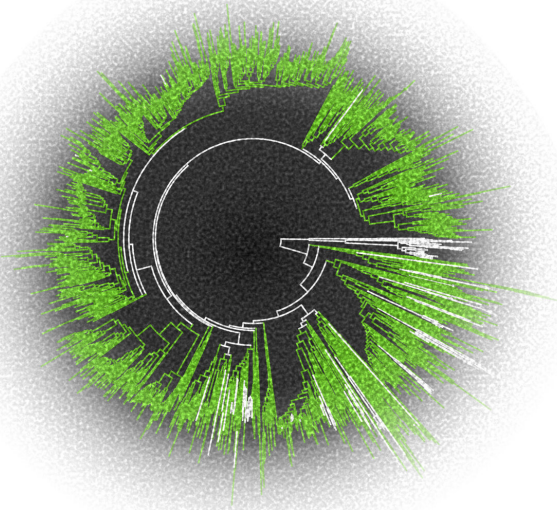


# Team significantly expands the global diversity of large and giant viruses

23 January 2020



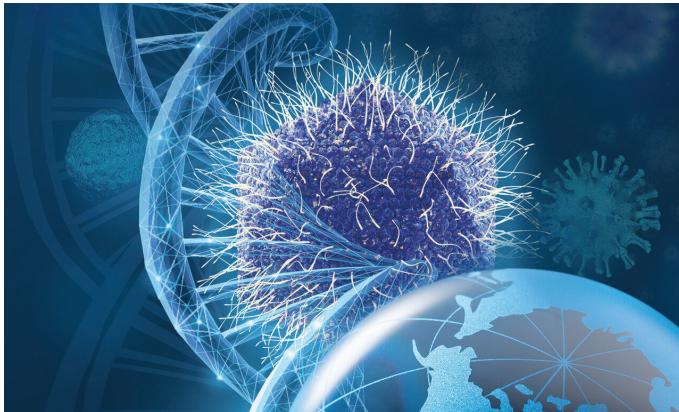
Metagenomic expansion of the diversity of the Nucleocytoplasmic Large DNA Viruses. The phylogenetic tree shows 2,074 giant virus metagenome-assembled genomes (in green) together with 205 previously published viral genomes (in white). Credit: Frederik Schulz

While the microbes in a single drop of water could outnumber a small city's population, the number of viruses in the same drop—the vast majority not harmful to humans—could be even larger. Viruses infect bacteria, archaea and eukaryotes, and they range in particle and genome size from small, to large and even giant. The genomes of giant viruses are on the order of 100 times the size of what has typically been associated with viruses, while the genomes of large viruses may be only 10 times larger. And yet, while they are found everywhere, comparatively little is known about viruses, much less those considered large and giant.

In a recent study published in the journal *Nature*, a team led by researchers at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), a DOE Office of Science User Facility located at Lawrence Berkeley National Laboratory (Berkeley Lab) uncovered a broad diversity of large and [giant viruses](#) that belong to the nucleocytoplasmic large DNA viruses (NCLDV) supergroup. The expansion of the diversity for large and giant viruses offered the researchers insights into how they might interact with their hosts, and how those interactions may in turn impact the host communities and their roles in carbon and other nutrient cycles.

"This is the first study to take a more global look at giant viruses by capturing genomes of uncultivated giant viruses from environmental sequences across the globe, then using these sequences to make inferences about the biogeographic distribution of these viruses in the various ecosystems, their diversity, their predicted metabolic features and putative hosts," noted study senior author Tanja Woyke, who heads JGI's Microbial Program.

The team mined more than 8,500 publicly available metagenome datasets generated from sampling sites around the world, including data from several DOE-mission relevant proposals through JGI's Community Science Program. Proposals from researchers at Concordia University (Canada), University of Michigan, University of Wisconsin-Madison, and the Georgia Institute of Technology focused on microbial communities from freshwater ecosystems, including, respectively, the northern Lakes of Canada, the Laurentian Great Lakes, Lake Mendota and Lake Lanier were of particular interest.



Art illustration capturing giant virus genomic diversity. This image complements a January 22, 2020, Nature paper led by researchers at the Department of Energy Joint Genome Institute uncovering a broad diversity of large and giant viruses that belong to the nucleocytoplasmic large DNA viruses (NCLDV) supergroup. The team reconstructed 2,074 genomes of large and giant viruses from more than 8,500 publicly available metagenome datasets generated from sampling sites around the world, and virus diversity in this group expanded 10-fold from just 205 genomes, redefining the phylogenetic tree of giant viruses. Credit: Zosia Rostomian/Berkeley Lab

### Sifting Out and Reconstructing Virus Genomes

Much of what is known about the NCLDV group has come from viruses that have been co-cultivated with amoeba or with their hosts, though metagenomics is now making it possible to seek out and characterize uncultivated viruses. For instance, a 2018 study from a JGI-led team uncovered giant viruses in the soil for the first time. The current study applied a multi-step approach to mine, bin and then filter the data for the major capsid protein (MCP) to identify NCLDV viruses. JGI researchers previously applied this approach to uncover a novel group of giant viruses dubbed "Klosneuviruses."

Previously known members of the viral lineages in the NCLDV group infect mainly protists and algae, and some of them have genomes in the megabase range. The study's lead and co-corresponding author Frederik Schulz, a research scientist in Woyke's group, used the MCP as a barcode to sift

out virus fragments, reconstructing 2,074 genomes of large and giant viruses. More than 50,000 copies of the MCP were identified in the metagenomic data, two-thirds of which could be assigned to viral lineages, and predominantly in samples from marine (55%) and freshwater (40%) environments. As a result, the giant virus protein space grew from 123,000 to over 900,000 proteins, and virus diversity in this group expanded 10-fold from just 205 genomes, redefining the phylogenetic tree of giant viruses.

### Metabolic Reprogramming a Common Strategy for Large and Giant Viruses

Another significant finding from the study was a common strategy employed by both large and giant viruses. Metabolic reprogramming, Schulz explained, makes the host function better under certain conditions, which then helps the virus to replicate faster and produce more progeny. This can provide short- and long-term impact on host metabolism in general, or on host populations impacted by adverse environmental conditions. Function prediction on the 2,000 new giant virus genomes led the team to uncover a prevalence of encoded functions that could boost host metabolism, such as genes that play roles in the uptake and transport of diverse substrates, and also photosynthesis genes including potential light-driven proton pumps. "We're seeing that this is likely a common strategy among the large and giant viruses based on the predicted metabolism that's encoded in the viral genomes," he said. "It seems to be way more common than had been previously thought."

Woyke noted that despite the number of metagenome-assembled genomes (MAGs) reconstructed from this effort, the team was still unable to link 20,000 major capsid proteins of large and giant viruses to any known virus lineage. "Getting complete, near complete, or partial giant virus genomes reconstructed from environmental sequences is still challenging and even with this study we are likely to just scratch the surface of what's out there. Beyond these 2,000 MAGs extracted from 8,000 metagenomes, there are still a lot of giant [virus](#) diversity that we're missing in the various ecosystems. We can detect a lot more

MCPs than we can extract MAGs, and they don't fit in the [genome](#) tree of viral diversity—yet."

"We expect this to change with not only new metagenome datasets becoming available but also complementary single-cell sorting and sequencing of viruses together with their unicellular hosts," Schulz added.

**More information:** Frederik Schulz et al, Giant virus diversity and host interactions through global metagenomics, *Nature* (2020). [DOI: 10.1038/s41586-020-1957-x](#)

Provided by Lawrence Berkeley National Laboratory

APA citation: Team significantly expands the global diversity of large and giant viruses (2020, January 23) retrieved 21 January 2021 from <https://phys.org/news/2020-01-team-significantly-global-diversity-large.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.*