Communicative & Integrative Biology 6:6, e25685; November/December 2013; © 2013 Landes Bioscience

Giant virus in the sea

Extending the realm of Megaviridae to Viridiplantae

Jean-Michel Claverie*

Structural and Genomic Information Laboratory (IGS-UMR7256 and Mediterranean Institute of Microbiology (FR3479)); Centre National de la Recherche Scientifique; Aix-Marseille University; Marseille, France

The viral nature of the first "giant virus," Mimivirus, was realized in 2003, 10 y after its initial isolation from the water of a cooling tower in Bradford, UK. Soon after its genome was sequenced, the mining of the Global Ocean Sampling environmental sequence database revealed that the closest relatives of Mimivirus, only known to infect Acanthamoeba, were to be found in the sea. These predicted marine Mimivirus relatives remained elusive until 2010, with the first genomic characterization of a virus infecting a heterotrophic unicellular eukaryote, the microflagellate grazer *Cafeteria roenbergensis*. The genome analysis of a virus (PgV) infecting the common unicellular algae *Phaeocystis globosa* now shows that it is a bona fide member of the Mimivirus family (i.e., the Megaviridae), extending the realm of these giant viruses to abundant blooming phytoplankton species. Despite its smaller genome size (460 kb encoding 434 proteins), PgV exhibits the most intriguing feature of the previously characterized Megaviridae: an associated virophage. However, the 19-kb virophage genome, devoid of a capsid gene, is packaged in the PgV particle and propagated as a "viral plasmid," the first ever described. The PgV genome also exhibits the duplication of "core genes," normally present as single copies and a putative new type of mobile element. In a DNA polymerase phylogeny including representatives of the three cellular domains, PgV and the other Megaviridae cluster into their own clade deeply branching between domains Archaea and Eukarya domains, thus exhibiting the topology of a fourth domain in the Tree of Life.

Keywords: prymnesiophyceae, haptophyta, virophage evolution, viral plasmid, fourth domain, Tree of Life

Submitted: 06/17/13

Accepted: 07/09/13

Citation: Claverie JM. Giant Virus in the sea: Extending the realm of Megaviridae to Viridiplantae. Commun Integr Biol 2013; 6: e 25685; http://dx.doi.org/10.4161/cib.25685

*Correspondence to: Jean-Michel Claverie; Email: Jean-Michel.Claverie@univ-amu.fr

Addendum to: Santini S, Jeudy S, Bartoli J, Poirot O, Lescot M, Abergel C, et al. Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. Proc Natl Acad Sci USA 2013; 110:10800-5; PMID:23754393; http://dx.doi.org/10.1073/pnas.1303251110

"Giant viruses"[1] were initially defined on the basis of their particle (virion) size—more than 0.5 μm in diameter—(making them observable under a light microscope) incorporating a DNA genome more than a million base pairs (Mb) in length, encoding about 1000 proteins (hence the proposed name "Megaviridae" for their family). Most of them replicate in Acanthamoeba including 4 that have been fully sequenced[2-5] (**Table 1**). Their replication entirely proceed in the host cytoplasm through the initial building of a large virion factory within which their genome is replicated and transcribed, and from the periphery of which new particles emerge at a later stage.[6-8] Long after the bioinformatics prediction that Mimivirus relatives should be abundant in the sea,[9,10] the first Megaviridae family member not infecting an Acanthamoeba (classified in the Amoebozoa, Unikonta) was identified in *Cafeteria roenbergensis*, a heterotrophic unicellular plankton species, belonging to a very distant phylum (Heterokontophyta, Bikonta).[11] Despite its smaller genome (about 700 kb), Cafeteria roenbergensis virus (CroV)[12] shared a large proportion of best matching orthologous proteins with Mimivirus, and possessed many of its unique features (**Table 1**), including its own type of virophage,[13] a specific type of mismatch repair protein (MutS7),[14] and an amino-acyl tRNA synthetase, a type of enzyme normally specific of cellular

Table 1. Specific features in the 6 fully sequenced genomes Megaviridae

Virus	Genome size (GenBank ID)	Virophage	tRNA ligase	DNA Mismatch Repair (MutS7)	Bifunctional Thy/DHFR	Intein (DNA pol)	Host taxonomy
Megavirus chilensis	1,259,197 (NC_016072)	+	7	+	+	+	Acanthamoeba Amoebozoa Unikonta
Moumouvirus	1,021,348 (NC_020104)	+	6	+	+	+	Acanthamoeba
Mimivirus	1,181,549 (NC_014649)	+	4	+	+	+	Acanthamoeba
Mimivirus (M4)	981,813 (JN036606)	?	3	+	+	+	Acanthamoeba
Cafeteria roenbergensis virus	617,453 (NC_014637)	+	1	+	+	+	Heterokontophyta AH/SAR megagroup Bikonta
Phaeocystis globosa virus (16T)	459,984 (NC_021312)	+	No	+	No	No	Haptophyta AH/SAR megagroup Bikonta

organisms.[15] The genome analysis of PgV-16T,[16] a virus known to regulate the population of the bloom-forming microalgae *Phaeocystis globosa*[17] (Haptophyta, Bikonta) now indicates that the Megaviridae family includes viruses infecting dominant phytoplankton species. With its 459 kb-genome encoding 434 proteins, PgV-16T is the most complex virus known to infect a photosynthetic organism.

Virophages might be common in the viral world

Although global gene content and detailed sequence similarity comparisons with the previously described Megaviridae members clearly classify PgV within the family (including the partially sequenced Organic Lake Phycodnaviruses (OLPV), Chrysochromulina erecina virus, and Phaoecystis pouchetti virus),[4] the finding of a PgV-associated virophage came as a surprise. Eight complete virophage genomes have now been described, but only 3 correspond to identified and isolated "hosts"—the 2 closely related "Sputnik" 1 and 2[18,19] infecting Mimivirus, and "Mavirus" infecting CroV.[20] The PgV virophage (PgVV)[16] is the third, and presumably not the last, of a rapidly growing series.[21-23] Except for their genome size in the 20 kb range, these new types of satellites viruses have little in common in terms of gene content, although they all code for a major capsid protein and one DNA primase.[18-22] Consistent with the fact that it was never observed in PgV-infected cultures, PgVV is the first example of a virophage lacking the information to make a capsid protein, the essential building block of a virus particle. Once multiplied in the PgV virus factory, the PgVV genome is thus packaged (in multiple copies[16]) alongside the PgV genome, and propagated through the PgV virion, either as an integrated or free viral "plasmid."[16] The precise molecular structure(s) that can be adopted by the PgVV genome are under further studies.

The finding of a virophage associated to PgV is already teaching us the important lesson that virophages are not solely associated to DNA viruses with micron-sized particles and 1 Mb genome sizes, but can be found with large DNA viruses of more reasonable (Poxvirus-like?) proportion. It is thus likely that many of such associations have been overlooked in the past, and that virophages might have played a fundamental role in the evolution of many more viruses than just the Megaviridae. If this is true, they could be (and have been) the main vehicle of gene transfers between Eukaryotic viruses, and indirectly between these viruses and their hosts. They may also be responsible for the sporadic occurrences of mobiles elements such as self-splicing introns, inteins, and transpovirons.[19]

Another lesson is that, like all parasites, the virophages are submitted to the irreversible phenomenon of reductive evolution, condemning them to disappear as individual biological entities, eventually saving some of their genes by integrating them into the genome of their companion virus, themselves undergoing a similar process vis-à-vis their cellular host.

PgV clusters with the previously described Megaviridae in an apparent "fourth domain" in the tree of life

The first published phylogenetic tree including Mimivirus 2 and using a concatenation of 7 universally conserved protein sequences, already pointed out that: "we could now build a tentative tree of life, within which Mimivirus appears to define a new branch distinct from the three other domains."[2] Elsewhere in the same article, Mimivirus was also shown to be part of the broad family of the nucleocytoplasmic large DNA viruses (NCLDV), branching near the middle of the previously defined Iridovirus, Phycodnaviruses, Poxviruses, and Asfarviruses lineages. Put together, these 2 results suggested in a subliminal way that all the large DNA viruses were in fact defining a domain distinct from the 3 established cellular domains. As additional genomes of Megaviridae became available, molecular phylogenies computed with an increasing number of universal proteins associated to

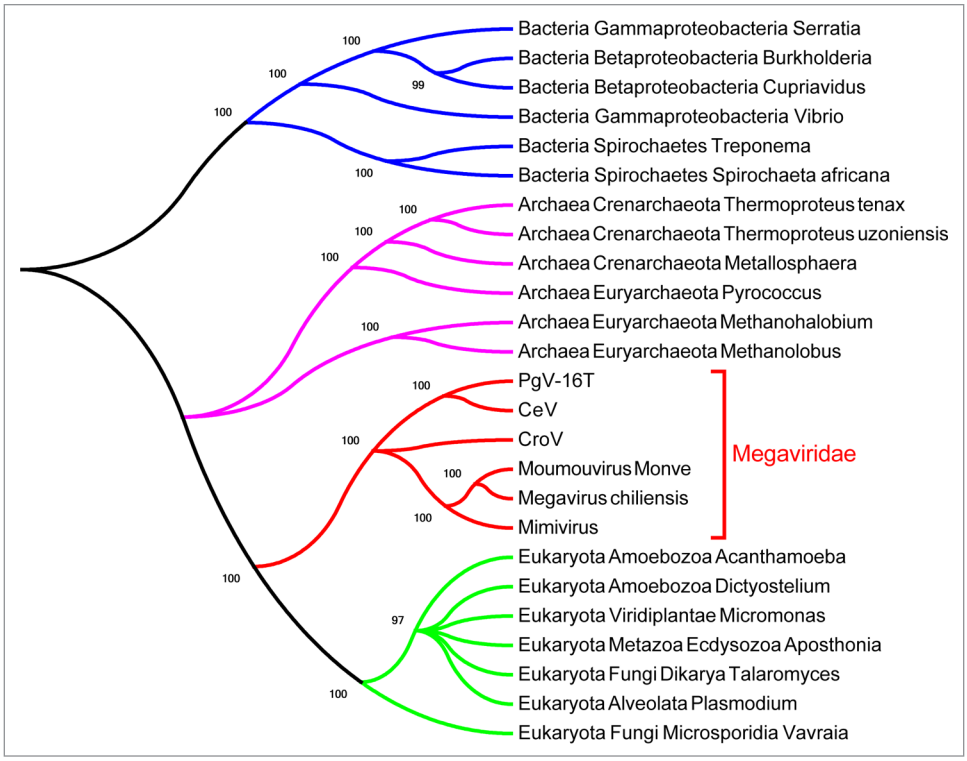

Figure 1. Simplified Tree of Life including the 6 fully sequenced megaviridae. The tree was produced using the default option on the MAFFT server (http://mafft.cbrc.jp) from the multiple alignment of 25 DNA polymerase B sequences (510 ungapped positions, excluding the inteins). Branches with bootstrap values < 80 are collapsed. Despite infecting eukaryotic hosts from vastly divergent phyla, the viruses do not show any phylogenetic affinity with a specific eukaryotic group, and cluster (in red) separately from the 3 cellular domains: Eukarya (green), Archaea (purple), and Eubacteria (blue). This strongly supported topology suggests that the common ancestor of the Megaviridaelargely predated the radiation of the Eukaryotes (CeV: *Chrysochromulinaericina* virus, unpublished).

basic functions (DNA clamp loaders, Ribonucleotide reductases, Aminoacyl-tRNA synthetases, DNA polymerases) kept clustering the megaviruses in their own clade, clearly separate from the cellular domains, thus contributing an additional domain rooted in between Archaea and Eukarya in the Tree of Life.[4,24,25] Despite receiving increasing support from various authors,[26-28] others remain strongly opposed to this view.[29,30] In agreement with the notion of a fourth domain of life anchored by the largest DNA viruses, the DNA polymerase of PgV nicely clusters with the other Megaviridae homologs, exhibiting no affinity with any of the main cellular lineages (**Fig. 1**). Based on the presence of a vestigial protein translation system, I previously argued that the genome of today's giant viruses derived from an ancestral cellular organism through the irreversible process of reductive evolution experienced by all parasites. DNA viruses exhibiting a whole range of genomic complexity could have been generated through this continuous process: viruses would have lost translation first (the ribosome), then transcription (the RNA polymerase), then DNA replication (The DNA polymerase), following an evolutionary scenario whereby they become increasingly dependent from their host.[15] The vast range of size and complexity among today's DNA viruses might thus be the results of differences in their evolution rates (the largest ones experiencing the least evolutionary pressure). In this context, the lack of affinity of the giant viruses with any of today's cellular domains, as well as the huge proportion of the viral genes without cellular homolog, would suggest that the lineage of the ancestral cellular organism that gave rise to giant viruses became extinct as a cellular life form, and only managed to survive as a parasitic "fourth domain." The recently described giant Pandoravirus might represent an independent instance of the same scenario.[31]

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Claverie JM, Ogata H, Audic S, Abergel C, Suhre K, Fournier PE. Mimivirus and the emerging concept of "giant" virus. Virus Res 2006; 117:133-44; PMID:16469402; http://dx.doi.org/10.1016/j.virusres.2006.01.008
2. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of Mimivirus. Science 2004; 306:1344-50; PMID:15486256; http://dx.doi.org/10.1126/science.1101485
3. Boyer M, Azza S, Barrassi L, Klose T, Campocasso A, Pagnier I, et al. Mimivirus shows dramatic genome reduction after intraamoebal culture. Proc Natl Acad Sci USA 2011; 108:10296-301; PMID:21646533; http://dx.doi.org/10.1073/pnas.1101118108
4. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc Natl Acad Sci USA 2011; 108:17486-91; PMID:21987820; http://dx.doi.org/10.1073/pnas.1110889108
5. Yoosuf N, Yutin N, Colson P, Shabalina SA, Pagnier I, Robert C, et al. Related giant viruses in distant locations and different habitats: Acanthamoeba polyphaga moumouvirus represents a third lineage of the Mimiviridae that is close to the megavirus lineage. Genome Biol Evol 2012; 4:1324-30; PMID:23221609; http://dx.doi.org/10.1093/gbe/evs109

6. Claverie JM, Abergel C. Mimivirus and its virophage. Annu Rev Genet 2009; 43:49-66; PMID:19653859; http://dx.doi.org/10.1146/annurev-genet-102108-134255
7. Mutsafi Y, Zauberman N, Sabanay I, Minsky A. Vaccinia-like cytoplasmic replication of the giant Mimivirus. Proc Natl Acad Sci USA 2010; 107:5978-82; PMID:20231474; http://dx.doi.org/10.1073/pnas.0912737107
8. Mutsafi Y, Shimoni E, Shimon A, Minsky A. Membrane assembly during the infection cycle of the giant Mimivirus. PLoS Pathog 2013; 9:e1003367; PMID:23737745; http://dx.doi.org/10.1371/journal.ppat.1003367
9. Ghedin E, Claverie JM. Mimivirus relatives in the Sargasso sea. Virol J 2005; 2:62; PMID:16105173; http://dx.doi.org/10.1186/1743-422X-2-62
10. Monier A, Claverie JM, Ogata H. Taxonomic distribution of large DNA viruses in the sea. Genome Biol 2008; 9:R106; PMID:18598358; http://dx.doi.org/10.1186/gb-2008-9-7-r106
11. Minge MA, Silberman JD, Orr RJ, Cavalier-Smith T, Shalchian-Tabrizi K, Burki F, et al. Evolutionary position of breviate amoebae and the primary eukaryote divergence. Proc Biol Sci 2009; 276:597-604; PMID:19004754; http://dx.doi.org/10.1098/rspb.2008.1358
12. Fischer MG, Allen MJ, Wilson WH, Suttle CA. Giant virus with a remarkable complement of genes infects marine zooplankton. Proc Natl Acad Sci USA 2010; 107:19508-13; PMID:20974979; http://dx.doi.org/10.1073/pnas.1007615107
13. Fischer MG, Suttle CA. A virophage at the origin of large DNA transposons. Science 2011; 332:231-4; PMID:21385722; http://dx.doi.org/10.1126/science.1199412
14. Ogata H, Ray J, Toyoda K, Sandaa RA, Nagasaki K, Bratbak G, et al. Two new subfamilies of DNA mismatch repair proteins (MutS) specifically abundant in the marine environment. ISME J 2011; 5:1143-51; PMID:21248859; http://dx.doi.org/10.1038/ismej.2010.210
15. Claverie JM, Abergel C. Open questions about giant viruses. Adv Virus Res 2013; 85:25-56; PMID:23439023; http://dx.doi.org/10.1016/B978-0-12-408116-1.00002-1
16. Santini S, Jeudy S, Bartoli J, Poirot O, Lescot M, Abergel C, et al. Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. Proc Natl Acad Sci USA 2013; 110:10800-5; PMID:23754393; http://dx.doi.org/10.1073/pnas.1303251110
17. Brussaard CP, Short SM, Frederickson CM, Suttle CA. Isolation and phylogenetic analysis of novel viruses infecting the phytoplankton *Phaeocystis globosa* (Prymnesiophyceae). Appl Environ Microbiol 2004; 70:3700-5; PMID:15184176; http://dx.doi.org/10.1128/AEM.70.6.3700-3705.2004
18. La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, et al. The virophage as a unique parasite of the giant mimivirus. Nature 2008; 455:100-4; PMID:18690211; http://dx.doi.org/10.1038/nature07218
19. Desnues C, La Scola B, Yutin N, Fournous G, Robert C, Azza S, et al. Provirophages and transpovirons as the diverse mobilome of giant viruses. Proc Natl Acad Sci USA 2012; 109:18078-83; PMID:23071316; http://dx.doi.org/10.1073/pnas.1208835109
20. Fischer MG, Suttle CA. A virophage at the origin of large DNA transposons. Science 2011; 332:231-4; PMID:21385722; http://dx.doi.org/10.1126/science.1199412
21. Yau S, Lauro FM, DeMaere MZ, Brown MV, Thomas T, Raftery MJ, et al. Virophage control of antarctic algal host-virus dynamics. Proc Natl Acad Sci USA 2011; 108:6163-8; PMID:21444812; http://dx.doi.org/10.1073/pnas.1018221108
22. Zhou J, Zhang W, Yan S, Xiao J, Zhang Y, Li B, et al. Diversity of virophages in metagenomic data sets. J Virol 2013; 87:4225-36; PMID:23408616; http://dx.doi.org/10.1128/JVI.03398-12
23. Gaia M, Pagnier I, Campocasso A, Fournous G, Raoult D, La Scola B. Broad spectrum of mimiviridae virophage allows its isolation using a mimivirus reporter. PLoS ONE 2013; 8:e61912; PMID:23596530; http://dx.doi.org/10.1371/journal.pone.0061912
24. Legendre M, Arslan D, Abergel C, Claverie JM. Genomics of Megavirus and the elusive fourth domain of Life. Commun Integr Biol 2012; 5:102-6; PMID:22482024; http://dx.doi.org/10.4161/cib.18624
25. Claverie JM, Abergel C, Ogata H. Mimivirus. Curr Top Microbiol Immunol 2009; 328:89-121; PMID:19216436; http://dx.doi.org/10.1007/978-3-540-68618-7_3
26. Boyer M, Madoui MA, Gimenez G, La Scola B, Raoult D. Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. PLoS ONE 2010; 5:e15530; PMID:21151962; http://dx.doi.org/10.1371/journal.pone.0015530
27. Nasir A, Kim KM, Caetano-Anolles G. Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. BMC Evol Biol 2012; 12:156; PMID:22920653; http://dx.doi.org/10.1186/1471-2148-12-156
28. Colson P, de Lamballerie X, Fournous G, Raoult D. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 2012; 55:321-32; PMID:22508375; http://dx.doi.org/10.1159/000336562
29. Moreira D, López-García P. Ten reasons to exclude viruses from the tree of life. Nat Rev Microbiol 2009; 7:306-11; PMID:19270719
30. Williams TA, Embley TM, Heinz E. Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses. PLoS ONE 2011; 6:e21080; PMID:21698163; http://dx.doi.org/10.1371/journal.pone.0021080
31. Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, et al. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science 2013;341:281-6; PMID: 23869018; http://dx.doi.org/10.1126/science.1239181