MODULE 1: **TITLE, AUTHORS, etc**

Code assigned:	***2012.002a-hF***	(to be completed by ICTV officers)

Short title: *Marseilleviridae*, a new family of giant viruses that infect *Acanthamoeba* spp.

Modules attached (modules 1 and 9 are required)	**1** ☒	**2** ☒	**3** ☒	**4** ☐	**5** ☒
	6 ☐	**7** ☐	**8** ☐	**9** ☒	

Author(s) with e-mail address(es) of the proposer:

Didier RAOULT (didier.raoult@gmail.com), Bernard LA SCOLA (bernard.lascola@univmed.fr), Isabelle PAGNIER (isabelle.pagnier@univ-amu.fr), Gibert GREUB (gilbert.greub@chuv.ch); Philippe COLSON (philippe.colson@univ-amu.fr)

List the ICTV study group(s) that have seen this proposal:

A list of study groups and contacts is provided at http://www.ictvonline.org/subcommittees.asp . If in doubt, contact the appropriate subcommittee chair (fungal, invertebrate, plant, prokaryote or vertebrate viruses)	

ICTV-EC or Study Group comments and response of the proposer:

The ICTV-EC considered your proposals on the creation of two new species and a new family Below are the comments that need to be addressed:

1. The two proposed species names refer to laboratory, and not necessarily, natural hosts. This is a problem not only for your viruses, but also for other viruses that are isolated from the environment. This issue is currently being considered by EC and the relevant Study Groups.
Answer: Regarding the name of the proposed species, there is no taxonomic rule for naming and choosing the name belongs to the discoverers, as has been the case in the history of mankind.

2. The terms "Marseillevirus" and "Lausannevirus" should not be capitalized. There are instances where capitals are retained in species names, but these two should not be among them. "Marseille" is a place (capitalized) but "marseillevirus" is not, rather a derivative term.
Answer: This has been considered.

3. It is unsatisfactory that the question of genera is not addressed adequately. Taxonomy is preferably done from the bottom up, and at least one of the viruses could be put into a genus at this stage. In this respect, the species names are again somewhat problematic. If you intend to indicate two eventual genera (Marseillevirus and Lausannevirus), one can see the sense. If it

will be one genus (Marseillevirus), the use of "lausannevirus" may confuse.
Answer: We are currently proposing one genus: "Marseillevirus" in the proposed family "Marseilleviridae".

4. For stability reasons, it would be better to supply GenBank, rather than RefSeq, accession numbers.
Answer: As advised, we supply GenBank, rather than RefSeq, accession numbers in the revised version of our proposal.

5. Despite what the legends say, the trees lack posteriors at the nodes. Also, it would help to have the families marked on Fig. 1.
Answer: This has been modified accordingly.

6. Other than marseillevirus and lausannevirus, has the pertinent information on any of the other isolates (Cannes 8 virus, Cannes 9 virus, etc) been published?
Answer: Articles that describes senegalvirus and tunisvirus have been published (they are referenced in this revised version of the proposal). Cannes 8 virus is not quoted here.

ICTV-EC comments

Provide GenBank accession numbers for all four proposed species. Remove references to *Megavirales*. Change species names, e.g. by adding spaces: *Acanthamoeba polyphaga Marseille virus, Acanthamoeba polyphaga Senegal virus, Acanthamoeba castellanii Lausanne virus, Acanthamoeba polyphaga Tunis virus*.

Response of Authors:

Provide GenBank accession numbers for all four proposed species.
This has been done.

Remove references to *Megavirales*.
This has been done.

Change species names, e.g. by adding spaces: *Acanthamoeba polyphaga Marseille virus, Acanthamoeba polyphaga Senegal virus, Acanthamoeba castellanii Lausanne virus, Acanthamoeba polyphaga Tunis virus*.
We chose to propose *Marseillevirus marseillevirus* and *Senegalvirus marseillevirus*.

Date first submitted to ICTV:	June 2012
Date of this revision (if different to above):	June 2013

MODULE 2: **NEW SPECIES**

creating and naming one or more new species.

Code	***2012.002aF***	(assigned by ICTV officers)

To create 2 new species within:

Genus:	***Marseillevirus***
Subfamily:	
Family:	***Marseilleviridae*** (new)
Order:	

And name the new species:	**GenBank sequence accession number(s) of reference isolate:**
Marseillevirus marseillevirus *Senegalvirus marseillevirus*	isolate T19: NC_013756.1 isolate SSV: JF909596-602

Reasons to justify the creation and assignment of the new species:

Marseillevirus marseilleviru isolate T19 (MVMV-T19) has been discovered in 2007 in water collected from a cooling tower in Paris, France, and has been isolated by culturing on *Acanthamoeba polyphaga* (Boyer et al., 2009). This discovery of MVMV-T19 has corresponded to the identification of a putative new viral family, the "Marseilleviridae" (Boyer et al., 2009; Yutin et al., 2009; Yutin et al., 2012). Phylogenetic and phyletic analyses previously published showed that MVMV-T19 is a member of a new viral family amongst the nucleocytoplasmic large DNA viruses (NCLDVs), a monophyletic viral group first described in 2001 (Iyer et al., 2001; Iyer et al., 2006; Boyer et al., 2009; Yutin et al., 2009; Koonin and Yutin, 2010; Colson et al., 2011; Yutin et al., 2012) (Figure 1). This has been well established in several studies using several conserved proteins including NCLDV core genes. Thus, MM-T19 represents the first described species for the proposed family "Marseilleviridae", and its founding member.
MVMV-T19 virions have an icosahedral shape and a diameter of approximately 250 nm (Boyer et al., 2009) with circular, double-stranded DNA molecule of 368,453 bp (Accession number: NC_013756.1). A total of 457 ORFs have been predicted to encode proteins. It has been highlighted that the MM-T19 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 intraamoebal lifestyle (Raoult and Boyer, 2010).

Senegalvirus marseillevirus isolate SSV (SVMV-SSV**)** has been described in 2012 (Lagier et al., 2012). This virus is the first isolate of a giant virus of amoeba from a human sample, which was a stool sample from a young Senegalese; isolation was performed on *Acanthamoeba polyphaga*.
The genome of this giant virus (Accession number: JF909596-JF909601) has a size around 372,690 base pairs (Lagier et al., 2012; Colson et al., 2013). The genome annotation of SM-SSV shows that it is a close relative to MVMV-T19, and a bona fide new marseillevirus. Nonetheless, the SVMV-SSV genome displays some differences compared to those of MM-T19, and Lausannevirus VT (LV-VT). Thus, 351 and 253 of its 479 predicted proteins are bona

fide orthologs to MVMV-T19 and LV-VT proteins, respectively.
The mean (±standard deviation (SD)) amino acid identity between SM-SSV and MM-T19 protein in pairs is 97±7%, whereas the mean identity for SVMV-SSV and L-VT protein pairs is 59±16%. Thus, the level of homology for the SM-SSV proteins is greater with their marseillevirus counterparts than with LV-VT. Congruently with comparative genomics, phylogeny reconstruction based on the family-B DNA polymerase shows that SM-SSV is clustered with MM-T19 within the proposed family “Marseilleviridae” (Figure 2).

MODULE 2: **NEW SPECIES**

creating and naming one or more new species.

Code	***2012.002bF***	(assigned by ICTV officers)

To create 2 new species within:

Genus:	unassigned
Subfamily:	
Family:	***Marseilleviridae*** (new)
Order:	

And name the new species:	**GenBank sequence accession number(s) of reference isolate:**
Lausannevirus *Tunisvirus*	isolate VT: NC_015326.1 isolate fontaine2 : KF483846

Reasons to justify the creation and assignment of the new species:

Lausannevirus isolate VT (LV-VT) has been described in 2011 (Thomas et al., 2011). This virus has been isolated using amoebal co-culture from freshwater collected in 2005 from the Seine river, France. The genome of LV-VT (Accession number: NC_015326.1) is 346,754 bp in length. This genome harbors 450 ORFs that cover 93% of the genome and have a mean length of 716 bp. Phylogenetic and phyletic analyses previously published showed that MM-T19 and LV-VT are members of a new viral family amongst the nucleocytoplasmic large DNA viruses (NCLDVs), a monophyletic viral group first described in 2001 (Iyer et al., 2001; Iyer et al., 2006; Boyer et al., 2009; Thomas et al., 2011; Yutin et al., 2009; Koonin and Yutin, 2010; Colson et al., 2011; Yutin et al., 2012) (Figure 1). This has been well established in several studies using several conserved proteins including NCLDV core genes.
Criteria for species demarcation are based on phylogenetic analysis and comparative genomics: although phylogeny reconstructions and comparative genomics have shown that LV-VT to be a close relative of MVMV-T19 and that both viruses belong to a same family, the genomes of the two giant viruses display considerable differences (Thomas et al., 2011). Indeed, a total of 332 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 a MM-T19 protein as the best BLASTp hit. In addition, comparative analysis of the genomes of LV-VT and MVMV-T19 shows a 150-kilo base pairs region with poor synteny with many hypothetical proteins followed by a 200-kilo base pairs region with a higher level of synteny, while only two-third of the LV-VT and MVMV-T19 proteins share a best reciprocal BLAST hit.

Tunisvirus isolate fontaine2 (TuV-font.2) has been described in 2012 (Boughalmi et al., 2012). This giant virus has been isolated from water from of fountain in Tunis, Tunisia, by culturing on *Acanthamoeba* spp. The viral genome (submitted to GenBank) is 380,011 base pairs in length (currently making this genome the largest among marseilleviruses).
The TuV-font.2 genome shows that this virus is a relative to MVMV-T19 and LV-VT and a bona fide new marseillevirus, although TuV-Font.2 genome displays some differences compared to those of MVMV-T19, LV-VT and SVMV-SSV. A total of 484 proteins were predicted from the TuV-font.2 genome, and 320 and 358 T-font.2 proteins found a hit against

MVMV-T19 and LV-TV protein, including 259 and 299 best reciprocal hits, respectively. The mean amino acid identity between TuV-font.2 and MVMV-T19 proteins in pairs is 60%, whereas the mean identity for TuV-font.2 and LV-TV protein pairs is 76%. Congruently with comparative genomics, phylogeny reconstruction based on the family-B DNA polymerase shows that TuV-font.2 is clustered with MVMV-T19 within the proposed family "Marseilleviridae" (Figure 2).

MODULE 3: **NEW GENUS**

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	***2012.002cF***	(assigned by ICTV officers)

To create a new genus within:

Subfamily:	
Family:	***Marseilleviridae*** (new)
Order:	

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no family is specified, enter "**unassigned**" in the family box

naming a new genus

Code	***2012.002dF***	(assigned by ICTV officers)

To name the new genus: *Marseillevirus*

Assigning the type species and other species to a new genus

Code	***2012.002eF***	(assigned by ICTV officers)

To designate the following as the type species of the new genus

Marseillevirus marseillevirus	Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered

The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of species (including the type species) that the genus will contain:**

2: *Marseillevirus marseillevirus*
Senegalvirus marseillevirus

Reasons to justify the creation of a new genus:

Additional material in support of this proposal may be presented in the Appendix, Module 9

The recent discovery of MVMV-T19 has corresponded to the identification of a putative new viral family, the *Marseilleviridae* (Boyer et al., 2009; Yutin et al., 2009). MVMV-T19 has been discovered in 2007 in water collected from a cooling tower in Paris, France, and has been isolated by culturing on *Acanthamoeba polyphaga* (Boyer et al., 2009). This virus has an icosahedral shape and a diameter of approximately 250 nm Its genome (Accession number: NC_013756.1) is a circular, double-stranded DNA molecule of 368,453 bp. A total of 457 ORFs have been predicted to encode proteins. It has been highlighted that the MVMV-T19 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 intraamoebal lifestyle (Raoult and Boyer, 2010). Later, LV-VT, a close relative to marseillevirus, has been described (Thomas et al., 2011). LV-VT has been isolated using amoebal co-culture from freshwater collected in 2005 from the Seine river, France. The genome of LV-VT (Accession number: NC_015326.1) is 346,754 bp in length. This genome harbors 450 ORFs that cover 93% of the genome and have a mean length of 716 bp. A total of 332 proteins (74% of the putative

proteome) display significant similarity to proteins in the NCBI non-redundant sequence database, among which 320 (71%) have a MM-T19 protein as the best BLASTp hit.

Phylogenetic and phyletic analyses previously published showed that marseillevirus and lausannevirus compose a new viral family amongst the nucleocytoplasmic large DNA viruses (NCLDVs), an apparently monophyletic group of viruses infecting eukaryotes that was first described in 2001 (Iyer et al., 2001; Iyer et al., 2006; Boyer et al., 2009; Thomas et al., 2011; Yutin et al., 2009; Koonin and Yutin, 2010; Colson et al., 2011; Yutin et al., 2012) (Figure 1). This has been well established in several studies using several conserved proteins including NCLDV core genes. Thus, *Marseillevirus marseillevirus* represents the first described species for the proposed family "Marseilleviridae", and its founding member. The NCLDV encompass the families *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, and *Phycodnaviridae* and the two distinct groups of giant viruses that have been isolated from *Acanthamoeba* giving rise to the now established family *Mimiviridae* and the proposed family "Marseilleviridae" (Colson et al., 2013).

Several giant viruses isolated in our laboratory using amoebal cultures have been found to be closely related to MVMV-T19 and LV-VT based on phylogeny reconstructions (La Scola et al., 2010; Boughalmi et al., 2012). Moreover, we have obtained the first isolate of a giant virus of amoeba from a human sample, which was a stool sample from a young Senegalese (Lagier et al., 2012; Colson et al., 2013). A "Marseillevirus family" is described in the NCBI GenBank database with the taxonomy ID: 944644, and marseillevirus (Boyer et al., 2009), lausannevirus (Thomas et al., 2011), senegalvirus (Lagier et al., 2012; Colson et al., 2013), tunisvirus (Boughalmi et al. 2012) are species linked to this family (Figure 2).

Although not as large as mimiviruses, these marseilleviruses have very large genomes and particle sizes (approximately 350 kilobase pairs and 200 nm, respectively). Between marseillevirus, lausannevirus, senegalvirus and tunisvirus, numbers of bona fide orthologs are 210-233, which represent 45-55% of the gene repertoire of these viruses, and mean amino acid identity between pairs of bona fide orthologs is 60-76%. In addition to these findings based on the analysis of their genomes, the "marseilleviruses" share similar morphological features including the size of their capsid that ranges from 190 to 250 nm. Moreover, viral factories are seen during the replication cycle and have a different aspect compared to those observed for *Acanthamoeba polyphaga* mimivirus and *Acanthamoeba castellanii* mamavirus. In addition, all the currently identified marseilleviruses have *Acanthamoeba* spp. as host and were isolated through culturing on these amoebae.

As a family *Marseilleviridae* is proposed and four members of this proposed family have been described (and the description of three new viruses is on-going), we propose to create a new viral genus named *Marseillevirus*.

Origin of the new genus name:

Name originates from that of the leading member of the proposed family and genus: *Marseillevirus*, which has been named in reference to Marseille, the French city where this giant virus has been discovered.

Reasons to justify the choice of type species:

Marseillevirus marseillevirus T19 is the first virus discovered (2007) and founding member of this proposed genus *Marseillevirus* and the proposed family *Marseilleviridae*.

Species demarcation criteria in the new genus:

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

The criteria being used for species demarcation are genome size, number of predicted genes, number of proteins involved in pairs of reciprocal best hits with another species of the genus, number of orthologs, amino acid identity between pairs of proteins being reciprocal best hits and orthologs.

The G+C% and genome size of the marseilleviruses ranges from 42.9 to 44.7 and from 346,754 to 380,011, respectively. Between marseillevirus, lausannevirus, senegalvirus and tunisvirus, numbers of bona fide orthologs are 210-351, which represent 45-73% of the viral gene repertoires, and mean amino acid identity between pairs of bona fide orthologs ranges from 60% to 97%.

MODULE 5: **NEW FAMILY**

creating and naming a new family

Code	***2012.002fF***	(assigned by ICTV officers)
To create a new family containing the subfamilies and/or genera listed below		

Code	***2012.002gF***	(assigned by ICTV officers)
To name the new family: *Marseilleviridae*		

assigning subfamilies, genera and unassigned species to a new family

Code		(assigned by ICTV officers)
To assign the following subfamilies (if any) to the new family:		

Code	***2012.002hF***	(assigned by ICTV officers)

To assign the following genera to the new family:

You may list several genera here. For each genus, please state whether it is new or existing.

- If the genus is new, it must be created in Module 3

If the genus already exists, please state whether it is currently unassigned or is to be removed from another family. If the latter, complete Module 7 to 'REMOVE' it from that family

Marseillevirus

Please enter here the TOTAL number of unassigned species that the family will contain (those NOT within any of the genera or subfamilies listed above):

2: *Lausannevirus*
Tunisvirus

Reasons to justify the creation of the new family:

The recent discovery of species *Marseillvirus marseillvirus* has corresponded to the identification of a putative new viral family, the *Marseilleviridae* (Boyer et al., 2009; Yutin et al., 2009). MVMV-T19 has been discovered in 2007 in water collected from a cooling tower in Paris, France, and has been isolated by culturing on *Acanthamoeba polyphaga* (Boyer et al., 2009). This virus has an icosahedral shape and a diameter of approximately 250 nm Its genome (Accession number: NC_013756.1) is a circular, double-stranded DNA molecule of 368,453 bp. A total of 457 ORFs have been predicted to encode proteins. It has been highlighted that the Marseillevirus 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 intraamoebal lifestyle (Raoult and Boyer, 2010). Later, LV-VT, a close relative to MVMV-T19, has been described (Thomas et al., 2011). LV-VT has been isolated using amoebal co-culture from freshwater collected in 2005 from the Seine river, France. The LV-VT genome (Accession number: NC_015326.1) is 346,754 bp in length. This genome harbors 450 ORFs that cover 93% of the genome and have a mean length of 716 bp. A total of 332 proteins (74% of the putative proteome) display significant similarity to proteins in the NCBI non-redundant sequence database, among which 320 (71%) have a marseillevirus protein as the best BLASTp hit.

Phylogenetic and phyletic analyses previously published showed that MVMV-T19 and L-VT compose a new viral family amongst the nucleocytoplasmic large DNA viruses (NCLDVs), an apparently monophyletic group of viruses infecting eukaryotes that was first described in 2001 (Iyer et al., 2001; Iyer et al., 2006; Boyer et al., 2009; Thomas et al., 2011; Yutin et al., 2009; Koonin and Yutin, 2010; Colson et al., 2011; Yutin et al., 2012) (Figure 1). This has been well established in several studies using several conserved proteins including NCLDV core genes. Thus, *Acanthamoeba polyphaga Marseille virus* represents the first described species for the proposed family "Marseilleviridae", and its founding member. The NCLDV encompass the families *Poxviridae, Asfarviridae, Iridoviridae, Ascoviridae,* and *Phycodnaviridae* and the two distinct groups of giant viruses that have been isolated from *Acanthamoeba* giving rise to the now established family *Mimiviridae* and the proposed family "Marseilleviridae" (Colson et al., 2013).
The common origin of all these viruses from the same ancestral virus relied on phylogenetic and phyletic analyses (Yutin et al., 2009; Yutin et al., 2012). All of the NCLDV share 5 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 NCLDV, were assigned, with high confidence, to the common ancestor of the entire group.

Several other giant viruses have been isolated in our laboratory from freshwater samples using amoebal cultures, and these viruses have been found to be closely related to MVMV-T19 and LV-VT based on phylogeny reconstructions, including TuV-font.2 (La Scola et al., 2010; Boughalmi et al., 2012) (Figure 2). Moreover, we have obtained the first isolate of a giant virus of amoeba from a human sample, which was a stool sample from a young Senegalese (Lagier et al., 2012; Colson et al., 2013). The genome of this giant virus (Accession number: JF909596-JF909601) has a size around 372,690 bp (currently making this genome the largest among marseilleviruses) and its analysis shows that the virus from Senegal is a close relative to MVMV-T19, and a bona fide new marseillevirus (Figure 2). Nonetheless, SM-SSV genome displays some differences compared to those of MVMV-T19 and LV-VT; 351 and 253 of its predicted proteins are bona fide orthologs of MVMV-T19 and LV-VT proteins, respectively.
A "Marseillevirus family" is described in the NCBI GenBank database with the taxonomy ID: 944644, and *Marseillevirus marseillevirus* (Boyer et al., 2009), lausannevirus (Thomas et al., 2011), *Senegalvirus marseillevirus* (Lagier et al., 2012; Colson et al., 2013), and tunisvirus (Boughalmi et al. 2012) are species linked to this family (Figure 2).
Although not as large as mimiviruses, these marseilleviruses have very large genomes and particle sizes (approximately 350 kilobase pairs and 200 nm, respectively). The G+C% and genome size of the marseilleviruses ranges from 42.9 to 44.7 and from 346,754 to 380,011, respectively. Between MVMV-T19, L-VT, SVMV-SSV and TuV-font.2, numbers of bona fide orthologs are 210-351, which represent 45-73% of the viral gene repertoires, and mean amino acid identity between pairs of bona fide orthologs ranges from 60% to 97%.
In addition to results based on the analysis of their genomes, marseillevirus, lausannevirus, senegalvirus and tunisvirus share similar morphological features including the size of their capsid that ranges from 190 to 250 nm. Moreover, viral factories are seen during the replication cycle and have a different aspect compared to those observed for *Acanthamoeba polyphaga* mimivirus and *Acanthamoeba castellanii* mamavirus. In addition, all the currently identified marseilleviruses have *Acanthamoeba* spp. as host and were isolated through culturing on these amoebae.

As four members of the proposed family “Marseilleviridae” have been described (and the description of three new viruses is on-going), we believe it is needed that these viruses can be linked to a viral family. As stated in the ICTV principles of nomenclature, “the primary purpose of naming a taxon is to supply a means of referring to the taxon”. Besides, the remarkable features of *Acanthamoeba polyphaga* mimivirus and MM-T19 have led to a considerable increase in interest for nucleocytoplasmic large DNA viruses and have largely contributed to better delineate this group of giant viruses.
Origin of the new family name:
Name originates from genus name of the type member: *Marseillevirus marseillevirus*, which has been named in reference to Marseille, the French city where it has been discovered.

MODULE 9: **APPENDIX**: supporting material

References:

Boughalmi M, Saadi H, Pagnier I, Colson P, Fournous G, Raoult D, La Scola B. High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. Environ Microbiol 2012 Dec 12.

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(12):e15530.

Boyer M, Yutin N, Pagnier I, et al. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci U S A 2009;106:21848-53.

Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, Yoosuf N, Million M, La Scola B, Raoult D. Evidence of the megavirome in humans. J Clin Virol 2013;57(3):191-200.

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(4):e18935.

Colson P, Pagnier I, Yoosuf N, Fournous G, La Scola B, Raoult D. "Marseilleviridae", a new family of giant viruses infecting amoebae. Arch Virol 2013;158(4):915-20.

Iyer LM, Aravind L, Koonin EV. Common origin of four diverse families of large eukaryotic DNA viruses. J Virol 2001;75:11720-34.

Iyer LM, Balaji S, Koonin EV, Aravind L. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res 2006;117(1):156-84.

Koonin EV, Yutin N. Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology 2010;53(5):284-92.

La Scola B, Campocasso A, N'Dong R, et al. Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology 2010;53:344-53.

Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, Trape JF, Koonin EV, La Scola B, Raoult D. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18(12):1185-93.

Raoult D, Boyer M. Amoebae as genitors and reservoirs of giant viruses. Intervirology. 2010;53(5):321-9.

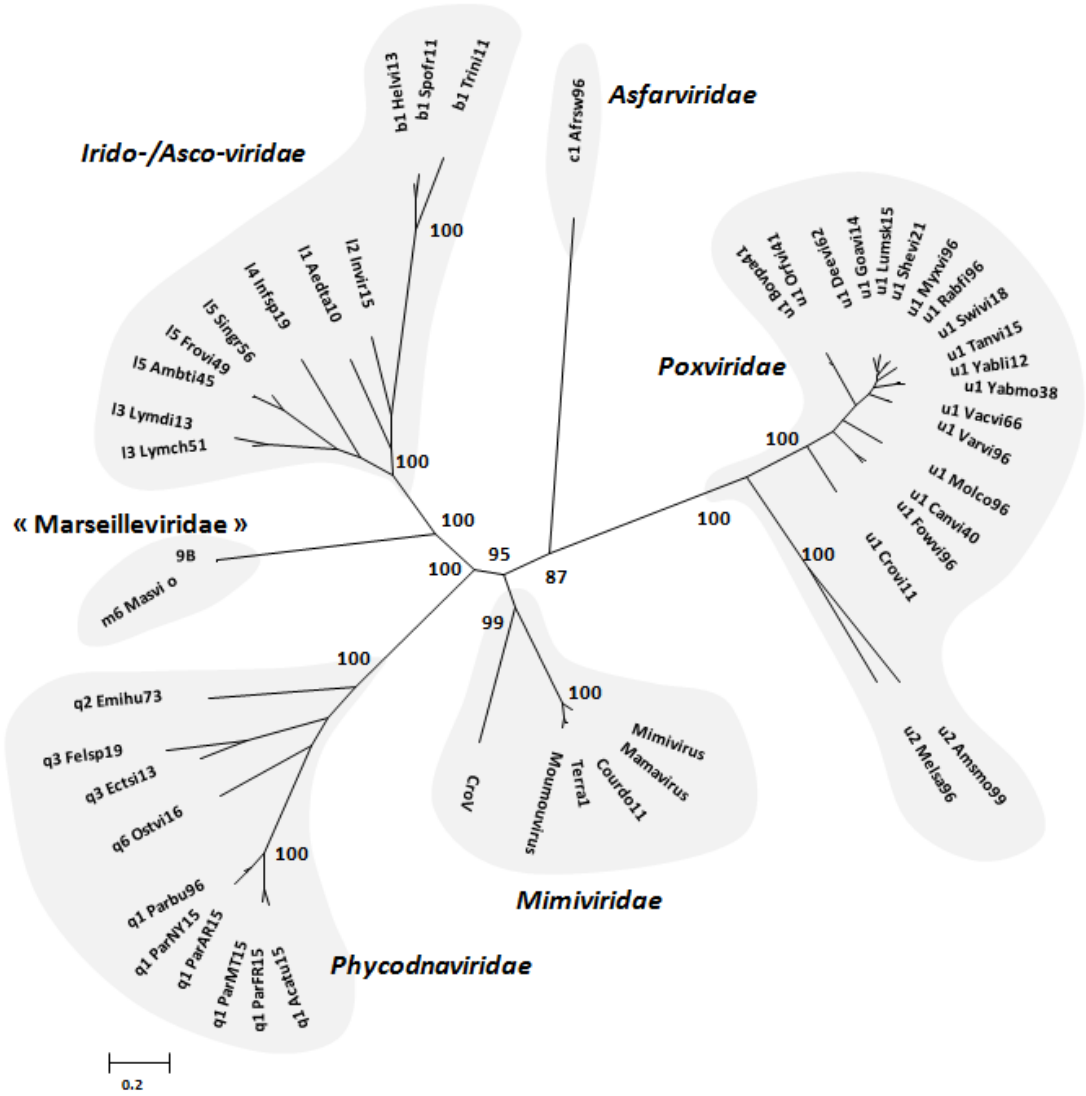
Thomas V, Bertelli C, Collyn F, et al. Lausannevirus, a giant amoebal virus encoding histone doublets. Environ Microbiol 2011;13:1454-66.

Yutin N, Koonin EV. Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. Virol J. 2012 Aug 14;9:161.

Yutin N, Wolf YI, Raoult D, et al. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol J 2009;6:223.

Annexe:

Figure 1. Bayesian phylogeny reconstruction from a cured concatenated alignment of 4 universal NCVOGs (496 conserved positions), including CroV corresponding proteins: primase-helicase (NCVOG0023), DNA polymerase (NCVOG0038), packaging ATPase (NCVOG0249), and A2L-like transcription factor (NCVOG0262).



Legend to figure 1: Bayesian posterior 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. Abbreviated names for NCLDVs: b1_Helvi, Heliothis virescens ascovirus 3e; b1_Spofr, Spodoptera frugiperda ascovirus 1a; b1_Trini, Trichoplusia ni ascovirus 2c; c1_Afrsw, African swine fever virus; l1_Aedta, Aedes taeniorhynchus iridescent virus (Invertebrate iridescent virus 3); l2_Invir, Invertebrate iridescent virus 6; l3_Lymch, Lymphocystis disease virus - isolate China; l3_Lymdi, Lymphocystis disease virus 1; l4_Infsp, Infectious spleen and kidney necrosis virus; l5_Ambti, Ambystoma tigrinum virus; l5_Frovi, Frog virus 3; l5_Singr, Singapore grouper iridovirus; m6_Masvi, Marseillevirus marseillevirus; q1_Acatu, Acanthocystis turfacea Chlorella virus 1; q1_ParAR, Paramecium bursaria Chlorella virus AR158; q1_Parbu, Paramecium bursaria Chlorella virus 1; q1_ParFR, Paramecium bursaria Chlorella virus FR483; q1_ParMT, Paramecium bursaria chlorella virus MT325; q1_ParNY, Paramecium bursaria Chlorella virus NY2A; q2_Emihu, Emiliania huxleyi virus 86; q3_Ectsi, Ectocarpus siliculosus virus 1; q3_Felsp, Feldmannia species virus; q6_Ostvi, Ostreococcus virus OsV5; u1_Bovpa, Bovine papular stomatitis virus; u1_Canvi, Canarypox virus; u1_Crovi, Crocodilepox virus; u1_Deevi, Deerpox virus W-848-83; u1_Fowvi, Fowlpox virus; u1_Goavi, Goatpox virus Pellor; u1_Lumsk, Lumpy skin disease virus NI-2490; u1_Molco, Molluscum contagiosum virus; u1_Myxvi, Myxoma virus;

u1_Orfvi, Orf virus, complete genome; u1_Rabfi, Rabbit fibroma virus; u1_Shevi, Sheeppox virus 17077-99; u1_Swivi, Swinepox virus; u1_Tanvi, Tanapox virus; u1_Vacvi, Vaccinia virus; u1_Varvi, Variola virus (smallpox virus); u1_Yabli, Yaba-like disease virus. (Adapted from [Colson et al., 2011]).

Figure 2. Phylogenic tree for family B DNA polymerase for the "Marseilleviridae" inferred by using the Maximum Likelihood method using mimiviruses as outgroup