

# COMPUTATIONAL RESEARCH in BOSTON and BEYOND SEMINAR

## Statistically Identifying Mechanisms of Phage-host Interactions in the Nahant Collection

JOY YANG

Massachusetts Institute of Technology

### ABSTRACT:

The Polz Lab maintains the Nahant Collection - 243 *Vibrio* strains challenged by 241 unique phage, all with sequenced genomes. This is the largest phylogenetically resolved host-range cross test available to date. These host strains match to 19 populations that coexist but are ecologically differentiated, and the phage fall into around 18 phylogenetically distinct groups with diverse infection strategies and morphologies.

This rich data set offers the opportunity to glean mechanistic insights from sequencing data, but doing so comes with a few challenges: (1) While the diverse population structure of phage and hosts is an interesting feature of the data, it means that statistical independence does not apply. Ignoring phylogenetic relationships can result in spurious correlations drowning out relevant signals. (2) A reasonable model should capture the generally lock-and-key nature of infection specificities due to protein interactions. For example, a specific methylase may evade a specific restriction modification system.

Here, in order to simplify computation, we first screen for genes of interest using generalized least squares to account for phylogenetic confounding. Then we build a multivariate model with statistical interaction terms that loosely represent putative interactions of host and phage genes. Finally, because our ultimate goal is to facilitate the process of generating testable hypotheses about biological mechanisms based on large-scale sequencing data, it is key to ensure that the model is interpretable and that the data is explorable by the general scientific community. To this end, we've written an interactive web visualization that anyone interested in the Nahant Collection will be able to access.

**FRIDAY, OCTOBER 5, 2018**

**1:00 PM – 2:00 PM**

**Building 32, Room 155**

STATA

*Pizza and beverages will be provided.*

<http://math.mit.edu/crib/>