

VIBRANT: Automated recovery, annotation and curation of microbial viruses


Kristopher Kieft
Zhichao Zhou
Karthik Anantharaman

Video Byte

Keywords: Virome, virus, bacteriophage, metagenome, machine learning, auxiliary metabolism, software, community structure, microbe, Microbiome, metagenomics

Posted Date: June 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-37677/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Viruses that infect bacteria and archaea are abundant in human and environmental microbiomes. Their roles in manipulating, killing, and recycling microbes makes them key players in environmental processes and human health and disease, including inflammatory bowel diseases. In spite of their importance, the tools available for analyzing viral genomes are limited. Now, a new tool allows researchers to identify viruses and predict their functions using genomic data. VIBRANT (Virus Identification By iteRative ANnoTation), is the first software to use a hybrid machine learning and protein similarity approach. going beyond traditional limitations to maximize the identification of highly diverse viruses. In validation experiments with reference datasets, VIBRANT recovered higher-quality virus sequences and reduced false identification of non-viral genome fragments compared to other identification programs. Tests with Crohn's disease-associated viromes revealed a higher abundance of certain viruses and more dysbiosis-associated viral proteins than in healthy individuals. VIBRANT also considers annotation information, allowing researchers to identify links between viral functions and their effects on metabolic processes. This tool represents an important step toward using omics data to understand how viruses shape human health and disease.