

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

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | **2022.002M** |  |
| **Short title:** Create one new genus (*Thriprhavirus*) and 14 new species in the subfamily *Alpharhabdovirinae* (*Mononegavirales*: *Rhabdoviridae*) | | |
|  | | |

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

|  |  |
| --- | --- |
| Walker PJ, Freitas-Astúa J, Bejerman N, Dietzgen RG, Fooks A, Kondo H, Kurath G, Ramos-González PL, Tesh RB, Tordo N, Vasilakis N, Whitfield AE | peter.walker@uq.edu.au; juliana.astua@embrapa.br; nicobejerman@gmail.com; r.dietzgen@uq.edu.au; tony.fooks@apha.gsi.gov.uk; hkondo@rib.okayama-u.ac.jp; gkurath@usgs.gov; plrg1970@gmail.com; rbtesh22@gmail.com; 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**

|  |
| --- |
| Minor typographical errors only |

**ICTV Study Group votes on proposal**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Group** | **Number of members** | | |
| **Votes support** | **Votes against** | **No vote** |
| *Rhabdoviridae* SG | 12 | 0 | 2 |

**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 | 27 May 2022 |
| 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**

|  |
| --- |
| 2022.002M.N.v1.Alpharhabdovirinae\_1ngen14nsp.xlsx |

**Abstract**

|  |
| --- |
| One new genus (*Thriprhavirus*) in the subfamily *Alpharhabdovirinae* is proposed to accommodate two new species for rhabdoviruses detected in thrips. In addition, we propose the creation of two new species in the genus *Tupavirus*, one new species in the genus *Almendravirus*, two new species in the genus *Alphanemrhavirus,* one new species in the genus *Tibrovirus*, one new species in the genus *Alpharicinrhavirus,* two new species in the genus *Ledantevirus* and three new species in the genus *Sigmavirus.* All viruses to be assigned to new species were detected by metagenomic sequencing and no virus isolates are yet available. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | 1. **Create one new genus *Thriprhavirus* including two new species**   Thrips tabaci associated dimarhabdovirus 1 (TtaDRV-1; strain Tamono1) was detected by metagenomic sequencing of onion thrips (*Thrips tabaci*) collected in Italy in 2018 [1]. Hangzhou Frankliniella intonsa rhabdovirus 1 (HFinRV-1; strain JM1FY86115) was detected by metagenomic sequencing of European flower thrips (*Frankliniella intonsa*) collected in China in 2016. Hubei dimarhabdovirus 4 (HbDRV-4; strain SCM46012) was detected by metagenomic sequencing of a mixed library of flies (Diptera) collected in China in 2013 [5].  We propose to establish a new genus (*Thriprhavirus*) to accommodate these viruses with TtaDRV-1 to be assigned to the new species *Thriprhavirus* *tabaci* and HFinRV-1 to be assigned to the new species *Thriprhavirus intonsa*. HbDRV-4 is considered to be a separate detection of HFinRV-1 (see below).  Genome organizations  The near-complete genome sequences of TtaDRV-1 (12,836 nt) and HFinRV-1 (13,196 nt) are available, lacking only extreme 3' and 5' termini (**Figure 1**). Each contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) and an alternative ORF at the start of the *G* gene encoding a small predicted transmembrane protein with the structural characteristics of a class I viroporin. The genomes also feature a very long 3' untranslated region in the G gene which varies significantly in length from 393 nt in TtaDRV-1 to 1,111 nt in HFinRV-1. Only a partial genome sequence (6,918 nt) is available for HbDRV-4 (*L* gene only) [5].  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, TtaDRV-1, HFinRV-1 and HbDRV-4 cluster form a distinct well-supported clade within the *Alpharhabdovirinae* but separate from all other clades of viruses assigned to other genera (**Figure 3**). Poor bootstrap support indicates that its exact position within the *Alpharhabdovirinae* clade is currently ambiguous.  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that HFinRV-1 and TtaDRV-1 share 72.6% identity in the L protein, 47.2% identity in the G protein and 51.1% identity in the N protein (**Table 1).** HbDRV-4 shares99.8% identity withHFinRV-1 in the L protein and so these are considered to be separate detections of the same virus.  Ecology  These viruses weredetected in thrips of different genera collected from different continents (Europe and Asia). No other rhabdovirus has yet been reported from thrips.  Species demarcation criteria  Similar to the criteria used for several other genera within the *Alpharhabdovirinae*, we propose that viruses assigned to different species within the genus *Thriprhavirus* should have several of the following characteristics: A) minimum amino acid sequence divergence of 12% in the G protein; B) minimum amino acid sequence divergence of 8% in the L protein; C) minimum amino acid sequence divergence of 4% in the N protein; D) can be distinguished in virus neutralization tests; E) exhibit significant differences in genome organization as evidenced by numbers and locations of ORFs; and F) occupy different ecological niches as evidenced by differences in hosts and or arthropod vectors.  HFinRV-1 and TtaDRV-1 meet criteria A, B, C and E. No neutralization test data are yet available as there are currently no isolates of these viruses. The genome organisations of HFinRV-1 and TtaDRV-1 are to be similar but feature long untranslated regions in the G mRNA which vary greatly in length. As the viruses have been detected only by metagenomic sequencing, their natural ecology is uncertain but they  have been detected in thrips of different genera and from different continents.   1. **Create two new species in the genus *Tupavirus***   Wufeng Rhinolophus pearsonii tupavirus 1 (WRpeTV-1; strain WFB\_Rpear) was detected by metagenomic sequencing of Pearson’s horseshoe bats (*Rhinolophus pearsonii*) collected in 2016 in China. Wenzhou Myotis laniger tupavirus 1 (WMlaTV-1; strain YJB\_HuaNan) was detected by metagenomic sequencing of Chinese water myotis bats (*Myotis laniger*) collected in 2016 in China.  We propose to assign WRpeTV-1to the new species *Tupavirus* *pearsonii* and WMlaTV-1 to the new species *Tupavirus laniger*.  Genome organizations  The near-complete genome sequences of WRpeTV-1 (11,829 nt) and WMlaTV-1 (11,434 nt) are available, lacking only extreme 3' and 5' termini (**Figure 1**). Each contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*), an alternative ORF in the *P* gene and a gene encoding a small protein between the *M* gene and *G* gene. Like Klamath virus (KLAV; species *Tupavirus klamath*), WRpeTV-1 also has a gene encoding a small protein between the *G* gene and *L* gene.  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, WRpeTV-1 and WMlaTV-1 cluster within the tupavirus clade (**Figure 3**) and are most closely related to tupaia rhabdovirus (TUPV; species *Tupavirus tupaia*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that WRpeTV-1 and WMlaTV-1 share 63.6% identity in the L protein, 50.9% identity in the G protein and 66.7% identity in the N protein. Each is most closely related to TUPV with which they display 61.7% – 62.8% identity in the L protein, 40.3% -– 41.9% identity in the G protein and 67.2% – 70.2% identity in the N protein (**Tables 2–4**).  Ecology  Tupaviruses have been isolated from various mammals and birds. These represent the first detection of tupaviruses in bats. There were detected in chiroped bats of two different genera, each from China.  Species demarcation criteria  Viruses assigned to different species within the genus *Tupavirus* 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 G proteins; D) significant differences in genome organization as evidenced by numbers and locations of ORFs; E) can be distinguished in serological tests; and F) occupy different ecological niches as evidenced by differences in hosts and or arthropod vectors.  Both proposed members of the genus meet demarcation criteria A, B and C. Neutralisation tests have not been conducted as there are currently no isolates of these viruses. Their genome organisations are similar to those of other tupaviruses. As the viruses have been detected only by metagenomic sequencing, their natural ecology is uncertain but they have been detected in bats of different genera.   1. **Create two new species in the genus *Alphanemrhavirus***   Sodak rhabdovirus 1 (SDRV-1; strain 20-13111) was detected by metagenomic sequencing of visceral homogenates from big brown bats (*Eptesicus fuscus*) collected in 2020 in South Dakota, USA [3]. It has been recognized by those reporting the virus that it may well have originated from internal nematodes parasitizing the bats [3].  Rattus tanezumi rhabdovirus 1 (RtaRV-1; strain TB2018B) was detected by metagenomic sequencing of the lung tissue of Asian house rats (*Rattus tanezumi*) collected in 2018 from Bangkok, Thailand [6]. A second strain of the virus (TB2018A) was detected in rats of the same species from the same location [6]. Coding complete sequences are available for each of these strains. Partial genome sequences (complete or partial L gene) were obtained for several other strains isolated in Thailand from rats of the same species or from brown rats (*Rattus norvegicus*) [6]. These all display high sequence identity in the available sequence fragments and appear to represent strains of the same virus. Partial L protein sequence has also been reported for a related but distinct virus detected by metagenomic sequencing of the lung tissue of a chestnut white-bellied rat (*Niviventer fulvescens*) collected in 2006 from Loei, Thailand [6]; this virus is not proposed for classification in the absence of the coding-complete sequence.  We propose to assign SDRV-1 to the new species *Alphanemrhavirus sodak* and RtaRV-1 to the new species *Alphanemrhavirus bangkok*.  Genome organizations  The near-complete genome sequences of SDRV-1 (11,221 nt) and RtaRV-1 (10,660 nt) are available, lacking only extreme 3' and 5' termini (**Figure 1**). Each contains only the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*).  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, SDRV-1 and RtaRV-1 cluster within the alphanemrhavirus clade (**Figure 3**) and are most closely related to Xingshan nematode virus 4 (XsNV-4; species *Alphanemrhavirus xingshan*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that SDRV-1 shares 46.1% – 53.0% identity with other alphanemrhaviruses in the L protein, 18.3% – 20.3% identity in the G protein and 34.1% – 47.2% identity in the N protein. RtaRV-1 shares 47.0% – 52.5% identity with other alphanemrhaviruses in the L protein, 19.8% – 23.0% identity in the G protein and 38.0% – 45.3% identity in the N protein. The two strains of RtaRV-1 share >99% identity in all three proteins. The L and N proteins of unclassified Xinzhou dimarhabdovirus 1 (XzDRV-1), which lacks a coding complete genome sequence, are also included for comparative purposes (**Tables 5–7**).  Ecology  Alphanemrhaviruses have been detected in nematode worms. The sequences of SDRV-1 and RtaRV-1 were obtained from visceral homogenates of bats and rats, respectively, and may well have originated from parasitizing nematode worms.  Species demarcation criteria  Viruses assigned to different species within the genus *Alphanemrhavirus* 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 organization 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.  Both proposed members of the genus meet demarcation criteria A, B and C. Neutralisation tests have not been conducted as there are currently no isolates of these viruses. The genome organisations are similar to those of other alphanemrhaviruses. As the viruses have been detected only by metagenomic sequencing, their natural ecology is uncertain.   1. **Create one new species in the genus *Tibrovirus***   Mundri virus (MUNV; strain A14) was detected in the plasma and cerebrospinal fluid of a 15 year-old female human with new-onset nodding syndrome in South Sudan in 2018 [2]. Surveys using RT-qPCR to detect the virus and antibody surveys using anti-N protein IgG did not reveal any clear causal association of the virus with new-onset nodding syndrome [2].  We propose to assign MUNV to the new species *Tibrovirus mundri*.  Genome organization  The near-complete genome sequence of MUNV (12,445 nt) is available, lacking only extreme 3' and 5' termini (**Figure 1**). Like other tibroviruses, it contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) as well as two additional genes (*U1* and *U2*) of unknown function between the *M* gene and *G* gene, and one additional gene (*U3*) encoding a small protein with the predicted structure of a class I viroporin between the *G* gene and *L* gene. Unlike most other tibroviruses, MUNV lacks an alternative ORF in the *P* gene but uniquely contains an alternative long ORF in the *M* gene.  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, MUNV clusters within the tibrovirus clade (**Figure 3**) and are most closely related to Ekpoma virus 2 (EKV2; species *Tibrovirus betaekpoma*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that MUNV is most closely related to EKV2 with which it shares 58.6% identity in the L protein, 49.9% identity in the G protein and 57.6% identity in the N protein. (**Tables 8–10**).  Ecology  Based on L protein sequences, tibroviruses fall phylogenetically into two clades comprising: (a) viruses associated with cattle and/or biting midge (*Culicoides* sp.) vectors; and (b) viruses detected in human tissue samples by metagenomic sequencing. MUNV falls into the second clade and represents the fourth distinct tibrovirus detected in humans.  Species demarcation criteria  Tibrovirus species demarcation criteria are based on low-level or no cross-reaction in virus-neutralisation tests supported by phylogenetic analysis and genetic diversity estimations using L and N gene sequences to establish that the species represents a distinct lineage. Typically, there will be <5% amino acid sequence diversity (divergence) within species and >20% diversity (divergence) between species.  The proposed new member of the genus meets the sequence divergence demarcation criteria. Neutralisation tests have not been conducted as there is currently no isolate of the virus.   1. **Create one new species in the genus *Alpharicinrhavirus***   Hubei tick rhabdovirus 1 (HbTRV-1; strain RhV/SZWH3) was detected by metagenomic sequencing of Asian long-horned ticks (*Haemaphysalis longicornis*) collected in 2019 in Hubei Province, China [7].  We propose to assign HbTRV-1 to the new species *Alpharicinrhavirus hubei*.  Genome organization  The near-complete genome sequence of HbTRV-1 (11,712 nt) is available, lacking only extreme 3' and 5' termini (**Figure 1**). Each contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) and, uniquely amongst alpharicinrhaviruses, has an alternative long ORF overlapping the end of the G ORF.  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, HbTRV-1 clusters within the alpharicinrhavirus clade (**Figure 3**) and is most closely related to Bole tick virus 2 (BlTV-2; species *Alpharicinrhavirus bole*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that HbTRV-1 is most closely related to BlTV-2 with which it shares 63.2% identity in the L protein, 43.1% identity in the G protein and 42.2% identity in the N protein. (**Tables 11–13**).  Ecology  Alpharicinrhaviruses have been identified by metagenomic sequencing only and all have been detected in hard ticks (Acari: Ixodidae). HbTRV-1 is the first alpharicinrhavirus to have been detected in Asian long-horned ticks (*Haemaphysalis longicornis*); two probable members of the genus (incomplete genome sequences) have been detected in ticks of other *Haemaphysalis* species in China.  Species demarcation criteria  Viruses assigned to different species within the genus *Alpharicinrhavirus* 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 organization 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 new member of the genus meets demarcation criteria A, B and C. Although there is, uniquely, an alternative ORF overlapping the end of the G ORF, it is not known if this is expressed. Neutralisation tests have not been conducted as there is currently no isolate of the virus. As the virus has been detected only by metagenomic sequencing, its natural ecology is uncertain but it has been detected in bats of a novel species.   1. **Create one new species in the genus *Almendravirus***   Xiangshan rhabdo-like virus 1 (XsRLV-1; strain Novel\_23) was detected by metagenomic sequencing of a mixed sample of insects (Diptera, Hymenoptera, Lepidoptera) collected in 2020 in Beijing, China (BioProject PRJNA728541).  We propose to assign XsRLV-1 to the new species *Almendravirus xianshan*.  Genome organization  The near-complete genome sequence of XsRLV-1 (10,798 nt) is available, lacking only extreme 3' and 5' termini (**Figure 1**). Like those of other almendraviruses, the genome contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) and an additional gene between the *G* gene and *L* gene encoding a small protein with the predicted structure of a class I viroporin.  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, XsRLV-1 clusters within the almendravirus clade (**Figure 3**) and are most closely related to Rio Chico virus (RCHV; species *Almendravirus chico*.  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that XsRLV-1 is most closely related to RCHV with which it shares 62.9% identity in the L protein, 37.8% identity in the G protein and 57.7% identity in the N protein (**Tables 14–16**).  Ecology  Almendraviruses have been isolated exclusively from mosquitoes in China and the Americas. The detection of XsRLV-1 in a mixed sample of insects (Diptera, Hymenoptera, Lepidoptera) suggests that it may also be of mosquito origin but the specific insect host has not been identified.  Species demarcation criteria  Viruses assigned to different species within the genus *Almendravirus* 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 G proteins; D) significant differences in genome organization 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 hosts and or arthropod vectors.  The proposed new member of the genus meets demarcation criteria A, B and C. The genome organization is similar to those of other almendraviruses. Neutralisation tests have not been conducted as there is currently no isolate of the virus. As the virus has been detected only by metagenomic sequencing of a mixed pool of insects, its natural ecology is uncertain.   1. **Create two new species in the genus *Ledantevirus***   Wenzhou Rhinolophus pusillus ledantevirus 1 (WRpuLV-1; strain YJB\_DanJiao) was detected by metagenomic sequencing of least horseshoe bats (*Rhinolophus pusillus*) collected in China in 2016. Longquan Niviventer coninga ledantevirus 1 (LNco LV-1; strain LQS\_baifu) was detected by metagenomic sequencing of Edward’s long-tailed giant rats (*Leopoldamys edwardsi*) collected in China in 2016.  We propose to assign WRpuLV-1 to the new species *Ledantevirus wenzhou* and LNcoLV-1 to the new species *Ledantevirus longquan*.  Genome organizations  The near-complete genome sequences of WRpuLV-1 (11,195 nt) and LNcoLV-1 (10,853 nt) are available, lacking only extreme 3' and 5' termini (**Figure 2**). Like most other ledanteviruses of phylogroups A and C, each contains only the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*).  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, WRpuLV-1 and LNcoLV-1 cluster within the ledantevirus clade (**Figure 3**). WRpuLV-1 is most closely related to Oita virus (OITAV; species *Ledantevirus oita*); LNcoLV-1 clusters with Nkolbisson virus (NKOV; species *Ledantevirus nkolbisson*) and Yongjia tick virus 2 (YjTV-2; species *Ledantevirus yongjia*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that WRpuLV-1 is most closely related to OITAV with which it displays 74.1% identity in the L protein, 58.0% identity in the G protein and 71.2% identity in the N protein (**Tables 17–19**). LNcoLV-1 is most closely related to NKOV, Barur virus (BARV; species *Ledantevirus barur*), Fukuoka virus (FUKV; species *Ledantevirus fukuoka*) and Nishimuro virus (NISV; species *Ledantevirus nishimuro*) with which it displays 55.1% – 56.0% identity in the L protein, 41.2% – 43.4% identity in the G protein and 52.0% – 53.0% identity in the N protein (**Tables 17–19**).  Ecology  WRpuLV-1 (detected in least horseshoe bats) falls within ledantevirus phylogroup C which comprises viruses that have been isolated primarily from bats (Chiroptera). LNcoLV-1 (detected in Edward’s long-tailed giant rats) falls within ledantevirus phylogroup A which comprises viruses isolated from rats, pigs, cattle or associated arthropod vectors (mosquitoes, midges and ticks).  Species demarcation criteria  Viruses assigned to different species within the genus *Ledantevirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 7% in L; B) minimum amino acid sequence divergence of 15% in G; C) significant differences in genome organisation as evidenced by numbers and locations of ORFs; D) can be distinguished in serological tests; and E) occupy different ecological niches as evidenced by differences in hosts and/or arthropod vectors.  WRpuLV-1 and LNcoLV-1 meet criteria A and B. The genome organisations of WRpuLV-1 and LNcoLV-1 are similar to those of ledanteviruses from the same phylogroups. Neutralization tests have not been conducted as there are currently no isolates of these viruses. As the viruses have been detected only by metagenomic sequencing their natural ecology is uncertain but ledanteviruses have not been reported previously in mammals of these species.   1. **Create three new species in the genus *Sigmavirus***   Yushu rhabdovirus (YsRV; strain YSN900) was detected by metagenomic sequencing of a pool of bird fecal samples collected in 2019 in the Yushu Tibetan Autonomous Prefecture, Qinghai Province, China (BioProject PRJNA706129). Apis rhabdovirus 3 (ApRV-3; strain Sichuan/2019) was detected by metagenomic sequencing of eastern honey bees (*Apis cerana*) collected in 2019 in Sichuan Province, China (BioProject PRJNA706851). Aksy-Durug Melophagus sigmavirus (ADMSV; strain 13577) was detected by metagenomic sequencing of sheep keds (*Melophagus ovinus*) collected in 2012 in the Republic of Tuva, Russia [4].  We propose to assign YsRV to the new species *Sigmavirus yushu,* ApRV-3 to the new species *Sigmavirus sichuan* and ADMSV to the new species *Sigmavirus tuva*.  Genome organizations  The near-complete genome sequences of YsRV (12,485 nt), ApRV-3 (13,131 nt) and ADMSV (11,712 nt) are available, lacking only extreme 3' and 5' termini (**Figure 2**). Each contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*). YsRv and ApRV-3 also contain the *X* gene between the *M* gene and *G* gene which is characteristic of many sigmaviruses. Like several other closely related sigmaviruses, ADMSV lacks the *X* gene.  Phylogenetic analysis  Based on ML trees generated from complete L protein sequences, YsRV, ApRV-3 and ADMSV cluster within the sigmavirus clade (**Figure 3**). YsRV clusters with Shayang fly virus 2 (SyFV-2; species *Sigmavirus shayang*) and Wuhan fly virus 2 (WhFV-2; species *Sigmavirus domestica*); ApRV-3 is more deeply rooted in the same clade; ADMSV lies on a more distantly related clade and clusters with Wuhan louse fly virus 9 (WhLFV-9; species *Sigmavirus hippoboscid*) and Wuhan louse fly virus 10 (WhLFV-10; species *Sigmavirus lousfly*).  Amino acid sequence identities  Pairwise sequence identities (p-distances) calculated in MEGA7 from ClustalW amino acid sequence alignments indicated that TsRV is most closely related to SyFV-2 and WhFV-2 with which it displays 73.6% – 74.0% identity in the L protein, 58.4% – 59.4% identity in the G protein and 65.1% – 65.5%% identity in the N protein (**Tables 20–22**). ApRV-3 is also most closely related to these three viruses as well as Hubei diptera virus 9 (HbDV-9; species *Sigmavirus hubei*) and Hubei diptera virus 10 (HbDV-10; species *Sigmavirus myga*) with which it displays is 45.8% – 49.1% identity in the L protein, 27.9% – 31.4% identity in the G protein and 23.6% – 26.7% identity in the N protein (**Tables 20–22**). ADMSV is most closely related to WhLFV-10 with which it displays is 51.7% identity in the L protein, 22.5% identity in the G protein and 29.4% identity in the N protein (**Tables 20–22**).  Ecology  Sigmaviruses identified to date infect flies (Diptera). The natural host of YsRV is not known but is most likely to be an insect that has been devoured by birds. The natural host of ApRV-3 appears to be bees; if so, this would represent the first detection of a sigmavirus in hymenopteran insects. ADMSV has been detected in sheep keds, a fly of the order Diptera which is a parasite of sheep.  Species demarcation criteria  Viruses assigned to different species within the genus *Sigmavirus*have one or both of the following characteristics: A) minimum amino acid sequence divergence of 10% in L; and B) occupy different ecological niches as evidenced by differences in hosts.  All proposed members of the genus meet demarcation criterion A. As the viruses have been detected only by metagenomic sequencing, their natural ecology is uncertain but bees (ApRV-3) and sheep kerds (ADMSV) appear to represent novel hosts. | |

**Supporting evidence**

****

**Figure 1.** Schematic representation of alpharhabdovirus genomes of five established genera (*Tupavirus*, *Alphanemrhavirus*, *Tibrovirus*, *Alpharicinrhavirus* and *Almendravirus*) and the proposed new genus *Thriprhavirus*. N, P, M, G and L represent ORFs encoding the structural proteins. Additional ORFs encoding homologous proteins are shown in the same colour. ORFs encoding proteins with the structural characteristic of class I viroporins are shown in yellow. Alternative ORFs (shaded grey) of significant length (>180 nucleotides) also occur in some genes but the significance of these is unknown. Viruses representing proposed new species are listed in red text.

****

**Figure 2.** Schematic representation of ledantevirus and sigmavirus genomes. N, P, M, G and L represent ORFs encoding the structural proteins. Ledantevirus genogroups determined from L protein phylogeny are shown. An additional gene that occurs only in phylogroup B ledanteviruses is coloured brown. An alternative long ORF that occurs in the M gene of FUKV only is shown in dark blue. Homologous genes (*X*) that occur in the genomes of most sigmaviruses are shown in light blue. Viruses representing proposed new species are listed in red text.

****

**Figure 3.** The evolutionary history was inferred from a MUSCLE alignment of complete L protein sequences of 189 rhabdoviruses that are currently assigned to species in the subfamily *Alpharhabdovirinae* as well as 16 viruses proposed to be assigned to 14 new species in the subfamily. Phylogenetically informative sites were selected from the alignment using Gblocks resulting in 930 positions in the final dataset. The tree was inferred in MEGA11 by using the Maximum Likelihood method based on the best-fit Le and Gascuel model with gamma distribution of evolutionary rates and invariable sites. The tree with the highest log likelihood (-131580.18) 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 values. 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 identities (p-distance) of a CLUSTAL W alignments of thriprhavirus N, G and L protein sequences.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Viruses | N proteins | | G proteins | | L proteins | | |
| HFiRV-1 | TTaDRV-1 | HFiRV-1 | TTaDRV-1 | HFiRV-1 | TTaDRV-1 | HbDRV-4 |
| HFiRV-1 |  |  |  |  |  |  |  |
| TTaDRV-1 | 51.1 |  | 47.2 |  | 72.6 |  |  |
| HbDRV-4 | n.a. | n.a. | n.a. | n.a. | 99.8 | 72.6 |  |

n.a. – not available

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Virus | DURV | KLAV | TUPV | WMlaTV-1 | WRpeTV-1 |
| DURV |  |  |  |  |  |
| KLAV | 52.5 |  |  |  |  |
| TUPV | 52.6 | 56.5 |  |  |  |
| WMlaTV-1 | 53.5 | 58.9 | 61.7 |  |  |
| WRpeTV-1 | 53.7 | 57.7 | 62.8 | 63.6 | 00.0 |

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Virus | DURV | KLAV | TUPV | WMlaTV-1 | WRpeTV-1 |
| DURV |  |  |  |  |  |
| KLAV | 19.2 | 10.0 |  |  |  |
| TUPV | 25.5 | 28.2 |  |  |  |
| WMlaTV-1 | 24.9 | 26.0 | 40.3 |  |  |
| WRpeTV-1 | 26.9 | 29.4 | 41.9 | 50.9 |  |

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Virus | DURV | KLAV | TUPV | WMlaTV-1 | WRpeTV-1 |
| DURV |  |  |  |  |  |
| KLAV | 47.1 | 100. |  |  |  |
| TUPV | 57.1 | 55.1 |  |  |  |
| WMlaTV-1 | 53.6 | 52.1 | 70.2 |  |  |
| WRpeTV-1 | 55.2 | 52.6 | 67.2 | 66.7 | 100.0 |

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Virus | XsNV-4 | SDRV-1 | RtaRV-1\_A | RtaRV-1\_B | XzDRV-1 | XzNV-4 |
| XsNV-4 |  |  |  |  |  |  |
| SDRV-1 | 53.0 |  |  |  |  |  |
| RtaRV-1\_A | 51.7 | 52.4 |  |  |  |  |
| RtaRV-1\_B | 51.7 | 52.5 | 99.7 |  |  |  |
| XzDRV-1 | 47.3 | 48.9 | 49.9 | 50.0 |  |  |
| XzNV-4 | 45.8 | 46.1 | 47.0 | 47.0 | 49.1 |  |

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Virus | XsNV-4 | SDRV-1 | RtaRV-1\_A | RtaRV-1\_B | XzDRV-1 | XzNV-4 |
| XsNV-4 |  |  |  |  |  |  |
| SDRV-1 | 18.3 |  |  |  |  |  |
| RtaRV-1\_A | 23.0 | 20.0 |  |  |  |  |
| RtaRV-1\_B | 22.8 | 20.2 | 99.4 |  |  |  |
| XzDRV-1 | 21.7 | 17.8 | 19.8 | 19.8 |  |  |
| XzNV-4 | 21.0 | 20.3 | 22.1 | 22.3 | 22.2 |  |

**Table 7.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alphanemrhavirus N protein sequences.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Virus | XsNV-4 | SDRV-1 | RtaRV-1\_A | RtaRV-1\_B | XzNV-4 |
| XsNV-4 |  |  |  |  |  |
| SDRV-1 | 47.2 |  |  |  |  |
| RtaRV-1\_A | 45.3 | 41.9 |  |  |  |
| RtaRV-1\_B | 45.3 | 41.9 | 100 |  |  |
| XzNV-4 | 35.0 | 34.1 | 38.0 | 38.0 |  |

**Table 8.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of tibrovirus L protein sequences.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BHV | TIBV | SWBV | CPV | EKV1 | BASV | EKV2 | MUNV |
| BHV | 10.0 |  |  |  |  |  |  |  |
| TIBV | 87.1 |  |  |  |  |  |  |  |
| SWBV | 65.3 | 65.1 |  |  |  |  |  |  |
| CPV | 56.2 | 57.3 | 57.0 |  |  |  |  |  |
| EKV1 | 46.0 | 46.0 | 46.7 | 48.6 |  |  |  |  |
| BASV | 43.1 | 43.3 | 43.6 | 43.0 | 44.4 |  |  |  |
| EKV2 | 43.3 | 43.5 | 45.6 | 43.8 | 45.6 | 51.2 |  |  |
| MUNV | 43.9 | 44.1 | 43.6 | 44.3 | 46.3 | 49.4 | 58.6 | ## |

**Table 9.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of tibrovirus G protein sequences.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BHV | TIBV | SWBV | CPV | EKV1 | BASV | EKV2 | MUNV |
| BHV |  |  |  |  |  |  |  |  |
| TIBV | 81.3 |  |  |  |  |  |  |  |
| SWBV | 62.8 | 63.3 |  |  |  |  |  |  |
| CPV | 54.5 | 56.2 | 57.1 |  |  |  |  |  |
| EKV1 | 34.1 | 34.6 | 34.4 | 32.4 |  |  |  |  |
| BASV | 29.1 | 28.6 | 27.9 | 27.6 | 30.0 |  |  |  |
| EKV2 | 30.9 | 30.6 | 29.7 | 30.1 | 28.0 | 34.0 |  |  |
| MUNV | 28.2 | 28.4 | 27.1 | 26.8 | 28.5 | 30.8 | 49.9 |  |

**Table 10.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of tibrovirus N protein sequences.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BHV | TIBV | SWBV | CPV | EKV1 | BASV | EKV2 | MUNV |
| BHV |  |  |  |  |  |  |  |  |
| TIBV | 94.4 |  |  |  |  |  |  |  |
| SWBV | 74.8 | 75.2 |  |  |  |  |  |  |
| CPV | 67.1 | 66.8 | 68.7 |  |  |  |  |  |
| EKV1 | 46.7 | 47.4 | 45.8 | 46.8 |  |  |  |  |
| BASV | 38.9 | 39.9 | 39.4 | 40.6 | 40.5 |  |  |  |
| EKV2 | 39.0 | 39.4 | 38.0 | 39.7 | 41.2 | 45.5 |  |  |
| MUNV | 39.9 | 40.6 | 39.9 | 39.0 | 41.3 | 46.4 | 57.6 |  |

**Table 11.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alpharicinrhavirus L protein sequences.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Virus | WhTV-1 | TcTV-3 | BCOV | BlTV-2 | HbTRV-1 | MLYV |
| WhTV-1 |  |  |  |  |  |  |
| TcTV-3 | 64.1 |  |  |  |  |  |
| BCOV | 46.4 | 48.8 |  |  |  |  |
| BlTV-2 | 47.3 | 48.2 | 56.5 |  |  |  |
| HbTRV-1 | 48.9 | 48.7 | 55.5 | 63.2 |  |  |
| MLYV | 50.3 | 51.1 | 48.2 | 50.1 | 49.3 |  |

**Table 12.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alpharicinrhavirus G protein sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Virus | BCOV | BlTV-2 | HbTRV-1 | MLYV |
| BCOV |  |  |  |  |
| BlTV-2 | 36.1 |  |  |  |
| HbTRV-1 | 36.9 | 43.1 |  |  |
| MLYV | 32.5 | 30.1 | 30.7 |  |

**Table 13.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of alpharicinrhavirus N protein sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Virus | WhTV-1 | BCOV | BlTV-2 | HbTRV-1 |
| WhTV-1 |  |  |  |  |
| BCOV | 24.2 |  |  |  |
| BlTV-2 | 25.5 | 36.7 |  |  |
| HbTRV-1 | 27.0 | 36.8 | 42.2 |  |

**Table 14.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of almendravirus L protein sequences.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | ABTV | PTAMV | CBV | MRV | BALV | RCHV | XsRLV-1 |
| ABTV |  |  |  |  |  |  |  |
| PTAMV | 63.9 |  |  |  |  |  |  |
| CBV | 41.9 | 42.5 |  |  |  |  |  |
| MRV | 41.6 | 41.8 | 58.3 |  |  |  |  |
| BALV | 44.1 | 45.5 | 44.6 | 43.9 |  |  |  |
| RCHV | 45.9 | 46.1 | 45.2 | 45.2 | 54.3 |  |  |
| XsRLV-1 | 45.9 | 46.8 | 44.8 | 44.7 | 52.5 | 62.9 |  |

**Table 15.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of almendravirus G protein sequences.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | ABTV | PTAMV | CBV | MRV | BALV | RCHV | XsRLV-1 |
| ABTV |  |  |  |  |  |  |  |
| PTAMV | 37.0 |  |  |  |  |  |  |
| CBV | 20.7 | 21.9 |  |  |  |  |  |
| MRV | 22.9 | 22.0 | 36.9 |  |  |  |  |
| BALV | 19.7 | 22.8 | 22.3 | 22.8 |  |  |  |
| RCHV | 25.6 | 24.3 | 25.5 | 24.0 | 24.7 |  |  |
| XsRLV-1 | 23.3 | 24.3 | 25.5 | 25.3 | 26.6 | 37.8 |  |

**Table 16.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of almendravirus N protein sequences.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | ABTV | PTAMV | CBV | MRV | BALV | RCHV | XsRLV-1 |
| ABTV |  |  |  |  |  |  |  |
| PTAMV | 61.4 |  |  |  |  |  |  |
| CBV | 24.2 | 23.5 |  |  |  |  |  |
| MRV | 25.1 | 23.0 | 34.7 |  |  |  |  |
| BALV | 25.9 | 25.9 | 26.4 | 28.6 |  |  |  |
| RCHV | 23.3 | 24.5 | 25.5 | 27.5 | 32.5 |  |  |
| XsRLV-1 | 24.5 | 22.7 | 27.9 | 27.8 | 33.5 | 57.7 |  |

**Table 17.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of ledantevirus L protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BARV | FUKV | NISHV | NKOV | LNcoLV-1 | YjTV-2 | KEUV | LDV | VAPV | KCV | MEBV | TYBV | KYAV | BUGV | WhFLV-5 | OITAV | KRV | WRpuLV-1 | FKRV | KOLEV |
| BARV |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FUKV | 92.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NISHV | 87.6 | 89.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NKOV | 61.4 | 61.8 | 61.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LNcoLV-1 | 55.4 | 55.1 | 55.6 | 56.0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YjTV-2 | 54.0 | 53.9 | 54.2 | 54.4 | 53.8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KEUV | 50.5 | 51.2 | 51.0 | 51.1 | 49.2 | 50.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LDV | 50.0 | 50.2 | 49.9 | 50.4 | 49.1 | 50.2 | 80.9 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| VAPV | 50.8 | 50.7 | 50.9 | 50.9 | 49.6 | 49.2 | 69.9 | 69.0 |  |  |  |  |  |  |  |  |  |  |  |  |
| KCV | 50.2 | 50.6 | 50.8 | 51.1 | 49.7 | 50.4 | 62.9 | 62.8 | 63.1 |  |  |  |  |  |  |  |  |  |  |  |
| MEBV | 48.9 | 49.0 | 49.1 | 49.2 | 48.5 | 47.5 | 49.2 | 48.7 | 49.8 | 49.0 |  |  |  |  |  |  |  |  |  |  |
| TYBV | 49.2 | 49.2 | 49.0 | 49.6 | 48.3 | 48.3 | 48.8 | 48.6 | 50.0 | 48.3 | 75.5 |  |  |  |  |  |  |  |  |  |
| KYAV | 49.0 | 48.8 | 48.7 | 48.2 | 48.1 | 48.4 | 48.3 | 47.7 | 48.8 | 48.4 | 63.8 | 63.1 |  |  |  |  |  |  |  |  |
| BUGV | 46.7 | 47.0 | 46.8 | 47.5 | 47.7 | 47.2 | 47.7 | 47.4 | 47.8 | 47.5 | 62.8 | 62.0 | 71.6 |  |  |  |  |  |  |  |
| WhLFV-5 | 47.9 | 47.9 | 48.1 | 47.9 | 47.8 | 46.8 | 49.1 | 48.9 | 50.0 | 48.6 | 61.6 | 61.1 | 59.1 | 57.0 |  |  |  |  |  |  |
| OITAV | 48.0 | 47.9 | 47.8 | 48.1 | 46.9 | 46.8 | 47.7 | 47.6 | 49.9 | 48.7 | 54.7 | 54.9 | 53.5 | 52.6 | 55.7 |  |  |  |  |  |
| KRV | 47.6 | 47.4 | 46.8 | 47.1 | 46.4 | 48.2 | 48.9 | 48.9 | 50.0 | 48.6 | 55.6 | 55.4 | 54.8 | 53.6 | 55.2 | 63.2 |  |  |  |  |
| WRpuLV-1 | 47.2 | 47.7 | 47.6 | 47.0 | 46.8 | 47.0 | 48.6 | 48.1 | 49.7 | 48.0 | 54.9 | 55.1 | 54.2 | 52.9 | 56.2 | 74.1 | 63.2 |  |  |  |
| FKRV | 47.7 | 47.5 | 47.7 | 47.0 | 46.6 | 46.7 | 47.7 | 46.9 | 48.6 | 48.3 | 54.8 | 53.5 | 53.1 | 52.3 | 53.7 | 62.2 | 61.6 | 62.75 |  |  |
| KOLEV | 47.2 | 47.9 | 47.5 | 48.1 | 47.2 | 47.9 | 49.3 | 48.4 | 49.5 | 49.2 | 55.1 | 54.6 | 54.6 | 53.4 | 54.2 | 62.2 | 62.6 | 62.99 | 75.3 | 100 |

**Table 18.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of ledantevirus G protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BARV | FUKV | NISHV | NKOV | LNcoLV-1 | YjTV-2 | KEUV | LDV | VAPV | KCV | MEBV | TYBV | KYAV | BUGV | WhFLV-5 | OITAV | KRV | WRpuLV-1 | FKRV | KOLEV |
| BARV |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FUKV | 84.7 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NISHV | 77.9 | 78.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NKOV | 47.1 | 46.8 | 49.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LNcoLV-1 | 41.2 | 41.7 | 41.6 | 43.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YjTV-2 | 38.8 | 40.8 | 38.6 | 39.8 | 42.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KEUV | 30.2 | 31.1 | 31.2 | 32.3 | 33.5 | 34.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LDV | 30.6 | 31.2 | 30.7 | 31.1 | 32.4 | 33.3 | 69.8 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| VAPV | 33.5 | 33.5 | 33.0 | 33.8 | 31.0 | 36.0 | 57.0 | 58.2 |  |  |  |  |  |  |  |  |  |  |  |  |
| KCV | 34.2 | 35.3 | 33.6 | 32.8 | 32.8 | 36.5 | 48.5 | 49.1 | 51.4 |  |  |  |  |  |  |  |  |  |  |  |
| MEBV | 27.5 | 28.1 | 28.5 | 28.6 | 28.0 | 30.5 | 29.9 | 30.7 | 31.6 | 31.1 |  |  |  |  |  |  |  |  |  |  |
| TYBV | 29.4 | 31.9 | 29.7 | 30.3 | 27.4 | 29.0 | 29.8 | 29.5 | 33.1 | 30.3 | 64.4 |  |  |  |  |  |  |  |  |  |
| KYAV | 29.9 | 29.4 | 29.6 | 29.8 | 27.1 | 29.4 | 29.3 | 29.2 | 29.4 | 30.1 | 47.0 | 47.0 |  |  |  |  |  |  |  |  |
| BUGV | 29.6 | 31.1 | 29.5 | 29.9 | 28.7 | 29.4 | 30.5 | 29.8 | 32.8 | 31.7 | 45.8 | 47.4 | 70.0 |  |  |  |  |  |  |  |
| WhLFV-5 | 28.7 | 28.7 | 29.0 | 30.9 | 29.5 | 28.9 | 29.2 | 29.3 | 30.0 | 30.1 | 43.6 | 44.5 | 43.2 | 41.7 |  |  |  |  |  |  |
| OITAV | 25.5 | 25.0 | 26.1 | 27.5 | 27.3 | 27.4 | 26.2 | 24.3 | 27.1 | 25.5 | 29.6 | 30.3 | 29.6 | 28.7 | 29.0 |  |  |  |  |  |
| KRV | 24.2 | 24.7 | 23.8 | 24.2 | 22.9 | 25.5 | 24.5 | 25.3 | 24.8 | 27.1 | 26.4 | 25.6 | 27.7 | 27.0 | 24.7 | 32.4 |  |  |  |  |
| WRpuLV-1 | 26.7 | 27.1 | 27.9 | 26.8 | 26.4 | 26.0 | 26.6 | 26.0 | 25.5 | 25.3 | 29.4 | 30.7 | 30.0 | 29.3 | 27.5 | 58.0 | 34.3 |  |  |  |
| FKRV | 25.0 | 23.5 | 24.5 | 26.2 | 26.2 | 26.5 | 26.8 | 27.1 | 27.1 | 25.1 | 31.0 | 31.9 | 29.3 | 28.8 | 28.1 | 43.3 | 32.4 | 41.4 |  |  |
| KOLEV | 26.3 | 25.8 | 25.8 | 27.5 | 26.7 | 26.9 | 27.3 | 27.6 | 25.6 | 26.2 | 29.1 | 30.2 | 32.1 | 29.9 | 30.5 | 39.1 | 33.2 | 40.8 | 50.7 |  |

**Table 19.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of ledantevirus N protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | BARV | FUKV | NISHV | NKOV | LNcoLV-1 | YjTV-2 | KEUV | LDV | VAPV | KCV | MEBV | TYBV | KYAV | BUGV | WhFLV-5 | OITAV | KRV | WRpuLV-1 | FKRV | KOLEV |
| BARV |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FUKV | 96.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NISHV | 94.1 | 95.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NKOV | 76.2 | 77.2 | 75.5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LNcoLV-1 | 52.5 | 53.0 | 52.3 | 52.0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YjTV-2 | 52.7 | 53.0 | 52.7 | 53.4 | 47.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| KEUV | 39.6 | 39.6 | 39.1 | 40.1 | 40.4 | 38.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LDV | 39.4 | 39.4 | 38.7 | 39.6 | 41.1 | 39.3 | 80.7 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| VAPV | 42.2 | 41.5 | 41.3 | 43.9 | 40.1 | 38.3 | 62.5 | 63.9 |  |  |  |  |  |  |  |  |  |  |  |  |
| KCV | 39.9 | 40.4 | 39.0 | 40.4 | 37.8 | 38.9 | 46.9 | 47.9 | 47.9 |  |  |  |  |  |  |  |  |  |  |  |
| MEBV | 41.4 | 41.0 | 40.7 | 40.0 | 42.3 | 39.4 | 41.4 | 40.7 | 37.1 | 37.7 |  |  |  |  |  |  |  |  |  |  |
| TYBV | 41.4 | 41.0 | 40.5 | 39.8 | 41.6 | 38.2 | 41.9 | 40.0 | 37.9 | 37.4 | 87.1 |  |  |  |  |  |  |  |  |  |
| KYAV | 40.0 | 40.2 | 39.5 | 39.5 | 40.4 | 38.5 | 41.2 | 38.8 | 36.9 | 35.8 | 75.2 | 77.3 |  |  |  |  |  |  |  |  |
| BUGV | 39.5 | 40.0 | 39.0 | 40.7 | 39.0 | 37.5 | 38.6 | 37.1 | 37.6 | 37.0 | 74.7 | 75.9 | 85.2 |  |  |  |  |  |  |  |
| WhLFV-5 | 38.6 | 38.6 | 38.8 | 39.0 | 39.2 | 36.3 | 39.8 | 38.8 | 38.3 | 37.0 | 72.6 | 72.6 | 67.4 | 69.3 |  |  |  |  |  |  |
| OITAV | 37.6 | 37.4 | 36.9 | 37.9 | 42.0 | 36.8 | 39.2 | 39.4 | 35.6 | 38.8 | 55.0 | 55.0 | 51.5 | 53.2 | 52.2 |  |  |  |  |  |
| KRV | 41.7 | 41.0 | 40.0 | 42.4 | 39.9 | 39.0 | 39.5 | 39.3 | 35.5 | 38.2 | 60.2 | 61.4 | 57.6 | 56.0 | 58.3 | 58.3 |  |  |  |  |
| WRpuLV-1 | 37.9 | 37.4 | 36.4 | 37.9 | 41.8 | 38.2 | 39.2 | 39.4 | 37.1 | 37.4 | 54.1 | 54.8 | 52.2 | 50.6 | 51.1 | 71.2 | 53.6 |  |  |  |
| FKRV | 40.0 | 39.8 | 40.0 | 40.2 | 39.2 | 41.3 | 36.7 | 37.1 | 35.0 | 36.7 | 60.4 | 60.0 | 58.1 | 57.8 | 56.7 | 56.4 | 60.0 | 55.0 |  |  |
| KOLEV | 40.1 | 40.1 | 40.3 | 40.6 | 42.1 | 39.8 | 38.7 | 39.9 | 37.5 | 38.7 | 62.4 | 61.3 | 60.1 | 58.2 | 62.4 | 57.7 | 63.4 | 55.6 | 73.7 |  |

**Table 20.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of sigmavirus L protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | SyFV-2 | WhFV-2 | YsRV | HbDV-10 | HbDV-9 | WhHFV-1 | ApRV-3 | DAnaSV | DStuSV | DAffSV | DObsSV | DImmSV | CCapSV | DMelSV | HbDRV-1 | ADMSV | WhLFV-9 | WhLFV-10 |
| SyFV-2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhFV-2 | 82.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YsRV | 74.0 | 73.6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-10 | 53.3 | 52.6 | 52.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-9 | 50.3 | 50.1 | 50.9 | 52.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhHFV-1 | 50.8 | 51.3 | 50.0 | 52.1 | 51.0 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ApRV-3 | 48.7 | 49.1 | 47.7 | 48.8 | 45.8 | 46.5 |  |  |  |  |  |  |  |  |  |  |  |  |
| DAnaSV | 44.0 | 43.7 | 43.0 | 45.6 | 43.8 | 42.6 | 41.5 |  |  |  |  |  |  |  |  |  |  |  |
| DStuSV | 41.7 | 42.2 | 42.0 | 43.4 | 42.5 | 41.7 | 39.6 | 41.5 |  |  |  |  |  |  |  |  |  |  |
| DAffSV | 41.1 | 41.8 | 41.5 | 43.6 | 41.6 | 41.4 | 40.5 | 40.8 | 60.3 |  |  |  |  |  |  |  |  |  |
| DObsSV | 43.2 | 43.6 | 43.2 | 42.7 | 44.1 | 41.5 | 41.8 | 41.9 | 41.1 | 40.9 |  |  |  |  |  |  |  |  |
| DImmSV | 42.8 | 42.2 | 42.0 | 42.1 | 43.4 | 42.7 | 41.2 | 42.4 | 39.6 | 39.1 | 55.1 |  |  |  |  |  |  |  |
| CCapSV | 43.5 | 43.6 | 43.1 | 44.7 | 43.2 | 43.9 | 41.2 | 42.3 | 40.7 | 40.9 | 40.9 | 41.7 |  |  |  |  |  |  |
| DMelSV | 42.1 | 41.9 | 42.1 | 43.1 | 42.9 | 41.3 | 40.7 | 40.9 | 40.6 | 39.2 | 41.2 | 41.4 | 48.1 |  |  |  |  |  |
| HbDRV-1 | 43.0 | 42.8 | 42.4 | 43.1 | 42.3 | 42.2 | 41.3 | 41.8 | 40.8 | 39.5 | 41.0 | 41.6 | 47.5 | 48.9 |  |  |  |  |
| ADMSV | 43.2 | 42.6 | 43.7 | 42.9 | 42.6 | 43.0 | 40.6 | 39.7 | 40.1 | 38.6 | 41.9 | 41.7 | 50.5 | 50.1 | 50.0 |  |  |  |
| WhLFV-9 | 42.0 | 41.3 | 41.9 | 42.8 | 41.9 | 42.1 | 39.8 | 40.9 | 41.1 | 40.3 | 40.9 | 40.5 | 49.8 | 51.2 | 50.8 | 52.0 |  |  |
| WhLFV-10 | 42.6 | 42.5 | 42.1 | 42.7 | 42.4 | 42.4 | 39.7 | 41.7 | 40.9 | 39.8 | 41.3 | 40.9 | 49.5 | 50.9 | 50.4 | 51.7 | 67.6 |  |

**Table 21.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of sigmavirus G protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | SyFV-2 | WhFV-2 | YsRV | HbDV-10 | HbDV-9 | WhHFV-1 | ApRV-3 | DAnaSV | DStuSV | DAffSV | DObsSV | DImmSV | CCapSV | DMelSV | HbDRV-1 | ADMSV | WhLFV-9 | WhLFV-10 |
| SyFV-2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhFV-2 | 67.9 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YsRV | 58.4 | 59.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-10 | 36.1 | 35.8 | 37.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-9 | 27.8 | 26.7 | 28.3 | 29.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhHFV-1 | 27.1 | 24.5 | 27.5 | 26.9 | 23.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ApRV-3 | 30.9 | 29.7 | 31.4 | 30.5 | 27.9 | 26.2 |  |  |  |  |  |  |  |  |  |  |  |  |
| DAnaSV | 22.6 | 21.6 | 21.1 | 24.3 | 21.1 | 22.8 | 20.6 |  |  |  |  |  |  |  |  |  |  |  |
| DStuSV | 25.0 | 25.5 | 25.8 | 24.0 | 21.5 | 23.3 | 22.4 | 19.4 |  |  |  |  |  |  |  |  |  |  |
| DAffSV | 23.3 | 26.0 | 25.2 | 24.9 | 21.1 | 22.5 | 20.6 | 20.5 | 46.2 |  |  |  |  |  |  |  |  |  |
| DObsSV | 19.0 | 21.2 | 21.3 | 20.4 | 18.2 | 21.8 | 18.8 | 20.2 | 21.7 | 20.4 |  |  |  |  |  |  |  |  |
| DImmSV | 20.3 | 21.3 | 20.6 | 21.4 | 20.2 | 24.5 | 19.4 | 18.8 | 21.0 | 19.0 | 36.2 |  |  |  |  |  |  |  |
| CCapSV | 21.6 | 22.5 | 20.4 | 23.9 | 20.6 | 22.2 | 20.3 | 21.5 | 20.1 | 19.9 | 20.3 | 20.1 |  |  |  |  |  |  |
| DMelSV | 21.3 | 22.3 | 23.7 | 24.1 | 19.8 | 22.2 | 21.7 | 21.7 | 22.6 | 23.8 | 22.1 | 19.9 | 26.4 |  |  |  |  |  |
| HbDRV-1 | 20.0 | 21.1 | 21.8 | 23.8 | 22.6 | 23.9 | 19.9 | 24.3 | 22.0 | 20.7 | 22.2 | 20.3 | 25.6 | 26.3 |  |  |  |  |
| ADMSV | 22.1 | 21.3 | 20.0 | 22.3 | 18.7 | 19.0 | 21.5 | 20.0 | 18.5 | 18.8 | 18.1 | 18.6 | 20.4 | 19.9 | 22.2 |  |  |  |
| WhLFV-9 | 19.1 | 19.5 | 20.9 | 19.7 | 19.9 | 24.0 | 21.6 | 21.0 | 21.1 | 19.4 | 19.4 | 20.5 | 22.5 | 22.7 | 24.1 | 19.8 |  |  |
| WhLFV-10 | 19.9 | 19.7 | 21.9 | 21.2 | 21.1 | 21.3 | 21.2 | 20.2 | 21.6 | 20.8 | 19.1 | 19.2 | 21.1 | 23.7 | 22.5 | 22.5 | 35.1 |  |

**Table 22.** Percentage amino acid identities (p-distance) of a CLUSTAL W alignment of sigmavirus N protein sequences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Virus | SyFV-2 | WhFV-2 | YsRV | HbDV-10 | HbDV-9 | WhHFV-1 | ApRV-3 | DAnaSV | DStuSV | DAffSV | DObsSV | DImmSV | CCapSV | DMelSV | HbDRV-1 | ADMSV | WhLFV-9 | WhLFV-10 |
| SyFV-2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhFV-2 | 83.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| YsRV | 65.5 | 65.1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-10 | 34.6 | 34.4 | 33.8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| HbDV-9 | 31.3 | 30.9 | 31.6 | 28.3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WhHFV-1 | 27.3 | 27.1 | 29.2 | 28.2 | 29.4 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ApRV-3 | 24.4 | 24.7 | 25.1 | 23.6 | 26.7 | 22.9 |  |  |  |  |  |  |  |  |  |  |  |  |
| DAnaSV | 27.3 | 27.1 | 26.6 | 25.9 | 23.2 | 22.6 | 18.8 |  |  |  |  |  |  |  |  |  |  |  |
| DStuSV | 26.7 | 27.8 | 27.2 | 25.5 | 23.6 | 22.0 | 22.5 | 24.2 |  |  |  |  |  |  |  |  |  |  |
| DAffSV | 27.5 | 27.1 | 28.3 | 24.9 | 23.6 | 23.7 | 25.7 | 23.7 | 43.3 |  |  |  |  |  |  |  |  |  |
| DObsSV | 20.1 | 20.8 | 21.7 | 19.8 | 22.1 | 19.7 | 18.1 | 22.0 | 20.9 | 22.3 |  |  |  |  |  |  |  |  |
| DImmSV | 20.8 | 20.6 | 22.0 | 20.5 | 22.8 | 20.5 | 20.5 | 21.3 | 18.8 | 18.3 | 43.5 |  |  |  |  |  |  |  |
| CCapSV | 22.8 | 22.5 | 21.6 | 21.9 | 23.5 | 19.3 | 19.6 | 21.1 | 21.3 | 20.0 | 22.4 | 20.7 |  |  |  |  |  |  |
| DMelSV | 23.0 | 22.5 | 24.6 | 22.8 | 23.7 | 23.5 | 23.8 | 20.6 | 20.6 | 20.4 | 22.2 | 20.3 | 30.5 |  |  |  |  |  |
| HbDRV-1 | 19.2 | 18.3 | 19.7 | 19.2 | 17.7 | 20.6 | 20.9 | 19.6 | 19.0 | 19.7 | 18.1 | 17.3 | 25.2 | 26.0 |  |  |  |  |
| ADMSV | 24.1 | 24.5 | 24.3 | 23.5 | 20.2 | 24.9 | 23.4 | 19.1 | 22.6 | 20.8 | 18.9 | 20.0 | 29.6 | 29.1 | 25.6 |  |  |  |
| WhLFV-9 | 23.7 | 25.1 | 24.0 | 22.6 | 21.4 | 20.8 | 19.9 | 20.9 | 22.8 | 22.0 | 20.8 | 19.2 | 27.6 | 33.9 | 21.5 | 27.1 |  |  |
| WhLFV-10 | 22.6 | 23.0 | 22.6 | 20.9 | 24.2 | 21.5 | 22.4 | 18.4 | 21.6 | 20.6 | 19.9 | 20.1 | 27.2 | 34.5 | 24.8 | 29.4 | 50.0 |  |

**References**

1. Chiapello M, Bosco L, Ciuffo M, Ottati S, Salem N, Rosa C, Tavella L, Turina M (2021) Complexity and local specificity of the virome associated with tospovirus-transmitting thrips species. J Virol 95:e0059721. doi: 10.1128/JVI.00597-21

2. Edridge AWD, Abd-Elfarag G, Deijs M, Jebbink MF, Boele van Hensbroek M, van der Hoek L (2022) Divergent rhabdovirus discovered in a patient with new-onset nodding syndrome. Viruses 14:e210. doi: 10.3390/v14020210

3. Hause BM, Nelson E, Christopher-Hennings J (2020) Novel and diverse non-rabies rhabdoviruses Identified in bats with human exposure, South Dakota, USA. Viruses 12:e1408. doi: 10.3390/v12121408

4. Litov AG, Belova OA, Kholodilov IS, Gadzhikurbanov MN, Gmyl LV, Oorzhak ND, Saryglar AA, Ishmukhametov AA, Karganova GG (2021) Possible arbovirus found in virome of *Melophagus ovinus*. Viruses 13:e2375. doi: 10.3390/v13122375

5. 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. doi: 10.1038/nature20167

6. Wu Z, Han Y, Liu B, Li H, Zhu G, Latinne A, Dong J, Sun L, Su H, Liu L, Du J, Zhou S, Chen M, Kritiyakan A, Jittapalapong S, Chaisiri K, Buchy P, Duong V, Yang J, Jiang J, Xu X, Zhou H, Yang F, Irwin DM, Morand S, Daszak P, Wang J, Jin Q (2021) Decoding the RNA viromes in rodent lungs provides new insight into the origin and evolutionary patterns of rodent-borne pathogens in Mainland Southeast Asia. Microbiome 9:e18. doi: 10.1186/s40168-020-00965-z

7. Xu L, Guo M, Hu B, Zhou H, Yang W, Hui L, Huang R, Zhan J, Shi W, Wu Y (2021) Tick virome diversity in Hubei Province, China, and the influence of host ecology. Virus Evolution 7:1-12. doi: 10.1093/ve/veab089