Virus Genes (2019) 55:87–94
https://doi.org/10.1007/s11262-018-1620-x

Genomic characterisation of Cuiaba and Charleville viruses: arboviruses (family *Rhabdoviridae*, genus *Sripuvirus*) infecting reptiles and amphibians

Nikos Vasilakis[1,2,3] · Robert B. Tesh[1,2,3] · Steven G. Widen[4] · Divya Mirchandani[5] · Peter J. Walker[6]

Received: 17 October 2018 / Accepted: 23 November 2018 / Published online: 3 December 2018

Abstract
Viruses in the family *Rhabdoviridae* are ecologically very diverse, infecting mammals, birds, reptiles, fish, plants and a wide range of other terrestrial and aquatic invertebrates. The genus *Sripuvirus* currently comprises five viruses that appear to circulate in transmission cycles involving reptiles and sandflies. Here, we report an analysis of the complete coding sequences of Cuiaba virus (CUIV), isolated from an amphibian (*Bufo marinus*) collected in Brazil, and Charleville virus (CHVV), isolated from sandflies (*Phlebotomus* sp.) and lizards (*Gehyra australis*), collected in Australia. CUIV and CHVV cluster phylogenetically with the sripuviruses in maximum likelihood trees generated from complete L protein (RdRp) sequences. They also share with sripuviruses unique features in genome organisation, including an additional gene (*U1*) between the matrix protein (*M*) gene and glycoprotein (*G*) gene, and an alternative long open reading frame near the start of the G ORF that encodes a predicted transmembrane protein. In view of their phylogenetic relationships, similar genome organisations and similar ecological characteristics, we propose the assignment of CUIV and CHVV as novel members of the genus *Sripuvirus*.

Keywords Rhabdovirus · Sprivivirus · Cuiaba virus · Charleville virus · Genome architecture

Introduction

The *Rhabdoviridae* is one of the most ecologically diverse families of RNA viruses with members infecting mammals, birds, reptiles, fish, plants and a wide range of other terrestrial and aquatic invertebrates [1, 2]. Many of the animal and plant rhabdoviruses are transmitted by arthropod vectors. The genus *Sripuvirus* currently comprises five viruses that appear to circulate in transmission cycles involving reptiles and sandflies [3, 4]. Chaco virus (CHOV; species *Chaco sripuvirus*) was isolated from lizards (*Ameiva ameiva ameiva* and *Kentropyx calcaratus*) collected in Brazil, in 1962 and 1963, respectively [5]. Sena Madureira virus (SMV; species *Sena Madureira sripuvirus*) was isolated from a lizard of the same Ameiva species collected in Brazil, in 1976 [6]. Sripur virus (SRIV; species *Sripur sripuvirus*) was isolated from sandflies (*Sergentomyia* sp.) collected in West Bengal, India in 1973 [6]. Niakha virus (NIAV; species *Niakha sripuvirus*) was isolated from a mixed pool of sandflies (*Phlebotomus duboscqi* and *Sergentomyia* sp.) collected in Senegal, in 1992 [7]. Almpiwar virus (ALMV; species *Almiwar sripuvirus*) was isolated from a skink (*Ablepharus boutonii virgatus*) trapped at Kowanyama, Queensland, Australia, in 1966 [8, 9]. Timbo virus (TIMV) was isolated from lizards (*Ameiva ameiva ameiva*) in Brazil, in 1962 and 1963 [5]. Based on serological cross-reactions and phylogenetic analysis using L protein sequences [5, 10, 11], TIMV is also likely to be a member of the genus *Sripuvirus* but has not yet been formally

Edited by Juergen A Richt.

✉ Peter J. Walker
peter.walker@uq.edu.au

[1] Department of Pathology, Center for Biodefence and Emerging Infectious Diseases, The University of Texas Medical Branch, Galveston, TX, USA

[2] Center for Tropical Diseases, The University of Texas Medical Branch, Galveston, TX, USA

[3] Institute for Human Infection and Immunity, The University of Texas Medical Branch, Galveston, TX, USA

[4] Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA

[5] Department of Microbiology and Immunology, The University of Texas Medical Branch, Galveston, TX, USA

[6] School of Biological Sciences, The University of Queensland, St Lucia, QLD, Australia

assigned. Hainan black-spectacled toad rhabdovirus (HnB-STRV), detected in a toad (*Duttaphrynus melanostictus*) collected in China [12], and Caiman lizard virus (CLV), isolated in 1999 from a lizard (*Dracaena guianensis*) imported to USA from Peru [10], are also related phylogenetically to the sripuviruses.

In this paper, we report the complete genome sequences of two additional rhabdoviruses that we propose as new members of the genus *Sripuvirus*. Cuiaba virus (CUIV) was isolated from a toad (*Bufo marinus*) captured in Para State, Brazil, in 1972 [6]. In a complement-fixation test, CUIV was shown to cross-react antigenically with Charleville virus (CHVV) [13]. CHVV was first isolated from 4 pools of sandflies (*Phlebotomus* sp.) collected in February 1969 at Charleville in south-western Queensland [14]. 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 [14]. 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 [14].

Materials and methods

Viruses

CUIV strain BeAn 227841 was originally obtained from the Instituto Evandro Chagas, Belem, Para State, Brazil. It had been serially passaged 15 times by intracerebral (IC) injection of newborn mice. CHVV strain Ch9824 was obtained from the World Reference Collection of Emerging Viruses and Arboviruses (WRCEVA) at the University of Texas Medical Branch, Galveston. CHVV had been serially passaged 5 times by IC injection of newborn mice. Both CUIV and CHVV are lethal to newborn mice within 4–6 days after IC inoculation [6]. Some of the earlier mouse passages had been conducted in Brazil (Instituto Evandro Chagas), Australia (Queensland Institute of Medical Research) and USA (Yale University).

RNA preparation and next-generation sequencing

Lyophilized virus stocks were reconstituted in 0.5 ml of PBS and viral RNA (vRNA) was extracted by the TRIZOL method, resuspended in 50 µl of RNAse-free water and quantified using a Nanopore instrument. The vRNA (~ 1 µg) was fragmented by incubation at 94 °C for 8 min in 19.5 µl of fragmentation buffer (Illumina 15016648). A sequencing library was prepared from the sample RNA using an Illumina TruSeq RNA v2 kit following the manufacturer's

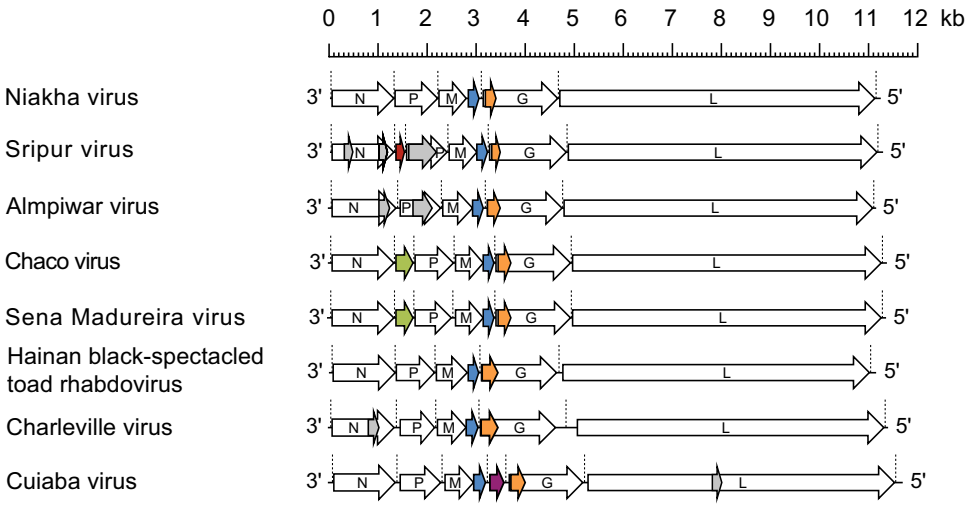


Fig. 1 Sripuvirus genome organisations. Open arrows indicate the locations of long open reading frames (ORFs) N, P, M, G and L, each of which is located within a transcriptional unit bounded by conserved transcription initiation and transcription termination/polyadenylation sequences. The Mx ORFs (shaded 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. The approximate boundaries of transcriptional units (gene junctions) are shown by vertical broken lines

protocol. The samples were sequenced on a NextSeq 550 using the 2 × 75 paired-end protocol. Reads in fastq format were quality-filtered and any adapter sequences were removed using Trimmomatic software [15]. The de novo assembly program ABySS version 2.0.2 [16] was used to assemble the reads into contigs using subsets of reads from 25,000 to 2 million pairs and kmer values from 20 to 40. The longest CHVV and CUIV viral contigs were obtained from 1 million read pairs. The reads were mapped back to the selected contigs using bowtie2 [17] and visualised with the Integrated Genomics Viewer [18] to verify that the assembled contigs were correct. Of ~ 16 million total read pairs, 170,070 read pairs (~ 1.07%) mapped to CHVV, and 1.3 million of 13.2 million total (~ 10%) mapped to CUIV.

Results

The complete genome sequence was determined for CUIV (11,700 nt) and the near-complete genome sequence (11,434 nt) was determined for CHVV, lacking nucleotides only at the extreme 3′- and 5′-termini. The sequences have been deposited in GenBank under accession numbers MH899110 (CUIV) and MH899109 (CHVV). BLAST searches (blastx) of the National Center for Biotechnology Information (NCBI) non-redundant protein sequence (nr) database indicated that each virus was most closely related to the five taxonomically assigned sripuviruses (ALMV, NIAV, SRIV, SMV and CHOV), as well as to unassigned Hainan black-spectacled toad rhabdovirus (HnBSTRV). The negative-sense (−) RNA genomes of CHVV and CUIV each contain the five canonical rhabdovirus structural protein genes (*N, P, M, G* and *L*) as individual transcriptional units with long open reading frames (ORFs) bounded by conserved transcription initiation (UUGU) and transcription termination/

A

```
CUIV        AAAAGGAGGGAAGAUGAUUACGAUUCCGAUUGAGUUUUUUGUCCCUUUUCUUGUUGUG
CHVV        CAACUGUGGGAACAAGCGAGUUGCUCUGCUGAAGUCAUAAUGUUUAAGAUUGCUUUUG
SRIV        GAGGCGUGGGAAAAAUUCUACUCAGAUGUGCAGAAGACCAUGAACUUGAUUUCAAUGG
NIAV        GAGGAGUGGGAAAUCACGUACUCAGAUCUGUGCAAGACCAUGAUUGUGUUGUCAGUGG
SMV         UAAAUUUGAGAUUAAUAAACGUCAUAUUUUGAUUUCACAGAAUGAUUGUCUUUCCUUU
CHOV        CCAAUUUGAGAUUCAAAAGAAUAAGAUUUUAAUAUCUCCUAAAUGAUUAUAUUACCCA
ALMV        CUCUCAUGGGAACAAAAGAGUUCAAAUCAAGAAGAUACAAUGAUUACCUUUAGUUUUG
HnBSTRV     GUAGAGUGGGGGACAAAGAGGCUAGUUCUGUACAAGGCAAAAUGAUCAAAAUUGGAGUG
```

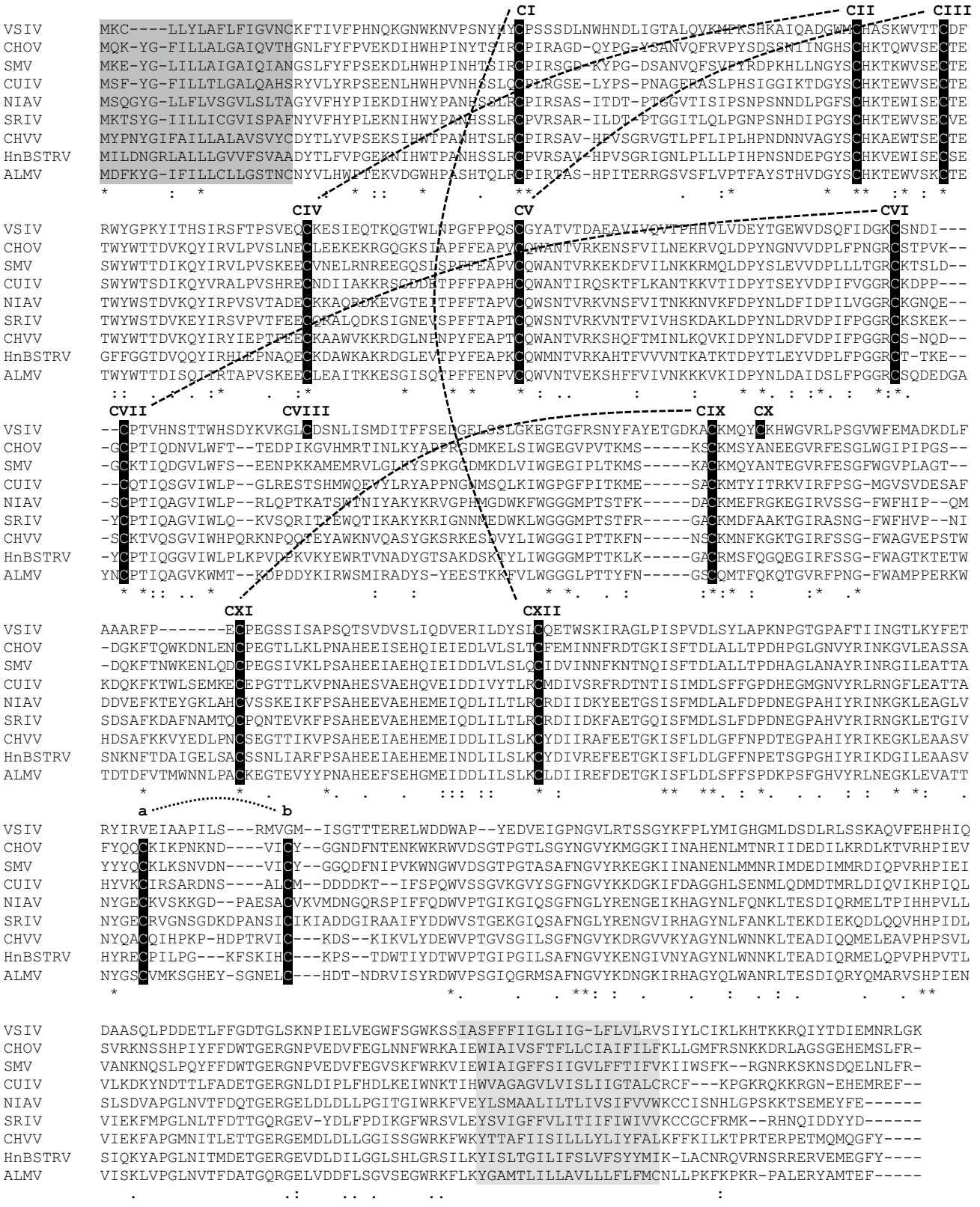
B

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--------
            *   :* *:  .
```

Fig. 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)


CI
CII
CIII
CIV
CV
CVI
CVII
CVIII
CIX
CX
CXI
CXII
a
b
VSIV
CHOV
SMV
CUIV
NIAV
SRIV
CHVV
HnBSTRV
ALMV

◂**Fig. 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 (C_I–C_{XII}; C_{II}–C_{IV}; C_{III}–C_V; C_{VI}–C_{VII}; C_{VIII}–C_X and C_{IX}–C_{XI}) and these bridges are conserved to various extents amongst rhabdoviruses in patterns that are somewhat genus-specific. Residues forming four of these bridges (C_I–C_{XII}; C_{III}–C_V; C_{VI}–C_{VII}; and C_{IX}–C_{XI}) are conserved in sripuviruses. Cysteines forming the C_{VIII}–C_X bridge are absent in all sripuviruses which have two additional conserved cysteines (*a, b*) downstream of C_{XII} that likely form an additional disulphide bridge. Predicted N-terminal signal domains (dark grey) and C-terminal transmembrane domains (light grey) are also shown

polyadenylation ($AC[U]_7$) sequences (Fig. 1). Like the sripuviruses and HnBSTRV, CHVV and CUIV also each contain an additional long open reading frame (249 nt) within the *M* gene (Mx), just overlapping the end of the M ORF (Fig. 1). In CHVV, the overlap occurs at the site of the M ORF termination codon; in CUIV, there is a short overlap of 14 nucleotides (Fig. 2). As observed previously for the sripuviruses [3], 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) that is complementary to the loop region of helix 26 of 18S ribosomal RNA [19]. Also as in the sripuviruses [3], there is an additional long ORF in the *G* genes of CHVV (351 nt) and CUIV (321 nt), in alternative reading frames commencing just downstream of the G ORF initiation codon (Fig. 1). These features also occur in HnBSTRV (Figs. 1, 2a). Conservation of each of the additional long ORFs in all of these viruses indicates that they are highly likely to be expressed. In CUIV, there is also an additional transcriptional unit (*U1* gene) between the *M* and *G* genes containing a long ORF (303 nt) (Fig. 1). There is no corresponding gene in CHVV, HnBSTRV or any of the classified sripuviruses.

Proteins encoded in the CHVV and CUIV *N, P, M, G* and *L* genes align with the corresponding structural proteins of the classified sripuviruses. 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 (C_I–C_{XII}; C_{III}–C_V; C_{VI}–C_{VII}; and C_{IX}–C_{XI}) (Fig. 3) [20, 21]. Like the other sripuviruses, CHVV and CUIV lack two cysteine residues that form another known disulphide bridge in VSIV (C_{VIII}–C_X) which links α-helix C to the *lm* loop in the pleckstrin homology domain (DIII) [21]. However, they 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) (Fig. 3) [21].

The CHVV and CUIV Mx ORFs encode proteins of 9.7 kDa and 9.8 kDa, respectively. A Clustal X alignment with the Mx proteins of the classified sripuviruses and HnBSTRV indicated that, although there is little global sequence identity, there is a high level of pairwise homology across the set with highest sequence conservation in the N-terminal domains. The CHVV and CUIV Gx ORFs encode proteins with calculated molecular weights of 12.9 kDa and 11.8 kDa, respectively. Like each of the classified sripuvirus (and HnBSTRV) Gx proteins, the TMHMM algorithm (http://www.cbs.dtu.dk/services/TMHMM-2.0) predicts that each contains a single transmembrane helix with highest probability ($p = 0.78$–0.65) of an N-terminal exodomain. The Gx proteins also display high levels of pairwise homology in a Clustal X alignment (Fig. 2b). The Mx and Gx proteins have not yet been identified in infected cells and their functions are unknown. The CUIV U1 ORF encodes a 11.2 kDa protein that is relatively rich in proline (9.9%) and basic amino acids (17.8%). As this ORF lies within a unique transcriptional unit bounded by conserved transcription initiation and transcription termination/polyadenylation sequences, it is highly likely to be expressed in infected cells. However, the sequence of the U1 protein has no obvious features that may suggest its function.

A phylogenetic tree was inferred from a Clustal X alignment of the L proteins of CHVV, CUIV, HnBSTRV and 110 other animal rhabdoviruses that are currently classified to species in the family *Rhabdoviridae*. The tree indicated that the CHVV, CUIV and HnBSTRV cluster with the sripuviruses in a distinct and well-supported monophyletic clade (Fig. 4). By this analysis, CHVV is most closely related to HnBSTRV which also clusters with ALMV; CUIV clusters with CHOV and SMV.

Pairwise sequence identities (p-distances) were calculated in MEGA7 from Clustal X alignments of the N, G and L proteins of CHVV, CUIV, HnBSTRV and the classified sripuviruses. By this analysis, 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, 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.

Discussion

The data presented here indicate that both CHVV and CUIV, as well as HnBSTRV, should be classified as members of the genus *Sripuvirus*. They cluster in a monophyletic clade with the sripuviruses in trees inferred from L protein sequences. Uniquely amongst known rhabdoviruses, the viruses represented in this clade have a strong ecological association with reptiles and amphibians, as well as phlebotomine sandflies. The viruses also have similar genome organisations that

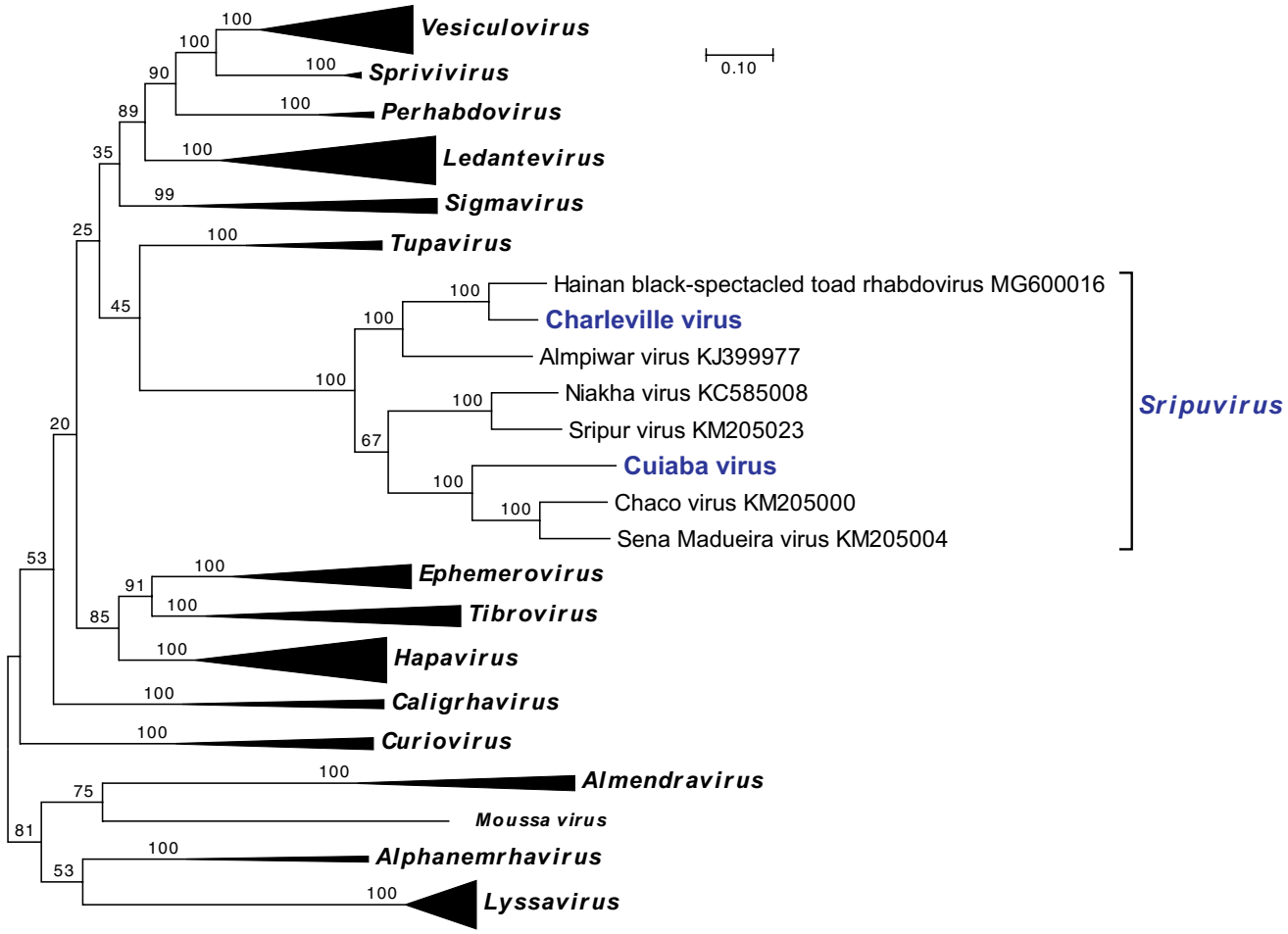


Fig. 4 The evolutionary history was inferred from a Clustal W alignment of complete L protein sequences of 3 proposed sripuviruses and 110 other animal rhabdoviruses currently assigned or recently proposed for assignment to species. Phylogenetically informative sites were selected from the alignment 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

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	

Values shown in italics indicate those viruses that display highest levels of amino acid sequence identity

are unique and characteristic of sripuviruses, including an additional long ORF (Mx) in the *M* gene, overlapping the distal end of the M ORF, and an alternative long ORF (Gx), encoding a predicted transmembrane protein, at the start of

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	

Values shown in italics indicate those viruses that display highest levels of amino acid sequence identity

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	

Values shown in italics indicate those viruses that display highest levels of amino acid sequence identity

the *G* gene. Expression of the Mx ORF appears likely to occur by a stop-start mechanism of translation due to the presence in all sripuviruses of a TURBS-like sequence immediately upstream of the region of overlap with the M ORF. TURBS-like sequences preceding similar ORF overlaps have also been reported to occur in viruses assigned to several other genera of the *Rhabdoviridae*, including hapaviruses (*G* gene), ephemeroviruses (*U1* or *α* gene) and curioviruses (*U1* gene) [3, 22]. Expression of the Gx proteins is likely to occur by leaky ribosomal scanning, allowing initiation of transcription by a proportion of ribosomes on the alternative start codon. Expression of rhabdovirus proteins from an alternative ORF by using alternative initiation codons near the start of the *P* gene has been demonstrated for vesicular stomatitis Indiana virus [23, 24]. In rabies virus, leaky ribosomal scanning is employed to express several N-terminal truncated forms of the P protein that are involved in nuclear transport [25]. Expression of the sripuvirus Mx and Gx ORFs has not yet been demonstrated experimentally.

Species demarcation criteria published by the International Committee on Taxonomy of Viruses (ICTV) for rhabdoviruses in the genus *Sripuvirus* require several (but not all) of the following: (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 (https://talk.ictvonline.org/ictv-reports/ictv_online_report/negative-sense-rna-viruses/mononegavirales/w/rhabdoviridae) [4]. 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. Neutralisation test results are not currently available. Overall, the data support the classification of CUIV, CHVV and HnBSTRV as separate new species in the genus *Sripuvirus*, and this will be the subject of a formal taxonomic proposal to the ICTV.

Acknowledgements Work at UTMB was supported by Grant R24 AI120942 (National Institute of Allergy and Infectious Diseases) and contract HHSN272004/D04 from the National Institutes of Health.

Author contributions PJW, NV and RBT conceived and planned experiments. NV, SGW and DM performed experiments. PJW and NV analysed data. PJW, NV and RBT wrote the manuscript. All authors reviewed and approved the manuscript.

References

1. Kuzmin IV, Novella IS, Dietzgen RG, Padhi A, Rupprecht CE (2009) The rhabdoviruses: biodiversity, phylogenetics, and evolution. Infect Genet Evol 9:541–553
2. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, Chen LJ, Qin XC, Xu J, Holmes EC, Zhang YZ (2015) Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife 4:e05378
3. Walker PJ, Firth C, Widen SG, Blasdell KR, Guzman H, Wood TG, Paradkar PN, Holmes EC, Tesh RB, Vasilakis N (2015) Evolution of genome size and complexity in the *Rhabdoviridae*. PLoS Pathog 11:e1004664
4. Walker PJ, Blasdell KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, Longdon B, Stone DM, Tesh RB, Tordo N, Vasilakis N, Whitfield AE, Consortium IR (2018) ICTV virus taxonomy profile: *Rhabdoviridae*. J Gen Virol 99:447–448
5. Causey OR, Shope RE, Bensabath G (1966) Marco, timbo, and chaco, newly recognized arboviruses from lizards of Brazil. Am J Trop Med Hyg 15:239–243
6. Karabatsos N (ed) (1985) International catalogue of arboviruses including certain other viruses of vertebrates. American Society for Tropical Medicine and Hygiene, San Antonio
7. Vasilakis N, Widen S, Mayer SV, Seymour R, Wood TG, Popov V, Guzman H, Travassos da Rosa AP, Ghedin E, Holmes EC, Walker PJ, Tesh RB (2013) Niakha virus: a novel member of the family *Rhabdoviridae* isolated from phlebotomine sandflies in Senegal. Virology 444:80–89
8. McAllister J, Gauci PJ, Mitchell IR, Boyle DB, Bulach DM, Weir RP, Melville LF, Davis SS, Gubala AJ (2014) Genomic characterisation of Almpiwar virus, Harrison Dam virus and Walkabout Creek virus; three novel rhabdoviruses from northern Australia. Virol Rep 3:1–17
9. Doherty RL, Whitehead RH, Wetters EJ, Gorman BM, Carley JG (1970) A survey of antibody to 10 arboviruses (Koongol group, Mapputta group and ungrouped) isolated in Queensland. Trans R Soc Trop Med Hyg 64:748–753
10. 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*). Vet Microbiol 158:274–279
11. Monath TP, Cropp CB, Frazier CL, Murphy FA, Whitfield SG (1979) Viruses isolated from reptiles: identification of three new members of the family *Rhabdoviridae*. Arch Virol 60:1–12
12. 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
13. Tesh RB, Travassos da Rosa APA, Travassos da Rosa JS (1983) Antigenic relationship among rhabdoviruses infecting terrestrial vertebrates. J Gen Virol 64:169–176
14. 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. Trans R Soc Trop Med Hyg 67:536–543
15. Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res 40:W622–W627
16. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19:1117–1123
17. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359
18. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26
19. Firth AE, Brierley I (2012) Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409
20. Walker PJ, Kongsuwan K (1999) Deduced structural model for animal rhabdovirus glycoproteins. J Gen Virol 80:1211–1220
21. 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
22. Allison AB, Mead DG, Palacios GF, Tesh RB, Holmes EC (2014) Gene duplication and phylogeography of North American members of the Hart Park serogroup of avian rhabdoviruses. Virology 448:284–292
23. Spiropoulou CF, Nichol ST (1993) A small highly basic protein is encoded in overlapping frame within the P gene of vesicular stomatitis virus. J Virol 67:3103–3110
24. Peluso RW, Richardson JC, Talon J, Lock M (1996) Identification of a set of proteins (C' and C) encoded by the bicistronic P gene of the Indiana serotype of vesicular stomatitis virus and analysis of their effect on transcription by the viral RNA polymerase. Virology 218:335–342
25. Pasdeloup D, Poisson N, Raux H, Gaudin Y, Ruigrok RW, Blondel D (2005) Nucleocytoplasmic shuttling of the rabies virus P protein requires a nuclear localization signal and a CRM1-dependent nuclear export signal. Virology 334:284–293