

Rapid taxonomic categorization of short, abundant virus sequences for ecological analyses

Anna Sjodin¹, Michael Willig¹, Armando Rodríguez-Durán², and Simon Anthony³

¹University of Connecticut

²Inter American University of Puerto Rico

³Columbia University

October 4, 2023

Abstract

Public health concerns about recent viral epidemics have motivated researchers to seek novel ways to understand pathogen infection in native, wildlife hosts. With its deep history of tools and perspectives for understanding the abundance and distribution of organisms, ecology can shed new light on viral infection dynamics. However, datasets allowing deep explorations of viral communities from an ecological perspective are lacking. We sampled 1,086 bats from two, adjacent Puerto Rican caves and tested them for infection by herpesviruses, resulting in 3,131 short, viral sequences. Using percent identity of nucleotides and a machine learning algorithm (affinity propagation), we categorized herpesviruses into 43 operational taxonomic units (OTUs), to be used in place of species in subsequent ecological analyses. Herpesvirus metacommunities demonstrated long-tailed rank frequency distributions at all analyzed levels of host organization (i.e., individual, population, and community). Although 13 herpesvirus OTUs were detected in more than one host species, OTUs generally exhibited host specificity by infecting a single core host species at a significantly higher prevalence than in all satellite species combined. We describes the natural history of herpesvirus metacommunities in Puerto Rican bats and suggest that viruses follow the general law that communities comprise few common and many rare species. To guide future efforts in the field of viral ecology, hypotheses are presented regarding mechanisms that contribute to these patterns.

Hosted file

Sjodin_etal_27Sept2023.docx available at <https://authorea.com/users/657158/articles/670802-rapid-taxonomic-categorization-of-short-abundant-virus-sequences-for-ecological-analyses>