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:** | ***2019.014S*** | |  |
| **Short title:**  Create 2 new species (*Parechovirus E* and *Parechovirus F*) in the genus *Parechovirus* | | | |
|  | | | |
| **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 | |
| Zell R, Gorbalenya AE, Hovi T, Knowles NJ, Lindberg M, Oberste S, Palmenberg AC, Reuter G, Simmonds P, Skern T, Tapparel C, Wolthers K, Woo P | | roland.zell@med.uni-jena.de; a.e.gorbalenya@lumc.nl; tapani.hovi@thl.fi; nick.knowles@pirbright.ac.uk; michael.lindberg@lnu.se; soberste@cdc.gov; acpalmen@wisc.edu; reuter.gabor@gmail.com; peter.simmonds@ndm.ox.ac.uk; timothy.skern@meduniwien.ac.at; caroline.tapparel@unige.ch; k.c.wolthers@amc.uva.nl; pcywoo@hkucc.hku.hk | |
| **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]) | | Jena University Hospital [RZ]  Leiden University Medical Center [AEG]  National Institute for Health and Welfare [TH]  The Pirbright Institute [NJK]  Linnaeus University Kalmar [ML]  Centers for Disease Control and Prevention [SO]  University of Wisconsin [ACP]  University of Pécs [GR]  University of Oxford [PS]  Medical University of Vienna [TS]  University of Geneve [CT]  Universiteit van Amsterdam [KW]  University of Hong Kong [PW] | | | | |
| **Corresponding author** | | | |
| **Roland Zell** (roland.zell@med.uni-jena.de) | | | |
| **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) | | ***Picornaviridae* Study Group** | |
| **ICTV Study Group comments (if any) and response of the proposer:** | | | |
|  | | | |
|  | | | |
| Date first submitted to ICTV: | | | 21/05/2019 |
| Date of this revision (if different to above): | | | 20/08/2019 |

|  |
| --- |
| **ICTV-EC comments and response of the proposer:** |
| Incorrect GenBank accession number (KY855435) used for Parechovirus E1.  Response: accession number corrected to KY645497. |

**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:** 2019.014S.A.v1.2newsp\_Parechovirus\_E-F.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. |

**Create two new species (*Parechovirus E* and *Parechovirus F*) in the genus *Parechovirus***

The genus *Parechovirus* presently consists of 4 species:

*Parechovirus A* 19 types

*Parechovirus B*  6 types

*Parechovirus C*  1 type

*Parechovirus D* 1 type

Two novel, parecho-like picornaviruses were detected in a pool of faecal samples from common kestrels (*Falcon tinnunculus*) and red-footed falcons (*F. vespertinus*) in Hungary (Pankovics et al., 2017) and in gut tissue of geckos (*Teratoscincus roborowskii*) captured in the Xinjiang province, China (Shi et al., 2018). No virus was isolated yet.

**Relation of novel parechoviruses to *Parechovirus A* to *D* and other picornaviruses:**

- Genome layout of falcon parechovirus:

5'-UTRIRES-II[1AB-1C-1D-2A1npgp/2A2npgp/2A3H-box/NC-2B-2Chel/3A-3BVPg-3Cpro-3Dpol]3'-UTR

genome layout of gecko parechovirus:

5'-UTR[1AB-1C-1D-2A1npgp/2A2H-box/NC-2B-2Chel/3A-3BVPg-3Cpro-3Dpol]3'-UTR

(compare Fig. 1 of supporting material)

- The novel parechoviruses have typical hallmarks of picornaviruses:

- **capsid proteins** 1AB, 1C, 1D have **rhv** domains with drug-binding site,

- falcon parechovirus: **2A1 and 2A2** have a **NPGP**-motif, 2A3 has a H-box/NC motiv,

gecko parechovirus: **2A1** has a **NPGP**-motif, 2A2 has a H-box/NC motiv

- **2Chel** with **GxxGxGKS** motif of helicases,

- **3BVPg** peptide with **Y-3** residue,

- **3Cpro** with **GxCGx14GxH** motif,

- **3Dpol** with **KDE**, **PSG**, **YGDD** and **FLKR** motifs,

- P1 and 3CD sequences cluster with the parechovirus sequences in **phylogenetic analyses** (compare Figs. 2, 3 of supporting material).

**Distinguishing features of falcon parechovirus compared to parechoviruses:**

- Falcon parechovirus has two 2A proteins with **NPGP**-motif

- **Sequence divergence** (uncorrected p-distance) of complete genomes suggest two novel parechovirus species, ***Parechovirus E*** and ***Parechovirus F***: genetic distances of *between-species* comparisons is greater 30% (compare Table 1):

Table 1. Estimates of evolutionary divergence between amino acid sequences

[1] L02971, Parechovirus A1 Harris

[2] AF327920, parechovirus B1 (Ljungan virus) 87-012

[3] KY432929, parechovirus B6 RtMrut-PicoV/JL2014-2

[4] HF677705, parechovirus C1 (Sebokele virus) An/B/1227/d

[5] KF006989, parechovirus D1 (ferret parechovirus) MpPeV1

[6] KY645497, parechovirus E1 (falcon picornavirus) falcon/HA18-080/2014/HUN

[7] MG600084, parechovirus F1 (Yili Teratoscincus roborowskii picornavirus 2) LPWC210215

[ 1 2 3 4 5 6 7 ]

[1]

[2] 0.521

[3] 0.524 0.183

[4] 0.538 0.367 0.366

[5] 0.610 0.585 0.590 0.585

[6] 0.526 0.367 0.360 0.399 0.580

[7] 0.617 0.578 0.585 0.596 0.607 0.594

- **Sequence divergences** (uncorrected p-distances) of orthologous proteins in pairwise comparisons of parechoviruses E1 and F1 with representative picornavirus supergroup 4 viruses (*Aalivirus/Aquamavirus/Avihepatovirus/Avisivirus/Crohivirus/Kunsagivirus/ Limnipivirus/Orivirus/Pasivirus/Potamipivirus/Shanbavirus*) justify creation of two new species (compare Table 2).

**Table 2: Amino acid divergence\***

P1 2Chel 3Cpro 3Dpol

**falcon parechovirus vs. ...**

*within-genus comparisons*:

parechovirus A1 52.8% 48.2% 48.5% 53.0%

parechovirus B1 36.8% 37.2% 39.1% 36.0%

parechovirus C1 35.5% 39.6% 40.2% 39.4%

parechovirus D1 42.7% 55.8% 61.1% 55.5%

parechovirus F1 55.2% 58.7% 66.1% 58.6%

*between-genus comparisons*: 58-78% 60-79% 56-86% 58-72%

**gecko parechovirus vs. ...**

*within-genus comparisons*:

parechovirus A1 58.8% 58.6% 67.0% 62.7%

parechovirus B1 53.6% 55.4% 64.6% 56.1%

parechovirus C1 56.3% 57.3% 65.6% 58.0%

parechovirus D1 57.6% 55.2% 66.5% 59.0%

parechovirus E1 55.2% 58.7% 66.1% 58.6%

*between-genus comparisons*: 55-80% 59-77% 66-86% 61-70%

\* number of amino acid differences per site

**Exemplar:**

*Parechovirus E*, parechovirus E1 (falcon parechovirus) strain falcon/HA18-080/2014/HUN, GenBank acc. no. KY645497

*Parechovirus F*, parechovirus F1 (gecko parechovirus, Yili Teratoscincus reborowskii picornavirus 2) strain LPWC210215, GenBank acc. no. MG600084

**Species demarcation criteria:**

Members of a species of the genus *Parechovirus*:

- are less than 30% divergent in polyprotein aa sequence,

-are less than 30% divergent in P1 aa sequence,

-are less than 20% divergent in 2C+3CD aa sequence,

- share a common genome organization,

| **References:** |
| --- |
| 1. Pankovics et al. 2017. Ljungan/Sebokele-like picornavirus in birds of prey, common kestrel (*Falco tinnunculus*) and red-footed falcon (*F. vespertinus*). Inf Genet Evol 55:14-19.  2. Shi et al. 2018. The evolutionary history of vertebrate RNA viruses. Nature 556:197-202. |

**Supporting Material**



**Figure 1:** Genome organization of parechovirus E1 and F1 (schematic depiction). The open reading frame is indicated by a box. Positions of putative 3Cpro cleavage sites are indicated by a ▼ and the site of termination/reinitiation of RNA translation at the NPGP sequence motif is indicated by a hash (#). The names and lengths of the deduced proteins are presented. The 5'-UTRs may be incomplete.



**Legend to Figure 2:**  Phylogenetic analysis of picornavirus **P1** using Bayesian tree inference (MrBayes 3.2). Forty-five picornavirus sequences of the *Aalivirus/Aquamavirus/Avihepatovirus/Avisivirus/Crohivirus/ Kunsagivirus/Limnipivirus/Orivirus/Parechovirus/Pasivirus/Potamipivirus/Shanbavirus* supergroup were retrieved from GenBank; the cardiovirus sequence served as outgroup. [Note: the supergroup does not imply a taxonomic entity but reflects phylogenetic clustering of the respective genera observed in different tree inference methods (NJ, ML, Bayesian MCMC).] Presented are GenBank accession numbers, ***genus*** ***names***, *species names*, type and—if available—common names in round brackets. Designations of isolates are given in square brackets. Yet unassigned viruses are printed in blue. The proposed name is printed in red and indicated by a dot (●). Numbers at nodes indicate posterior probabilities obtained after 2,000,000 generations. The optimal substitution model (GTR+G+I) was determined with MEGA 5. The scale indicates substitutions/site.



**Legend to Figure 3:**  Phylogenetic analysis of picornavirus **3CD** using Bayesian tree inference (MrBayes 3.2). Forty-six picornavirus sequences of the *Aalivirus/Aquamavirus/Avihepatovirus/Avisivirus/ Crohivirus/Kunsagivirus/Limnipivirus/Orivirus/Parechovirus/Pasivirus/Potamipivirus/Shanbavirus* supergroup were retrieved from GenBank; the cardiovirus sequence served as outgroup. [Note: the supergroup does not imply a taxonomic entity but reflects phylogenetic clustering of the respective genera observed in different tree inference methods (NJ, ML, Bayesian MCMC).] Presented are GenBank accession numbers, ***genus*** ***names***, *species names*, type and—if available—common names in round brackets. Designations of isolates are given in square brackets. Yet unassigned viruses are printed in blue. The proposed name is printed in red and indicated by a dot (●). Numbers at nodes indicate posterior probabilities obtained after 2,000,000 generations. The optimal substitution model (GTR+G+I) was determined with MEGA 5. The scale indicates substitutions/site.