Arch Virol (2013) 158:915–920
DOI 10.1007/s00705-012-1537-y

VIROLOGY DIVISION NEWS

"Marseilleviridae", a new family of giant viruses infecting amoebae

Philippe Colson · Isabelle Pagnier · Niyaz Yoosuf · Ghislain Fournous · Bernard La Scola · Didier Raoult

Received: 3 August 2012 / Accepted: 3 October 2012 / Published online: 29 November 2012

Abstract The family "Marseilleviridae" is a new proposed taxon for giant viruses that infect amoebae. Its first member, *Acanthamoeba polyphaga* marseillevirus (APMaV), was isolated in 2007 by culturing on amoebae a water sample collected from a cooling tower in Paris, France. APMaV has an icosahedral shape with a diameter of ≈250 nm. Its genome is a double-stranded circular DNA that is 368,454 base pairs (bp) in length. The genome has a GC content of 44.7 % and is predicted to encode 457 proteins. Phylogenetic reconstructions showed that APMaV belongs to a new viral family among nucleocytoplasmic large DNA viruses, a group of viruses that also includes *Acanthamoeba polyphaga* mimivirus (APMV) and the other members of the family *Mimiviridae* as well as the members of the families *Poxviridae*, *Phycodnaviridae*, *Iridoviridae*, *Ascoviridae*, and *Asfarviridae*. In 2011, *Acanthamoeba castellanii* lausannevirus (ACLaV), another close relative of APMaV, was isolated from river water in France. The ACLaV genome is 346,754 bp in size and encodes 450 genes, among which 320 have an APMaV protein as the closest homolog. Two other giant viruses closely related to APMaV and ACLaV have been recovered in our laboratory from a freshwater sample and a human stool sample using an amoebal co-culture method. The only currently identified hosts for "marseilleviruses" are *Acanthamoeba* spp. The prevalence of these viruses in the environment and in animals and humans remains to be determined.

P. Colson · I. Pagnier · N. Yoosuf · G. Fournous · B. La Scola · D. Raoult
URMITE UM63 CNRS 7278 IRD 198 INSERM U1905, Aix-Marseille Université, Facultés de Médecine et de Pharmacie, 27 boulevard Jean Moulin, 13385 Marseille Cedex 05, France

P. Colson · B. La Scola · D. Raoult (✉)
Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, IHU Méditerranée Infection, Centre Hospitalo-Universitaire Timone, Assistance Publique - Hôpitaux de Marseille, 264 rue Saint-Pierre, 13385 Marseille Cedex 05, France
e-mail: didier.raoult@gmail.com

Introduction

Acanthamoeba polyphaga marseillevirus (APMaV) was isolated in 2007 from water collected from a cooling tower in Paris, France, using a method based on *Acanthamoeba polyphaga* culture [1]. The name of this virus originates from the name of its amoebal host and the name of the French city, Marseille, where it was discovered. APMaV was described five years after the discovery of *Acanthamoeba polyphaga* mimivirus (APMV), the first giant virus identified using an amoebal co-culture method. APMV was revealed to be the largest known virus [2, 3]. APMaV was found to be smaller than APMV with respect to the sizes of the capsid and the genome. Nonetheless, with a capsid diameter of approximately 250 nm (Fig. 1) and a genome composed of 368,454 base pairs (bp) encoding 457 genes, APMaV represents a new giant virus. After the discovery of APMaV, other giant viruses were isolated from freshwater using the amoebal co-culture method and were briefly described in 2010 [4]. Among these new viruses, Cannes 8 virus (Ca8V) is a close relative of APMaV based on the phylogeny of the B-family DNA polymerase gene [4]. In 2011, another large DNA virus, *Acanthamoeba castellanii* lausannevirus (ACLaV), was described, and this virus was determined to be a close relative of APMaV. ACLaV was isolated by culturing freshwater collected in 2005 from the Seine River in France on amoebae [5]. An additional close relative of APMaV was recently recovered in our laboratory from the stool of a young Senegalese

man, and this new virus was named Senegal virus (SNGV) (Fig. 2) [6]. The genomes of Ca8V and SNGV have been sequenced on a 454-Roche GS20 instrument (Roche, USA) as described previously [1]. Additionally, the genome of the Ca8V isolated in our laboratory was sequenced on a SOLiD instrument (Life Technologies Corporation).

Genomics of "marseilleviruses"

The APMaV genome is a circular double-stranded DNA molecule of 368,454 base pairs (Table 1) [1]. Its GC content is 44.7 %. The APMaV genome harbors 457 open reading frames (ORFs) predicted to encode proteins with a size ranging from 50 to 1,537 amino acids. These ORFs represent 89 % of the genome. APMaV was identified as representing a unique nucleocytoplasmic large DNA virus (NCLDV) family [1, 7]. NCLDVs were described in 2001 as a monophyletic group of large viruses with a DNA genome. This group of viruses comprises the families *Poxviridae*, *Asfarviridae*, *Iridoviridae* and *Phycodnaviridae*, which were grouped together based on a set of core genes shared by all of the member viruses [8]. Later, APMV and then APMaV were found to be related to this group of viruses [1, 3], for which we recently proposed reclassification in a new viral order, "Megavirales" (talk.ictvonline.org/files/proposals/taxonomy_proposals_fungal1/m/fung01/4261.aspx) [9]. All of these giant viruses share a common and very early ancestor based on phylogenetic and phyletic analysis of conserved and informational genes [7, 10, 11]. Among the NCLDVs, APMaV branched deeply with irido-/ascoviruses on the basis of the phylogenetic reconstruction of conserved

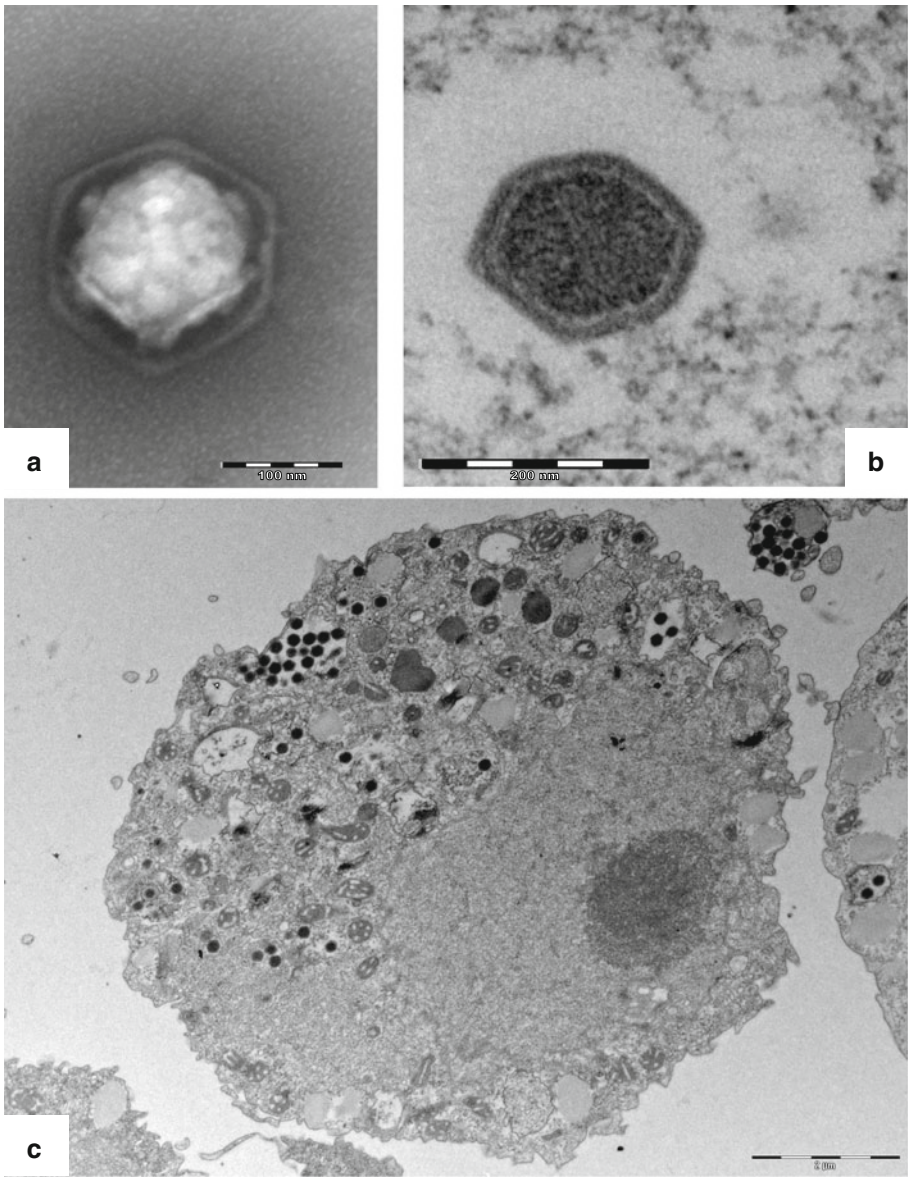

Fig. 1 Electron microscopy images of APMaV particles in a culture supernatant (*scale bar* represents 100 nm) (**a**) and in *Acanthamoeba* sp (*scale bar* represents 200 nm) (**b**), and of *Acanthamoeba* sp infected with APMaV (*scale bar* represents 2 µm) (**c**)

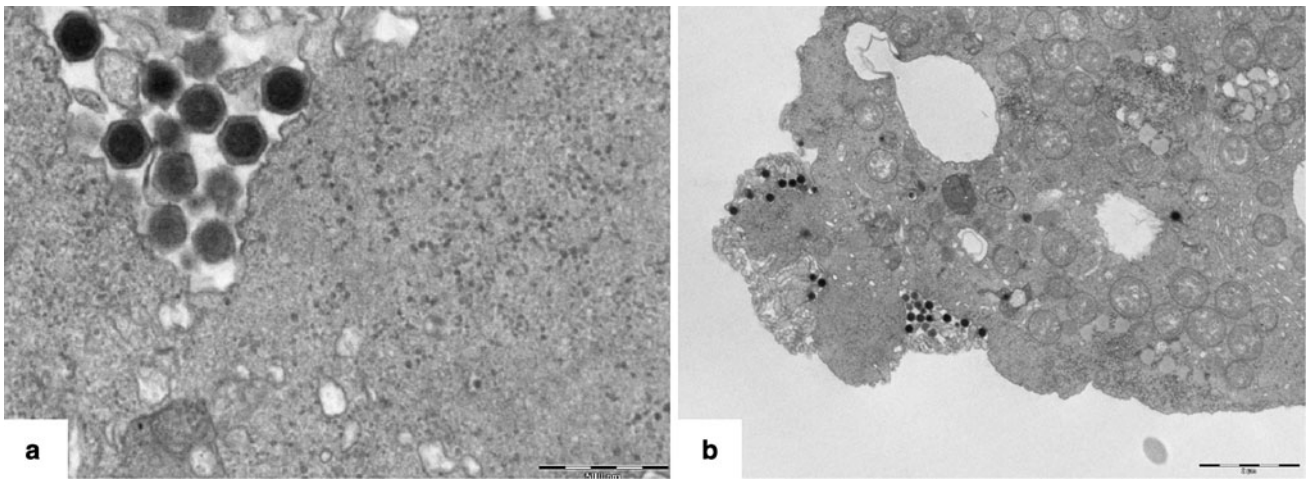

Fig. 2 Electron microscopy images of SNGV in *Acanthamoeba polyphaga*. **a** The *scale bar* represents 500 nm. **b** The *scale bar* represents 2 μm

Table 1 Description of the primary features of the "Marseilleviridae" members

Name	Source	Country/ region	Capsid size (nm)	Genome GenBank accession no.	Date of creation	Genome topology	Genome size (bp)	Number of genes	References
APMaV	Cooling tower	France (Paris)	250	NC_013756	25/01/ 2010	Circular	368,454	457	[1]
ACLaV	River (Seine)	France	190–220 nm	NC_015326	01/04/ 2011	Linear/ circular	346,754	450	[5]
Ca8V	Cooling tower	France (Cannes)	180	JF979175.1[a]	30/06/ 2012	–	374,039	–	[4]
SNGV	Human stool sample	Senegal	210	JF909596-JF909602	13/09/ 2011	–	372,690	–	[6]

[a] GenBank accession no. corresponds to the B-family DNA polymerase gene

genes [1, 7]. In contrast, a comparison of the NCLDV gene repertoires instead grouped APMaV with APMV and *Acanthamoeba polyphaga* mamavirus (APMV2). The analysis of the APMaV genome has highlighted its mosaicism and the role of the amoeba as a biological niche for gene acquisition and exchange between sympatric bacteria, viruses and their amoebal hosts [1, 12]. Thus, on the basis of phylogenetic analysis, the APMaV genome contains 51 genes (11 %) of probable NCLDV origin, 49 (11 %) of probable bacterial or bacteriophage origin, and 85 (19 %) of probable eukaryotic origin [1]. A total of 49 proteins have been identified in purified APMaV virions [1]. These proteins have been linked to several functional categories and include NCLDV core proteins, including the capsid protein. Of note, APMaV messenger RNAs, including transcripts encoding the DNA polymerase and the capsid, were found to be encapsidated in the virions.

The ACLaV genome is 346,754 bp in length, and its GC content is 42.9 % [5]. It can be circular molecule or a linear DNA molecule with terminal repeats. This genome harbors 450 ORFs that cover 93 % of the genome and have a mean length of 716 bp. ACLaV encodes homologs for all of the NCLDV core genes detected in APMaV. The phylogenetic analyses published previously showed that APMaV and ACLaV make up a new viral family among the nucleocytoplasmic large DNA viruses (NCLDVs) [1, 5, 7, 10]. This family structure has been well established in several studies using several conserved proteins, including those encoded by NCLDV core genes. Although comparative genomics and phylogenetic reconstructions have shown that ACLaV is a close relative of APMaV and that both viruses belong to the same family [5], the genomes of these two giant viruses display considerable differences [5] (Figs. 3, 4). Indeed, a total of 332 ACLaV proteins (73.8 % of the putative proteome) display significant similarity to proteins in the NCBI non-redundant sequence database, and among those proteins, only 320 (71.1 %) have an APMaV protein as the best BLASTp hit. In addition, comparative analysis of the ACLaV and APMaV genomes revealed a 150-kb region with poor synteny with many hypothetical proteins, followed by a 200-kb region with a higher level of synteny (Figs. 3, 4), [5].

Only two-thirds of the ACLaV and APMaV proteins share a best reciprocal BLAST hit.

Another giant virus, Cannes 8 virus (Ca8V), has been isolated in our laboratory from a freshwater sample using amoebal culture, and this virus has been found to be closely related to APMaV and ACLaV based on phylogenetic reconstructions (Table 1) [4]. Moreover, we obtained the first isolate of a giant virus infecting amoebae from a human sample, a stool sample from a young Senegalese man [6]. The genome (accession numbers JF909596-JF909601) of this giant virus, named Senegal virus (SNGV), has a size of approximately 373 kbp (in the same range as those of APMaV and ACLaV). The analysis of the genomes of SNGV and Ca8V demonstrated that they are bona fide new members of the proposed family "Marseilleviridae" (talk.ictvonline.org/files/proposals/taxonomy_proposals_fungal1/m/fung01/4262.aspx). Nonetheless, the genomes of SNGV and Ca8V display some differences compared with the genomes of APMaV and ACLaV (Fig. 4). The number of bidirectional best hits for APMaV and other members of the family "Marseilleviridae" tentatively ranges from 300 to 399. Overall, the ranges in size and in the number of genes for these new members of the family "Marseilleviridae" are similar to those of APMaV and ACLaV. At the present time, we propose defining only one genus, named "Marseillevirus". The species "Marseillevirus marseillevirus" is assigned to this genus and has one member, APMaV, while the "marseilleviruses" ACLaV, SNGV and Ca8V remain presently unassigned until additional "marseilleviruses" are described.

Members of the family "Marseilleviridae" (the "marseilleviruses"), like those of the family *Mimiviridae* (the mimiviruses) and other NCLDV families, do not meet the usual criteria quoted by Lwoff to define viruses [13], and the outstanding characteristics of these viruses led us recently to propose a new order made up of these giant viruses [9].

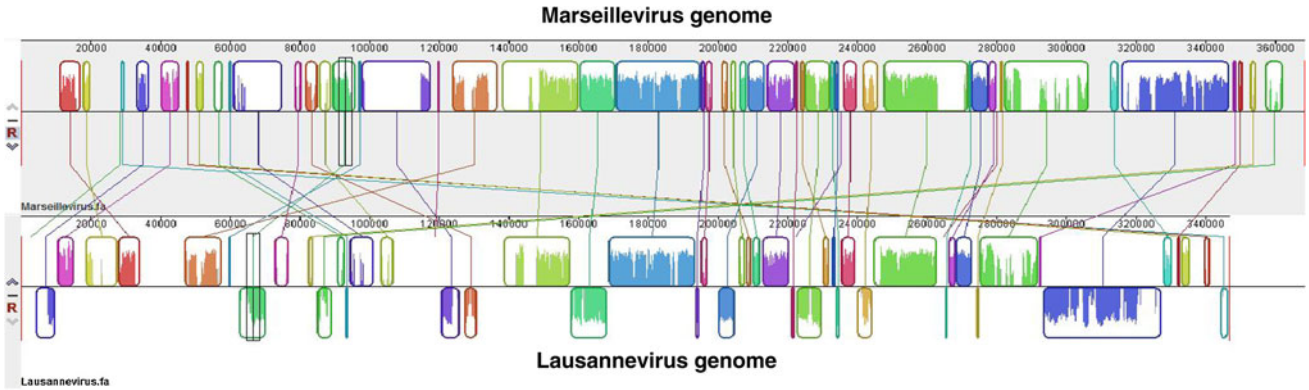

Fig. 3 Comparison and gene alignment of the genomes of APMaV and ACLaV using Mauve software [14]. Colored outlined blocks surround regions of the genome sequence that aligned to part of the other genome. The *colored bars* inside the blocks are related to the level of sequence similarity. Lines link blocks with homology between two genomes. Regions that are inverted relative to the other genome are shifted below a genome's center axis

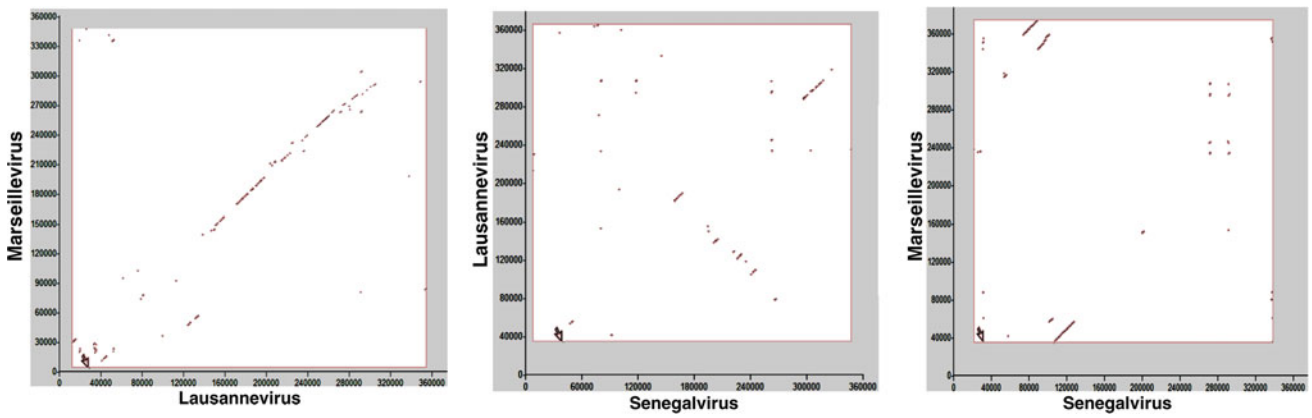

Fig. 4 *Dot plots* for the comparisons of the APMaV, ACLaV, and SNGV genomes using Owen software [15]

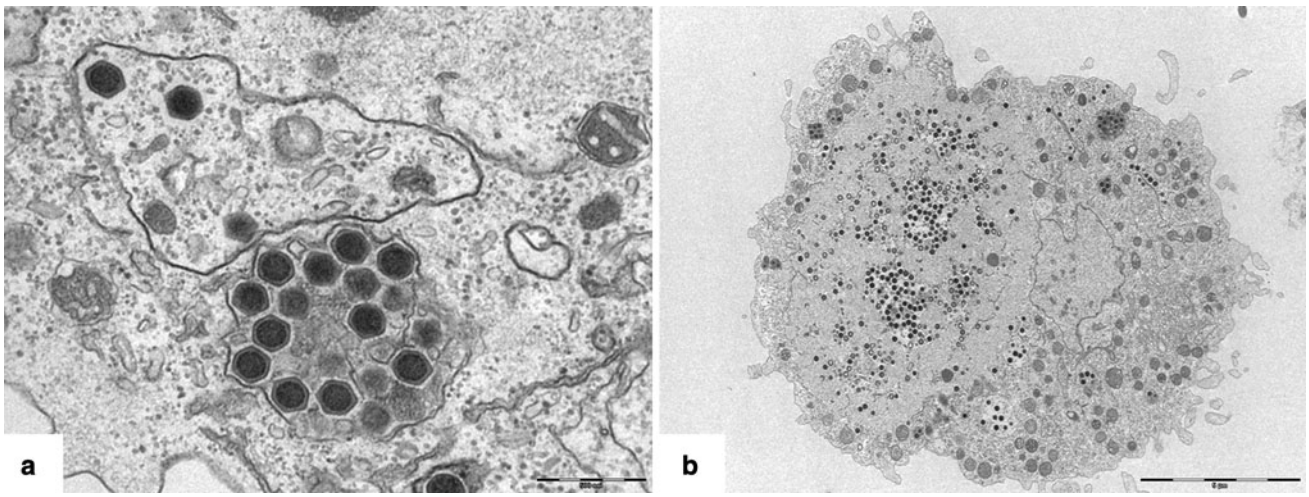

Fig. 5 Electron microscopy images of Ca8V in *Acanthamoeba polyphaga* (*scale bar* represents 500 nm) (**a**) and of *Acanthamoeba polyphaga* infected with Ca8V (*scale bar* represents 5 μm) (**b**)

Morphological properties

APMaV, ACLaV, SNGV and Ca8V share similar morphological features, including the size of their capsids, which ranges from 190 to 250 nm (Fig. 1, 2, 5; Table 1). APMaV has an icosahedral shape and a diameter of ≈250 nm (Table 1; Fig. 1) [1]. The capsid shell has a thickness of ≈10 nm, and 12-nm-long fibers with globular ends are present at the viral surface. A membrane may surround the nucleocapsid, which is separated from the capsid shell by a gap of ≈52 nm. For all members of the family "Marseilleviridae", viral factories can be observed during the replication cycle. These viral factories have different appearances than those observed for APMV and APMV2, tending to be more widely distributed in the amoebal cytoplasm.

Properties in culture

All of the currently identified members of the family "Marseilleviridae" have *Acanthamoeba* spp. as their hosts and were isolated by culturing samples on these amoebae. In amoebal culture, APMaV enters the amoeba 30-60 min post-infection (p.i.) [1]. Later, a viral factory appears close to the nucleus of the amoeba. Capsid assembly and viral genome encapsidation are observed simultaneously in these viral factories, leading to mature and immature APMaV particles. The replication cycle is complete at 5 h p.i, which is a short period of time compared to that observed for APMV. The morphology of the host-cell nucleus changes considerably between 30 min and 2.5 h p.i. Regarding ACLaV, a few viral particles are present 30 min p.i [5]. After an eclipse phase, viruses can be observed again, in large vesicles, at 4 h p.i., and they fill the entire amoeba at 8 h p.i. before amoebal lysis at 16 h p.i.

Prevalence, host, and pathogenicity

The prevalence of "marseilleviruses" in environmental samples is currently unknown. Of note, four of these viruses were recently recovered from 103 water samples [4]. The only currently identified hosts for "marseilleviruses" are *Acanthamoeba* spp. [1, 4, 5]. No data are currently available on the prevalence of "marseilleviruses" in human or animal samples, and no pathogenic role has been demonstrated to date, but one virus belonging to the family "Marseilleviridae", SNGV, has been isolated from a human stool sample [6].

Conclusion

Acanthamoeba polyphaga marseillevirus (APMaV) and its close relatives exhibit remarkable features that are shared by mimiviruses and have contributed to a considerable increase in the interest in NCLDVs and to the better delineation of this group of giant viruses, for which we have recently proposed a new viral order named "Megavirales" [1, 5, 9]. The family "Marseilleviridae" would be included in the order "Megavirales". Further isolates will most likely be described that will be closely related to APMaV. We are performing comparative genomics analysis of the genomes of new putative "marseilleviruses" and will submit these new genomes to sequence databases. We believe that these viruses should be linked to a viral family.

Conflict of interest All of the authors declare that they have no potential conflict of interest.

References

1. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, 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
2. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie JM, Raoult D (2003) A giant virus in amoebae. Science 299:2033
3. 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
4. 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
5. 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
6. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, Trape JF, Koonin E, Koonin EV, La Scola B, Raoult D (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect (in press)
7. Koonin EV, Yutin N (2010) Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology 53:284–292
8. Iyer LM, Aravind L, Koonin EV (2001) Common origin of four diverse families of large eukaryotic DNA viruses. J Virol 75:11720–11734
9. 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
10. 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
11. Boyer M, Madoui MA, Gimenez G, La Scola B, Raoult D (2010) Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. PLoS One 5:e15530
12. Raoult D, Boyer M (2010) Amoebae as genitors and reservoirs of giant viruses. Intervirology 53:321–329
13. Lwoff A (1957) The concept of virus. J Gen Microbiol 17:239–253
14. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403
15. Ogurtsov AY, Roytberg MA, Shabalina SA, Kondrashov AS (2002) OWEN: aligning long collinear regions of genomes. Bioinformatics 18:1703–1704