This form should be used for all taxonomic proposals. Please complete all those modules that are applicable.



For guidance, see the notes written in blue and the separate document “Help with completing a taxonomic proposal”

Please try to keep related proposals within a single document.

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Code assigned:** | ***2019.013M*** | | | | (to be completed by ICTV officers) |
| **Short title:** Create three new species in the genus *Sripuvirus,* family *Rhabdoviridae* | | | | | |
| **Modules attached** | | **1**  **2  3  4** | | | |
| **Author(s):** | | | | | |
| Members of the ICTV *Rhabdoviridae* Study Group:  Peter J. Walker  Kim R. Blasdell  Ralf G. Dietzgen  Anthony R. Fooks  Juliana Freitas-Astúa  Hideki Kondo  Gael Kurath  Ivan Kuzmin  Robert B. Tesh  Nikos Vasilakis  Anna E. Whitfield | | | | | |
| **Corresponding author with e-mail address:** | | | | | |
| peter.walker@uq.edu.au | | | | | |
| **List the ICTV study group(s) that have seen this proposal:** | | | | | |
| A list of study groups and contacts is provided at <http://www.ictvonline.org/subcommittees.asp> . If in doubt, contact the appropriate subcommittee chair (there are six virus subcommittees: animal DNA and retroviruses, animal ssRNA-, animal ssRNA+, fungal and protist, plant, bacterial and archaeal) | | | *Rhabdoviridae* Study Group | | |
| **ICTV Study Group comments (if any) and response of the proposer:** | | | | | |
| Proposal supported by a majority of Study Group members (11 supporters and 2 non-responders). | | | | | |
|  | | | | | |
| Date first submitted to ICTV: | | | | 25 January 2019 | |
| Date of this revision (if different to above): | | | |  | |

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

**Part 2**: **PROPOSED TAXONOMY**

|  |
| --- |
| Present the proposed new taxonomy on accompanying spreadsheet |
| **Name of accompanying spreadsheet:**  2018.013M.A.v1.Sripuvirus\_3newsp.xlsx |

**Part 4:** **APPENDIX**: supporting material

| **References:** |
| --- |
| 1. **Karabatsos N (ed).** 1985. International Catalogue of Arboviruses Including Certain other Viruses of Vertebrates. American Society for Tropical Medicine and Hygiene, San Antonio.  2. **Tesh RB, Travassos da Rosa APA, Travassos da Rosa JS.** 1983. Antigenic relationship among rhabdoviruses infecting terrestrial vertebrates. Journal of General Virology **64:**169-176.  3. **Vasilakis N, Tesh RB, Widen SG, Mirchandani D, Walker PJ.** 2018. Genomic characterisation of Cuiaba and Charleville viruses: arboviruses (family *Rhabdoviridae*, genus *Sripuvirus*) infecting reptiles and amphibians. Virus Genes doi:10.1007/s11262-018-1620-x.  4. **Doherty RL, Carley JG, Standfast HA, Dyce AL, Kay BH, Snowdon WA.** 1973. Isolation of arboviruses from mosquitoes, biting midges, sandflies and vertebrates collected in Queensland, 1969 and 1970. Transactions of the Royal Society for Tropical Medicine Hygiene **67:**536-543.  5. **Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, Wang W, Eden JS, Shen JJ, Liu L, Holmes EC, Zhang YZ.** 2018. The evolutionary history of vertebrate RNA viruses. Nature **556:**197-202.  6. **Wellehan JF, Jr., Pessier AP, Archer LL, Childress AL, Jacobson ER, Tesh RB.** 2012. Initial sequence characterization of the rhabdoviruses of squamate reptiles, including a novel rhabdovirus from a caiman lizard (*Dracaena guianensis*). Veterinary Microbiology **158:**274-279.  7. **Walker PJ, Kongsuwan K.** 1999. Deduced structural model for animal rhabdovirus glycoproteins. Journal of General Virology **80:**1211-1220.  8. **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. |

|  |
| --- |
| **Annex:**  **Proposed new members of the genus *Sripuvirus***  Cuiaba virus (CUIV; strain BeAn 227841) was isolated from a toad (*Bufo marinus*) captured in Para State, Brazil, in 1972 (1). In complement-fixation tests, CUIV was shown to cross-react antigenically with Charleville virus (2). The complete CUIV genome sequence (11,700 nt) has been determined (3). We propose CUIV be assigned to the new species *Cuiaba sripuvirus*.  Charleville virus (CHVV; strain Ch9824) was first isolated from sandflies (*Phlebotomus* sp.) collected in 1969 at Charleville in south-western Queensland (4). It was subsequently isolated on three separate occasions from geckos (*Gehyra australis*) collected at Kowanyama (formerly Mitchell River Aboriginal community) on the western shore of the Gulf of Carpentaria in far northern Queensland (4). CHVV was also isolated from a pool of midges (*Lasiohelea* sp.) collected at the same location but re-isolation from the same insect pool was unsuccessful (4). The near-complete CHVV genome sequence (11,434 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini (3). We propose CHVV be assigned to the new species *Charleville sripuvirus.*  Hainan black-spectacled toad rhabdovirus (HnBSTRV) was detected in a toad (*Duttaphrynus melanostictus*) collected in China (5). The near-complete HnBSTRV genome sequence (11,100 nt) has been determined, including complete coding sequences but incomplete 3' and 5' termini (5). We propose HnBSTRV be assigned to the new species *Hainan sripuvirus.*  Other related viruses  Caiman lizard virus (CLV), isolated in 1999 from a lizard (*Dracaena guianensis*) imported to USA from Peru (6), are also related phylogenetically to the sripuviruses. Bullet-shaped virions were observed in intracytoplasmic inclusions of erythrocytes from an infected lizard. Only partial sequences of the L gene have been reported (6). Phylogenetic analysis using the translated partial L protein sequence indicated that the virus clusters with members of the genus *Sripuvirus* (6). CLV is considered to be a likely candidate for classification in the genus but more complete genome sequence data is required.  **Genome organisations**  The genomes CUIV, CHVV and HnBSTRV range in size from approximately 11.1 kb to 11.7 kb in length. Like other sripuviruses, each contains the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G* and *L*) and additional long open reading frame (249 nt) within the *M* gene (Mx), just overlapping the end of the M ORF (**Figure 1**) (3). As observed for other sripuviruses, expression of this additional ORF from a bicistronic mRNA appears to occur by a stop-start mechanism facilitated by ‘termination upstream ribosome binding site’ (TURBS) which contains a short sequence motif (AGGGA in CHVV; UGGGA in CUIV; UGGGG in HnBSTRV) that is complementary to the loop region of helix 26 of 18S ribosomal RNA (**Figure 2A**) (3). Like other sripuviruses, there is also an additional long ORF (Gx) in the *G* genes of CHVV (351 nt), CUIV (321 nt) and HnBSTRV (351 nt) in alternative reading frames commencing just downstream of the G ORF initiation codon (**Figure 1**) (3). As in other sripuviruses, the Gx proteins display high levels of pairwise homology and are predicted to be transmembrane proteins with a single transmembrane helix with an N-terminal exodomain (**Figure 2B**) (3). In CUIV, there is also an additional transcriptional unit (*U1* gene) between the *M* and *G* genes containing a long ORF (303 nt) (**Figure 1**) (3). There is no corresponding gene in CHVV, HnBSTRV or any of the classified sripuviruses.  **Glycoprotein structures**  A Clustal X alignment of the G proteins of CHVV, CUIV, HnBSTRV and the classified sripuviruses with the G protein of vesicular stomatitis Indiana virus (VSIV) indicated that 10 of the 12 cysteine residues in the ectodomain are fully conserved, and these correspond to five established disulphide bridges (CI-CXII; CIII-CV; CVI-CVII; and CIX-CXI) (**Figure 3**) (3, 7). Like the other sripuviruses, CHVV, CUIV and HnBSTRV lack two cysteine residues that form another known disulphide bridge in VSIV (CVIII-CX) which links α-helix C to the *lm* loop in the pleckstrin homology domain (DIII) (8). However, they all have two additional cysteine residues in the ectodomain which are likely to form an additional disulphide bridge between β strands s and t in the lateral domain (D1) (3, 8).  **Phylogenetic analysis**  Based on ML trees generated from complete L protein sequences, CHVV, CUIV and HnBSTRV cluster with the sripuviruses in a distinct and well-supported monophyletic clade (**Figure 4**) (3). By this analysis, CHVV is most closely related to HnBSTRV which also clusters with ALMV; CUIV clusters with CHOV and SMV.  **Amino acid sequence identities**  Pairwise sequence identities (p-distances) calculated in MEGA7 from Clustal X alignments of the N, G and L proteins indicated that CHVV was also most closely related to HnBSTRV with sequence divergence of 33.0% in the N protein, 39.2% in the G protein and 26.7% in the L protein (**Tables 1, 2 and 3**) (3). CUIV was most closely related to CHOV and SMV with sequence divergence of 55.3% and 57.3% in the N protein, 56.5% and 55.1% in the G protein, and 43.8% and 43.4% in the L protein, respectively.  **Species demarcation criteria**  Viruses assigned to different species within the genus *Sripuvirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 5% in N; B) minimum sequence divergence of 10% in L; C) minimum amino acid sequence divergence of 15% in G; 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 hosts and/or arthropod vectors.  CHVV, CUIV and HnBSTRV meet all sequence divergence criteria (A-C). They also appear to occupy different ecological niches (criterion F) with respect to geographic location and hosts in which they were detected. Although CHVV has a similar genome organisation to HnBSTRV, the genome organisation of CUIV is unique, containing an additional transcriptional unit between the *M* and *G* genes (criterion D). Neutralisation test results are not currently available (criterion E). Overall, the data support the classification of CUIV, CHVV and HnBSTRV as separate new species in the genus *Sripuvirus*.  . |

**Figure 1.** Sripuvirus genome organisations. Open arrows indicate the locations of long open reading frames (ORFs) N, P, M, G and L, each of which is located within a transcriptional unit bounded by conserved transcription initiation and transcription termination/polyadenylation sequences. The Mx ORFs (shaded blue) overlap the end of the M ORFs within the same transcriptional units and encode small proteins of unknown function. The Gx ORFs (shaded orange) occur in alternative reading frames within the *G* genes and encode predicted membrane spanning proteins. Additional ORFs occur in independent transcriptional units between the *N* and *P* genes (shaded red and green), and between the *M* and *G* genes (shaded purple). Alternative ORFs (shaded grey) of significant length (>180 nucleotides) also occur in the *N, P* and *L* genes but the significance of these is unknown.

[Reproduced under license from Springer, *Virus Genes,* Vasilakis *et al.,* doi:10.1007/s11262-018-1620-x, copyright 2018.]

**A**

CUIV AAAAGG**AGGGA**AG**AUG**AUUACGAUUCCGAU**UGA**GUUUUUUGUCCCUUUUCUUGUUGUG

CHVV CAACUG**UGGGA**ACAAGCGAGUUGCUCUGCUGAAGUCA**UAAUG**UUUAAGAUUGCUUUUG

SRIV GAGGCG**UGGGA**AAAAUUCUACUCAGAUGUGCAGAAGACC**AUGA**ACUUGAUUUCAAUGG

NIAV GAGGAG**UGGGA**AAUCACGUACUCAGAUCUGUGCAAGACC**AUGA**UUGUGUUGUCAGUGG

SMV UAAAUU**UGAGA**UUAAUAAACGUCAUAUUUUGAUUUCACAGA**AUGA**UUGUCUUUCCUUU

CHOV CCAAUU**UGAGA**UUCAAAAGAAUAAGAUUUUAAUAUCUCCUAA**AUGA**UUAUAUUACCCA

ALMV CUCUCA**UGGGA**ACAAAAGAGUUCAAAUCAAGAAGAUACA**AUGA**UUACCUUUAGUUUUG

HnBSTRV GUAGAG**UGGGG**ACAAAGAGGCUAGUUCUGUACAAGGCAAA**AUGA**UCAAAAUUGGAGUG

**B** **Transmembrane domain**

HnBSTRV\_Gx MEDLPSFWALCFQLLLITHCSSLVRRTFIGLLLTILVCAVLSGVLSILYQEGSETSPSYSLYIQTVMTSLDTVAIKLNGFLNV

CHVV\_Gx MESLPSFWHLLFQYIVITHCMSPVRKVFIGLLLIILVCVALSGVLSILCPEESELCHSSSHYIQTTTMLLDTVAIRLNGLLSV

ALMV\_Gx ME-FSSYCACLGQLIVIMFFTGPLRKLMVGTLLVILSCGVLSELLRILSLNAEDQSLFWSQLLHTLLTLTDTAATKLNGFLNV

SRIV\_Gx ME-SFSLFVELFRLLLIMYFTTPLKRIFIGIQLIIVVCGVLYGVLGYLT-HLQVESLYNFPAIQATMIFLDTVVTKLSGFQSV

NIAV\_Gx MD-YFFLLVEFSRLLLVMFFTTPLKKIFIGIQLIIVVFGALYAVLASLTP-LQVGSLFQYPQIQATMIFLDSVVTKLNGFQSV

SMV\_Gx MD-LSCWLLELFKLLMAVFFISPVKRIFIGTQLTILVLGALLGLATSIL-GTVQMSSSLSHIVTLSISLMDTAVTKLSGFRSV

CHOV\_Gx MD-LSCWLSELFKLLMAIFFISPLKRIFIGTQLIILVFGALLGLAINIP-DTVQMSSSVYHIVTQAISLMDTAVTRLSGFQSV

CUIV\_Gx MD-LSCLLSELFKLIAVMFFIGPARKIFIGIQLTIVLFSALYGVASYILP-LMPASEPLYHTLLEVLKLTDIAVTRLSGFRSV

\*: : : . :: ::\* \* \*: .\* : : . : \* .. :\*.\*: .\*

HnBSTRV\_Gx LKDFLGALMSSNTSDIWNLMLRSAKMLGRQREMG

CHVV\_Gx LKHGIGLLMSNSTSDILSQLLKSARLPGLRREMA

ALMV\_Gx LRHGIGQLIFPK----------------------

SRIV\_Gx LKHGIGALMLRNTSGQSQ----------------

NIAV\_Gx LKRGIGALMSSSTSDQ------------------

SMV\_Gx LNRGIGQLILNSISESCQ----------------

CHOV\_Gx LRRGIGQLMLNNISGSCQ----------------

CUIV\_Gx LSRGIGHLILNNMSEPCQSLTENVMT--------

\* :\* \*: .

**Figure 2.** **A.** TURBS-like sequence motifs in the *M* genes of sripuviruses at the junction of the M and Mx ORFs. The motif features the consensus sequence UGGGA (highlighted grey) upstream of overlapping or consecutive initiation codons (Mx ORF) and termination codons (M ORF) (double underlined). Variations in the TURBS sequence (AGGGA, UGAGA, UGGGG) occur in CUIV, CHOV, SMV and HnBSTRV. **B.** Clustal X alignment of the sripuvirus Gx proteins. Predicted transmembrane domains are shaded (grey). [Reproduced under license from Springer, *Virus Genes,* Vasilakis *et al.,* doi:10.1007/s11262-018-1620-x, copyright 2018.]

**CI CII CIII**

VSIV MKC----LLYLAFLFIGVNCKFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPKSHKAIQADGWMCHASKWVTTCDF

CHOV MQK-YG-FILLALGAIQVTHGNLFYFPVEKDIHWHPINYTSIRCPIRAGD-QYPG-YSANVQFRVPYSDSSNIINGHSCHKTQWVSECTE

SMV MKE-YG-LILLAIGAIQIANGSLFYFPSEKDLHWHPINHTSIRCPIRSGD-KYPG-DSANVQFSVPYRDPKHLLNGYSCHKTKWVSECTE

CUIV MSF-YG-FILLTLGALQAHSRYVLYRPSEENLHWHPVNHSSLQCPLRGSE-LYPS-PNAGFRASLPHSIGGIKTDGYSCHKTEWVSECTE

NIAV MSQGYG-LLFLVSGVLSLTAGYVFHYPIEKDIHWYPANHSSLRCPIRSAS-ITDT-PTGGVTISIPSNPSNNDLPGFSCHKTEWISECTE

SRIV MKTSYG-IILLICGVISPAFNYVFHYPLEKNIHWYPANHSSLRCPVRSAR-ILDT-PTGGITLQLPGNPSNHDIPGYSCHKTEWVSECVE

CHVV MYPNYGIFAILLALAVSVYCDYTLYVPSEKSIHWTPANHTSLRCPIRSAV-HPVSGRVGTLPFLIPLHPNDNNVAGYSCHKAEWTSECTE

HnBSTRV MILDNGRLALLLGVVFSVAADYTLFVPGEKNIHWTPANHSSLRCPVRSAV-HPVSGRIGNLPLLLPIHPNSNDEPGYSCHKVEWISECSE

ALMV MDFKYG-IFILLCLLGSTNCNYVLHWPTEKVDGWHPASHTQLRCPIRTAS-HPITERRGSVSFLVPTFAYSTHVDGYSCHKTEWVSKCTE

\* : \* : \* :: \* . \*\* . :\* \* \*\* :\* : \*

**CIV CV CVI**

VSIV RWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPPQSCGYATVTDAEAVIVQVTPHHVLVDEYTGEWVDSQFIDGKCSNDI---

CHOV TWYWTTDVKQYIRVLPVSLNECLEEKEKRGQGKSIAPFFEAPVCQWANTVRKENSFVILNEKRVQLDPYNGNVVDPLFPNGRCSTPVK--

SMV SWYWTTDIKQYIRVLPVSKEECVNELRNREEGQSLSPFFEAPVCQWANTVRKEKDFVILNKKRMQLDPYSLEVVDPLLLTGRCKTSLD--

CUIV SWYWTSDIKQYVRALPVSHRECNDIIAKKRSGDDETPFFPAPHCQWANTIRQSKTFLKANTKKVTIDPYTSEYVDPIFVGGRCKDPP---

NIAV TWYWSTDVKQYIRPVSVTADECKKAQRDKEVGTEITPFFTAPVCQWSNTVRKVNSFVITNKKNVKFDPYNLDFIDPILVGGRCKGNQE--

SRIV TWYWSTDVKEYIRSVPVTFEECQRALQDKSIGNEVSPFFTAPTCQWSNTVRKVNTFVIVHSKDAKLDPYNLDRVDPIFPGGRCKSKEK--

CHVV TWYWTTDVKQYIRYIEPTFEECKAAWVKKRDGLNPNPYFEAPTCQWANTVRKSHQFTMINLKQVKIDPYNLDFVDPIFPGGRCS-NQD--

HnBSTRV GFFGGTDVQQYIRHLEPNAQECKDAWKAKRDGLEVTPYFEAPKCQWMNTVRKAHTFVVVNTKATKTDPYTLEYVDPLFPGGRCT-TKE--

ALMV TWYWTTDISQIIRTAPVSKEECLEAITKKESGISQTPFFENPVCQWVNTVEKSHFFVIVNKKKVKIDPYNLDAIDSLFPGGRCSQDEDGA

:: . : . :\* . :\* \* \* \* \* : .. : : \* \*. : :\*. : \*:\*.

**CVII CVIII CIX CX**

VSIV --CPTVHNSTTWHSDYKVKGLCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGDKACKMQYCKHWGVRLPSGVWFEMADKDLF

CHOV -GCPTIQDNVLWFT--TEDPIKGVHMRTINLKYAPPRGDMKELSIWGEGVPVTKMS-----KSCKMSYANEEGVRFESGLWGIPIPGS--

SMV -GCKTIQDGVLWFS--EENPKKAMEMRVLGLKYSPKGGDMKDLVIWGEGIPLTKMS-----KACKMQYANTEGVRFESGFWGVPLAGT--

CUIV --CQTIQSGVIWLP--GLRESTSHMWQEVYLRYAPPNGNMSQLKIWGPGFPITKME-----SACKMTYITRKVIRFPSG-MGVSVDESAF

NIAV -SCPTIQAGVIWLP--RLQPTKATSWTNIYAKYKRVGPHMGDWKFWGGGMPTSTFK-----DACKMEFRGKEGIRVSSG-FWFHIP--QM

SRIV -YCPTIQAGVIWLQ--KVSQRITIEWQTIKAKYKRIGNNMEDWKLWGGGMPTSTFR-----GACKMDFAAKTGIRASNG-FWFHVP--NI

CHVV -SCKTVQSGVIWHPQRKNPQQTEYAWKNVQASYGKSRKESDVYLIWGGGIPTTKFN-----NSCKMNFKGKTGIRFSSG-FWAGVEPSTW

HnBSTRV -YCPTIQGGVIWLPLKPVDPKVKYEWRTVNADYGTSAKDSKTYLIWGGGMPTTKLK-----GACRMSFQGQEGIRFSSG-FWAGTKTETW

ALMV YNCPTIQAGVKWMT--KDPDDYKIRWSMIRADYS-YEESTKKFVLWGGGLPTTYFN-----GSCQMTFQKQTGVRFPNG-FWAMPPERKW

\* \*:: .. \* : : \* \*. . : :\*:\* : :\* .\*

**CXI CXII**

VSIV AAARFP-------ECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNPGTGPAFTIINGTLKYFET

CHOV -DGKFTQWKDNLENCPEGTLLKLPNAHEEISEHQIEIEDLVLSLTCFEMINNFRDTGKISFTDLALLTPDHPGLGNVYRINKGVLEASSA

SMV -DQKFTNWKENLQDCPEGSIVKLPSAHEEIAEHQIEIDDLVLSLQCIDVINNFKNTNQISFTDLALLTPDHAGLANAYRINRGILEATTA

CUIV KDQKFKTWLSEMKECEPGTTLKVPNAHESVAEHQVEIDDIVYTLRCMDIVSRFRDTNTISIMDLSFFGPDHEGMGNVYRLRNGFLEATTA

NIAV DDVEFKTEYGKLAHCVSSKEIKFPSAHEEVAEHEMEIQDLILTLRCRDIIDKYEETGSISFMDLALFDPDNEGPAHIYRINKGKLEAGLV

SRIV SDSAFKDAFNAMTQCPQNTEVKFPSAHEEVAEHEMEIQDLILTLRCRDIIDKFAETGQISFMDLSLFDPDNEGPAHVYRIRNGKLETGIV

CHVV HDSAFKKVYEDLPNCSEGTTIKVPSAHEEIAEHEMEIDDLILSLKCYDIIRAFEETGKISFLDLGFFNPDTEGPAHIYRIKEGKLEAASV

HnBSTRV SNKNFTDAIGELSACSSNLIARFPSAHEEIAEHEMEINDLILSLKCYDIVREFEETGKISFLDLGFFNPETSGPGHIYRIKDGILEAASV

ALMV TDTDFVTMWNNLPACKEGTEVYYPNAHEEFSEHGMEIDDLILSLKCLDIIREFDETGKISFLDLSFFSPDKPSFGHVYRLNEGKLEVATT

\* \* . \*. . . ::: :: \* : \*\* \*\*. : \*. . . : : \* \*: .

**a b**

VSIV RYIRVEIAAPILS---RMVGM--ISGTTTERELWDDWAP--YEDVEIGPNGVLRTSSGYKFPLYMIGHGMLDSDLRLSSKAQVFEHPHIQ

CHOV FYQQCKIKPNKND----VICY--GGNDFNTENKWKRWVDSGTPGTLSGYNGVYKMGGKIINAHENLMTNRIIDEDILKRDLKTVRHPIEV

SMV YYYQCKLKSNVDN----VICY--GGQDFNIPVKWNGWVDSGTPGTASAFNGVYRKEGKIINANENLMMNRIMDEDIMMRDIQPVRHPIEI

CUIV HYVKCIRSARDNS----ALCM--DDDDKT--IFSPQWVSSGVKGVYSGFNGVYKKDGKIFDAGGHLSENMLQDMDTMRLDIQVIKHPIQL

NIAV NYGECKVSKKGD--PAESACVKVMDNGQRSPIFFQDWVPTGIKGIQSGFNGLYRENGEIKHAGYNLFQNKLTESDIQRMELTPIHHPVLL

SRIV NYGECRVGNSGDKDPANSICIKIADDGIRAAIFYDDWVSTGEKGIQSAFNGLYRENGVIRHAGYNLFANKLTEKDIEKQDLQQVHHPIDL

CHVV NYQACQIHPKP-HDPTRVIC---KDS--KIKVLYDEWVPTGVSGILSGFNGVYKDRGVVKYAGYNLWNNKLTEADIQQMELEAVPHPSVL

HnBSTRV HYRECPILPG---KFSKIHC---KPS--TDWTIYDTWVPTGIPGILSAFNGVYKENGIVNYAGYNLWNNKLTEADIQRMELQPVPHPVTL

ALMV NYGSCVMKSGHEY-SGNELC---HDT-NDRVISYRDWVPSGIQGRMSAFNGVYKDNGKIRHAGYQLWANRLTESDIQRYQMARVSHPIEN

\* \*. . . \*\*: : . . : . : . . . \*\*

VSIV DAASQLPDDETLFFGDTGLSKNPIELVEGWFSGWKSSIASFFFIIGLIIG-LFLVLRVSIYLCIKLKHTKKRQIYTDIEMNRLGK

CHOV SVRKNSSHPIYFFDWTGERGNPVEDVFEGLNNFWRKAIEWIAIVSFTFLLCIAIFILFKLLGMFRSNKKDRLAGSGEHEMSLFR-

SMV VANKNQSLPQYFFDWTGERGNPVEDVFEGVSKFWRKVIEWIAIGFFSIIGVLFFTIFVKIIWSFK--RGNRKSKNSDQELNLFR-

CUIV VLKDKYNDTTLFADETGERGNLDIPLFHDLKEIWNKTIHWVAGAGVLVISLIIGTALCRCF---KPGKRQKKRGN-EHEMREF--

NIAV SLSDVAPGLNVTFDQTGERGELDLDLLPGITGIWRKFVEYLSMAALILTLIVSIFVVWKCCISNHLGPSKKTSEMEYFE------

SRIV VIEKFMPGLNLTFDTTGQRGEV-YDLFPDIKGFWRSVLEYSVIGFFVLITIIFIWIVVKCCGCFRMK--RHNQIDDYYD------

CHVV VIEKFAPGMNITLETTGERGEMDLDLLGGISSGWRKFWKYTTAFIISILLLYLIYFALKFFKILKTPRTERPETMQMQGFY----

HnBSTRV SIQKYAPGLNITMDETGERGEVDLDILGGLSHLGRSILKYISLTGILIFSLVFSYYMIK-LACNRQVRNSRRERVEMEGFY----

ALMV VISKLVPGLNVTFDATGQRGELVDDFLSGVSEGWRKFLKYGAMTLILLAVLLLFLFMCNLLPKFKPKR-PALERYAMTEF-----

. .: .. . .. :

**Figure 3.** ClustalX multiple sequence alignment of the G proteins of sripuviruses and vesicular stomatitis Indiana virus (VSIV). Twelve cysteine residues in VSIV form six disulphide bridges (CI–CXII; CII–CIV; CIII–CV; CVI–CVII; CVIII–CX and CIX–CXI) and these bridges are conserved to various extents amongst rhabdoviruses in patterns that are somewhat genus-specific. Residues forming four of these bridges (CI–CXII; CIII–CV; CVI–CVII; and CIX–CXI) are conserved in sripuviruses. Cysteines forming the CVIII–CX bridge are absent in all sripuviruses which have two additional conserved cysteines (*a, b*) downstream of CXII that likely form an additional disulphide bridge. Predicted N-terminal signal domains (dark grey) and C-terminal transmembrane domains (light grey) are also shown. [Reproduced under license from Springer, *Virus Genes,* Vasilakis *et al.,* doi:10.1007/s11262-018-1620-x, copyright 2018.]



**Figure 4.** The evolutionary history was inferred from a Clustal W alignment of complete L protein sequences of 3 proposed sripuviruses and 110 other animal rhabdoviruses currently assigned or recently proposed for assignment to species. Phylogenetically informative sites were selected from the alignment using GBLOCKS resulting in 1069 positions in the final dataset. The tree was inferred in MEGA7 by using the maximum likelihood method based on the Whelan and Goldman (WAG) + Freq. model. The tree with the highest log likelihood (−105224.3294) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying neighbour-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. [Reproduced under license from Springer, *Virus Genes,* Vasilakis *et al.,* doi:10.1007/s11262-018-1620-x, copyright 2018.]

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

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NIAV | SRIV | ALMV | CHOV | SMV | HnBSTRV | CHVV | CUIV |
| NIAV |  |  |  |  |  |  |  |  |
| SRIV | 65.0 |  |  |  |  |  |  |  |
| ALMV | 35.7 | 36.0 |  |  |  |  |  |  |
| CHOV | 30.3 | 30.5 | 30.5 |  |  |  |  |  |
| SMV | 31.3 | 31.8 | 30.3 | 74.4 |  |  |  |  |
| HnBSTRV | 34.7 | 34.7 | 39.2 | 29.8 | 30.0 |  |  |  |
| CHVV | 34.0 | 34.5 | 40.2 | 29.3 | 29.8 | 67.0 |  |  |
| CUIV | 32.8 | 32.3 | 33.0 | 44.7 | 42.7 | 34.5 | 32.3 |  |

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

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NIAV | SRIV | ALMV | CHOV | SMV | HnBSTRV | CHVV | CUIV |
| NIAV |  |  |  |  |  |  |  |  |
| SRIV | 63.5 |  |  |  |  |  |  |  |
| ALMV | 45.8 | 47.0 |  |  |  |  |  |  |
| CHOV | 40.8 | 40.0 | 41.2 |  |  |  |  |  |
| SMV | 41.0 | 41.6 | 40.6 | 67.2 |  |  |  |  |
| HnBSTRV | 49.5 | 48.7 | 46.0 | 38.4 | 37.9 |  |  |  |
| CHVV | 51.1 | 48.2 | 50.1 | 41.0 | 39.8 | 60.8 |  |  |
| CUIV | 39.6 | 40.8 | 37.3 | 43.5 | 44.9 | 35.9 | 39.0 |  |

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

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | NIAV | SRIV | ALMV | CHOV | SMV | HnBSTRV | CHVV | CUIV |
| NIAV |  |  |  |  |  |  |  |  |
| SRIV | 67.3 |  |  |  |  |  |  |  |
| ALMV | 48.6 | 49.0 |  |  |  |  |  |  |
| CHOV | 47.9 | 48.1 | 48.1 |  |  |  |  |  |
| SMV | 47.3 | 47.6 | 48.1 | 70.8 |  |  |  |  |
| HnBSTRV | 48.0 | 48.5 | 56.0 | 47.2 | 47.3 |  |  |  |
| CHVV | 49.4 | 50.0 | 55.5 | 48.3 | 48.4 | 73.3 |  |  |
| CUIV | 47.1 | 46.5 | 47.9 | 56.2 | 56.6 | 45.3 | 46.2 |  |

[Tables 1-3 reproduced under license from Springer, *Virus Genes,* Vasilakis *et al.,* doi:10.1007/s11262-018-1620-x, copyright 2018.]