This Word module should be used for all taxonomic proposals.

Please complete **Part 1** and:

either **Part 3** for proposals to create new taxa or change existing taxa

or **Part 2** for proposals of a general nature.

Submit the completed Word module, together with the accompanying Excel module named in Part 3, to the appropriate ICTV Subcommittee Chair.

The Word module explains and justifies your proposal. The Excel module is a critical document that will be used to implement the proposed taxonomic changes once they are approved and ratified. If proposals presented in the Word module are not presented accurately in the Excel module, the taxonomic changes cannot proceed.

For guidance, see the notes written in blue, below, and the Help Notes in file Taxonomic\_Proposals\_Help\_2019.

**Part 1:** **TITLE, AUTHORS, etc**

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2019.103B*** |  |
| **Short title:** Create one new family (*Autographiviridae*) including nine subfamilies and one hundred and thirty-two genera in the order *Caudovirales* |
|  |
| **Author(s) and email address(es):**  |
| List authors in a single line *Archives of Virology* citation format (e.g. Smith AB, Huang C-L, Santos, F) | Provide email address for each author in a single line separated by semi-colons |
| Turner D, Kropinski AM, Alfernas-Zerbini P, Buttimer C, Lavigne R, Bister JR, Tolstoy I, Morozova VV, Babkin IV, Kozlova YN, Tikunov AY, Tikunova NV, Adriaenssens EM  | Phage.Canada@gmail.com; Evelien.Adriaenssens@quadram.ac.uk; Dann2.Turner@uwe.ac.uk; palfenas@uvf.br;tolstoy@ncbi.nlm.nih.gov; rob.lavigne@kuleuven.be;colin.buttimer@mycit.ie; mba@ufv.br; morozova@niboch.nsc.ru; i\_babkin@mail.ru; ulona@ngs.ru; arttik@ngs.ru; tikunova@niboch.nsc.ru |
| **Author(s) institutional address(es) (optional):**

|  |
| --- |
| Provide institutional addresses, each on a single line followed by author(s) initials (e.g. University of Woolloomooloo [SAB, HCL]) |
| Department of Applied Sciences, University of the West of England, Bristol [DT]University of Guelph, Canada [AMK]Quadram Institute Bioscience, UK [EMA]Laboratory of Gene Technology, KU Leuven, Belgium [RL]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA [IT]Gut Phageomics Lab, Biosciences Institute, University College Cork, Cork, Ireland [CB]Department of Microbiology, Universidade Federal de Viçosa, Viçosa, Brazil [PA-Z]Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia [MVV, BIV, KYN, TAY, TNV] |

 |
| **Corresponding author** |
| Dann Turner |
| **List the ICTV study group(s) that have seen this proposal:** |
| A list of study groups and contacts is provided at <http://www.ictvonline.org/subcommittees.asp> . If in doubt, contact the appropriate subcommittee chair (there are six virus subcommittees: animal DNA and retroviruses, animal ssRNA-, animal ssRNA+, fungal and protist, plant, bacterial and archaeal) | **ICTV Bacterial and Archaeal Viruses Subcommittee*****Caudovirales* Study Group** |
|

|  |
| --- |
| **Authority to use the name of a living person:**Please provide the following information if you propose taxon name(s) which are derived from the name of a living person or persons*.* Please attach documents to verify that permission has been obtained.  |
| **Taxon name** | **Person from whom the name is derived** | **Permission obtained (Y/N)** |
| *Studiervirinae* | F. William Studier | Y |
| *Molineuxvirinae* | Ian J Molineux | Y |
| *Colwellvirinae* | Rita R. Colwell | Y |
| *Chatterjeevirus* | Smriti Narayan Chatterjee | Y |
|  |

**ICTV Study Group comments (if any) and response of the proposer:** |
|       |
|  |
| Date first submitted to ICTV: |       |
| Date of this revision (if different to above): |       |

|  |
| --- |
| **ICTV-EC comments and response of the proposer:** |
|       |

**Part 2:** **NON-STANDARD**

Template for any proposal regarding ICTV procedures, rules or policy, not involving the creation of new taxonomy.

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

**Part 3:** **PROPOSED TAXONOMY**

|  |
| --- |
| **Name of accompanying Excel module:** 2019.103B.A.v1.Autographiviridae\_1fam9subfam132gen.xlsx |

The taxonomic changes you are proposing should be presented on an accompanying Excel module, 2019\_TP\_Template\_Excel\_module. Please enter the file name of the completed module in this box.

**Supporting material:**

| additional material in support of this proposal |
| --- |
| Please explain the reasons for the taxonomic changes you are proposing and provide evidence to support them. The following information should be provided, where relevant:* **Species demarcation criteria**: Explain how new species differ from others in the genus and demonstrate that these differences meet the criteria previously established for demarcating between species. If no criteriahave previously been established, and if there will now be more than one species in the genus, please state the demarcation criteria you are proposing.
* **Higher taxa**:
	+ There is no formal requirement to state demarcation criteria when proposing new genera or other higher taxa. However, a similar concept should apply in pursuit of a rational and consistent virus taxonomy.
	+ Please indicate the **origin of names** assigned to new taxa at genus level and above.
	+ For each new genus a **type species** must be designated to represent it. Please explain your choice.
* **Supporting evidence**: The use of Figures and Tables is strongly recommended (note that copying from publications will require permission from the copyright holder). For phylogenetic analysis, please provide a tree where branch length is **proportional to genetic** distance, generated using an appropriate algorithm (Neighbour-Joining, Maximum Likelihood, or Bayesian) and provide evidence of the reliability of the branching (e.g., by bootstrapping).

Please refer to the Help Notes file (Taxonomic\_Proposals\_Help\_2019) for more information.**Species demarcation criteria:** We have chosen 95% DNA sequence identity as the criterion for demarcation of species in this new genus. Each of the proposed species differs from the others with more than 5% at the DNA level as confirmed with the BLASTN algorithm. **History:**In 2008 the International Committee on Viruses Taxonomy (ICTV) Virus Taxonomy (Carstens & Ball, 2009) defined the family *Podoviridae* as consisting of four genera, one of which was called the “T7-like viruses.” The T7-like viruses contained three species; Enterobacteria phage T7, *Kluyvera* phage Kvp1, and *Pseudomonas* phage gh-1. That same year Lavigne et al., (2008) re-examined the taxonomy of the *Podoviridae* and based upon shared protein homologs and defined three genera of T7-like phages within a single subfamily, the *Autographivirinae*. The defining characteristic of the subfamily was the presence of a virion-encoded RNA polymerase from which the subfamily derived its name; “Auto” and “Graphein” derived from the Greek meaning “self-writing” or “self-transcribing”. In the 2018b ICTV taxonomic release, the *Autographivirinae* is comprised of ten genera; *Drulisvirus* [2015.007a-dB]*, Friunavirus* [2016.013a-dB]*, Napahaivirus* [2018.096B]*, Phikmvvirus* [2007.110-113B]*, Phimunavirus* [2018.110B], *Pollyceevirus* [2018.086B]*, Pradovirus* [2016.079a-dB]*, Przondovirus* [2016.023a-dB]*, Tseptimavirus* [2008.020-023B] and *Zindervirus* [2007.115-119B], that collectively encompass a total of 54 species. A further three species are included within the subfamily, but which lack an assigned genus, *Prochlorococcus* virus PSSP7 [2008.066B] and *Synechococcus* viruses P60 [2008.053B] and Syn5 [2008.070B]The defining morphological characteristics of all these viruses are that they possess a small (ca. 60 nm in diameter) icosahedral head attached to a short tail. The genomes are comprised of linear terminally redundant dsDNA of approximately 41 kb and all encode a large (>100 kDa) single subunit RNA polymerase which is responsible for middle and late transcription. Further common characteristics of these phages included nucleotide similarity, conservation of gene arrangement and apparently genus-specific lysis cassettes and RNAP specificity loops.**New higher taxa and naming origins:** **Family:*****Autographiviridae*:** named after the original subfamily proposed by Lavigne et al., (2009). The name is derived from the Greek “Auto” and “Graphein” derived from the Greek meaning “self-writing” or “self-transcribing” and denotes that all these bacterial viruses encode a large single subunit RNA polymerase.**Subfamilies:*****Okabevirinae:*** The name derives from the Japanese microbiologist Norio Okabe (d. 2014) who studied bacteriophages with activity against a variety of plant pathogenic bacteria at the Faculty of Agriculture, Shizuoka University, Iwata, Japan.***Studiervirinae*:** named in honour of F. William Studier, (Emeritus Professor, Biology Department, Brookhaven National Laboratory, USA) for developing the techniques for studying the intracellular development of bacteriophages through denaturing polyacrylamide gel electrophoresis and autoradiography. His achievements have been recognized by election to the American Academy of Arts and Sciences in 1990 and the National Academy of Sciences in 1992. Studier has worked extensively on the biology of *Escherichia* virus T7 since the mid-1960s.***Krylovirinae:*** named in honour of Victor Krylov for his extensive work on the genetics of *Pseudomonas* and bacteriophages infecting this genus.***Beijerinckvirinae*:** named in honour of Martinus Willem Beijerinck (1851 – 1931) the Dutch microbiologist and botanist who first isolated *Acinetobacter* from soil (Beijerinck. 1911).***Molineuxvirinae*:** named in honour ofIan J Molineux (Professor, Molecular Biosciences, The University of Texas at Austin). His major research interest is understanding how nucleic acids pass through lipid bilayers. The model system used is bacteriophage T7, which uses three different motor proteins to transport its DNA into the cell at the initiation of infection. Professor Molineux has isolated many T7-like viruses.***Corkvirinae*:** Named after the city in the Republic of Ireland. ***Slopekvirinae*:**Named after Stefan Slopek (1914 – 1995) who published a series of papers on the effectiveness of bacteriophage) s against several bacterial pathogens including *Klebsiella* spp. Professor Slopek held the position of the head of Faculty and Department of Clinical Microbiology of the Wroclaw School of Medicine.***Melnykvirinae:*** The name derives from the name of the Russian scientist Moisei Mel’nyk who performed early work in phage therapy for the prophylaxis and treatment of dysentery at the Kharkov Mechnikov Institute (Myelnikov, 2018).***Colwellvirinae:*** This subfamily is named in honour of Dr. Rita R. Colwell (b.1934) who is an American environmental microbiologist and scientific administrator, having been from 1998 to 2004 the Director of the National Science Foundation. She isolated and characterized numerous marine phages, including those against *Vibrio cholerae*.**Genera:**The names of existing genera have not been altered but an explanation of the origins of their names are provided.***Wuhanvirus:*** Named after the capital city of Hubei province in China where the type isolate of this genus, *Pasteurella* phage PHB02, was discovered.***Maculvirus:*** The name derives from thecommune of Macul in Chile where the type isolate of this genus was discovered.***Minipunavirus:*** The name derives directly from the first isolate of this type, *Morganella* phage MmP1.***Bifseptimavirus:*** The name derives directly from the first isolate of this type, *Pseudomonas* phage Bf7.***Poseidonvirus:*** The name derives from Poseidon, the Greek god of the sea.***Tritonvirus:*** the name derives from Triton, the fish-tailed son of Poseidon in Greek mythology.***Risjevirus:*** The name derives directly from the first isolate of this type, *Ralstonia* phage RSJ2.***Kotilavirus:*** The name derives from J. Kotila who, alongside G. Coons, performed early studies in the mid-1920s of bacteriophages of plant pathogens*.****Gajwadongvirus:*** The name derives from the Gajwa-dong parish in the province of Gyeongsangnam of South Korea where the type isolate of this genus *Escherichia* phage ECB5, was isolated.***Eracentumvirus:*** The name derives directly from the first isolate of this type, *Erwinia* phage Era103.***Novosibovirus:*** The name derives from Novosibirsk where the type isolate of this genus, *Proteus* phage PM16, was discovered.***Ahpunavirus:*** The name derives directly from the first isolate of this type, *Aeromonas* phage Ahp1.***Ermolevavirus*:** The name derives from the Russian microbiologist Zinaida Ermol’eva (1898 – 1974) who established a bacteriophage laboratory at the All-Union Institute of Experimental Medicine in Moscow. ***Pokrovskaiavirus:*** The name recognises the work of Magdalina Pokrovskaia (b. 1901) for her early work on bacteriophages of *Yersinia pestis* (Myelnikov, 2018).***Atuphdovirus:*** The name derives directly from the first isolate of this type, *Agrobacterium* phage Atu\_ph02.***Gyeongsanvirus:*** The name derives from the city Gyeongsan in South Korea where the first isolate of this type, *Ralstonia* phage DU\_RP\_I, was isolated.***Ampunavirus:*** The name derives directly from the first isolate of this type, *Burkholderia* phage Bp-AMP1.***Axomammavirus*:** Named after Axomamma, a goddess of potatoes in Inca mythology as the host bacterial genus *Pectobacterium* is a pectinolytic plant pathogen causing blackleg, stem rot and soft rot in a variety of plant hosts including potato.***Sednavirus:*** The name derives from Sedna, goddess of the sea and its creatures in Inuit mythology.***Tangaroavirus:*** The name derives from Tangaroa, the god of the sea in Maori mythology.***Kaohsiungvirus:*** The name derives from the city Kaohsiung in Taiwan where the first isolate of this type, *Vibrio* phage phi-A318, was isolated.***Murciavirus:*** The name derives from the city of Murcia where the type isolate of this genus, *Marinomonas* phage CPP1m, was discovered.***Lauvirus*:** The name derives directly from the first isolate of this type, *Pseudomonas* phage Lau218.***Phimunavirus:*** The name derives directly from the first isolate of this type, *Pectobacterium* phage phiM1. ***Pektosvirus:*** The name derives from the Greek adjective ‘pektos’ from which the host bacterial genus *Pectobacterium* derives its name.***Teseptimavirus:*** The name derives directly from the first isolate of this type, *Escherichia* phage T7.***Teetrevirus:*** The name derives directly from the first isolate of this type, *Escherichia* phage T3.***Pifdecavirus:*** The name derives directly from the first isolate of this type, *Pseudomonas* phage pf-10.***Acadevirus:*** The name derives from Academgorodok, a district of the city of Novosibirsk in Russia where the type isolate, *Proteus* phage PM85 was discovered.***Zindervirus:*** The name recognises the work of Norton David Zinder who discovered the process of bacteriophage-mediated transduction in *Salmonella* [2018.007B]***Vectrivirus:*** The name derives directly from the first isolate of this type, *Escherichia* phage Vec3.***Cuernavacavirus:*** The name derives from the city of Cuernavaca in Mexico where the first isolate of this type, *Rhizobium* phage RHEph02, was studied.***Drulisvirus:*** The name recognises the work of Dr Zuzanna Drulis-Kawa [2018.007B]***Kayfunavirus:*** The name derives directly from the first isolate of this type, *Escherichia* phage K1F.***Berlinvirus:*** The name derives directly from the first isolate of this type, *Yersinia* phage Berlin.***Phikmvvirus:*** The name derives directly from the first isolate of this type, *Pseudomonas* phage phiKMV.***Friunavirus:*** The name derives directly from the first isolate of this type, *Acinetobacter* phage Fri1***Przondovirus:*** The name recognises the work of Przondo-Hessek who performed early work on bacteriophages of *Klebsiella* species [2018.007B].***Helsettvirus:*** The name derives from the Northern Sami name for the city of Helsinki in Finland, where the bacteriophages constituting this genus were studied.***Chatterjeevirus:*** This genus is named in honour of Senior Professor Smriti Narayan Chatterjee (b. 1932) who is the former senior Professor & Director (Offg), Saha Institute of Nuclear Physics (Kolkata) and former Senior Scientist (Emeritus) of the Indian National Science Academy. He is also the founding President of the DNA Society of India. Being the recipient of many awards, he carried out a number of early studies on the biophysics of Vibrio phages and showed that the phages belonging to any serological group had distinct morphology.***Aqualcavirus:*** The name is derived from the host genera that this phage infects i.e. ***Aqua****microbium* + ***Alca****ligenaceae* [2018.046B].***Napahaivirus:*** The name derives from the place (Napahai wetland, Kunming, Yunnan, China) where the first isolate of this type, *Pseudomonas* phage VSW-3, was isolated [2018.096B].***Pollyceevirus:*** The name derives directly from the first isolate of this type, *Pseudomonas* phage PollyC [2018.086B].***Aerosvirus:*** The name derives from the Greek “aeros” meaning “air” or “gas” from which the host bacterial genus *Aeromonas* derives its name.***Cronosvirus****:* Named after the Titan of Greek mythology and the root origin of the genus name *Cronobacter* infected by both bacteriophages comprising the proposed genus.***Uliginvirus:*** The name derives from the Greek “uligin” meaning “in marshes” after the first isolate of this type, *Pseudomonas* phage Uligo.***Bonnellvirus:*** The name derives from one of the highest points in Austin, Texas, in recognition that the type isolate of this genus, *Escherichia* phage J8-65, was studied at the University of Austin.***Pelagivirus:*** The name stems from the Latin noun pelagus “sea” from which the host bacterial genus Pelagibacter derives its name.***Higashivirus:*** Named after the Japanese word ‘Higashi’ meaning “east” in recognition that the type isolate of this genus, *Ralstonia* phage RSB1 was isolated at Hiroshima University located in Higashi-Hiroshima district, Japan.***Aarhusvirus:*** Named after the city in Denmark where the first isolate of this type, *Dickeya* phage Dagda, was isolated.***Ningirsuvirus:*** Named after the Sumerian deity also known as Ninurta, after the first isolate of this type, *Dickeya* phage Ninurta.***Wanjuvirus* –** named after the county located within North Jeolla Province in South Korea, where the first isolate of this type was studied.***Aegirvirus:*** The name derives from Aegir (or Ægir), the god of the sea in Norse mythology.***Tiamatvirus:*** The name derives from Tiamat, a goddess of the salt sea in Mesopotamian mythology.***Lirvirus:*** The name derives from the Old Irish “Lir” meaning “Sea” and is the god of the sea in Gaelic mythology.A number of new genera comprised of single species were named using words for “foot” or “stump” plus “virus.” These included: Albanian – Kembe (***Kembevirus***), Trung (***Trungvirus***); Amharic – Igiri (***Igirivirus***), Guto (***Gutovirus***); Arabic – Qadam (***Qadamvirus***); Armenian – Votkov (***Votkovvirus***), Aghby (***Aghbyvirus***); Azerbaijani – Ayaq (***Ayaqvirus***); Basque – Oinez (***Oinezvirus***); Belarusian – Nohi (***Nohivirus***), Pien (***Pienvirus***); Bosnian – Stopala (***Stopalavirus***), Panj (***Panjvirus***); Bulgarian – Krak (***Krakvirus***); Cebuano – Tiil (***Tiilvirus***), Tuod (***Tuodvirus***); Dutch – Voet (***Voetvirus***), Stomp (***Stompvirus***); Finnish – Jalka (***Jalkavirus***), Kanto (***Kantovirus***); French – Pied (***Piedvirus***); Frisian – Foet (***Foetvirus***); Georgian – Pekhit (***Pekhitvirus***); German – Fuss (***Fussvirus***), Stumpf (***Stumpfvirus***); Greek – Podi (***Podivirus***), Koutsouro (***Koutsourovirus***); Gugarati – Paga (***Pagavirus***), Stampa (***Stampavirus***); Hausa – Kafa (***Kafavirus***); Hindi – Pair (***Pairvirus***); Hmong - Taw; Icelandic - Fotur, Stubbur; Indonesian - Kaki, Tunggul; Irish - Chos; Japanese - Ashi, Kirikabu; Kannada (***Kannadavirus***) – Paada (***Paadavirus***); Khmer – Cheung (***Cheungvirus***), Daem (***Daemvirus***); Kurdish – Ling (***Lingvirus***), Serkor (***Serkorvirus***); Lithuanian – Pedos (***Pedosvirus***), Kelmas (***Kelmasvirus***); Luxembourgish - Fouss (***Foussvirus***), Stompel (***Stompelvirus***); Malayalam - Kalppathi (***Kalppathivirus***); Maltese – Sieq (***Sieqvirus***); Maori – Waewae (***Waewaevirus***); Mardarin – Laoyazi (***Laoyazivirus***); Polish – Stopa (***Stopavirus***); Russian – Stupnya (***Stupnyavirus***); Sudanese – Suku (***Sukuvirus***); Swahili – Mguu (***Mguuvirus***); Tajic – Pijola (***Pijolavirus***); Tamil – Paadam (***Paadamvirus***); Telugu – Phut (***Phutvirus***); Turkish – Ayak (***Ayakvirus***); Urdu – Pow (***Powvirus***); Vietnamese – Banchan (***Banchanvirus***); Welsh – Troed (***Troedvirus***); Xhosa – Unyawo (***Unyawovirus***)**Supporting evidence:** Recent network and phylogenetic analyses have shown that the families *Podoviridae, Siphoviridae* and *Myoviridae* are polyphyletic (Iranzo et al., 2016; Aiewsakun et al., 2018; Low et al., 2019; Barylski et al., 2019). The phages related to the *Autographivirinae* represent a distinct cluster within the dsDNA viruses (Iranzo et al., 2016) and a number of studies have highlighted issues with the current ICTV taxonomic designations where certain genera are not robustly monophyletic (Low et al., 2019; Jang et al., 2019). Since the inception of the subfamily, a substantial number of Autographivirinae-like phages have been deposited in the International Nucleotide Sequence Database collaboration which lack appropriate taxonomic designations, necessitating a systematic examination of these genomes. Genome sequences of bacterial viruses related to the *Autographivirinae* were retrieved from Genbank using both 2018 ICTV and NCBI taxonomic bins and by searches of the literature. The identification of candidate *Autographivirinae*-like genomes was performed using hidden markov models of the major capsid protein, TBLASTX and SymBets BLASTP (Kristensen et al., 2010) approaches implement at NCBI. A total of 471 candidate phage genomes were investigated using comparative genomics, proteomics and marker gene phylogenies.Genome level comparisons were performed using the Gegenees fragmented aligner in BLASTN and TBLASTX modes (Ågren et al., 2012), average nucleotide identities were calculated using the ANIb method in pyANI and by calculation of tBLASTx DICE scores. Predicted proteome comparisons were performed with Roary (Page et al., 2015) using submitted annotations and GET\_HOMOLOGUES (Contreras-Moreira & Vinuesa, 2013). All genome and proteome comparisons were hierarchically clustered by the complete-linkage method using the hclust package in R (R Core Team, 2013). Maximum-likelihood phylogenetic trees were inferred for marker genes after alignment with Clustal Omega using IQTree with SH-aLRT test and ultrafast bootstrap with 1000 replicates using the best-fit model automatically selected by ModelFinder (Trifinopoulos et al., 2016; Kalyaanamoorthy et al., 2017; Hoang et al., 2018). Lastly, the 471 genomes were analysed using the network-based virus classification tool vConTACT v2.0 with the following settings: Protein-protein similarity method (diamond); Reference database (Refseq database v88 with ICTV taxonomy); PC generation method (MCL); VC generation method (ClusterONE); Inflation value for PCs (2); Minimum Density for VCs (0.3); Minimum VC size (2); Max overlap for VCs (0.8); Penalty for VC (2); Haircut for VC (0.65); Merge method for VCs (single); Similarity method during VC merging (match); Seed method (ClusterONE only; unused\_nodes).The different methods used to assess this new family provided results that were predominantly concordant and fit well with the proposed taxonomy. One inconsistency was that two members of the genus Vectrevirus were mixed with the Zindervirus clade for trees constructed using the tail tubular protein B and major capsid protein. For trees constructed using the DNA polymerase, the proposed *Tspetimavirus* and *Teetrevirus* clades were combined but clustered separately for all other analyses. The existing genus-level classifications *Drulisvirus, Friunavirus, Napahaivirus, Phikmvvirus, Pollyceevirus, Pradovirus, Przondovirus, Tseptimavirus* and *Zindervirus* are supported, with some specific exceptions. *Pseudomonas* virus gh1 is removed from the *Tseptimavirus* and *Pantoea* phages Limelight and Limezero are removed from the *Phikmvvirus*. *Pseudomonas* phage LKA1 is designated as an outlier of the *Phikmvvirus* and is removed from the genus. These alterations are congruent with observations made in the literature that these genera were not robustly monophyletic (Iranzo et al., 2016; Bolduc et al., 2017; Jang et al., 2019, Low et al., 2019). The viral clusters defined by vConTACT v2.0 exhibited some differences between the genera proposed here, suggesting more inclusive viral clades. For example, the T7- and T3-like phages and the gh-1 and Pf10-like phages were combined into single VCs. In the absence of a whole genome genus criterion, at this time we have chosen to split these clades into genera using a threshold of >60% nucleotide identity and the presence of core gene sets. Based on the results of this study, we propose that the *Autographvirinae* and Autographivirinae-like viruses be removed from the family *Podoviridae* and be assigned a family rank, “Autographiviridae” that encompasses 9 subfamilies and 132 genera based upon the genome, proteome and marker-gene based analyses. Of the 471 genomes, 95 exhibit >95% nucleotide sequence identity to a species and are thus designated as strains. Forty-eight phage genomes assigned to 21 genera are left unassigned at the subfamily level. It is likely that these unassigned viruses represent clades at both genus and subfamily ranks that are currently under-sampled. For this reason, we have chosen not to define genera based on single virus isolates within the family at this time. **Supporting Evidence:**Due to the number of phage genomes analysed, supporting evidence is provided as attachments.**File 1: MCP\_tree.pdf** Maximum likelihood tree based on Clustal Omega alignment of the major capsid protein using IQTree. Branch support values were calculated from 1000 ultrafast bootstrap replicates. The scale bar represents the number of substitutions per site. The tree is rooted at the major capsid protein of *Pseudomonas* phage LUZ7 [CAZ66220] Branches corresponding to proposed genera are delineated using coloured blocks, while suggested subfamilies are delineated by grey boxes with dashed outer lines.**File 2: RNAP\_tree.pdf** Maximum likelihood tree based on Clustal Omega alignment of the DNA-dependent RNA polymerase protein using IQTree. Branch support values were calculated from 1000 ultrafast bootstrap replicates. The tree is rooted using the RNA polymerase of *Shewanella* phage Spp001 [AHJ10567]. The scale bar represents the number of substitutions per site. Branches corresponding to proposed genera are delineated using coloured blocks, while suggested subfamilies are delineated by grey boxes with dashed outlines.­**File 3: DICE\_hmap.pdf** Heatmap representing TBLASTX DICE distance scores. The suggested subfamilies are delineated by grey rectangles with dashed outlines.**File 4: DICE\_tree.pdf** TBLASTX DICE score tree. The scale bar represents the calculated distance metric. Coloured blocks and grey boxes with dashed outlines respectively delineate the proposed genera and subfamilies. |
|  |
| **References:** |
| Carstens EB, Ball LA, (2009) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2008). Arch. Virol., 154(7):1181-1188.Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM (2008). Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. Res. Microbiol. 159(5):406-414.Adriaenssens EM, Wittmann J, Kuhn JH, Turner D, Sullivan MB, Dutilh BE,, Jang HB, van Zyl LJ, Klumpp J, Lobocka M, Moreno Switt AI, Rumnieks J, Edwards RA, Uchiyama J, Alfenas-Zerbini P, Petty NK, Kropinski AM, Barylski J, Gillis A, Clokie MRC, Prangishvili D, Lavigne R, Aziz RK, Duffy S, Krupovic M, Poranen MM, Knezevic P, Enault F, Tong Y, Oksanen HM, Rodney Brister J (2018). Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. Arch Virol. 163(4):1125-1129. Beijerinck MW, (1911) Über Pigmentbildung bei Essigbakterien. Centr. Bakteriol. Parasitenk. Abt. 167-76Myelnikov D, (2018) An Alternative Cure: The Adoption and Survival of Bacteriophage Therapy in the USSR, 1922–1955. J Hist Med Allied Sci. 73(4):385-411Iranzo J, Krupovic M, Koonin E V. The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. MBio. 2016;7: e00978-16Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P (2018). Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. J. Gen. Virol. 99: 1331-1343Barylski J, Enault F, Dutilh BE, Schuller MBP, Edwards RA, Gillis A, Klumpp J, Knezevic P, Krupovic M, Kuhn JH, Lavigne R, Oksanen HM, Sullivan MB, Jang HB, Simmonds P, Aiewsakun P, Wittmann J, Tolstoy I, Brister JR, Kropinski AM, Adriaenssens EM (2019). Analysis of Spounaviruses as a Case Study for the Overdue Reclassification of Tailed Phages. Syst. Biol. doi: 10.1093/sysbio/syz036Jang HB, Bolduc B, Zablocki O, Khun JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat. Biotech. 37:632-639Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, Koonin EV, Mushegian A (2010). A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics (Oxford, England), 26(12), 1481–1487. doi:10.1093/bioinformatics/btq229Ågren J, Sundström A, Håfström T, Segerman B (2012). Gegenees: Fragmented Alignment of Multiple Genomes for Determining Phylogenomic Distances and Genetic Signatures Unique for Specified Target Groups. PLoS ONE 7(6): e39107.Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J, "Roary: Rapid large-scale prokaryote pan genome analysis", Bioinformatics, 2015;31(22):3691-3693Contreras-Moreira B, and P Vinuesa (2013) GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl.Environ.Microbiol. 79:7696-7701.R Core Team (2013) R: A Language and Environment for Statistical Computing. http://www.R-project.org R Foundation for Statistical Computing, Vienna, Austria, ISBN3-900051-07-0Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res., 44:W232-W235.Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods, 14:587-589. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol., 35:518–522. Bolduc B, Jang HB, Doulcier G, You Z-Q, Roux S, Sullivan MB (2017). vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. Peer J 5:e3243Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. (2019). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res.47(D1):D23-D28. Tolstoy I, Kropinski AM, Brister JR. Bacteriophage Taxonomy: An Evolving Discipline. Methods Mol Biol. 2018;1693:57-71. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res.44(D1):D733-45. |