

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | **2020.014M** |  |
| **Short title:**  Create one new genus (*Alphapaprhavirus*) including two new species (*Mononegavirales*: *Rhabdoviridae*) | | |
|  | | |

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

|  |  |
| --- | --- |
| Walker PJ, Blasdell KR, Dietzgen RG, AR, Fooks, Freitas-Astúa J, Kondo H, Kuzmin IV, Tesh RB, Vasilakis N, Whitfield AE | peter.walker@uq.edu.au; kim.blasdell@csiro.au; r.dietzgen@uq.edu.au; tony.fooks@apha.gsi.gov.uk; juliana.astua@embrapa.br; hkondo@rib.okayama-u.ac.jp; ivkuzmin@yandex.ru; rtesh@utmb.edu; nivasila@utmb.edu; awhitfi@ncsu.edu |

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

|  |
| --- |
|  |

**Corresponding author**

|  |
| --- |
| Peter Walker |

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

|  |
| --- |
| *Rhabdoviridae* Study Group |

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

|  |
| --- |
| Approved by all responding SG members (11 of 14) with minor revisions. |

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

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

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chair | 2 August 2020 |
| Date of this revision (if different to above) |  |

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

|  |
| --- |
|  |

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

**Text of proposal**

|  |
| --- |
|  |

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

**Name of accompanying Excel module**

|  |
| --- |
| 2020.001M\_014M\_015M\_016M.R.Rhabdoviridae.xlxs |

**Abstract**

|  |
| --- |
| The new genus *Alphapaprhavirus* is proposed to accommodate three currently unassigned rhabdoviruses that have been detected in lepidopteran insects (butterflies, moths).  The viruses will be assigned to two new species within the new genus: *Pararge alphapaprhavirus and Hubei alphapaprhavirus.* Each member virus was detected by metagenomic sequence analysis. No virus isolates are available at this time. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | **Viruses to be assigned to the new genus**  Pararge aegeria rhabdovirus (PAeRV) was detected in a laboratory population of butterflies (*Pararge aegeria*) in Belgium, in 2013 [3]. The near-complete genome sequence (13,062 nt) of the virus has been determined, including complete coding sequences but incomplete 3' and 5' termini [3]. We propose to assign PAeRV to the new species *Pararge alphapaprhavirus*.  Lepidopteran rhabdo-related virus OKIAV-12 was detected as a sequence read in the transcriptome of butterflies (*Pararge aegeria*) collected in North Rhine-Westphalia, Germany, in 2011 [1]. OKIAV-12 is the number assigned to the assembled genome. The near-complete genome sequence (13,045 nt) of the virus has been determined, including complete coding sequences but incomplete 3' and 5' termini [1]. Based on the very high levels of amino acid sequence identity (**Tables 1-3**), we consider OKIAV-12 to be a naturally circulating variant of PAeRV which had been reported previously [3].  Hubei lepidoptera virus 2 (HbLV-2; strain LCM101902) was detected in insects (Lepidoptera) collected in Hubei Province, China, in 2013 [6]. The near-complete genome sequence (13,109 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [6]. We propose to assign HbLV-2 to the new species *Hubei alphapaprhavirus*.  **Genome organisations**  Alphapaprhavirus genomes contain the four canonical rhabdovirus structural protein genes (N, P, M, and L) and two genes encoding class I transmembrane glycoproteins (G1 and G2) (**Figure 1, Figure 2**).  A Clustal X alignment indicates that alphapaprhavirus G1 and G2 proteins are similar in structure and length, and share identifiable sequence identity, with each containing all 12 conserved cysteine residues that in vesicular stomatitis Indiana virus (VSIV) form six disulphide bonds in the folded protein (**Figure 2**). The G1 proteins have four additional cysteine residues (**a, b, c and d in Figure 2**), that are likely to form two additional disulphide bonds in the ‘stalk’ region of the glycoproteins; two of these cysteine residues (**c and d in Figure 2**) also occur in G2. The similarities in structure indicate that G1 and G2 are likely to have arisen by gene duplication and it is not possible on the available evidence to determine whether either or both proteins is functionally homologous to the G proteins of other rhabdoviruses. Consecutive class I transmembrane glycoprotein genes also occur in the genomes of other rhabdoviruses assigned to other genera, i.e., ephemeroviruses and Ngaingan virus (species *Ngaingan hapavirus*), but these are phylogenetically distant from the alphapaprhaviruses.  **Phylogeny and amino acid sequence relationships**  Based on ML trees generated from complete L protein sequences, alphapaprhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 3**).  Amino acid sequence divergence between PAeRV and HbLV-2 in pair-wise alignments (p-distances) are 51.5% in the N protein, 30% in the G1 protein and 59.5% in the L protein (**Tables 1-3**).  **Other related viruses**   1. Two other rhabdoviruses that were detected in lepidopterans appear to be related to the alphapaprhaviruses.   Gata virus (GATV; strain M4) was detected in a moth (*Orgyia pseudotsugata*) collected in Washington State, USA, in 2016 [4]. The near-complete genome sequence (11,238 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [4].  Orgi virus (ORGIV; strain SP1) was detected in a moth (*Orgyia pseudotsugata*) collected in Washington State, USA, in 2016 [4]. The near-complete genome sequence (11,308 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [4].  GATV and ORGIV have been excluded from the genus at this time due to poor bootstrap support for the node linking them to PAeRV and HBLV-2 (**Figure 3**). GATV and ORGIV also lack the gene encoding a second class I transmembrane glycoprotein (G2).   1. Three other viruses detected in insects also appear to be related to the alphapaprhaviruses.   Tetrastichus brontispae RNA virus 1 (TbRV-1) was detected in 2017 in a colony of wasps (*Tetrastichus brontispae*) reared in Fujian Province, China [5]. The near-complete genome sequence (12,165 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [5].  Wuhan insect virus 7 (WhIV-7) was detected in aphids (*Hyalopterus pruni*) collected in 2013 in Hubei Province, China [2]. The near-complete genome sequence (11,165 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2].  Hubei dimarhabdovirus 2 (HbDRV-2) was detected in 2013 in insects of the order Odonta (dragonflies, damselflies) in Hubei Province, China [6]. The near-complete genome sequence (11,332 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [6].  TbRV-1, WhIV-7 and HbDRV-2 have not been classified at this stage due to the diversity of insect hosts in which they were detected and as bootstrap support for the node linking them to the clade of lepidopteran viruses is poor.  It is anticipated that the detection of additional related rhabdoviruses in the future may provide a clearer basis for the classification of these five viruses.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Alphapaprhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 5% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G1 proteins; D) significant differences in genome organisation as evidenced by numbers and locations of ORFs; E) can be distinguished in virus neutralisation tests; and F) occupy different ecological niches as evidenced by differences in vertebrate hosts and or arthropod vectors.  PAeRV and HbLV-2 meet demarcation criteria A, B, C and F.  **Derivation of the genus name**  *Alphapaprhavirus* is derived from the alpha group of lepidopteran (*papilionem*, latin butterfly) rhabdoviruses. [A beta group of lepidopteran rhabdoviruses may be the subject of a future proposal.]  **Type species**  *Pararge alphapaprhavirus* is designated as the type species of the genus as Pararge aegeria rhabdovirus was the first of proposed members of the genus to have been reported. | |

**Supporting evidence**

****

**Figure 1.** Alphapaprhavirus genome organisations. Each genome contains N, P, M and L genes corresponding to those of other rhabdoviruses and two genes (G1 and G2) encoding encoding class I transmembrane glycoproteins. Each glycoprotein is structurally homologous to the G proteins of other rhabdoviruses.

****

**Figure 2**. A Clustal X alignment of the VSIV G protein with the alphapaprhavirus G1 and G2 proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII) as are other cysteine residues (a-d) that may form additional disulphide bonds. The HbLRV-2 G2 protein contains an apparently unpaired cysteine residue in the ectodomain.

****

**Figure 3.** The evolutionary history was inferred from a Clustal W alignment of 175 complete L protein sequences of 167 animal rhabdoviruses currently assigned or recently proposed for assignment to species, two proposed members of the genus *Alphapaprhavirus* (PAeRV and HbLV-2), a virus considered a variant of PAeRV (OKIAV-12 and five other related rhabdoviruses (Gata virus, Orgi virus, Wuhan insect virus 7, Hubei dimarhabdovirus 2 and Tetrastichus brontispae RNA virus 1). Phylogenetically informative sites were selected from the alignment using Gblocks resulting in 957 positions in the final dataset. The tree was inferred in MEGA7 by using the Maximum Likelihood method based on the Whelan and Goldman + Freq. model. The tree with the highest log likelihood (-147281.07) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap values (100 iterations) are shown for each node.

**Table 1.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of alphapaprhavirus N proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | PAeRV | OKIAV12 | HbLV-2 |
| PAeRV |  |  |  |
| OKIAV12 | 100 |  |  |
| HbLV-2 | 51.5 | 51.5 |  |

**Table 2.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of alphapaprhavirus G1 proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | PAeRV | OKIAV12 | HbLV-2 |
| PAeRV |  |  |  |
| OKIAV12 | 98.7 |  |  |
| HbLV-2 | 30.0 | 30.0 |  |

**Table 3.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of alphapaprhavirus L proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | PAeRV | OKIAV12 | HbLV-2 |
| PAeRV |  |  |  |
| OKIAV12 | 99.1 |  |  |
| HbLV-2 | 59.5 | 59.7 |  |

**References**

1. Kafer S, Paraskevopoulou S, Zirkel F, Wieseke N, Donath A, Petersen M, Jones TC, Liu S, Zhou X, Middendorf M, Junglen S, Misof B, Drosten C (2019) Re-assessing the diversity of negative strand RNA viruses in insects. PLoS Pathogens 15:e1008224. PMID: 31830128; DOI: 10.1371/journal.ppat.1008224
2. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, Chen LJ, Qin XC, Xu J, Holmes EC, Zhang YZ (2015) Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife 4:e05378. PMID: 25633976; DOI: 10.7554/eLife.05378
3. Longdon B, Murray GG, Palmer WJ, Day JP, Parker DJ, Welch JJ, Obbard DJ, Jiggins FM (2015) The evolution, diversity, and host associations of rhabdoviruses. Virus Evolution 1:e014. PMID: 27774286; DOI: 10.1093/ve/vev014
4. Makhsous N, Shean RC, Droppers D, Guan J, Jerome KR, Greninger AL (2017) Genome sequences of three novel bunyaviruses, two novel rhabdoviruses, and one novel nyamivirus from Washington State moths. Genome Announcements 5:e01668-01616. PMID: 28209840; DOI: 10.1128/genomeA.01668-16
5. Meng E, Tang B, Sanchez-Garcia FJ, Qiao T, Fu L, Wang Y, Hou YM, Wu JL, Chen ZM (2019) The first complete genome sequence of a novel Tetrastichus brontispae RNA Virus 1 (TbRV-1). Viruses 11:e257. PMID: 30871248; DOI: 10.3390/v11030257
6. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ (2016) Redefining the invertebrate RNA virosphere. Nature 540:539-543. PMID: 27880757 DOI: 10.1038/nature20167