"*Mamonoviridae*", a proposed new family of the phylum *Nucleocytoviricota*

Ruixuan Zhang[a], Masaharu Takemura[b], Kazuyoshi Murata[c], Hiroyuki Ogata[a]

Affiliations

[a]Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji 611-0011, Japan

[b]Laboratory of Biology, Institute of Arts and Sciences, Tokyo University of Science, Shinjuku, Tokyo 162-8601, Japan

[c]Exploratory Research Center on Life and Living Systems (ExCELLS) & National Institute for Physiological Sciences, National Institutes of Natural Sciences, 38 Nishigonaka, Okazaki, Aichi, 444-8585, Japan

Corresponding author at: ogata@kuicr.kyoto-u.ac.jp

Abstract

Acanthamoeba castellanii medusavirus J1 is a giant virus isolated from a hot spring in Japan in 2019. Recently, a close relative of this virus was also isolated in Japan and named medusavirus stheno T3. Here, we describe their morphological, genomic and gene content similarities and also propose to create a new family "*Mamonoviridae*", a new genus, "*Medusavirus*", and two species, "*Medusavirus medusae*" and "*Medusavirus sthenus*" to classify these two viruses within the phylum *Nucleocytoviricota*.

Introduction

Various amoeba-infecting giant viruses have been isolated during the last 20 years [1–6]. They are characterized by the large size of their genomes and particles. This group of viruses have been classified in the phylum *Nucleocytoviricota* [7]. Nucleo-Cytoplasmic Virus Orthologous Groups (NCVOGs) and Giant Virus Orthologous Groups (GVOGs) have been widely used for core gene identification and to conduct comprehensive classification of these viruses [8, 9]. As a result, the structure of the phylum *Nucleocytoviricota* has been recently challenged and an expansion from seven currently recognized families (*Mimiviridae*, *Phycodnaviridae*, *Ascoviridae*, *Iridoviridae*, *Marseilleviridae*, *Asfarviridae*, and *Poxviridae*) to 32 families has been proposed [8, 10, 11].

Recently, two viruses were discovered by co-culture with *Acanthamoeba castellanii*. The first one was isolated from a hot spring in Japan and was named Acanthamoeba castellanii medusavirus J1 (ACMV-J1), because the host amoebae tend to form cysts upon infection with this virus and this phenomenon is reminiscent of Medusa in Greek mythology [4]. The second isolate was a close relative of the first virus and named medusavirus stheno T3 (MVS-T3), because Stheno is a sister of Medusa [5]. Both viruses show substantial morphological and genomic similarities with the members of the phylum *Nucleocytoviricota*. However, these viruses were not phylogenetically close to any member of established families within *Nucleocytoviricota*. Thus, in order to officially classify these two viruses within the ICTV framework Virus Taxonomy, we propose to create two species, "*Medusavirus medusae*" (typified by ACMV-J1) and "*Medusavirus sthenus*" (with an exemplar isolate MVS-T3), to classify them in the new genus and family, "*Medusavirus*"

and "*Mamonoviridae*" respectively, which belongs to the class *Megaviricetes* of the phylum *Nucleocytoviricota*.

Etymology of taxa nomenclature

Species and genus nomenclature was inspired by the two Gorgon sisters from Greek mythology (Medusa and Stheno), while the family name originates from Japanese word "mamono" (魔物), meaning "monster".

Infection cycle

ACMV-J1 was shown to enter the host cell by endocytosis and then, enter the host nucleus approximately one hour post infection (hpi). The virus gradually transformed the host nucleus into a viral factory without disrupting the nuclear membrane. At around 10 hpi, the cytoplasm was filled with empty viral capsids and eventually, viral particles were released outside the cell in a non-lytic way at around 14 hpi [4, 12] (Fig. 1a).

Genomic and proteomic features

The ACMV-J1 virion shows an icosahedral shape with a diameter of approximately 260 nm, including surface spikes, as revealed by cryo-electron microscopy (cryo-EM) [12] (Fig 1b&1c). It encapsidates a linear, double-stranded DNA (dsDNA) genome of 381,277 bp with a high G+C content (61.7%) [4]. A total of 461 open reading frames (ORFs) have been predicted in the genome. ACMV-J1 genome encodes five of the seven core genes of *Nucleocytoviricota* that are frequently used in phylogeny [8]. These are major capsid protein (MCP), superfamily II helicase (SFII), DNA polymerase family B (PolB), A32-like packaging ATPase (A32), and virus late transcription factor (VLTF3). However, the virus is unique among other amoeba-infecting giant viruses in encoding a full set of histone proteins (i.e., linker histone H1, and core histones H2A, H2B, H3, and H4) and lacking two of the core genes, namely, RNA polymerase and DNA topoisomerase II (TopoII).

MVS-T3 was isolated in 2021 [5]. This virus shows icosahedral particles similar to those of ACMV-J1, and has a G+C-rich (62.64%), 362,811 bp-long dsDNA genome. The average

nucleotide identity (ANI) between ACMV-J1 and MVS-T3 is 79.5%. MVS-T3 has the same set of core genes as ACMV-J1 and also encodes a full set of histones, but the genes for H3 and H4 are fused into a single gene.

Phylogenomics

To clarify the relationship between medusaviruses and other members of the *Nucleocytoviricota*, we used the seven core genes from GVOGs, which have been argued to have the optimum performance for phylogenetic analysis of *Nucleocytoviricota* (i.e., PolB, SFII, A32, VLTF3, TopoII, TFIIB, RNAPL) [8]. The two medusaviruses formed a clade in the phylum *Nucleocytoviricota* with a high branch support (SH-aLRT = 100%, Ultrafast bootstrap = 100%) (Fig. 2), consistent with a previous study that demonstrated that ACMV-J1 does not belong to any virus group identified so far [4]. In the tree, medusaviruses are close to Feldmannia species virus, Ectocarpus siliculosus virus 1, coccolithoviruses, pandoraviruses and molliviruses, which were previously suggested to form a putative and yet non-recognized order, "*Pandoravirales*" [8]. However, the branch support for this clade was weak (SH-aLRT = 97.9% and Ultrafast bootstrap = 58%). Thus, here we focus on position of the two medusaviruses and propose to create two new species in a new genus and a new family.

Relationship between medusaviruses and clandestinovirus

Recently, another giant virus named clandestinovirus was isolated by co-culture with another host, *Vermamoeba vermiformis*, in France [6]. The clandestinovirus shows a larger genome, more genes and a lower G+C content (581 kbp, 617 genes, 43.5%) than medusaviruses. In terms of core genes, clandestinovirus encodes all core genes that medusaviruses have and additionally encodes RNA polymerase and TopoII. Alike medusaviruses, clandestinovirus also induces a nucleo-cytoplasmic infection, and enters and turns the host nucleus into the viral factory. A previous study has shown that the closest relative of clandestinovirus is ACMV-J1 in terms of the core genes [6].

Here, we used a quantitative way to draw a family-level boundary to figure out the relationship between clandestinovirus and medusaviruses. We compared these three viruses

in terms of the nucleotide level similarity, including ANI and tetra-nucleotide similarity (TETRA), phylogenomic distance by calculating the distance between tips on the phylogenomic tree, and number of shared OGs. We then compared these metrics between them to the inter- and intra-family metrics for other virus families. As a result, the relationship between medusaviruses and clandestinovirus lies in the middle of inter-family and intra-family levels.

In terms of phylogenomic tree, the clandestinovirus branched together with medusaviruses with a high branch support (Ultrafast bootstrap = 100%, SH-aLRT = 98.8%) (Fig. 2). However, the tip distances (3.92 to ACMV-J1 and 3.95 MVS-T3) lay between mean values for intra-family (2.46) and inter-family distances (7.30) (Fig. 3a).

In terms of genome-level nucleotide similarity, ANI and TETRA were calculated by python package pyani [13]. The ANI between clandestinovirus and the two medusaviruses were both 0, whereas the average of intra- and inter-family were 0.36 and 0.01, respectively. In addition, only kaumoebavirus had non-zero ANI (0.68) against clandestinovirus. The TETRA between medusaviruses and clandestinovirus were both 0.32, which was lower than the average of inter-family TETRA (0.38). In addition, TETRA between clandestinovirus and medusaviruses only ranked 134th and 139th among 220 comparisons between clandestinovirus and other viruses. (Fig. 3b, 3c).

We then used Orthofinder v.2.5.2 to identify OGs and calculated the gene-sharing level S_{ij} based on the number of shared OGs between viral genomes [14]. The number of shared OGs was normalized by the total number of OGs of each virus under comparison using the following formula:

$$S_{ij} = \frac{OG_{ij}}{\sqrt{OG_i \times OG_j}}$$

Here, OG_{ij} is the number of shared OGs between virus i and j, and OG_i is the total number of OGs in virus i. The gene-sharing level between clandestinovirus and medusaviruses (0.16 to ACMV-J1, ranked 30th among all comparisons between clandestinovirus and other

viruses; 0.17 to MVS-T3, 25th) lay between the mean values for intra- and inter-family levels (0.47 and 0.07, respectively) (Fig. 3d).

Among known viruses, clandestinovirus is the closest relative of medusaviruses. However, they show large divergence that places their phylogenetic relationships in the middle of intra- and inter-family levels. Thus, at this moment we do not include the clandestinovirus into the proposed new family "*Mamonoviridae*".

Finally, we propose the following simple and ready-made criteria for species, genus and family demarcations under the family "*Mamonoviridae*". If a virus shares >95% ANI, similar morphology, and comparable genome size to the members of two proposed species (e.g., "*Medusavirus medusae*" and "*Medusavirus sthenus*") in the genus "*Medusavirus*", it should be classified in one of these two taxa. The average of intra-genus ANI is 70% within five families of the phylum *Nucleocytoviricota* (i.e., *Mimiviridae*, *Ascoviridae*, *Phycodnaviridae*, *Poxviridae* and *Iridoviridae*). By taking this statistic in consideration, we propose that if a virus shares >70% ANI, similar morphology, and comparable compositions of core genes to the members of the proposed genus "*Medusavirus*", it should be classified in this genus. For a virus distantly related to the members of this proposed family "*Mamonoviridae*", its inclusion in or exclusion from the family should be considered based on phylogenomic analyses like those we presented in this study. We acknowledge that these criteria are subject to updated according to the progress of analytical methods and discoveries of new traits in viruses.

Conclusion

Medusaviruses are amoeba-infecting giant viruses that carry out a nucleo-cytoplasmic infection cycle and are unique among known viruses by encoding a full set of histone genes. Currently, there are two well-characterized but not yet officially classified medusaviruses (ACMV-J1 and MVS-T3). Our phylogenomic analysis revealed that this group of viruses does not branch within any groups of viruses. Thus, based on overall characteristics of the two currently known medusaviruses, in particular genome features and phylogenomics, here we propose creation of two species, "*Medusavirus medusae*" and "*Medusavirus sthenus*" in

a new genus, "*Medusavirus*" and a new family "*Mamonoviridae*" to classify ACMV-J1 and MVS-T3, respectively. We propose that the new family is included in the class *Megaviricetes* of the phylum *Nucleocytoviricota*. This article is related to a taxonomic proposal, recently officially submitted to the ICTV for consideration, but not yet approved/ratified at the time of publication. Therefore, taxa proposed in this paper are not part of the official ICTV taxonomy.

Acknowledgements

Computational time was provided by the SuperComputer System, Institute for Chemical Research, Kyoto University.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI [grant numbers 18H02279, 20H03078, 22H00384]; and the Kyoto University Foundation, and the International Collaborative Research Program of Institute for Chemical Research, Kyoto University [grant number 2020-31].

References

1. la Scola B, Audic S, Robert C, et al (2003) A Giant Virus in Amoebae. Science (New York, NY) 299:2033. https://doi.org/10.1126/science.1081867

2. Boyer M, Yutin N, Pagnier I, et al (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proceedings of the National Academy of Sciences 106:21848–21853. https://doi.org/10.1073/pnas.0911354106

3. Philippe N, Legendre M, Doutre G, et al (2013) Pandoraviruses: Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes. Science 341:281–286. https://doi.org/10.1126/science.1239181

4. Yoshikawa G, Blanc-Mathieu R, Song C, et al (2019) Medusavirus, a Novel Large DNA Virus Discovered from Hot Spring Water. Journal of Virology 93:e02130-18. https://doi.org/10.1128/JVI.02130-18

5. Yoshida K, Zhang R, Garcia KG, et al (2021) Draft Genome Sequence of Medusavirus Stheno, Isolated from the Tatakai River of Uji, Japan. Microbiol Resour Announc 10:. https://doi.org/10.1128/MRA.01323-20

6. Rolland C, Andreani J, Sahmi-Bounsiar D, et al (2021) Clandestinovirus: A Giant Virus With Chromatin Proteins and a Potential to Manipulate the Cell Cycle of Its Host Vermamoeba vermiformis. Front Microbiol 12:715608. https://doi.org/10.3389/fmicb.2021.715608

7. Koonin EV, Dolja VV, Krupovic M, et al (2020) Global Organization and Proposed Megataxonomy of the Virus World. Microbiol Mol Biol Rev 84:e00061-19. https://doi.org/10.1128/MMBR.00061-19

8. Aylward FO, Moniruzzaman M, Ha AD, Koonin EV (2021) A phylogenomic framework for charting the diversity and evolution of giant viruses. PLOS Biology 19:e3001430. https://doi.org/10.1371/journal.pbio.3001430

9. 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. https://doi.org/10.1186/1743-422X-6-223

10. Walker PJ, Siddell SG, Lefkowitz EJ, et al (2021) Changes to virus taxonomy and to the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2021). Arch Virol 166:2633–2648. https://doi.org/10.1007/s00705-021-05156-1

11. Aylward FO, Abrahão J, Brussaard C, Fischer MG, Moniruzzaman M, Ogata H, Suttle CA (2022) Create 3 new families, 3 subfamilies, 13 genera, and 20 new species within the order Imitervirales (phylum Nucleocytoviricota) and rename two existing species. https://talk.ictvonline.org/files/proposals/taxonomy_proposals_fungal1/m/fung01/13591

12. Watanabe R, Song C, Kayama Y, et al Particle Morphology of Medusavirus Inside and Outside the Cells Reveals a New Maturation Process of Giant Viruses. Journal of Virology 0:e01853-21. https://doi.org/10.1128/jvi.01853-21

13. Pritchard L, Glover RH, Humphris S, et al (2015) Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods 8:12–24. https://doi.org/10.1039/C5AY02550H

14. Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology 20:238. https://doi.org/10.1186/s13059-019-1832-y

15. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973. https://doi.org/10.1093/bioinformatics/btp348

16. Katoh K (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Research 33:511–518. https://doi.org/10.1093/nar/gki198

17. Minh BQ, Schmidt HA, Chernomor O, et al (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37:1530–1534. https://doi.org/10.1093/molbev/msaa015

18. Kalyaanamoorthy S, Minh BQ, Wong TKF, et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285

19. Hoang DT, Chernomor O, von Haeseler A, et al (2018) UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution 35:518–522. https://doi.org/10.1093/molbev/msx281

Figures

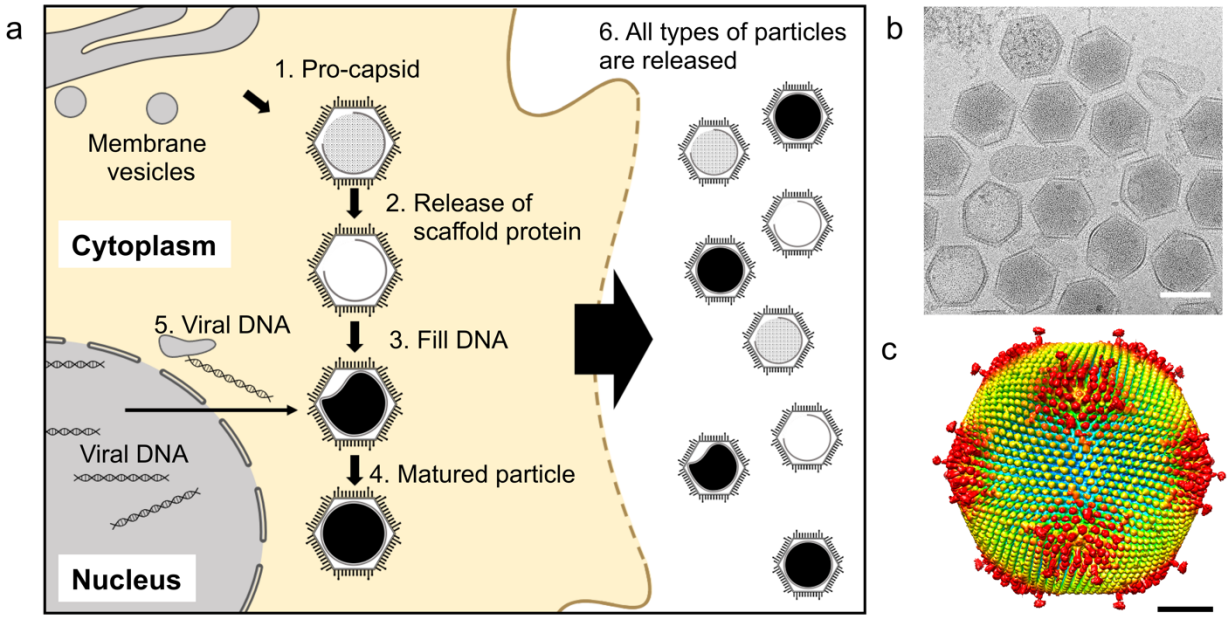
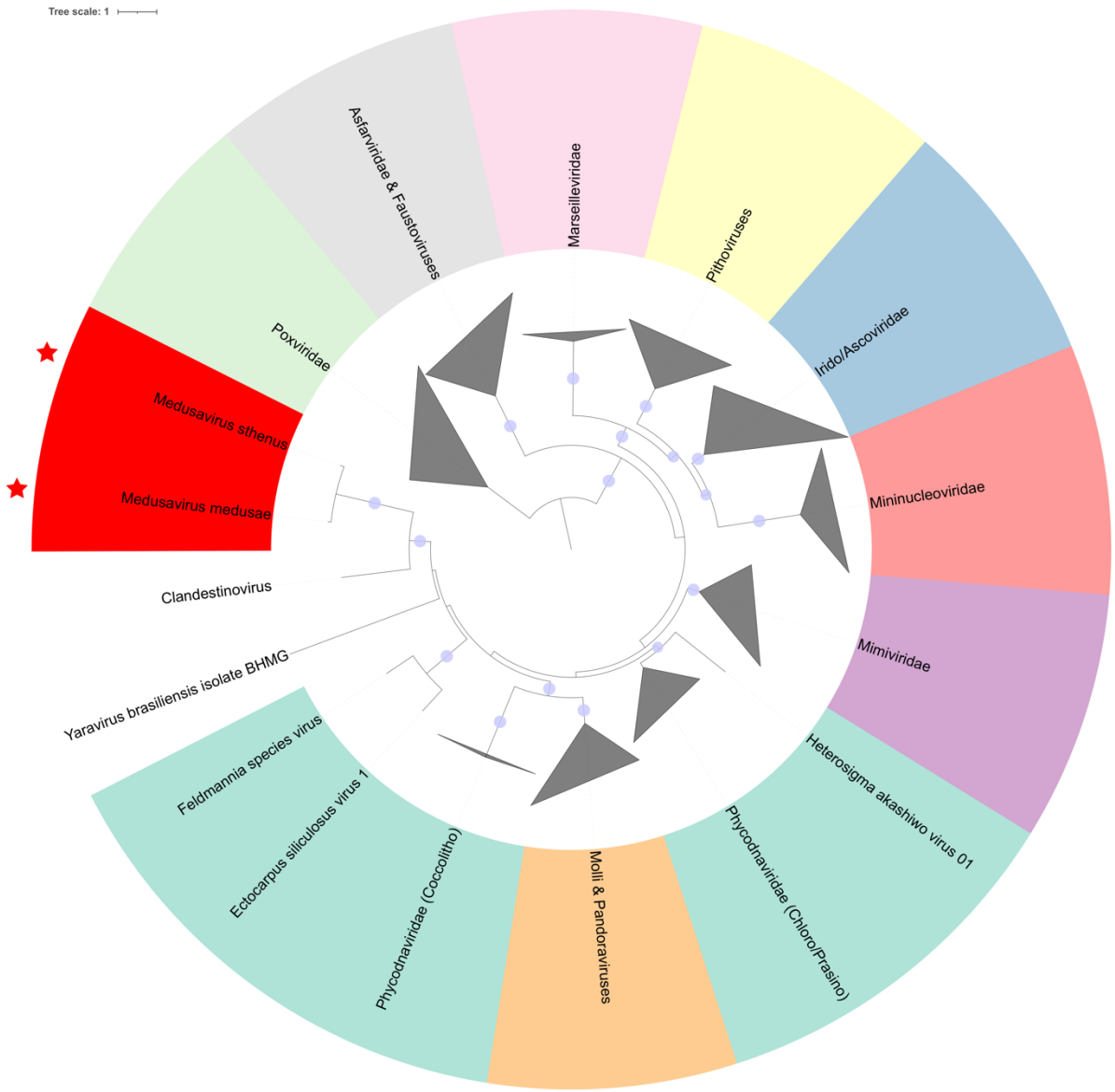


Fig. 1 Acanthamoeba castellanii medusavirus J1 (ACMV-J1) replication and its particle feature [12]. (a) ACMV-J1 replication in amoeba cell after infection. (b) A cryo-EM image of ACMV-J1. Scale 200 nm. (c) A 3D reconstruction of ACMV-J1 virion. Scale 50 nm.



230

231 Fig 2 Maximum-likelihood phylogenetic tree of *Nucleocytoviricota*. The tree was based on a
232 concatenated amino acid sequence alignment of seven marker genes constructed using MAFFT
233 (v.7.471) and trimAl (v.1.4.1) and was built using IQ-TREE 2 (v.2.1.3) [15–17]. The model was
234 LG+F+R8 selected by the built-in Modelfinder of IQ-TREE 2 [18]. The branch supports were
235 computed by 1000 ultrafast bootstrap and SH-aLRT [19]. The tree was visualized by iTOL,
236 the round labels on branches represent high confidence supports with Ultrafast bootstrap ≥

95%, SH-aLRT ≥ 80%. Position of proposed family "*Mamonoviridae*" is reported in red background and marked with stars.

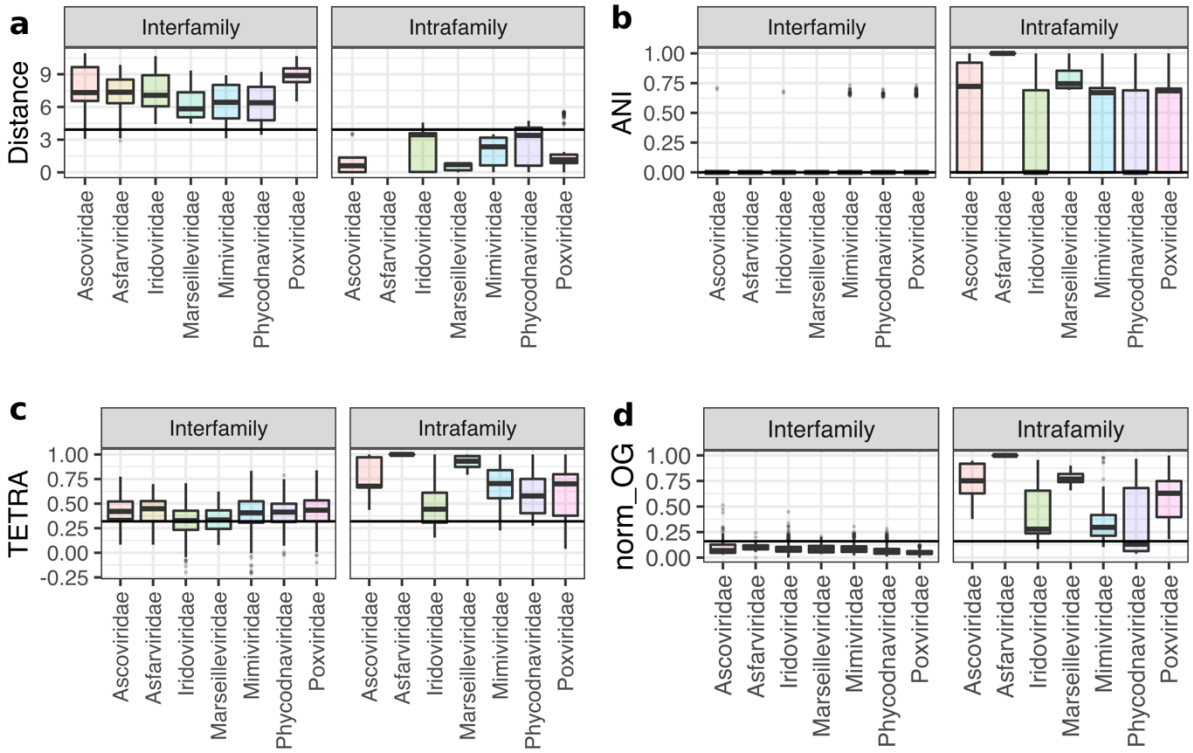


Fig. 3 Boxplots for (a) tip distance, (b) ANI, (c) TETRA, and (d) normalized OGs sharing level. The horizontal black line represents the value between clandestinovirus and Acanthamoeba castellanii medusavirus J1 (ACMV-J1), a member of proposed species "*Medusavirus medusae*".