Network analysis reveals statistical associations between cyanophage host genes and cyanobacterial ecotypes across ocean basins

Jacob A. Cram, David García-Prieto, Matthew D. Hays, Clara A. Fuchsman

University of Maryland
Center for Environmental Science
Horn Point Laboratory

Introduction

- Cyanophage harbor genes for host functions (Breitbart 2012).
  - Eg. Photosystem II (psb), Purine synthesis (pur)
- Different ecotypes of *Prochlorococcus* are adapted to different conditions (Johnson et al. 2006). Do cyanophage that infect different ecotypes have different host genes?

Methods

- Cyanophage host genes and cyanobacterial ecotype abundances were obtained from publicly available metagenome libraries (Figure 1).
- Viral host genes were identified by placing sequences on phylogenetic trees.
- Associations were identified between each viral gene and all cyanobacterial ecotypes + environmental and location parameters
  - With many individual Lasso regressions
  - Each lasso has the same manually selected penalty coefficient (λ = 0.2).

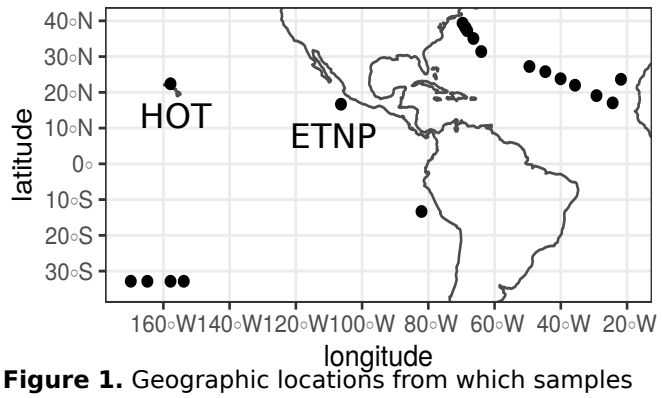

Figure 1. Geographic locations from which samples are collected. Samples were from geotraces cruises GA03 (North Atlantic, East Coast North Atlantic, West Coast North Atlantic) and GP13 (South Pacific) (Biller et al. 2018); from the Hawaii Ocean Time-series (year Aug 2015; Luo et al 2020). ETNP is Fuchsman et al 2017. ETSP is Fuchsman et al (in revision).

Acknowledgements

The authors thank JL Weissman and S Gleich for helpful conversations and advice.

Results

Cyanobacterial Ecotypes

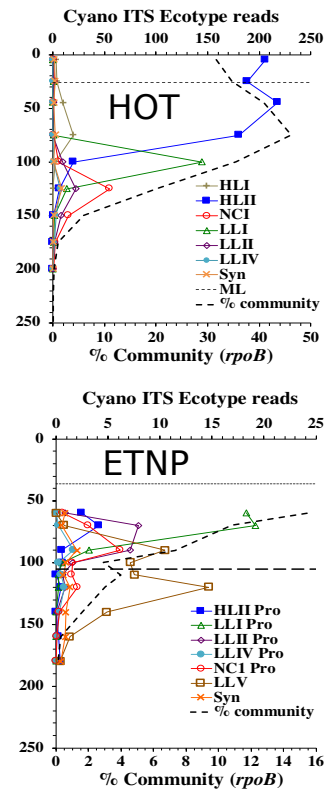

Figure 2. Prochloroccus ecotype relative abundances at HOT (Aug 2015; oxic), and ETNP (anoxic). Grey dotted line····base of mixed layer. Black dashed line––top of anoxic zone (<10 nM O_2).

Example Correlations

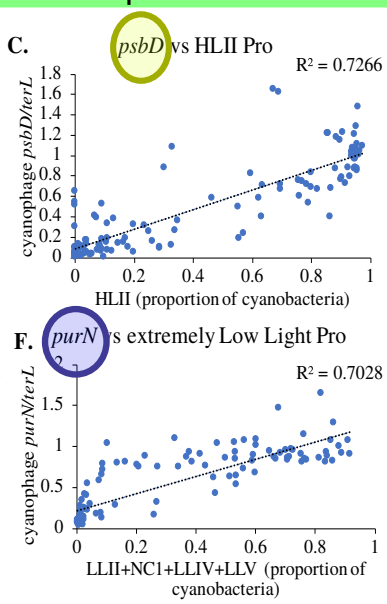

Figure 4. Example correlations between viral host gene ratios and cyanobacterial ecotype abundance. Data are from all stations in Figure 1, not just those shown in Figure 2.

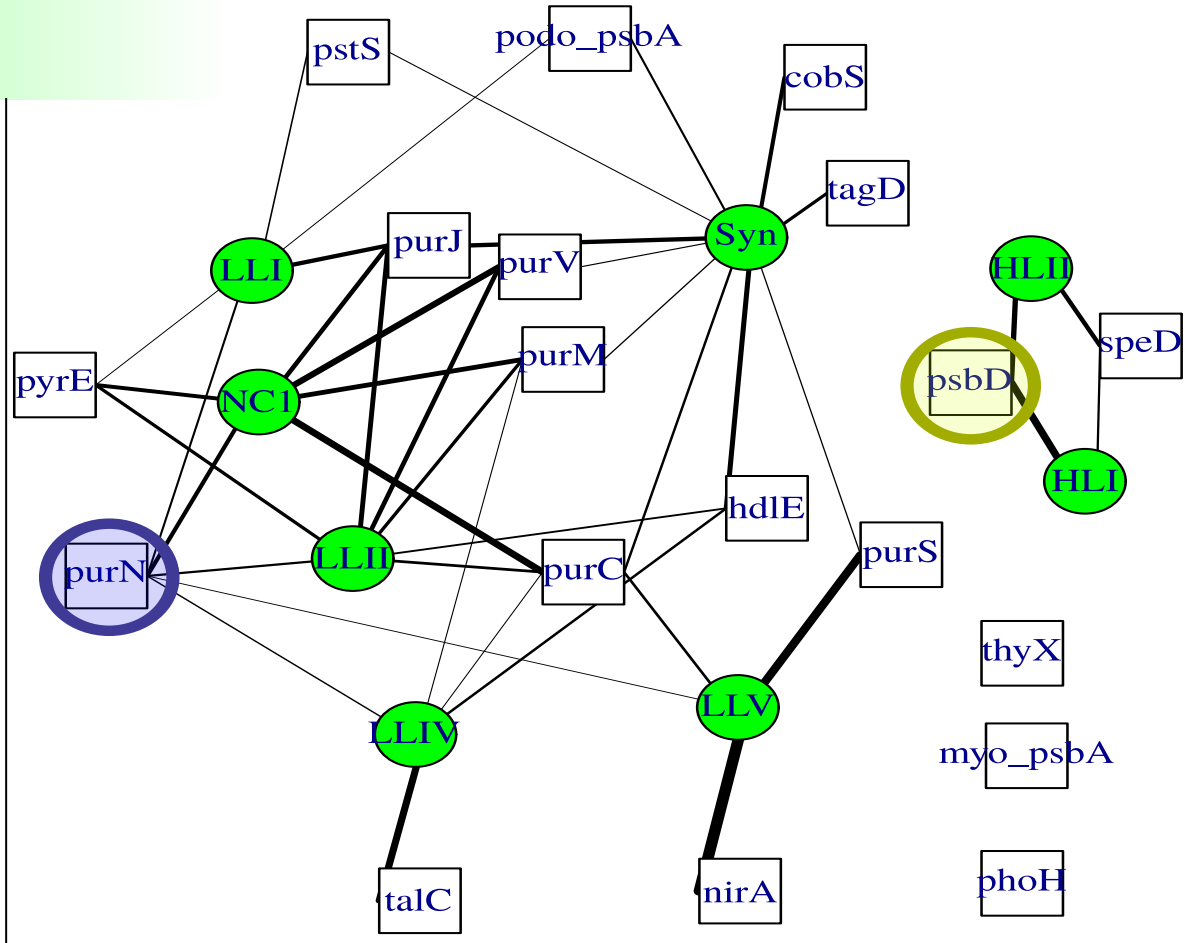

Figure 5. Association network between cyanobacterial ecotypes and host gene ratios. Edge widths are associated with lasso coefficient scores and indicate stronger associations.

Cyanophage Host Gene Ratios

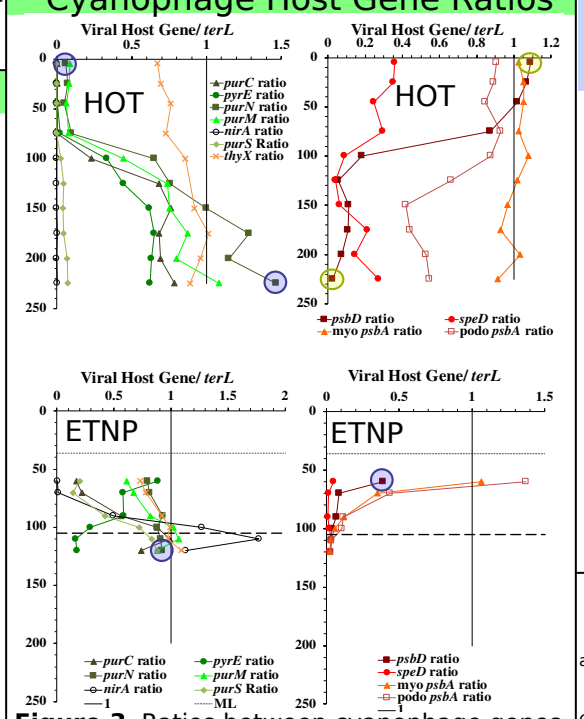

Figure 3. Ratios between cyanophage genes coding for different host functions and the cyanophage *terL* core gene. Higher values indicate greater expression of these genes by cyanobacterial viral factories.

Discussion

- We have a system for identifying putative assocations between viral host gene abundance and cyanobacterial ecotypes.
- It appears that there are different cyanophage ecotypes that associate with different cyanobacterial ecotypes.
- Cyanophage appear to use different genes to adapt to the different cyanobacterial

References

- Biller SJ. et al. Marine microbial metagenomes sampled across space and time. Sci Data. 2018.
- Breitbart M. Marine Viruses: Truth or Dare. An. Rev. of Mar. Sci. 2012
- Fuchsman CA. et al. Niche Partitioning of the N Cycling Microbial Community of an Offshore Oxygen Deficient Zone. F. in Micro. 2017.
- Fuchsman CA. et al. An Analysis of Protists in Pacific Oxygen Deficient Zones: Implications for Prochlorococcus and N2 producing bacteria. In Rev.
- Johnson et al. Niche Partitioning Among Prochlorococcus Ecotypes Along Ocean-Scale Environmental Gradients. Science. 2006.
- Luo E et al. Double-stranded DNA virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water. ISME J. 2020.