2013

"Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses

Philippe Colson
Aix-Marseille Universite

Xavier De Lamballerie
EHSP French School of Public Health, Aix-Marseille Universite

Natalya Yutin
National Institutes of Health

Sassan Asgari
The University of Queensland

Yves Bigot
International Committee on Taxonomy of Viruses

See next page for additional authors

Arch Virol (2013) 158:2517–2521
DOI 10.1007/s00705-013-1768-6

ORIGINAL ARTICLE

"Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses

Philippe Colson · Xavier De Lamballerie · Natalya Yutin · Sassan Asgari · Yves Bigot · Dennis K. Bideshi · Xiao-Wen Cheng · Brian A. Federici · James L. Van Etten · Eugene V. Koonin · Bernard La Scola · Didier Raoult

Received: 3 August 2012 / Accepted: 7 May 2013 / Published online: 29 June 2013

Abstract The nucleocytoplasmic large DNA viruses (NCLDVs) comprise a monophyletic group of viruses that infect animals and diverse unicellular eukaryotes. The NCLDV group includes the families *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Mimiviridae* and the proposed family "Marseilleviridae". The family *Mimiviridae* includes the largest known viruses, with genomes in excess of one megabase, whereas the genome size in the other NCLDV families varies from 100 to 400 kilobase pairs. Most of the NCLDVs replicate in the cytoplasm of infected cells, within so-called virus factories. The NCLDVs share a common ancient origin, as demonstrated by evolutionary reconstructions that trace approximately 50 genes encoding key proteins involved in viral replication and virion formation to the last common ancestor of all these viruses. Taken together, these characteristics lead us to propose assigning an official taxonomic rank to the NCLDVs as the order "Megavirales", in reference to the large size of the virions and genomes of these viruses.

Introduction

The nucleocytoplasmic large DNA viruses (NCLDVs) are an apparently monophyletic group of viruses infecting eukaryotes that was first described in 2001 [1]. The NCLDVs encompass the families *Poxviridae*, *Asfarviridae, Iridoviridae*, *Ascoviridae*, and *Phycodnaviridae* [1–3] and two groups of distinct giant viruses that have been isolated from *Acanthamoeba*, giving rise to the now established family *Mimiviridae* [4–7] and the proposed family "Marseilleviridae" [8–10] (Table 1).

S. Asgari, Y. Bigot, D. K. Bideshi, X.-W. Cheng and B. A. Federici belong to the *Ascoviridae* study group of the International Committee on Taxonomy of Viruses (ICTV).

P. Colson · B. La Scola · D. Raoult (✉)
Unité des Rickettsies, URMITE UMR CNRS 7278 IRD 198 INSERM U1095, Facultés de Médecine et de Pharmacie, IHU Méditerranée Infection, Aix-Marseille Université, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 05, France
e-mail: didier.raoult@gmail.com

P. Colson · X. De Lamballerie · B. La Scola · D. Raoult
Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, IHU Méditerranée Infection, Assistance Publique, Hôpitaux de Marseille, Centre Hospitalo-Universitaire Timone, 264 Rue Saint-Pierre, 13385 Marseille Cedex 05, France

X. De Lamballerie
Unité des Virus Emergents, UMR190 Emergence des pathologies virales, Faculté de Médecine, Institut de Recherche pour le Développement, EHSP French School of Public Health, Aix-Marseille Université, Marseille, France

N. Yutin · E. V. Koonin
National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health, Bldg. 38A, 8600 Rockville Pike, Bethesda, MD 20894, USA

S. Asgari
School of Biological Sciences, The University of Queensland, St Lucia, QLD 4072, Australia

Y. Bigot
UMR INRA-CNRS 7247, PRC, Centre INRA de Nouzilly, 37380 Nouzilly, France

D. K. Bideshi
Molecular and Developmental Biology, University of California, Riverside, CA 92521, USA

The viruses of the family *Mimiviridae* possess by far the largest virions and genomes among all currently known viruses. Strikingly, mimivirus genomes are larger than those of many parasitic and symbiotic bacteria and approach the size and complexity of the smallest known free-living bacteria and archaea. Moreover, mimiviruses encode many genes that have not been found in other viruses, in particular multiple components of the translation system such as aminoacyl-tRNA synthetases [4, 11, 12]. A distant relative of mimiviruses, *Cafeteria roenbergensis* virus (CroV), has been isolated from a marine stramenopile [13], and numerous homologs of mimivirus genes have been detected in metagenomic samples [14–17], indicating that the actual diversity of the giant viruses remains largely untapped. These unusual features of the mimiviruses have attracted strong interest of many researchers and revitalized the study of molecular biology and biochemistry of the NCLDVs.

A reclassification of the members of the NCLDV families into a new virus order "Megavirales" has been suggested recently [18], based upon several defining features. In the present proposal, we succinctly summarize these unique traits of the NCLDV and make the formal case for the distinctness of the NCLDVs from other large DNA viruses.

The defining features of the NCLDVs

The monophyly of the NCLDVs, i.e., the common origin of all these viruses from the same ancestral virus, was inferred from the results of phylogenetic and phyletic analyses [1–3, 19]. All of the NCLDVs share five core genes, namely those encoding the major capsid protein (poxvirus D13 gene), helicase-primase (D5), DNA polymerase elongation subunit family B, DNA-packaging ATPase (A32), and viral late transcription factor 3 (A2L). Moreover, approximately 50 genes, although missing in some of the NCLDVs, were assigned, with high confidence, to the common ancestor of the entire group [3]. The maximum-likelihood evolutionary reconstruction that led to this conclusion relied upon a phylogenetic tree of the universally conserved NCLDV genes and the patterns of presence-absence of other genes as derived from the clusters of orthologous genes of the NCLDVs (NCVOGs). Although a comprehensive phylogenetic analysis of the putative ancestral NCLDV genes revealed a complex picture of evolution that involved multiple non-orthologous gene displacements, on the whole, the results of this analysis were compatible with the descent of (nearly) all of these genes from an ancestral virus [19]. Moreover, the inferred ancestral NCLDV genes encode proteins that perform key functions in virus genome replication and expression as well as virion morphogenesis and structure, suggesting that the putative ancestral virus already possessed the main biological features of the extant NCLDVs.

Importantly, the set of the approximately 50 ancestral genes sharply partitions the NCLDVs from all other groups of viruses, including large DNA viruses infecting eukaryotes, such as nudiviruses, herpesviruses and baculoviruses, as well as large DNA viruses of bacteria and archaea. Although some of these viruses share with the NCLDVs one or more of the "virus hallmark genes", such as the DNA polymerase, the helicase-primase or the packaging ATPase, none come close to possessing the entire set of the ancestral NCLDV genes [20].

The apparent origin of the NCLDVs from a common ancestral virus is buttressed by shared genomic, structural and biological features. With the exception of the members of the virus families *Poxviridae* and *Ascoviridae*, all the NCLDVs form large, icosahedral capsids (more than 100 nm in diameter) that are comprised of a single, homologous double β barrel jelly roll protein [21] and encapsidate a single double-stranded DNA molecule ranging in size from 100 kilobase pairs to over one megabase pairs (Table 1). The exceptions to this conserved virion architecture are the poxviruses and the ascoviruses, with their unique brick-shaped virions and allantoid capsids, respectively [22, 23]. However, at least in poxviruses, this appears to be a derived trait, because an intermediate icosahedral structure in poxvirus virion morphogenesis [24] contains the D13 protein that is the poxvirus homolog of the major capsid protein of the rest of the NCLDVs [1] and adopts a similar jelly roll fold [22].

A notable feature of the gene repertoires of many NCLDVs is the presence of genes that appear to have been derived from the host and encode proteins involved in virus-host interactions. The specific compositions of these variable portions of the NCLDV genomes strongly depend on the host. Thus, poxviruses and asfarviruses that infect vertebrates possess multiple genes that interfere with host immunity and programmed cell death [25–27]. In contrast,

X.-W. Cheng
Department of Microbiology, Miami University, Oxford, OH 45056, USA

B. A. Federici
Department of Entomology, University of California, Riverside, CA 92521, USA

J. L. Van Etten
Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583-0722, USA

J. L. Van Etten
Nebraska Center for Virology, University of Nebraska, Lincoln, NE 68583-0900, USA

Table 1 Viral families and genera in the proposed order 'Megavirales'

Family	Subfamily	Genus	Genome size (bp)		Host range
			Min	Max	
Ascoviridae		*Ascovirus*	119,343	186,262	Insects
Asfarviridae		*Asfivirus*	170,101	182,284	Mammals, dinoflagellates
Iridoviridae		*Chloriridovirus*	191,100	191,100	Insects
		Iridovirus	205,791	212,482	Insects
		Lymphocystivirus	102,653	186,250	Fishes
		Megalocytivirus	111,362	111,362	Fishes
		Ranavirus	105,890	140,131	Amphibia
Mimiviridae		-	617,453	1,259,197	Amoeba, green algae, heterokonts, haptophyta
		Mimivirus	1,021,348	1,259,197	Amoeba
"Marseilleviridae"		"Marseillevirus"	346,754	368,454	Amoeba
Phycodnaviridae		*Chlorovirus*	288,047	368,683	Green algae
		Coccolithovirus	407,339	407,339	Haptophyta
		Phaeovirus	154,641	335,593	Heterokonts
		Prasinovirus	184,095	198,519	Green algae
		Raphidovirus	-	-	Heterokonts
Poxviridae	*Chordopoxvirinae*	*Avipoxvirus*	288,539	359,853	Birds
		Capripoxvirus	149,599	150,773	Mammals
		Cervidpoxvirus	166,259	170,560	Mammals
		Crocodylidpoxvirus	190,054	190,054	Reptiles
		Leporipoxvirus	159,857	161,773	Mammals
		Molluscipoxvirus	190,289	190,289	Human
		Orthopoxvirus	175,699	224,499	Mammals
		Parapoxvirus	134,431	145,289	Mammals
		Suipoxvirus	146,454	146,454	Mammals
		Yatapoxvirus	134,721	144,575	Primates
		Unassigned	190,054	190,054	Animals
	Entomopoxvirinae	*Alphaentomopoxvirus*	n.a.	n.a.	Insects
		Betaentomopoxvirus	232,392	232,392	Insects
		Gammaentomopoxvirus	n.a.	n.a.	Insects
		Unassigned	236,120	236,120	Insects

n.a., not available

mimiviruses, phycodnaviruses and marseilleviruses that infect unicellular eukaryotes encompass numerous genes encoding proteins that can be predicted to modulate core cellular functions and intracellular signaling [4, 8].

The NCLDVs either replicate entirely within the cytoplasm of the infected cells or at least undergo essential parts of their reproduction cycles in the cytoplasm. The cytoplasm of cells infected with viruses from each of the NCLDV families (with the possible exception of some phycodnaviruses) contains distinct compartments known as virus factories, which are the sites of viral genome replication and expression as well as virion morphogenesis [28–30]. Generally similar virus factories have been described in cells infected with RNA viruses that replicate in the cytoplasm, e.g., picornaviruses [31, 32]. However, among viruses with DNA genomes, this feature is unique to the NCLDVs and sharply differentiates the NCLDVs from other large DNA viruses of eukaryotes, such as herpesviruses and baculoviruses, that replicate in the cell nucleus.

Diversity, host range and evolution of the NCLDVs

The complexities of the evolutionary histories of the core NCLDV genes [19] notwithstanding, the phylogenetic signals are coherent enough to obtain a consensus phylogeny (Figure 1) [19]. Each of the NCLDV families comes across as a clade, with the clades assembled into two major branches, one of which encompasses poxviruses and

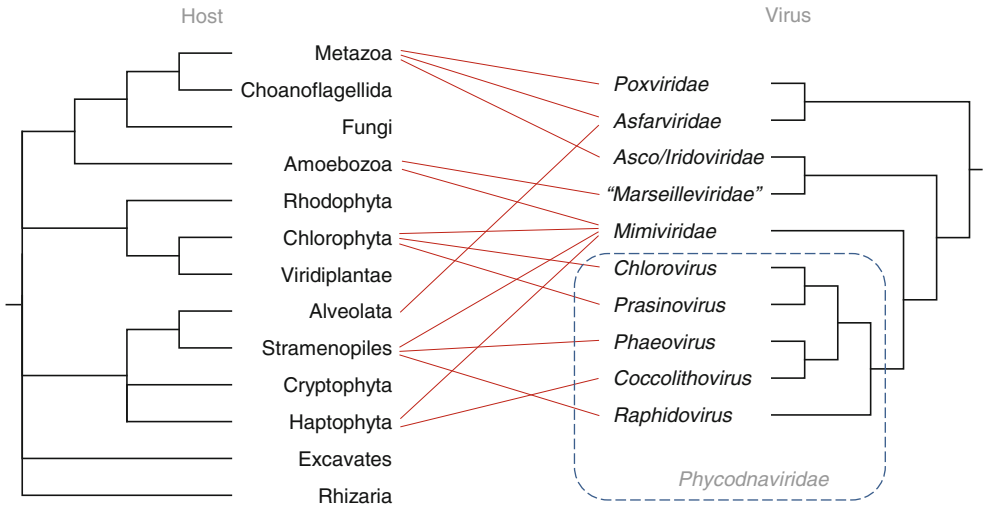

Fig. 1 Coevolution of the viruses in the proposed order "Megavirales" and their hosts. The schematic family-level evolutionary tree of the NCLDVs represents the consensus of the phylogenies of the core NCLDV genes (superfamily II helicase (NCVOG0076), A2L-like transcription factor (NCVOG0262), RNA polymerase α subunit (NCVOG0274), RNA polymerase β subunit (NCVOG0271), mRNA capping enzyme, A32-like packaging ATPase (NCVOG0249), small subunit of ribonucleotide reductase (NCVOG0276), myristoylated envelope protein (NCVOG0211), primase-helicase (NCVOG0023), and DNA polymerase (NCVOG0038)). For the highly diverse family *Phycodnaviridae*, a more detailed, genus-level phylogeny is shown. The schematic supergroup-level evolutionary tree of the eukaryotes shows a multifurcation, given the lack of resolution at the deepest level. Lines connect virus families (and genera of the family *Phycodnaviridae*) and their known hosts

asfarviruses, and the second one consists of the remaining five NCLDV families (Figure 1).

Of the five currently recognized supergroups of eukaryotes [33–35], known NCLDV hosts belong to three, namely unikonts (*Metazoa* and *Amoebozoa*), *Plantae* (green algae but not vascular plants) and *Chromalveolata* (*Haptophyta*, Stramenopiles); the remaining two supergroups, *Rhizaria* and *Excavata*, have not been studied biologically in sufficient detail to rule out the possibility that these organisms harbor NCLDVs as well.

Superposition of the host ranges of the NCLDVs over the consensus phylogenetic tree of the conserved genes of these viruses reveals a maze of virus-host relationships in which representatives of the same supergroup of eukaryotes are infected by members of multiple NCLDV families, whereas viruses of the same family infect hosts of multiple eukaryotic supergroups (Figure 1). Conceivably, this complex picture results from the ancient origin of the NCLDVs, which might have been concomitant with eukaryogenesis [20], and transfer of viruses between taxonomically distant hosts.

The proposed order "Megavirales'"

In summary, the NCLDVs encompass an extremely broad range of viruses with large DNA genomes that infect hosts across (almost) the entire range of eukaryotic diversity.

All these viruses are united by:

- common origin that is manifest in the existence of a large set of ancestral genes that are responsible for key viral functions;
- common virion architecture;
- common major biological features, in particular virus reproduction within cytoplasmic factories.

Taken together, these shared features strongly support the classification of the seven families of the NCLDV into a new viral order (http://talk.ictvonline.org/files/proposals/taxonomy_proposals_fungal1/m/fung01/4261.aspx). We propose to name this order "Megavirales" in reference to the characteristic large or giant size of the virions and genomes of these viruses.

Acknowledgments Natalya Yutin and Eugene V. Koonin are supported by intramural funds of the US Department of Health and Human Services (to the National Library of Medicine). James Van Etten is partially supported by NIH Grant P20 RR15635 from the COBRE Program of the National Center for Research Resources.

Conflict of interest There are no potential conflicts of interest or financial disclosures for any of the authors.

References

1. Iyer LM, Aravind L, Koonin EV (2001) Common origin of four diverse families of large eukaryotic DNA viruses. J Virol 75:11720–11734. doi:10.1128/JVI.75.23.11720-11734.2001
2. Iyer LM, Balaji S, Koonin EV, Aravind L (2006) Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res 117:156–184. doi:10.1016/j.virusres.2006.01.009

3. Yutin N, Wolf YI, Raoult D, Koonin EV (2009) Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol J 6:223. doi:10.1186/1743-422X-6-223
4. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie JM (2004) The 1.2-megabase genome sequence of Mimivirus. Science 306:1344–1350. doi:10.1126/science.1101485
5. Claverie JM, Abergel C, Ogata H (2009) Mimivirus. Curr Top Microbiol Immunol 328:89–121
6. La Scola B, Campocasso A, N'Dong R, Fournous G, Barrassi L, Flaudrops C, Raoult D (2010) Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology 53:344–353. doi:10.1159/000312919
7. Yoosuf N, Yutin N, Colson P, Shabalina SA, Pagnier I, Robert C, Azza S, Klose T, Wong J, Rossmann MG, La SB, Raoult D, Koonin EV (2012) 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 4:1324–1330. doi:10.1093/gbe/evs109
8. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La SB, Koonin EV, Raoult D (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci USA 106:21848–21853. doi:10.1073/pnas.0911354106
9. Thomas V, Bertelli C, Collyn F, Casson N, Telenti A, Goesmann A, Croxatto A, Greub G (2011) Lausannevirus, a giant amoebal virus encoding histone doublets. Environ Microbiol 13:1454–1466. doi:10.1111/j.1462-2920.2011.02446.x
10. Colson P, Pagnier I, Yoosuf N, Fournous G, La Scola B, Raoult D (2012) Marseilleviridae, a new family of giant viruses infecting amoebae. Arch Virol 158(4):915–920. doi:10.1007/s00705-012-1537-y
11. Claverie JM, Ogata H, Audic S, Abergel C, Suhre K, Fournier PE (2006) Mimivirus and the emerging concept of "giant" virus. Virus Res 117:133–144. doi:10.1016/j.virusres.2006.01.008
12. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM (2011) Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc Natl Acad Sci USA 108:17486–17491
13. Fischer MG, Allen MJ, Wilson WH, Suttle CA (2010) Giant virus with a remarkable complement of genes infects marine zooplankton. Proc Natl Acad Sci USA 107:19508–19513. doi:10.1073/pnas.1007615107
14. Monier A, Claverie JM, Ogata H (2008) Taxonomic distribution of large DNA viruses in the sea. Genome Biol 9:R106. doi:10.1186/gb-2008-9-7-r106
15. Monier A, Larsen JB, Sandaa RA, Bratbak G, Claverie JM, Ogata H (2008) Marine mimivirus relatives are probably large algal viruses. Virol J 5:12. doi:10.1186/1743-422X-5-12
16. Yamada T (2011) Giant viruses in the environment: their origins and evolution. Curr Opin Virol 1:58–62
17. Kristensen DM, Mushegian AR, Dolja VV, Koonin EV (2010) New dimensions of the virus world discovered through metagenomics. Trends Microbiol 18:11–19. doi:10.1016/j.tim.2009.11.003
18. Colson P, de Lamballerie X, Fournous G, Raoult D (2012) Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 55(5):321–332. doi:10.1159/000336562
19. Yutin N, Koonin EV (2012) Hidden evolutionary complexity of nucleo-cytoplasmic large DNA viruses of eukaryotes. Virol J 9:161
20. Koonin EV, Yutin N (2010) Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology 53:284–292. doi:10.1159/000312913
21. Krupovic M, Bamford DH (2008) Virus evolution: how far does the double beta-barrel viral lineage extend? Nat Rev Microbiol 6:941–948
22. Bahar MW, Graham SC, Stuart DI, Grimes JM (2011) Insights into the evolution of a complex virus from the crystal structure of vaccinia virus D13. Structure 19:1011–1020
23. Bigot Y, Asgari S, Bideshi D, Cheng XW, Federici BA et al. (2011) Family Ascoviridae. In: CM Fauquet, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) Virus taxonomy. Ninth report of the international committee on taxonomy of viruses, San Diego, pp 147–152
24. Szajner P, Weisberg AS, Lebowitz J, Heuser J, Moss B (2005) External scaffold of spherical immature poxvirus particles is made of protein trimers, forming a honeycomb lattice. J Cell Biol 170:971–981
25. Seet BT, Johnston JB, Brunetti CR, Barrett JW, Everett H, Cameron C, Sypula J, Nazarian SH, Lucas A, McFadden G (2003) Poxviruses and immune evasion. Annu Rev Immunol 21:377–423
26. Werden SJ, Rahman MM, McFadden G (2008) Poxvirus host range genes. Adv Virus Res 71:135–171. doi:10.1016/S0065-3527(08)00003-1
27. Dixon LK, Abrams CC, Bowick G, Goatley LC, Kay-Jackson PC, Chapman D, Liverani E, Nix R, Silk R, Zhang F (2004) African swine fever virus proteins involved in evading host defence systems. Vet Immunol Immunopathol 100:117–134. doi:10.1016/j.vetimm.2004.04.002
28. Condit RC (2007) Vaccinia, Inc.—probing the functional substructure of poxviral replication factories. Cell Host Microbe 2:205–207
29. Mutsafi Y, Zauberman N, Sabanay I, Minsky A (2010) Vaccinia-like cytoplasmic replication of the giant Mimivirus. Proc Natl Acad Sci USA 107:5978–5982. doi:10.1073/pnas.0912737107
30. Netherton CL, Wileman TE (2013) African swine fever virus organelle rearrangements. Virus Res 173(1):76–86. doi:10.1016/j.virusres.2012.12.014
31. Novoa RR, Calderita G, Arranz R, Fontana J, Granzow H, Risco C (2005) Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol Cell 97:147–172. doi:10.1042/BC20040058
32. Netherton CL, Wileman T (2011) Virus factories, double membrane vesicles and viroplasm generated in animal cells. Curr Opin Virol 1:381–387
33. Keeling PJ (2007) Genomics. Deep questions in the tree of life. Science 317:1875–1876
34. Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, Pearlman RE, Roger AJ, Gray MW (2005) The tree of eukaryotes. Trends Ecol Evol 20:670–676
35. Koonin EV (2010) The origin and early evolution of eukaryotes in the light of phylogenomics. Genome Biol 11:209–211