Giant viruses come of age
Matthias G Fischer

Viruses with genomes up to a few megabases in length are a common occurrence in nature, even though they have escaped our notice until recently. These giant viruses infect mainly single-celled eukaryotes and isolation efforts concentrating on amoebal hosts alone have spawned hundreds of viral isolates, featuring viruses with previously unseen virion morphologies and the largest known viral genomes and particles. One of the challenges that lie ahead is to analyze and categorize the available data and to establish an approved classification system that reflects the evolutionary relationships and biological properties of these viruses. Extensive sampling of Acanthamoeba-infecting mimiviruses and initial characterization of their virophage parasites have provided a first blueprint of the genetic diversity and composition of a giant virus clade that will facilitate the taxonomic grouping of these fascinating microorganisms.

Address
Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany
Corresponding author: Fischer, Matthias G
(mfischer@mpimf-heidelberg.mpg.de)

New additions to the NCLDV tree from past and present
The largest and most diverse group of DNA viruses infecting eukaryotic hosts are the nucleocytoplasmic large DNA viruses (NCLDV) [11,12] that currently include seven families (Asociiridae, Asfarviridae, Iridoviridae, Marsilleviridae, Mimiviridae, Phycodnaviridae, and Poxxviridae), and for which a novel viral order named ‘Megavirales’ has been suggested [13]. The family relationships among NCLDV members are represented by the phylogenetic reconstruction of DNA polymerase B sequences shown in Figure 1. The genome lengths in this group of large DNA viruses vary considerably, from ≈100 kb in the smallest iridoviruses up to ≈2500 kb in the pandoraviruses, which feature the largest viral genomes reported to date. Among the 2556 predicted proteins encoded by the provisionally named ‘Pandoravirus salinus’, only ≈7% have detectable similarity to previously described sequences [14**, including other NCLDVs. Strikingly, and in contrast to the large pseudo-icosahedral particles produced by AMPV (most giant virus particles are not perfectly icosahedral due to the presence of a unique vertex), pandoravirions have a novel type of morphology, consisting of 1 μm by 0.5 μm ovoid particles with a single apical pore for genome delivery (Figure 2e). Another difference to AMPV, which replicates in the host cytoplasm, is the finding that pandoraviruses have a nuclear phase in their replication cycle [14**]. Comparative genomic analyses confirmed that pandoraviruses are de facto unrelated to mimiviruses, but instead share a few genes with algae-infecting phycodnaviruses [14**,15]. Whereas the origins of these viral lineages are still deeply shrouded in the fog of evolution, two surprising studies revealed the longevity of individual giant virus particles. From Siberian permafrost soil that was dated to be 30 000 years old, two novel types of giant viruses were isolated on Acanthamoeba hosts. Remarkably, these virions had retained their infectivity during the long hibernation allowing them to infect and lyse the amoeba once thawed. One of these viruses, provisionally named ‘Pithovirus sibericum’, features an elongated particle with a length of 1.5 μm and a diameter of 0.5 μm (Figure 2f),
which at first glance resembles pandoravirions, but is in fact structurally unrelated \[16\]. The pithovirions are close in dimensions to an *Escherichia coli* cell and represent the current size record-holder among viral containers \[16\]. However, despite its enormous particle size and a shape similar to pandoravirions, the ‘P. sibericum’ genome is a ‘mere’ 610 kb and only about a third of its 467 proteins have recognizable counterparts in public databases. Phylogenetically, this virus shows a remote relatedness to irido- and ascoviruses. The second novel type of giant virus that was reawakened from permafrost soil was provisionally called ‘Mollivirus sibericum’ and produces virions with previously unseen morphology that are spherical, 0.6 μm in diameter, and covered with a
Figure 2

Representative particles of six different types of giant viruses infecting amoebae. (a) 'Port-Miou virus', a yet-to-be classified marseillevirus [69]. (b) An unclassified 'faustovirus'. (c) 'Megavirus chilensis', a yet-to-be classified mimivirus. (d) 'Mollivirus sibericum', unclassified. (e) 'Pandoravirus salinus', unclassified. (f) 'Pithovirus sibericum', unclassified. All scale bars are 200 nm. Electron micrographs in panels (a), (c)-(f) were kindly provided by Chantal Abergel (Aix-Marseille University); the EM image in panel (b) was kindly provided by Jacques Bou Khalil & Bernard La Scola (Aix-Marseille University).

'hairy' tegument (Figure 2d) [17]. Its 651 kb genome contains 523 predicted protein-coding genes, 35% of which have database homologs. Among the latter is a significant fraction (83 genes) with phylogenetic ties to pandoraviruses; however, this fraction represents only 16% of the 'M. sibericum' coding potential. The interested reader is referred to a recent review [18] for a side-by-side comparison of mimi-, pandora-, pitho-, and mollivirus. Another recently described addition to the NCLDV group is 'faustovirus', which was isolated on
Vermamoeba (formerly Hartmanella) vermiciformis, an amoeba that is commonly found in human environments [19*]. Although the ‘faustovirus’ capsids have the typical icosahedral symmetry (Figure 2b) also found in the amoeba-infecting mimiviruses or marseilleviruses (Figure 2a,c), their 466 kb genome encodes several proteins with phylogenetic affinity to asfarviruses (African swine fever virus). So far, a dinoflagellate-specific virus infecting Heterocapsa circularisquama was the only known asfar-like virus with a non-vertebrate host [20]. A common theme for the viruses described above is that once a new prototype of giant virus has been characterized, subsequent reports on related viruses appear rapidly. This was the case for pandoraviruses [21–23], ‘P. sibericum’ [23,24], and ‘faustovirus’ [19*,25], proving that giant viruses are widespread microbivorous members in various environments, even though they remained hidden from our eyes until recently.

Deep sampling of mimiviruses and the challenge of viral classification

In particular the decade-long research focus on APMV has led to significant improvements in isolation procedures [26*]. For instance, mimiviruses are now routinely isolated from various sources such as oysters [27] or leeches [28] using an amoebal reporter assay. More than 100 strains have been reported so far [22,29–31], increasingly from Brazilian waters [22,32,33], and many more are awaiting their characterization in laboratory freezers [26*,27]. Reports of giant viruses associated with humans suggest that mammals can provide suitable habitats for these viruses [34–36]. Although studies of APMV infecting human macrophages [37] and potential links between APMV and cases of pneumonia [38] have been published, conclusive evidence are missing to establish giant viruses as human pathogens. However, giant viruses are probably more common in the human microbiome than initially assumed, and they can for instance infect protozoa that cause human diseases such as amoebic keratitis [21,34].

The enormous success in isolating giant viruses on the Acanthamoeba system has interesting consequences. We are currently facing a phylogenetic distribution where particular branches of the NCLDV tree experience high saturation, whereas giant viruses infecting non-amoebal hosts have much fewer representatives (Figure 1). This offers unique challenges and opportunities. On the one hand, we still do not know how many major clades of giant viruses there are. Given the recent discoveries of pandora-, pitho-, moll-, and faustoviruses (all of which were isolated on amoebae), we can safely assume that exploration of other protist hosts is guaranteed to uncover a plethora of new giant viruses and that so far, we have only seen the tip of the iceberg when it comes to giant virus diversity. A related question concerns the distinctiveness of particular groups of viruses in phylogenetic reconstructions, and whether increased sampling will eventually transform a well-contoured phylogenetic tree into a continuum, a phylogenetic mist. On the other hand, the genetic diversity of mimivirus isolates may provide a preview of what to expect for other viral clades as well. Despite the large number of available isolates, mimiviruses still cluster into three distinct clades (based on DNA polymerase B phylogeny) that are currently designated lineages A, B, C (Figure 1). Lineage A comprises the original APMV strain [1], the type member of lineage B is ‘Acanthamoeba polyphaga moumouivirus’ [39], and ‘Megavirus chilensis’ is the type member of lineage C [40]. Based on this situation, the Mimiviridae could be subclassified into three genera, one for each lineage. But there are more hurdles to overcome. One of the biggest conceptual challenges in the field is to establish a meaningful and sustainable taxonomy for giant viruses, a topic that was debated intensively at the 2nd Ringberg Symposium on Giant Virus Biology, which was held near Tegernsee, Germany in November 2015 (Figure 3). A particular source of confusion stems from the term ‘Megavirus’, which was introduced with the publication of ‘M. chilensis’ to describe viruses with >1 Mbp genomes [40], but was thenceforth used to refer to a steadily increasing group of viruses encompassing the already established family Mimiviridae as well as more distantly related algae-infecting viruses with much smaller genomes (e.g. Aurococcus anophagefferens virus [41*] and Phaeocystis globosa virus 16T [42*]). The suffix -viridae is used by the International Committee on the Taxonomy of Viruses (ICTV) as a family taxon designation and implies that the ‘Megaviridae’ have official taxonomic status. Unfortunately, this is not the case, and in the absence of a taxonomic proposal (TaxoProp), the purpose of which is to define the membership criteria for a proposed taxon, the scope of the ‘Megaviridae’ remains unclear and its use should be discouraged. Instead, separate families will probably have to be established for algal viruses with mimivirus-like properties as well as for Cafeteria roenbergensis virus (CroV [43]), a virus belonging to one of the two currently approved species of the Mimiviridae (Cafeteria roenbergensis virus in the genus Cafeteriaivirus and Acanthamoeba polyphaga mimivirus in the genus Mimivirus). But also other parts of the NCLDV tree are in need of a spring-cleaning. One discussion topic initiated by J-M Claverie during the 2nd Ringberg Symposium dealt with the taxonomy of members of the Phycodnaviridae, a family of algae-infecting viruses with large, though mostly not giant, DNA genomes [44]. This family includes six genera, whose members have genetic differences that are sufficiently large to justify a status upgrade for these genera to the family level. In particular the grouping of coccolithoviruses (viruses infecting Emiliana huxleyi [45]) with other phycodnaviruses has become questionable due to a lack of demonstrable monophyly (see Figure 1). Overall, the classification of giant viruses into meaningful taxa will have to be based on more than just DNA polymerase B as a genetic proxy. The identification of a set of conserved genes that are shared...

by many NCLDVs [11,46] and are presumed to represent the minimal gene set of the hypothesized NCDLV ancestor [47], provides several targets for taxonomic classification. Additional markers such as the DNA repair gene MutS [48] and asparagine synthetase [49], which are present in a subset of NCLDVs, will prove useful in resolving certain subgroups. Although classification should mainly be based on comparative genomics, other distinguishing features such as host range, capsid structure, mode of genome delivery, and nuclear involvement in the replication cycle may provide additional criteria for classification. Even parasitism by virophages may aid in establishing a taxonomic scheme for giant viruses.

Virophage parasitism of giant viruses
Virophages are small DNA viruses of the family Laciviridae [50] that depend on and parasitize giant viruses of the family Mimiviridae as well as their smaller relatives (labeled ‘extended Mimiviridae’ in Figure 1). Virophages have circular or linear 15–30 kbp dsDNA genomes and icosahedral particles with a diameter of 50–80 nm [51,52]. Initially discovered with a lineage A mimivirus isolate [53], virophages have since been found in association with lineage C mimiviruses [52], algal viruses [42], and giant viruses infecting heterotrophic nanoflagellates [54]. Whereas isolation of virophages in laboratory cultures proves difficult owing to the need for simultaneously cultured cellular and giant virus hosts, computational discovery of virophages in metagenomics datasets hints at a wide geographic distribution and a diverse genetic landscape for these viruses [55–58,59,60]. So far, virophages have not been found for giant viruses outside the Mimiviridae and their algae-infecting relatives. Although the biology of virophages is largely unknown, it is assumed that virophages depend on the cytoplasmic transcription apparatus encoded by the co-infecting giant virus [54,61,62]. Based on this model, virophages are expected to exclusively parasitize viruses that are transcriptionally independent from the host cell, permitting a cytoplasmic replication cycle of the virus. Virophages that have been isolated so far (Figure 1) comply with this prediction. Interestingly, no virophages associated with poxviruses have been found to date, despite their cytoplasmic replication mode. This may be due to insufficient sampling, or additional factors may be needed to support a virophage infection. For instance, the presence of fibrils on the outside of mimivirions is essential for successful replication of Sputnik virus (genus Sputnikirus, species
**Mimivirus-dependent virus Sputnik**, probably because this virus enters the amoebal host cell as a composite by attaching to the mimivirus fiber [63]. The three mimivirus lineages differ in their susceptibility to infection by virophages. Whereas Sputnik virus replicates with mimiviruses of all three lineages, Zamilon virus (genus *Sputnikvirus*, species *Mimivirus-dependent virus Zamilon*) is only able to grow in *Acanthamoeba* hosts co-infected with lineage B or C mimiviruses [52*]. Recently, this difference has been attributed to a proposed CRISPR-Cas-like system in mimiviruses called MIMIVIRE [64**]. According to this study, lineage A mimiviruses contain in their genomes four repeats of a small (15 nt) oligonucleotide stretch that has a perfect match in the Zamilon genome. In contrast, lineage B and C mimiviruses have only one copy of this Zamilon-specific sequence in a non-orthologous location. Furthermore, the genes in the vicinity of the repeats in lineage A mimiviruses encode a nuclelease and a helicase that could be involved in sequence-specific Zamilon genome degradation. Silencing of these genes rendered APMV susceptible to Zamilon infection [64**]. Future studies will show whether virophage-specific resistance mechanisms are also present in other giant viruses.

**Future perspective**

Giant viruses are highly diverse in genomic composition and particle morphology. Giant viruses of six phylogenetic clades were isolated in recent years on amoebae alone (mimi-, marseill-, pandora-, pitho-, moll-, and fausto-viruses) and viruses of additional clades are likely to exist. The exploration of giant viruses infecting other hosts, in particular photosynthetic protists, has just begun [41*, 42*, 65, 66], and metagenomic surveys [49*, 67, 68] foreshadow a wealth of novel viruses to crowd the NCLDV tree in the near future. This diversity poses a significant challenge for the scientific community to establish a biologically meaningful classification system that permits easy access to an ever-increasing dataset of giant viruses. A proper classification of the three mimivirus lineages would mark a good starting point for such a taxonomic endeavor.

**Acknowledgements**

The author gratefully acknowledges support from the Max Planck Society and thanks J.H. Kuhn, J.-M. Claverie and J. Reinstein for helpful comments on the manuscript. Many thoughts presented in this paper were derived from or inspired by presentations and comments by J. Van Etten, J.-M. Claverie, S.W. Wilhelm and other participants of the 2nd Ringberg Symposium on Giant Virus Biology.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

15. The first giant viruses with amphora-shaped particles and the longest known viral genomes.
18. First representative of a new clade of giant viruses from Siberian permafrost.
20. Yet another virus recovered from ancient permafrost that represents a new giant virus clade.
A giant virus related to African swine fever virus isolated on Vermamoeba vermiformis.


A streamlined technique for isolating giant viruses on amoebae.


This study describes an algae-infecting virus that shares several features with mimiviruses despite having a much smaller genome.


First genome analysis of a cultured mimi-like algal virus including an associated virophage.


Computational discovery of new giant viruses by metagenomic data mining.


This paper evaluated virophages and satellite viruses from ‘subviral agents’ to legitimate viruses.


The first finding of virophage genomes integrated in a eukaryotic genome.


64. Levasseur A, Bekliz M, Chabrière E, Pontarotti P, La Scola B, Raoult D: MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. Nature 2016 http://dx.doi.org/10.1038/nature17140.

This paper describes a CRISPR-Cas-like system against virophages in APMV.


