

# Reclassification of Giant Viruses Composing a Fourth Domain of Life in the New Order *Megavirales*

Philippe Colson<sup>a, b</sup> Xavier de Lamballerie<sup>b, c</sup> Ghislain Fournous<sup>a</sup> Didier Raoult<sup>a, b</sup>

<sup>a</sup>Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), CNRS UMR 6236 – IRD 198 – INSERM 1095, Méditerranée Infection, Facultés de Médecine et de Pharmacie, <sup>b</sup>Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Assistance publique des hôpitaux de Marseille, et <sup>c</sup>Unité des Virus Emergents, UMR190 Emergence des pathologies virales, Aix-Marseille Université, Institut de Recherche pour le Développement, EHSP French School of Public Health, Faculté de Médecine, Marseille, France

## Key Words

*Megavirales* · Giant virus · *Acanthamoeba polyphaga* Mimivirus · *Mimiviridae* · Nucleocytoplasmic large DNA viruses · Amoeba · *Acanthamoeba* spp. · Phagocytic protist · Classification · Domains of life

## Abstract

Interest in giant viruses has risen sharply since 2003, following the discovery of the Mimivirus and four other protist-infecting giant viruses that are linked to the nucleocytoplasmic large DNA viruses (NCLDVs). Despite considerable heterogeneity in hosts and genome sizes, the NCLDVs have been shown to be monophyletic based on analyses of their sequences and gene repertoires and recent studies have proposed that these viruses share a common ancient ancestor and compose a fourth domain of life. In addition, several characteristics of these giant viruses contradict or do not match the criteria used for the canonical definition of viruses, and the NCLDV denomination is not completely appropriate. We propose here to define a new viral order named *Megavirales*.

Copyright © 2012 S. Karger AG, Basel

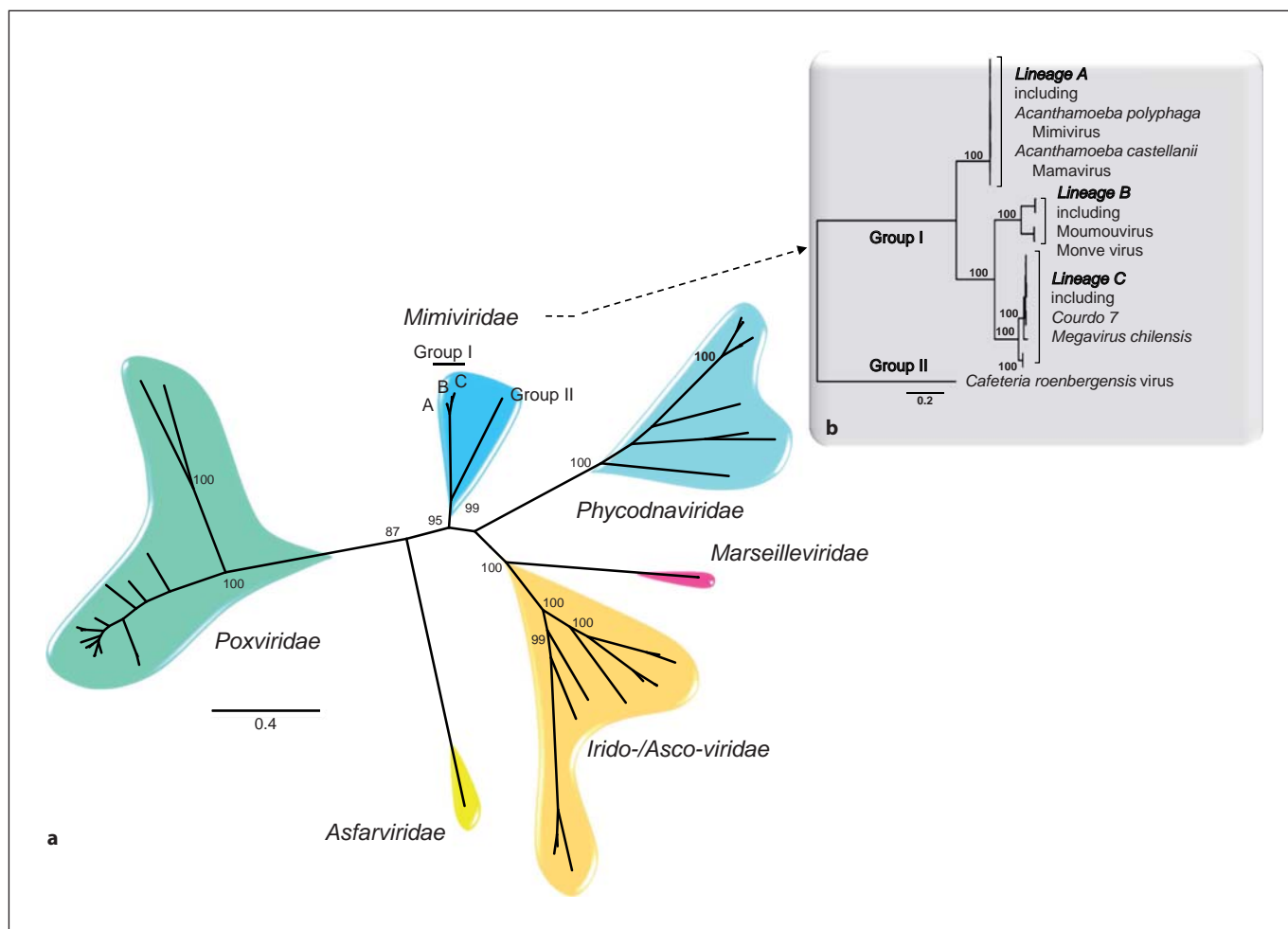
## Introduction

The existence of viruses with singularly large particle and genome sizes has been appreciated since the discovery of jumbo bacteriophages in the 1970s and the phycodnaviruses in the early 1980s [1, 2]. The interest in giant viruses increased dramatically in 2003 with the discovery of *Acanthamoeba polyphaga* Mimivirus, whose genome was the largest ever described among viruses (1,181 kb). It encodes more than 900 proteins, including some never identified previously in viruses [3, 4]. Overall, the Mimivirus discovery has led to considerable breakthroughs in our understanding of the definition, origin, and evolution of viruses [4–7]. Consequently, the number of publications and citations related to giant viruses has increased by more than 1 log (online suppl. fig. S1; for all online suppl. material, see [www.karger.com/doi/10.1159/000336562](http://www.karger.com/doi/10.1159/000336562)). Since 2008, several new giant viruses including close relatives to Mimivirus (Mamavirus, Terra2, Moumou, Courdo 11, *Megavirus chilensis*) and others more distantly related (*Cafeteria roenbergensis* virus (CroV), Marseillevirus, and Lausannevirus) have been recovered from different phagocytic protists and water samples by four teams (table 1; online suppl. table S1; fig. 1a) [5, 8–14].

**Table 1.** Main features of nucleocytoplasmic large DNA viruses whose genome is available in the NCBI GenBank genome database

Family	Subfamily/genus	Number of members	Name of the representing member whose genome has the most ancient date of creation (GenBank accession No.)	Most ancient date of creation	Genome topology	Genome size (nt)	GC content (%)	% coding	Number of coding genes	Number of protein coding genes	Number of structural RNA
<i>Ascoviridae</i>	<i>Ascovirus</i>	4	<i>Spodoptera frugiperda</i> ascovirus 1a (NC_008361)	13/09/2006	linear/circular	119,343–186,262	35–49	68–86	123–180	123–180	0
<i>Asfarviridae</i>	<i>Asfarvirus</i>	1	African swine fever virus (NC_001659)	22/04/1995	linear	170,101	38	87	151	151	0
<i>Iridoviridae</i>	<i>Chloriridovirus</i>	2	<i>Aedes taeniorhynchus</i> iridescent virus (NC_008187)	20/06/2006	linear	191,100	47	68	126	126	0
	<i>Iridovirus</i>	2	Invertebrate iridescent virus 6 (NC_003038)	31/07/2001	circular	212,482	28	90	463	468	0
	<i>Lymphocystivirus</i>	2	Lymphocystis disease virus 1 (NC_001824)	24/07/1997	linear	102,653–186,250	27–29	67–92	110–239	110–239	0
	<i>Megalocytivirus</i>	1	Infectious spleen and kidney necrosis virus (NC_003494)	28/03/2002	linear	111,362	54	93	125	125	0
	<i>Ranavirus</i>	4	<i>Ambystoma tigrinum</i> virus (NC_005832)	24/03/2004	linear	105,890–140,131	48–55	77–89	95–162	95–162	0
<i>Marseilleviridae</i>		2	Marseillevirus (NC_013756)	25/01/2010	circular	346,754–368,454	42–44	83–92	444–457	428–444	0
<i>Mimiviridae</i>	Group I										
	Lineage A	2	<i>Acanthamoeba polyphaga</i> Mimivirus (NC_014649)	12/11/2010	linear	1,181,549	23–28	88–90	544–1,059	544–1,023	22–39
	Lineage B	1	Monve virus (JN885994-JN886001)	17/01/2012	linear	1,015,033 <sup>a</sup>	23	–	–	–	–
	Lineage C	2	Courdo 7 virus (JN885990-JN885993)	17/01/2012	linear	1,170,106 <sup>a</sup>	25	–	–	–	–
		1	<i>Megavirus chilensis</i> (JN258408)	19/10/2011	linear	1,259,197	25	91	1,120	1,120	3
	Group II	1	<i>Cafeteria roenbergensis</i> virus (NC_014637)	10/11/2010	linear	617,453	23	90	544	544	22
<i>Phycodnaviridae</i>	Unclassified	3	<i>Ostreococcus</i> virus OsV5 (NC_010191)	28/12/2007	linear	184,409–191,761	41–44	89–92	237–264	230–264	0–5
	<i>Chlorovirus</i>	5	<i>Paramecium bursaria</i> Chlorella virus 1 (NC_000852)	22/12/1995	linear	288,047–368,683	39–49	92–93	699–886	689–886	7–12
	<i>Coccolithovirus</i>	1	<i>Emiliania huxleyi</i> virus 86 (NC_007346)	10/08/2005	circular	407,339	40	90	478	472	6
	<i>Phaeovirus</i>	2	<i>Ectocarpus siliculosus</i> virus 1 (NC_002687)	27/02/2001	circular/circular/linear	154,641–335,593	51	70–84	150–240	150–240	0
	<i>Prasinovirus</i>	3	<i>Bathycoccus</i> sp. RCC1105 virus BpV1 (NC_014765)	09/12/2010	linear	184,095–198,519	37–40	93–94	203–255	203–250	3–6
<i>Poxviridae</i>	<i>Chordopoxvirinae</i>										
	<i>Avipoxvirus</i>	1	Canarypox virus (NC_005309)	03/01/2004	linear	359,853	30	90	328	328	0
	<i>Capripoxvirus</i>	3	Lumpy skin disease virus NI-2490 (NC_003027)	22/07/2001	linear	149,599–150,773	25	93–95	148–156	148–156	0
	<i>Cervidpoxvirus</i>	2	Deerpox virus W-848-83 (NC_006966)	08/04/2005	linear	166,259–170,560	26–27	90–93	169–170	169–170	0
	<i>Leporipoxvirus</i>	1	Rabbit fibroma virus (NC_001266)	15/12/1999	linear	159,857	39	93	165	165	0
	<i>Molluscipoxvirus</i>	1	<i>Molluscum contagiosum</i> virus subtype 1 (NC_001731)	15/08/1996	linear	190,289	63	85	163	163	0
	<i>Orthopoxvirus</i>	7	Variola virus (NC_001611)	04/05/1993	linear	185,578–224,499	32–33	86–91	173–233	173–233	0
	<i>Parapoxvirus</i>	4	Bovine papular stomatitis virus (NC_005337)	21/01/2004	unknown/linear	134,431–145,289	63–64	89–94	130–134	130–134	0
	<i>Suipoxvirus</i>	1	Swinepox virus (NC_003389)	31/01/2002	linear	146,454	27	96	150	150	0
	Unclassified	1	Crocodilepox virus (NC_008030)	16/05/2006	linear	190,054	61	93	173	173	0
	<i>Yatapoxvirus</i>	4	Myxoma virus (NC_001132)	04/12/1999	linear	134,721–144,575	27–43	94–95	140–170	140–170	0
	<i>Entomopoxvirinae</i>	2	<i>Melanoplus sanguinipes</i> entomopoxvirus (NC_001993)	22/12/1998	linear	232,392–236,120	17–18	87–91	267–297	267–294	0
	Unclassified	1	Fowlpox virus (NC_002188)	21/03/2000	linear	288,539	30	84	260	261	0

<sup>a</sup> Approximate size: genome submitted as contigs.



**Fig. 1.** Phylogeny reconstruction from a cured concatenated alignment of universal NCVOGs [including primase-helicase (NCVOG0023), DNA polymerase (NCVOG0038), packaging ATPase (NCVOG0249), and A2L-like transcription factor (NCVOG0262)] for the giant viruses currently classified as NCLDVs (**a**) [modified from 14] and the *Mimiviridae* (**b**). Probabilities are mentioned near branches as a percentage and are used as confidence values of tree branches. Only probabilities at major nodes are shown. Scale bar represents the number of estimated changes per position for a unit of branch length.

All of the previously mentioned protist-associated giant viruses have been linked to nucleocytoplasmic large DNA viruses (NCLDVs) (tables 1–4) [15–18]. However, this grouping is not completely appropriate and several unique features of the NCLDVs do not match the criteria for the canonical definition of viruses [6, 32, 33]. In addition, these giant viruses were suggested to share a common ancestral origin and compose a new domain of life, aside *Bacteria*, *Archaea*, and *Eukarya* [17, 18, 34]. Therefore, we propose here to define a new viral order named *Megavirales*.

### Rationale and Argument Supporting the Definition of a New Viral Order

#### *The Current Definition and Classification of Giant Viruses Are Inappropriate*

The canonical definition of viruses was described by Lwoff [33] during the pregenomic era in 1957 and was historically based on negative criteria (online suppl. table S2). Later, genomics failed to identify any common gene in the virosphere that could be equivalent to universal proteins or ribosomal RNA for *Eukarya*, *Archaea*, and *Bacteria* [6, 35–37]. Thus, viruses remained separate from

**Table 2.** Brief description of the main features of NCLDV

Family name	Main features
<i>Poxviridae</i>	<i>Poxviridae</i> include member species infecting insects ( <i>Entomopoxvirinae</i> ) and vertebrates ( <i>Chordopoxvirinae</i> ) [19]. Several genera of poxviruses cause human illness, including <i>Orthopoxvirus</i> and <i>Parapoxvirus</i> , and a majority of the infections involving these viruses are zoonoses. Poxviruses are complex viruses whose genome contains inverted terminal repeats between <0.1 and 12.4 kb in size, and strands of DNA at the genome termini that are covalently linked to produce hairpins [19, 20].
<i>Iridoviridae</i>	Iridoviruses infect invertebrates and poikilothermic vertebrates, including insects, fish, amphibians, and reptiles [21]. Their viral genomes are both circularly permuted and terminally redundant, which is a unique feature among eukaryote-infecting virus genomes [21, 22]. They are classified into five genera: <i>Iridovirus</i> , <i>Chloriridovirus</i> , <i>Ranavirus</i> , <i>Lymphocystivirus</i> , and <i>Megalocyctivirus</i> . Their genomes contain frequent complex repeat sequences.
<i>Ascoviridae</i>	The ascoviruses were discovered in the 1970s and classified as a single genus, <i>Ascovirus</i> , within the family <i>Ascoviridae</i> [23, 24]. They infect the lepidoptera (moths and butterflies). These viruses harbor large membrane-bound vesicles whose formation is unique in the viral world and which are responsible for their name. At least 12 different proteins, ranging in size from 11 to 200 kDa, are present in virions [24]. The ascovirus genome has a complex organization [23, 24]. <i>Ascoviridae</i> represent a unique family of insect viruses and are most closely related to the invertebrate iridoviruses [16, 24].
<i>Asfarviridae</i>	The <i>Asfarviridae</i> family is represented exclusively by the African swine fever virus, the agent of a fatal hemorrhagic disease in domestic swine [25, 26].
<i>Phycodnaviridae</i>	The <i>Phycodnaviridae</i> are lytic and lysogenic viruses and are represented by three major viruses: <i>Paramecium bursaria Chlorella virus 1</i> (PBCV-1), <i>Emiliana huxleyi</i> virus (EhV), and <i>Ectocarpus siliculosus</i> virus (EsV), which belong to the genera <i>Chlorovirus</i> , <i>Coccolithovirus</i> , and <i>Phaeovirus</i> , respectively [27, 28]. The PBCV-1 genome has covalently closed hairpin termini flanked by approximately 2.2-kb inverted repeats [28, 29]. The EsV-1 genome has a relatively low gene density and inverted repeats at each end that allow circularization [27–29]. It is integrated into the host genome and passes to the daughter cells during cell division. The EhV-86 genome may have linear and circular phases. <i>Ectocarpus siliculosus</i> virus 1 infects the gametes or spores of filamentous marine brown macroalgae. In contrast to the <i>Chlorella</i> viruses and EsV-1 that depend on host transcription machinery (PBCV-1 lacks a recognizable RNA polymerase gene), EhV has six RNA polymerase-encoding genes [30]. It is noteworthy that the <i>Phycodnaviridae</i> hosts are widely distributed. Indeed, <i>Chlorella</i> are green algae that are among the most widely distributed and most frequently encountered algae [28, 29]. <i>Emiliana huxleyi</i> , the host of EhV, is also a ubiquitous unicellular marine phytoplankton that is thought to be important in ocean carbon and sulfur cycles [31]. <i>Ectocarpus siliculosus</i> is a member of benthic communities in the near-shore coastal environments of all the world's oceans [28].
<i>Mimiviridae</i>	<i>Acanthamoeba polyphaga</i> Mimivirus has been described in 2003 and was the largest virus [3, 4]. It has been assigned to a new viral family: the <i>Mimiviridae</i> [4]. <i>Acanthamoeba castellanii</i> Mamavirus, another strain of Mimivirus, was described in 2008 [5]. In 2010, several new giant viruses recovered by culture on amoebae have been described and phylogeny reconstructions based on highly conserved genes delineate three lineages (referred to as A, B and C) that compose a first group among the <i>Mimiviridae</i> (fig. 1b). One of these lineages (A) is composed of Mimivirus and closely related viruses [9]. A new <i>Mimiviridae</i> member, <i>Megavirus chilensis</i> , was recently described [10]. It falls undoubtedly within lineage C, being closely related to a previously recovered giant virus, Courdo 11 virus [9]. In 2010, <i>Cafeteria roenbergensis</i> virus was described and linked to the <i>Mimiviridae</i> family, away from the group consisting of three lineages A, B and C [11, 14]. This virus infects a widespread marine heterotrophic flagellate that belongs to the Chromalveolata phylum, and not <i>Acanthamoeba</i> sp. as Mimivirus and Mamavirus.
<i>Marseilleviridae</i> <sup>a</sup>	The <i>Marseilleviridae</i> family is a putative new viral family among the NCLDVs. Marseillevirus was discovered in 2007 in water collected from a cooling tower in Paris, France [12]. In 2011, a close relative to Marseillevirus, named Lausannevirus, was described [13]. The Lausannevirus genome shares 89% of genes with the Marseillevirus genome.

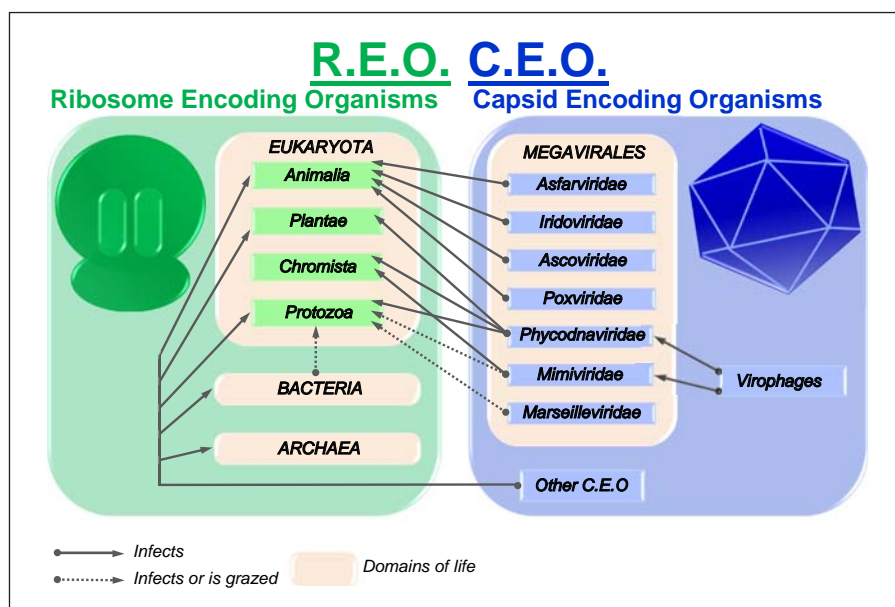
<sup>a</sup> Putative new viral family.

these biological entities. Recently, a new classification was proposed, which defines them as capsid-encoding organisms as opposed to ribosome-encoding organisms that compose the three canonical domains of life and are used to complete the viral life cycle (fig. 2). Besides, major monophyletic classes of viruses were tentatively defined [15, 37]. Most giant viruses were linked to one of these classes, NCLDV, which includes *Poxviridae*, *Asfarvi-*

*dae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Mimiviridae*, and *Marseilleviridae* (Marseillevirus and Lausannevirus) (tables 1–3; fig. 1a; online suppl. table S1) [13, 17]. Regarding *Mimiviridae*, new giant viruses infecting amoebae have been described by La Scola et al. [9] in 2010 and phylogeny reconstructions based on highly conserved genes enable delineating three lineages, referred to as A, B, and C. One of these lineages (A) is composed of

**Table 3.** A brief history of key steps in the definition of the NCLDV superfamily and the Megavirales order

Reference	Major findings
Iyer et al., 2001 [15]	<b>The NCLDV group was conceived</b> in 2001 by Iyer et al. [15] on the basis of comparative analyses of viral genomes that aimed to delineate a complete set of conserved genes. The NCLDVs originally consisted of four families of viruses, <i>Poxviridae</i> , <i>Asfarviridae</i> , <i>Iridoviridae</i> , and <i>Phycodnaviridae</i> , which harbor a large double-stranded DNA genome greater than 200,000 bp in size and sharing 9 genes (and 22 additional genes found in at least three of these four viral families).
Raoult et al., 2004 [4]	The genome of <b>Mimivirus</b> was described in 2004 and was the largest for a virus, being larger than that of several bacterial and archaeal parasites [3, 4]. This genome is a double-stranded linear DNA molecule of 1,181,404 bp. Raoult et al. [4] have identified 1,262 ORFs, of which 911 were predicted to be protein-coding genes. A total of 298 ORFs (24% of the predicted genes) were able to be associated with functional attributes. Among the predicted ORFs, 194 matched significantly with 108 clusters of orthologous groups. Mimivirus led to the definition of a new family among NCLDVs: <i>Mimiviridae</i> .
Iyer et al., 2006 [16]	In a work examining the Mimivirus genome and several other genomes recently available for three of the previously defined NCLDV families, Iyer et al. [16] tentatively defined the <b>gene complement of the ancestral NCLDV</b> and found that it harbored <b>at least 41 core genes</b> . They also identified a core gene set of 11 conserved genes in all the NCLDVs. Significant matches with sequence databases were obtained and conserved domains were identified for 188 ORFs.
La Scola et al., 2008 [5]	The concept of <b>virophage</b> was described in 2008. La Scola et al. [5] presented the recovery and analysis of Sputnik, a representative of a new family of viruses that grew in the giant virus factory found in amoebae coinfecting with <b>Mamavirus</b> , a new Mimivirus strain. The virophage presented functional analogy with bacteriophages; it replicated only in the presence of Mamavirus and its growth was deleterious to the giant virus. The Mamavirus genome was described in detail in 2011 [8]. It is highly similar to that of Mimivirus, has a length of 1,191 kb and harbors 1,023 predicted protein-coding genes.
Boyer et al., 2009 [12]	<b>The Marseillevirus</b> was isolated using amoebal culture of water collected in a cooling tower [12]. Its genome is a circular, double-stranded DNA molecule of 368,453 bp. A total of 457 ORFs have been predicted to encode proteins. Phylogenetic analyses have shown that Marseillevirus belongs to a new viral family of NCLDVs. It has been highlighted that its genome includes genes of likely bacterial, archaeal, eukaryotic, and viral origins. This mosaic gene content has supported the model of amoebae as hot spots for gene gain and exchange between entities with a sympatric intra-amoebal lifestyle. The genome of Lausannevirus, a close relative of the Marseillevirus, was described in 2011 [13]. It was recovered using amoebal co-culture from a water sample collected in the Seine river that runs through Paris, France.
Yutin et al., 2009 [17]	In 2009, Yutin et al. [17] updated the core NCLDV gene set on the basis of constructing clusters of orthologous groups of proteins with an extended number of NCLDV genomes, including those of Mamavirus and Marseillevirus. They identified 1,445 <b>NCLDV orthologous groups, so-called NCVOGs</b> , including 177 represented in more than one NCLDV family and five including proteins from all 45 analyzed viruses. Additionally, they used a maximum-likelihood reconstruction of the evolution of NCLDVs to define a set of 47 conserved genes that were probably present in the genome of the NCLDV common ancestor. Yutin et al. proposed that NCVOGs might provide a helpful platform for genome analysis and functional annotation of newly characterized NCLDVs.
Koonin and Yutin, 2010 [18]	Koonin and Yutin [18] addressed the origins and evolution of NCLDVs. The NCLDVs infect diverse animals and protists and the NCLDV core genes appear to have various probable origins including eukaryotes, bacteria, and bacteriophages (fig. 3). Their results suggest that the <b>NCLDVs originated at an early stage in the evolution of eukaryotes</b> .
La Scola et al., 2010 [9]	19 new giant viruses have been recovered by improved amoebal culture protocols from 105 environmental samples, mostly water [9]. For those linked to the <i>Mimiviridae</i> , phylogeny reconstruction of highly conserved proteins revealed <b>three main lineages</b> named A (that includes Mimivirus), B (whose leading member is Moumouvirus), and C composed of several members including Courdo 11 and Terra1 (table 1; fig. 1b). In 2011, <i>Megavirus chilensis</i> , another <i>Mimiviridae</i> that falls within group C, was described [10]. It is very closely related to Courdo 11 and has the largest genome amongst those released so far (table 1; fig. 1b).
Boyer et al., 2010 [34]	In 2010, phylogenetic and phyletic analyses based on the presence/absence patterns of informational genes (the genes involved in DNA biosynthesis and processing, including nucleotide biosynthesis, DNA replication and repair, recombination, and transcription) shared by <i>Eukarya</i> , <i>Bacteria</i> , <i>Archaea</i> , and the NCLDVs, allowed the delineation of a <b>fourth domain of life</b> (consisting of the NCLDVs) that supports the monophyly and common origin of these giant viruses [34].
Fischer et al., 2010 [11]	The genome of the <b><i>Cafeteria roenbergensis virus</i> (CroV)</b> was described in 2010 [11]. This virus infects a widespread marine heterotrophic flagellate that is phylogenetically distant from the amoebal host of the Mimivirus and Marseillevirus ( <i>Acanthamoeba</i> sp.). The CroV genome has an estimated size of 730 kb and contains 544 putative ORFs. Phylogenetic reconstructions indicated that CroV is related to the <i>Mimiviridae</i> family, apart from the group composed by three lineages (fig. 1b). A new virophage that infects the <i>Cafeteria roenbergensis virus</i> , Mavirus, was subsequently described [35]; it was hypothesized that Maverick/Polinton transposons may have originated from ancient relatives of the Mavirus.



**Fig. 2.** Schematic illustrating the relationships between ribosome-encoding organisms and capsid-encoding organisms, including the *Megavirales* members.

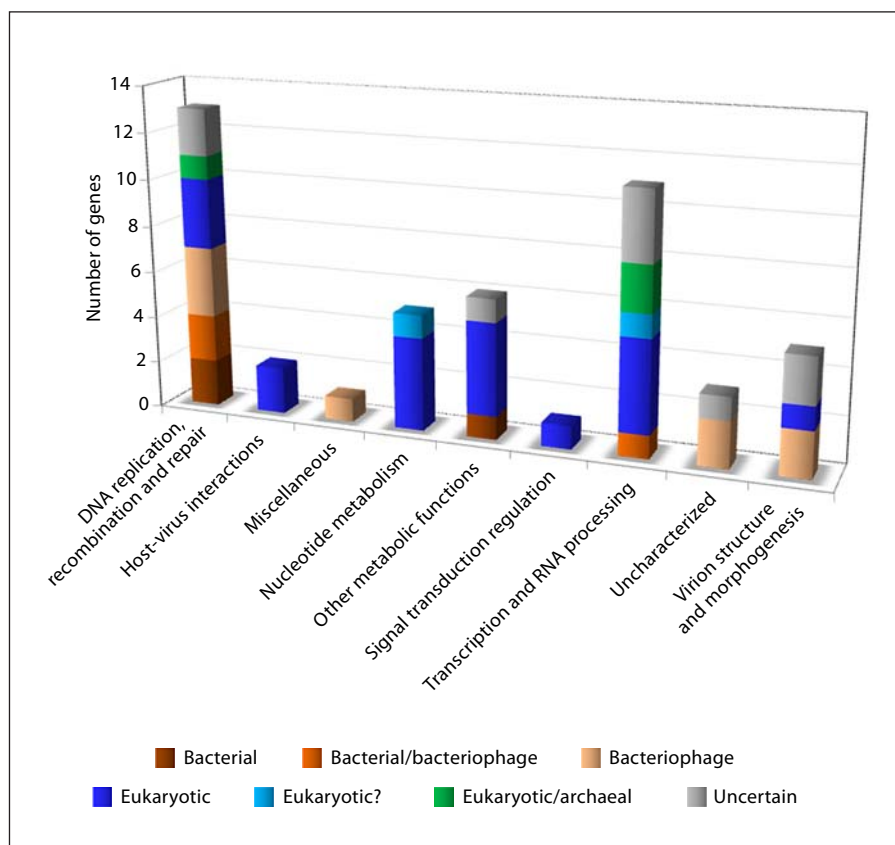
**Table 4.** Range of genome sizes, % GC content, coding proportions of genomes, and numbers of genes for the seven nucleocytoplasmic large DNA virus family members whose genome sequences are available in the NCBI GenBank genome database

Family	Genome size, nt		GC content, %		% coding		Number of genes	
	min	max	min	max	min	max	min	max
<i>Ascoviridae</i>	102,653	119,343	29	55	79	92	99	110
<i>Asfarviridae</i>	170,101	170,101	38	38	87	87	151	151
<i>Poxviridae</i>	134,431	359,853	17	64	80	96	130	328
<i>Marseilleviridae</i>	346,754	368,454	42	44	83	92	444	457
<i>Iridoviridae</i>	102,653	212,482	27	55	67	93	95	463
<i>Phycodnaviridae</i>	154,641	407,339	37	51	70	94	150	886
<i>Mimiviridae</i>	617,453	1,259,197	23	28	88	91	544	1,120

Mimivirus and closely related viruses (table 1; fig. 1b). The recently described *Megavirus chilensis* [10] is closely related to a giant virus previously recovered and classified within lineage C (table 1; fig. 1b). CroV has been also classified among the *Mimiviridae*, apart from the group composed by the lineages A, B, and C (table 1; fig. 1b) [11, 14].

Despite large heterogeneity in their hosts and genome sizes, the monophyly of the NCLDVs has been attested by phylogenetic and phyletic analyses, and the gene repertoires of these viruses distinguished them from bacteria, archaea, and eukaryotes [15, 18, 37]. The NCLDVs were originally defined as sharing nine genes found in all families, including three viral hallmark genes (table 3) [15]. Later, Yutin et al. [17] identified a set of 1,445 NCLDV

clusters of orthologous groups of proteins, referred to as NCVOGs, that included 177 represented in two or more NCLDV families and 5 present in all viruses. Other viruses, including *Myoviridae*, *Nimaviridae*, *Herpesviridae*, and *Polydnaviridae*, exhibit large genome and particle sizes, but their gene content precludes their incorporation within the NCLDVs [15, 37]. Some viral hallmark genes are shared between the NCLDVs and other large DNA viruses, as exemplified by the B-family DNA polymerases that are shared with herpesviruses and baculoviruses, but there are considerable numbers of other genes shared by the NCLDVs to the exclusion of all other large viruses [17, 37]. Moreover, the DNA replication and transcription of herpesviruses and baculoviruses occur exclusively in the



**Fig. 3.** Functional annotation and probable origin of the reconstructed core gene set of the common ancestor of the NCLDVs (47 NCVOGs) [adapted from 18].

nucleus, in contrast to NCLDVs [15]. Regarding *Myoviridae*, they are tailed bacteriophages [28].

Based on current knowledge, giant viruses and other canonical viruses differ in many aspects, which is not consistent with the concept (conveyed by Lwoff's classification) that the viral world is a homogeneous class of entities (online suppl. table S2) [6, 16, 32]. As an example, the huge gap between the Mimivirus and the hepatitis C virus is striking. The specific NCLDV features that strongly challenge the canonical definition of viruses are listed below (online suppl. table S2).

The NCLDVs have a capsid diameter that ranges between 150 and 500 nm, which contradicts the historical concept of viruses as small, ultrafilterable entities [32, 38–40]. In addition, the NCLDVs have large genomes that range in size between 103 and 1,259 kb and harbor 95–1,120 genes (table 4).

Viral messenger RNAs were detected in Mimivirus and Marseillevirus particles [4, 12]. These transcripts encode notably for capsid protein, DNA polymerase, or TFII-like transcription factor. The presence of RNA in the vaccinia virus particles has also been reported [41]. This contradicts

a key point of Lwoff's viral classification, which stated that viruses only harbor one type of nucleic acid [32, 33].

The *Mimiviridae* and *Marseilleviridae* genomes encode proteins involved in translation, which represents a unique feature of these viruses [4, 12, 32]. Besides, the genomes of *Mimiviridae* and *Phycodnaviridae* exhibit tRNAs [4, 28].

The NCLDVs were suggested to have a common ancestral origin dating back to an early stage of *Eukarya* evolution (table 3) [15, 17, 34, 37]. Thus, Yutin et al. [17] used maximum-likelihood reconstruction to delineate a set of 47 conserved genes that were probably present in the genome of the NCLDV common ancestor, which may have been a giant virus (fig. 3). Additionally, the NCLDVs infect a considerable diversity of hosts that belong to the three canonical domains of life [17, 28, 42]. Moreover, cross-mapping of the NCLDV and host eukaryotic trees generated a complex network in which members of the same NCLDV branch exhibited relationships with eukaryotic organisms of different supergroups [17]. For example, despite the relationship between them, irido-/ascoviruses infect animals, while Marseillevirus infects a

protist. Yutin et al. [17] have proposed the hypothesis of a 'Big Bang-like' event concomitantly with eukaryogenesis for the origin of the NCLDV [43].

Furthermore, it was proposed in 2010 that NCLDVs might define a fourth domain of life. This has been based on phylogenetic and phyletic studies of the repertoires of genes involved in information storage and processing and nucleotide transport and metabolism, and shared by *Eukarya*, *Bacteria*, *Archaea*, and the NCLDVs [34]. This work provided additional data supporting the monophyly and common origin of these giant viruses. In addition, it supports the hypotheses that the core genome of the NCLDVs may be as ancient as those of the three current canonical domains of life and that NCLDVs may have emerged as ancient roots from the rhizome of life [34, 44]. It was claimed in a recent work that the methodology used by Boyer et al. [34] for phylogenetic reconstructions was not the most appropriate to avoid spurious tree topologies generated by compositional heterogeneity and homoplasy, and alternative informational gene phylogenies did not support a fourth domain of life for NCLDVs [45]. Nevertheless, these trees fail to show a monophyly of *Eukarya* as well. In addition, other recent findings based on extensive analysis of metagenomic data suggest the existence of domains other than *Eukarya*, *Archaea*, and *Bacteria* [46].

#### *Other Major Features of Giant Viruses Classified Along with NCLDVs*

The NCLDVs can be characterized by other peculiar features in addition to those that radically classify them as separate from other viruses.

*Poxviridae*, *Iridoviridae*, and *Asfarviridae* can build viral factories [47], also reported in the case of Mimivirus, Megavirus, Marseillevirus, and Lausannevirus [10, 12, 13, 48]. These factories are associated with a massive production of virions.

The NCLDVs display a high level of genomic plasticity. Indeed, lineage-specific gene expansion and horizontal gene transfer have played a major role in the shaping of their genomes [4, 42, 49–53]. The proportion of duplicated genes in these viruses was found to range between 8 and 44%, with the highest proportions observed in Mimivirus [50, 51]. In addition, horizontal gene transfer has generated considerable genome plasticity and mosaicism, although the direction or source of the transfers and the fraction of gene content involved remain controversial [7, 51]. According to Filée [51], 0.8–11.9% of the genes were exchanged with the viral hosts, with the highest proportion being observed in *Poxviridae*, and up to 9.6% of the gene content was exchanged with bacteria, with the high-

est proportion being in Mimivirus. The potential mechanisms by which *Poxviridae* shape their genome through transfers of genes of host or viral origin have been particularly described [54–57]. In addition, the numbers of gene transfers with bacteria are the greatest for the Mimiviruses, Marseilleviruses, and phycodnaviruses that infect hosts feeding on bacteria [42]. The sympatric lifestyle within a phagocytic protist that grazes on bacteria, giant viruses, and virophages provides many opportunities indeed for these pathogens to gain and exchange genes. Thus, amoebae have been described as hot spots for gene transfer that may lead to the emergence of chimeric viruses and even the creation of new species [12, 58]. Interestingly, a reduction in genome size by approximately 16% was recently observed for the Mimivirus when subcultured 150 times in a germ-free amoebal host [59].

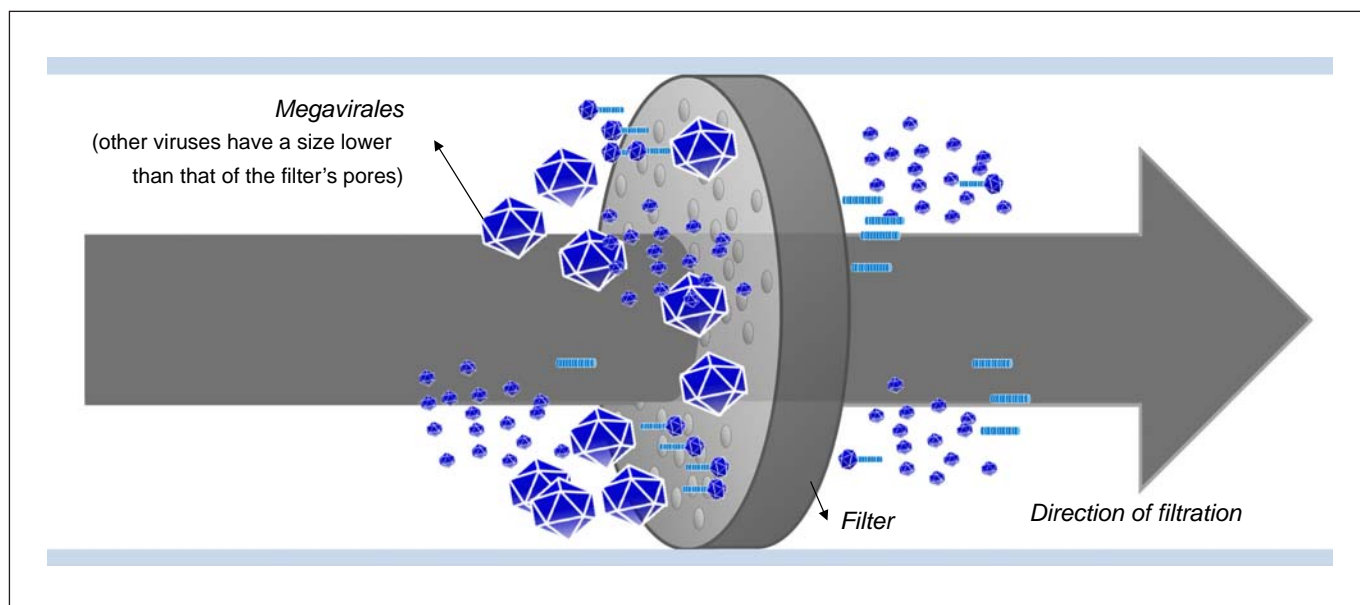
In addition to the core gene set, NCLDV genomes contain open reading frames (ORFs) without detectable homologs, also known as ORFans [60]. Strikingly, 2.8–75.2% of ORFs in the NCLDV genomes lack homologs in the NCBI GenBank reference protein sequence database. Moreover, 0.3–10.4% of these ORFs have homologs in the GenBank environmental protein sequence database.

The NCLDVs themselves can be infected by viruses [61], as has been previously shown for eukaryotes, bacteria, and archaea. In 2008, La Scola et al. [5] identified Sputnik, a virus infecting Mamavirus, which led to the creation of the virophage concept. Since then three new virophages have been described in association with *Mimiviridae* and *Phycodnaviridae* [9, 35, 62]. Importantly, virophages may be involved in gene transfer [5, 35].

#### *Giant Viruses Classified with the NCLDVs Are Probably Common Inhabitants of Our Biosphere*

According to our current knowledge, the NCLDVs remain a minority in the virosphere. Nevertheless, several findings indicate that they are common inhabitants of our biosphere. It is noteworthy that their presence has probably been largely underestimated up to this point because most metagenomic studies have adhered to the dogma of the small size of viruses by filtering samples prior to analysis (fig. 4) [63–65]. Notwithstanding, sequences similar to those from Mimivirus, African swine fever virus, and iridoviruses have already been identified in marine environmental samples or human serum and sewage [66–71]. Furthermore, giant viruses have been recovered from five different geographical areas worldwide and they have been isolated from approximately 20% of water samples in one study by optimizing amoebal culture protocols [9].





**Fig. 4.** Schematic illustrating how the giant viruses may have been excluded during the assessment of viromes by metagenomic studies that have filtered samples prior to analysis. Such procedures are inevitably preventing the detection of viruses larger than the pores of the filters used, i.e. 0.2–0.45  $\mu\text{m}$ .

#### *NCLDV Is Not an Appropriate Denomination and Has No Recognized Taxonomic Meaning*

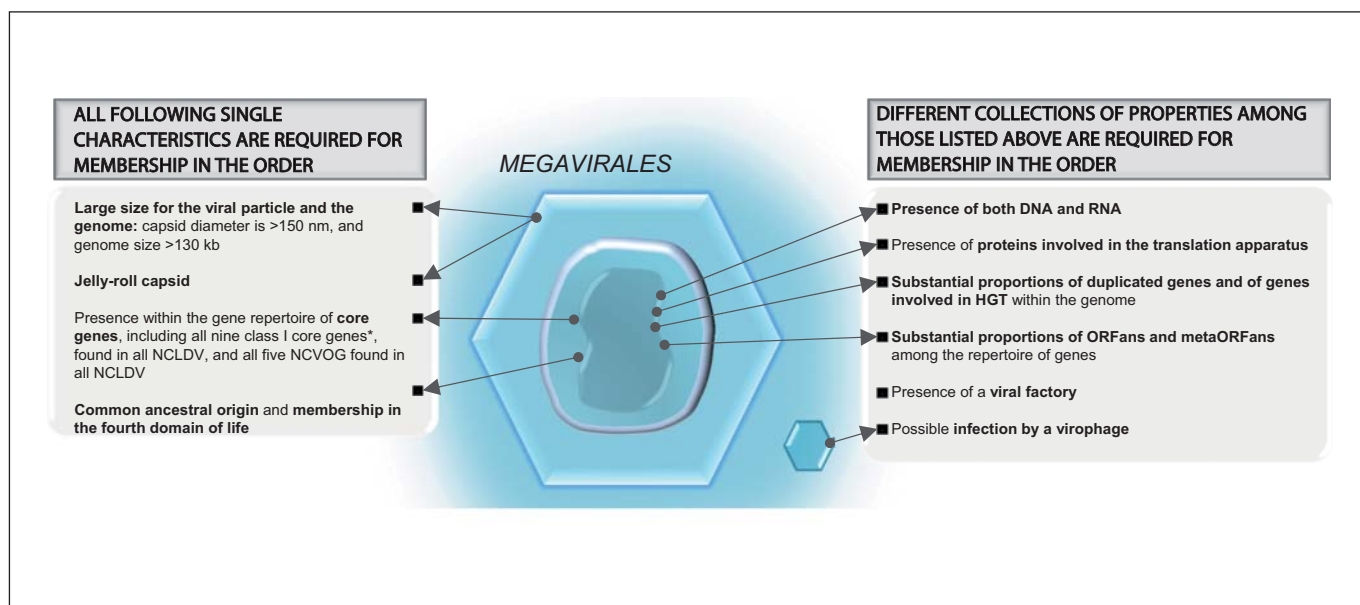
Finally, the NCLDV denomination does not take into account that Mimivirus and Marseillevirus harbor both DNA and RNA. Moreover, the NCLDVs compose a superfamily, a grouping that has no formally recognized taxonomic meaning according to the International Committee on Taxonomy of Viruses (ICTV) (<http://www.ictvonline.org/virusTaxonomy.asp?bhcp=1>). In the current ICTV classification, none of the NCLDV families are assigned to a viral order. We propose that these viruses should be assigned to a newly defined order (a group of families sharing certain common characteristics according to the ICTV) named *Megavirales*, in reference to the uncommon size of both the members' particles and their genomes.

#### **Definition of the *Megavirales***

Viral members of the new *Megavirales* order correspond to the giant viruses previously classified within the NCLDVs (table 1; online suppl. table S1). *Megavirales* can be defined by the criteria mentioned below, as illustrated in figure 5.

All of the following single characteristics are required for membership in the order (the monothetical system [72]):

- Giant viral particle and genome, capsid diameter  $>150$  nm and genome size  $>100$  kb (or in that order of magnitude).
- Presence in the gene content of all nine class I NCLDV core genes, i.e. VV D5-type ATPase (superfamily III helicase), DNA polymerase (B family), VV A32 virion packaging ATPase, VV A18 helicase (superfamily II), capsid protein D13L, thiol oxidoreductase, VV D6R/D11L-like helicase (superfamily II), S/T protein kinase, transcription factor VLTF2 [15] and all five NCVOGs found in all NCLDVs (i.e. NCLDV major capsid protein, D5-like helicase-primase, DNA polymerase elongation subunit family B, A32-like packaging ATPase and Poxvirus Late Transcription Factor VLTF3-like) [18]. These genes have various functions and origins.
- Common ancestral origin and membership in the proposed fourth domain of life.
- A jelly-roll capsid protein, which is a hallmark viral protein [6, 37]. The capsid is icosahedral in all NCLDVs except in poxviruses, where it forms intermediate structures during virion morphogenesis, but is not a protein of the virion [37, 73].



**Fig. 5.** Major features of *Megavirales* members and criteria required for membership in the *Megavirales* order.

Different combinations of properties among those listed above are required for membership in the order (the polythetical system): presence of both DNA and RNA; presence of proteins involved in the translation apparatus; substantial proportions of genes duplicated and involved in horizontal gene transfer within the genome; substantial proportions of ORFans and metaORFans among the gene repertoire; presence of viral factories; some or all steps of DNA replication and transcription occurring in the host cytoplasm, and possible infection by a virophage.

## Conclusion

The tremendous recent increase in knowledge about giant viruses has generated divergence rather than reinforced the borders of the previously defined viral world. *Megavirales* gather viral entities that appear to be incompatible within the framework of the virosphere as it has been defined since the beginning of virology. Moreover, they lay the foundation for a new understanding in which viruses consolidate their status as early protagonists in evolution.

## References

- Donelli G, Dore E, Frontali C, Grandolfo ME: Structure and physico-chemical properties of bacteriophage G. III. A homogeneous DNA of molecular weight 5 times 10(8). *J Mol Biol* 1975;94:555–565.
- Van Etten JL, Meints RH, Kuczmarski D, Burbank DE, Lee K: Viruses of symbiotic *Chlorella*-like algae isolated from *Paramecium bursaria* and *Hydra viridis*. *Proc Natl Acad Sci USA* 1982;79:3867–3871.
- La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie JM, Raoult D: A giant virus in amoebae. *Science* 2003;299:2033.
- Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie JM: The 1.2-megabase genome sequence of Mimivirus. *Science* 2004;306:1344–1350.
- La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, et al: The virophage as a unique parasite of the giant mimivirus. *Nature* 2008;455:100–104.
- Raoult D, Forterre P: Redefining viruses: lessons from Mimivirus. *Nat Rev Microbiol* 2008;6:315–319.
- Forterre P: Giant viruses: conflicts in revisiting the virus concept. *Intervirology* 2010;53:362–378.
- Colson P, Yutin N, Shabalina S, Robert C, Fournous G, La Scola B, Raoult D, Koonin EV: Viruses with over 1,000 genes: *Acanthamoeba castellanii* Mamavirus, a new mimivirus strain, and reannotation of Mimivirus genes. *Genome Biol Evol* 2011;3:737–742.
- La Scola B, Campocasso A, N'Dong R, Fournous G, Barrassi L, Flaudrops C, Raoult D: Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. *Intervirology* 2010;53:344–353.

- 10 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;108:17486–17491.
- 11 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;107:19508–19513.
- 12 Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, et al: Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc Natl Acad Sci USA* 2009;106:21848–21853.
- 13 Thomas V, Bertelli C, Colly F, Casson N, Telenti A, Goesmann A, Croxatto A, Greub G: Lausannevirus, a giant amoebal virus encoding histone doublets. *Environ Microbiol* 2011;13:1454–1466.
- 14 Colson P, Gimenez G, Boyer M, Fournous G, Raoult D: The giant *Cafeteria roenbergensis* virus that infects a widespread marine phagocytic protist is a new member of the fourth domain of Life. *PLoS One* 2011;6:e18935.
- 15 Iyer LM, Aravind L, Koonin EV: Common origin of four diverse families of large eukaryotic DNA viruses. *J Virol* 2001;75:11720–11734.
- 16 Iyer LM, Balaji S, Koonin EV, Aravind L: Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res* 2006;117:156–184.
- 17 Yutin N, Wolf YI, Raoult D, Koonin EV: Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol J* 2009;6:223.
- 18 Koonin EV, Yutin N: Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. *Intervirology* 2010;53:284–292.
- 19 Lefkowitz EJ, Wang C, Upton C: Poxviruses: past, present and future. *Virus Res* 2006;117:105–118.
- 20 Baroudy BM, Venkatesan S, Moss B: Incompletely base-paired flip-flop terminal loops link the two DNA strands of the vaccinia virus genome into one uninterrupted polynucleotide chain. *Cell* 1982;28:315–324.
- 21 Williams T: The iridoviruses. *Adv Virus Res* 1996;46:345–412.
- 22 Jakob NJ, Müller K, Bahr U, Darai G: Analysis of the first complete DNA sequence of an invertebrate iridovirus: coding strategy of the genome of *Chilo iridescent virus*. *Virol J* 2001;286:182–196.
- 23 Federici BA: Enveloped double-stranded DNA insect virus with novel structure and cytopathology. *Proc Natl Acad Sci USA* 1983;80:7664–7668.
- 24 Federici BA, Vlak JM, Hamm JJ: Comparative study of virion structure, protein composition and genomic DNA of three ascovirus isolates. *J Gen Virol* 1990;71:1661–1668.
- 25 Dixon LK, Abrams CC, Bowick G, Goatley LC, Kay-Jackson PC, Chapman D, Liverani E, Nix R, Silk R, Zhang F: African swine fever virus proteins involved in evading host defence systems. *Vet Immunol Immunopathol* 2004;100:117–134.
- 26 de Villiers EP, Gallardo C, Arias M, da Silva M, Upton C, Martin R, Bishop RP: Phylogenomic analysis of 11 complete African swine fever virus genome sequences. *Virology* 2010;400:128–136.
- 27 Wilson WH, Van Etten JL, Allen MJ: The Phycodnaviridae: the story of how tiny giants rule the world. *Curr Top Microbiol Immunol* 2009;328:1–42.
- 28 Van Etten JL, Lane LC, Dunigan DD: DNA viruses: the really big ones (gigaviruses). *Annu Rev Microbiol* 2010;64:83–99.
- 29 Van Etten JL: Unusual life style of giant chlorovirus. *Annu Rev Genet* 2003;37:153–195.
- 30 Wilson WH, Schroeder DC, Allen MJ, Holden MT, Parkhill J, Barrell BG, Churcher C, Hamlin N, Mungall K, Norbertczak H, et al: Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* 2005;309:1090–1092.
- 31 Fuhrman JA: Marine viruses and their biogeochemical and ecological effects. *Nature* 1999;399:541–548.
- 32 Raoult D, La Scola B, Birtles R: The discovery and characterization of Mimivirus, the largest known virus and putative pneumonia agent. *Clin Infect Dis* 2007;45:95–102.
- 33 Lwoff A: The concept of virus. *J Gen Microbiol* 1957;17:239–253.
- 34 Boyer M, Madoui M-A, Gimenez G, La Scola B, Raoult D: Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4th domain of life including giant viruses. *PLoS One* 2010;5:e15530.
- 35 Fischer MG, Suttle CA: A virophage at the origin of large DNA transposons. *Science* 2011;332:231–234.
- 36 Woese CR, Kandler O, Wheelis ML: Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 1990;87:4576–4579.
- 37 Koonin EV, Senkevich TG, Dolja VV: The ancient virus world and evolution of cells. *Biol Direct* 2006;1:29.
- 38 Ivanovski D: Über die Mosaikkrankheit der Tabakspflanze. *St Petersburg Acad Imp Sci Bull* 1892;35:67–70.
- 39 Beijerinck MW: Über ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabakblätter. *Verhandlungen der Koninklyke akademie van Wetenschappen te Amsterdam*; in Johnson J (ed): *Phytopathological Classics No 7*. St Paul, American Phytopathological Society, 1942, pp 33–52.
- 40 Lustig A, Levine AJ: One hundred years of virology. *J Virol* 1992;66:4629–4631.
- 41 Roening G, Holowczak JA: Evidence for the presence of RNA in the purified virions of vaccinia virus. *J Virol* 1974;14:704–708.
- 42 Filée J, Chandler M: Gene exchange and the origin of giant viruses. *Intervirology* 2010;53:354–361.
- 43 Koonin EV, Wolf YI, Nagasaki K, Dolja VV: The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat Rev Microbiol* 2008;6:925–939.
- 44 Raoult D: The post-Darwinist rhizome of life. *Lancet* 2010;375:104–105.
- 45 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.
- 46 Wu D, Wu M, Halpern A, Rusch DB, Yooseph S, Frazier M, Venter JC, Eisen JA: Stalking the fourth domain in metagenomic data: searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees. *PLoS One* 2011;6:e18011.
- 47 Novoa RR, Calderita G, Arranz R, Fontana J, Granzow H, Risco C: Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol Cell* 2005;97:147–172.
- 48 Suzan-Monti M, La Scola B, Barrassi L, Espinosa L, Raoult D: Ultrastructural characterization of the giant volcano-like virus factory of *Acanthamoeba polyphaga* Mimivirus. *PLoS One* 2007;2:e328.
- 49 Suzan-Monti M, La Scola B, Raoult D: Genomic and evolutionary aspects of Mimivirus. *Virus Res* 2006;117:145–155.
- 50 Suhre K: Gene and genome duplication in *Acanthamoeba polyphaga* Mimivirus. *J Virol* 2005;79:14095–14101.
- 51 Filée J: Lateral gene transfer, lineage-specific gene expansion and the evolution of nucleocytoplasmic large DNA viruses. *J Invertebr Pathol* 2009;101:169–171.
- 52 Moreira D, Brochier-Armanet C: Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC Evol Biol* 2008;8:12.
- 53 Colson P, Raoult D: Gene repertoire of amoeba-associated giant viruses. *Intervirology* 2010;53:330–343.
- 54 Shackelton LA, Holmes EC: The evolution of large DNA viruses: combining genomic information of viruses and their hosts. *Trends Microbiol* 2004;12:458–465.
- 55 McLysaght A, Baldi PF, Gaut BS: Extensive gene gain associated with adaptive evolution of poxviruses. *Proc Natl Acad Sci USA* 2003;100:15655–15660.
- 56 Evans DH, Stuart D, McFadden G: High levels of genetic recombination among co-transfected plasmid DNAs in poxvirus-infected mammalian cells. *J Virol* 1988;62:367–375.
- 57 Pickup DJ, Ink BS, Parsons BL, Hu W, Joklik WK: Spontaneous deletions and duplications of sequences in the genome of cowpox virus. *Proc Natl Acad Sci USA* 1984;81:6817–6821.

- 58 Raoult D, Boyer M: Amoebae as genitors and reservoirs of giant viruses. *Intervirology* 2010;53:321–329.
- 59 Boyer M, Azza S, Barrassi L, Klose T, Campocasso A, Pagnier I, Fournous G, Borg A, Robert C, Zhang X, et al: Mimivirus shows dramatic genome reduction after intramoebal culture. *Proc Natl Acad Sci USA* 2011;108:10296–10301.
- 60 Boyer M, Gimenez G, Suzan-Monti M, Raoult D: Classification and determination of possible origins of ORFans through analysis of nucleocytoplasmic large DNA viruses. *Intervirology* 2010;53:310–320.
- 61 Desnues C, Boyer M, Raoult D: Sputnik, a virophage infecting the viral domain of life. *Adv Vir Res* 2012;82:63–89.
- 62 Yau S, Lauro FM, Demaere MZ, Brown MV, Thomas T, Raftery MJ, Andrews-Pfannkoch C, Lewis M, Hoffman JM, Gibson JA, et al: Virophage control of antarctic algal host-virus dynamics. *Proc Natl Acad Sci USA* 2011; 108:6163–6168.
- 63 Edwards RA, Rohwer F: Viral metagenomics. *Nat Rev Microbiol* 2005;3:504–510.
- 64 Angly FE, Willner D, Prieto-Davó A, Edwards RA, Schmieder R, Vega-Thurber R, Antonopoulos DA, Barott K, Cottrell MT, Desnues C, et al: The GAAS metagenomic tool and its estimations of viral and microbial average genome size in four major biomes. *PLoS Comput Biol* 2009;5:e1000593.
- 65 Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F: Laboratory procedures to generate viral metagenomes. *Nat Protoc* 2009;4: 470–483.
- 66 Ghedin E, Claverie JM: Mimivirus relatives in the Sargasso Sea. *Virology* 2005;2:62.
- 67 Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K, et al: The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* 2007;5:e77.
- 68 Monier A, Larsen JB, Sandaa RA, Bratbak G, Claverie JM, Ogata H: Marine mimivirus relatives are probably large algal viruses. *Virology* 2008;5:12.
- 69 Ogata H, Toyoda K, Tomaru Y, Nakayama N, Shirai Y, Claverie JM, Nagasaki K: Remarkable sequence similarity between the dinoflagellate-infecting marine girus and the terrestrial pathogen African swine fever virus. *Virology* 2009;6:178.
- 70 Loh J, Zhao G, Presti RM, Holtz LR, Finkbeiner SR, Droit L, Villasana Z, Todd C, Pipas JM, Calgua B, et al: Detection of novel sequences related to African swine fever virus in human serum and sewage. *J Virol* 2009;83:13019–13025.
- 71 Kristensen DM, Mushegian AR, Dolja VV, Koonin EV: New dimensions of the virus world discovered through metagenomics. *Trends Microbiol* 2010;18:11–19.
- 72 Condit RC: Principles of virology; in Knipe DM, Howley PM (eds): *Fields Virology*, ed 5, Philadelphia, Lippincott Williams & Wilkins, 2007.
- 73 Moss B: Poxviridae: the viruses and their replication; in Fields BN, Knipe DM, Howley PM (eds): *Fields Virology*, ed 3, Philadelphia, Lippincott-Raven, 1996, pp 2637–2671.