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:** | ***2019.076B*** | |  |
| **Short title:** Create three new genera including four new species in the family *Podoviridae* | | | |
|  | | | |
| **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 | |
| Kropinski AM, Adriaenssens EM | | [Phage.Canada@gmail.com](mailto:Phage.Canada@gmail.com);  evelien.adriaenssens@quadram.ac.uk | |
| **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]) | | University of Guelph, Canada [AMK]  Quadram Institute Bioscience, UK [EMA] | | | | |
| **Corresponding author** | | | |
| Andrew M. Kropinski | | | |
| **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) | | **Bacterial and Archaeal Viruses Subcommittee**  **Caudovirales Study Group** | |
| **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:** 2019.076B.N.v1.Podoviridae\_4gen.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. |

| **References:** |
| --- |
| 1: Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2019;47(D1):D23-D28.  2: Tolstoy I, Kropinski AM, Brister JR. Bacteriophage Taxonomy: An Evolving Discipline. Methods Mol Biol. 2018;1693:57-71.  3: 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.  4: Agren J, Sundström A, Håfström T, Segerman B. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One. 2012;7(6):e39107.  5: Chan PP, Lowe TM. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol Biol. 2019;1962:1-14.  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.  7: Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010;5(6):e11147.  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.  9: Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol. 2006;55(4):539-52. |

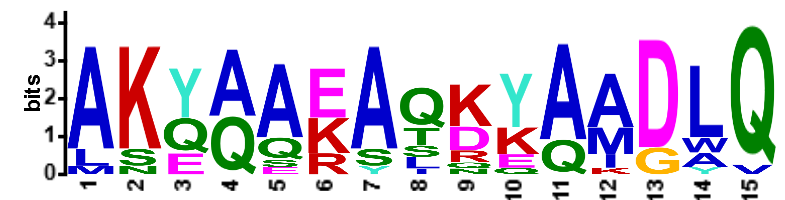
**Introduction:** We have generally assumed that the morphological classification provided in GenBank is correct. In the absence of a full taxonomy we have resorted to examination of the genome for diagnostic proteins. These include “tail sheath protein” and “base plate protein” homologs which identify myoviruses; long “tail tapemeasure proteins”, and lack of the above two proteins, which are diagnostic of the siphoviruses; and, lack of all three in the case of the podoviruses. We wanted a positive selection and have investigates P22 GP7 and GP20 homologs. These two proteins are found in the phage head and contribute to DNA injection.

MEME analysis (<http://alternate.meme-suite.org>) of phage GP7 and GP20 homologs revealed that the phages in this proposal each possess high scoring motifs which are not found in T7 proteins or those of myoviruses or siphoviruses.

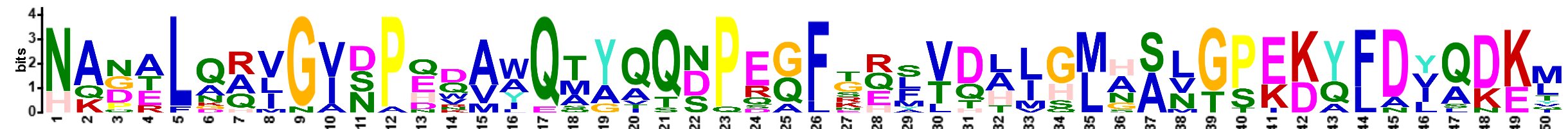
**GP7 - Motif 1**



**GP7 – Motif 6**



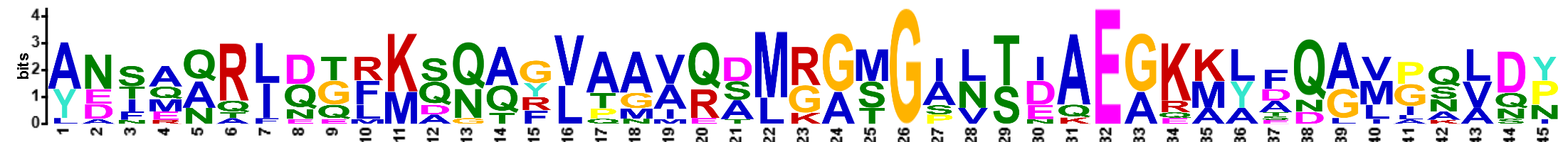
**GP20 – Motif 1**

****

**GP20 – Motif 2**

****

**GP20 – Motif 4**

****

**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.

**Recognition of new species and genera:** This was based upon phage genomes showing ≈70 - ≤95% DNA sequence identity as revealed by BLASTN analysis at NCBI. The phages described below belong to a higher taxon, mostly likely a subfamily, but we do not choose to propose one at this time.

**Phylogeny:** The phylogenetic tree was constructed using the large subunit terminase protein homologs of this group of phages with phylogeny.fr in “one click” mode [8]. "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." It also includes the use of Gblocks to eliminate poorly aligned positions and divergent regions. "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 [9] for details."



**Recognition of new species and genera:** This was based upon phage genomes showing ≈70 - ≤95% DNA sequence identity as revealed by BLASTN analysis at NCBI.

**Proposal 1:** To create a new genus, *Skarprettervirus*, consisting of a single species in the family *Podoviridae*.

**Source of the name of this taxon:** The name is directly derived from the name of the first virus of its type, Escherichia phage Skarpretter.

**History:** Escherichia phage Skarpretter was isolated from wastewater in Denmark using Escherichia coli MG1655 K12 as the host bacterium. In GenBank it is listed as a member of the *Siphoviridae*. No publications.

**GenBank Summary:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Phage name | RefSeq No. | INSDC | Size (Kb) | GC% | Protein | tRNA |
| Escherichia phage Skarpretter |  | [MK105855.1](https://www.ncbi.nlm.nih.gov/nuccore/MK105855.1) | 42.04 | 55.8 | 63 | 0 |

**BLASTN homologs:** The next most closely related phage is Escherichia phage C130\_2 which displays 40.7% DNA sequence identity to phage Skarpretter.

**Electron micrograph:** None available

**Proposal 2:** To create a new genus, *Giessenvirus*, consisting of a single species in the family *Podoviridae*.

**Source of the name of this taxon:** The name is directly derived from the name of the city in Germany where the first virus of its type, Escherichia phage C130\_2, was isolated.

**History:** Escherichia phage C130\_2 was isolated from cheese in Germany and lyses both Escherichia coli and Shigella strains. The publication describes it as a member of the *Myoviridae*.

**Publication:**

Sváb D, Falgenhauer L, Rohde M, Chakraborty T, Tóth I. Complete genome sequence of C130\_2, a novel myovirus infecting pathogenic Escherichia coli and Shigella strains. Arch Virol. 2019;164(1):321-324.

**GenBank Summary:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Phage name | RefSeq No. | INSDC | Size (Kb) | GC% | Protein | tRNA |
| Escherichia phage C130\_2 |  | [MH363708.1](https://www.ncbi.nlm.nih.gov/nuccore/MH363708.1) | 41.78 | 55.4 | 59 | 0 |

**BLASTN homologs:** The next most closely related phage is Escherichia phage Skarpretter which displays 41.4% DNA sequence identity to phage C130\_2.

**Electron micrograph:** None available

**Proposal 3:** To create a new genus, *Sortsnevirus*, consisting of two species in the family *Podoviridae*.

**Source of the name of this taxon:** The name is directly derived from the name of the first virus of its type, Escherichia phage Sortsne.

**History:** Escherichia phage Sortsne was isolated from wastewater in Denmark using Escherichia coli K-12 MG1655 as the host.

**Publication:** None

**GenBank Summary:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Phage name | RefSeq No. | INSDC | Size (Kb) | GC% | Protein | tRNA | Overall DNA sequence identity (\*\*) | % common proteins (\*\*) |
| Escherichia phage Sortsne |  | [MK651787.1](https://www.ncbi.nlm.nih.gov/nuccore/MK651787.1) | 41.91 | 60.0 | 62 | 0 | 100% | 100% |
| Klebsiella phage vB\_KpnS\_IME279 |  | [MF614100.1](https://www.ncbi.nlm.nih.gov/nuccore/MF614100.1) | 42.52 | 59.3 | 59 | 0 | 69.1 | 82.3 |

****

**BLASTN homologs:** The next most closely related phage is Enterobacteria phage IME\_EC2 which displays 61.9% DNA sequence identity to phage Sortsne.