Poster for Evolution meeting (Providence, R. I. June 21-25, 2019)

May 03, 2019

Title

Dolichospermum evolution during during cyanobacterial bloom: insights from metatranscriptomics.

Authors & Affiliations

Sébastien Renaut[1,2] Nathalie Fortin[3] Charles W. Greer[4] Jesse B. Shapiro[2]

[1]Institut de Recherche en Biologie Végétale, Université de Montréal, Montréal, Québec, Canada
[2]Department of Biological Sciences, Université de Montréal, Montreal, Quebec, H3T 1J4, Canada
[3]

[4]Energy Mining and Environment, National Research Council Canada, Montreal, QC, Canada

Abstract

Specific environmental conditions, such as elevated levels of nutrients from human activities, warmer temperatures, still water, and sunlight promote the growth of cyanobacteria, which at high density, form harmfull algae blooms. Each summer, cyanobacteria make an appearance in Lake Champlain and pose health risks to humans and animals due to the production of cyanotoxins. Here, we used a metatranscriptomics approach that targets genes expressed in the water to identify changes in the bacterial community during the summer. At the height of the bloom (mid-September), chlorophyll A increased by nearly 100 fold and communities were dominated by the cyanobacteria *Dolichospermum circinale* and to a lesser extent *Microcystis aeruginosa*. 7,000 *Dolichospermum* SNPs (10% of all SNPs) showed changes in allele frequencies in at least one of the sampling site, from the begining of the bloom to the end of the season, showing a highly dynamic evolution of the system. 19 genes showed repeatable patterns of evolution in all three sampling sites. These represent good candidates that may help to predict blooms in the future.

Introduction

- Specific environmental conditions, such as elevated levels of nutrients from human activities, warmer temperatures, still water, and sunlight promote the growth of cyanobacteria, which at high density, form harmfull algae blooms.
- Each summer, cyanobacteria make an appearance in Lake Champlain and pose health risks to humans and animals due to the production of cyanotoxins.

Objectives

- Use a metatranscriptomics approach that targets genes expressed in the water to identify changes in the bacterial community during the summer.
- Identify *Dolichospermum* (main bloom causing cyanobacteria) genes that change in allele (SNPs) frequency from the beginning to the end of the bloom. These represent good candidates to help to predict blooms in the future.

Method

1. Water sampled at three location in Lake Champlain (St1, St2, PRM, **Figure 1: A map of sampling sites**) during a cyanobacterial bloom in the summer of 2016 (June 1st - October 10th).

2. RNA extracted, followed by High Throughput sequencing (100bp paired end Illumina HiSeq).
3. Metatranscriptomes processed (trimming, merging, removal of rRNA, annotation) using SAMSA2 pipeline[1].
4. Sequences annotated to *Dolichospermum* (the main cyanobacteria responsible for the bloom) extracted and used to produce a *de novo* reference transcriptome assemblies using Trinity[2]
5. *Dolichospermum* sequences aligned to transcriptome using bowtie2[3]. SNP called using SAMtools[4].
6. Changes in allele frequency quantified using logistic regressions in R.

7. Candidate genes showing significant changes in allele frequency matched to reference database and compared against reference transcriptome.

Results

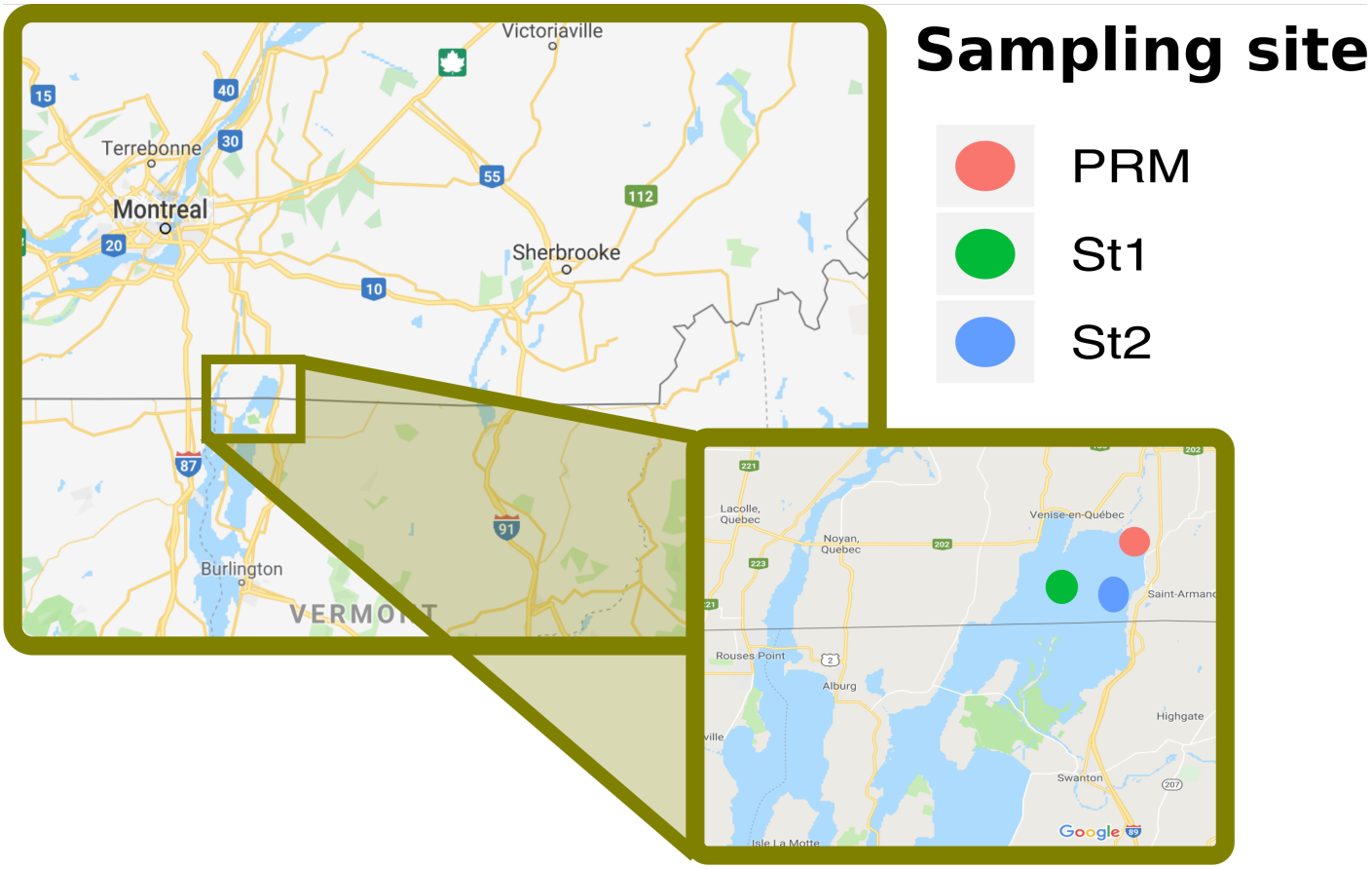

Figure 1: A map of sampling sites

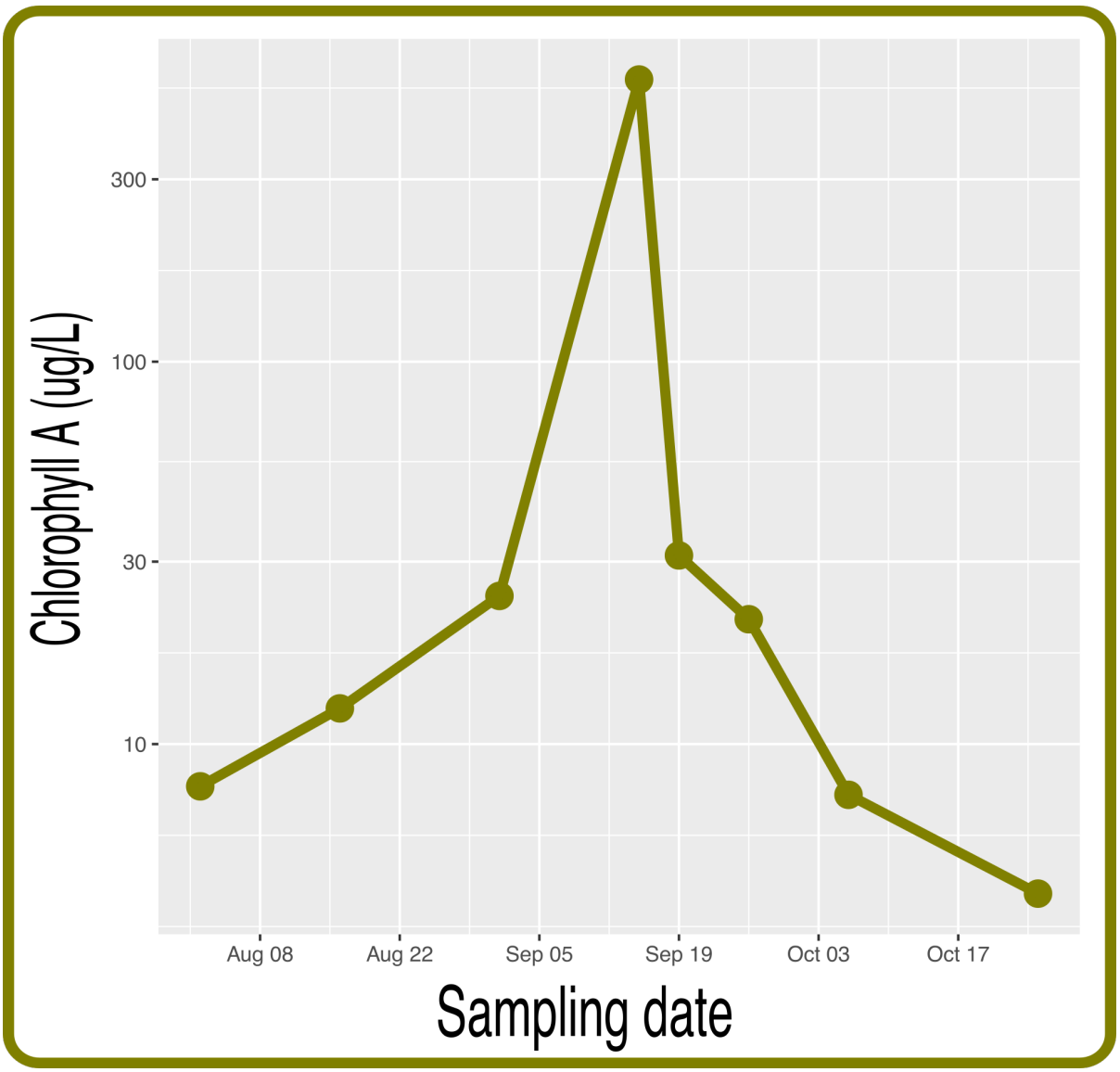

Figure 2: Chlorophyll concentration peaked in mid-September (August 1st: 7.8 ug/L, Sept 15th: 547 ug/L)

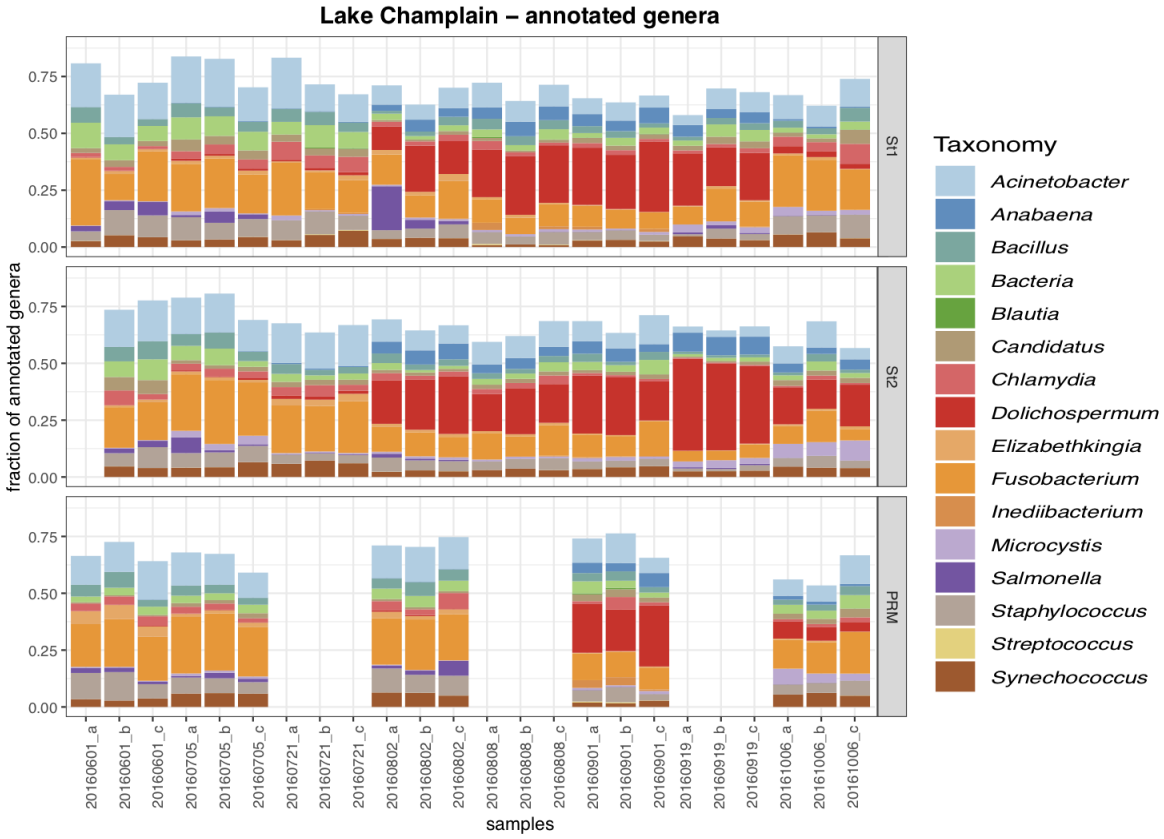

Figure 3: Communities dominated by Gram negative *Fusobacterium* & *Acinetobacter* prior to bloom. During Bloom, *Dolichospermum circinale* rose in frequency, followed to a lesser extent by *Microcystis aeruginosa* in September.

(An image of Dolicho anc Microcystis could be usefull)

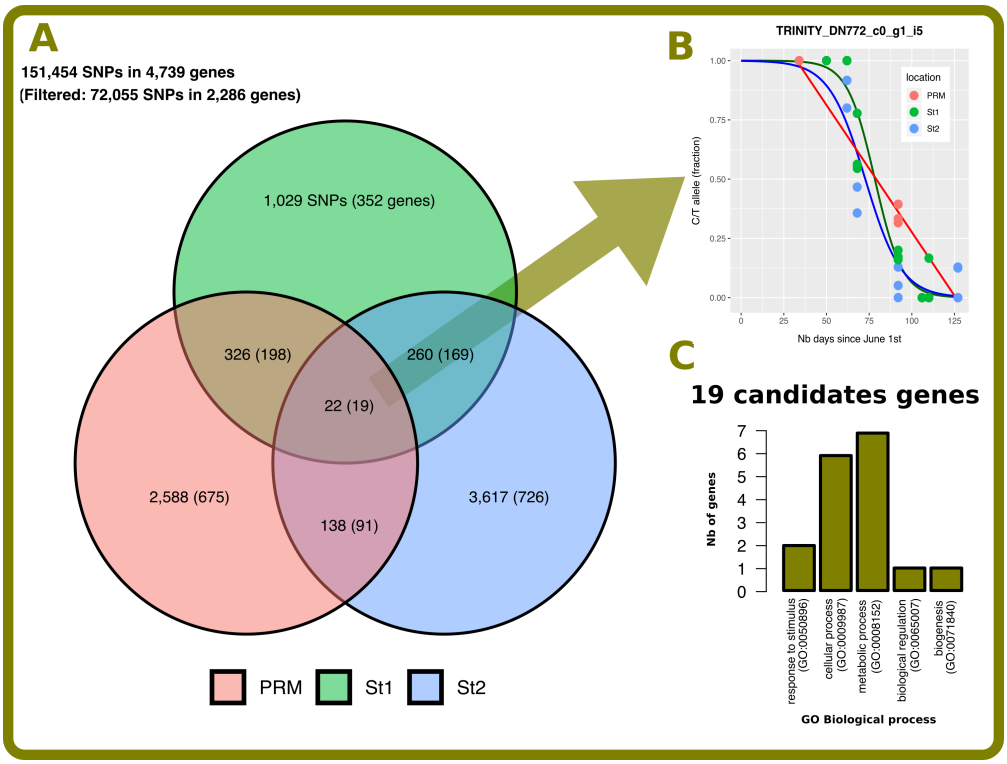

Figure 4: A. Over 150k SNPs identified (<50% missing data: 72k). Venn diagram: Number of SNPs (genes) showing significant changes in allele frequency from beginning (June 1st) to end of bloom (October 6th). B. Example of a significant SNP in both St1, St2 (logistic regressions) and PRM (linear regression) sites (Total: 22 SNPs, 19 SNPs). C. Gene ontology groups for candidates (no significant over-represented groups compared to reference transcriptome)

Conclusions

- Metatranscriptomics can identify cyanobacteria known to cause blooms (*Dolichospermum circinale*, *Microcystis aeruginosa*).
- ~7,000 SNPs (10% of all SNPs) show changes in allele frequencies during the bloom in at least one of the sampling site, showing a highly dynamic evolution of the system. 19 genes show repeatable changes in allele frequencies in all three sampling sites. These represent good candidates to predict bloom.
- Further sampling and sequencing in progress in other lakes will help to confirm patterns observed here.

References

1. Westreich, S. T., Treiber, M. L., Mills, D. A., Korf, I. & Lemay, D. G. SAMSA2: A standalone metatranscriptome analysis pipeline. *BMC bioinformatics* **19**, 175 (2018).

2. Grabherr, M. G. *et al.* Trinity: Reconstructing a full-length transcriptome without a genome from rna-seq data. *Nature biotechnology* **29**, 644 (2011).

3. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nature methods* **9**, 357 (2012).

4. Li, H. *et al.* The sequence alignment/map format and samtools. *Bioinformatics* **25**, 2078–2079 (2009).