

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2025

<https://ictv.global/taxonomy/templates>**Part 1a: Details of taxonomy proposals**

|  |  |
| --- | --- |
| **Title:** | Reclassifying the order *Crassvirales* to establish two sister orders and one sister family, with the creation of eight new families, 159 new genera, and 603 new species |
| **Code assigned:** | 2025.013B.Crassvirales\_reorganisation | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | | |
| **Given name (+middle initial(s))** | **Surname** | **Affiliation** | **Email address** | **Corr. author(s)** |
| Fabian TS | Bastiaanssen | University College Cork, Ireland | fbastiaanssen@ucc.ie |  |
| Rémi | Denise | University College Cork, Ireland | rdenise@ucc.ie |  |
| George S | Bouras | The University of Adelaide, Australia | george.bouras@adelaide.edu.au |  |
| Steven R | Stockdale | Novogene, Ireland | stephen.stockdale@novogene-europe.com |  |
| Robert A | Edwards | Flinders University, Australia | raedwards@gmail.com |  |
| Colin | Hil | University College Cork, Ireland | c.hill@ucc.ie |  |
| Andrey N | Shkoporov | University College Cork, Ireland | andrey.shkoporov@ucc.ie | X |
|  |  |  |  |  |

**Part 1b: Taxonomy Proposal Submission**

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

|  |
| --- |
| **List the ICTV Study Group(s) that have seen or have been involved in creating this proposal:** <https://ictv.global/sc> |
| Crassvirales Study Group |

|  |  |  |  |
| --- | --- | --- | --- |
| **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:** |  |

**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:** |  |

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

<https://ictv.global/taxonomy/templates>

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

|  |  |
| --- | --- |
| **Etymology (origin) of proposed taxonomic names:** | |
| Taxon name | Etymology of the term |
| *Paracrassvirales* | Meaning "near" or "beside" *Crassvirales*, this clade is highly similar to *Crassvirales*. |
| *Metacrassvirales* | Meaning “beyond” or “other” *Crassvirales*, this clade is highly similar to *Crassvirales*. |
| *Tinaiviridae* | From the Central Dusun “tinai”, meaning intestines |
| *Darmviridae* | From the Dutch “darm”, meaning intestines |
| *Jelitioviridae* | From the Polish “jelito”, meaning intestines |
| *Tripviridae* | From the French “tripes” meaning intestines |
| *Meiviridae* | From the Hebrew “מְעִי”, (me’i) meaning intestines |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| **Permission for use of names derived from a living person:** | | |
| **Taxon name** | **Full name of person from whom the name is derived** | **Attached** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

|  |
| --- |
| **Abstract of Taxonomy Proposal:** |
| *Taxonomic rank(s) affected*:  *Realm: Duplodnaviria; Kingdom: Heunggongvirae; Phylum: Uroviricota; Class: Caudoviricetes;* Order: *Crassvirales*  *Description of current taxonomy*:  *Crassvirales* was established under taxonomic proposal 2021.022B  *Proposed* *taxonomic change(s):*  We propose:   1. The establishment of demarcation criteria for the order *Crassvirales* 2. The establishment of the order “*Paracrassvirales”* containing one novel family 3. The establishment of the order “*Metacrassvirales*” containing three novel families 4. The adjustment of demarcation criteria for families within *Crassvirales* 5. The creation of one new families within *Crassvirales* 6. The moving and renaming of one genus within *Crassvirales* 7. The adjustment of demarcation criteria for subfamilies within *Crassvirales* 8. The abolition of 11 subfamilies 9. The adjustment of demarcation criteria for genera within *Crassvirales* to reflect ICTV guidelines 10. The creation of 80 new genera 11. The adjustment of demarcation criteria for species within *Crassvirales* to reflect ICTV guidelines 12. The creation of 601 new species   *Justification*:  Previously suspected members of Crassvirales, Epsiloncrassviridae and Zetacrassviridae are currently unable to be either placed within or outside of the Crassvirales order, due to a lack of demarcation criteria within Crassvirales. Phylogenetic trees utilizing structural and maximum likelihood approaches based on marker genes reveal the formation of unique clades that align with the proposed orders and support the further proposed changes to family, genus and species demarcation criteria. |
|  |

|  |
| --- |
| **Text of Taxonomy proposal:** |

|  |
| --- |
| *Taxonomic rank(s) affected:*  *Realm: Duplodnaviria; Kingdom: Heunggongvirae; Phylum: Uroviricota; Class: Caudoviricetes;* Order: Crassvirales  *Description of current taxonomy*:  Crassvirales was established under taxonomic proposal 2021.022B  *Proposed* *taxonomic change(s)*:  We propose:   1. The establishment of demarcation criteria for the order *Crassvirales* 2. The establishment of the order “*Paracrassvirales”* containing one novel family 3. The establishment of the order “*Metacrassvirales*” containing three novel families 4. The adjustment of demarcation criteria for families within Crassvirales 5. The creation of one new family within Crassvirales 6. The adjustment of demarcation criteria for subfamilies within Crassvirales 7. The abolition of 11 subfamilies 8. The adjustment of demarcation criteria for genera within Crassvirales to reflect ICTV guidelines 9. The creation of 80 new genera 10. The adjustment of demarcation criteria for species within Crassvirales to reflect ICTV guidelines 11. The creation of 601 new species   *Demarcation criteria:*  **Genus and species**:  The demarcation of genus and species was performed by calculating total average nucleotide identity with Vclust1, adhering to the ICTV2 suggested thresholds, species are categorized as genomes showing ≤95% similarity, while genera are characterised by ≥70% similarity.  **Subfamilies:**  No meaningful demarcation criteria for subfamilies were found; we propose the removal of the current criteria.  **Families***:*  The demarcation of Families is based on deep-branching monophyletic clades in the gene and consensus species trees generated using both structural phylogenetic and maximum-likelihood phylogenetic trees, all derived from four well-conserved core genes.  **Orders:**  The demarcation of Orders is based on deep-branching monophyletic clades in the gene and consensus species trees generated using both structural phylogenetic and maximum-likelihood phylogenetic trees, all derived from four well-conserved core genes. As well as the presence or absence of two key core marker genes.  ***Justification***:  Currently, *Crassvirales* lacks demarcation criteria. The existing criteria within *Crassvirales* are ambiguous, unintuitive and based on outdated methods and results. Phylogenetic trees utilising structural and maximum likelihood approaches based on marker genes reveal the formation of unique clades that align with the proposed orders and support the proposed demarcation criteria. The criteria proposed for family demarcation establish monophyletic clades through proteomic and genomic analysis, enhancing clarity and reproducibility. The proposed demarcation criteria for genera and species align with current ICTV guidelines and are supported by phylogenetic analysis.  **Orders**:  We propose the creation of two new orders, *Paracrassvirales* and *Metacrassvirales.* Structural phylogenetic analysis of both single protein and core proteomics using FoldTree3 reveals the creation of three separate clades (Figure 1-6), further supported by maximum likelihood phylogenies of the same proteins (Figure 7). Furthermore, we propose the use of two marker genes discovered through PhaMMseqs v1.0.4 4 to delineate the sister orders. One of these genes (NC\_049977 Gp42) is found in over 99% of all Crassvirales genomes, while the other (NC\_049977 Gp80) is present in more than 99% of genomes from both *Crassvirales* and *Metacrassvirales*.  *Metacrassvirales* and *Paracrassvirales* were chosen as the names for these sister-clades to represent their similarity to *Crassvirales*.  **Families:**  We propose the creation of seven new families, “*Tinaiviridae”* belonging to *Paracrassvirales*, *Meiviridae*, *Tripiviridae* and *Jelitoviridae* belonging to *Metacrassvirales* and *Darmviridae* belonging to *Crassvirales*. The proposed families form monophyletic clades in structural phylogenetic analysis (Figure 1-6) of both single-protein and core proteomics, as well as maximum likelihood phylogenies (Figure 7). And create separate clusters based on shared genes clusters (Figure 9) and according to tANI (Figures 8). While Darmviridae could be placed as a subfamily within Suoliviridae, we have chosen to establish it as a distinct family. This decision aligns with our broader move to eliminate all previously defined subfamilies, whose demarcation criteria were overly broad and inconsistently applied. Recognizing Darmviridae at the family level helps preserve meaningful hierarchical resolution and avoids compressing the distinction between family and subfamily to the point of diminishing its taxonomic value.  **Subfamilies:**  We propose eliminating the existing 11 subfamilies and the demarcation criteria used to define them. We believe the current criteria are too broad, as "27-79% of proteins shared within each subfamily in all possible pairwise comparisons” does not appear to identify any meaningful groupings. Although we couldn't find any criteria that produce more biologically relevant clades, we aim to re-evaluate this situation once more members of each family are identified, and potential new methods are developed.  **Genus and species:**  We propose the creation of 159 new genera and 603 new species. A table with the total, global and average nucleotide identity, including coverage between all genomes, is available in the supplementary data.  Genera were named using the previously established method for algorithmically creating names (2021.022B). Each letter of the term to be mutated was assigned a random number, 1-3. Characters were then passed through a three step script to create the following changes:   1. '*a1'→'ah', 'b1'→'p', 'c1'→'k', 'd1'→'t', 'e1'→'eh', 'g1'→'j', 'i1'→'ih', 'k1'→'c', 'o1'→'oh', 't1'→'d', 'u1'→'uh'* 2. *'a2'→'e', 'e2'→'i', 'i2'→'o', 'o2'→'u', 'u2'→'a'* 3. *'a3'→'i', 'e3'→'o', 'i3'→'u', 'o3'→'a', 'u3'→'e', 'j3'→'g', 'k3'→'c', 'p3'→'b', 't3'→'d'*   In order to maximize the likelihood that the mutated term was pronounceable, only the first occurrence of a repeating letter was kept. All long terms were cut after the 7th character and shortened further if needed to the last occurring vowel. This was to prevent a hard consonant before the genus level suffix '*-virus*'. Each mutated word needed to be a minimum of 5 characters in length, have two consonants and two vowels. |
|  |

|  |
| --- |
| **References:** |
| 1. Zielezinski, A. *et al.* Ultrafast and accurate sequence alignment and clustering of viral genomes. 2024.06.27.601020 Preprint at https://doi.org/10.1101/2024.06.27.601020 (2024).  2. Turner, D., Kropinski, A. M. & Adriaenssens, E. M. A Roadmap for Genome-Based Phage Taxonomy. *Viruses* **13**, 506 (2021).  3. Moi, D. *et al.* Structural phylogenetics unravels the evolutionary diversification of communication systems in gram-positive bacteria and their viruses. 2023.09.19.558401 Preprint at https://doi.org/10.1101/2023.09.19.558401 (2023).  4. Gauthier, C. H., Cresawn, S. G. & Hatfull, G. F. PhaMMseqs: a new pipeline for constructing phage gene phamilies using MMseqs2. *G3 GenesGenomesGenetics* **12**, jkac233 (2022).  5. Zhang, C. & Mirarab, S. ASTRAL-Pro 2: ultrafast species tree reconstruction from multi-copy gene family trees. *Bioinformatics* **38**, 4949–4950 (2022).  6. Tabatabaee, Y., Zhang, C., Warnow, T. & Mirarab, S. Phylogenomic branch length estimation using quartets. *Bioinformatics* **39**, i185–i193 (2023).  7. Bryant, D. & Charleston, M. MAD roots for large trees. Preprint at https://doi.org/10.48550/arXiv.1811.03174 (2018).  8. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).  9. Steenwyk, J. L., Buida, T. J., Li, Y., Shen, X.-X. & Rokas, A. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLOS Biol.* **18**, e3001007 (2020).  10. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).  11. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).  12. Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C. & Gascuel, O. Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. *Syst. Biol.* **60**, 685–699 (2011).  13. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017). |

|  |  |
| --- | --- |
| **Accompanying files:** | |
| **Filename** | **Description of contents** |
| **Taxa\_names.xlsx** | **Contains all genomes, their accession and their taxonomic rank** |
| **Distance.xslx** | **Contains all genomes and their distances to each other genome, using tANI, gANI, ANI and metadata.** |

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

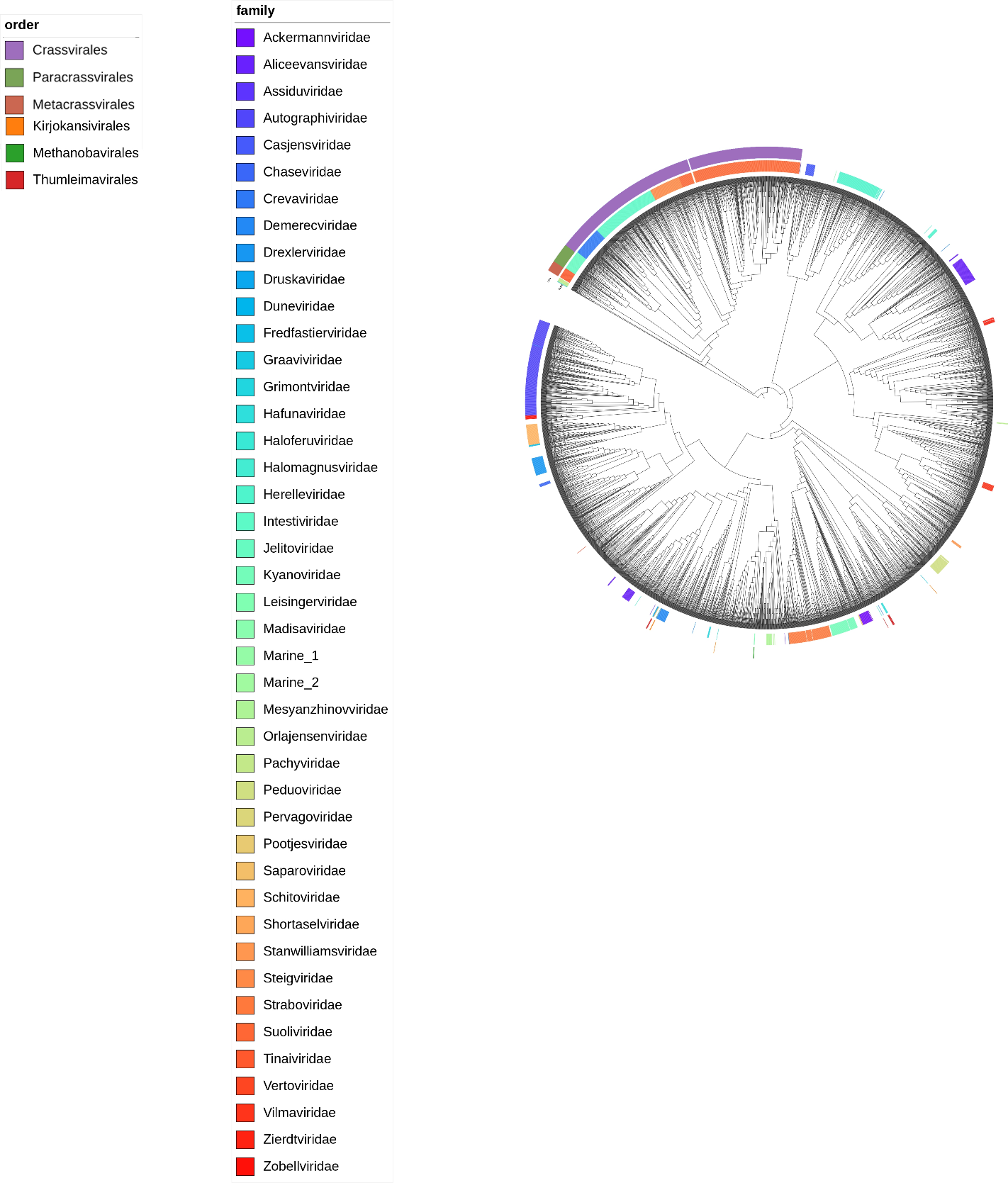


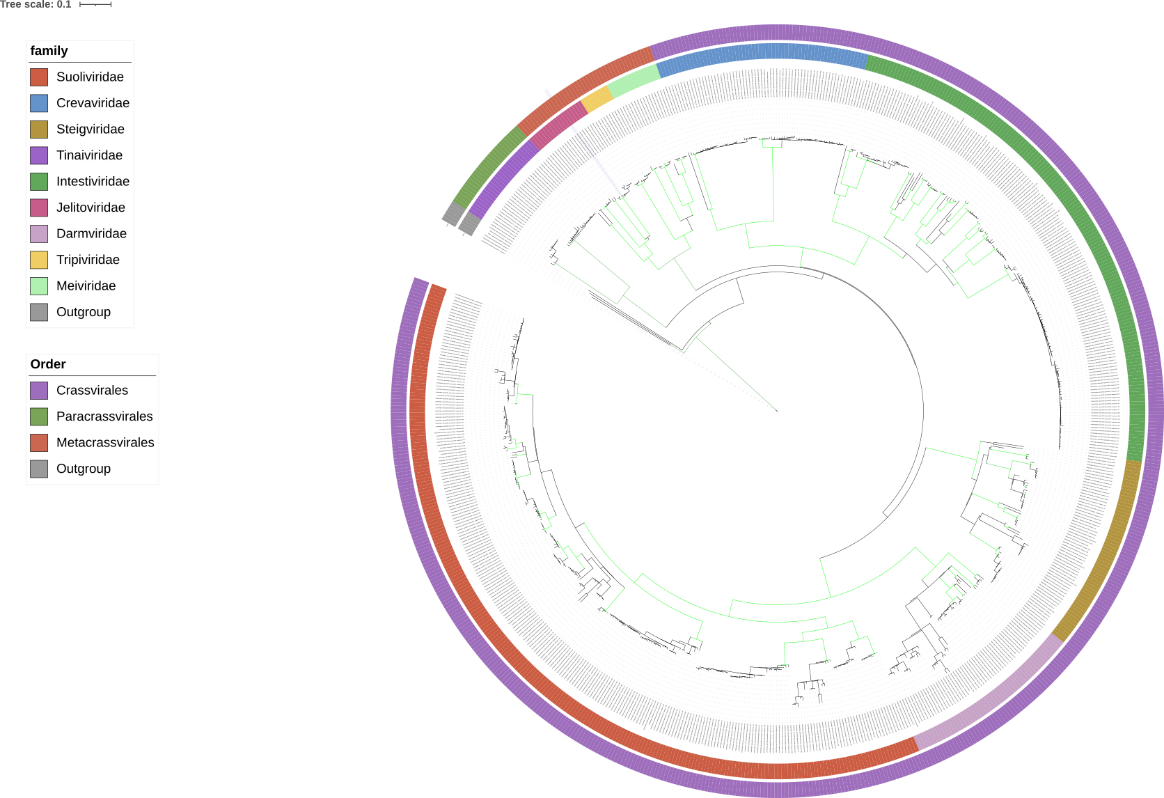
Figure 1: Foldtree structural phylogenetic tree based on the large terminase structure of 3896 members of *Caudoviricetes*. The classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring. Gaps in the annotation are due to a lack of classification for that level. The tree is midpoint rooted. Marine and Pachyviridae sequences closely related to the orders were selected and used as outgroups for all other figures( see suplementary figures).

Figure 2: Core-gene structural phylogenetic species tree of the three orders based on gene trees created using ASTRAL-Pro3 v1.2.25. Branch lengths are computed using integrated CASTLES-Pro v1.2.26. 16 rounds of searching and subsampling were performed to ensure that maximum exploration of tree configurations was attempted. The tree was rooted on the outgroups. All branches with at least 95% bootstrapping are colored green, the classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring. 

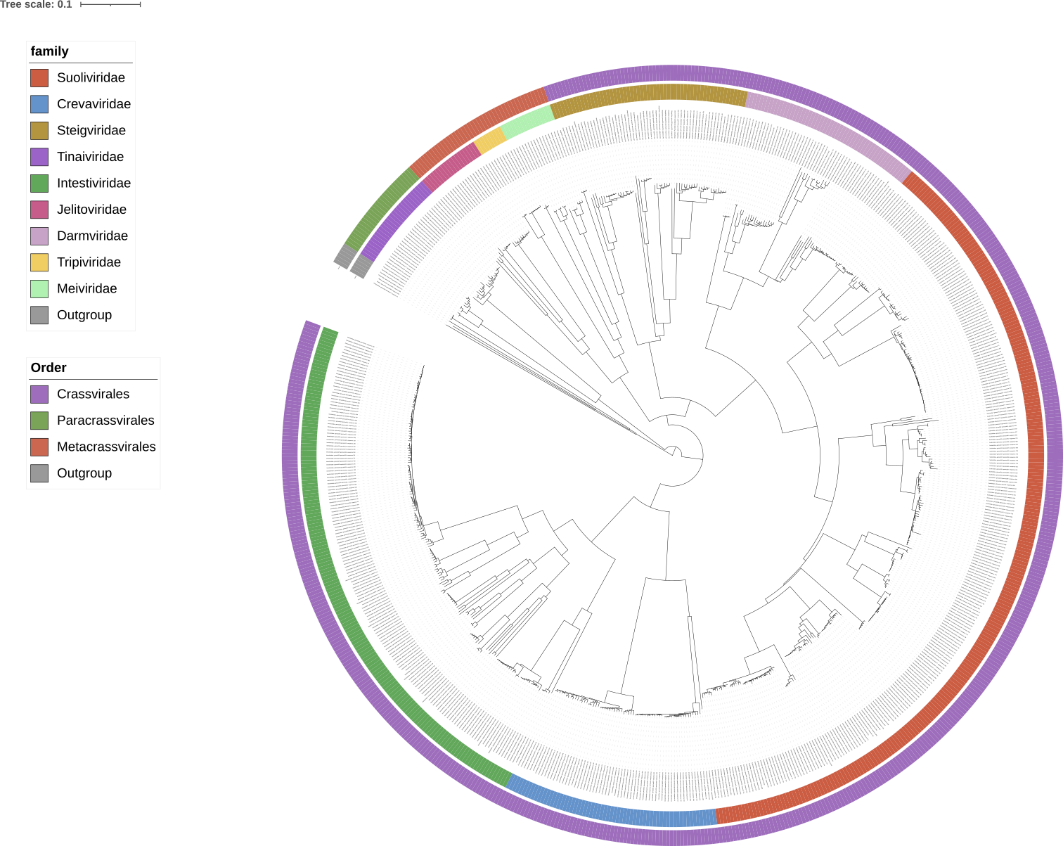
Figure 3: Single gene structural phylogeny of the DNA primase protein using foldtree. The tree was rooted using Madroot7 to ensure Minimal Ancestor Deviation(MAD). The classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring. 

Figure 4: Single gene structural phylogeny of the large terminase subunit using foldtree. The tree was rooted using Madroot7 to ensure Minimal Ancestor Deviation(MAD). The classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring

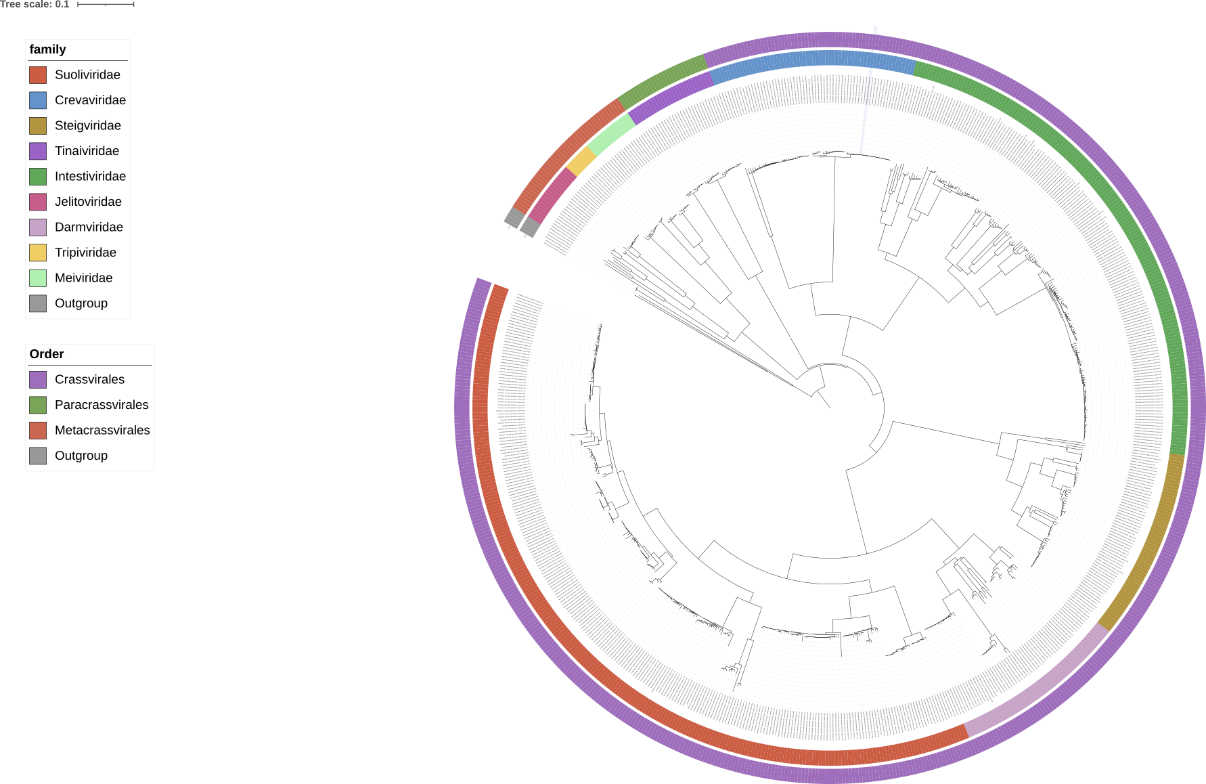
Figure 5: Single gene structural phylogeny of the portal protein using foldtree. The tree was rooted using Madroot7 to ensure Minimal Ancestor Deviation(MAD). The classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring

Figure 6:Single gene structural phylogeny of the major capsid protein using foldtree. The tree was rooted using Madroot7 to ensure Minimal Ancestor Deviation(MAD). The classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ringA colorful circle with a black background

AI-generated content may be incorrect.

Figure 7: Core-gene phylogenetic tree of the three orders based on concatenated protein created with MAFFT v7.2738 using the L-INS-i algorithm. Multiple alignments were curated using the ClipKit9 Kpic method to select informative sites. The concatenation of these alignments was used to infer maximum likelihood trees with IQTree version 2.3.410 (options -allnni -nm 4000). We evaluated the node supports using the option -bb 1,000 for ultrafast bootstraps with UFBoot211 and -alrt 1,000 for SH-aLRT12. The best evolutionary model was selected with ModelFinder13. The tree was rooted on the outgroups. All branches with at least 95% bootstrapping are colored green, the classified and proposed families are shown in the inner coloured ring, and the proposed and classified orders in the outer coloured ring. A close-up of a grid

AI-generated content may be incorrect.

Figure 8:Heatmap of the total average nucleotide identity between families of the *Crassvirales*, *Paracrassvirales* and *Metacrassvirales*. Values above 10% are set to 10%, as including higher values reduces visual granularity and obscures key patterns. The heatmap was visualized using Tidyheamaps and hierarchically clustered using the complete linkage method of R. Coloured bars indicate the family of each genome.A diagram of a graph

AI-generated content may be incorrect.

Figure 9: Heatmap of the shared protein cluster percentage between families of the Crassvirales, Paracrassvirales and Metacrassvirales. The heatmap was visualized using Tidyheamaps and hierarchically clustered using the complete linkage method of R. Coloured bars indicate the family of each genome.