# Fourth Year Committee Report

Sarah Stevens, McMahon Lab
Meeting: October 27th, 2015 in MSB 6503

## Research Progress

The manuscript I submitted last year has been published in ISMEJ. This work shows evidence for both of the two major models of bacterial speciation by tracking single-nucleotide variants in populations of bacteria using 30 metagenome-assembled genomes (MAGs) and a metagenomic timeseries containing 63 samples over 6 years. We found that one population showed a genome-wide loss of diversity where others had seen a reduction in diversity for only a particular region of the genome. This suggests that co-existing populations in the lake have a high selection to recombination ratio where others have a low selection to recombination ratio.

In the past year, I have also worked on the population dynamics of many dominant freshwater bacteria, focusing on two groups (which represent roughly family level) for which we have multiple single-amplified genomes (SAGs) and sequence-discrete populations. Genetically distinct populations within roughly species level groups of the acI actinobacterial lineage living in the same lake had different seasonal abundance patterns, suggesting these populations were also ecologically distinct. In contrast, sympatric LD12 populations were much less genetically differentiated and had similar temporal abundance patterns. This suggests that within one lake, some freshwater lineages harbor genetically discrete (but still closely related) and ecologically distinct populations, while other lineages are composed of less differentiated populations with overlapping niches. I have submitted the manuscript for this project on bioRxiv, a biology preprint server and it will soon be submitted to Nature Communications.

The manuscripts for both of these projects are attached to this report.

## New Questions

I continue to be interested in discerning the forces shaping wild bacterial populations and would like to turn my focus from the whole genomes of populations to their gene content.

*Does selection affect homologs of the same gene differently? Which traits are under strong selection?* First I will identify the homologous genes in the reference MAGs and SAGs using BLAST and clustering. Then I will mapping the metagenomic reads from the same lake back the reference genomes and determine the sequence-discrete populations each reference belongs to based on its coverage discontinuity. For each of these populations and their genes, I will calculate the relative abundance, call single nucleotide variants, and calculate the degree of selection.

*Are there related sequence-discrete populations in TB and CB? How closely related are they? Do they share a common gene pool?* We can now bin MAG's from Crystal Bog (CB), which is of similar location and trophic status to Trout Bog (TB). I will bin MAGs from the CB assemblies and use new techniques to get more MAGs from the TB assemblies. With these genomes, I will ask if there are related genomes and how similar they are across their genomes. For genomes that are very closely related (>95% nucleotide identity across their whole genomes), I will also investigate if the associated populations share a common gene pool between the two lakes. Using the cross mapping I will identify if there are regions or genes that are present in only one of the lakes and if the diversity of shared genes is different between the two lakes.

## Publications

Garcia, S. L.*, **Stevens, S. L. R.***, Crary, B., Martinez-Garcia, M., Stepanauskas, R., Woyke, T., Tringe, S. G., Andersson, S., Bertilsson, S., Malmstrom, R., McMahon, K. D. (*in prep for Nature Communications*). Contrasting patterns of genome-level diversity across distinct co-occurring populations. doi:http://dx.doi.org/10.1101/080168
Bendall, M. L.*, **Stevens, S. L. R.***, Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W., Martin, J., Pati, A., Bushnell, B., Froula, J., Kang, D., Tringe, S. G., Bertilsson, S., Moran, M. A., Shade, A., Newton, R. J., McMahon, K. D., Malmstrom, R. R. (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. ISMEJ. doi:http://dx.doi.org/10.1038/ismej.2015.241 **featured article**
*Equal contributors

## Oral Presentations

**Stevens, S. L. R.**, Bendall, M. L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., ... McMahon, K. D. Malmstrom, R. R. Tracking Microbial Populations Through Time Using Single-cell Genomes and Metagenomics. UW Center for Limnology Seminar. December 2015. Madison, WI. https://goo.gl/0ge2LZ
**Stevens, S. L. R.**, Bendall, M. L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., ... McMahon, K. D. Malmstrom, R. R. Genome-wide and Gene-specific Selective Sweeps in Freshwater Bacterial Populations Revealed Using Metagenomics. JF Crow Institute for the Study of Evolution Seminar Series. October 2015. Madison, WI. https://goo.gl/oSnDYG

## Poster Presentations

**Stevens, S. L. R.**, Garcia, S. L. ... McMahon, K. D. Contrasting patterns of genome-level diversity across distinct co-occurring populations. 16th International Symposium on Microbial Ecology. August 2016. Montreal, Canada. https://goo.gl/6iunz0
**Stevens, S. L. R.**, Garcia, S. L. ... McMahon, K. D. Tracking distinct freshwater populations through time by mapping metagenomes to single-cell genomes. DOE Joint Genome Institute User Meeting 2016. Walnut Creek, CA. https://goo.gl/ShUQVn

## Professional Development

- Software/Data Carpentry
  - Taught 4 Software Carpentry Workshops (2 on-campus, 2 off-campus)
  - Helped with 2 Data Carpentry Workshops on-campus
  - Helped with Instructor Training on-campus
- Computational Biology, Ecology, and Evolution(ComBEE) support group - Started Nov. 2014
- ComBEE Python Study Group - Started Dec. 2014
- ComBEE R Study Group - Started Fall 2015