

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2024

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

|  |  |  |
| --- | --- | --- |
| **Title:** | Create one (1) new family (*Felixviridae*), including one (1) new subfamily (*Maevirinae*), three (3) genera (two new: *Nakavirus, Chronisvirus;* one existent*: Certevirus*), including two (2) new species (*Nakavirus sapi* and *Chronisvirus chronis*). | |
| **Code assigned:** | 2024.044B |

|  |  |  |  |
| --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | |
| **Name** | **Affiliation** | **Email address** | **Corresponding author(s)** X |
| Nobrega F.L. | School of Biological Sciences, University of Southampton, Southampton, United Kingdom | f.nobrega@soton.ac.uk | X |
| Rothschild-Rodriguez, D. | School of Biological Sciences, University of Southampton, Southampton, United Kingdom | d.rothschild-rodriguez@soton.ac.uk |  |
| Lambon, K. | School of Biological Sciences, University of Southampton, Southampton, United Kingdom | ksl1u20@soton.ac.uk |  |

**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:** |
| Caudoviricetes 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:** | 19/06/2024 |

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

|  |  |
| --- | --- |
| **Executive Committee Meeting Decision code:** | **X** |
| A – Accept | **X** |
| 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| W - Withdrawn |  |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:** | 27/06/2024 |

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

|  |
| --- |
| **Name of accompanying Excel module:** |
| 2024.044B.A.v1.Felixviridae\_1nf.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*:  Viruses*; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes*  *Description of current taxonomy*:  Unclassified *Caudoviricetes.*  *Proposed* *taxonomic change(s):*  We’d like to add a new family, *Felixviridae*, which includes at least one new subfamily, *Maevirinae,* and at least three genera, *Nakavirus* (new), *Chronisvirus* (new) and *Certevirus* (already existent). The *Nakavirus* and *Chronisvirus* genera include at least one newly proposed species each, *Nakavirus sapi* (for phages RothC and RothD), and *Chronisvirus chronis* (for phage vB\_Kpn\_Chronis), respectively.  *Justification*:  This family has not been taxonomically characterised, with only one previously cultured *Klebsiella* phage relative, vB\_Kpn\_Chronis, and a Protoea phage, PdC23. We isolated 53 phages as part of the Klebsiella Phage Collection, two of which, RothC and RothD were not able to be assigned to any existing viral families, leading us to propose a new family, *Felixviridae.* We additionally found this family of phages to be present in metagenomes coming from human stool, a proxy for the human gut microbiota, which are also currently unclassified, but previously identified by the Gut Phage Database (GPD). Genomic analyses provide sufficient evidence to suggest the *Felixviridae* family is currently composed of at least one subfamily, *Maevirinae*, comprising RothC, D and Chronis; at least three genera, *Nakavirus, Chronisvirus,* and previously existent *Certevirus.* The *Nakavirus* genus holds phages RothC and RothD. High similarity between RothC and RothD places them under the same species proposed to be *Nakavirus sapi*. The *Chronisvirus* genus holds the phage vB\_Kpn\_Chronis, forming the *Chronisvirus chronis* species. |

|  |
| --- |
| * **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*:  Viruses*; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes.*  *Description of current taxonomy*:  Unclassified *Caudoviricetes.*  *Proposed* *taxonomic change(s)*:  **Family ‘*Felixviridae*’**  The family ‘Felixviridae’ was selected as the Roth phages C and D were all found by ‘luck’ considering their temperate lifestyle, where ‘Felix’ is latin for ‘luck’.  **Subfamily ‘*Maevirinae’***  The subfamily ‘Maevirinae’ was named after the Spanish-language colloquialism ‘mae’, commonly used in Costa Rica, where PhD student DRR is from. The colloquialism is often used in the context of calling someone (e.g. ‘hey man’) and is applicable to any gender.  **Genus ‘*Nakavirus’***  We propose that the main phages of our work (RothC and RothD) are assigned to a new genus named ‘Nakavirus’, where ‘Naka’ is Japanese for ‘inside’, which can also be interpreted as ‘stomach’ or ‘belly’. Proposed by Master student K.L. who has Japanese heritage and played a key contribution in identifying the conserved region that differentiates this genus. The founding members of this genus, RothC and RothD, are of the proposed *Nakavirus sapi* species, where sapi comes from *sapiens*.  **Genus ‘*Chronisvirus*’**  We propose the other cultured phage that falls under the *Maevirinae* subfamily, to be assigned a new genus, named after its only phage so far, vB\_Kpn\_Chronis, with the species *Chronisvirus chronis.*  **Genus ‘*Certevirus’***  While investigating RothC and RothD, we identified isolate Protoea phage PdC23 shares high similarity at the protein level. PdC23 was previously assigned to the former *Siphoviridae* family (*Certevirus* genus, *Certevirus C23* species), we propose it should now be under the *Felixviridae* family.  *Demarcation criteria:*  A threshold of 46.6% nucleotide identity to RothC and RothD was used for assigning at the subfamily level. A 70% nucleotide identity was used for assigning at the genus level, and 95% was set for same species classification. This was based on complete genome nucleotide comparisons.  *Justification*:  Temperate phages RothC and RothD were isolated in January 2023 using sewage filtrate from Southampton, UK, England, against ST323 (capsule KL21) *Klebsiella pneumoniae,* as part of the 53 phages in the Klebsiella Phage Collection open-source resource ([www.klebphacol.org](http://www.klebphacol.org)). While all other Roth phages belonged to classified taxa, RothC and RothD were left unassigned. Upon an initial BLASTn to the NCBI database, no closely related phage genomes were within the top 10 hits. Filtering to the virus taxa ID (10239) revealed the top hit was a metagenome-derived partially complete phage assembly: ctlJz2 (accession [BK029112.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK029112.1?report=genbank&log$=nucltop&blast_rank=1&RID=6K3X35RW016))1, followed by a previously cultured phage: vB\_Kpn\_Chronis (accession [MN013086.1](https://www.ncbi.nlm.nih.gov/nuccore/MN013086))2 currently ‘*unclassified*’, and several more partial metagenomes, most of which came from the same human gut study as the first hit (Table 1)1. Only phage Chronis was kept for downstream comparisons due to its complete genome (see Table 2 for the characteristics of these phages).  Network clustering analyses on vContact23 against its own database showed RothC and RothD shared edges with some phages (Figure 1a), however at very low intergenomic similarity such that RothC and RothD were assigned into their own viral cluster by the tool (Figure 1b). We therefore, proceeded to expand the database by adding 83 curated *Klebsiella* phages (INPHARED accessed January 20244 and NCBI virus accessed 29th of January 2024). Similarly, some phages showed shared edges (Figure 1c), but only phage Chronis showed a higher (48.8%) intergenomic similarity at the nucleotide level (Figure 1d). Gene-based comparisons using Clinker5 on RothC, RothD and Chronis, led us to identify a highly conserved region (1bp – 20,241bp, encoding 24 genes) on these three phages of ≈46kbp genome length (Figure 1e). The limited homology to phages in recognized families, and the high core gene similarity between RothC, RothD and Chronis, leads us to suggest these phages belong to a new family, with the proposed name of *Felixviridae.*  As mentioned above, RothC and RothD have relatives in gut metagenome (Table 1). Moreover, they were isolated to target a gut-associated pathobiont (ST323 *K. pneumoniae*)6, we therefore wanted to determine if we could find other closely related phages in publicly available gut phage databases. We searched for relatives in the Gut Phage Database (GPD)7 via BLASTn. Searching the entire genomes of RothC and RothD, we found 500 other closely related metagenomic assemblies of phages belonging to this family (Evalues of 0 to 9E-132; Figure 1f); only 45 remained when searching for the conserved region. However, due to the presence of many low quality (as assigned by the GPD’s metadata using checkv) and/or incomplete assemblies from GPD hits, we proceeded our comparisons retaining only “high-quality” assemblies with a mean SD genome completeness of 98.0% 2.84 (n=154). We used VipTree8 to infer order and family rank, which uses a proteomic approach. We selected for comparison with other dsDNA viruses of the VipTree database. The 154 phages in addition to RothC, RothD and Chronis, were separated into 10 groups (Figure 2a). We therefore continued with the characterization of group number VI, which is where RothC, RothD and Chronis lie, with an additional 19 phages from the GPD (total n=22 phages; Figure 2b). This revealed Pantoea phage PdC23 (NC\_071008/OL396571) was under the same clade as the 22 phages of new proposed *Felixviridae* family (Figure 2b), and we therefore decided to add it to the downstream analyses (n=23 phages). PdC23 phage belongs to an unclassified family (previously *Siphoviridae*) and the *Certevirus* genus (species *Certevirus C23*); it shares 24% intergenomic similarity to RothC and RothD (figure 2c). At the nucleotide level, intergenomic comparisons using VIRIDIC9 between the 23 phages of interest show >22% similarity (Figure 2c), and viral phylogenic predictor VICTOR (based on the Genome BLAST Distance Phylogeny (GBDP) tailored to prokaryotic viruses)10, predicts all 22 phage belong to the same genus, contrasting with traditional thresholds of 70% similarity for genus classification which would suggest only phages RothC and RothD share the same genus (Figure 2c,d). Additionally, these comparisons suggest a new subfamily encompassing 11/23 phages, based on a similarity of 46.6% to RothC and RothD (Figure 2c). Gene synteny plots from Clinker, demonstrate the conserved region is maintained in all 23 phages (Figure 2e). We therefore proceeded to characterize the core genome of these phages to infer taxonomy at and below the rank of family. Firstly, annotations of the conserved region for the representative phages of the proposed family (RothC and RothD), show 22/24 coding sequences are hypothetical proteins (20/23 have no known sequence homologs in other phage genomes when a BLASTp search was conducted on the NCBI database), one is the terminase large subunit, and another is a tape measure protein. Intergenomic similarities of only the extracted conserved region demonstrates all phages fall within the same genus (>70%), except for PdC23. These core genome analyses would therefore suggest PdC23 remains under its *Certevirus* genus but should move to the *Felixviridae* family, and that all other phages fall under the new *Nakavirus* genus (Figure 2f). However, maintaining a conservative classification, we propose all 23 phages fall under *Felixviridae,* 11 of the phages belong to a new proposed subfamily, *Maevirinae.* Within the *Maevirinae* subfamily, we propose two new genera, *Nakavirus* and *Chronisvirus* encompassing phages RothC, RothD and vB\_Kpn\_Chronis, respectively (as per Figure 2c). High intergenomic similarity (>95%) between RothC and RothD leads us to propose a new species for them, *Nakavirus sapi* (Figure 2f, Table 3); and lastly, a new species for vB\_Kpn\_Chronis, *Chronisvirus chronis*.  Annotated phages RothC and RothD were deposited to GenBank on the 18th of June 2024 via the Geneious portal (BankIT ID: 2841377), accession numbers: PP934563 and PP934564, respectively. The complete dataset is also available on figshare: <https://figshare.com/articles/dataset/RothC_and_RothD_genomes_and_annotations/26058466>. The manuscript entailing the complete characterization of these phages is currently under preparation. |

|  |
| --- |
| **References:** |
| 1. Tisza, M. J. & Buck, C. B. A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *Proc Natl Acad Sci U S A* **118**, e2023202118 (2021).  2. Thurgood, T. L. *et al.* Genome Sequences of 12 Phages That Infect Klebsiella pneumoniae. *Microbiol Resour Announc* **9**, (2020).  3. Bin Jang, H. *et al.* Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology 2019 37:6* **37**, 632–639 (2019).  4. Cook, R. *et al.* INfrastructure for a PHAge REference Database: Identification of Large-Scale Biases in the Current Collection of Cultured Phage Genomes. *Phage* **2**, 214 (2021).  5. Gilchrist, C. L. M. & Chooi, Y. H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473–2475 (2021).  6. Federici, S. *et al.* Targeted suppression of human IBD-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell* **185**, 2879-2898.e24 (2022).  7. Camarillo-Guerrero, L. F., Almeida, A., Rangel-Pineros, G., Finn, R. D. & Lawley, T. D. Massive expansion of human gut bacteriophage diversity. *Cell* **184**, 1098-1109.e9 (2021).  8. Nishimura, Y. *et al.* ViPTree: the viral proteomic tree server. *Bioinformatics* **33**, 2379–2380 (2017).  9. Moraru, C., Varsani, A. & Kropinski, A. M. VIRIDIC—A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses 2020, Vol. 12, Page 1268* **12**, 1268 (2020).  10. Meier-Kolthoff, J. P. & Göker, M. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* **33**, 3396–3404 (2017).  11. Hockenberry, A. J. & Wilke, C. O. BACPHLIP: Predicting bacteriophage lifestyle from conserved protein domains. *PeerJ* **9**, (2021).  12. Grami, E. *et al.* Isolation, Characterization, and Comparative Genomic Analysis of vB\_Pd\_C23, a Novel Bacteriophage of Pantoea dispersa. *Curr Microbiol* **80**, 1–11 (2023). |

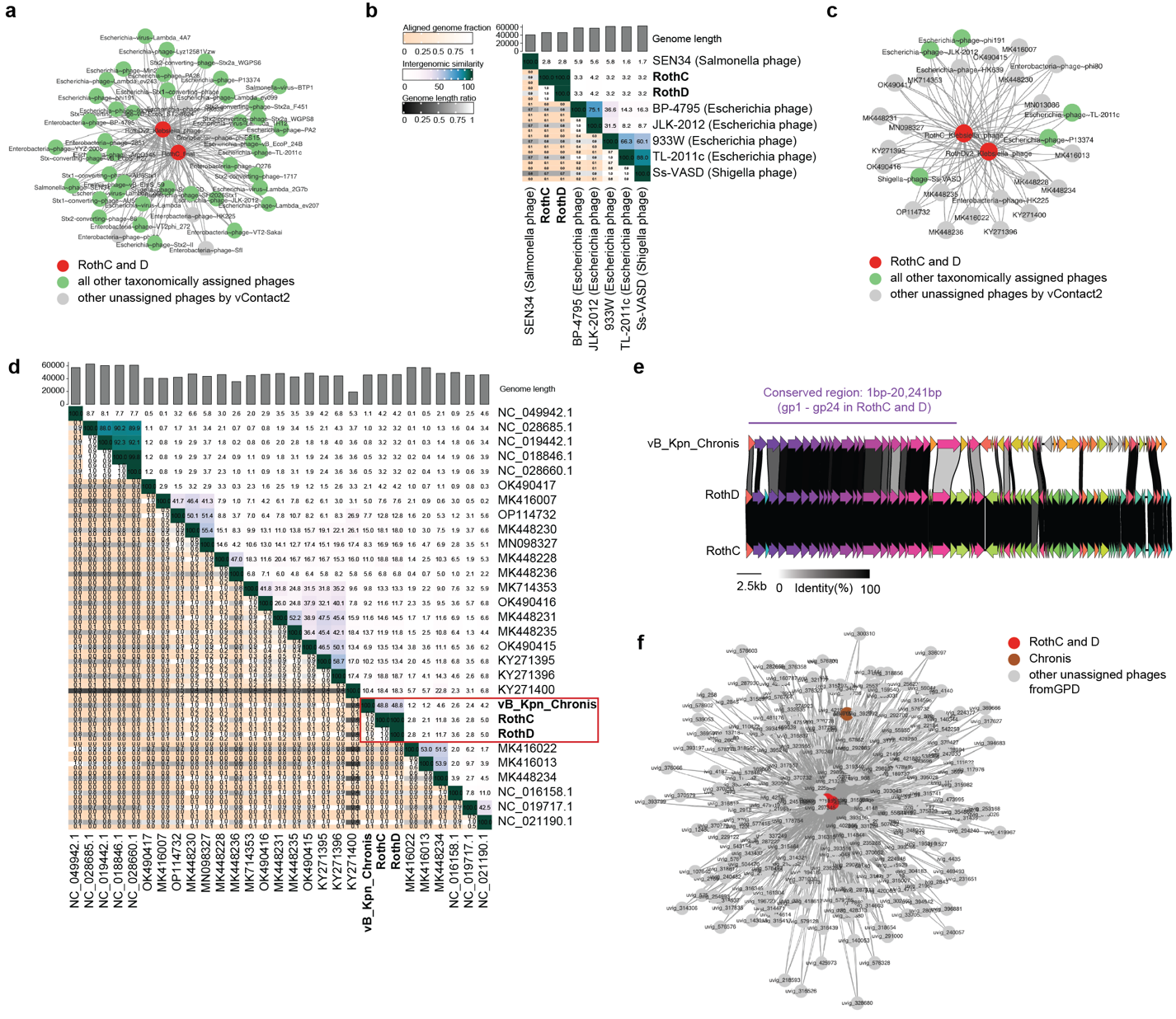
|  |
| --- |
| **Tables, Figures:** |

**Table 1**: Top 10 BLASTn hits (filtered for ID10293-viruses) for RothD (shares >99.9% similarity to RothC).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Description** | **Name** | **Max**  **Score** | **Total**  **Score** | **Query**  **Cover** | **Evalue** | **% identity** | **Acc. Length** | **Accession** |
| TPA: Caudoviricetes sp. isolate **ctlJz2**, partial genome | Caudoviricetes sp. | 18918 | 38755 | 56% | 0 | 93.45 | 47040 | [BK029112.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK029112.1?report=genbank&log$=nucltop&blast_rank=1&RID=6K8BDSUS013) |
| Klebsiella phage **vB\_Kpn\_Chronis**, complete genome | Klebsiella phage vB\_Kpn\_Chronis | 18873 | 33244 | 48% | 0 | 93.39 | 45702 | [MN013086.1](https://www.ncbi.nlm.nih.gov/nucleotide/MN013086.1?report=genbank&log$=nucltop&blast_rank=2&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ctGkX1, partial genome | Caudoviricetes sp. | 18397 | 44468 | 61% | 0 | 95.75 | 42898 | [BK057285.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK057285.1?report=genbank&log$=nucltop&blast_rank=3&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ctxCl4, partial genome | Caudoviricetes sp. | 18308 | 33973 | 46% | 0 | 95.61 | 27613 | [BK044657.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK044657.1?report=genbank&log$=nucltop&blast_rank=4&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ctEHP1, partial genome | Caudoviricetes sp. | 8818 | 9337 | 14% | 0 | 91.37 | 9900 | [BK024740.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK024740.1?report=genbank&log$=nucltop&blast_rank=5&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ctWLo4, partial genome | Caudoviricetes sp. | 8538 | 11773 | 29% | 0 | 80.32 | 46345 | [BK017775.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK017775.1?report=genbank&log$=nucltop&blast_rank=6&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ctOhM6, partial genome | Caudoviricetes sp. | 7974 | 10426 | 27% | 0 | 80.7 | 21321 | [BK020257.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK020257.1?report=genbank&log$=nucltop&blast_rank=7&RID=6K8BDSUS013) |
| TPA: Bacteriophage sp. isolate ct2EK6, partial genome | Bacteriophage sp. | 7631 | 10728 | 28% | 0 | 80.65 | 26206 | [BK032332.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK032332.1?report=genbank&log$=nucltop&blast_rank=8&RID=6K8BDSUS013) |
| TPA: Siphoviridae sp. ctpLW14, partial genome | Siphoviridae sp. ctpLW14 | 6427 | 7950 | 26% | 0 | 78.12 | 29937 | [BK015100.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK015100.1?report=genbank&log$=nucltop&blast_rank=9&RID=6K8BDSUS013) |
| TPA: Caudoviricetes sp. isolate ct2Kd5, partial genome | Caudoviricetes sp. | 5166 | 9478 | 13% | 0 | 95.9 | 20732 | [BK033248.1](https://www.ncbi.nlm.nih.gov/nucleotide/BK033248.1?report=genbank&log$=nucltop&blast_rank=10&RID=6K8BDSUS013) |

**Table 2:** Characteristics of the three cultured phages to date of the proposed *Felixviridae* family.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phage name** | **Accession No.** | **Genome length (bp)** | **GC %** | **tRNAs (\*by tRNAscan)** | **Lifestyle (\*by Bacphlip**11**)** |
| Klebsiella phage RothC | PP934563 | 46,433 | 53.1 | \*2 (tRNA-Ile; tRNA-Arg) | \*Temperate (score: 0.96) |
| Klebsiella phage RothD | PP934564 | 46,356 | 53.1 | \*2 (tRNA-Ile; tRNA-Arg) | \*Temperate (score: 0.96) |
| Klebsiella phage vB\_Kpn\_Chronis2 | [MN013086.1](https://www.ncbi.nlm.nih.gov/nucleotide/MN013086.1?report=genbank&log$=nucltop&blast_rank=2&RID=6K8BDSUS013) | 45,702 | 52.3 | 1 (tRNA-Asn) | Temperate |
| Pantoea phage PdC312 | [NC\_071008.1](https://www.ncbi.nlm.nih.gov/nuccore/NC_071008.1), [OL396571.1](https://www.ncbi.nlm.nih.gov/nuccore/OL396571.1) | 44,715 | 49.7% | unknown | Virulent |

****

**Figure 1. Preliminary data supporting the proposal of a the new *Felixviridae* Family. a)** vContact2 RothC and RothD cluster using the default database with no filters. **b)** VIRIDIC intergenomic similarity heatmap of RothC, D and six randomly selected phages from the cluster in (a). **c)** vContact2 RothC and RothD cluster adding our own curated set of *Klebsiella* phages from INPHARED and NCBI virus databases (accessed January 2024). **d)** VIRIDIC intergenomic similarity heatmap of RothC, RothD and other the other *Klebsiella* phages from the cluster in (c). Red box highlights the only phage closely related to RothC and RothD, Chronis. **e)** Clinker synteny plot of RothC, RothD and Chronis, highlighting the identified conserved region. **f)** vContact2 RothC and RothD cluster of the 500 Gut Phage Database (GPD) hits, with the addition of phage Chronis.

A close-up of a chart

Description automatically generated

**Figure 2. Nucleotide and proteome analyses of the new *Felixviridae* candidates. a)** VipTree proteomic tree of the 154 GPD hits in addition to RothC, RothD and vB\_Kpn\_Chronis, against the dsDNA viruses from the VipTree database. The red arrows indicate where the 157 interrogated phages fall, suggesting their existence in 10 groups, some of which entail new families. Family number VI, marked by a star, indicates the new proposed *Felixviridae* family where RothC, RothD and Chronis lie. **b)** Pruned tree from (a) showing only family VI (n=22 phages) and sister clades, showing Pantoea phage PdC23 as a close relative. **c)** Intergenomic similarity matrix between the 22 phages of interest plus PdC23 (complete genome), as the proposed *Felixviridae* family, *Maevirinae* subfamily, and *Nakavirus* and *Chronisvirus* genera. Calculated by VIRIDIC. **d)** Phylogenetic tree calculated by VICTOR at the nucleotide level for the 22 phages in question plus PdC23, visualized on iTOL. The phages’ lifestyle was predicted by Bacphlip (all with a score 0.7). **e)** Clinker-built synteny plot of the phages showing the conserved region (or core genome) encompassing 24 coding sequences (as per the reference phages of this family: RothC and RothD). **f)** as (c) but with only the core genome shown in (e). These analyses argue a different classification. The trees in this figure were rooted at the mid-point.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 3:** Proposed new taxonomy. | | | | |
| Family | *Felixviridae* | | | |
| Subfamily | *Maevirinae* | | |  |
| Genus | *Nakavirus* | *Chronisvirus* | Others(not being proposed) | *Certevirus*  (already existent) |
| Species | *Nakavirus sapi* | *Chronisvirus chronis* | Others (not being proposed) | *Certevirus C23*  (already existent) |
| Phage | RothC  RothD | vB\_Kpn\_Chronis | uvig\_267671  ivig\_3440  uvig\_257380  uvig\_321813  uvig\_337249  uvig\_394187  uvig\_396516  uvig\_420063 | Pantoea phage PdC23 |