

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | **2020.023M** |  |
| **Short title:**  Create seven new genera (*Alphacrustrhavirus*, *Alphadrosrhavirus, Alphahymrhavirus*, *Betahymrhavirus*, *Betanemrhavirus, Betapaprhavirus* and *Betaricinrhavirus*), including 16 new species (*Mononegavirales*: *Rhabdoviridae*) | | |
|  | | |

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

|  |  |
| --- | --- |
| Walker PJ, Blasdell KR, Dietzgen RG, Freitas-Astúa J, Kondo H, Kurath G, Kuzmin IV, Tesh RB, Tordo N, Vasilakis N, Whitfield AE | peter.walker@uq.edu.au; kim.blasdell@csiro.au; r.dietzgen@uq.edu.au; juliana.astua@embrapa.br; hkondo@rib.okayama-u.ac.jp; gkurath@usgs.gov; ivkuzmin@yandex.ru; rtesh@utmb.edu; ntordo@pasteur.fr; 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 | 24 August 2020 |
| Date of this revision (if different to above) | 2 December 2020 |

**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**

|  |
| --- |
| 2020.023M.R.Rhabdoviridae\_7ngen\_16nsp.xlxs |

**Abstract**

|  |
| --- |
| We propose the establishment of seven new genera in the *Rhabdoviridae* to accommodate 16 new species for viruses discovered in invertebrates. These viruses are phylogenetically divergent from those assigned to genera that have been proposed to be classified to the subfamilies *Alpharhabdovirinae*, *Betarhabdovirinae* and *Gammarhabdovirinae*. The new genera will be *Alphacrustrhavirus* (2 species for crustacean viruses), *Alphadrosrhavirus* (2 species for fly viruses), *Alphahymrhavirus* (4 species for wasp viruses), *Betahymrhavirus* (2 species for wasp viruses), *Betapaprhavirus* (2 species moth viruses), *Betanemrhavirus* (2 species for roundworm viruses) and *Betaricinrhavirus* (2 species for hard tick viruses). |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | We have proposed elsewhere (2020.016M) the establishment of three new subfamilies to accommodate all existing and currently proposed genera of rhabdoviruses. In that proposal, we identified a diverse and phylogenetically distinct set of invertebrate rhabdoviruses, many recently discovered by metagenomics, that do not fall within the three proposed subfamilies. Here, we address the classification of some of those viruses by proposing the formation of seven new genera to accommodate 16 new species. Viruses assigned to each genus form a distinct monophyletic group based on well-supported Maximum Likelihood trees inferred from complete L sequences.   1. **Genus *Alphacrustrhavirus***   The new genus *Alphacrustrhavirus* comprises viruses that have been detected in marine crustaceans.    **Viruses to be assigned**  Wenling crustacean virus 10 (WlCV-10; strain WLJQ101844) and Wenling crustacean virus 11 (WlCV-11; strain WLJQ201798) were each discovered by HTS in a pool of marine crustaceans (multiple families) collected in Zhezhang Province, China, in 2014 [8]. The near-complete genome sequence of each virus has been determined (WlCV-10, 11,938 nt; WlCV-11, 11,560 nt), including complete coding sequences but incomplete 3' and 5' termini [8]. We propose to assign Wenling crustacean virus 10 to the new species *Wenling alphacrustrhavirus* and Wenling crustacean virus 11 to the new species *Zhezhang alphacrustrhavirus*.  No isolates are currently available for either of these viruses.  **Genome organization and expression**  Alphacrustrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) (**Figure 1**). In WlCV-11, the *P* gene contains an alternative open reading frame (ORF) of 297 nt encoding a highly basic 11.9 kDa protein (Px) (**Figure 1**). The Px ORF commences near the start of the P ORF but it is not known if the protein is expressed.  A Clustal X alignment indicates that alphacrustrhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates conservation of many of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 3**). Disulphide bridge CII-CIV appears to be absent and cysteine CIX appears to have formed a new unique pairing.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, alphacrustrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergence between WlCV-10 and WlCV-11 in pair-wise alignments (p-distances) are 71.7% in the N proteins, 79.0% in the G proteins and 57.4% in the L protein.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Alphacrustrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C and D. As no virus isolates are currently available neutralisation tests have not been conducted (criterion E). As the viruses were each detected in the same mix of diverse crustaceans, no information is yet available on any possible differences in ecology (criterion F).  **Derivation of the genus name**  *Alphacrustrhavirus* is derived from the alpha group of crustacean rhabdoviruses.  **Type species**  *Wenling alphacrustrhavirus* is designated as the type species of the genus (for no particular reason).   1. **Genus *Alphadrosrhavirus***   The new genus *Alphadrosrhavirus* comprises viruses that have been detected in flies (Drosophilidae). They are distant phylogenetically from those rhabdoviruses detected in flies of various species that have been assigned to an existing genus (*Sigmavirus*).  **Viruses to be assigned**  Shayang fly virus 3 (SyFV-3; strain SYY1-1) was discovered by HTS in a pool of oriental latrine flies (*Chrysomya megacephala*) collected in Hubei Province, China, in 2012 [4]. The near-complete genome sequence (15,462 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [4]. We propose to assign Shayang fly virus 3 to the new species *Shayang alphadrosrhavirus.*  Wuhan house fly virus 2 (WhHFV-2; strain SYY4-5) was discovered by HTS in a pool of houseflies (*Musca domestica*) collected in Hubei Province, China, in 2013 [4]. The near-complete genome sequence (14,731 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [4]. We propose to assign Wuhan house fly virus 2 to the new species *Hubei alphadrosrhavirus.*  No isolates are currently available for either of these viruses.  **Other related viruses**  Drosophila sturtevanti rhabdovirus 1 (DstuRV-1; strain 1) discovered by HTS in a pool of flies (*Drosophila sturtevanti*) collected in Grenada, Spain, in 2009 [5]. Only partial genome sequence (7,658 nt) has been determined, including the complete *L* gene [5].  Wuhan fly virus 3 (WhFV-3; strain SYY2-5) discovered by HTS in a pool of houseflies (*Musca domestica*) collected in Hubei Province, China, in 2012 [5]. Only partial genome sequence (7,785 nt) has been determined, including the complete *L* gene [5].  Based on the available L protein sequences, each of these viruses clusters phylogenetically with the alphadrosrhaviruses. However, as only partial genome sequences are currently available, they are not proposed for classification at this time.  **Genome organization and expression**  Alphadrosrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) (**Figure 1**). They share the common characteristic of an additional gene (*U1* in SYFV-3 and *U2* in WhHFV-2) between the *G* and *L* genes in which there are two overlapping ORFs. Each of these ORFs (U1a and U1b in SYFV-3; U2a and U2b in WhHFV-2) encodes a small hydrophobic protein with a strongly predicted transmembrane domain (see **Appendix A**). In WhHFV-2, there is also an additional gene (U1) between the *P* gene and *M* gene containing a single ORF of 492 nt encoding an 18.2 kDa protein (**Figure 1**). Small alternative or overlapping ORFs also occur in the *P* genes of each virus but it is not known if they are expressed.  A Clustal X alignment indicates that alphadrosrhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates conservation of 10 of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 4**). Only disulphide bridge CIX-CXI appears to be absent. Two additional cysteine residues in each virus appear to be available to form an additional disulphide bridge.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, alphadrosrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergence between SyFV-3 and WhHFV-2 in pair-wise alignments (p-distances) are 73.1% in the N proteins, 64.1% in the G proteins and 53.4% in the L protein.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Alphadrosrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C, D and F. As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Alphadrosrhavirus* is derived from the alpha group of fly (Drosophilidae) rhabdoviruses.  **Type species**  *Shayang alphadrosrhavirus* is designated as the type species of the genus (for no particular reason).   1. **Genus *Alphahymrhavirus***   The new genus *Alphahymrhavirus* comprises viruses that have been detected in hymenopteran insects (Hymenoptera). They are distant phylogenetically from those rhabdoviruses detected in hymenopteran insects of various species that will be assigned to another new genus (*Betahymrhavirus;* see below).  **Viruses to be assigned**  Lasius neglectus virus 2 (LnegV-2; strain Cambridge) was discovered in the transcriptome of invasive garden ants (*Lasius neglectus*) collected in the United Kingdom, in 2016 [3]. The complete genome sequence (12,041 nt) has been determined [3]. We propose to assign Lasuis neglectus virus 2 to the new species *Neglectus alphahymrhavirus.*  Hymenopteran rhabdo-related virus 38 (HyRRV-38; strain OKIAV38) was discovered in the transcriptome shotgun assembly (TSA) of a leaden spider wasp (*Pompilus cinereus*) collected in Germany, in 2011 [2]. The near-complete genome sequence (12,368 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. We propose to assign Hymenopteran rhabdo-related virus 38 to the new species *Cinereus alphahymrhavirus.*  Hymenopteran rhabdo-related virus 46 (HyRRV-46; strain OKIAV46) was discovered in the TSA of a leaden cuckoo wasp (*Chrysura radians*) collected in Turlin, Italy, in 2012 [2]. The near-complete genome sequence (11,807 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. We propose to assign Hymenopteran rhabdo-related virus 46 to the new species *Radians alphahymrhavirus.*  Hymenopteran rhabdo-related virus 109 (HyRRV-109; strain OKIAV109) was discovered in the TSA of leaden cricket-hunting wasps (*Chlorion hirtum*) collected in Israel, in 2013 [2]. The near-complete genome sequence (11,801 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. We propose to assign Hymenopteran rhabdo-related virus 109 to the new species *Hirtum alphahymrhavirus.*  **Other related viruses**  Wuhan ant virus (WhAV; strain WHMY02) was discovered in pool of a Japanese carpenter ants (*Camponotus japonicus*) collected in Hubei Province, China, in 2013 [4]. Only partial genome sequence (7,702 nt) has been determined, including the complete *L* gene [4].  Based on the available L protein sequences, WhAV clusters phylogenetically with the alphahymrhaviruses. However, as only partial genome sequence is currently available, the virus not proposed for classification at this time.  **Genome organization and expression**  Alphahymrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) (**Figure 1**). In HyRRV-109, the *P* gene contains an alternative open reading frame (ORF) of 207 nt encoding a highly acidic 8.0 kDa protein (Px) (**Figure 1**). The Px ORF commences near the start of the P ORF but it is not known if the protein is expressed.  A Clustal X alignment indicates that alphahymrhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates conservation of six of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 5**). Disulphide bridge CII-CIV, CVI-CVII and CIX-CXI appears to be absent. Various other cysteine residues in each virus appear to be available to form additional disulphide bridges.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, alphahymrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Pairwise sequence identities (p-distances) calculated in MEGA7 from Clustal W alignments indicate amino acid sequence divergence of 41.9% – 76.1% in the N proteins, 57.5% – 81.3% in the G proteins and 35.3% – 60.0% in the L proteins (**Tables 1, 2 and 3**).  **Species demarcation criteria**  Viruses assigned to different species within the genus *Alphahymrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C, D and F. As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Alphahymrhavirus* is derived from the alpha group of hymenopteran (Hymenoptera) rhabdoviruses.  **Type species**  *Neglectus alphahymrhavirus* is designated as the type species of the genus as complete genome sequence appears to be available.   1. **Genus *Betahymrhavirus***   The new genus *Betahymrhavirus* comprises viruses that have been detected in hymenopteran insects (Hymenoptera). They are distant phylogenetically from those rhabdoviruses detected in hymenopteran insects of various species that will be assigned to another new genus (*Alphahymrhavirus;* see above).  **Viruses to be assigned**  Hymenopteran rhabdo-related virus 23 (HyRRV-23; strain OKIAV23) was discovered in the TSA of a cuckoo wasp (*Chrysura austriaca*) collected in Germany, in 2010 [2]. The near-complete genome sequence (13,158 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. We propose to assign Hymenopteran rhabdo-related virus 23 to the new species A*ustriaca betahymrhavirus.*  Hymenopteran rhabdo-related virus 22 (HyRRV-22; strain OKIAV22) was discovered in the TSA of a cuckoo wasp (*Chrysis* sp.) collected in Israel, in 2012 [2]. The near-complete genome sequence (13,040 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. Based on amino acid sequence identities (see below) we consider HyRRV-22 to be a second isolate of HyRRV-23 and so it should be assigned to the same species.  Hymenopteran rhabdo-related virus 24 (HyRRV-24; strain OKIAV24) was discovered in the TSA of a leaden spider wasp (*Heterodontonyx* sp.) collected in Western Australia, in 2011 [2]. The near-complete genome sequence (12,564 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2]. We propose to assign Hymenopteran rhabdo-related virus 24 to the new species *Heterodontonyx betahymrhavirus.*  **Genome organization and expression**  Betahymrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) (**Figure 1**). They also each have an additional gene between the *M* gene and *G* gene in with two overlapping reading frames with a “slippery” sequence in the overlap region that would allow expression of the second ORF by ribosomal frame-shift. (**Figure 1**).  A Clustal X alignment indicates that betahymrhavirus G proteins share identifiable sequence identity and are similar in structure and length. They share 14 conserved cysteine residues which are likely to form seven disulphide bridges in the folded protein [7, 10]. However, alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates that few if any of these cysteine residues are commonly conserved (**Figure 6**).  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, betahymrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergence in pair-wise alignments (p-distances) between viruses representing the two species are 47.9% – 48.2% in the N proteins, 52.3% – 52.5% in the G proteins and 42.0% – 42.3% in the L protein (**Tables 4, 5 and 6**).  **Species demarcation criteria**  Viruses assigned to different species within the genus *Betahymrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C and F. The viruses share a similar genome organisation (criterion D). As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Betahymrhavirus* is derived from the beta group of hymenopteran (Hymenoptera) rhabdoviruses.  **Type species**  *Austriaca betahymrhavirus* is designated as the type species (for no particular reason).   1. **Genus *Betanemrhavirus***   The new genus *Betanemrhavirus* comprises viruses that have been detected in roundworms (Nematoda). They are distant phylogenetically from those rhabdoviruses detected in nematodes of various species that have been assigned to the existing genus *Alphanemrhavirus*.  **Viruses to be assigned**  Shayang ascaridia galli virus 2 (SyAGV-2; strain HC21241) was detected in a pool of roundworms (*Ascaridia galli* and *Ascaris suum*) sampled in Hubei Province, China, in 2014 [8]. The near-complete genome sequence (13,277 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [8]. We propose to assign Shayang ascaridia galli virus 2 to the new species *Shayang betanemrhavirus*.  Hubei rhabdo-like virus 9 (HbRLV-9; strain WHZHC73015) was detected in a pool of large pig roundworms (*Ascaris suum*) sampled in Hubei Province, China, in 2014 [8]. The near-complete genome sequence (13,776 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [8]. We propose to assign Hubei rhabdo-like virus 9 to the new species *Hubei betanemrhavirus*.  No isolates are currently available for either of these viruses.  **Genome organization and expression**  Betanemrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) (**Figure 1**) and an additional gene (*U1*) between the *P* and *M* genes (**Figure 1**). Proteins encoded in the U1 genes vary in size but appear to be homologous. In HbRLV-9, there is also a unique additional gene (*U2*) between the *M* and *G* genes (**Figure 1**).  A Clustal X alignment indicates that betanemrhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates likely conservation of 10 of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 7**). Only disulphide bridge CIX-CXI does not appear to be conserved in the betanemrhaviruses. Conservation of eight additional cysteine residues in only the betanemrhavirus ectodomains suggests there may be four additional disulphide bridges that are not present in VSIV.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, betanemrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergence between SyAGV-2 and HbRLV-9 in pair-wise alignments (p-distances) are 72.0% in the N proteins, 59.2% in the G proteins and 54.8% in the L protein.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Betanemrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C and F. The viruses have similar genome organisations (criterion D). As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Betanemrhavirus* is derived from the beta group of nematode rhabdoviruses. The alpha group of nematode rhabdoviruses form the genus *Alphanemaprhavirus* and are phylogenetically distant from the betanemrhaviruses.  **Type species**  *Shayang betanemrhavirus* is designated as the type species of the genus (for no particular reason).   1. **Genus *Betapaprhavirus***   The new genus *Betapaprhavirus* comprises viruses that have been detected in lepidopteran insects (Lepidoptera). They are distant phylogenetically from those rhabdoviruses detected in lepidopterans of various species that are to be assigned to the genus *Alphapaprhavirus* (see proposal 2020.014M).  **Viruses to be assigned**  Spodoptera frugiperda rhabdovirus (SfruRV; strain Sf) was detected initially in the Sf9 cell line derived from the fall armyworm moth (Lepodoptera: Noctuidae) and then in the corresponding parental cell line (Sf21) [6]. The near-complete genome sequence (13,534 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [6]. We propose to assign Spodoptera frugiperda rhabdovirus to the new species *Frugiperda betapaprhavirus*.  Lepidopteran rhabdo-related virus 34 (LeRRV-34; strain OKIAV34) was detected in the TSA of the orange moth (*Triodia sylvina*) collected in Germany in 2011 [2, 11]. The near-complete genome sequence (13,262 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [2, 11]. We propose to assign Lepidopteran rhabdo-related virus 34 to the new species *Sylvina betapaprhavirus*.  No isolates are currently available for either of these viruses.  **Other related viruses**  Lepidopteran rhabdo-related virus 32 (LeRRV-32; strain OKIAV32) was detected in the TSA of the longhorn moth (*Nemophora degeerella*) collected in Austria, in 2011 [2]. Only partial genome sequence (6,486 nt) has been determined, including the complete L gene [2].  Lepidopteran rhabdo-related virus 35 rhabdovirus (LeRRV-35; strain OKIAV35) was detected in the TSA of the pine-tree lappet (*Dendrolimus pini*) collected in Germany, in 2012 [2]. Only partial genome sequence (4,109 nt) has been determined, including the partial L gene [2].  Lepidopteran rhabdo-related virus 33 (LeRRV-33; strain OKIAV33) was detected in the TSA of the speckled wood butterfly (*Pararge aegeria*) collected in Germany, in 2011 [2]. Only partial genome sequence (2,016 nt) has been determined, including the partial L gene [2].  Based on the available L protein sequences, each of these viruses clusters phylogenetically with the betapaprhaviruses. However, as only partial genome sequences are currently available, they are not proposed for classification at this time.    **Genome organization and expression**  Betapaprhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) as well as an additional gene (*U1*) between the *G* and *L* genes (**Figure 1**). The *U1* genes encode small basic proteins (12.6 -13.6 kDa) with identifiable sequence homology (**Figure 1**).  A Clustal X alignment indicates that betapaprhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates conservation of six of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 8**). Only disulphide bridges CI-CXII, CIII-CV and CVIII-CX appear to be conserved in the betapaprhaviruses.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, betapaprhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergence between SfruRV and LeRRV-34 in pair-wise alignments (p-distances) are 68.6% in the N proteins, 64.9% in the G proteins and 54.0% in the L protein.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Betapaprhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C and F. The viruses have similar genome organisations (criterion D). As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Betapaprhavirus* is derived from the beta group of lepidopteran (*papilionem*, latin butterfly) rhabdoviruses.  **Type species**  *Frugiperda betapaprhavirus* is designated as the type species of the genus as Spodoptera frugiperda rhabdovirus virus was the first of proposed members of the genus to have been reported and is present in insect cell cultures.   1. **Genus *Betaricinrhavirus***   The new genus *Betaricinrhavirus* comprises viruses that have been detected in hard ticks (Ixodidae). They are distant phylogenetically from those rhabdoviruses detected in hard ticks of various species that are to be assigned to the genus *Alpharicinrhavirus* (see proposal 2020.0101M).  **Viruses to be assigned**  Chimay rhabdovirus (CRV; strain Chimay-1) was detected in a pool of hard ticks (*Ixodes ricinus*) sampled in Belgium in 2009. The near-complete genome sequence (13,706 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini. We propose to assign Chimay rhabdovirus to the new species *Chimay betaricinrhavirus*.  Blacklegged tick rhabdovirus 1 (BLTRV-1; strain RTS95.16) was detected in pools of hard ticks (*Ixodes scapularis*) sampled in Connecticut and New York, USA, in 2016 [9]. The near-complete genome sequence (13,841 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini [9]. We propose to assign blacklegged tick rhabdovirus 1 to the new species *Scapularis betaricinrhavirus*.  No isolates are currently available for either of these viruses.  **Other related viruses**  Fairlight virus (FLTV) and Quarantine Head virus (QHV) were each detected in pools of hard ticks (*Amblyomma moreliae*) sampled from a bluetongue lizard in New South Wales, Australia in 2016 [1]. Only partial genome sequences comprising complete or near-complete G gene and L gene sequences are presently available for these viruses [1].  Although QHV and FLTV cluster phylogenetically with CRV and BLTRV-1, they are not proposed for classification at this time.  **Genome organization and expression**  Betaricinrhavirus genomes contain the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*). In the *N* genes of each virus, a small open reading frame (Nx) overlaps the N ORF. In the *P* genes, an alternative open reading frame (Px) occurs within the P ORF (**Figure 1**).  A Clustal X alignment indicates that betaricinrhavirus G proteins share identifiable sequence identity and are similar in structure and length. Alignment with the G protein of vesicular stomatitis Indiana virus (VSIV) indicates likely conservation of 10 of the 12 cysteine residues that form disulphide bonds in the folded protein [7, 10] (**Figure 9**). Only disulphide bridge CIX-CXI does not appear to be conserved in the betaricinrhaviruses. Additional cysteine residues in CRV, QHV and FLTV G proteins appear to be unpaired.  **Phylogeny and sequence relationships**  Based on ML trees generated from complete L protein sequences, betaricinrhaviruses form a well-supported monophyletic clade that is distinct from all currently assigned genera and other currently unassigned rhabdoviruses (**Figure 2**). Amino acid sequence divergences between CRV and BLTRV-1 in pair-wise alignments (p-distances) are 58.6% in the N proteins, 56.5% in the G proteins and 23.1% in the L protein (**Table1**).  **Species demarcation criteria**  Viruses assigned to different species within the genus *Betaricinrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 10% in N proteins; B) minimum sequence divergence of 10% in the L proteins; C) minimum amino acid sequence divergence of 15% in G 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.  The proposed members of the new genus meet demarcation criteria A, B, C and F. The viruses have similar genome organisations (criterion D). As no virus isolates are currently available neutralisation tests have not been conducted (criterion E).  **Derivation of the genus name**  *Betaricinrhavirus* is derived from the beta group of tick rhabdoviruses derived from *ricinus* (Latin, tick). The alpha group of tick rhabdoviruses which form the proposed new genus *Alpharicinrhavirus* (in the proposed new subfamily *Alpharhabdovirinae*) are phylogenetically distant from the betaricinrhaviruses.  **Type species**  *Chimay betaricinrhavirus* is designated as the type species of the genus as Chimay rhabdovirus was the first of proposed members of the genus to have been reported. | |

**Supporting evidence**

****

**Figure 1.** Genome organisations of viruses to be assigned to seven new genera. Each genome contains long open reading frames (ORFs) in the N, P, M, G and L genes (open arrows). Additional long ORFs occur in some viruses either as additional genes or as alternative or overlapping reading frames. ORFs encoding cognate proteins are shown in the same colours.

****

**Figure 2.** The evolutionary history was inferred from a Clustal W alignment of 247 complete L protein sequences of 165 animal rhabdoviruses currently assigned or recently proposed for assignment to species in the proposed subfamilies *Alpharhabdovirinae*, *Betarhabdovirinae* and *Gammarhabdovirinae*, and unclassified rhabdoviruses that are proposed here to be assigned to seven new genera. Phylogenetically informative sites were selected from the alignment using Gblocks resulting in 398 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 (-78271.98) 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-Joining 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.

****

**Figure 3.** A Clustal X alignment of the VSIV G protein with the alphacrustrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and alphacrustrhavirus G proteins are shaded in blue.

****

**Figure 4.** A Clustal X alignment of the VSIV G protein with the alphadrosrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and alphadrosrhavirus G proteins are shaded in blue.

****

**Figure 5.** A Clustal X alignment of the VSIV G protein with the alphahymrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and alphahymrhavirus G proteins are shaded in blue.

****

**Figure 6.** A Clustal X alignment of the VSIV G protein with the betahymrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and betahymrhavirus G proteins are shaded in blue.

****

**Figure 7**. A Clustal X alignment of the VSIV G protein with the betanemrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and betanemrhavirus G proteins are shaded in blue.

****

**Figure 8**. A Clustal X alignment of the VSIV G protein with the betapaprhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and betapaprhavirus G proteins are shaded in blue.

** **

****

**Figure 9**. A Clustal X alignment of the VSIV G protein with the betaricinrhavirus G proteins. Conserved cysteine residues in the VSIV G protein are marked (CI-CXII). Cysteine residues that are likely to be conserved in VSIV and betaricinrhavirus G proteins are shaded in blue. The N-terminal region of the QHV G protein sequence appears to be incomplete.

**Table 1.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alphahymrhavirus N protein sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | LnegV-2 | HyRRV-38 | HyRRV-46 | HyRRV-109 |
| LnegV-2 |  |  |  |  |
| HyRRV-38 | 36.2 |  |  |  |
| HyRRV-46 | 27.3 | 24.2 |  |  |
| HyRRV-109 | 24.2 | 23.9 | 58.1 |  |

**Table 2.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alphahymrhavirus G protein sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | LnegV-2 | HyRRV-38 | HyRRV-46 | HyRRV-109 |
| LnegV-2 |  |  |  |  |
| HyRRV-38 | 24.4 |  |  |  |
| HyRRV-46 | 18.7 | 21.8 |  |  |
| HyRRV-109 | 18.7 | 20.3 | 42.5 |  |

**Table 3**. Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alphahymrhavirus L protein sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | LnegV-2 | HyRRV-38 | HyRRV-46 | HyRRV-109 |
| LnegV-2 |  |  |  |  |
| HyRRV-38 | 49.5 |  |  |  |
| HyRRV-46 | 41.9 | 40.0 |  |  |
| HyRRV-109 | 41.3 | 40.2 | 64.7 |  |

**Table 4**. Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of betahymrhavirus N protein sequences.

|  |  |  |  |
| --- | --- | --- | --- |
|  | HyRRV-22 | HyRRV-23 | HyRRV-24 |
| HyRRV-22 |  |  |  |
| HyRRV-23 | 93.9 |  |  |
| HyRRV-24 | 51.8 | 52.1 |  |

**Table 5**. Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of betahymrhavirus G protein sequences.

|  |  |  |  |
| --- | --- | --- | --- |
|  | HyRRV-22 | HyRRV-23 | HyRRV-24 |
| HyRRV-22 |  |  |  |
| HyRRV-23 | 89.2 |  |  |
| HyRRV-24 | 47.5 | 47.7 |  |

**Table 6**. Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of betahymrhavirus L protein sequences.

|  |  |  |  |
| --- | --- | --- | --- |
|  | HyRRV-22 | HyRRV-23 | HyRRV-24 |
| HyRRV-22 |  |  |  |
| HyRRV-23 | 92.8 |  |  |
| HyRRV-24 | 58.0 | 57.7 |  |

**References**

1. Harvey E, Rose K, Eden JS, Lo N, Abeyasuriya T, Shi M, Doggett SL, Holmes EC (2019) Extensive diversity of RNA viruses in Australian ticks. J Virol 93:e01358-01318. PMID: 30404810 DOI: 10.1128/JVI.01358-18

2. 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

3. Kleanthous E, Olendraite I, Lukhovitskaya NI, Firth AE (2019) Discovery of three RNA viruses using ant transcriptomic datasets. Arch Virol 164:643-647. PMID: 30415391 DOI: 10.1007/s00705-018-4093-2

4. 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

5. 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

6. Ma H, Galvin TA, Glasner DR, Shaheduzzaman S, Khan AS (2014) Identification of a novel rhabdovirus in *Spodoptera frugiperda* cell lines. J Virol 88:6576-6585. PMID: 24672045 DOI: 10.1128/JVI.00780-14

7. Roche S, Bressanelli S, Rey FA, Gaudin Y (2006) Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. Science 313:187-191. PMID: 16840692 DOI: 10.1126/science.1127683

8. 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

9. Tokarz R, Sameroff S, Tagliafierro T, Jain K, Williams SH, Cucura DM, Rochlin I, Monzon J, Carpi G, Tufts D, Diuk-Wasser M, Brinkerhoff J, Lipkin WI (2018) Identification of novel viruses in *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks. mSphere 3:e00614-00617. PMID: 29564401 DOI: 10.1128/mSphere.00614-17

10. Walker PJ, Kongsuwan K (1999) Deduced structural model for animal rhabdovirus glycoproteins. Journal of General Virology 80:1211-1220. PMID: 10355768 DOI: 10.1099/0022-1317-80-5-1211

11. Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, Quintana JF, Brouqui JM, Bayne EH, Longdon B, Buck AH, Lazzaro BP, Akorli J, Haddrill PR, Obbard DJ (2015) The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. PLoS Biology 13:e1002210. PMID: 26172158 DOI: 10.1371/journal.pbio.1002210