

'Fishing expedition' nets nearly tenfold increase in number of sequenced virus genomes

August 14 2015, by Emily Caldwell

Using a specially designed computational tool as a lure, scientists have netted the genomic sequences of almost 12,500 previously uncharacterized viruses from public databases.

The finding doubles the number of recognized virus genera - a biological classification one step up from species - and increases the number of sequenced virus genomes available for study almost tenfold.

The research group studies viruses that infect [microbes](#), and specifically bacteria and archaea, single-cell microorganisms similar to bacteria in size, but with a different evolutionary history.

Microbes are essential contributors to all life on the planet, and viruses have a variety of influences on microbial functions that remain largely misunderstood, said Matthew Sullivan, assistant professor of microbiology at The Ohio State University and senior author of the study.

Sullivan partners with scientists studying microbes in the human gut and lung, as well as natural environments like soils and oceans. Most recently, he reported on the diversity of oceanic viral communities in a special issue of the journal *Science* featuring the Tara Oceans Expedition, a global study of the impact of climate change on the world's oceans.

"Virus-bacteria and virus-archaea interactions are probably quite important to the dynamics of that microbe, so if researchers are studying a microbe in a specific environment, they've been missing a big chunk of its interaction dynamics by ignoring the viruses," Sullivan said. "This work will help researchers recognize the importance of viruses in a lot of different microbes."

"In all of our studies, we're working with people who know the microbes well, and we help them decide how viruses might be helpful to the microbial system. The projects range from fundamental, basic science to applied medical science."

The research is published in the online journal *eLife*.

Finding a treasure trove of new virus genome sequences has opened the door to using those data to identify previously unknown microbial hosts, as well. These new possibilities are attributed to VirSorter, a [computational tool](#) developed by study lead author Simon Roux, a postdoctoral researcher in Sullivan's lab.

The sorter scoured public databases of sequenced [microbial genomes](#), looking for fragments of genomes that resembled [virus genomes](#) that had already been sequenced - for starters. VirSorter also "fished" for sequences by looking for genes known to help produce a protein shell that all viruses have, called a capsid.

"The idea is that bacteria don't use capsids or produce them, so any capsid gene should come from a virus," Roux said. The sorter then associated capsid genes with unfamiliar genes - those considered new, small or organized differently - that are unlikely to be produced by bacteria.

"None of these genomic features is really a smoking gun per se, but

combining them led to a robust detection of 'new' viruses - viruses we did not have in the database, but can identify because they have capsid genes and a viral organization," he said.

Using microbial genomes as a data source meant researchers could link newly identified virus sequences to the proper microbial host. The scientists then tried a reverse maneuver on the data to see if virus sequences alone could be used to identify unknown hosts - and this way of analyzing the sequences could predict the host with up to 90 percent accuracy.

"We can survey a lot of environments to find new viruses, but the challenge has been answering, who do they infect?" Sullivan said. "If we can use computational tricks to predict the host, we can explore that viral-host linkage. That's a really important part of the equation."

Though viruses are generally thought to take over whatever organism they invade, Sullivan's lab has identified a few viruses, called prophages, which coexist with their host microbes and even produce genes that help the host cells compete and survive.

Viruses can't survive without a host, and the most-studied viruses linked to disease are lytic in nature: They get inside a cell and make copies of themselves, destroying the cell in the process.

But the genome sequences revealed in this study suggest that there are many more prophage-like viruses that are different in one important respect: Their genome remains separate from their microbial hosts' genome.

"The extrachromosomal form of this virus type appears quite widespread, and virtually nobody is studying these kinds of viruses," said Sullivan, who also has an appointment in civil, environmental and

geodetic engineering. "That is a really different and largely unexplored phenomenon, and it's important to understand those [viruses](#)' ability to interact and tie into the function of those cells."

More information: *eLife*, [elifesciences.org/content/earl ...
15/07/22/eLife.08490](https://elifesciences.org/content/earl...15/07/22/eLife.08490)

Science, www.sciencemag.org/content/348/6237/1261498.full

Provided by The Ohio State University

Citation: 'Fishing expedition' nets nearly tenfold increase in number of sequenced virus genomes (2015, August 14) retrieved 4 October 2025 from <https://phys.org/news/2015-08-fishing-nets-tenfold-sequenced-virus.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
