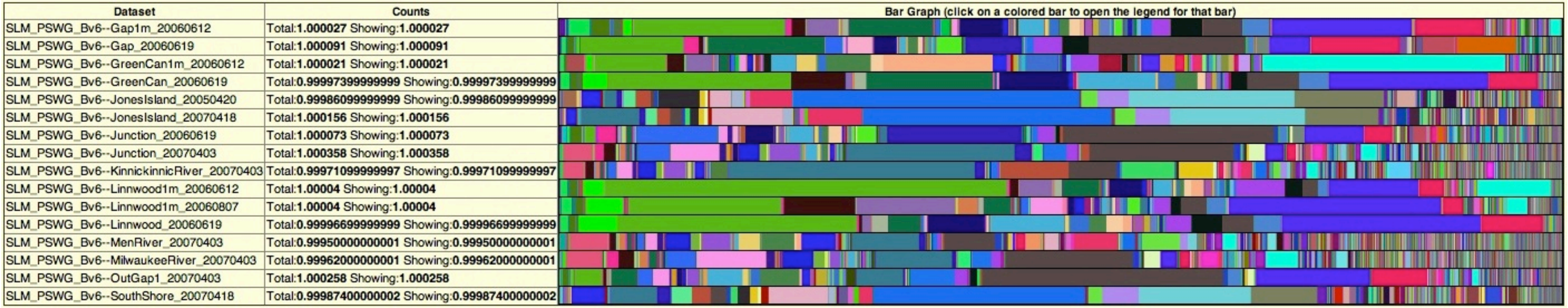
🟩	8.3E-5	? Actinobacteria_03_14
🟩	0.000667	? Actinobacteria_03_16
🟪	4.2E-5	? Actinobacteria_03_1941
🟦	0.000542	? Actinobacteria_03_212

Cluster Fraction Min= 1 % Max = 100 %



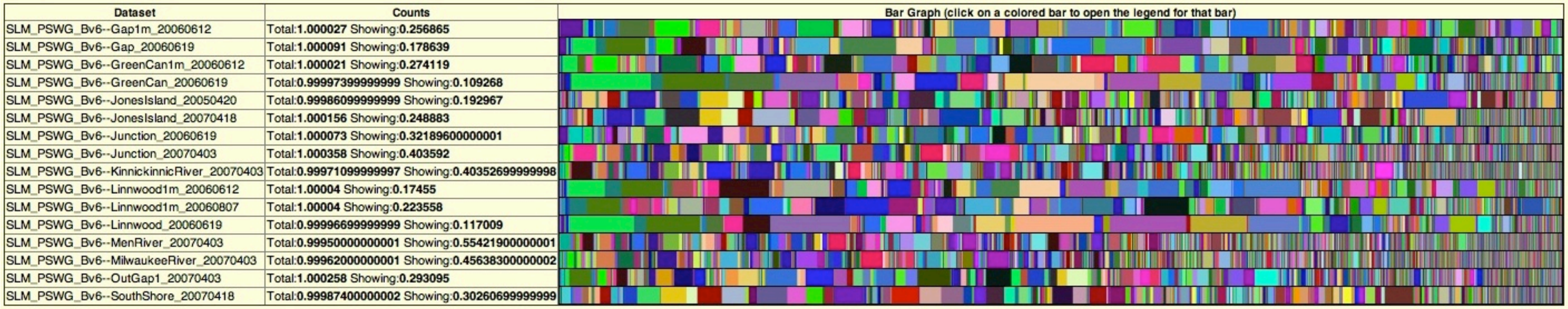
Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
SLM_PSWG_Bv6--Gap1m_20060612	Total:**1.000027** Showing:**0.743162**	
SLM_PSWG_Bv6--Gap_20060619	Total:**1.000091** Showing:**0.821452**	
SLM_PSWG_Bv6--GreenCan1m_20060612	Total:**1.000021** Showing:**0.725902**	
SLM_PSWG_Bv6--GreenCan_20060619	Total:**0.99997399999999** Showing:**0.890706**	
SLM_PSWG_Bv6--JonesIsland_20050420	Total:**0.99986099999999** Showing:**0.806894**	
SLM_PSWG_Bv6--JonesIsland_20070418	Total:**1.000156** Showing:**0.751273**	
SLM_PSWG_Bv6--Junction_20060619	Total:**1.000073** Showing:**0.678177**	
SLM_PSWG_Bv6--Junction_20070403	Total:**1.000358** Showing:**0.596766**	
SLM_PSWG_Bv6--KinnickinnicRiver_20070403	Total:**0.99971099999997** Showing:**0.596184**	
SLM_PSWG_Bv6--Linnwood1m_20060612	Total:**1.00004** Showing:**0.82549**	
SLM_PSWG_Bv6--Linnwood1m_20060807	Total:**1.00004** Showing:**0.776482**	
SLM_PSWG_Bv6--Linnwood_20060619	Total:**0.99996699999999** Showing:**0.882958**	
SLM_PSWG_Bv6--MenRiver_20070403	Total:**0.99950000000001** Showing:**0.445281**	
SLM_PSWG_Bv6--MilwaukeeRiver_20070403	Total:**0.99962000000001** Showing:**0.543237**	
SLM_PSWG_Bv6--OutGap1_20070403	Total:**1.000258** Showing:**0.707163**	
SLM_PSWG_Bv6--SouthShore_20070418	Total:**0.99987400000002** Showing:**0.697267**	

Cluster Fraction Min= 0 % Max = 100 %

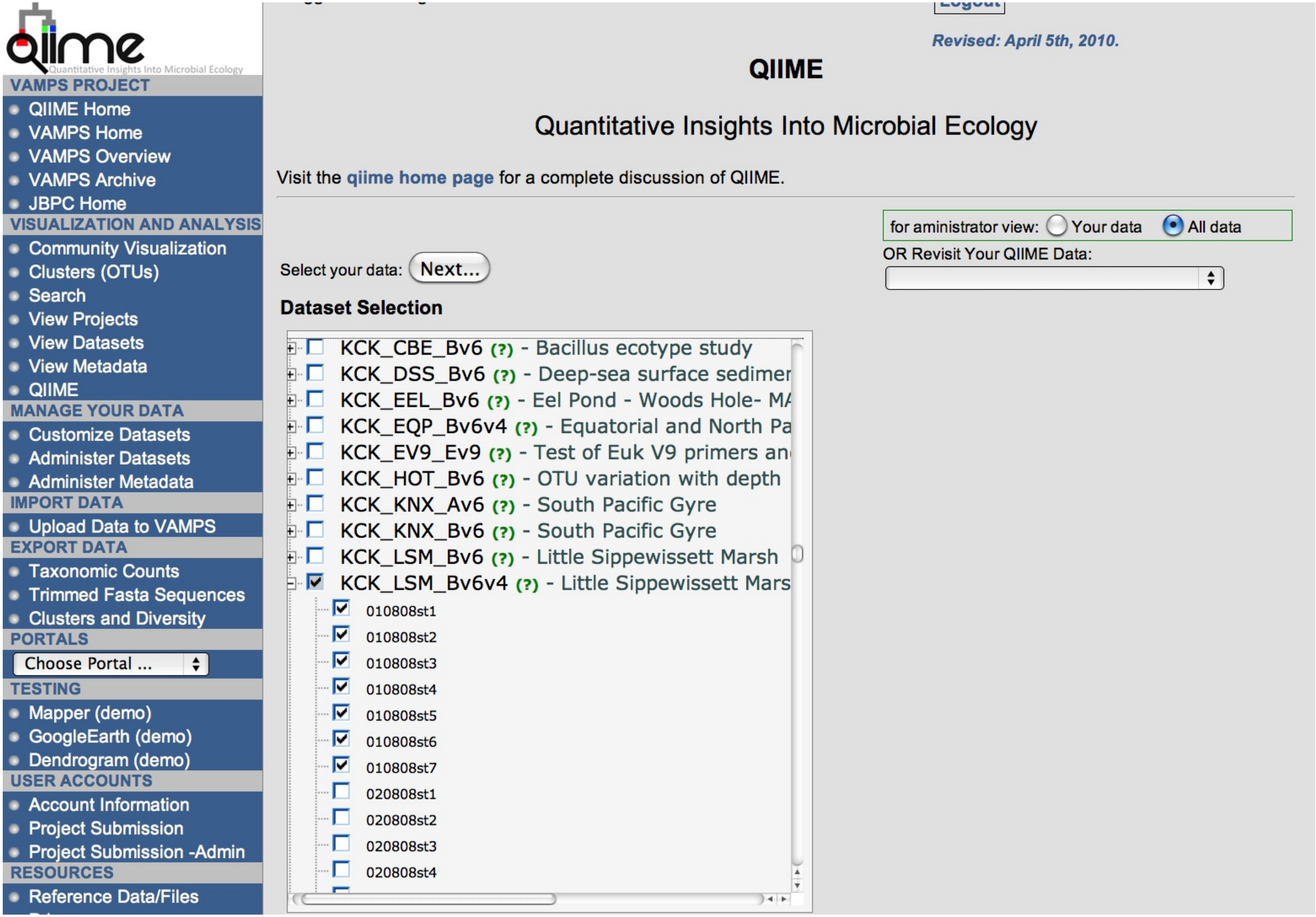


Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
SLM_PSWG_Bv6--Gap1m_20060612	Total:**1.000027** Showing:**1.000027**	
SLM_PSWG_Bv6--Gap_20060619	Total:**1.000091** Showing:**1.000091**	
SLM_PSWG_Bv6--GreenCan1m_20060612	Total:**1.000021** Showing:**1.000021**	
SLM_PSWG_Bv6--GreenCan_20060619	Total:**0.99997399999999** Showing:**0.99997399999999**	
SLM_PSWG_Bv6--JonesIsland_20050420	Total:**0.99986099999999** Showing:**0.99986099999999**	
SLM_PSWG_Bv6--JonesIsland_20070418	Total:**1.000156** Showing:**1.000156**	
SLM_PSWG_Bv6--Junction_20060619	Total:**1.000073** Showing:**1.000073**	
SLM_PSWG_Bv6--Junction_20070403	Total:**1.000358** Showing:**1.000358**	
SLM_PSWG_Bv6--KinnickinnicRiver_20070403	Total:**0.99971099999997** Showing:**0.99971099999997**	
SLM_PSWG_Bv6--Linnwood1m_20060612	Total:**1.00004** Showing:**1.00004**	
SLM_PSWG_Bv6--Linnwood1m_20060807	Total:**1.00004** Showing:**1.00004**	
SLM_PSWG_Bv6--Linnwood_20060619	Total:**0.99996699999999** Showing:**0.99996699999999**	
SLM_PSWG_Bv6--MenRiver_20070403	Total:**0.99950000000001** Showing:**0.99950000000001**	
SLM_PSWG_Bv6--MilwaukeeRiver_20070403	Total:**0.99962000000001** Showing:**0.99962000000001**	
SLM_PSWG_Bv6--OutGap1_20070403	Total:**1.000258** Showing:**1.000258**	
SLM_PSWG_Bv6--SouthShore_20070418	Total:**0.99987400000002** Showing:**0.99987400000002**	

Cluster Fraction Min= 0 % Max = 1 %



Dataset	Counts	Bar Graph (click on a colored bar to open the legend for that bar)
SLM_PSWG_Bv6--Gap1m_20060612	Total:**1.000027** Showing:**0.256865**	
SLM_PSWG_Bv6--Gap_20060619	Total:**1.000091** Showing:**0.178639**	
SLM_PSWG_Bv6--GreenCan1m_20060612	Total:**1.000021** Showing:**0.274119**	
SLM_PSWG_Bv6--GreenCan_20060619	Total:**0.99997399999999** Showing:**0.109268**	
SLM_PSWG_Bv6--JonesIsland_20050420	Total:**0.99986099999999** Showing:**0.192967**	
SLM_PSWG_Bv6--JonesIsland_20070418	Total:**1.000156** Showing:**0.248883**	
SLM_PSWG_Bv6--Junction_20060619	Total:**1.000073** Showing:**0.32189600000001**	
SLM_PSWG_Bv6--Junction_20070403	Total:**1.000358** Showing:**0.403592**	
SLM_PSWG_Bv6--KinnickinnicRiver_20070403	Total:**0.99971099999997** Showing:**0.40352699999998**	
SLM_PSWG_Bv6--Linnwood1m_20060612	Total:**1.00004** Showing:**0.17455**	
SLM_PSWG_Bv6--Linnwood1m_20060807	Total:**1.00004** Showing:**0.223558**	
SLM_PSWG_Bv6--Linnwood_20060619	Total:**0.99996699999999** Showing:**0.117009**	
SLM_PSWG_Bv6--MenRiver_20070403	Total:**0.99950000000001** Showing:**0.55421900000001**	
SLM_PSWG_Bv6--MilwaukeeRiver_20070403	Total:**0.99962000000001** Showing:**0.45638300000002**	
SLM_PSWG_Bv6--OutGap1_20070403	Total:**1.000258** Showing:**0.293095**	
SLM_PSWG_Bv6--SouthShore_20070418	Total:**0.99987400000002** Showing:**0.30260699999999**	



Revised: April 5th, 2010.

QIIME

Quantitative Insights Into Microbial Ecology

Visit the **qiime home page** for a complete discussion of QIIME.

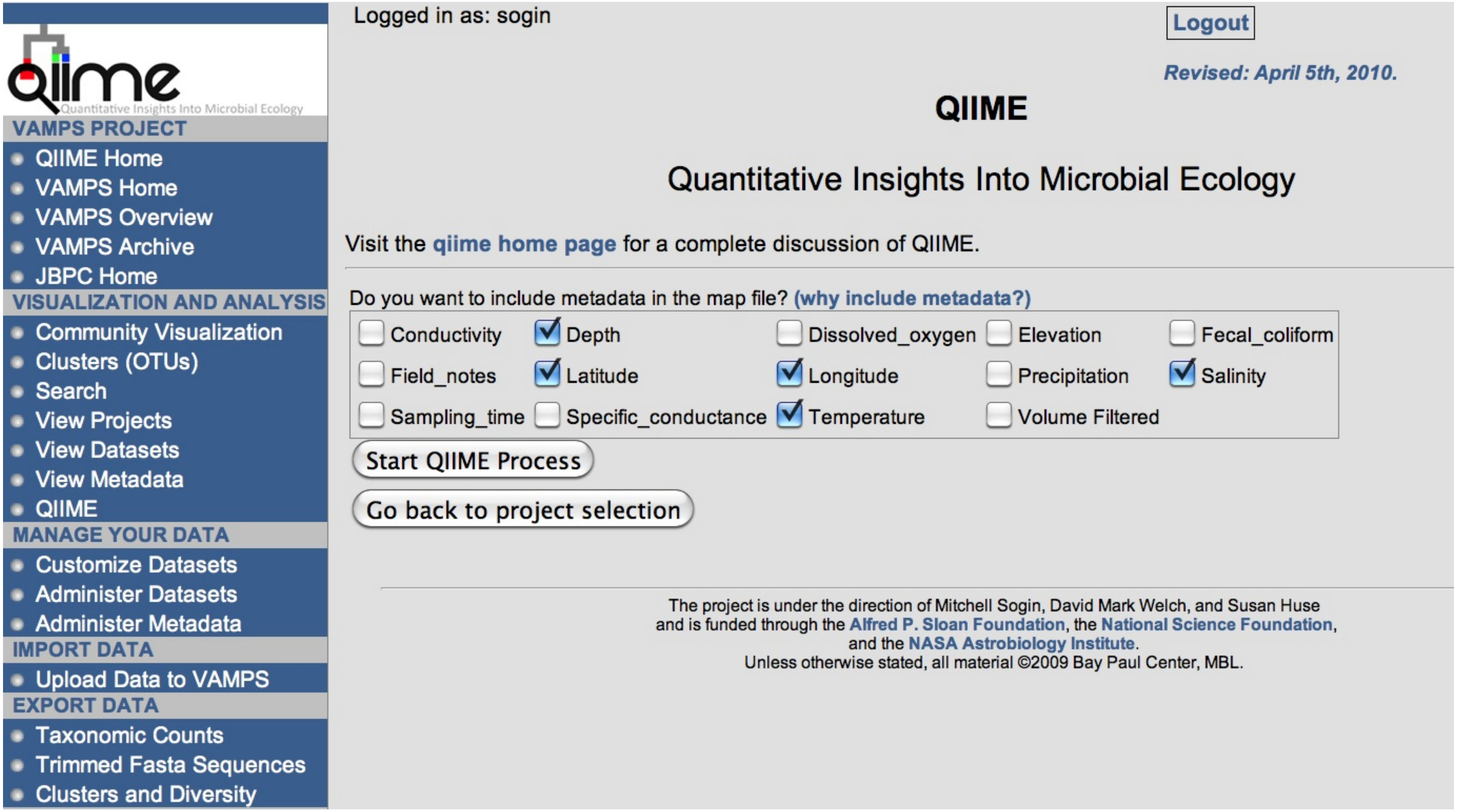
for aministrator view: ○ Your data ● All data

OR Revisit Your QIIME Data:

Select your data: Next...

Dataset Selection

☐ KCK_CBE_Bv6 (?) - Bacillus ecotype study
☐ KCK_DSS_Bv6 (?) - Deep-sea surface sedimer
☐ KCK_EEL_Bv6 (?) - Eel Pond - Woods Hole- MA
☐ KCK_EQP_Bv6v4 (?) - Equatorial and North Pa
☐ KCK_EV9_Ev9 (?) - Test of Euk V9 primers an
☐ KCK_HOT_Bv6 (?) - OTU variation with depth
☐ KCK_KNX_Av6 (?) - South Pacific Gyre
☐ KCK_KNX_Bv6 (?) - South Pacific Gyre
☐ KCK_LSM_Bv6 (?) - Little Sippewissett Marsh
☑ KCK_LSM_Bv6v4 (?) - Little Sippewissett Mars
  - ☑ 010808st1
  - ☑ 010808st2
  - ☑ 010808st3
  - ☑ 010808st4
  - ☑ 010808st5
  - ☑ 010808st6
  - ☑ 010808st7
  - ☐ 020808st1
  - ☐ 020808st2
  - ☐ 020808st3
  - ☐ 020808st4



Logged in as: sogin

Revised: April 5th, 2010.

QIIME

Quantitative Insights Into Microbial Ecology

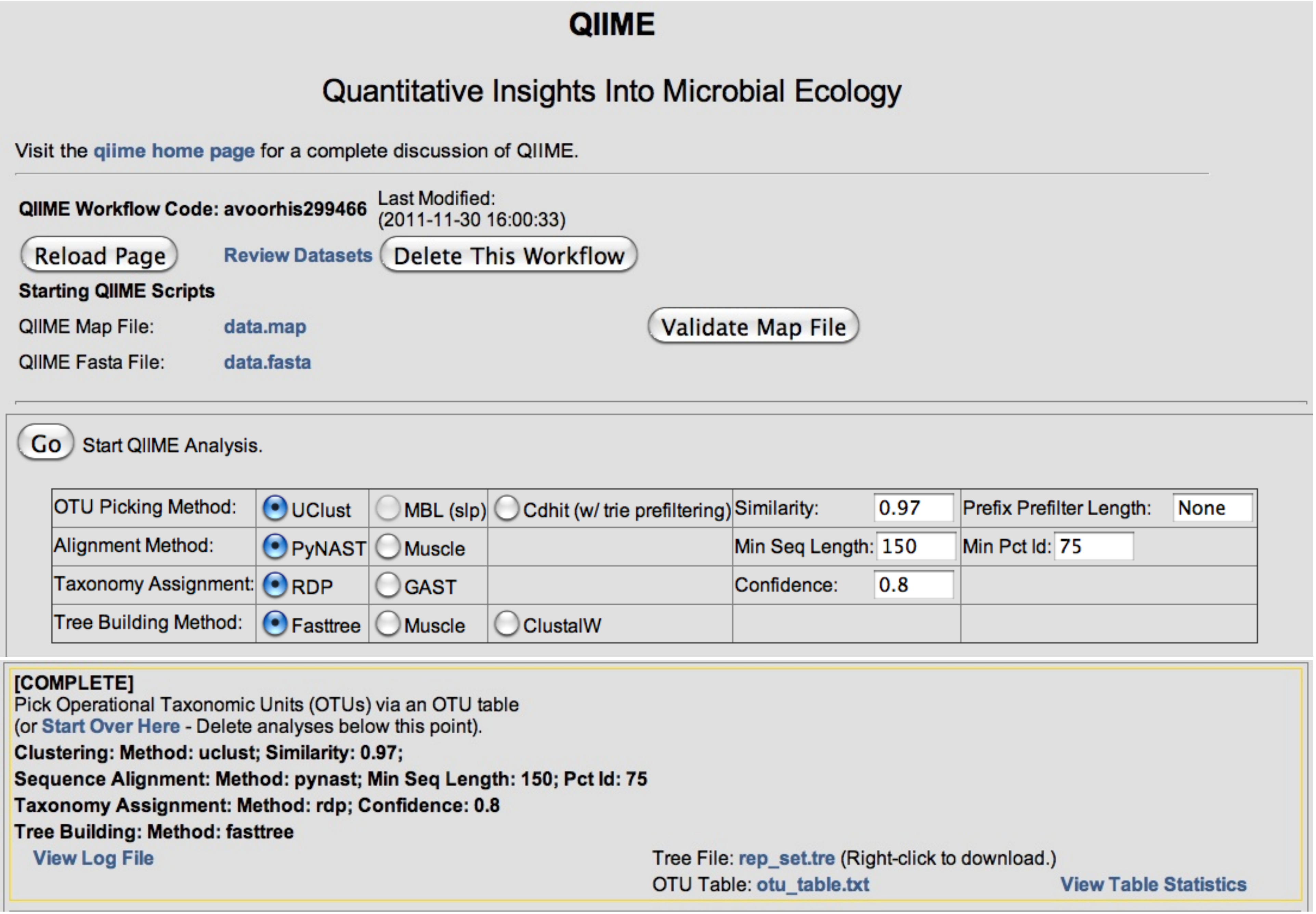
Visit the **qiime home page** for a complete discussion of QIIME.

Do you want to include metadata in the map file? **(why include metadata?)**

☐ Conductivity	☑ Depth	☐ Dissolved_oxygen	☐ Elevation	☐ Fecal_coliform
☐ Field_notes	☑ Latitude	☑ Longitude	☐ Precipitation	☑ Salinity
☐ Sampling_time	☐ Specific_conductance	☑ Temperature	☐ Volume Filtered	

The project is under the direction of Mitchell Sogin, David Mark Welch, and Susan Huse and is funded through the **Alfred P. Sloan Foundation**, the **National Science Foundation**, and the **NASA Astrobiology Institute**.

Unless otherwise stated, all material ©2009 Bay Paul Center, MBL.



QIIME

Quantitative Insights Into Microbial Ecology

Visit the qiime home page for a complete discussion of QIIME.

QIIME Workflow Code: avoorhis299466 Last Modified: (2011-11-30 16:00:33)

Starting QIIME Scripts

QIIME Map File: data.map

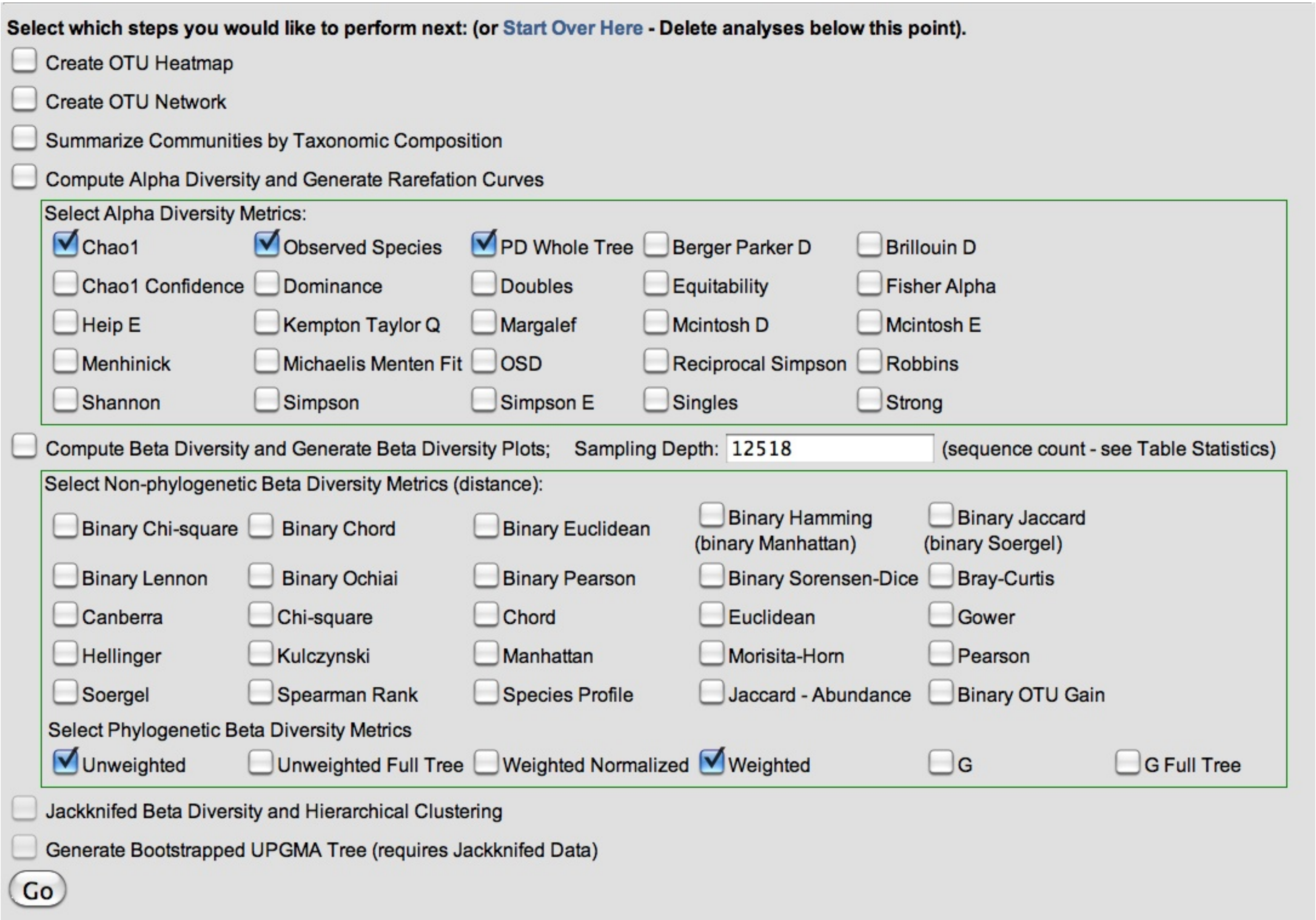
QIIME Fasta File: data.fasta

Start QIIME Analysis.

OTU Picking Method:	● UClust	○ MBL (slp)	○ Cdhit (w/ trie prefiltering)	Similarity:	0.97	Prefix Prefilter Length:	None
Alignment Method:	● PyNAST	○ Muscle		Min Seq Length:	150	Min Pct Id:	75
Taxonomy Assignment:	● RDP	○ GAST		Confidence:	0.8		
Tree Building Method:	● Fasttree	○ Muscle	○ ClustalW				

[COMPLETE]
Pick Operational Taxonomic Units (OTUs) via an OTU table
(or Start Over Here - Delete analyses below this point).
Clustering: Method: uclust; Similarity: 0.97;
Sequence Alignment: Method: pynast; Min Seq Length: 150; Pct Id: 75
Taxonomy Assignment: Method: rdp; Confidence: 0.8
Tree Building: Method: fasttree

Tree File: rep_set.tre (Right-click to download.)
OTU Table: otu_table.txt



Select which steps you would like to perform next: (or Start Over Here - Delete analyses below this point).

☐ Create OTU Heatmap

☐ Create OTU Network

☐ Summarize Communities by Taxonomic Composition

☐ Compute Alpha Diversity and Generate Rarefation Curves

Select Alpha Diversity Metrics:

☑ Chao1 ☑ Observed Species ☑ PD Whole Tree ☐ Berger Parker D ☐ Brillouin D
☐ Chao1 Confidence ☐ Dominance ☐ Doubles ☐ Equitability ☐ Fisher Alpha
☐ Heip E ☐ Kempton Taylor Q ☐ Margalef ☐ Mcintosh D ☐ Mcintosh E
☐ Menhinick ☐ Michaelis Menten Fit ☐ OSD ☐ Reciprocal Simpson ☐ Robbins
☐ Shannon ☐ Simpson ☐ Simpson E ☐ Singles ☐ Strong

☐ Compute Beta Diversity and Generate Beta Diversity Plots; Sampling Depth: 12518 (sequence count - see Table Statistics)

Select Non-phylogenetic Beta Diversity Metrics (distance):

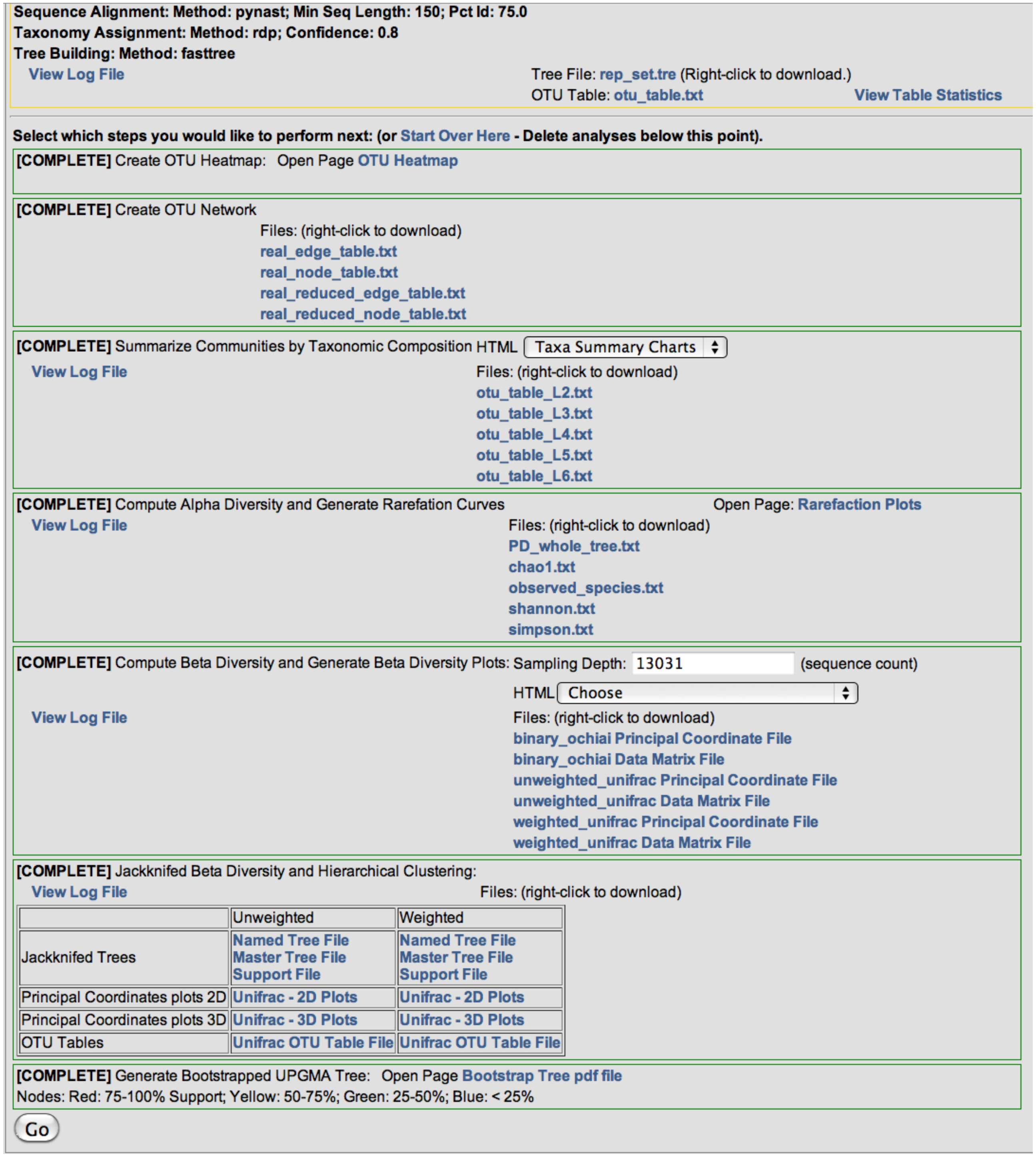
☐ Binary Chi-square ☐ Binary Chord ☐ Binary Euclidean ☐ Binary Hamming (binary Manhattan) ☐ Binary Jaccard (binary Soergel)
☐ Binary Lennon ☐ Binary Ochiai ☐ Binary Pearson ☐ Binary Sorensen-Dice ☐ Bray-Curtis
☐ Canberra ☐ Chi-square ☐ Chord ☐ Euclidean ☐ Gower
☐ Hellinger ☐ Kulczynski ☐ Manhattan ☐ Morisita-Horn ☐ Pearson
☐ Soergel ☐ Spearman Rank ☐ Species Profile ☐ Jaccard - Abundance ☐ Binary OTU Gain

Select Phylogenetic Beta Diversity Metrics

☑ Unweighted ☐ Unweighted Full Tree ☐ Weighted Normalized ☑ Weighted ☐ G ☐ G Full Tree

☐ Jackknifed Beta Diversity and Hierarchical Clustering

☐ Generate Bootstrapped UPGMA Tree (requires Jackknifed Data)



Sequence Alignment: Method: pynast; Min Seq Length: 150; Pct Id: 75.0
Taxonomy Assignment: Method: rdp; Confidence: 0.8
Tree Building: Method: fasttree
View Log File

Tree File: **rep_set.tre** (Right-click to download.)
OTU Table: **otu_table.txt**

View Table Statistics

Select which steps you would like to perform next: (or Start Over Here - Delete analyses below this point).

[COMPLETE] Create OTU Heatmap: Open Page **OTU Heatmap**

[COMPLETE] Create OTU Network
Files: (right-click to download)
real_edge_table.txt
real_node_table.txt
real_reduced_edge_table.txt
real_reduced_node_table.txt

[COMPLETE] Summarize Communities by Taxonomic Composition HTML [Taxa Summary Charts]
View Log File
Files: (right-click to download)
otu_table_L2.txt
otu_table_L3.txt
otu_table_L4.txt
otu_table_L5.txt
otu_table_L6.txt

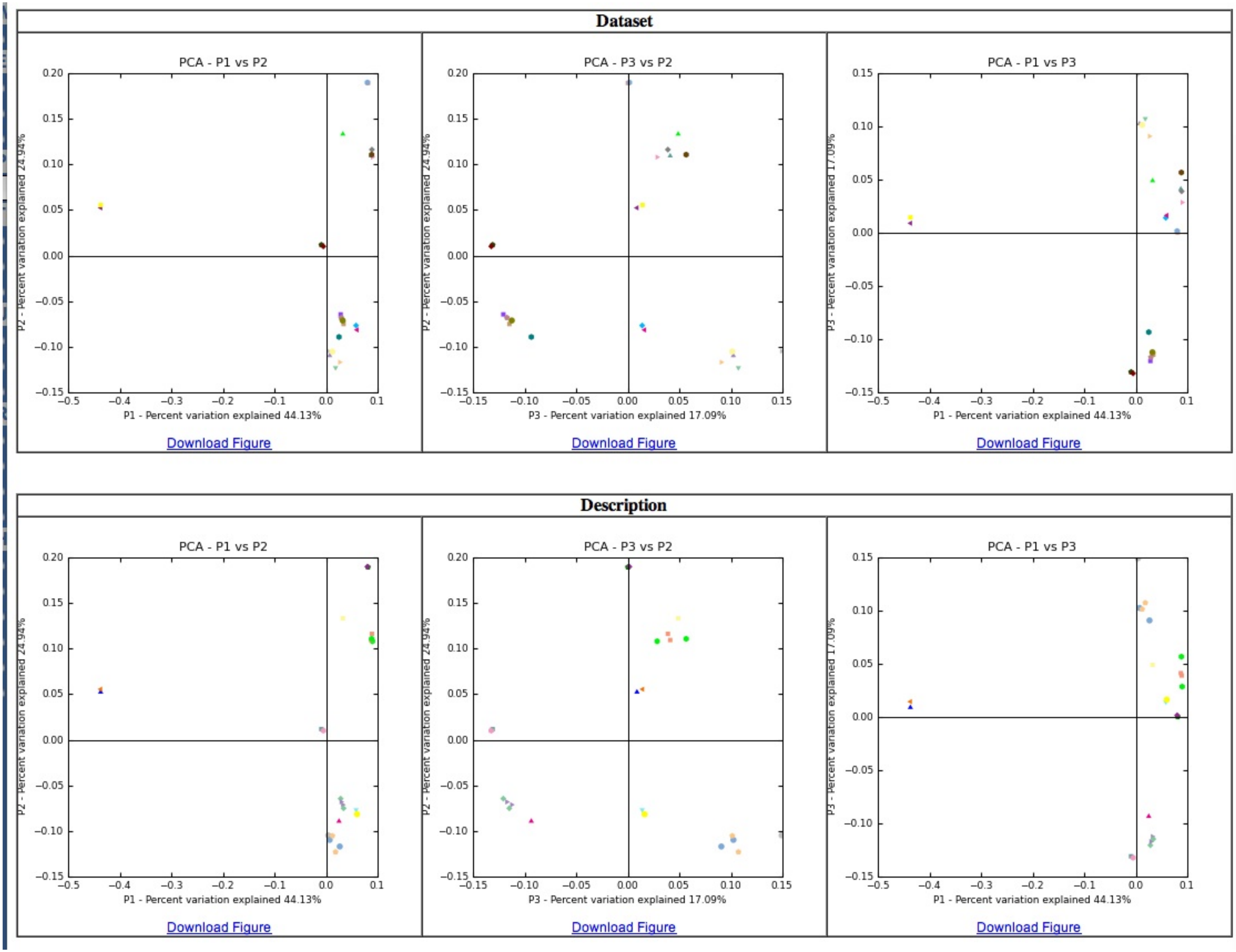
[COMPLETE] Compute Alpha Diversity and Generate Rarefation Curves
Open Page: **Rarefaction Plots**
View Log File
Files: (right-click to download)
PD_whole_tree.txt
chao1.txt
observed_species.txt
shannon.txt
simpson.txt

[COMPLETE] Compute Beta Diversity and Generate Beta Diversity Plots: Sampling Depth: 13031 (sequence count)
HTML [Choose]
View Log File
Files: (right-click to download)
binary_ochiai Principal Coordinate File
binary_ochiai Data Matrix File
unweighted_unifrac Principal Coordinate File
unweighted_unifrac Data Matrix File
weighted_unifrac Principal Coordinate File
weighted_unifrac Data Matrix File

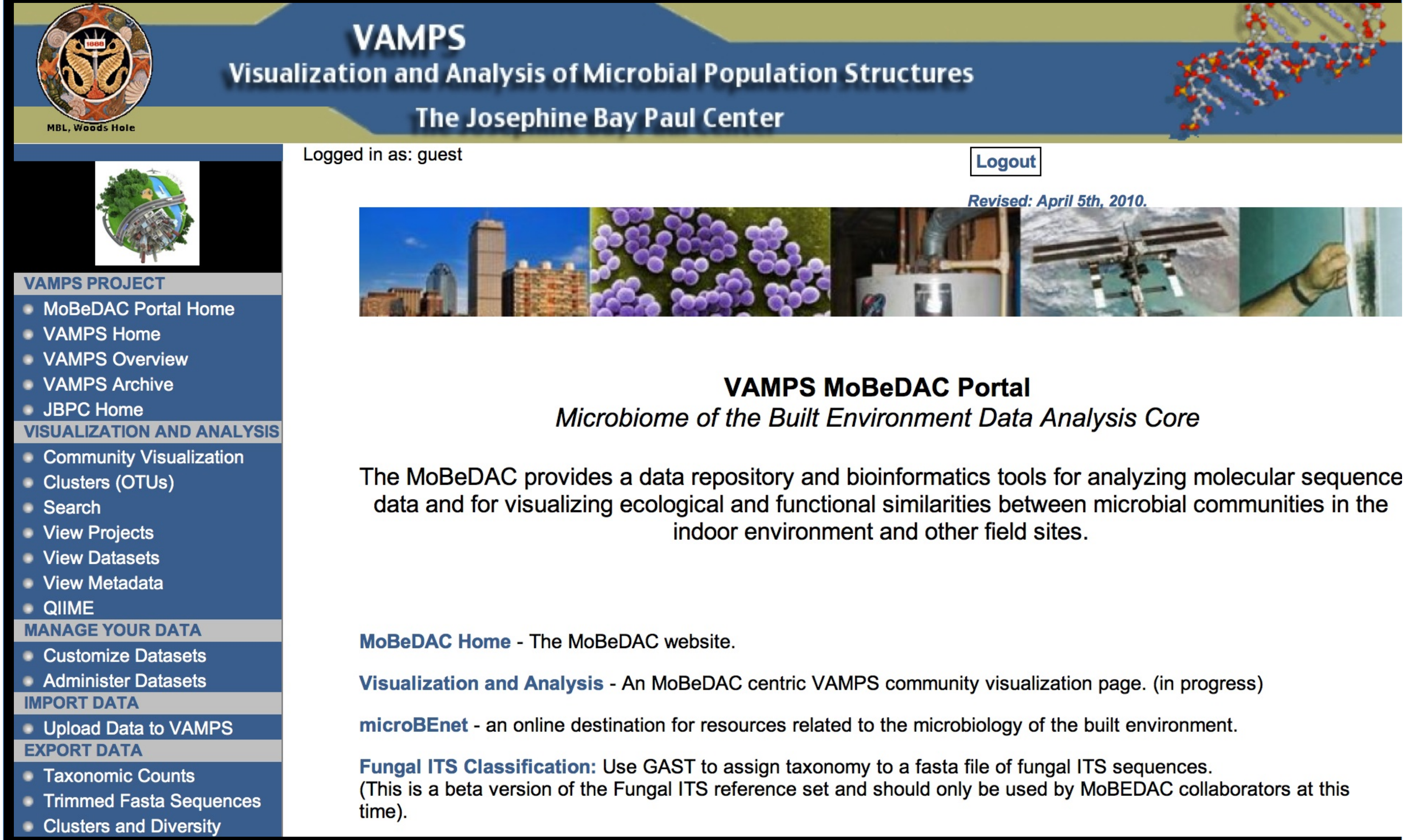
[COMPLETE] Jackknifed Beta Diversity and Hierarchical Clustering:
View Log File
Files: (right-click to download)

	Unweighted	Weighted
Jackknifed Trees	**Named Tree File** **Master Tree File** **Support File**	**Named Tree File** **Master Tree File** **Support File**
Principal Coordinates plots 2D	**Unifrac - 2D Plots**	**Unifrac - 2D Plots**
Principal Coordinates plots 3D	**Unifrac - 3D Plots**	**Unifrac - 3D Plots**
OTU Tables	**Unifrac OTU Table File**	**Unifrac OTU Table File**

[COMPLETE] Generate Bootstrapped UPGMA Tree: Open Page **Bootstrap Tree pdf file**
Nodes: Red: 75-100% Support; Yellow: 50-75%; Green: 25-50%; Blue: < 25%


Dataset
PCA - P1 vs P2
P2 - Percent variation explained 24.94%
P1 - Percent variation explained 44.13%
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
0.1
Download Figure
PCA - P3 vs P2
P2 - Percent variation explained 24.94%
P3 - Percent variation explained 17.09%
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
Download Figure
PCA - P1 vs P3
P3 - Percent variation explained 17.09%
P1 - Percent variation explained 44.13%
Download Figure
Description
PCA - P1 vs P2
P2 - Percent variation explained 24.94%
P1 - Percent variation explained 44.13%
Download Figure
PCA - P3 vs P2
P2 - Percent variation explained 24.94%
P3 - Percent variation explained 17.09%
Download Figure
PCA - P1 vs P3
P3 - Percent variation explained 17.09%
P1 - Percent variation explained 44.13%
Download Figure

http://vamps.mbl.edu



Logged in as: guest

Revised: April 5th, 2010.

VAMPS MoBeDAC Portal

Microbiome of the Built Environment Data Analysis Core

The MoBeDAC provides a data repository and bioinformatics tools for analyzing molecular sequence data and for visualizing ecological and functional similarities between microbial communities in the indoor environment and other field sites.

MoBeDAC Home - The MoBeDAC website.

Visualization and Analysis - An MoBeDAC centric VAMPS community visualization page. (in progress)

microBEnet - an online destination for resources related to the microbiology of the built environment.

Fungal ITS Classification: Use GAST to assign taxonomy to a fasta file of fungal ITS sequences. (This is a beta version of the Fungal ITS reference set and should only be used by MoBEDAC collaborators at this time).