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:** | ***2019.020S*** | |  |
| **Short title:** Create one new genus (*Nuarterivirus)*;move the existing subgenus *Pedartevirus* to the genus *Iotaarterivirus;* rename one species from the subgenus *Pedartevirus*; create one new species in the new genus *Nuarterivirus*;create one new subgenus and two new species in the existing genus *Betaarterivirus* | | | |
|  | | | |
| **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 | |
| Brinton MA, Gulyaeva A, Balasuriya UBR, Dunowska M, Faaberg KS, Goldberg T, Leung FC-C, Nauwynck HJ, Snijder EJ, Stadejek T, Gorbalenya AE | | [mbrinton@gsu.edu](mailto:mbrinton@gsu.edu)  [A.Gulyaeva@lumc.nl](mailto:A.Gulyaeva@lumc.nl)  [ubalasuriya@uky.edu](mailto:ubalasuriya@uky.edu)  M.Dunowska@massey.ac.nz kay.faaberg@ars.usda.gov  tony.goldberg@wisc.edu  fcleung@hku.hk  [hans.nauwynck@ugent.be](mailto:hans.nauwynck@ugent.be)  E.J.Snijder@lumc.nl [tomasz.stadejek@outlook.com](mailto:tomasz.stadejek@outlook.com)  [A.E.Gorbalenya@lumc.nl](mailto:A.E.Gorbalenya@lumc.nl) | |
| **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]) | |  | | | | |
| **Corresponding author** | | | |
| Margo A. Brinton, [mbrinton@gsu.edu](mailto:mbrinton@gsu.edu)  Alexander E. Gorbalenya, [A.E.Gorbalenya@lumc.nl](mailto:A.E.Gorbalenya@lumc.nl) | | | |
| **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) | | This proposal is filed by the **Arterivirus Study Group** in consultation with:  *Nidovirales* Study Group  *Coronaviridae* Study Group | |
| **ICTV Study Group comments (if any) and response of the proposer:** | | | |
|  | | | |
|  | | | |
| Date first submitted to ICTV: | | | 1 June 2019 |
| Date of this revision (if different to above): | | | 15 October 2019 |

|  |
| --- |
| **ICTV-EC comments and response of the proposer:** |
| **Nick Knowles**:  Most things were approved, but there were a few issues that were picked up:   * 2019.020S-023S.N.v1.Nidovirales   + First line (species *Deltaarterivirus pejah*) - says move species (and also rename in comments), but it should be two different entries - a move of the subgenus (from one genus to another) and a rename only of the species (it stays within the subgenus *Pedartevirus*)   + *Shingleback nidovirus 1* - says move species, but it is a move of the subgenus   The other major thing was the use of partial (non-full-coding) genome sequences being used to propose new species.    Hebius snake nidovirus 1 (MG600021) 12.5 Kb  Murina bat coronavirus JTAC2 (KU182966) 25.7 Kb  Tropidophorus coronavirus 118981 (MG600026) 22.4 Kb  Sectovirus 2 (MG600031) 26 Kb    I realise that you might consider that each of these has enough data for classification, but the committee was adamant that (from this point on) new species should have at least complete coding sequences. These four species proposals need to be removed from the proposals (we had the same problem for a few picornavirus sequences).    Please can you let me have your updated proposals as soon as possible?  **All SGs concerned with nidoviruses**:   * The excel sheet with the *Nidovirales* taxonomy has been updated as requested. * The four nidovirus TPs submitted on June 1st, 2019 (including this TP) have been updated to indicate that four nidoviruses with incomplete coding sequences are not part of the proposed taxonomy. |

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

|  |
| --- |
| **Name of accompanying Excel module:** 2019.020S-023S.A.v1.Nidovirales.xlxs |

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  This and three accompanying proposals are based on analyses of the genomic diversity of viruses in the *Nidovirales* order and most closely related unclassified viruses performed by A.A. Gulyaeva, D. V. Samborskiy, I.A.Sidorov, and A.E.Gorbalenya (Gulyaeva et al & Gorbalenya, in preparation). The general framework of this analysis and its specific application to a group of viruses included in this particular proposal are separately summarized below.  **Computational Taxonomy Framework**. The analyses of the order *Nidovirales* included all publicly available (>3500; beginning of April 2019) (near) complete genome sequences of nidoviruses and most closely related unclassified viruses from diverse vertebrate and invertebrate hosts; many of these sequences are currently the only source of information about the respective viruses. Sequences were analyzed in the computational comparative genomics framework DEmARC (DivErsity pArtitioning by hieRarchical Clustering) using profiles of multiple sequence alignments (MSA), Bayesian and Maximum-likelihood phylogenetic trees, and the profiles of clustering cost (CC) function that were produced for weighted hierarchical clustering of pairwise patristic distances (PPD); DEmARC 1.41 was used for analysis (software is available upon request to AEG). In profiles of CC function, all local minima (smallest CC values in a range of PPD values) were considered as candidate thresholds for ranks because they satisfied two requirements, (i) the clusters formed under these thresholds were monophyletic in the ML tree of the respective nidovirus subset, and (ii) all intra- and inter-cluster PPDs were (predominantly) smaller and (predominantly) larger, respectively, than the respective threshold. If *all* intra- and inter-cluster PPDs, respectively, were smaller and larger than the respective threshold, such clustering has a cost of zero, CC=0, according to DEmARC. We have also measured persistence of a clustering as a range of PPD values over which this clustering was favored with the support of CC=0. The respective “threshold PPD ranges” were considered best candidates for demarcation. Those thresholds supported independently by several datasets were predominantly used to set demarcation criteria of a rank, as these assignments were less likely to be fortuitous due to biased virus sampling and/or residue selection.  Genome sequences were assigned to the nidovirus taxa using either Haygens tool (<http://veb.lumc.nl/HAYGENS/>) or by the authors who described the viruses. Assignments were verified by alignments and phylogenetic analyses of five replicative protein domains characteristic of nidoviruses (synteny of molecular markers), namely the 3CLpro, NiRAN, RdRp, ZBD and HEL1. As shown in **Fig. 1**, 15 groups of nidovirus lineages, ranging from separate families to the entire order plus outgroup, were analyzed. For each group, the MSAs of concatenated five replicative domains conserved in the *Nidovirales* were generated and used in the phylogenetic and DEmARC analyses. For analyses involving viruses of an outgroup, tentatively called *Protonidovirales*, the MSA included RdRp and HEL1 domains. These viruses also share the 3CLpro domain, which was *not* included in the analysis due to its extreme divergence. Data from these analyses provided support for monophyletic clusters, levels and clusters of classification, agreement between phylogeny and classification for each virus group, and inter-group agreement regarding classification levels.  **Application of the computational taxonomy framework to the *Arteriviridae***. The current taxonomy of the *Arteriviridae* family includes six subfamilies, 12 genera, 10 subgenera, and 20 species, mostly delineated during prior DEmARC analyses of the genetic divergence in this family. Analyses used for this proposal included 333 new genome sequences. They were classified using a dataset that was limited to the *Arteriviridae* family, with WPDV being the most divergent arterivirus (Ar3 group in **Fig. 1**). Location of the 3CLpro, NiRAN, RdRp, ZBD and HEL1 domains used for DEmARC analysis are shown for EAV, as a representative of the family(**Fig. 2**). Among candidate thresholds identified, mostly supported with a CC=0 or close, those preferentially selected were those which were observed in already established taxa. This was the case for all taxa at the species, subgenera, and subfamily ranks. Under the selected thresholds, three new species and one new genus were created for three rat arteriviruses. One new species (*Nuarterivirus guemel*) was assigned to a new genus (*Nuarterivirus*), while two other new species (*Betaarterivirus* *sheoin* and *Betaarterivirus timiclar*) were assigned to the *Betaarterivirus* genus. Likewise, two established species (SHFV and PBJV) of the *Deltaarterivirus* genus were separated to different genera: SHFV was left in the *Deltaarterivirus* and PBJV reassigned to *Iotaarterivirus*. This was done because a monophyletic cluster of SHFV and PBJV, poorly supported in 2017 (SH<70%), was no longer observed in the arterivirus tree in 2019. Rather, SHFV is basal to the MYBV-1 and KKCBV cluster (*Thetaarterivirus* genus), while PBJV is basal to the KRTGV and DeMAV cluster (*Iotaarterivirus* genus) in the current arterivirus tree (**Fig. 3-5**).  ***Demarcation criteria***. We used either a range or a particular value of patristic pairwise distances (PPD) calculated for FastTree 2.1.4 SSE3 ML phylogeny of the MSA of five concatenated domains 3CLpro, NiRAN, RdRp, ZBD and HEL1 as the demarcation criterion to taxa at each of four ranks: subfamily, genus, subgenus, and species (**Table 1**). *They were selected as local minima in the CC distribution, commonly corresponding to the CC=0* (see above).  **Table 1** Demarcation thresholds for four ranks of the *Arteriviridae* family   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **rank** | **Taxa #**  **current new** | **PPD range1** | **PUD (%) range2** | **Dataset used3** | | subfamily | 6 0 | 0.952-1.502 | 0.397-0.497 | Ar3 | | genus4 | 12 1 | 0.578-0.637 | 0.290-0.310 | Ar3 | | subgenus5 | 10 (17) 1 (2) | 0.281-0.361 | 0.165-0.203 | Ar3 | | species | 20 3 | 0.174-0.196 | 0.107-0.120 | Ar3 |   **1**Demarcation threshold depicted as a range of PPD values for which the number of clusters (taxa) remained constant and CC=0. PPD values accounting for repeated replacements of amino acid residues.  **2**Demarcation threshold depicted as a range of PUD, deduced from PPD valuesfor which the number of clusters (taxa) remained constant and the CC=0. PUD values are calculated as the % of different residues in compared proteins.  **3**See Figure 1.  4PBJV was reassigned from the *Deltaarterivirus* to the *Iotaarterivirus* genus. In order to keep the taxonomy stable, the SG decided to demarcate genera at the depicted threshold rather than use another one favored by DEmARC with two extra clusters.  5In brackets, the number of DEmARC-based clusters are shown; some of these clusters were not defined as taxa since the SG decided to recognize the subgenera only for those genera which have two or more species.    **Fig. 1**. Nidovirus phylogeny and subsets used to advance taxonomy. Depicted are phylogenetic trees of nidoviruses representing 113 established and newly proposed species (left) and 14 subsets of nidoviruses, which were analysed to build taxonomy (right). For virus acronyms, see the accompanying spreadsheet; black and green, established and newly delineated species, respectively. The tree was reconstructed by IQ‑Tree 1.5.5 based on MSAs of five domains (3CLpro, NiRAN, RdRp, ZBD and HEL1) with the evolutionary model for each domain selected independently; subsequently the tree was midpoint-rooted. Branch support was estimated using Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) with 1000 replicates and is depicted by shaded circles. DEmARC-based classifications of 14 subsets of nidoviruses and one group including all nidoviruses plus an outgroup including four viruses (protonidoviruses) were analyzed to verify and advance the taxonomy of the order. The taxonomic assignment of four nidoviruses included in this computational analysis, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, is deferred until complete coding sequences of the respective viruses become available.    **Fig. 2**. Domain combination used for phylogenetic and DEmARC analyses of the *Arteriviridae* family. Shown are the locations of the conserved replicative domains (5d, 5 domains) used in the analysis of the Ar3 group (see Fig. 1) and depicted relative to the genome, open reading frames, and replicative domains of Equine arteritis virus (EAV). The results shown in **Figs 3-5** were obtained using the MSA of this domain combination.    **Fig. 3.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted on **Figure 1**) of the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed subgenus structure of the group is detailed on the right.    **Fig. 4.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted on **Figure 1**) of the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed genus structure of the group is detailed on the right.    **Fig. 5.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted on **Figure 1**) of the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed subfamily structure of the group is detailed on the right. |
| ` |

| **References:** |
| --- |
| 1. Brinton MA, et al. (2017) ICTV proposal 2017.012S Expansion of the rank structure of the family Arteriviridae and renaming its taxa.  2. Vanmechelen B, Vergote V, Laenen L, Kuhn JH, & Maes P (2017) ICTV proposal 2017.001S One new subfamily, genus and species in the family Arteriviridae (Nidovirales).  3. Gorbalenya AE, et al. (2017) ICTV proposal 2017.015S Reorganization and expansion of the order Nidovirales at the family and sub-order ranks.  4. Lauber C & Gorbalenya AE (2012) Partitioning the genetic diversity of a virus family: approach and evaluation through a case study of picornaviruses. J. Virol 86(7):3890-3904.  5. Lauber C & Gorbalenya AE (2012) Toward genetics-based virus taxonomy: comparative analysis of a genetics-based classification and the taxonomy of picornaviruses. J Virol 86(7):3905-3915.  6. Wu Z, et al. (2018) Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. Microbiome 6(1):178. |