



MICROBIOLOGY

Learning the diversity of giant viruses

Researchers identify a broad phylogenetic diversity of nucleocytoplasmic large DNA viruses by mining global-scale metagenomics datasets.

The nucleocytoplasmic large DNA viruses (NCLDV) typically have genome sizes ranging from a hundred kilobases to several megabases, and they can replicate in the host's nucleus and cytoplasm. Historically, isolation and cocultivation with hosts has been the main method for studying NCLDVs. Tanja Woyke, who leads the Microbial Genomics Program at the US Department of Energy's Joint Genome Institute (JGI), says that cocultivated viruses, in an "analog to bacterial and archaeal diversity, represent only a minor fraction of extant giant virus diversity." Woyke's group uses single-cell genomic approaches to uncultivated organisms: bacteria, archaea and now giant viruses.

Recent advances in metagenomic methods offer a way to discover previously undescribed NCLDVs from environmental samples in a cultivation-independent

manner. To study NCLDVs globally, Woyke's team surveyed NCLDV major capsid protein sequences across publicly available datasets. Thanks to the wealth of metagenome data and bioinformatic tools developed for data mining, binning and assembly, the researchers were able to reconstruct about 2,000 NCLDV genomes from metagenome datasets archived in JGI's Integrated Microbial Genomes and Metagenomes system, IMG/M. The resulting genomes have expanded the known phylogenetic diversity of NCLDV about 11-fold.

The group further studied the predicted functions encoded in the NCLDV genomes and showed that the identified viral genomes encode proteins with roles in energy generation and nutrient acquisition across different habitats, indicating NCLDVs may have shared strategies to affect host

metabolism. Woyke notes that "to see it [host metabolic reprogramming] at a larger scale and across the various new clades of NCLDVs across different habitats was most exciting."

Yet they also detected more pieces of large and giant viruses in unbinned metagenome sequences that they could not link to known virus lineages, suggesting that a large fraction of the viruses are still not taxonomically defined. "We are only at the very tip of the iceberg," says Woyke.

Lei Tang

Published online: 4 March 2020
<https://doi.org/10.1038/s41592-020-0778-z>

Research paper
Schulz, F. et al. Giant virus diversity and host interactions through global metagenomics. *Nature* <https://doi.org/10.1038/s41586-020-1957-x> (2020).


COMMUNICATIONS
BIOLOGY

A new open access biological sciences journal from Nature Research.

Communications Biology publishes high-quality research, reviews and commentary in all areas of the biological sciences.

Research papers published by the journal represent significant advances for a specialized area of research.

Submit your research and benefit from:

- **Fast decisions and easy submission**
- **Rigorous and balanced peer review**
- **High Nature Research editorial standards**
- **Global visibility, fully open access**

nature.com/commsbio

 @CommsBio

natureresearch

A82211