

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2024

**Part 1a: Details of taxonomy proposals**

|  |  |  |
| --- | --- | --- |
| **Title:** | Create a phylum within kingdom ‘*Trapavirae’* (realm *Monodnaviria*) for classification of hyperthermophilic archaeal viruses with pleomorphic virions | |
| **Code assigned:** | 2024.004A.N.v1.Thalassapleoviridae\_newphylum |

|  |  |  |  |
| --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | |
| **Name** | **Affiliation** | **Email address** | **Corresponding author(s)** X |
| Baquero DP | Institut Pasteur, Université de Paris, Archaeal Virology Unit, Paris, France | dp.baquero645@gmail.com |  |
| Bignon EA | Institut Pasteur, Université de Paris, Archaeal Virology Unit, Paris, France | eduardo.bignon@pasteur.fr |  |
| Krupovic M | Institut Pasteur, Université de Paris, Archaeal Virology Unit, Paris, France | mart.krupovic@pasteur.fr | X |

**Part 1b: Taxonomy Proposal Submission**

|  |  |  |  |
| --- | --- | --- | --- |
| **ICTV Subcommittee:** | | | |
| Animal DNA Viruses and Retroviruses |  | Bacterial viruses |  |
| Animal minus-strand and dsRNA viruses |  | Fungal and protist viruses |  |
| Animal positive-strand RNA viruses |  | Plant viruses |  |
| Archaeal viruses | **X** | General - |  |

|  |
| --- |
| **List the ICTV Study Group(s) that have seen or have been involved in creating this proposal:** |
|  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Optional – complete only if formally voted on by an ICTV Study Group:** | | | |
| **Study Group** | **Number of members** | | |
| **Votes in support** | **Votes against** | **No vote** |
|  |  |  |  |
|  |  |  |  |

|  |  |
| --- | --- |
| **Submission date:** | DD/MM/YYYY |

**Part 1c: Feedback from ICTV Executive Committee (EC) meeting**

|  |  |
| --- | --- |
| **Executive Committee Meeting Decision code:** | **X** |
| A – Accept |  |
| Ac – Accept subject to revision by relevant subcommittee chair. No further vote required |  |
| U – Accept without revision but with re-evaluation and email vote by the EC |  |
| Uc – Accept subject to revision and re-evaluation and email vote by the EC |  |
| Ud – Deferred to the next EC meeting, with an invitation to revise based on EC comments |  |
| J - Reject |  |
| W - Withdrawn |  |

|  |
| --- |
| **Comments from the Executive Committee:** |
|  |

**Part 1d: Revised Taxonomy Proposal Submission**

|  |
| --- |
| **Response of proposer:** |
|  |

|  |  |
| --- | --- |
| **Revision date:** | DD/MM/YYYY |

**Part 3:** **TAXONOMIC PROPOSAL**

|  |
| --- |
| **Name of accompanying Excel module:** |
| 2024.004A.N.v1.Thalassapleoviridae\_newphylum.xlsx |

|  |  |  |  |
| --- | --- | --- | --- |
| **Taxonomic changes proposed:** | | | |
| Establish new taxon | **X** | Split taxon |  |
| Abolish taxon |  | Merge taxon |  |
| Move taxon |  | Promote taxon |  |
| Rename taxon |  | Demote taxon |  |
| Move and rename |  |

|  |  |  |
| --- | --- | --- |
| **Is any taxon name used here derived from that of a living person:** | | **N** |
| **Taxon name** | **Person from whom the name is derived** | **Attached X** |
|  |  |  |
|  |  |  |
|  |  |  |

|  |
| --- |
| **Abstract of Taxonomy Proposal:** |
| *Taxonomic rank(s) affected*: *Monodnaviria*, *Trapavirae*  *Description of current taxonomy*: Monodnavirian kingdom *Trapavirae* currently comprises a single family, *Pleolipoviridae*, which includes haloarchaeal viruses with enveloped pleomorphic virions and single-stranded (ss) or double-stranded (ds)DNA genomes.  *Proposed* *taxonomic change(s):* Here we propose to classify viruses infecting hyperthermophilic marine archaea, distantly related to pleolipovirids, into a new family, “Thalassapleoviridae”, and include it into a new phylum within the kingdom *Trapavirae*.  *Justification*: Whole-genome phylogenomic analysis and maximum likelihood phylogenetic analysis based on the membrane fusion protein characteristic of members of the kingdom *Trapavirae* show that members of the proposed family “Thalassapleoviridae” form a monophyletic group separate from the haloarchaeal pleolipovirids and currently unclassified related viruses of methanogenic archaea. |

|  |
| --- |
| * **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*: *Monodnaviria*, *Trapavirae*  *Description of current taxonomy*: Monodnavirian kingdom *Trapavirae* currently comprises a single family, *Pleolipoviridae*, which includes viruses with enveloped pleomorphic virions and single-stranded (ss) or double-stranded (ds)DNA genomes. All members of this family infect hyperhalophilic archaea of the order Halobacteriales (Liu et al., 2022). Pleolipoviruses encode two major structural proteins, a small hydrophobic protein with 3 transmembrane domains, exemplified by VP3 of HRPV1, and a membrane fusion protein, exemplified by VP4 of HRPV1 (Pietilä et al., 2009). The latter protein is unrelated to those described in other classified viruses and has a unique V-shaped structural fold (El Omari et al., 2019). Recently, three additional groups of archaeal viruses encoding homologs of pleolipoviral VP3 and VP4 were identified. These viruses infect widely different archaeal hosts, namely, methanogens of the Methanomassiliicoccales and Methanonatronarchaeia (Medvedeva et al., 2024), and hyperthermophiles of the class Archaeoglobi (Baquero et al., 2024). Here, we propose classification of pleolipovirus relatives infecting Archaeoglobi hosts.  *Proposed* *taxonomic change(s)*: Analysis of the *Archaeoglobus veneficus* genome (NC\_015320) revealed the presence of a putative provirus of 17.9 kbp flanked by direct repeats of 27 nucleotides, corresponding to putative attachment (att) sites, which was named as Archaeoglobus veneficus pleomorphic virus 1 (AvPV1; Baquero et al., 2024). One of the att sites overlapped a tRNA-Gly gene, whereas the other one was adjacent to the integrase gene. It was subsequently confirmed by PCR analysis that the virus is active and is released into the cell-free culture supernatants of *A. veneficus*.Pleomorphic virus-like particles of ~53 nm in diameter, resembling those produced by haloarchaeal pleolipoviruses, were also observed in the cell-free supernatants (Fig. 1A).  The genome of AvPV1 contains 31 open reading frames (ORFs) (Fig. 1B). Sensitive profile–profile comparisons with HHpred showed that AvPV1 carries a block of genes characteristic of pleolipoviruses. In particular, ORF4 and ORF6 encode homologs of the two major structural pleolipovirus proteins, VP3-like integral membrane protein and VP4-like fusogen, respectively. Structural modeling of the AvPV1 VP4-like protein confirmed that it has the same V-shaped fold as the fusogens of bona fide pleolipoviruses (Fig. 2). In addition, ORF8 and ORF10 encode homologs of the conserved HRPV1 protein ORF6 of unknown function and the putative NTPase, respectively (Fig. 1B). Similar to gammapleolipovirus His2, AvPV1 appears to encode a second divergent copy (ORF5) of a putative spike/fusion protein (HHpred probability of 87.9%). Overall, these results indicated that AvPV1 is distantly related to pleolipoviruses of halophilic archaea and is the first representative of this virus lineage associated with marine archaea (Baquero et al., 2024).  Searches for AvPV1 VP4-like homologs in the Whole Genome Shotgun and non-redundant protein sequence databases at the National Center for Biotechnology Information (NCBI) and the Integrated Microbial Genomes/Virus database using TBLASTN or BLASTP (E-value≤1e-5) yielded 19 contigs originating from geographically remote hydrothermal vents (Baquero et al., 2024). Eight of the contigs corresponded to complete or nearly complete viral genomes (Fig. 3). Three of the complete viral genomes were detected as proviruses integrated in the genomes of *Archaeoglobus profundus* DSM 5631, *Geoglobus acetivorans* SBH6, and *Geoglobus ahangari* 234, suggesting that hyperthermophilic pleomorphic viruses are primarily associated with members of the class Archaeoglobi. The identified viruses exhibit a similar arrangement of the core structural genes to that of AvPV1, but all of them lack the second copy of the spike protein gene (Fig. 3). Similar to AvPV1, the VP4-like fusogen of the identified viruses also exhibits a V-shaped fold (Fig. 2). Here, we propose to classify AvPV1-like viruses into a new virus family, “Thalassapleoviridae”, after Thalassa, the primordial Greek goddess of the sea.  *Demarcation criteria:* we propose using 95% sequence identity as a species demarcation criterion, to be consistent with the classification of other bacterial and archaeal DNA viruses, including pleolipoviruses (Liu et al., 2022). Viruses will be assigned to different genera based on differences in the gene contents and if they form monophyletic groups in phylogenomic analyses of the whole genome proteomes.  *Justification*:  *Justification for species*: Five complete genomes of AvPV1-like viruses are available in GenBank (Table 1). Pairwise comparison of the genome sequences showed that all of them share less than 95% sequence identity across the genome length and thus represent distinct species.  *Justification for genera*: There are notable differences between viruses associated with *Archaeoglobus* and *Geoglobus* hosts. In particular, the two groups of viruses encode at least two non-orthologous groups of integrases of the tyrosine recombinase superfamily. Furthermore, viruses of *Geoglobus* (GacPV1, GahPV1 and ThalV2) encode a putative rolling-circle replication endonucleases (RCRE) of the HUH superfamily, likely responsible for the genome replication initiation. By contrast, obvious candidates for genome replication proteins could not be identified in AvPV1 and ApPV1. A characteristic feature of AvPV1 is that it encodes a second copy of a putative spike/fusion protein, which distinguishes it from ApPV1. Indeed, AvPV1 and ApPV1 do not share genes outside of the morphogenesis and integration modules. Furthermore, AvPV1 consistently formed an outgroup to the rest of the viruses in both phylogenomic analysis and maximum likelihood phylogenetic analysis of the VP4-like fusogens (Figs. 4 and 5). Thus, we suggest creating 3 genera to accommodate the 5 AvPV1-like viruses (Table 1).  *Justification for family*: In whole-proteome-based phylogenomic analysis (Fig. 4) and maximum likelihood phylogenetic analysis of the VP4-like fusogens (Fig. 5), AvPV1-like viruses formed a monophyletic group outside of the family *Pleolipoviridae* and separate from the two groups of related viruses associated with methanogenic archaea. Accordingly, a separate family should be created for the classification of AvPV1-like viruses.  *Justification for higher level taxonomy*: In the absence of an objective outgroup, we rooted the phylogenetic tree with VP4 from Methanomassiliicoccales, a group of archaea assigned to a different phylum (Thermoplasmatota) than the other three archaeal lineages, all in Halobacteriota. Archaeoglobi VP4 homologs were at the base of the clade including Halobacteria, Archaeoglobi, and Methanonatronarchaeia, with viruses of haloarchaea (family *Pleolipoviridae*) and unclassified viruses of Methanonatronarchaeia forming monophyletic sister groups (Fig. 5). A congruent grouping of *Pleolipoviridae* withviruses of Methanonatronarchaeia was obtained also in whole-proteome-based phylogenomic analysis (Fig. 4). Given the vast diversity of haloarchaeal pleolipoviruses, most of which are unclassified (see Fig. 5), it is expected that the family will be elevated to an order level in the near future, and hence haloarchaeal and methanonatronarchaeal viruses would be grouped into a taxon at the class level. Thus, we propose including the proposed family of AvPV1-like viruses, “Thalassapleoviridae”, into the kingdom *Trapavirae* at the phylum level, which we propose naming “Calorviricota”. The intermediate taxa, class and order, would be named “Caminiviricetes” and “Ageovirales”, respectively.  *Etymology*:   * *Calor*, heat in Latin, referring to the environmental conditions under which these viruses replicate. * *Camini*, chimney in Latin, referring to the chimney-like deep sea hydrothermal vents in which Archaeoglobales and their viruses thrive. * *Ageo*, truncation of *Archaeoglobus* and *Geoglobus*, two genera of archaea that host these viruses.   Species epithets:   * *atlanticense*, after Atlantic Ocean, source of origin. * *guaymasense*, after Guaymas Basin, source of origin. * *pacificense*, after Pacific Ocean, source of origin. |

|  |
| --- |
| **References:** |
| Baquero DP, Bignon EA, Krupovic M. Pleomorphic viruses establish stable relationship with marine hyperthermophilic archaea. ISME J. 2024; 18(1):wrae008. doi: 10.1093/ismejo/wrae008. PMID: 38366050  El Omari K, Li S, Kotecha A, Walter TS, Bignon EA, Harlos K, Somerharju P, De Haas F, Clare DK, Molin M, Hurtado F, Li M, Grimes JM, Bamford DH, Tischler ND, Huiskonen JT, Stuart DI, Roine E. The structure of a prokaryotic viral envelope protein expands the landscape of membrane fusion proteins. Nat Commun. 2019; 10(1):846. doi: 10.1038/s41467-019-08728-7. PMID: 30783086  Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017; 14(6):587-589. doi: 10.1038/nmeth.4285. PMID: 28481363  Liu Y, Dyall-Smith M, Oksanen HM. ICTV Virus Taxonomy Profile: Pleolipoviridae 2022. J Gen Virol. 2022; 103(11):001793. doi: 10.1099/jgv.0.001793. PMID: 36374180  Medvedeva S, Borrel G, Krupovic M, Gribaldo S. A compendium of viruses from methanogenic archaea reveals their diversity and adaptations to the gut environment. Nat Microbiol. 2023; 8(11):2170-2182. doi: 10.1038/s41564-023-01485-w. PMID: 37749252  Meier-Kolthoff JP, Göker M. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. Bioinformatics. 2017; 33(21):3396-3404. doi: 10.1093/bioinformatics/btx440. PMID: 29036289  Pietilä MK, Roine E, Paulin L, Kalkkinen N, Bamford DH. An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Mol Microbiol. 2009 Apr;72(2):307-19. doi: 10.1111/j.1365-2958.2009.06642.x. PMID: 19298373  Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016; 44:W232-5. doi: 10.1093/nar/gkw256. PMID: 27084950 |

|  |
| --- |
| **Tables, Figures:** |

<Start here>

Table 1. Properties of thalassapleovirids.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Genus** | **Species** | | **Virus name** | | **Virus name abbreviation** | | **Genome length (bp)** | | **Accession number** | |
| “Avenivirus” | | “Avenivirus atlanticense” | | Archaeoglobus veneficus pleomorphic virus 1 | | AvPV1 | | 17950 | | BK065155 | |
| “Aprofuvirus” | | “Aprofuvirus guaymasense” | | Archaeoglobus profundus pleomorphic virus 1 | | ApPV1 | | 14699 | | BK065154 | |
| “Geogavirus” | | “Geogavirus atlanticense” | | Geoglobus acetivorans pleomorphic virus 1 | | GacPV1 | | 13362 | | BK065156 | |
| “Geogavirus” | | “Geogavirus guaymasense” | | Geoglobus ahangari pleomorphic virus 1 | | GahPV1 | | 16765 | | BK065157 | |
| “Geogavirus” | | “Geogavirus pacificense” | | Thalassapleovirus 2 | | ThalV2 | | 16206 | | BK065158 | |

A close-up of a round object

Description automatically generated

**Figure 1.** Archaeoglobus veneficus pleomorphic virus 1 (AvPV1). A, Electron micrographs of pleomorphic VLPs negatively stained with 2% uranyl acetate. (Scale bar, 100 nm). B, Genome map of AvPV1. The open reading frames (ORFs) are represented with arrows indicating the direction of transcription.

A group of colorful spirals

Description automatically generated

**Figure 2.** Structural models of VP4-like fusogens. Comparison of the published crystal structure of the fusion protein of the haloarchaeal pleolipovirus HRPV6 (PDB: 6QGL; El Omari et al., 2019) to the AlphaFold2 structural model for the VP4-like fusion protein of AvPV1, ApPV1, GacPV1 and GahPV1. Terminal ends containing the transmembrane domains were trimmed for the convenience of presentation. Protein structures are colored using the rainbow scheme from blue (N-terminus) to red (C-terminus).

A close-up of a graph

Description automatically generated

**Figure 3.** Alignment of complete or nearly-complete genomes of pleomorphic viruses originating from marine deep-sea hydrothermal vents. ORFs are depicted by arrows that indicate the direction of transcription. Functional annotations are depicted above the corresponding ORFs. Homologous genes are shown using the same colors and are connected by shading in grayscale, with intensity reflecting the amino acid sequence identity. HJR, Holliday junction resolvase; NT, nucleotidyltransferase; PH, pleckstrin homology domain; PIN, PIN domain protein; RHH, ribbon-helix-helix protein; RCRE, rolling-circle replication endonuclease; TGM, transglutaminase; wHTH, winged helix-turn-helix domain; Zn finger, zinc finger domain-containing protein.

A screenshot of a computer

Description automatically generated

**Figure 4.** Inferred phylogenomic tree of pleomorphical viruses associated with hyperthermophilic (Archaeoglobi), halophilic (Halobacteria) and methanogenic (Methanonatronarchaeia and Methanomassiliicoccales) archaea based on whole genome VICTOR analysis at the amino acid level (Meier-Kolthoff and Göker, 2017). The tree is midpoint rooted and the branch length is scaled in terms of the Genome BLAST Distance Phylogeny (GBDP) distance formula D6. The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications.

A diagram of a virus

Description automatically generated

**Figure 5.** Maximum-likelihood analysis of VP4-like fusogens. Clades of VP4 homologs encoded by viruses associated with different archaeal lineages are indicated with different colors. Pleolipovirus isolates are labeled. The scale bar represents the number of substitutions per site. Circles at nodes indicate bootstrap support >90%. The maximum likelihood phylogeny was inferred using IQ-TREE v1.6.12 (Trifinopoulos et al., 2016). The substitution model best fitting the data was identified with ModelFinder (Kalyaanamoorthy et al., 2017) and was Blosum62+F+R8. The branch support was assessed using SH-like approximate likelihood ration test (aLRT) with 1000 replicates.