

phylosmith: an R-package for reproducible and efficient microbiome analyses

Schuyler D. Smith and Adina Howe



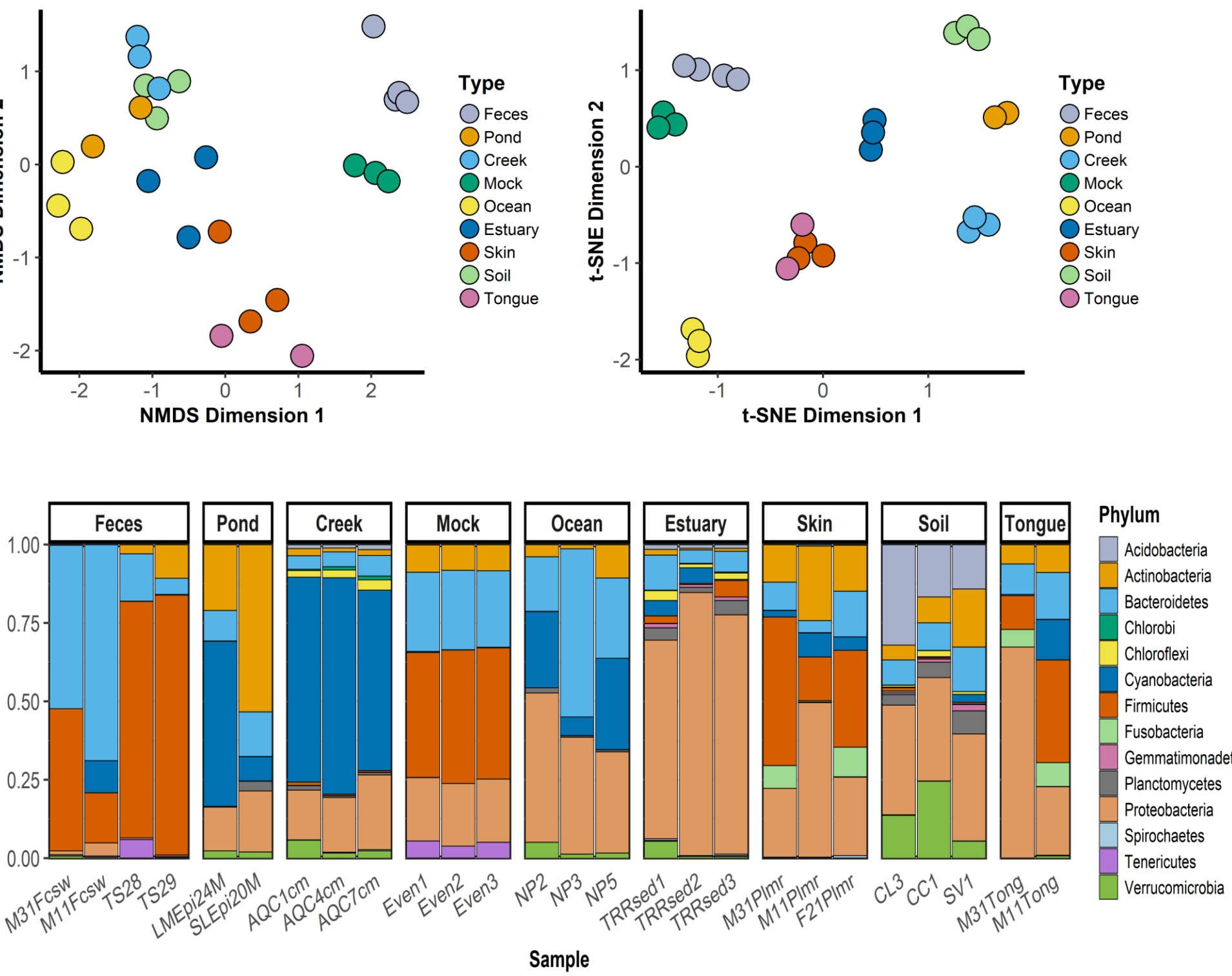
schuyler-smith.github.io/phylosmith

Agriculture and Biosystems Engineering
Bioinformatics and Computational Biology
IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Package Overview: phylosmith is designed to utilize the phyloseq data-standard for microbiome datasets. The analyses of microbiomes have developed a formulaic approach. For many biologists studying in this field, data analysis programs can be a time-barrier to learn. phylosmith aims to provide convenient and efficient implementation of many of the most common approaches in intuitive functions.

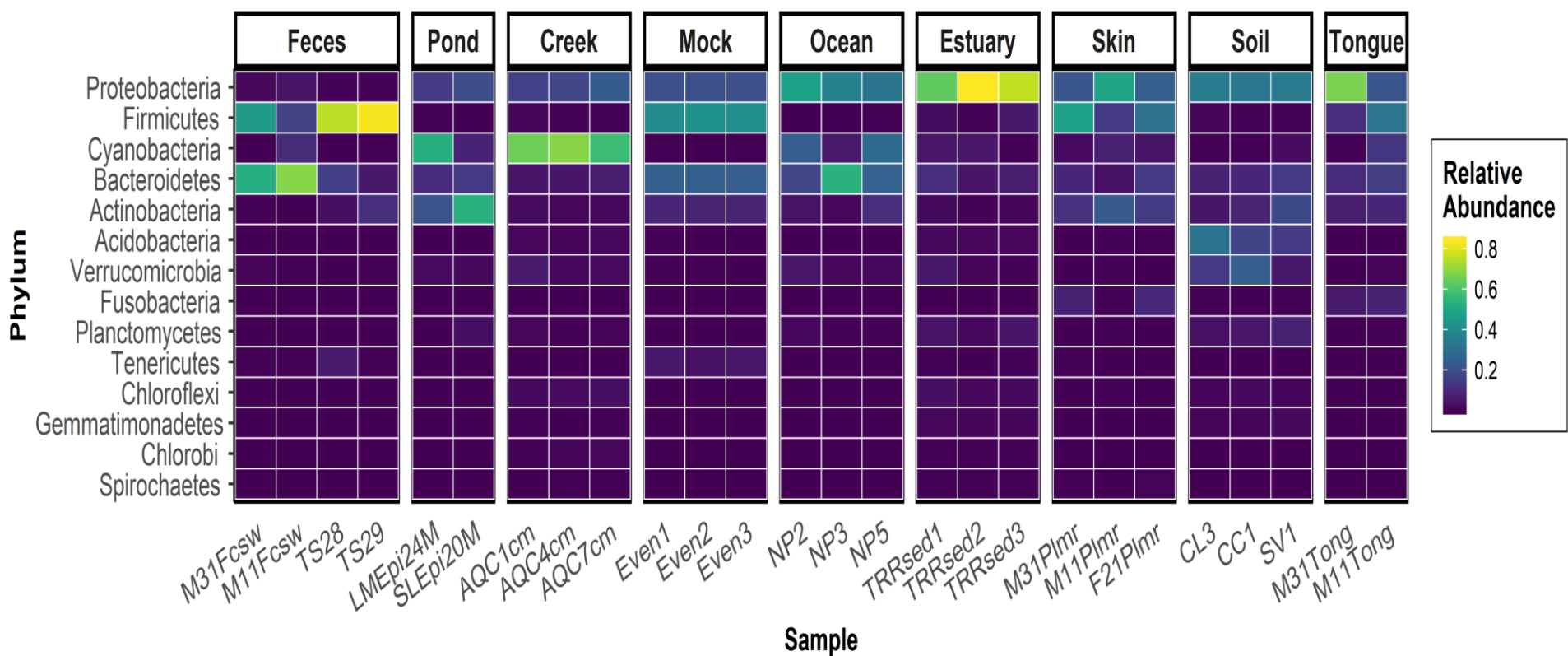
Graphs: commonly used visualizations for microbiome studies.

```
abundance_heatmap()  
abundance_lines()  
alpha_diversity_graph()  
nmads_phyloseq()  
phylogeny_profile()  
taxa_abundance_bars()  
taxa_core_graph()  
tsne_phyloseq()
```



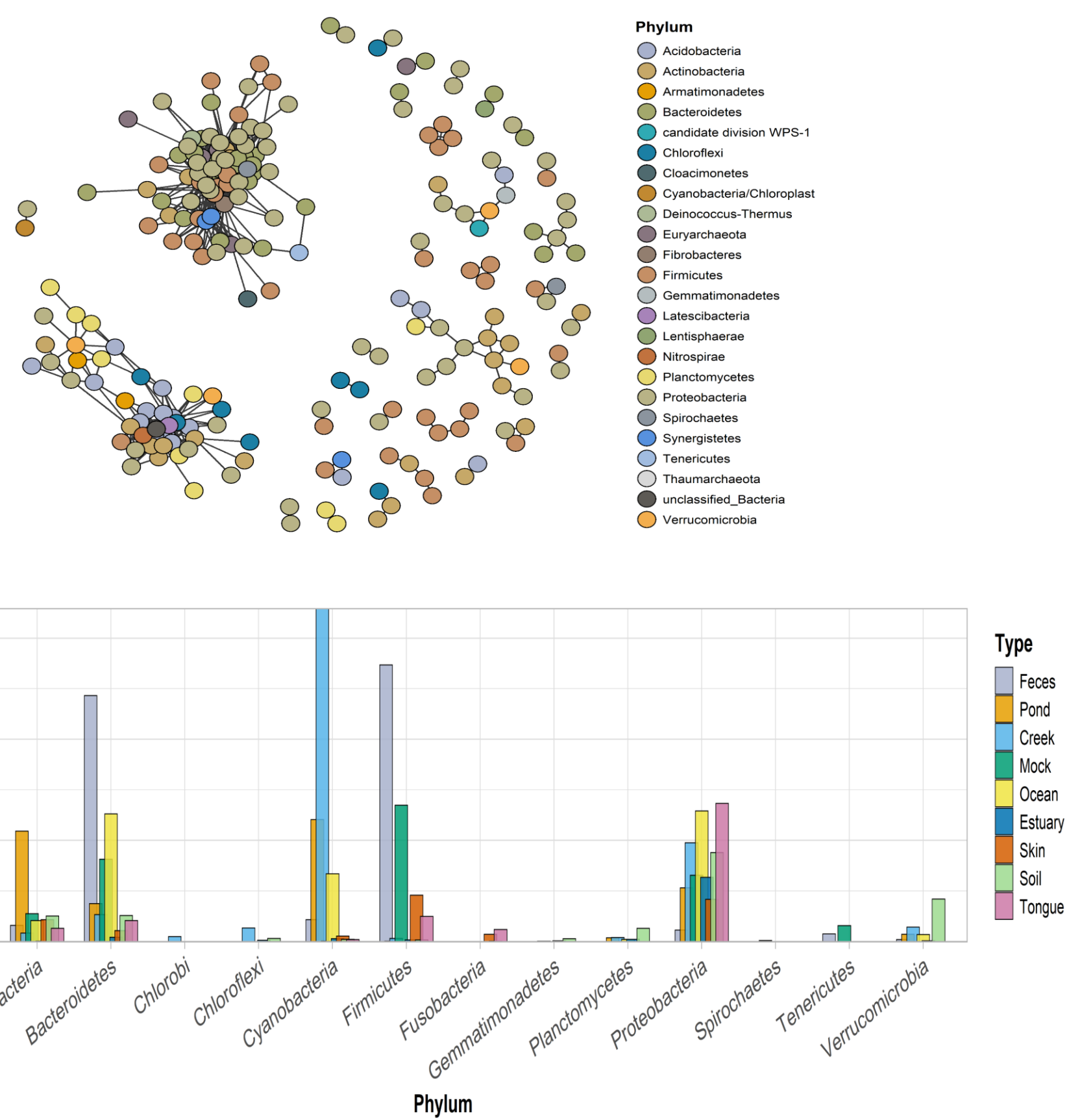
Wrangling: subset, transform, and reshape the data.

```
common_taxa()  
conglomerate_samples()  
conglomerate_taxa()  
melt_phyloseq()  
merge_treatments()  
relative_abundance()  
set_sample_order()  
set_treatment_levels()  
taxa_core()  
taxa_filter()  
taxa_proportions()  
unique_taxa()
```



Co-occurrence: calculated with Spearman's rank-correlation

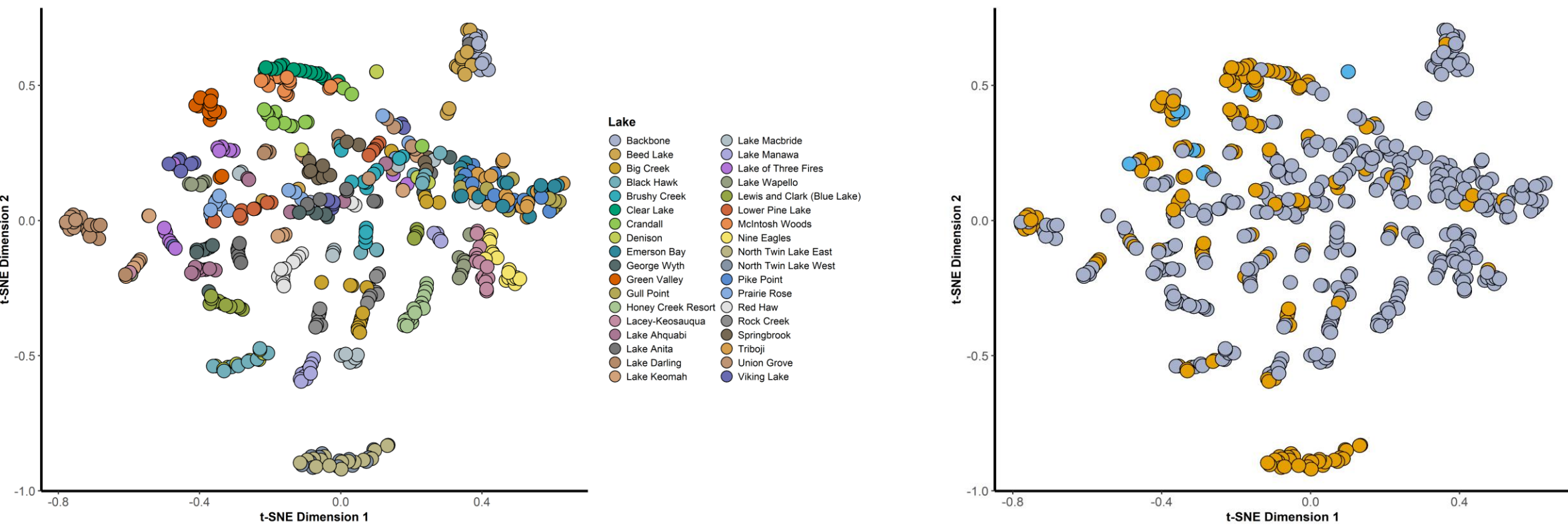
```
co_occurrence()  
permute_rho()  
histogram_permuted_rhos()  
quantile_permuted_rhos()  
network_layout_ps()  
co_occurrence_network()
```



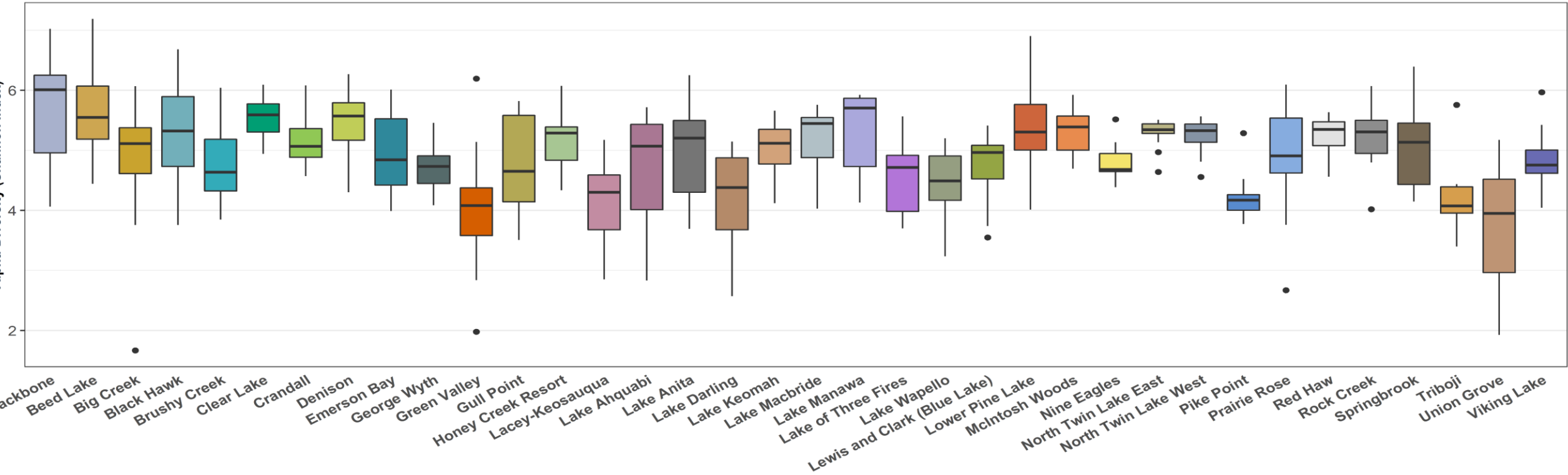
Example Workflow: Toxic algae blooms of Iowa lakes.

```
library(phyloseq)  
library(phylosmith)
```

```
tsne_phyloseq(Iowa_Lakes, treatment = 'Lake', circle = FALSE)  
tsne_phyloseq(Iowa_Lakes, 'Microcystin', circle = FALSE)
```



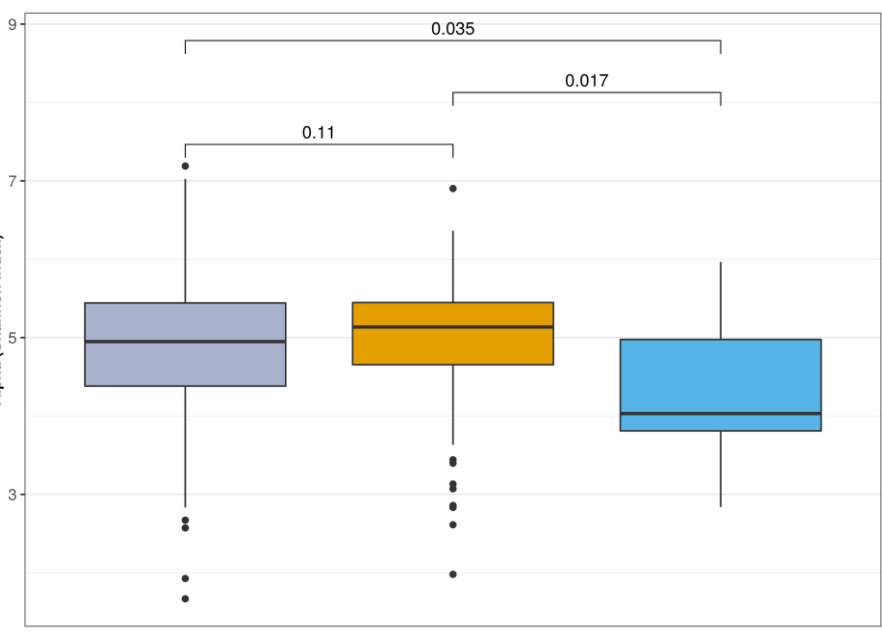
```
alpha_diversity_graph(Iowa_Lakes, treatment = 'Lake')
```



```
alpha_diversity_graph(Iowa_Lakes, treatment = 'Microcystin')
```

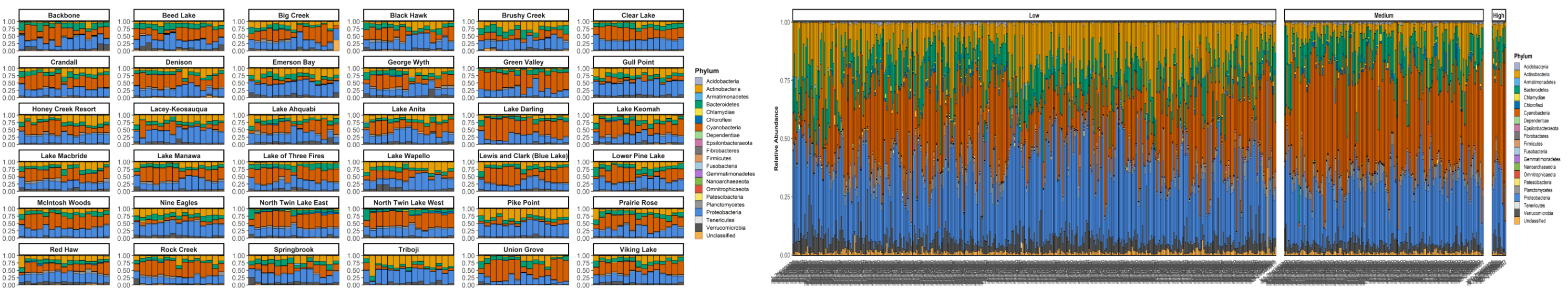
```
taxa_proportions(Iowa_Lakes, 'Phylum')
```

Phylum	Proportion
1: Cyanobacteria	0.312
2: Proteobacteria	0.299
3: Actinobacteria	0.139
4: Bacteroidetes	0.117
5: Verrucomicrobia	0.055
6: Planctomycetes	0.030
7: Chloroflexi	0.011
8: <NA>	0.013
9: Firmicutes	0.004
10: Armatimonadetes	0.003



```
phyla <- taxa_proportions(Iowa_Lakes, 'Phylum')$Phylum[1:20]  
Iowa_Lakes <- subset_taxa(Iowa_Lakes, Phylum %in% phyla)
```

```
phylogeny_profile(Iowa_Lakes, 'Lake', classification = 'Phylum',  
  relative_abundance = TRUE, grid = T)  
phylogeny_profile(Iowa_Lakes, 'Microcystin', classification = 'Phylum',  
  relative_abundance = TRUE, grid = F)
```



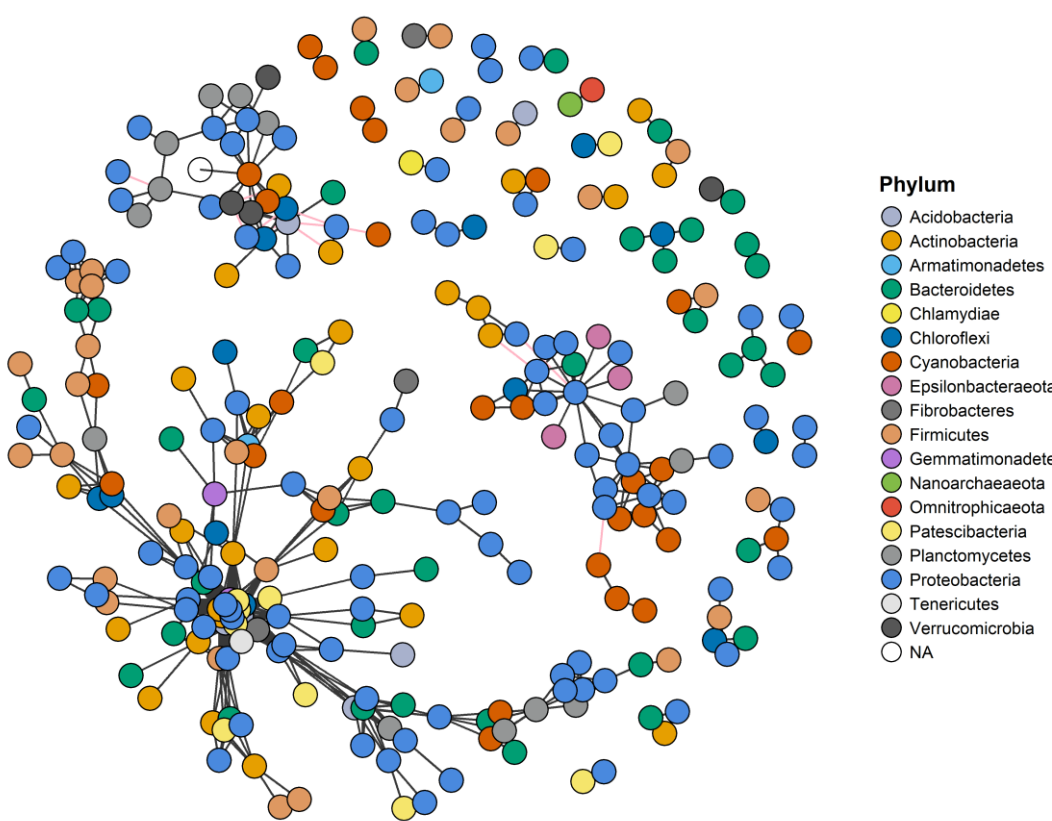
```
Toxic_Lakes <- taxa_filter(Iowa_Lakes, treatment = 'Microcystin',  
  subset = c('Medium', 'High'))  
Toxic_Lakes <- conglomerate_taxa(Toxic_Lakes, 'Genus')
```

```
permuted_rho <- permute_rho(Toxic_Lakes, permutations = 1000)  
q_rho <- quantile_permuted_rhos(permuted_rho, p = .001)
```

```
cooccurrence <- co_occurrence(Toxic_Lakes, p = 0.01)  
cooccurrence <- cooccurrence[rho >= 2*(q_rho$upper) |  
  rho <= 2*(q_rho$lower)]
```

```
layout <- network_layout_ps(Toxic_Lakes, co_occurrence_table = cooccurrence)
```

```
co_occurrence_network(Toxic_Lakes, classification = 'Phylum',  
  co_occurrence_table = cooccurrence, layout = layout)
```



Acknowledgements

This work was supported by the National Science Foundation Directorate of Biological Sciences under awards DEB 1737758 and DEB 1737765. Data for the Iowa lakes was supported by EPA STAR grant number 83927001.

References

Caporaso, J. G., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. PNAS, 108, 4516-4522. PMID: PMC3063599

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE, 8(4), 1-11. doi:10.1371/journal.pone.0061217

Smith, (2019). phylosmith: an R-package for reproducible and efficient microbiome analysis with phyloseq-objects. Journal of Open Source Software, 4(38), 1442, <https://doi.org/10.21105/joss.01442>

github.com/schuyler-smith/phylosmith