

The International Committee on Taxonomy of Viruses

Taxonomy Proposal Form, 2024

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

|  |  |  |  |
| --- | --- | --- | --- |
| **Title:** | Create a new family, *Ehrlichviridae*, for a group of *Bacillus* Andromeda-like phages (Class: *Caudoviricetes*) | |  |
| **Code assigned:** | 2024.012B |

|  |  |  |  |
| --- | --- | --- | --- |
| **Author(s), affiliation and email address(es):** | | | |
| **Name** | **Affiliation** | **Email address** | **Corresponding author(s)** X |
| Barylski J | Uniwersytet im. Adama Mickiewicza w Poznaniu, Poland | b54026@amu.edu.pl | **x** |
| Moraru C | Carl von Ossietzky Universität Oldenburg, Germany | liliana.cristina.moraru@uol.de |  |
| Tolstoy I | National Center for Biotechnology Information, MD, USA | tolstoy@ncbi.nlm.nih.gov |  |
| Kropinski AM | University of Guelph, Ontario, Canada [AMK] | Phage.Canada@gmail.com |  |

**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:** | 19/04/2024 |

**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| W - Withdrawn |  |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:** | DD/MM/YYYY |

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

|  |
| --- |
| **Name of accompanying Excel module:** |
| 2024.012B.A.v1.Ehrlichviridae\_nf.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:** | | **Y** |
| **Taxon name** | **Person from whom the name is derived** | **Attached X** |
| *Ehrlichviridae* | Melanie and Kenneth Ehrlich | Y |
|  |  |  |
|  |  |  |

|  |
| --- |
| **Abstract of Taxonomy Proposal:** |
| *Taxonomic rank(s) affected*:  Realm *Duplodnaviria*, kingdom *Heunggongvirae*, phylum *Uroviricota*, class *Caudoviricetes*  *Description of current taxonomy*:  The genus *Andromedavirus* currently exists as a floating genus in the class *Caudoviricetes*  *Proposed* *taxonomic change(s):*  A. To create a new genus *Suttonboningtonvirus* with one species  B. To create a new genus *Gettysburgvirus* with three species  C. To add three new species to the genus *Andromedavirus*  D. To create a new single species genus *Anathvirus*  E. To create a new single species genus *Dazunavirus*  F. To create a new single species genus *Chennaivirus*  G. To create a new single species genus *Nairobivirus*  H. To create a new family, *Ehrlichviridae*, for the above-mentioned taxa.  *Justification*:  The phages comprising these taxa form a deep branching clade using tblastx distances and single gene phylogeny. Core gene analysis shows the presence of 15 proteins conserved across all members of the proposed family. |

|  |
| --- |
| * **Text of Taxonomy proposal:** |
| *Taxonomic rank(s) affected*: Species, genus and Family  *Description of current taxonomy*:  At present only the genus *Andromedavirus* exists.  *Proposed* *taxonomic change(s)*:  A. To create a new genus *Suttonboningtonvirus* with one species  B. To create a new genus *Gettysburgvirus* with three species  C. To add three new species to the genus *Andromedavirus*  D. To create a new single species genus *Anathvirus*  E. To create a new single species genus *Dazunavirus*  F. To create a new single species genus *Chennaivirus*  G. To create a new single species genus *Nairobivirus*  H. To create a new family, *Ehrlichviridae*, for the above mentioned taxa.  *Demarcation criteria:*  **Species demarcation criteria:** Two phages are assigned to the same species if their genomes are more than 95% identical over their genome length for isolates.  These values can be calculated by a number of tools, such as BLASTn [1,2] – usually calculated using intergenomic distance calculator VIRIDIC [3].  **Genus demarcation criteria:** In search for criteria that create cohesive and distinct genera that are reproducible and monophyletic, the Bacterial Viruses Subcommittee has established 70% nucleotide identity of the genome length as the cut-off for genera. Genus-level groupings should always be monophyletic in the signature genes, as tested with a phylogenetic tree. [10]  **Subfamily demarcation criteria:** Subfamilies are to be created when two or more genera are related below the family level. In practical terms, this usually means that they share a low degree of sequence similarity (usually about 40-50%) and that the genera form a clade in a marker tree phylogeny. [10]  **Family demarcation criteria:** The family is represented by a cohesive and monophyletic group in the main predicted proteome-based clustering tools (VirClust, ViPTree, GRAViTy dendrogram, vConTACT2 network). Members of the family share a significant number of orthologous genes (the number will depend on the genome sizes and number of coding sequences of members of the family). [10]  *Justification*:  The phages comprising these taxa form a deep branching clade using tblastx distances and single gene phylogeny. Core gene analysis shows the presence of 15 proteins conserved across all members of the proposed family. |

|  |
| --- |
| **References:** |
| 1. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W, Marchler-Bauer A, Landrum M, Lathrop S, Lu Z, Madden TL, O'Leary N, Phan L, Rangwala SH, Schneider VA, Skripchenko Y, Wang J, Ye J, Trawick BW, Pruitt KD, Sherry ST. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2021 Jan 8;49(D1):D10-D17. doi: 10.1093/nar/gkaa892. PMID: 33095870  2. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):D733-45. doi: 10.1093/nar/gkv1189. PMID: 26553804.  3. Moraru C, Varsani A, Kropinski AM. VIRIDIC-A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. Viruses. 2020 Nov 6;12(11):1268. doi: 10.3390/v12111268. PMID: 33172115; PMCID: PMC7694805. <http://kronos.icbm.uni-oldenburg.de/viridic/>  4. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral proteomic tree server. Bioinformatics. 2017; 33(15):2379-2380. doi:10.1093/bioinformatics/btx157. PubMed PMID: 28379287. <https://www.genome.jp/viptree/>  5. Rohwer F, Edwards R. The Phage Proteomic Tree: a genome-based taxonomy for phage. J Bacteriol. 2002 Aug;184(16):4529-35. PubMed PMID: 12142423  6. Turner D, Reynolds D, Seto D, Mahadevan P. CoreGenes3.5: a webserver for the determination of core genes from sets of viral and small bacterial genomes. BMC Res Notes. 2013;6:140. doi: 10.1186/1756-0500-6-140. PMID: 23566564.  7. Davis P, Seto D, Mahadevan P. CoreGenes5.0: An Updated User-Friendly Webserver for the Determination of Core Genes from Sets of Viral and Bacterial Genomes. Viruses. 2022 Nov 16;14(11):2534. doi: 10.3390/v14112534. PMID: 36423143; PMCID: PMC9693508.  8. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008;36(Web Server issue):W465-9. doi: 10.1093/nar/gkn180. Epub 2008 Apr 19. PMID: 18424797.  9. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol. 2006;55(4):539-52. PMID: 16785212. DOI: 10.1080/10635150600755453.  10. Turner D, Kropinski AM, Adriaenssens EM. A Roadmap for Genome-Based Phage Taxonomy. Viruses. 2021 Mar 18;13(3):506. doi: 10.3390/v13030506. PMID: 33803862; PMCID: PMC8003253.  11. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol. 2019 Jun;37(6):632-639. doi: 10.1038/s41587-019-0100-8. Epub 2019 May 6. PMID: 31061483.  12. Bolduc B, Jang HB, Doulcier G, You ZQ, Roux S, Sullivan MB. vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. PeerJ. 2017 May 3;5:e3243. doi: 10.7717/peerj.3243. PMID: 28480138; PMCID: PMC5419219.  13. Moraru C. VirClust-A Tool for Hierarchical Clustering, Core Protein Detection and Annotation of (Prokaryotic) Viruses. Viruses. 2023 Apr 19;15(4):1007. doi: 10.3390/v15041007. PMID: 37112988; PMCID: PMC10143988.  14. Loessner MJ. Improved procedure for bacteriophage typing of Listeria strains and evaluation of new phages. Appl Environ Microbiol. 1991 Mar;57(3):882-4. doi: 10.1128/aem.57.3.882-884.1991. PMID: 2039238; PMCID: PMC182812.  15. Loessner MJ, Inman RB, Lauer P, Calendar R. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of Listeria monocytogenes: implications for phage evolution. Mol Microbiol. 2000 Jan;35(2):324-40. doi: 10.1046/j.1365-2958.2000.01720.x. PMID: 10652093.  16. Casey A, Jordan K, Coffey A, McAuliffe O. Complete Genome Sequences of vB\_LmoS\_188 and vB\_LmoS\_293, Two Bacteriophages with Specificity for Listeria monocytogenes Strains of Serotypes 4b and 4e. Genome Announc. 2015 Apr 9;3(2):e00040-15. doi: 10.1128/genomeA.00040-15. PMID: 25858822; PMCID: PMC4392134.  17: Dorscht J, Klumpp J, Bielmann R, Schmelcher M, Born Y, Zimmer M, Calendar R, Loessner MJ. Comparative genome analysis of Listeria bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. J Bacteriol. 2009 Dec;191(23):7206-15. doi: 10.1128/JB.01041-09. Epub 2009 Sep 25. PMID: 19783628; PMCID: PMC2786548.  16. Nguyen LT, Schmidt HA, von Haeseler A, and Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Molecular Biology and Evolution, 32:268-274. <https://doi.org/10.1093/molbev/msu300>  17. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the ultrafast bootstrap approximation. Molecular Biology and Evolution, 35:518–522. <https://doi.org/10.1093/molbev/msx281>  18. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, and Jermiin JS (2017) ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates, Nature Methods, 14:587–589. <https://doi.org/10.1038/nmeth.4285>  19. Loney RE, Delesalle VA, Chaudry BE, Czerpak M, Guffey AA, Goubet-McCall L, McCarty M, Strine MS, Tanke NT, Vill AC, Krukonis GP. A Novel Subcluster of Closely Related Bacillus Phages with Distinct Tail Fiber/Lysin Gene Combinations. Viruses. 2023 Nov 17;15(11):2267. doi: 10.3390/v15112267. PMID: 38005943; PMCID: PMC10674732. |

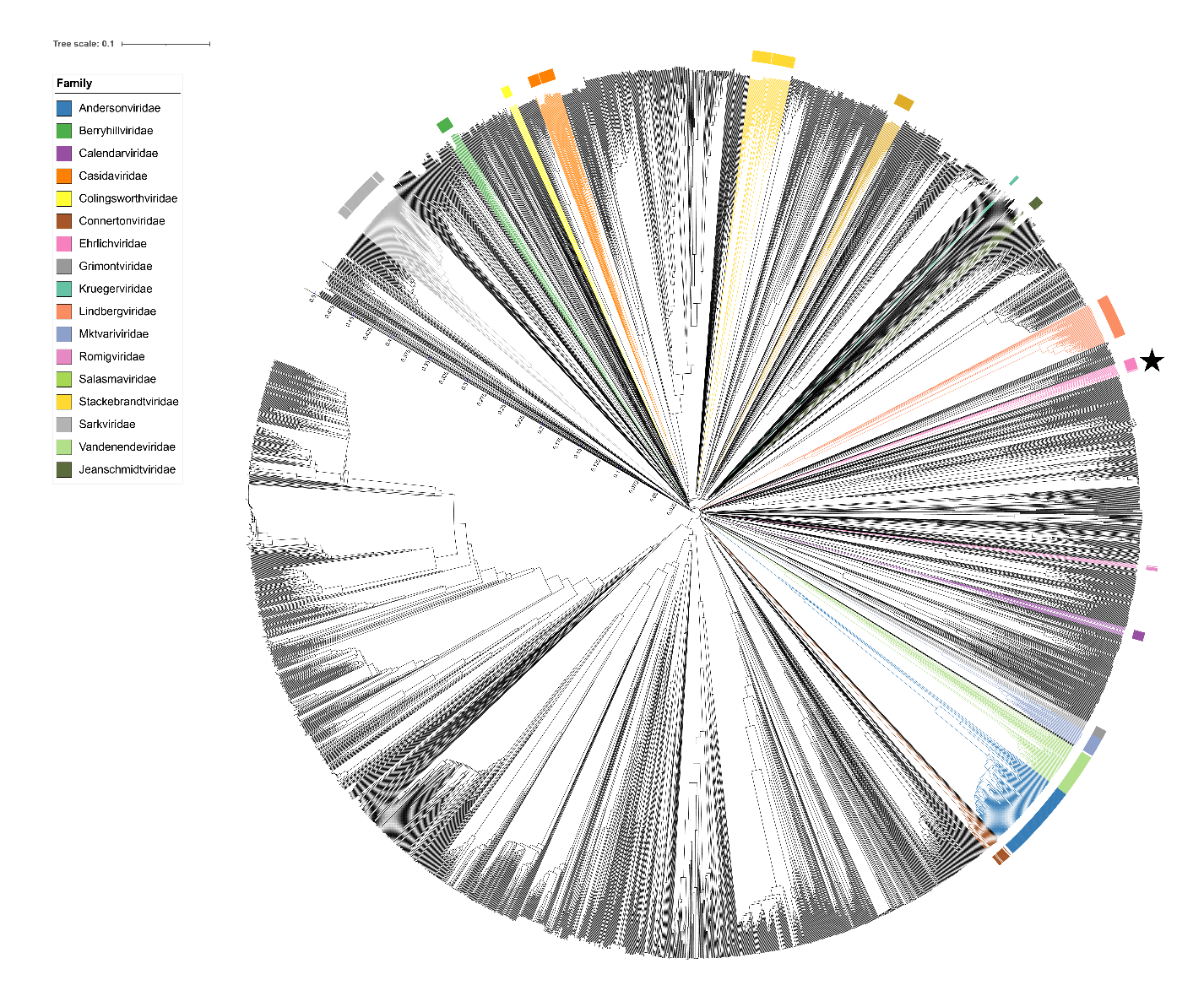
|  |
| --- |
| **Tables, Figures:** |

<Start here>



Figure 1. VIRIDIC heat map of a portion of the members of this family: VIRIDIC (Virus Intergenomic Distance Calculator; VIRIDIC (Virus Intergenomic Distance Calculator; [3]; http://rhea.icbm.uni-oldenburg.de/VIRIDIC/) computes pairwise intergenomic distances/similarities amongst phage genomes. Data values which are bordered in black correspond to strains. Abbreviations: phg = phage; Baci = *Bacillus*; Stap = *Staphylococcus*; Geob = *Geobacillus*. Since this figure is somewhat difficult to read we have appended the complete VIRIDIC heatmap (Ehrlichviridae\_2024\_VIRIDIC\_heatmap). The coloured accession numbers and phage names in Column A represent ICTV-recognized species.

The results indicate the existence of eight genera.

Figure 2. ViPTree [4] analysis Proteomic tree of 4,408 bacterial viruses with proposed viral families labeled by the coloured ring. The *Ehrlichviridae* are marked with a star symbol. The hierarchical tree was created using ViPTreeGen (version 1.1.2) [4] and annotated using iToL [15-16]. The tree is based on a dissimilarity matrix generated by pairwise tBLASTx scores between each of the genomes.

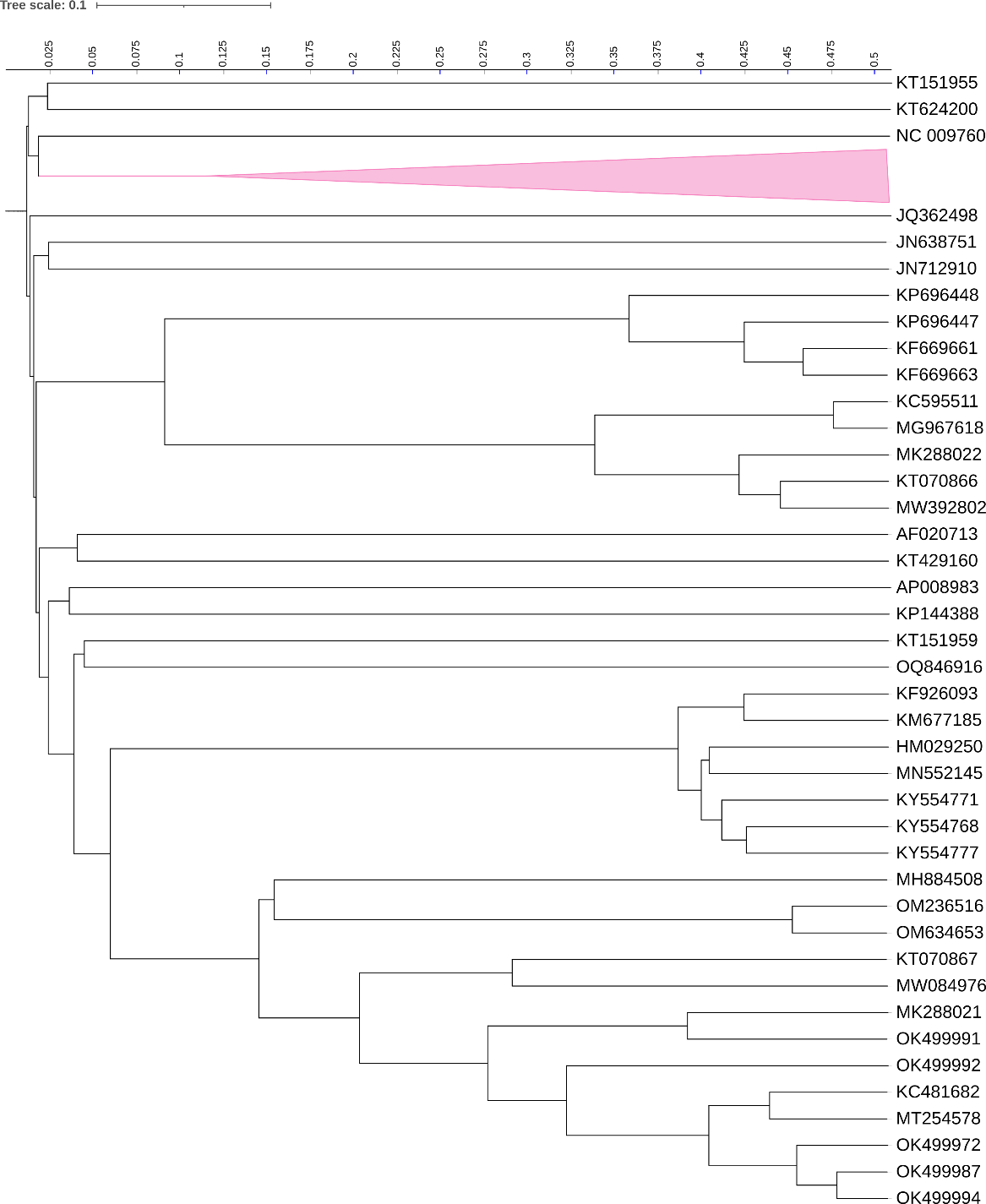
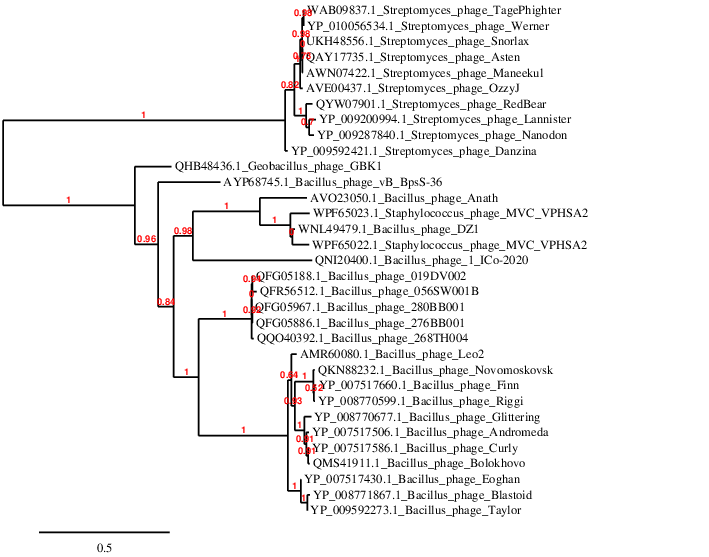


Figure 3. ViPTree [4] hierarchical tree pruned to show the proposed *Connertonviridae* as a pink-coloured collapsed clade alongside neighbouring clades.



Figure 4. VirClust protein heatmap: at the first level, proteins are grouped based on their reciprocal BLASTP similarities into protein clusters, or PCs. At the second level, PCs are grouped based on their Hidden Markov Model (HMM) similarities into protein superclusters, or PSCs. AT the third, still experimental level, PSCs are grouped based on their HMM similarities into protein super-superclusters, or PSSC [13].



**Figure 5. Phylogeny:** The phylogenetic tree was constructed using the DNA polymerase of these and related phages with phylogeny.fr in “one click” mode [6]. "The "One Click mode" targets users that do not wish to deal with program and parameter selection. By default, the pipeline is already set up to run and connect programs recognized for their accuracy and speed (MUSCLE for multiple alignment and PhyML for phylogeny) to reconstruct a robust phylogenetic tree from a set of sequences. The usual bootstrapping procedure is replaced by a new confidence index that is much faster to compute. See: Anisimova M., Gascuel O. Approximate likelihood ratio test for branches: A fast, accurate and powerful alternative [7] for details." The members of the *Ehrlichviridae* are indicated with a blue rectangle and show deep branching from the nearest neighbours – *Streptomyces* phages.

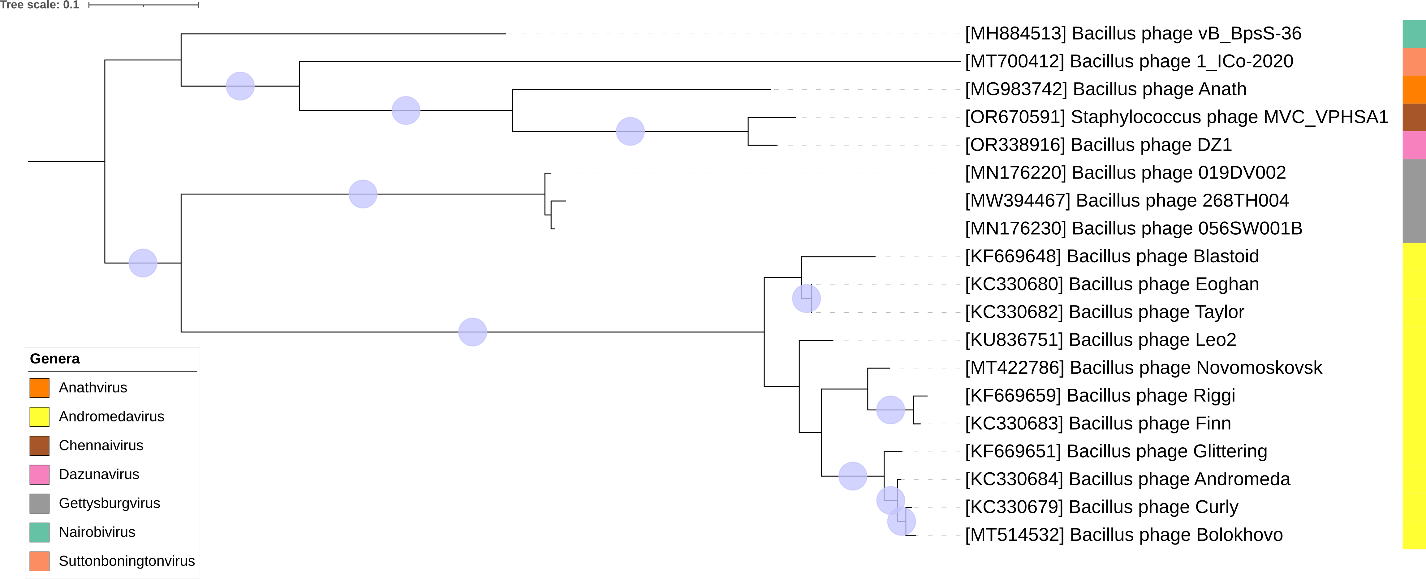


Figure 6. Core genome phylogeny of the proposed *Erlichviridae* family of bacterial viruses. A partitioned protein ML phylogeny was created from 15 genes present in all species of the proposed family. Alignments were performed using MAFFT in e-insi mode and trimmed using trimAl with a gap threshold of 0.5. The tree was calculated using IQ-Tree2 with 1000 ultrafast (UF) bootstrap replicates and SH-Alrt tests with -m TEST to optimise models for each alignment. The tree is rooted at the midpoint and UF bootstrap support ≥ 95% are shown. The coloured strips indicate proposed genera and subfamilies. Figure generated in iTOL, the Interactive Tree Of Life [14, 15]

Table 1. Signature genes in the proposed X family of bacterial viruses. Genes were identified by clustering with MMSeqs2, with thresholds of 35% sequence similarity and 50% coverage.

|  |  |  |  |
| --- | --- | --- | --- |
| **protein cluster** | **No. of genomes (19 total)** | **Percentage of genomes present in protein cluster** | **Predicted gene function** |
| 1 | 19 | 100% | scaffold protein |
| 2 | 19 | 100% | ssDNA-binding protein |
| 3 | 19 | 100% | head-to-tail adaptor |
| 4 | 19 | 100% | major capsid protein |
| 5 | 19 | 100% | hypothetical protein |
| 6 | 19 | 100% | endonuclease |
| 7 | 19 | 100% | DNA primase |
| 8 | 19 | 100% | portal vertex protein |
| 9 | 19 | 100% | major tail protein |
| 10 | 19 | 100% | transcriptional regulator |
| 11 | 19 | 100% | Terminase, large subunit |
| 12 | 19 | 100% | holin or phosphodiesterase |
| 13 | 19 | 100% | helicase |
| 14 | 19 | 100% | endonuclease |
| 15 | 19 | 100% | deoxynucleoside monophosphate kinase |
| 16 | 18 | 94.74% | head morphogenesis protein |
| 17 | 18 | 94.74% | DNA polymerase III alpha subunit |
| 18 | 18 | 94.74% | PD-(D/E)XK nuclease superfamily protein |
| 19 | 18 | 94.74% | virion morphogenesis protein |
| 20 | 18 | 94.74% | tail terminator |
| 21 | 18 | 94.74% | hypothetical protein |
| 22 | 18 | 94.74% | FtsK/SpoIIIE ATPase |

**Proposals Data:**

**A.** **To create a new genus *Suttonboningtonvirus* with one species**

**B.** **To create a new genus *Gettysburgvirus* with three species**

**C.** **To add three new species to the genus *Andromedavirus***

**D.** **To create a new single species genus *Anathvirus***

**E.** **To create a new single species genus *Dazunavirus***

**F.** **To create a new single species genus *Chennaivirus***

**G. To create a new single species genus *Nairobivirus***

**H. To create a new family, *Ehrlichviridae*, for the above mentioned taxa.**

**Taxonomic Proposals:**

1. **To create a new genus *Suttonboningtonvirus* with one species**

**Origin of the name of this taxon:** This taxon was named after The University of Nottingham’s Sutton Bonington Campus where *Bacillus* phage 1\_ICo-2020 was characterized.

**Historical aspects:** This phage was originally isolated from wastewater in Egypt.

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage 1\_ICo-2020 | MT700412.1 | 51.23 | 75 | 100 | 100 |

1. **To create a new genus *Gettysburgvirus* with three species**

**Origin of the name of this taxon:** This taxon is named after Gettysburg College, where the phage isolation, DNA preparation, and annotation analysis was performed.

**Historical aspects:** Soil was collected from various locations in the southwest United States and used to isolate these *Bacillus subtilis* phages [19].

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage 056SW001B | MN176230.1 | 52.69 | 83 | 100 | 100 |
| *Bacillus* phage 268TH004 | MW394467.1 | 52.83 | 82 | 86.8 | 86.7 |
| *Bacillus* phage 019DV002 | MN176220.1 | 52.75 | 84 | 93.3 | 91.6 |

**(\*) determined using VIRIDIC [3]**

**(\*\*) determined using CoreGenes 3.5 [6]**

**Conclusion:** The DNA sequence similarity value is consistent with membership in the same genus

**C.** **To add three new species to the genus *Andromedavirus***

**Origin of the name of this taxon:** NA

**Historical aspects:** This genus was created originally through TaxoProp 2013.037a-dB.A.v3.Andromedalikevirus

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage Andromeda | KC330684.1 | 49.26 | 79 | 100 | 100 |
| *Bacillus* phage Novomoskovsk | MT422786.1 | 49.26 | 81 | 73.7 | 87.3 |
| *Bacillus* phage Leo2 | KU836751.1 | 48.59 | 74 | 80.8 | 89.9 |
| *Bacillus* phage Bolokhovo | MT514532.1 | 49.68 | 81 | 91.8 | 88.6 |

**(\*) determined using VIRIDIC [3]**

**(\*\*) determined using CoreGenes 3.5 [6]**

**Conclusion:** The DNA sequence similarity value is consistent with membership in the same genus

**D.** **To create a new single species genus *Anathvirus***

**Origin of the name of this taxon:** The name of this taxon is derived directly from hat of Bacillus phage Anath

**Historical aspects:** Bacillus phage Anath was isolated from groundwater against *Bacillus mycoides* at Aarhus Universitet in Denmark.

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage Anath | MG983742.1 | 52.37 | 76 | 100 | 100 |

**E.** **To create a new single species genus *Dazunavirus***

**Origin of the name of this taxon:** The name of this taxon derived from the first virus of its type, *Bacillus* phage DZ1

**Historical aspects:** *Bacillus* phage DZ1 was isolated from sewage against *Bacillus cereus* by scientist in the Institute of Microbiology, Guangdong Academy of Sciences (China).

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage DZ1 | OR338916.1 | 49.87 | 67 | 100 | 100 |

**F.** **To create a new single species genus *Chennaivirus***

**Origin of the name of this taxon:** The name of this taxon is derived directly from that of the city that is the capital of the state of Tamil Nadu, India.

**Historical aspects:** *Staphylococcus* siphophage MVC\_VPHSA1 was isolated in 2019 from sewage against *Staphylococcus aureus* by the staff at Madras Veterinary

College, TANUVAS, Veppery, Chennai, Tamil Nadu 600007, India

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Staphylococcus* phage MVC\_VPHSA1 | OR670591.1 | 51.39 | 109 | 100 | 100 |

**G. To create a new single species genus *Nairobivirus***

**Origin of the name of this taxon:** The name of this taxon derives from the city, Nairobi, where at Jomo Kenyatta University of Agriculture and Technology, the first virus of its type, *Bacillus* phage vB\_BpsS-36 was isolated.

**Historical aspects:** *Bacillus* phage vB\_BpsS-36 was isolated from haloalkaliphilic Lake Elmenteita sediment against *Bacillus pseudalcaliphilus.*

**Genomic characterization:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phage name | INSDC | Size (Kb) | Protein | Overall % DNA sequence identity (\*) | Overall % homologous proteins (\*\*) |
| *Bacillus* phage vB\_BpsS-36 | MH884513.1 | 50.48 | 68 | 100 | 100 |

**H. To create a new family, *Ehrlichviridae*, for the above mentioned taxa.**

**Origin of the name of this taxon:** This taxon is named in honour of molecular biologist Melanie Ehrlich (b. 1946; New York City, New York, USA) and organic chemist/biochemist Kenneth Craig Ehrlich (b. 1944; New York City, New York, USA).

ME received her PhD in 1971 from the State University of New York, Stony Brook. During her postdoctoral training at Albert Einstein School of Medicine in New York, she began studies of the structure of *Bacillus* phage SP-15 DNA, which has a unique sensitivity to alkali. She continued that research in collaboration with her husband nine years later. They found that the novel base that replaces 62% of T residues in this phage DNA has a phosphoglucuronate that is attached to the 5-(4’,5’-dihydroxypentyl)uracil. This was, and still is, a novel finding of a natural DNA containing a uronic acid moiety and a phosphate that is not part of the phosphodiester backbone. Twenty-six years later, she gave a sample of four-decades old SP-15 DNA (saved as fibers in 50% ethanol) to Andrew Kropinski for him to sequence this extraordinary DNA. Surrounding this research were her studies on a *Xanthomonas* phage (XP-12), which has all its C residues 5-methylated and then on 5-methylcytosine in human DNA from normal and diseased individuals (cancer, a DNMT3B-deficiency syndrome, atherosclerosis, osteoporosis, and muscle diseases). She is the founder and current president of the Epigenetics Society.



KE received his PhD in organic chemistry in 1969 at the State University of New York, Stony Brook and did postdoctoral research at Columbia University. Together with his wife, Melanie Ehrlich, he developed an interest in understanding the chemical consequences of DNA base modifications on hydrogen bonding and the stability of the DNA helix. For example, they found that the naturally occurring methylation of all C residues in XP12 DNA, which makes the helix more hydrophobic, stabilizes it. In contrast, in SP-15 DNA, the unusual phosphoglucuronate attached to a modified T residue containing a 5-carbon side chain explains its low helix denaturation temperature as well as its highly unusual alkaline sensitivity. Besides using his chemistry background to increase our understanding of the biochemistry of highly modified phage DNA, he helped elucidate biochemical pathways of toxin production by fungi (*Aspergillus*) and the structure and chemistry of proteoglycans in aorta, among other diverse research areas.



**Conclusion:** All our analyses (genomic, proteomic and phylogenetic) reveal that these phages listed above belong to a new family which we have named in honour of the pioneering *Bacillus* phage scientists, Melanie and Kenneth Ehrlich.