

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | **2021.006S** |  |
| **Short title:** Create five new subfamilies (*Caphthovirinae*, *Kodimesavirinae*, *Ensavirinae*, *Paavivirinae* and *Heptrevirinae*) (*Picornavirales*: *Picornaviridae*) | | |
|  | | |

**Author(s) and email address(es)**

|  |  |
| --- | --- |
| Knowles NJ, Gorbalenya AE, Lindberg AM, Oberste MS, Palmenberg AC, Reuter G, Simmonds P, Skern T, Tapparel C, Wolthers KC, Woo PCY, Zell R | nick.knowles@pirbright.ac.uk; a.e.gorbalenya@lumc.nl; michael.lindberg@lnu.se; mbo2@cdc.gov; acpalmen@wisc.edu; reuter.gabor@ddr.antsz.hu; peter.simmonds@ndm.ox.ac.uk; timothy.skern@meduniwien.ac.at; caroline.tapparel@unige.ch; k.c.wolthers@amc.uva.nl; pcywoo@hku.hk; roland.zell@med.uni-jena.de |

**Author(s) institutional address(es) (optional)**

|  |
| --- |
| The Pirbright Institute, Ash Road, Pirbright, Woking, Surrey, GU24 0NF, UNITED KINGDOM. [NJK]  Department of Medical Microbiology, Leiden University Medical Center, P.O. Box 9600, E4-P, 2300 RC Leiden, THE NETHERLANDS [AEG]  Dept. of Chemistry and Biomedical Sciences, Linnaeus University, S-391 82 Kalmar, SWEDEN. [AML]  Division of Viral Diseases, Centers for Disease Control and Prevention (CDC), 1600 Clifton Rd. NE, Mailstop A34, Atlanta, GA 30333, USA. [MSO]  Dept. of Biochemistry, Institute for Molecular Virology, 527B Bock Laboratories, 1525 Linden Drive, Madison, WI 53706, USA. [ACP]  Department of Medical Microbiology and Immunology Medical School, University of Pécs, H-7624 Szigeti út 12., Pécs, HUNGARY. [GR]  Nuffield Department of Experimental Medicine, University of Oxford, Peter Medawar Building, Sout Parks Road, Oxford, OX1 3SY, UNITED KINGDOM. [PS]  Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1090 Vienna, AUSTRIA. [TS]  University of Geneva Faculty of Medicine, 1 Rue Michel-Servet, 1205 Geneva, SWITZERLAND. [CT]  Department of Medical Microbiology, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, THE NETHERLANDS. [KCW]  Dept. of Microbiology, The University of Hong Kong, HONG KONG SAR. [PCYW]  Jena University Hospital, Department of Virology and Antiviral Therapy, Hans-Knoell-Strasse 2, D-07745 Jena, GERMANY. [RZ] |

**Corresponding author**

|  |
| --- |
| Knowles NJ |

**List the ICTV Study Group(s) that have seen this proposal**

|  |
| --- |
| *Picornaviridae* Study Group |

**ICTV study group comments and response of proposer**

|  |
| --- |
|  |

**Authority to use the name of a living person**

|  |  |
| --- | --- |
| **Is any taxon name used here derived from that of a living person (Y/N)** | N |

|  |  |  |
| --- | --- | --- |
| **Taxon name** | **Person from whom the name is derived** | **Permission attached (Y/N)** |
|  |  |  |
|  |  |  |
|  |  |  |

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chair | 28/05/2021 |
| Date of this revision (if different to above) | 14/09/2021 |

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

|  |
| --- |
| Dear Nick and Study Group  Many thanks for submitting the sub-family taxonomy proposal for Picornaviridae. This was reviewed at the ICTV Executive Committee meeting yesterday as it was given a designation of Ac. This means that it is accepted pending minor changes as listed below:  a) As discussed, we can change the name of the Cafthovirinae sub-family  b) There is a formal check done of the spreadsheet and I attach the errors detected.  There is the same error flagged about not including the name of the order above Picornaviridae in the proposed taxonomy. The other errors originate from not including the previous taxonomy in the grey boxes on the left side of the spreadsheet.  Do email back if any of this is unclear. Can you send this back with 2 months so it can be included in this year’s ratification vote.  Best wishes  Peter  Proposer response:  The names of two of the proposed subfamilies have been changed (*Cafthovirinae* to *Caphthovirinae* and *Dikomesavirinae* to *Kodimesavirinae*) to conform with the published paper by Zell et al., 2021.  Errors flagged by the spreadsheet checker have been corrected. |

**Part 2:** **NON-TAXONOMIC PROPOSAL**

**Text of proposal**

|  |
| --- |
|  |

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

**Name of accompanying Excel module**

|  |
| --- |
| 2021.006S.R.Picornaviridae\_5nsfam.xlsx |

**Abstract**

|  |
| --- |
| The virus family *Picornaviridae* presently comprises 68 approved genera with 158 species plus many unassigned viruses. In order to correlate picornavirus taxonomy with the functional and genomic groupings between genera, the establishment of five subfamilies (*Caphthovirinae*, *Kodimesavirinae*, *Ensavirinae*, *Paavivirinae* and *Heptrevirinae*) is proposed. The subfamilies are defined by phylogenetic analyses of 3CD- (precursor of virus-encoded proteinase and polymerase) and P1- (capsid protein precursor) coding sequences and comprise between 7 and 22 currently approved virus genera. Members of the proposed subfamilies typically show some commonalities in their genome organisations, including VP1/2A cleavage mechanisms and possession of leader proteins. Other features, such as internal ribosomal entry site types, are more variable. The proposed addition of a subfamily layer to the taxonomy of picornaviruses provides a valuable additional organisational level to the family that acknowledges the existence of higher-level evolutionary groupings of its component genera. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | Picornaviruses are genetically highly diverse and infect members of six vertebrate classes in all continents. Members of the family *Picornaviridae* are currently classified into 68 assigned genera with 158 species. However, many yet unassigned viruses are awaiting classification. All picornaviruses share orthologous proteins, which exhibit the characteristic picornavirus hallmarks (i.e., rhv domains with a jelly-roll fold of the three major CPs, a Walker A motif of 2Chel, a GxCGx10-15GxH active site sequence of the 3Cpro, and the RdRP sequence motifs KDE, DxxxxD, (Y)GDD and FLK(R) of the 3Dpol). More variable and often genetically non-homologous picornavirus proteins include the leader protein (L) where present, 1A, 2A, 2B, 3A and 3B, although some may play similar functional roles in replication.  Phylogenetic relationships between the five proposed subfamilies are shown in Figs. 1 (3CD region) and 2 (P1 capsid). Generalised genome layouts are depicted in Fig. 3. GRAViTy analysis is shown in Fig. 4.  The genomic and functional properties of the supergroups/proposed picornavirus subfamilies are as follows:  ***Caphthovirinae*** (SG1). The name was derived from the two earliest described genera ***Ca****rdiovirus* and *A****phtho****virus*. Members: *Ailurivirus, Aphthovirus, Bopivirus, Cardiovirus, Cosavirus, Erbovirus, Hunnivirus, Malagasivirus, Mischivirus, Mosavirus, Mupivirus, Senecavirus, Teschovirus, Torchivirus, Tottorivirus* and *Marsupivirus* with altogether 35 species. Characteristic feature of virus genome expression in this subfamily is the cleavage of CP precursor 1AB (VP0) into 1A (VP4) and 1B (VP2) (Figure 3). Other typical but not invariant features are the presence of a leader protein (exceptions: *Bopivirus, Cosavirus, Marsupivirus*) and a 2A with NPG↓P motif (exceptions: *Cardiovirus C, Malagasivirus, Mupivirus, Tottorivirus*). Aphthoviruses and erboviruses possess a leader proteinase (Lpro). Mosaviruses have two copies of 3BVPg, while foot-and-mouth disease virus (genus *Aphthovirus*) has three copies. Viruses in the majority of genera possess a type II IRES.  ***Kodimesavirinae*** (SG2). The name was derived from the four earliest described genera of this subfamily, ***Di****cipivirus*, ***Ko****buvirus*, ***Me****grivirus* and ***Sa****livirus*. Members: *Danipivirus,* *Dicipivirus, Gallivirus, Hemipivirus, Kobuvirus, Livupivirus, Ludopivirus, Megrivirus, Myrropivirus, Oscivirus, Passerivirus, Pemapivirus, Poecivirus, Pygoscepivirus, Rafivirus, Rajidapivirus, Rosavirus, Sakobuvirus, Salivirus, Sicinivirus, Symapivirus* and *Tropivirus* (altogether 38 species). Characteristic feature: 2B protein is conserved (orthologous) among the members of the subfamily (Fig. 3, Suppl. Fig. 3). Other typical features: most members of this subfamily have a leader protein (exception: *Dicipivirus, Megrivirus A*, *Megrivirus B*, *Megrivirus C*, *Megrivirus D, Poecivirus, Rafivirus C, Rosavirus*), uncleaved 1AB (exception: *Danipivirus, Dicipivirus, Rosavirus*), presence of a 2A protein with an H-box/NC motif (exceptions: *Dicipivirus*, a few kobuviruses, *Oscivirus, Pygoscepivirus* and *Rajidapivirus,* ; some megriviruses exhibit a composite 2A gene region with three different 2A proteins, the last of which has a H-box/NC motif). Viruses of the *Kodimesavirinae* have a type II or type IV IRES.  ***Ensavirinae*** (SG3). The name was derived from the two earliest described genera, ***En****terovirus* and ***Sa****pelovirus*. Members: *Anativirus, Boosepivirus, Diresapivirus, Enterovirus, Felipivirus, Parabovirus, Rabovirus, Sapelovirus* (30 species). Characteristic features are the 1AB cleavage and the presence of a 2A protein with proven or assumed chymotrypsin-like cysteine proteinase activity (Fig. 3; exception: *Anativirus* which appears to lack a functional 2A protein). Beside the 1A and 2A proteins, 2B and 3A proteins are also conserved (orthologous) among the members of the subfamily (Suppl. Figs. 4, 5, 6). IRES types I, II and IV are found.  ***Paavivirinae*** (SG4). The name was derived from the two earliest described genera, ***Pa****rechovirus* and ***Avi****hepatovirus*. Members: *Aalivirus, Aquamavirus, Avihepatovirus, Avisivirus, Crohivirus, Grusopivirus, Kunsagivirus, Limnipivirus, Orivirus, Parechovirus, Pasivirus, Potamipivirus, Shanbavirus* (29 species). Characteristic features: polyprotein expressed by all members lack an L protein and 1AB remains uncleaved (Fig. 3). The 2B protein is conserved (orthologous) among all members of the subfamily (Suppl. Fig. 7). Viruses of this subfamily exhibit the highest variability of their 2A gene region. Most viruses have one of three composite 2A gene regions: (i) one to three 2A proteins with NPG↓P motif plus a protein with H-box/NC motif (crohiviruses, grusopi A viruses, parechoviruses B-F, potamipiviruses and shanbaviruses), (ii) one to four 2A proteins with NPG↓P motif plus a domain with homology to the AIG1-type guanine nucleotide-binding domain plus a domain with H-box/NC motif (aaliviruses, avihepatoviruses and avisiviruses), (iii) one to three 2A proteins with NPG↓P motif plus a unique protein domain with unknown function (aquamaviruses and kunsagiviruses). A few viruses of the subfamily have single 2A proteins: (i) members of *Parechovirus A* have only a protein with H-box/NC motif, (ii) pasiviruses a protein with NPG↓P motif, and (iii) oriviruses and grusopi B viruses have a 2A protein with unknown function. All known aquamaviruses have two copies of 3BVPg. The IRES belong to type II or IV.  ***Heptrevirinae*** (SG5). The name was derived from the two earliest described genera, ***Hep****atovirus* and ***Tre****movirus*. Members: *Caecilivirus,* *Crahelivirus, Fipivirus, Gruhelivirus, Hepatovirus, Rohelivirus,* and *Tremovirus* (21 species). Characteristic features: very short 1A peptide (14-35 amino acids) and 1AB cleavage occurs (Fig. 3). The 2A protein in hepatoviruses has unknown function and is believed to be cleaved from the VP1-2A precursor (pX) by an unknown host proteinase in a late stage of virion morphogenesis. Other viruses of the subfamily (caeciliviruses, craheliviruses, gruheliviruses, roheliviruses) may apply a similar mechanism. Fipiviruses and tremoviruses have 2A proteins with an H-box/NC motif. Hepatoviruses have a type V IRES and tremoviruses have a type IV IRES. The IRES type of the remaining viruses of this subfamily is undetermined.  ***Picornaviruses not assigned to a subfamily.*** In addition to themany picornavirus sequences assigned to the five proposed subfamilies, highly divergent virus sequences of the genera *Ampivirus* and *Harkavirus* as well as several further currently unclassified picornaviruses are candidates of additional subfamilies. However, the limited availability of sequence data for them prevents reliable assignment to one of the proposed subfamilies or the definition of a new subfamily.  Further analyses and figures may be found in Zell et al. (2021). | |

**Supporting evidence**



**Figure 1:** **Phylogenetic analysis picornavirus 3CD protein.** 578 sequences representing the 3CD region (3162 nt) of all known picornavirus species and types were analysed with MrBayes v3.2 (nucleotide substitution model HKY+G+I). Convergence was reached after 13 million generations. Sequences cluster in clades 1 to -8 (indicated in different colors). Clades 1 to -5 correspond to supergroups (SG). Posterior probabilities of major clades are presented. The scale indicates substitutions per nucleotide.



**Figure 2:** **Phylogenetic analysis picornavirus P1 protein.** 651 sequences representing the P1 region (5214 nt) of all known picornavirus species and types were analysed with MrBayes v3.2 (nucleotide substitution model GTR+G+I). Convergence was reached after 17 million generations. Sequences cluster in clades (indicated in different colors). Clades 1 to -5 correspond to supergroups. Posterior probabilities of major clades are presented. The scale indicates substitutions per nucleotide.



**Figure 3: Schematic presentation of picornaviral genome organization and expression (not drawn to scale).** Five supergroups (SG) are distinguished by phylogenetic relationships (compare Figures 1, 2) and different functional organisation of their genomes. The open reading frame encoding a polyprotein is indicated by a box. Only dicipiviruses of the *Kodimesavirinae* have dicistronic genomes (not shown). Localisation of processed proteins in the polyprotein is indicated. Protein designations follow the L434 rule. Members of the *Paavivirinae* and the megriviruses of the *Kodimesevirinae* may have a composite 2A gene region comprising up to six 2A domains with partly unknown functions. Superscripts indicate domains with known function, i.e., rhv, characteristic jelly roll folding of capsid proteins, pro, chymotrypsin-like cysteine proteinase domain, npgp, cis-active translational termination-reinitiation site, H-box/NC, a domain with similarity to the H-rev 107 family of proteins, hel, P-loop ATPase-like helicase domain, VPg, genome-linked viral peptide, pol, RNA-dependent RNA polymerase. \* indicates genome regions which are not shared by all members of the respective supergroup/subfamily (the predominant domain is shown); § copy number of 3BVPg peptide may vary in few members; † function of 2A protein unknown. Roman numerals indicate IRES types found in the 5'-untranslated region (5'-UTR). Orange boxes indicate the proteins of the Hel-Pro-Pol core replicative module common to all picornaviruses, green boxes those proteins which are orthologous within the respective supergroup/subfamily only, and yellow boxes highlight specific features characteristic to some supergroups/subfamilies.



**Figure 4: GRAViTy analysis.** Output from analysis by GRAViTy of seventy-five full-length candidate picornavirus sequences downloaded from GenBank with the RNA virus dataset (DB-4 - Baltimore Group III, IV, V, VI and VII - all RNA viruses and retroelements), using the web server hosted by the MRC-University of Glasgow (<http://gravity.cvr.gla.ac.un>). The section of the dendrogram containing the submitted sequences and classified picornaviruses within DB-4 is shown (with the exception of ampivirus which grouped elsewhere in the *Picornavirales* clade). Branches representing existing and proposed genera are condensed. Unassigned viruses are labelled by GenBank accession numbers. Booster bootstrap values [5] of ≥70% generated from 100 data re-samplings are shown on branches. Picornavirus supergroups (SG) are presented in different colours.

**References**

1. Zell R, Knowles NJ, Simmonds P (2021) A proposed division of the family *Picornaviridae* into subfamilies based on phylogenetic relationships and functional genomic organization. Arch Virol, 166(10): 2927-2935. doi: 10.1007/s00705-021-05178-9. Epub 2021 Aug 4.