

# Lawrence Berkeley National Laboratory

LBL Publications

## Title

Extreme dimensions — how big (or small) can tailed phages be?

## Permalink

<https://escholarship.org/uc/item/2p7437kd>

## Journal

Nature Reviews Microbiology, 19(7)

## ISSN

1740-1526

## Authors

Berg, Maureen

Roux, Simon

## Publication Date

2021-07-01

## DOI

10.1038/s41579-021-00574-z

Peer reviewed

# Extreme dimensions — how big (or small) can tailed phages be?

Maureen Berg & Simon Roux

5 This month's Genome Watch highlights the search for unusually large (or small) tailed phages driven by metagenomics.

10 Bacteriophages (phages) that is, viruses that infect bacteria, represent an extremely diverse, yet under-characterized, group of viruses. Although most known phages harbour genomes that are shorter than 200kb packaged into capsids with a diameter under 100 Å, more and more 'extremely large' phages are being discovered. Early reports of phages with genomes larger than 200kb defined them as 'jumbo phages' (Ref. 1). Phylogenetic analysis of these jumbo phages revealed that they form distinct and cohesive groups with multiple independent origins. This suggested that jumbo phages are not merely mistakes or aberrations of smaller phages that got too big, but that large genome size is generally a stable trait. In fact, a larger genome can be advantageous; almost all jumbo phages encode their own specific transcription factors, at least some parts of the replication machinery, and many also have their own tRNA genes, which enables increased independence from their host. Large genomes also provide more flexibility in terms of transcription strategy; smaller phage genomes tend to have a strict modular genome structure to increase efficiency, whereas jumbo phage genomes may be much less organized in the absence of a strong selection pressure to maintain a compact genome.

50 Early reports were based on a handful of isolated jumbo phages, but a recent study reported more than 350 (near-)complete large phage genomes assembled from metagenomes and revealed how extensive and pervasive large phages really are in nature<sup>2</sup>. These metagenome-derived large phages, the largest harbouring genomes of  $\geq 700$ kb, spanned across multiple clades, and were identified in various environments from freshwater lakes to human gut. Functional annotation further confirmed some of the initial jumbo phage work, showing that these large phages encode a large diversity of genes, including many involved in replication. Several CRISPR loci were detected in these large phages, although most lack the CRISPR machinery for spacer acquisition and tend to only include the spacer arrays. Notably, some large phages in the Biggiephage clade encode a special family of hypercompact Cas proteins that can target foreign DNA, such as that from competing phages, which suggests that some selection pressure towards gene coding optimization and compactness remains<sup>3</sup>.

85 More than a hundred years after the discovery of bacteriophages, we have still barely begun to characterize the full diversity, ecology and evolutionary history of jumbo phages. This probably results from several methodological biases, which render these ultra large phages challenging to

isolate and study. For instance, jumbo phages are usually retained by 0.2µm filters typically used to purify phage particles prior to isolation, and they can have a limited ability to diffuse through agar overlays<sup>1</sup>. Meanwhile, only recently have sequencing and bioinformatic technologies enabled the assembly of such large (>500kb) genomes from metagenomes. By providing reference sequences and ecological distribution for these clades of large phages, these metagenome assemblies now enable the establishment of modified and targeted isolation protocols to eventually obtain a better representation of large phages in isolate collections.

Strikingly, the same is true at the other end of the genome size spectrum, where the existence of ultra small dsDNA phages (<10kb) has been posited based on theoretical analysis of capsid structures and suggested by corresponding metagenome circular contigs, but not confirmed yet via laboratory cultivation<sup>4</sup>. Metagenomics can thus guide us in the exploration of the ultra large and ultra small phages to eventually better understand the eco-evolutionary and biophysical constraints operating on double-stranded DNA phage genomes and capsids.

Maureen Berg and Simon Roux  
DOE Joint Genome Institute, Berkeley, CA, USA.  
\*e-mails: mberg@lbl.gov, sroux@lbl.gov  
<https://doi.org/10.1038/s41579-XXX-XXXX-X>

1. M Iyer, L., Anantharaman, V., Krishnan, A., Burroughs, A. M. & Aravind, L. Jumbo phages: a comparative genomic overview of core functions and adaptations for biological conflicts. *Viruses*

---

13, 1–42 (2021).

2. Al-Shayeb, B. *et al.* Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).
3. Pausch, P. *et al.* Crispr-cas  $\Phi$  from huge phages is a hypercompact genome editor. *Science* **369**, 333–337 (2020).
4. Luque, A., Benler, S., Lee, D. Y., Brown, C. & White, S. The missing tailed phages: prediction of small capsid candidates. *Microorganisms* **8**, 1–18 (2020).

#### **Competing interests**

The authors declare no competing interests.

Credit: Philip Patenall/Springer Nature Limited

---