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:** | ***2019.023S*** | |  |
| **Short title:** Create five new families and a new suborder of vertebrate viruses in the order *Nidovirales* | | | |
|  | | | |
| **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 | |
| Gorbalenya AE, Brinton MA, de Groot RJ, Gulyaeva AA, Lauber C, Neuman BW, Ziebuhr J | | [A.E.Gorbalenya@lumc.nl](mailto:A.E.Gorbalenya@lumc.nl);  [mbrinton@gsu.edu](mailto:mbrinton@gsu.edu);  R.J.deGroot@uu.nl‎;  [A.Gulyaeva@lumc.nl](mailto:A.Gulyaeva@lumc.nl);  [chris.lauber@tu-dresden.de](mailto:chris.lauber@tu-dresden.de);  bneuman@tamut.edu‎;  [john.ziebuhr@viro.med.uni-giessen.de](mailto:john.ziebuhr@viro.med.uni-giessen.de) | |
| **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** | | | |
| 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 **Nidovirus Study Group** in consultation with:  *Arteriviridae* 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; * Four nidovirus TPs have been updated to indicate that four nidoviruses with incomplete genome sequences are not part of the proposed taxonomy |

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

|  |
| --- |
| **Name of accompanying Excel module:** 2019.020S-023S.A.v1.Nidovirales.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  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 the 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 profiles of clustering cost (CC) function that were produced for weighted hierarchical clustering of pairwise patristic distances (PPD); DEmARC 1.41 was used for the analysis (software is available upon request to AEG). All local minima of the CC profile (smallest CC values in a range of PPD values) were considered as candidate thresholds for demarcating 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, 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 zero CC (CC=0). The respective “threshold PPD ranges” were considered best candidates for demarcation. Those thresholds that were supported independently by several datasets were predominantly used to set the 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 nidovirus taxa using either the 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 an outgroup, were analyzed. For each group, the MSA of five concatenated replicative domains conserved in the *Nidovirales* were generated and used in phylogenetic and DEmARC analyses. For analyses involving viruses of an outgroup, tentatively called *Protonidovirales*, the MSA included the RdRp and HEL1 domains. These viruses also share the 3CLpro domain, which, however, 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 vertebrate viruses that represent sister lineages to the *Coronaviridae* and *Tobaniviridae* families**. The current taxonomy of vertebrate nidoviruses recognizes three families: *Arteriviridae*, *Coronaviridae* and *Tobaniviridae*. The analyses used for this proposal included six new genome sequences identified during a metagenomic study of vertebrates, which were found to be distantly related to all known vertebrate nidoviruses (Shi et al., 2018). These sequences and others were classified using a subset, CoTo3 (see **Fig. 1**), that combined all established and candidate vertebrate nidoviruses excluding the *Arteriviridae* viruses (Ar3 subset) as detailed in a separate taxonomy proposal (TP\_Arteriviridae\_Word\_module\_2019.v4.docx). Unlike arteriviruses, these viruses encode replicative ExoN and/or O-MT domains in ORF1b or its equivalent. Like other nidoviruses, these six distantly related viruses encode 3CLpro, NiRAN, RdRp, ZBD and HEL1 domains which define the order and were used for the DEmARC analysis. Their genome locations are shown for SARS-CoV and BRV, as representatives of the *Coronaviridae* and *Tobaniviridae* families as detailed in a separate proposal (TP\_CoronaTobaniviridae\_Word\_module\_2019.v4.docx) **(Fig. 2)**. Due to very large distances (which exceeded the subfamily threshold) between these new viruses and between each of them and the canonical nidoviruses, partitioning of the diversity of the *Coronaviridae* and *Tobaniviridae* families was used to establish demarcation criteria for species, genus and subfamily ranks. At these ranks, these six viruses were assigned to separate (six) taxa. Only at the family rank, defined jointly by these six viruses and the *Coronaviridae* and *Tobaniviridae* families, two new viruses - HHPAV and WJHAV - were assigned to a common family taxon. Thus, the six viruses prototype only five new nidovirus families. These assignments are detailed in a tree of the vertebrate nidoviruses (excluding the *Arteriviridae* viruses) for the genus, subfamily and family ranks (**Figs. 3-5**).  The viruses representing five new nidovirus families form two sister monophyletic clusters in the phylogeny excluding the *Arteriviridae* (**Figs. 3-5**): #1) HOFAV, GGGSAV and CBHPTAV; and #2) HHPAV, WJHAV and NGSAV. Viruses of the cluster #1 are basal to the *Arteriviridae* family and form a monophyletic lineage together with the *Arteriviridae* in the phylogeny of all nidoviruses (**Fig. 1**). Based on this association and the large distances, we assigned the three new virus families of cluster #1 to the suborder *Arnidovirineae* that includes currently only the family *Arteriviridae*. Viruses of the two new families of cluster #2 are separated from the vertebrate nidoviruses of two established suborders by distances that are typical for inter-suborder relations. Based on the distinct phylogenetic position of the cluster #2 viruses and their large distances to other vertebrate nidoviruses, we assigned them to a new suborder *Nanidovirineae*. Further research is required to determine the characteristics common for viruses of this new suborder.  Taxonomic assignment of four nidoviruses, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, which were used in the presented computational analysis, have been postponed till after full genome sequence of the respective viruses were determined.  ***Demarcation criteria***. We used either a range or a particular value of patristic pairwise distances (PPD) calculated using FastTree 2.1.4 SSE3 ML phylogeny based on an MSA of five concatenated domains (3CLpro, NiRAN, RdRp, ZBD and HEL1) as demarcation criterion for taxa at each of the following four ranks: subfamily, genus, subgenus, and species (**Tables 1 and 2**). They were selected as local minima in the CC distribution, commonly corresponding to the CC=0 or close to it (see above).  **Table 1** Demarcation thresholds for four ranks of five new families of the *Nidovirales* order   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **rank** | **Taxa #**  **current new** | **PPD range1** | **PUD (%) range2** | **Dataset used3** | | family**4** | 0 5 | 2.498-3.410 | 0.604-0.683 | CoTo3\_5d | | subfamily**5** | 0 6 | 1.404-1.779 | 0.461-0.519 | CoTo3\_5d | | genus**5** | 0 6 | 0.851-0.891 | 0.351-0.360 | CoTo3\_5d | | species**5** | 0 6 | 0.085-0.090 | 0.073-0.076 | CoTo3\_5d |   **1**The demarcation threshold is depicted as a range of PPD values for which the number of clusters (taxa) remained constant and the CC=0. PPD values account for repeated replacements of amino acid residues.  **2**The demarcation threshold is depicted as a range of PUD values, deduced from PPD values for which the number of clusters (taxa) remained constant and the CC=0. PUD values were calculated as the % of different residues in compared proteins.  **3**See Figure 1.  4The demarcation at this rank was determined by partitioning diversity of the *Coronaviridae* and *Tobaniviridae* families, and six new nidoviruses.  5The demarcation at this rank was determined by partitioning diversity of the *Coronaviridae* and *Tobaniviridae* families.    **Fig. 1**. Nidovirus phylogeny and subsets used to advance the taxonomy. Depicted is a phylogenetic tree of nidoviruses representing 113 established and newly proposed species (left) and 14 subsets of nidoviruses (right), which were analysed to build the taxonomy. For virus acronyms, see the accompanying spreadsheet; black and green indicate established and newly delineated species, respectively. The tree was reconstructed using IQ‑Tree 1.5.5 based on an MSAs of five domains (3CLpro, NiRAN, RdRp, ZBD and HEL1) with the best fitting evolutionary model selected independently for each domain. Subsequently, the tree was midpoint-rooted. Branch support was estimated using the 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 analysed to verify and advance the taxonomy of the order. Taxonomic assignment of four nidoviruses, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, which were used in the presented computational analysis, have been postponed till after full genome sequence of the respective viruses were determined.      **Fig. 2**. Domain combination used for phylogenetic and DEmARC analyses of nidoviruses. Shown are the locations of the conserved replicative domains (5d, 5 domains) used in this analysis of the CoTo3 group (see **Fig. 1**). They are depicted in relation to the genome, open reading frames, and (other) replicative domains of SARS coronavirus (SARS-CoV) and Breda torovirus (BRV). The results shown in **Figs 3-5** were obtained using an MSA of this domain combination for the vertebrate nidoviruses, excluding the *Arteriviridae*.    **Fig. 3.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted in **Figure 1**) of vertebrate nidoviruses, excluding the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed **genus** structure of the six viruses forming a sister lineage to the *Coronaviridae* and *Tobaniviridae* is detailed on the right. Taxonomic assignment of four nidoviruses, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, which were used in the presented computational analysis, have been postponed till after full genome sequence of the respective viruses were determined.    **Fig. 4.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted on **Figure 1**) of vertebrate nidoviruses, excluding the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed **subfamily** structure of the six viruses forming a sister lineage to the *Coronaviridae* and *Tobaniviridae* is detailed on the right. Taxonomic assignment of four nidoviruses, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, which were used in the presented computational analysis, have been postponed till after full genome sequence of the respective viruses were determined.    **Fig. 5.** Cluster partitioning of the phylogenetic tree (branch of the tree depicted on **Figure 1**) of vertebrate nidoviruses, excluding the *Arteriviridae* family. For virus abbreviations, see the accompanying spreadsheet. The current and proposed **family** structure of the six viruses forming a sister lineage to the *Coronaviridae* and *Tobaniviridae* is detailed on the right. Taxonomic assignment of four nidoviruses, HHPAV, Mu-BatCoV\_JTAC2, TsinCoV\_118981 and GMRSToV, which were used in the presented computational analysis, have been postponed till after full genome sequence of the respective viruses were determined. |
| 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:** |
| --- |
| Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, Holmes EC, Zhang YZ. 2018. The evolutionary history of vertebrate RNA viruses. Nature 556:197-202.  Gorbalenya AE, Brinton MA, Cowley J, de Groot R, Gulyaeva A, Lauber C, Neuman B, Ziebuhr J. 2017. ICTV taxonomy proposal 2017.014S. Establishing taxa at the ranks of subfamily, genus, sub-genus and species in six families of invertebrate nidoviruses. |