JOURNAL OF VIROLOGY, July 2011, p. 6185–6196 Vol. 85, No. 13
0022-538X/11/$12.00 doi:10.1128/JVI.00040-11

RNA Splicing in a New Rhabdovirus from *Culex* Mosquitoes▽†

Ryusei Kuwata,[1] Haruhiko Isawa,[1]* Keita Hoshino,[1] Yoshio Tsuda,[1] Tohru Yanase,[2] Toshinori Sasaki,[1] Mutsuo Kobayashi,[1] and Kyoko Sawabe[1]

Department of Medical Entomology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan,[1] and Kyushu Research Station, National Institute of Animal Health, NARO, 2702 Chuzan, Kagoshima 891-0105, Japan[2]

Received 7 January 2011/Accepted 8 April 2011

Among members of the order *Mononegavirales*, RNA splicing events have been found only in the family *Bornaviridae*. Here, we report that a new rhabdovirus isolated from the mosquito *Culex tritaeniorhynchus* replicates in the nuclei of infected cells and requires RNA splicing for viral mRNA maturation. The virus, designated *Culex tritaeniorhynchus* rhabdovirus (CTRV), shares a similar genome organization with other rhabdoviruses, except for the presence of a putative intron in the coding region for the L protein. Molecular phylogenetic studies indicated that CTRV belongs to the family *Rhabdoviridae*, but it is yet to be assigned a genus. Electron microscopic analysis revealed that the CTRV virion is extremely elongated, unlike virions of rhabdoviruses, which are generally bullet shaped. Northern hybridization confirmed that a large transcript (approximately 6,500 nucleotides [nt]) from the CTRV L gene was present in the infected cells. Strand-specific reverse transcription-PCR (RT-PCR) analyses identified the intron-exon boundaries and the 76-nt intron sequence, which contains the typical motif for eukaryotic spliceosomal intron-splice donor/acceptor sites (GU-AG), a predicted branch point, and a polypyrimidine tract. *In situ* hybridization exhibited that viral RNAs are primarily localized in the nucleus of infected cells, indicating that CTRV replicates in the nucleus and is allowed to utilize the host's nuclear splicing machinery. This is the first report of RNA splicing among the members of the family *Rhabdoviridae*.

Rhabdoviruses (family *Rhabdoviridae*) are enveloped viruses with a nonsegmented negative-strand (NNS) RNA genome in the order *Mononegavirales*, which also includes three other families—*Bornaviridae*, *Filoviridae*, and *Paramyxoviridae*. Rhabdoviruses have a diverse host range and include many serious pathogens that affect human and animal health as well as aquaculture and agricultural industries (68). At present, the International Committee on Taxonomy of Viruses recognizes six genera in the family *Rhabdoviridae*. Members of the genera *Nucleorhabdovirus* and *Cytorhabdovirus* are plant viruses mainly transmitted by arthropod vectors, while those of the genera *Novirhabdovirus*, *Lyssavirus*, *Ephemerovirus*, and *Vesiculovirus* have been isolated from various vertebrate and invertebrate hosts. Although to date more than 150 rhabdoviruses have been found from various hosts worldwide, most of the viruses remain to be assigned to a genus within the family because of the genetic diversity among them.

In general, the rhabdovirus genome with negative polarity is 11,000 to 15,000 nucleotides (11 to 15 knt) in size and contains at least five genes encoding viral structural proteins in the following order: 3′-nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and large protein (L)-5′ (68). Each gene is primarily regulated by *cis*-acting transcription start and stop signals located at the beginning and end of the gene, respectively. Consequently, the gene junction consists of highly conserved sequences of the stop signal of the upstream gene, an intergenic region (IGR), and the start signal of the downstream gene. Like other members of the order *Mononegavirales*, the 3′ leader and 5′ trailer regions at the genome termini are notable for their high A/U content with the potential for complementarity between the first 10 to 20 nt to form a stable panhandle structure, believed to contain the genomic and antigenomic promoters essential for viral replication and transcription (83). The three nucleotides (UGC) at the 3′ termini are perfectly conserved among most of the animal-infecting rhabdoviruses, except for novirhabdoviruses, in which they are believed to be essential for polymerase function (81).

The morphology of the rhabdovirus virion generally differs according to the host when fixed prior to negative staining; animal-infecting rhabdoviruses are commonly bullet or cone shaped with one rounded tip and one flat tip, while plant-infecting rhabdoviruses appear bacilliform with two rounded tips. The virions are mostly 100 to 350 nm in length and 45 to 100 nm in diameter, and plant-infecting rhabdovirus virions tend to be longer than animal-infecting rhabdovirus virions. Virion RNA is tightly encapsidated by the N protein to form a helical nucleocapsid, and the capsid is associated with the two proteins, P and L, to form ribonucleoprotein complexes (RNPs). The L protein serves as an RNA-dependent RNA polymerase (RdRp) for viral transcription and replication in cooperation with the P protein, which plays an important accessory role in the regulation of viral RNA synthesis. The M protein is the inner component of the virion and forms a layer between the nucleocapsid and the outer membrane. The outer surface of the envelope is covered with transmembrane spikes formed by trimerized G proteins, which are 5 to 10 nm in length and approximately 3 nm in diameter.

* Corresponding author. Mailing address: Department of Medical Entomology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Phone: 81 3 5285 1111. Fax: 81 3 5285 1147. E-mail: hisawa@nih.go.jp.

† Supplemental material for this article may be found at http://jvi.asm.org/.

▽ Published ahead of print on 20 April 2011.

Replication of rhabdoviruses mainly occurs in the cytoplasm of an infected cell. However, plant-infecting viruses of the genus *Nucleorhabdovirus* vary profoundly in their sites of replication and morphogenesis; they replicate in the nucleus, and the virions bud predominantly from the inner nuclear membrane (1, 21). After cellular invasion, rhabdoviruses are believed to associate with the endoplasmic reticulum for uncoating and releasing the nucleocapsid core into the cytoplasm, and at this point, the replication cycles differ in the *Nucleorhabdovirus* from those of other rhabdoviruses (26). According to the model of nucleorhabdovirus replication based mainly on the studies conducted in the sonchus yellow net virus (SYNV), the released nucleocapsid enters the cell nucleus through the nuclear pore complex and then undertakes transcription and replication of viral RNAs. The generated viral mRNAs are exported to the cytoplasm and translated to proteins. The synthesized viral proteins, which are essential for viral transcription and replication, are transported back to the nucleus to continue transcription of mRNA and replication of the progeny of genomic RNA.

Because many rhabdoviruses are mechanically or biologically transmitted by insect vectors to their vertebrate or plant hosts, insects could play a central role in the horizontal transmission of rhabdoviruses (21, 43, 44, 46, 65). In this study we characterized a new rhabdovirus from the mosquito *Culex tritaeniorhynchus* (Diptera: Culicidae) that is known to be a major vector of the Japanese encephalitis virus ([JEV] family *Flaviviridae*) in Asia. Taxonomic classification based on genomic structure and molecular phylogenetic studies indicated that the new mosquito virus belonged to the family *Rhabdoviridae* of the order *Mononegavirales*. Surprisingly, we found that this new rhabdovirus has an intron-like sequence ("putative intron") in the viral genomic RNA and replicates in the host cell nucleus for RNA splicing. Electron microscopic analyses revealed that the virion is elongated; this morphology is extremely different from the morphologies of other known rhabdoviruses. Our experimental data strongly suggested that the elongated virus, designated *Culex tritaeniorhynchus* rhabdovirus (CTRV), requires RNA splicing for viral mRNA maturation, and to our knowledge, this is the first report of viral RNA splicing among the members of the family *Rhabdoviridae*.

MATERIALS AND METHODS

Mosquito collection. Blood-fed *C. tritaeniorhynchus* mosquitoes were captured in Tateyama City, Chiba Prefecture, Japan, in pigpens. Collected mosquitoes were reared until the blood was fully digested. Then, they were sorted into pools and kept frozen at −80°C.

Cell cultures and virus isolation. The mosquito cell line C6/36 (Health Science Research Resources Bank, Osaka, Japan) was used for further study. Cell cultures and virus isolation were conducted by the method described previously (25). Briefly, mosquito homogenates were centrifuged, and the supernatants were passed through sterile 0.45-μm-pore-size filters. The filtrates were diluted and inoculated on monolayers of C6/36 cells, and the cells were incubated at 28°C in 0.5% CO_2 conditions for 7 days. After two additional blind passages, the supernatants were harvested and stored as viral stocks at −80°C.

Determination of the complete genome sequence. A cDNA library was constructed using RNA extracted from culture supernatant of infected cells using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA). For the initial amplification of viral sequences, a universal primer set for the NS5 regions of flaviviruses (34) was used, resulting in the unexpected amplification of a partial rhabdoviral RdRp-like sequence (see Results section). To amplify unknown sequences flanking the above sequence, a cDNA walking strategy based on cassette ligation-mediated PCR was performed using an LA PCR *in vitro* Cloning Kit (Takara Bio, Shiga, Japan). After a nearly full-length genome sequence was obtained, its accuracy and completeness were confirmed by a long PCR using the same cDNA library and TaKaRa LA *Taq* (Takara Bio). To determine the 5′- and 3′-terminal sequences of the viral genome, rapid amplification of the cDNA ends (RACE) was performed using a Thermoscript reverse transcription-PCR (RT-PCR) kit (Invitrogen) according to the method described previously (39). Nucleotide sequence analyses were performed using GENETYX software, version 9 (GENETYX Co., Tokyo, Japan) and the BLAST program on the National Center for Biotechnology Information website. Amino acid sequence identity and similarity scores were calculated between putative CTRV proteins and equivalent proteins of vesicular stomatitis Indiana virus (VSIV; GenBank accession number AF473864) or rabies virus (RABV; GenBank accession number GU345748). Protein sequences were analyzed and characterized using the ExPASy Proteomics server (http://au.expasy.org) and Phobius (http://phobius.cbr.su.se/).

Gene start/end and IGRs. The rhabdovirus genome contains conserved sequences serving as *cis*-acting transcription start/stop signals located at the beginning/end of each gene. To identify the *cis*-regulatory elements in CTRV genome, we examined the 5′ initiation and the 3′ polyadenylation site for each viral transcript by RACE. Briefly, total RNAs were extracted from CTRV-infected C6/36 cells using an RNeasy Mini Kit (Qiagen Inc.). Poly(A) RNAs were separated from total RNAs using a Poly(A)$^+$ Isolation Kit from total RNA (Nippon Gene Co., Tokyo, Japan) and then were subjected to 5′ and 3′ RACE analyses using a GeneRacer kit (Invitrogen).

Phylogenetic study. Phylogenetic studies were conducted based on the protein sequences of conserved domain block III in the L protein among selected members of the order *Mononegavirales* and the family *Rhabdoviridae* (52). In addition, we also constructed a phylogenetic tree based on the partial protein sequences of the N protein among selected members of the family *Rhabdoviridae* to avoid genetic bias (35). Multiple sequence alignment matrices were created using Clustal X, version 2.0.8 (37), in the default configuration. The aligned matrix data were analyzed using neighbor joining (NJ) algorithms using the programs PROTDIST, NEIGHBOR, SEQBOOT, and CONSENSE including the software package PHYLIP, version 3.69 (J. Felsenstein, Department of Genome Sciences, University of Washington, Seattle, WA). Distance matrices of each sequence for partial L and N proteins were established using the PMB (probability matrix from blocks) (76) and JTT (28) models, respectively. The statistical significance of the resulting NJ trees was evaluated using a bootstrap test with 1,000 replications (12). The trees were represented graphically using TreeView software, version 1.6.6 (version 1.6.6; R. D. M. Page, University of Glasgow, Glasgow, United Kingdom [http://taxonomy.zoology.gla.ac.uk/rod/treeview.html]).

Electron microscopic analyses. Culture supernatant from CTRV-infected C6/36 cells was sampled for examination of virions. First, the supernatant was centrifuged at 1,600 × *g* for 15 min, and then the upper fluid was centrifuged again at 12,000 × *g* for 30 min to remove cellular debris. Next, CTRV virions in the fluid were precipitated by ultracentrifugation at 77,000 × *g* for 2 h in a Beckman TLS55 rotor (Beckman Coulter, Cedar Grove, NJ). The resulting precipitate was suspended in and dialyzed against a buffer (20 mM Tris-HCl, 130 mM NaCl, and 1 mM EDTA, pH 7.0), fixed in 2% glutaraldehyde, and processed for negative-staining electron microscopy with 2% uranyl acetate. In addition, we conducted thin-section electron microscopic analysis of the CTRV-infected C6/36 cells inoculated with the virus stock for 8 days. CTRV-infected cells were fixed in 2% glutaraldehyde at 4°C overnight and then fixed in 2% osmic acid for 1 h on ice. The cell pellets were dehydrated and embedded in Epon 812, and thin sections were cut and stained with uranyl acetate.

Northern hybridization. The CTRV L gene transcript was examined by Northern hybridization using digoxigenin (DIG)-labeled sense and antisense RNA probes. A 859-bp region upstream of the putative intron (sequence position 4894 to 5752; probe LPB1) and a 720-bp region downstream of the putative intron (sequence position 10406 to 11125; probe LPB2) were selected as the probes, which were initially amplified by RT-PCR (primer sequences are shown in Table S1 in the supplemental material). To generate the template for *in vitro* transcription, T7 polymerase promoter sequence was added to the first PCR product by serial PCRs (primer sequences used are available upon request). The PCR product containing a complete T7 promoter sequence was used as the template for generating DIG-labeled RNA probes with a DIG RNA Labeling Kit (Roche Diagnostics, Tokyo, Japan). Total RNA extracted from the CTRV-infected C6/36 cells was electrophoresed in a 0.8% denaturing formaldehyde agarose gel, transferred to a nylon membrane (Hybond-N+; GE Healthcare, Uppsala, Sweden), and UV cross-linked to the membrane. After a 30-min prehybridization, the probe was hybridized to the blot in the DIG Easy Hyb (Roche Diagnostics) at 68°C overnight. The membrane was washed twice with 2× SSC (0.1% SDS) at room temperature for 10 min and with 0.1× SSC (0.1% SDS; 1× SSC is 0.15 M

NaCl plus 0.015 M sodium citrate) at 68°C for 20 min. The RNA-RNA hybrid was detected by immunodetection with alkaline phosphatase (AP)-conjugated anti-DIG antibody Fab fragments in blocking solution, and the hybridization signals were detected using CDP-Star (Roche Diagnostics).

Analysis of the L gene transcript. The coding region for the L gene of CTRV is interrupted by a putative 76-nt intron, raising the possibility that the L gene transcript undergoes splicing to produce a mature L mRNA. To examine this possibility, strand-specific RT-PCRs were performed to detect and distinguish several viral RNA species (viral genomic RNA, antigenomic RNA, and mRNA) in the infected cells. Viral first-strand cDNAs from the infected cells were synthesized using avian myeloblastosis virus (AMV) reverse transcriptase (Takara Bio) with each of the four gene-specific primers located in and around the putative intron: primers A and B for complementary to positive-sense viral RNAs and primers C and D for complementary to negative-sense viral RNA (primer sequences are shown in Table S1 in the supplemental material). After the reverse transcriptase was inactivated, each resultant cDNA was used as a template for PCR with different primer combinations (see Results section and Fig. 5A). The PCR products were separated by agarose gel electrophoresis, extracted from the gels, and sequenced.

***In situ* hybridization.** Because the infection of C6/36 cells with CTRV did not result in the destruction of cells, it was thus possible to subculture the infected cells and analyze infection dynamics. In order to determine the subcellular localization of viral RNA in the infected cells, RNA-RNA *in situ* hybridization (ISH) was performed against the C6/36 cells persistently infected with CTRV using DIG-labeled RNA probes. A 558-bp region within the M gene (sequence position 2499 to 3056) and three regions within the L gene (LPB1 and LPB2 regions and putative 76-nt intron region) were selected as the probes, which were initially amplified by RT-PCR (primer sequences are shown in Table S1 in the supplemental material). Preparation of the templates for *in vitro* transcription and DIG-labeled RNA probes were performed as previously described. CTRV-infected C6/36 cells mounted on glass slides were fixed in buffered 4% paraformaldehyde with 5% acetic acid for 30 min at room temperature, permeabilized with 1 μg/ml proteinase K (Takara Bio) at 37°C for 10 min, and fixed again for 10 min in fixation buffer with the same constituent. After 2 h of prehybridization, fixed cells were treated under coverslips with a 0.5 μg/ml DIG-labeled RNA probe in hybridization buffer (50% formamide, 2× SSC, 1 μg/μl tRNA, 1 μg/μl salmon sperm DNA, 1 μg/μl bovine serum albumin [BSA], and 10% dextran sulfate) at 42°C for 16 h. After hybridization, the slides were washed once with 50% formamide in 2× SSC at 42°C for 30 min, treated with 4 μg/ml of RNase A at 37°C for 30 min, and rinsed twice with both 2× SSC (0.1% SDS) and 0.2× SSC (0.1% SDS) at 42°C for 20 min. The RNA-RNA hybrid was detected by immunodetection with AP-conjugated anti-DIG antibody Fab fragments in blocking solution and visualized with 5-nitroblue tetrazolium/bromo-4-chloro-3-indolylphosphate stain (NTB/BCIP; Roche Diagnostics).

Nucleotide sequence accession number. The complete CTRV genome sequence has been deposited in the GenBank database under accession number AB604791 (see Fig. S1 in the supplemental material).

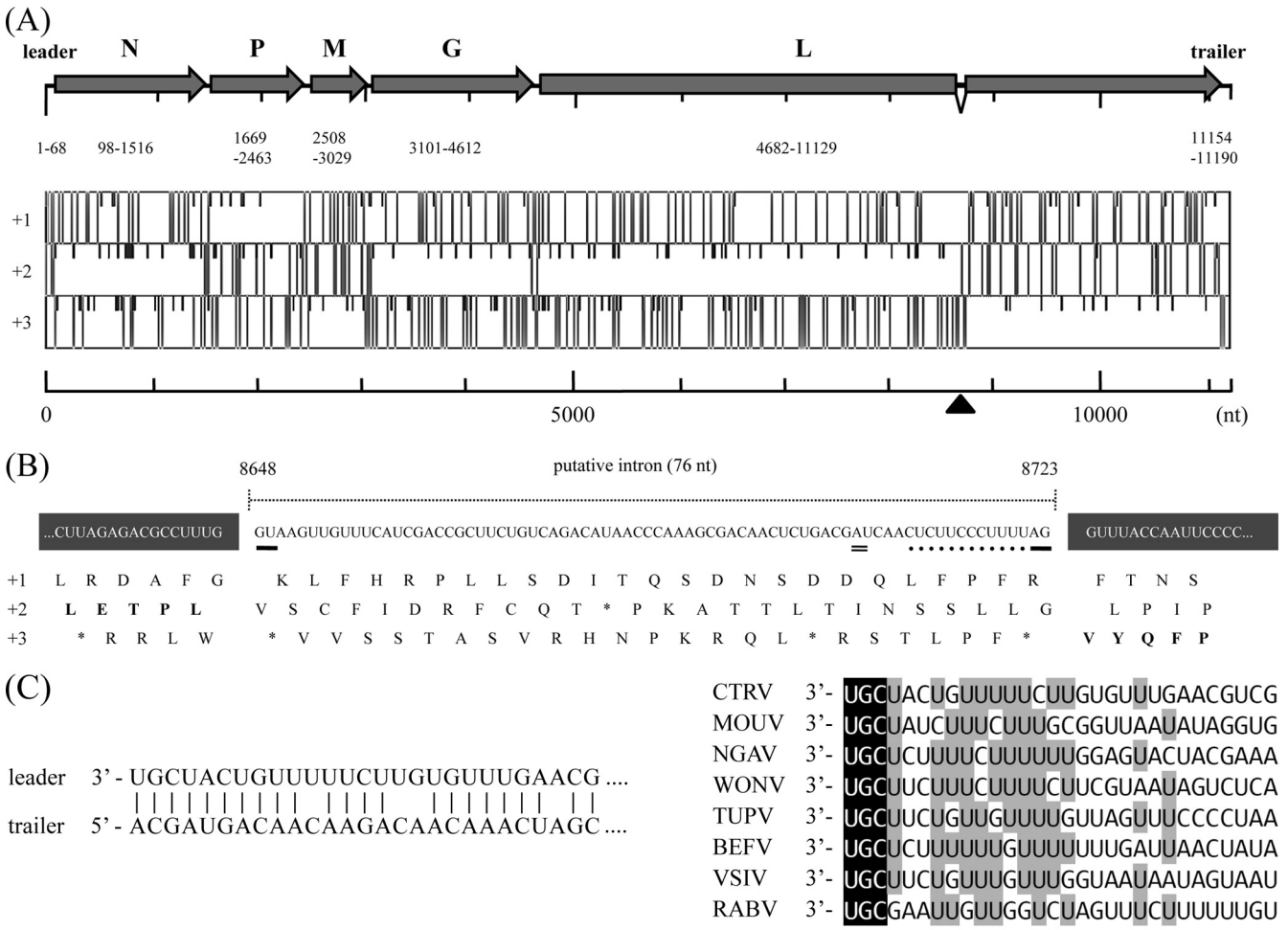
RESULTS

Isolation of CTRV. During the course of JEV isolation from *C. tritaeniorhynchus* mosquitoes, we observed that several mosquito homogenate samples from Tateyama City showed moderate cytopathic effects (CPE) on C6/36 cells. Because such CPE were speculated to be caused by infection with *Culex flavivirus* (CxFV) (24), RT-PCR with a primer set for detecting the flavivirus NS5 gene (34) was performed on RNAs extracted from cell culture supernatants. However, an approximately 180-nt fragment (sequence position 6309 to 6451) was amplified instead of the expected 260-nt fragment (data not shown). Sequence analysis revealed that the amplified fragment unexpectedly showed significant similarity to a portion of the L protein of the family *Rhabdoviridae* (*e.g.*, 72% identity to a portion of the L protein of Le Dantec virus, GenBank accession number AAZ43278), indicating that the infectious agent is not CxFV but a previously undescribed rhabdovirus. On the other hand, the viral stock did not cause any CPE on vertebrate cells (Vero and BHK-21), and no RT-PCR products derived from the viral genome were obtained with RNA from these inoculated cells (data not shown). The new virus was tentatively designated CTRV.

Genomic structure of CTRV. The genome sequence of CTRV was 11,190 nt in length and contained the predicted open reading frames (ORFs) encoding five structural proteins for rhabdoviruses in the following order: 3′-N-P-M-G-L-5′ (Fig. 1A). BLAST analyses indicated that ORFs located at sequence positions 98 to 1516 (Fig. 1A, frame +2), 3101 to 4612 (frame +2), and the following large portion of the genome had similarities to the N, G, and L proteins of other rhabdoviruses, respectively. However, the coding region for the putative L protein was divided into two different reading frames (frames +2 and +3) around sequence position 8700 (Fig. 1A). The unusual coding strategy for the CTRV L protein was further examined as described below. The remaining two ORFs located at sequence positions 1669 to 2463 (frame +1) and 2508 to 3029 (frame +3) could correspond to P and M genes of rhabdoviruses, respectively, although the sequences have no homologies to any known proteins, as is the case for other rhabdoviruses.

Putative intron of the CTRV L gene. Rhabdoviral L protein, which is responsible for most enzymatic activities as RdRp, is composed of more than 2,000 amino acids encoded by a single ORF (68). However, in CTRV, the coding region for the L protein was interrupted by an in-frame stop codon at sequence position 8681 (Fig. 1A and B, frame +2) and was divided into two different reading frames. BLAST analyses indicated that amino acid sequences of the two reading frames have significant similarities to the N- and C-terminal domains of rhabdovirus L protein, respectively. This finding raised the possibility that CTRV may have an unusual gene regulation mechanism for expressing a continuous L protein from the two reading frames. Ribosomal frameshifting is unlikely to occur because some in-frame stop codons exist between the posterior region of frame +2 and the anterior region of frame +3 (Fig. 1B). Interestingly, the GENETYX program predicted a putative 76-nt intron (sequence position 8648 to 8723) on the basis of the typical splice sites, GU-AG boundaries, and pyrimidine-rich tract (Fig. 1B). The putative 76-nt intron contained a predicted branch point sequence, CUGAC (the branch point is at A), which would be the core splicing signal lying approximately 18 nt upstream from the 3′ splice site (e.g., YUNAY for human [15]). If the putative 76-nt intron is spliced out, the coding region for the L protein is not interrupted by the in-frame stop codons and constitutes a single, large ORF encoding 2,123 amino acids, which showed significant similarities to L proteins of other rhabdoviruses. This raises the possibility that CTRV genomic RNA contains a 76-nt intron that is spliced out for generating mature L mRNA. This hypothesis was verified by experimental studies as described below.

Leader and trailer sequences. The genome terminal sequences of NNS RNA viruses exhibit inverse complementarity between the 3′ leader and 5′ trailer regions, which are believed to be essential for viral replication and transcription (83). The leader and trailer regions in the CTRV genome are 68 and 37 nt in length, respectively, and their terminal nucleotides (~27 nt) are highly complementary to each other (Fig. 1C). In general, the 3′ terminal trinucleotides (UGC) and the following U-rich domain are common features among many animal-



(D)

Gene junctions	Gene end	IGR	Gene start
leader/N			**AACAU**UAGUCA
N/P	AUUCA**UGAAAAAAA**	CAUC	**AACAU**UAAUCA
P/M	AAACG**UGAAAAAAA**	CUC	**AACAU**CACUCU
M/G	GCUCA**UGAAAAAAA**	CUCUU	**AACAU**GAAUCA
G/L	CAA**AACAU**UAACAUAACUAACUCGAACUCA**UGAAAAAAA**CAU		
L/trailer	UCACG**UGAAAAAAA**		

FIG. 1. Genome organization of CTRV. (A) The upper diagram represents the schematic transcription map of the 11,190-nt CTRV positive-strand antigenome. The shaded arrows indicate the positions of the ORFs for the five typical rhabdovirus proteins, N, P, M, G, and L. The numbers below the arrows indicate the sequence positions of the leader, ORFs, and trailer. The lower diagram represents the positions of start codons (short vertical lines) and stop codons (long vertical lines) in each of the three reading frames on the positive strand. Note that the coding region for the CTRV L protein is interrupted by an in-frame stop codon around sequence position 8700 (black arrowhead). (B) Antigenomic nucleotide sequences (upper) and predicted amino acid sequences of the three reading frames (lower) around the putative intron in the L gene. The 76-nt intron (8648 to 8723 nt) contains typical splice site motifs—GU-AG boundaries (underlines), a predicted branch point (double underline), and a polypyrimidine tract (dotted line). If the intron is spliced out, the translational reading frame shifts from +2 to +3 to generate a continuous L protein. (C) Complementarity of the 3′ and 5′ termini of the CTRV genome (left) and an alignment of the leader sequence of CTRV with those of other animal-infecting rhabdoviruses (right). The conserved 3′ terminal trinucleotides UGC are indicated in reverse type. The nucleotides shaded in gray indicate more than 50% sequence identity. (D) Sequences of the gene junctions in the CTRV antigenomic RNA. The upstream gene end sequences, IGRs, and the downstream gene start sequences are indicated. Consensus sequences of the predicted transcription start/stop signals are shown in bold. Note that the transcription start signal of the L gene precedes the transcription stop signal of G gene, resulting in a 36-nt gene overlap at the G/L junction.

TABLE 1. CTRV gene and predicted protein

Protein	Gene length (nt)	ORF length (nt)	5′ UTR length (nt)	3′ UTR length (nt)	Protein length (aa)	Protein mass (kDa)	pI
N	1,566	1,419	29	118	472	52.8	6.25
P	838	795	30	13	264	28.0	9.91
M	594	522	28	44	173	18.8	9.79
G	1,602	1,512	22	68	503	55.0	7.79
L	6,509	6,448[a]	37	24	2,123	239.2	6.18

[a] The number of nucleotides includes a putative 76-nt intron.

infecting rhabdoviruses, which were also observed in the terminal extracistronic region of CTRV (Fig. 1C).

Gene junctions. Gene expression in rhabdoviruses is regulated by *cis*-acting conserved sequences of transcription start/stop signals located at the beginning/end of each gene, and the gene junctions generally consist of the transcription stop (polyadenylation) signal of the upstream gene, IGR, and transcription start signal of the downstream gene. RACE analysis revealed that the consensus sequences AACAU and UGAAA AAAA were determined at the 5′ and 3′ termini of each predicted gene, respectively (Fig. 1D), indicating that these sequences are the transcription start/stop signals for CTRV. IGRs between N/P, P/M, and M/G genes showed variations in length and sequence (CAUC, CUC, and CUCUU, respectively). Interestingly, the junction between the G/L genes differed from the three former junctions in that the transcription start signal of the L gene preceded the transcription stop signal of the G gene, indicating a 36-nt overlap between the G/L genes (Fig. 1D).

N gene. The N protein encoded by the N gene is tightly associated with viral RNA to form a stable ribonucleocapsid complex that serves as a template for viral RNA synthesis. The CTRV N gene consisted of 1,566 nt, including 29 nt of the 5′ untranslated region (UTR) and 118 nt of the 3′ UTR, and encoded a protein of 472 amino acids (aa) with a calculated molecular mass of 52.8 kDa (Table 1). The protein sequence identities and similarities were 21.8% and 34.4% between CTRV and VSIV, respectively, and 17.6% and 33.5% between CTRV and RABV, respectively. Like many other rhabdoviruses, the CTRV N protein was predicted to contain protein kinase C (PKC), CK2, and TYR phosphorylation sites. A putative amidation site is found in the CTRV N protein, similar to that in the *Ngaingan virus* (NGAV; GenBank accession number NC_013955) although the C-terminal microbody targeting signal is not found (19). An RNA binding motif, GXS XKSPYSS, in the rhabdovirus N protein is highly conserved among the ephemeroviruses, vesiculoviruses, and lyssaviruses (33, 81). The sequence $_{284}$KLGNRSPYSA$_{293}$ in the CTRV N protein appeared to correspond to the motif.

P gene. The P protein encoded by the P gene is an essential cofactor of RNA polymerase for viral RNA synthesis. The P proteins of all NNS RNA viruses play similar roles in viral replication but have no apparent sequence similarity to each other even though the viruses belong to the same family. The CTRV putative P gene consisted of 838 nt, including 30 nt of the 5′ UTR and 13 nt of the 3′ UTR, and encoded a protein of 264 aa with a calculated molecular mass of 28.0 kDa (Table 1). The CTRV putative P protein was predicted to contain PKC, CK2, and TYR phosphorylation sites and eight N-myristoylation sites and lack cyclic AMP (cAMP)- and cGMP-dependent protein kinase phosphorylation sites.

M gene. The M protein encoded by the M gene is the smallest and most abundant protein in rhabdovirus virions, forming a layer between the nucleocapsid and the outer membrane. In addition to its structural roles, the M protein is assumed to play essential roles in virus assembly and budding and to cause CPE by inhibition of host gene expression and cytoskeletal disorganization (18, 27). The CTRV M gene consisted of 594 nt, including 28 nt of the 5′ UTR and 44 nt of the 3′ UTR, and encoded a protein of 173 aa with a calculated molecular mass of 18.8 kDa (Table 1). The CTRV M protein was predicted to contain several phosphorylation sites, an amidation site, and two N-myristoylation sites. The M proteins of VSV and RABV contain proline-rich motifs that constitute potential late domains (5) and is necessary for efficient virus budding (53, 84), but such a motif was not found in the CTRV M protein.

G gene. The G protein encoded by the G gene is a component of the projections on the virion surface and is involved in attachment to the cell receptor, virion assembly, and budding. The CTRV G gene consisted of 1,602 nt, including 22 nt of the 5′ UTR and 68 nt of the 3′ UTR, and encoded a protein of 503 aa with a calculated molecular mass of 55.0 kDa (Table 1). The protein sequence identities and similarities were 17.5% and 32.6% between CTRV and VSIV, respectively, and 17.3% and 32.1% between CTRV and RABV, respectively. The rhabdovirus G protein consisted of an N-terminal hydrophobic signal peptide, a major antigenic domain, and a C-terminal hydrophobic transmembrane domain followed by a short hydrophilic cytoplasmic domain (6, 80). Protein sequence analysis suggested that the CTRV G protein contains a signal peptide at residues 1 to 17, an ectodomain at residues 18 to 462, and a hydrophobic transmembrane domain at residues 463 to 484, followed by a hydrophilic cytoplasmic domain at residues 485 to 503. Although the G proteins of different rhabdovirus genera have low amino acid sequence similarity, two to six potential glycosylation sites and 12 cysteine residues are well conserved among them (6, 80). The CTRV G protein was predicted to contain five potential N-glycosylation sites and many myristoylation and phosphorylation sites. Twelve conserved cysteine residues were also found in the CTRV G protein.

L gene. The L protein encoded by the L gene is a component of viral RNPs and serves as the RdRp for viral RNA synthesis. The CTRV L gene consisted of 6,509 nt, including 37 nt of the 5′ UTR, 24 nt of the 3′ UTR, and 76 nt of the putative intron. Therefore, the L gene, excluding the putative 76-nt intron, encoded a protein of 2,123 aa with a calculated molecular mass

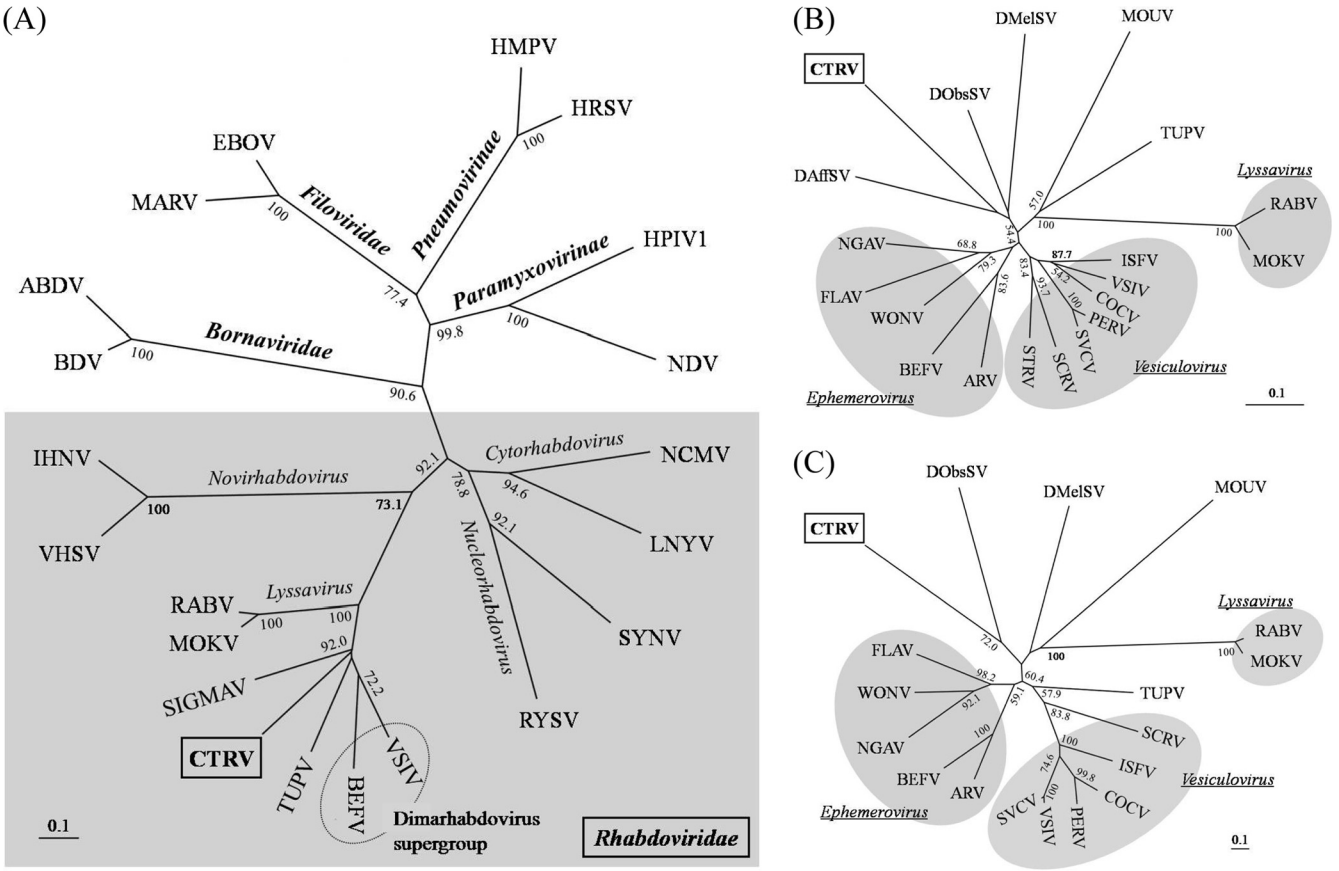


FIG. 2. Phylogenetic relationships between CTRV and other NNS RNA viruses. These dendrograms were constructed based on the protein sequence similarities of the highly conserved domain III of the L protein among selected members of the order *Mononegavirales* (A) and the family *Rhabdoviridae* (B) or based on the protein sequence similarities of partial N protein among selected members of the family *Rhabdoviridae* (C). In dendrogram A, family or subfamily names (bold face) and genus names are indicated on each appropriate branch. In dendrograms B and C, the genus names in the family *Rhabdoviridae* are underlined, and the viruses in the same genus are shaded. More than 50% of bootstrap values from 1,000 intermediate trees are indicated at the branch points. The bar represents the expected substitutions per site. Abbreviations of NNS RNA virus and sequence accession numbers (L and/or N protein) used in this study are follows: ABDV, avian Borna disease virus (ACS32310); ARV, Adelaide river virus (AF234534 and U10363); BDV, Borna disease virus (NC_001607); BEFV, bovine ephemeral fever virus (AF234533); COCV, Cocal virus (ACB47438 and ACB47434); DAffSV, *Drosophila affinis* sigma virus (ACU65445); DMelSV (SIGMAV), *Drosophila melanogaster* sigma virus (NC_013135); DObsSV, *Drosophila obscura* sigma virus (ACU65444 and ACU65439); EBOV, Zaire ebolavirus (NC_002549); FLAV, Flanders virus (AH012179); HMPV, human metapneumovirus (NC_004148); HPIV1, human parainfluenza virus 1 (NC_003461); HRSV, human respiratory syncytial virus (NC_001781); IHNV, infectious hematopoietic necrosis virus (NC_001652); ISFV, Isfahan virus (AJ810084); LNYV, lettuce necrotic yellows virus (NC_007642); MARV, Lake Victoria marburgvirus (NC_001608); MOKV, Mokola virus (NC_006429); MOUV, Moussa virus (FJ985748), NCMV, northern cereal mosaic virus (AB030277); NDV, Newcastle disease virus (NC_002617); NGAV, Ngaingan virus (NC_013955); PEFV, pike fry rhabdovirus (FJ872827); RABV, rabies virus (GU345748); RYSV, rice yellow stunt virus (AB011257); SCRV, *Siniperca chuatsi* rhabdovirus (NC_008514); SFRV, starry flounder rhabdovirus (AY450644); SVCV, spring viremia of carp virus (NC_002803); SYNV, sonchus yellow net virus, (NC_001615); TUPV, tupaia rhabdovirus (AY840978); VHSV, viral hemorrhagic septicemia virus (NC_000855); VSIV, vesicular stomatitis Indiana virus (NC_001560); WONV, Wongabel virus (NC_011639).

of 239.2 kDa (Table 1). The protein sequence identities and similarities were 35.8% and 54.3% between CTRV and VSIV, respectively, and 30.2% and 47.6% between CTRV and RABV, respectively. RdRps of all NNS RNA viruses share six highly conserved domains (designated block I to VI), which include characteristic sequence motifs and are essential for fundamental polymerase activities (52); the CTRV putative L protein also contains the predicted six conserved domains. For instance, the sequence $_{716}$LAQGDNQVL$_{724}$ could correspond to the motif C in block III, which is important for polymerase function (52, 58, 62). The sequence $_{1161}$GSMT-(69 aa)-HR$_{1235}$ could correspond to the conserved motif GSXT-(69–70aa)-HR in block V, which is essential for mRNA capping activity (38). The putative splice site in the CTRV L gene was located immediately below the predicted block V (52).

Phylogenetic study. First, the taxonomic position of CTRV among members of the order *Mononegavirales* was evaluated based on protein sequences of the conserved domain block III in the L protein (52). The resulting phylogenetic tree showed that CTRV certainly belongs to the family *Rhabdoviridae* and is closely related to the dimarhabdovirus supergroup, which includes the animal-infecting rhabdovirus of the genera *Vesiculovirus* and *Ephemerovirus* (4) (Fig. 2A). More detailed phylogenetic analysis among selected members of the family *Rhabdoviridae* indicated that the CTRV and *Drosophila* rhabdovirus group (sigma viruses [40]) form a distinct cluster from other

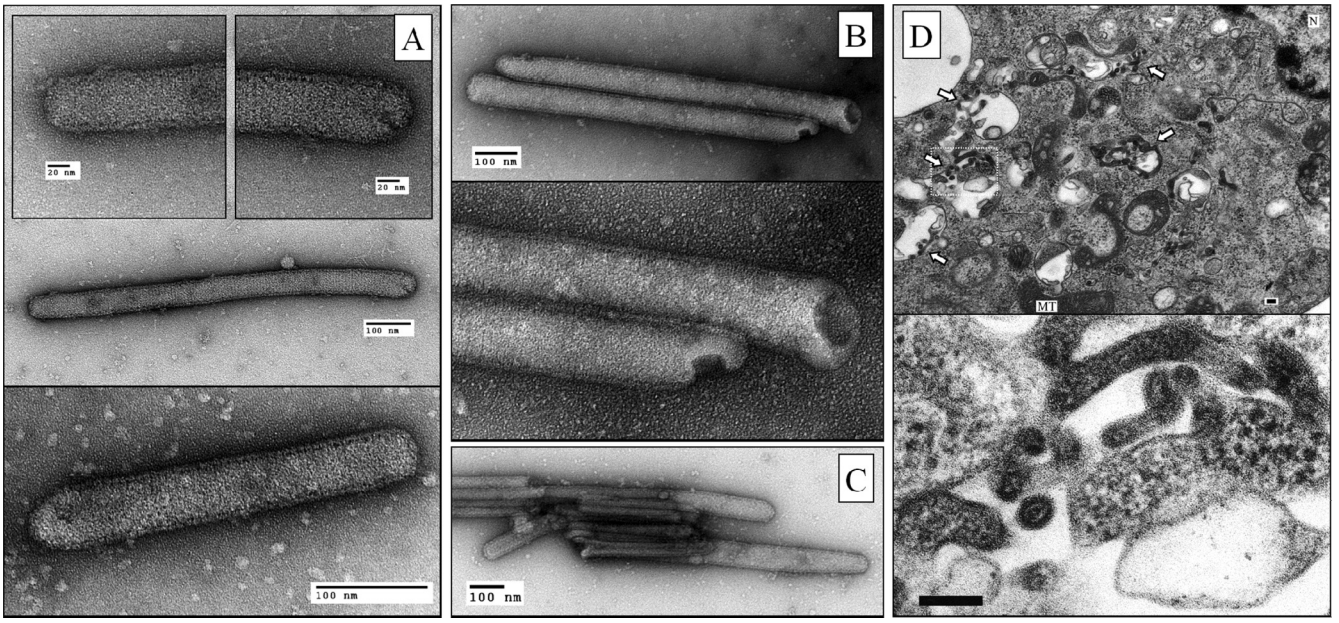


FIG. 3. Electron microscopic analyses of CTRV virions and thin-section of infected C6/36 cells. (A) The virions exhibited width uniformity and substantial variation in length, which are straight with two rounded tips and surrounded by a layer with projections entirely on the virion surface. The penetrating stain image showed the periodic fringe of the internal structure. (B) The truncated virion image shows the presence of a hollow cavity in the interior of the particle. (C) Some parts of the virion may be seen without the outer envelope, which seems to expose the naked inner structure. (D) Thin section of the C6/36 cells infected with CTRV. Abbreviations: N, nucleus; MT; mitochondrion. White arrows indicate the CTRV virions in the intracellular vacuoles, and the lower panel shows detail in dotted-line box in the upper panel. Scale bar, 100 nm.

members (Fig. 2B). Furthermore, the phylogenetic tree based on conserved N protein sequences led to a nearly similar result, except for tupaia rhabdovirus (TUPV) position (Fig. 2C). In this tree, CTRV formed a cluster with *Drosophila obscura* sigma virus (DObsSV) (40).

Electron microscopy analyses. We observed images of CTRV virions present in the culture supernatant from infected cells by negative-staining electron microscopy. Interestingly, CTRV virions were elongated, unlike those of other common rhabdoviruses (Fig. 3). The widths of the virions were almost the same at 59.4 ± 4.8 nm; however, the lengths of the virions varied significantly from 250 to 1,100 nm (Fig. 3A). The particles were straight with both tips rounded and surrounded by the outer envelope with the layer of projections that protrude 6 to 8 nm above the surface of the virion entirely. A penetrating stain image revealed a periodicity of the internal structure, and the internal striations had a pitch of 7.5 to 8.0 nm. Truncated forms of the virions were occasionally seen in the same preparation, and the interior of the truncated end appeared to form a hollow cavity (Fig. 3B). The outer envelope in some virions may be partially removed during sample preparation, and the width of the naked inner structure was about 40 nm in diameter (Fig. 3C). Thin-section electron microscopy images of CTRV-infected cells revealed that virions were present in the intracellular vacuoles (Fig. 3D), and some virions showed axial pores surrounded by two concentric layers. No such image was observed in mock-inoculated controls (not shown).

Northern hybridization. To examine the CTRV L gene transcript, Northern hybridization analysis was performed with RNA probes specific for the regions upstream (LPB1) or downstream (LPB2) of the putative intron in the L gene. The probes derived from regions LPB1 and LPB2 gave very similar hybridization patterns in the blots (Fig. 4). When sense RNA probes for detecting negative-strand (NS) viral RNA were used, a hybridizing band was detected at more than 9,000 nt in size, which could correspond to viral genomic RNA. On the other hand, when an antisense RNA probe was used for detecting positive-strand viral RNA, an intense hybridizing band was detected at approximately 6,500 nt in size, which almost corresponded to the expected size of viral full-length L mRNA. The result indicated that at least a large transcript derived from the L gene was present in the CTRV-infected C6/36 cells. Additionally, in the same blot, a weak band was also detected at the same size as that of viral genomic RNA, suggesting that the band was derived from viral antigenomic RNA which is present at a significantly lower level in the infected cell.

Analysis of the L gene transcript. As described above, sequence analysis strongly suggested that CTRV genomic RNA contains a putative 76-nt intron that is spliced out for generating mature L mRNA. In order to verify this hypothesis, strand-specific RT-PCR analyses with different primer combinations were performed on viral RNA extracted from the infected cells (Fig. 5A). When primer A or B for first-strand cDNA synthesis was used, only negative-strand viral RNA (genomic RNA) could serve as the template for RT-PCR, and the following PCR actually generated a single amplified product with the expected size (Fig. 5B, lanes 1 and 2). On the other hand, when primer C or D was used for first-strand cDNA synthesis, positive-strand viral RNAs (antigenomic RNA, pre-mRNA, and mature mRNA) could be used as templates for RT-PCR. In fact, when primer C was used, two amplified fragments were obtained (Fig. 5B, lane 3). The lower and

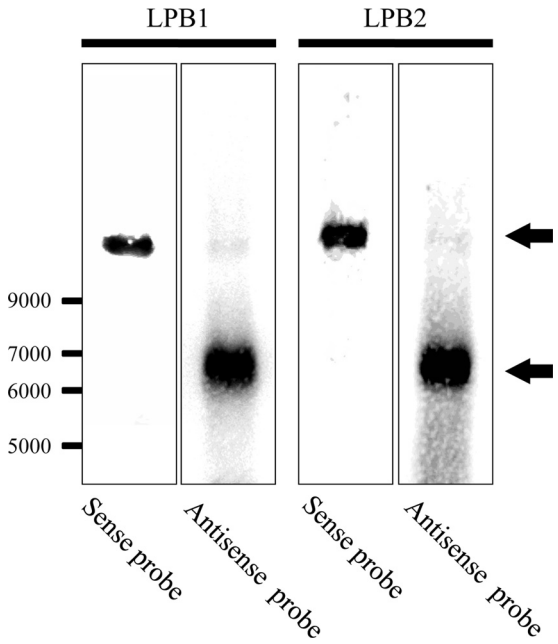


FIG. 4. Northern hybridization analysis of the CTRV L gene transcript in infected cells. Total RNA from CTRV-infected C6/36 cells was electrophoresed in a 0.8% denaturing formaldehyde agarose gel. Hybridization was performed using sense or antisense RNA probes specific for the regions upstream (LBP1; sequence position 4894 to 5752) or downstream (LBP2; sequence position 10172 to 11125) of the putative intron in the L gene. Black arrows indicate the hybridizing bands corresponding to the expected sizes of viral genomic or antigenomic RNA (upper) and viral full-length L mRNA (lower), respectively. Both a single-stranded RNA ladder (New England BioLabs, Ipswich, MA) and an RNA 6000 Ladder (Ambion Inc., Austin, TX) were used as molecular weight markers.

middle bands corresponded to the expected sizes of PCR products derived from spliced (mature mRNA) and unspliced RNAs (antigenomic RNA and pre-mRNA), respectively. The upper smear band (~480 nt) could be composed of the heteroduplex molecules formed by heterologous annealing of the two above-mentioned products. Furthermore, when primer D was used, cDNA synthesis could occur from unspliced RNAs (antigenomic RNA and pre-mRNA) but not from spliced RNA (mature mRNA), and the following PCR resulted in generation of a single amplified product (Fig. 5B, lane 4). All cDNA amplicons obtained in this study were sequenced individually and compared to each other (Fig. 5C). In conclusion, we confirmed that the smallest PCR product in lane 3 contains a deletion of 76 nt as expected, in accordance with in the sequence shown in Fig. 1B. These results provided experimental evidence that the RNA transcript spliced at the predicted donor and acceptor sites in the CTRV-infected cell, and the CTRV L gene requires RNA splicing for L mRNA maturation.

***In situ* hybridization.** If CTRV really requires RNA splicing for viral mRNA maturation, viral genomic RNA could be localized in the nucleus to utilize the host's nuclear splicing machinery. In order to verify this hypothesis, subcellular localizations of viral RNAs in the CTRV-infected C6/36 cells were examined by RNA-RNA ISH using DIG-labeled RNA probes. When sense probes were used for the M gene, intense hybridization signals were detected in the nucleus of infected cells; on the other hand, when antisense probes were used for the M gene, the hybridization signals were detected in the nucleus and diffusely present in the cytoplasm of infected cells (Fig. 6). Similar results were also obtained by using probes derived from both LPB1 and LPB2 regions of the L gene (Fig. 6). These results could suggest that viral genomic RNAs and also viral antigenomic RNAs are primarily localized in the cell nucleus and that viral gene transcripts (mRNA) are expressed in the cytoplasm. When sense probes were used for the 76-nt intron, hybridization signals were detected in the cell nucleus as in the case of other sense probes; however, when antisense probes were used for the 76-nt intron, no detectable signals were obtained (Fig. 6). These results might indicate that the unspliced pre-mRNAs are absent in the cytoplasm and that the positive-strand viral RNAs in the nucleus are undetectable by using the short antisense probe. In conclusion, viral genomic RNA is transported to the host cell nucleus where it replicates, which allows it to utilize the RNA splicing machinery for viral mRNA maturation.

DISCUSSION

In this study, we characterized a new rhabdovirus from *Culex* mosquitoes that replicates in the nucleus of the host cell and requires RNA splicing for viral mRNA maturation. Apart from many eukaryotes, only a fraction of negatively stranded RNA viruses require RNA splicing, which contributes both proteomic diversity and regulation of gene expression (9). In orthomyxoviruses such as influenza and Thogoto viruses, several genome segments often code for additional viral proteins, which are translated from spliced or bicistronic mRNAs (29, 50). In Borna disease virus (BDV), the prototype of the family *Bornaviridae*, the viral transcripts are spliced alternatively by a combination of three acceptor and two donor sites to generate various protein isoforms (66, 67). However, to date, such RNA splicing events have not been reported in any rhabdoviruses, even in nucleorhabdoviruses that replicate in the cell nucleus. As shown in this study, RNA splicing in CTRV could be absolutely required for expression of the full-length L protein because the coding region for the L protein is interrupted by an in-frame stop codon that separates the coding region into two different reading frames (Fig. 1A and B). This differs fundamentally from the in-frame alternative splicing observed in orthomyxoviruses and bornaviruses. Strand-specific RT-PCR analyses identified the intron-exon boundaries and the spliced intron sequence in the CTRV L gene (Fig. 5). The 76-nt intron has the conserved motif of eukaryotic spliceosomal introns—GU-AG boundaries, a predicted branch point, and a polypyrimidine tract (Fig. 1B)—suggesting that the conserved *cis*-element is probably fundamental for splice site recognition for L mRNA maturation. Because no rhabdovirus has been reported to have an intron in its genome, the origin and significance of the RNA splicing events in CTRV are unclear. It is speculated that the presence of the intron and the efficiency of RNA splicing could significantly influence the expression level and pattern of the L protein. A reverse genetic approach may help to clarify the regulatory mechanism and functional importance of RNA splicing in CTRV.

In orthomyxoviruses and bornaviruses, viral RNA is transported into the nuclei of host cells, where viral replication and

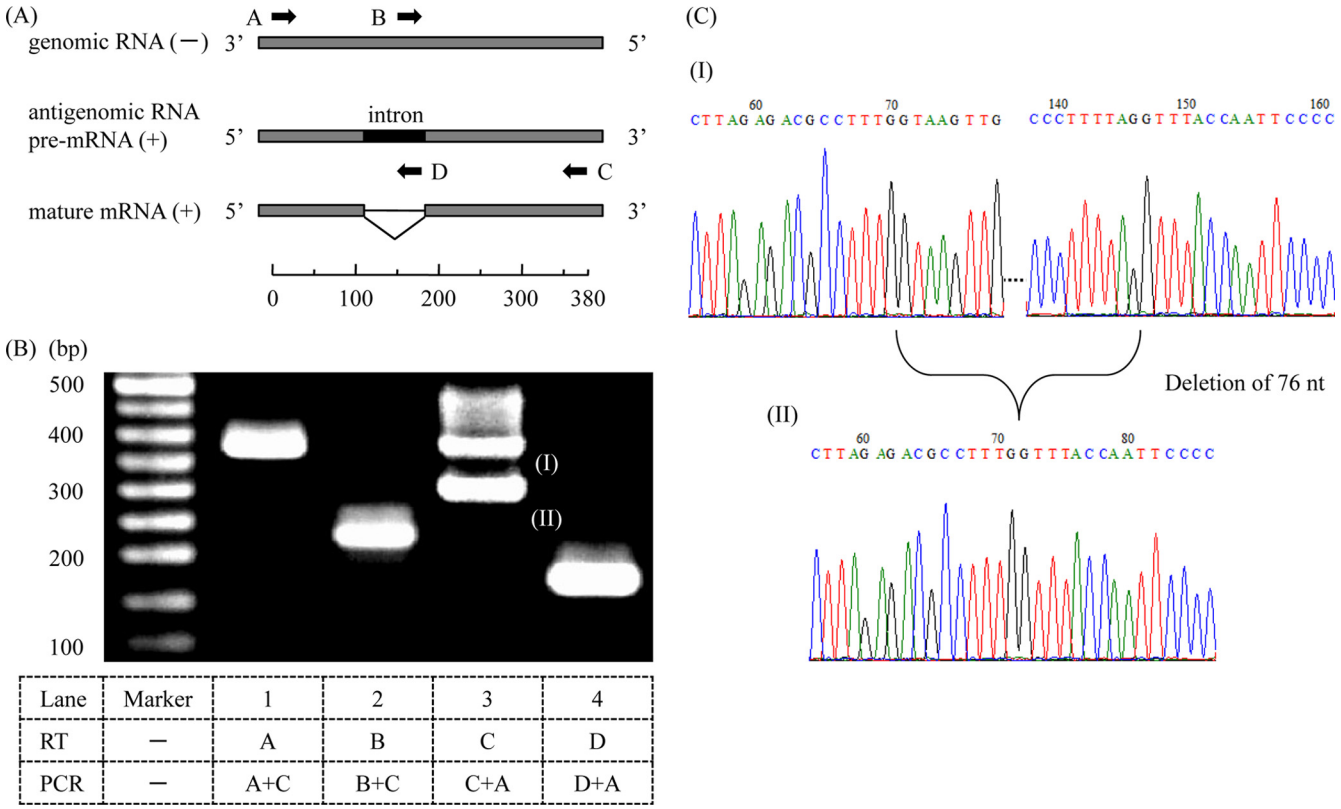


FIG. 5. Strand-specific RT-PCR analysis for the CTRV L gene transcript. (A) Schematic map of partial regions of CTRV genomic RNA (negative strand; −), antigenomic RNA, and mRNA (positive strand; +) around the intron in the L gene, as shown with a scale bar. The four black arrows indicate the positions and orientations of CTRV gene-specific primers used for strand-specific RT-PCR. For RT, primers A and B anneal specifically to negative-strand RNA, whereas primers C and D anneal specifically to positive-strand RNA, respectively. (B) Agarose gel electrophoresis of the strand-specific RT-PCR products (lanes 1 to 4). The lower table indicates the primers used for each RT and the following PCR. Note that the smallest fragment (II) on lane 3 is an amplified product from spliced mature mRNA with the expected size of 297 bp. (C) Sequence chromatograms of the two PCR products (fragments I and II) in panel B, which were directly sequenced with primer A. The upper chromatogram was obtained from fragment I, and the lower one was obtained from fragment II. Note that the sequence of fragment II contains a deletion of 76 nt, as expected in Fig. 1B.

transcription occur. The nucleocytoplasmic trafficking of viral RNA is mediated by the binding of the cargo molecules that possess nuclear localization signals (NLS) and nuclear export signals (NES), which are involved in RNP transport between the nucleus and cytoplasm (3, 10). In this study, we demonstrated that CTRV genomic RNA is predominantly localized in the nucleus of infected cells (Fig. 6), suggesting that CTRV RNPs may possess such signals. Protein subcellular localization predictions using the WoLF PSORT program (22) strongly suggested that the CTRV P protein may be located in the cell nucleus. The CTRV P protein actually contains three basic residue clusters that may serve as NLS as predicted by PSORT II program (47): $_{4}$PRKTRNK$_{10}$, $_{107}$PAQRRR$_{112}$, and $_{188}$PKR AKAA$_{194}$. In the BDV P protein, bipartite proline-rich NLS have been identified in the N- and C-terminal regions (59, 60). The CTRV as well as BDV P protein contained a remarkable proline-rich cluster with four basic residues in the N-terminal region ($_{3}$PRKTRNKTPAPPAPPAPGP$_{23}$), which is expected to serve as NLS. The NLS in the N protein have been identified in various negative-strand (NS) RNA viruses that replicate in the nucleus (3, 16, 17, 30, 42, 69). Although morbilliviruses in the family *Paramyxoviridae* replicate in the cytoplasm, their N proteins also possess NLS (55). The CTRV N protein contained one basic residue cluster, $_{343}$RRRHR$_{347}$, which is a candidate NLS, although the NLS score by PSORT II prediction is low. On the other hand, the leucine-rich cluster in both the N and/or P proteins has been identified as NES in some NS RNA viruses (3, 31, 49, 51, 55, 57, 77). In addition to the leucine-rich cluster, methionine-rich domains also serve as NES in the BDV P proteins (72). Analysis using an NES prediction program (NetNES server [36]) suggested that the CTRV L protein contains an NES that consists of a canonical leucine-rich cluster, $_{1811}$LLDLQQNIISE$_{1821}$. The sequences $_{174}$LEPSALAI$_{181}$ in the N protein and $_{71}$MFDGLGL$_{77}$ in the P protein were also predicted as candidate NES, but their NES scores were not so high.

Gene expression of NNS RNA virus is regulated by *cis*-acting transcription signals located at the beginning and end of the genes, and each gene is separated by an IGR, which also plays an important role in controlling gene expression (83). In the present study, we determined the transcription signals of each CTRV gene and found that IGRs show variation in length and sequence (Fig. 1D). In particular, the transcription start signal of the L gene precedes the transcription stop signal of

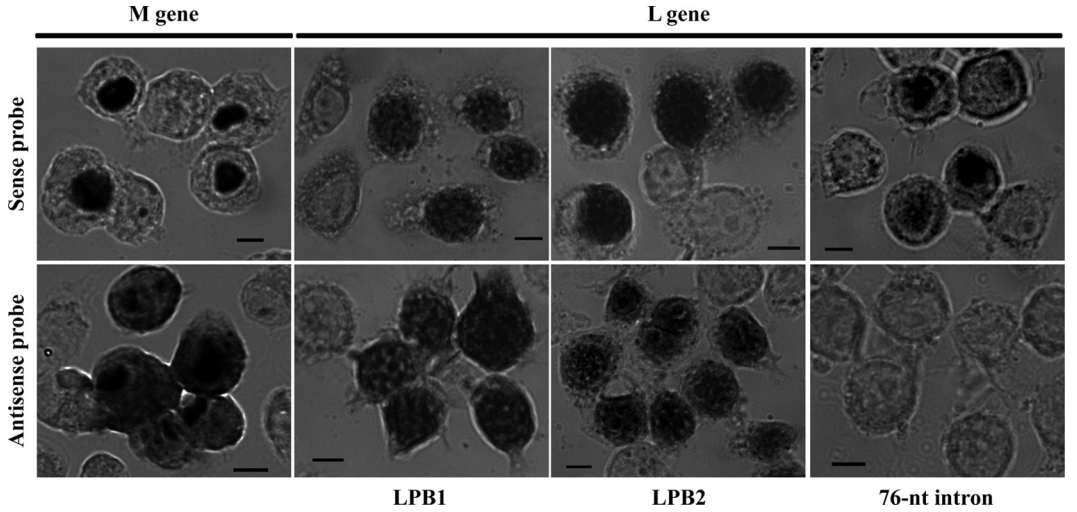


FIG. 6. Subcellular localization of CTRV RNA in the C6/36 cells persistently infected with CTRV. DIG-labeled sense or antisense RNA probe for CTRV M and L genes specifically recognized negative-sense or positive-sense viral RNA, respectively. Several uninfected cells were also seen in the visual field. Scale bar, 5 μm.

the G gene, resulting in a 36-nt overlap at the G/L gene junction (Fig. 1D). Such a gene overlap in NNS RNA viruses has been found in the genera *Pneumovirus* (7, 85), *Ebolavirus* (48), *Bornavirus* (56), and also the family *Rhabdoviridae* (45, 64). Among these, a gene overlap in human respiratory syncytial virus (HRSV; a pneumovirus) appears to be most similar to that of CTRV in position and length; namely, a start signal of the HRSV L gene is located in the terminal tail of the upstream M2-1 gene, resulting in a 68-nt gene overlap. Analyses of transcript profiles suggested that HRSV yields many short, truncated, premature L mRNAs attributed to gene overlap (7), but the synthesis of full-length L mRNA is probably supported by a compensatory mechanism for recycling polymerases (11). In general, the L gene expression of NNS RNA viruses is fundamentally expressed at the lowest level because the transcription level gradually decreases in the 5′ to 3′ direction of antigenomic RNA, which is due to sequential transcription from a single promoter in the 3′ leader region and attenuation of transcription across gene junctions. Furthermore, L gene expression in many NNS RNA viruses is downregulated by other *cis*-acting signals (13, 63). Transcription of CTRV L gene may be highly regulated by both RNA splicing and the G/L gene overlap.

Electron microscopic analysis revealed that the CTRV virion has an unusual elongated shape (Fig. 3) although other morphological features (e.g., diameter, shape of the tips, and internal and external structures) are similar to those of the family *Rhabdoviridae*. The virion morphologies of rhabdoviruses often showed pleomorphism, especially in length. However, to our knowledge, such extremely elongated virions have never been reported in members of the family *Rhabdoviridae*. The virion morphology of CTRV had something in common with that of Taastrup virus (TV) (41); however, TV was closely related but did not belong to the family *Rhabdoviridae* (2). We initially attempted to purify the CTRV virions with sucrose density gradient centrifugation; however, no clear virus band was obtained after centrifugation (data not shown). The difficulty in virion purification may be attributed to its unique shape and pleomorphisms in length. Thin-section electron microscopic analysis of CTRV-infected C6/36 cells revealed that the CTRV particles were present in intracellular vacuoles (Fig. 3D); hence, CTRV maturation may have occurred by assembly and budding from the plasma membrane, as in BDV (32), but not at the nuclear membrane, as in nucleorhabdoviruses (21). The recently detailed budding step of Zaire ebolavirus (EBOV), which appears as long flexuous filaments, has been shown by electron tomographic imaging (82). Such an imaging technique or immunoelectron microscopic studies are needed to visualize assembly and budding of the new rhabdovirus.

Based on phylogenetic study, CTRV is indicated to belong to the family *Rhabdoviridae* and is closely related to the dimarhabdovirus supergroup; however, it remains unassigned to a genus (Fig. 2). Although CTRVs share a characteristic with the plant-infecting nucleorhabdovirus at the site of replication, they are distantly related to each other (Fig. 2A). In other phylogenetic trees for selected members of the family *Rhabdoviridae*, CTRV forms a cluster with *Drosophila* rhabdoviruses (Fig. 2B and C), which are commonly termed sigma viruses (40). Sigma viruses are widespread in the natural populations of *Drosophila* flies, with a proportional morbidity and maintained by vertical transmission (8). Like many other rhabdoviruses, sigma viruses can be adapted to grow in mosquito cell line C6/36 (73). This indicated that CTRV and sigma viruses may not only have diverged from a common ancestral virus but may also cause similar physiological effects in hosts. For instance, it is well known that sigma virus infection leads to enhanced anesthetic sensitivity of host flies to carbon dioxide (CO_2) gas (8). This irreversible CO_2 sensitivity has been investigated in both flies and mosquitoes and is induced by infection of not only sigma virus but also some arboviruses (54, 71, 74). Interestingly, the inherited etiologic agent that induces CO_2 sensitivity had been previously reported in *Culex quinquefasciatus* (61), which is one of the most common and widespread *Culex* mosquitoes, although the virus seems not to be a rhabdovirus because of its morphological difference (75). Another aspect of interest is that sigma virus infection appears to re-

duce the overwintering fitness of the host flies (14). In Japan the breeding season of *C. tritaeniorhynchus* is generally from late March to late September or early October; however, locations of the overwintering sites of the mosquitoes still remain unclear (78, 79). Recently, a large number of *C. tritaeniorhynchus* in the reproductive diapause stage have migrated somewhere from late September to November, which represents the prediapause seasonal migration from breeding sites to overwintering sites (70). By analogy to sigma virus infection, CTRV infection may have an effect on the overwintering fitness of natural mosquito populations. Detailed field and laboratory studies will be valuable for understanding the transmission mode and life cycle of CTRV.

Studies of the "hijacking" of the host cell nucleus by RNA viruses will provide valuable insights and contribute to the understanding of cell biology and virus-host interaction (20). Recently, endogenization of BDV-derived elements in mammalian genomes has been reported (23). Despite marked differences in the host (invertebrate and vertebrate), CTRV and BDV share common features such as the establishment of persistent infection, replication in the cell nucleus, and use of the host's RNA splicing machinery. Further studies of the relationship between CTRV and the host mosquito could provide new information about NNS RNA virus-host interaction and evolution.

ACKNOWLEDGMENTS

We thank Shogo Tanaka of Kyushu Research Station, National Institute of Animal Health, NARO, for technical advice.

This work was supported in part by a Grant-in-aid of Japanese Ministry of Health, Labor and Welfare (H18-Shinko-Ippan-009 and H21-Shinko-Ippan-005). It was also supported in part by Grant-in-aids for Scientific Research (number 21406012) and a bilateral program from the Japan Society for the Promotion of Science.

REFERENCES

1. **Ammar, el-D., C. W. Tsai, A. E. Whitfield, M. G. Redinbaugh, and S. A. Hogenhout.** 2009. Cellular and molecular aspects of rhabdovirus interactions with insect and plant hosts. Annu. Rev. Entomol. **54:**447–468.
2. **Bock, J. O., T. Lundsgaard, P. A. Pedersen, and L. S. Christensen.** 2004. Identification and partial characterization of Taastrup virus: a newly identified member species of the *Mononegavirales*. Virology **319:**49–59.
3. **Boulo, S., H. Akarsu, R. W. Ruigrok, and F. Baudin.** 2007. Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. Virus Res. **124:**12–21.
4. **Bourhy, H., J. A. Cowley, F. Larrous, E. C. Holmes, and P. J. Walker.** 2005. Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. J. Gen. Virol. **86:**2849–2858.
5. **Chen, B. J., and R. A. Lamb.** 2008. Mechanisms for enveloped virus budding: can some viruses do without an ESCRT? Virology **372:**221–232.
6. **Coll, J. M.** 1995. The glycoprotein G of rhabdoviruses. Arch. Virol. **140:**827–851.
7. **Collins, P. L., R. A. Olmsted, M. K. Spriggs, P. R. Johnson, and A. J. Buckler-White.** 1987. Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. Proc. Natl. Acad. Sci. U. S. A. **84:**5134–5138.
8. **Contamine, D., and S. Gaumer.** 2008. Sigma rhabdoviruses, p. 576–581. *In* A. Granoff and R. G. Webster (ed.), Encyclopedia of virology, 3rd ed., vol. 5. Academic Press, San Diego, CA.
9. **Cros, J. F., and P. Palese.** 2003. Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. Virus Res. **95:**3–12.
10. **de la Torre, J. C.** 2006. Reverse-genetic approaches to the study of Borna disease virus. Nat. Rev. Microbiol. **4:**777–783.
11. **Fearns, R., and P. L. Collins.** 1999. Model for polymerase access to the overlapped L gene of respiratory syncytial virus. J. Virol. **73:**388–397.
12. **Felsenstein, J.** 1985. Confidence-limits on phylogenies—an approach using the bootstrap. Evolution **39:**783–791.
13. **Finke, S., J. H. Cox, and K. K. Conzelmann.** 2000. Differential transcription attenuation of rabies virus genes by intergenic regions: generation of recombinant viruses overexpressing the polymerase gene. J. Virol. **74:**7261–7269.
14. **Fleuriet, A.** 1981. Effect of overwintering on the frequency of flies infected by the rhabdovirus sigma in experimental populations of *Drosophila melanogaster*. Arch. Virol. **69:**253–260.
15. **Gao, K., A. Masuda, T. Matsuura, and K. Ohno.** 2008. Human branch point consensus sequence is yUnAy. Nucleic Acids Res. **36:**2257–2267.
16. **Ghosh, D., R. E. Brooks, R. Wang, J. Lesnaw, and M. M. Goodin.** 2008. Cloning and subcellular localization of the phosphoprotein and nucleocapsid proteins of Potato yellow dwarf virus, type species of the genus *Nucleorhabdovirus*. Virus Res. **135:**26–35.
17. **Goodin, M. M., et al.** 2001. Interactions and nuclear import of the N and P proteins of sonchus yellow net virus, a plant nucleorhabdovirus. J. Virol. **75:**9393–9406.
18. **Graham, S. C., et al.** 2008. Rhabdovirus matrix protein structures reveal a novel mode of self-association. PLoS Pathog. **4:**e1000251.
19. **Gubala, A., et al.** 2010. Ngaingan virus, a macropod-associated rhabdovirus, contains a second glycoprotein gene and seven novel open reading frames. Virology **399:**98–108.
20. **Hiscox, J. A.** 2007. RNA viruses: hijacking the dynamic nucleolus. Nat. Rev. Microbiol. **5:**119–127.
21. **Hogenhout, S. A., M. G. Redinbaugh, and el-D. Ammar.** 2003. Plant and animal rhabdovirus host range: a bug's view. Trends Microbiol. **11:**264–271.
22. **Horton, P., et al.** 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. **35:**W585–W587.
23. **Horie, M., et al.** 2010. Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature **463:**84–87.
24. **Hoshino, K., et al.** 2007. Genetic characterization of a new insect flavivirus isolated from *Culex pipiens* mosquito in Japan. Virology **359:**405–414.
25. **Hoshino, K., H. Isawa, Y. Tsuda, K. Sawabe, and M. Kobayashi.** 2009. Isolation and characterization of a new insect flavivirus from *Aedes albopictus* and *Aedes flavopictus* mosquitoes in Japan. Virology **391:**119–129.
26. **Jackson, A. O., R. G. Dietzgen, M. M. Goodin, J. N. Bragg, and M. Deng.** 2005. Biology of plant rhabdoviruses. Annu. Rev. Phytopathol. **43:**623–660.
27. **Jayakar, H. R., E. Jeetendra, and M. A. Whitt.** 2004. Rhabdovirus assembly and budding. Virus Res. **106:**117–132.
28. **Jones, D. T., W. R. Taylor, and J. M. Thornton.** 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8:**275–282.
29. **Kawaoka, Y., et al.** 2005. Orthomyxoviridae. p. 681–693. *In* C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (ed.), Virus taxonomy. Eighth report of the International Committee on Taxonomy of Viruses. Elsevier, London, United Kingdom.
30. **Kobayashi, T., et al.** 1998. Nuclear targeting activity associated with the amino terminal region of the Borna disease virus nucleoprotein. Virology **243:**188–197.
31. **Kobayashi, T., et al.** 2001. Borna disease virus nucleoprotein requires both nuclear localization and export activities for viral nucleocytoplasmic shuttling. J. Virol. **75:**3404–3412.
32. **Kohno, T., et al.** 1999. Fine structure and morphogenesis of Borna disease virus. J. Virol. **73:**760–766.
33. **Kouznetzoff, A., M. Buckle, and N. Tordo.** 1998. Identification of a region of the rabies virus N protein involved in direct binding to the viral RNA. J. Gen. Virol. **79:**1005–1013.
34. **Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp.** 1998. Phylogeny of the genus *Flavivirus*. J. Virol. **72:**73–83.
35. **Kuzmin, I. V., G. J. Hughes, and C. E. Rupprecht.** 2006. Phylogenetic relationships of seven previously unclassified viruses within the family *Rhabdoviridae* using partial nucleoprotein gene sequences. J. Gen. Virol. **87:**2323–2331.
36. **la Cour, T., et al.** 2004. Analysis and prediction of leucine-rich nuclear export signals. Protein Eng. Des. Sel. **17:**527–536.
37. **Larkin, M. A., et al.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948.
38. **Li, J., A. Rahmeh, M. Morelli, and S. P. Whelan.** 2008. A conserved motif in region V of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping. J. Virol. **82:**775–784.
39. **Li, Z., M. Yu, H. Zhang, H. Y. Wang, and L. F. Wang.** 2005. Improved rapid amplification of cDNA ends (RACE) for mapping both the 5′ and 3′ terminal sequences of paramyxovirus genomes. J. Virol. Methods **130:**154–156.
40. **Longdon, B., D. J. Obbard, and F. M. Jiggins.** 2010. Sigma viruses from three species of *Drosophila* form a major new clade in the rhabdovirus phylogeny. Proc. Biol. Sci. **277:**35–44.
41. **Lundsgaard, T.** 1997. Filovirus-like particles detected in the leafhopper *Psammotettix alienus*. Virus Res. **48:**35–40.
42. **Martins, C. R. F., et al.** 1998. Sonchus yellow net rhabdovirus nuclear viroplasms contain polymerase-associated proteins. J. Virol. **72:**5669–5679.
43. **Mavale, M. S., et al.** 2005. Vertical and venereal transmission of Chandipura virus (Rhabdoviridae) by *Aedes aegypti* (Diptera: Culicidae). J. Med. Entomol. **42:**909–911.
44. **Mavale, M. S., et al.** 2007. Experimental transmission of Chandipura virus by *Phlebotomus argentipes* (Diptera: Psychodidae). Am. J. Trop. Med. Hyg. **76:**307–309.
45. **McWilliam, S. M., K. Kongsuwan, J. A. Cowley, K. A. Byrne, and P. J.**

Walker. 1997. Genome organization and transcription strategy in the complex G_{NS}-L intergenic region of bovine ephemeral fever rhabdovirus. J. Gen. Virol. **78:**1309–1317.
46. **Mead, D. G., et al.** 2009. Experimental transmission of vesicular stomatitis New Jersey virus from *Simulium vittatum* to cattle: clinical outcome is influenced by site of insect feeding. J. Med. Entomol. **46:**866–872.
47. **Nakai, K., and P. Horton.** 1999. PSORT: a program for detecting sorting signals in proteins and determining their subcellular localization. Trends Biochem. Sci. **24:**34–35.
48. **Neumann, G., S. Watanabe, and Y. Kawaoka.** 2009. Characterization of Ebola virus regulatory genomic regions. Virus Res. **144:**1–7.
49. **Nishie, T., K. Nagata, and K. Takeuchi.** 2007. The C protein of wild-type measles virus has the ability to shuttle between the nucleus and the cytoplasm. Microbes Infect. **9:**344–354.
50. **Palese, P., and M. Shaw.** 2007. *Orthomyxoviridae*: the viruses and their replication, p. 1647–1690. *In* D. M. Knipe, et al. (ed.), Fields virology, 5th ed. Lippincott, Williams & Wilkins, Philadelphia, PA.
51. **Pasdeloup, D., et al.** 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.
52. **Poch, O., B. M. Blumberg, L. Bougueleret, and N. Tordo.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J. Gen. Virol. **71:**1153–1162.
53. **Raux, H., et al.** 2010. The matrix protein of vesicular stomatitis virus binds dynamin for efficient viral assembly. J. Virol. **84:**12609–12618.
54. **Rozen, L.** 1980. Carbon dioxide sensitivity in mosquitoes infected with sigma, vesicular stomatitis, and other rhabdoviruses. Science **207:**989–991.
55. **Sato, H., M. Masuda, R. Miura, M. Yoneda, and C. Kai.** 2006. Morbillivirus nucleoprotein possesses a novel nuclear localization signal and a CRM1-independent nuclear export signal. Virology **352:**121–130.
56. **Schneemann, A., P. A. Schneider, S. Kim, and W. I. Lipkin.** 1994. Identification of signal sequences that control transcription of Borna disease virus, a nonsegmented, negative-strand RNA virus. J. Virol. **68:**6514–6522.
57. **Schneider, J., and T. Wolff.** 2009. Nuclear functions of the influenza A and B viruses NS1 proteins: do they play a role in viral mRNA export? Vaccine **27:**6312–6316.
58. **Schnell, M. J., and K. K. Conzelmann.** 1995. Polymerase activity of *in vitro* mutated rabies virus L protein. Virology **214:**522–530.
59. **Schwemmle, M., C. Jehle, T. Shoemaker, and W. I. Lipkin.** 1999. Characterization of the major nuclear localization signal of the Borna disease virus phosphoprotein. J. Gen. Virol. **80:**97–100.
60. **Shoya, Y., et al.** 1998. Two proline-rich nuclear localization signals in the amino-and carboxyl-terminal regions of the Borna disease virus phosphoprotein. J. Virol. **72:**9755–9762.
61. **Shroyer, D. A., and L. Rosen.** 1983. Extrachromosomal inheritance of carbon dioxide sensitivity in the mosquito *Culex quinquefasciatus*. Genetics **104:**649–659.
62. **Sleat, D. E., and A. K. Banerjee.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. J. Virol. **67:**1334–1339.
63. **Stillman, E. A., and M. A. Whitt.** 1998. The length and sequence composition of vesicular stomatitis virus intergenic regions affect mRNA levels and the site of transcript initiation. J. Virol. **72:**5565–5572.
64. **Teninges, D., F. Bras, and S. Dezelee.** 1993. Genome organization of the sigma rhabdovirus: six genes and a gene overlap. Virology **193:**1018–1023.
65. **Tesh, R. B., B. N. Chaniotis, and K. M. Johnson.** 1972. Vesicular stomatitis virus (Indiana serotype): transovarial transmission by phlebotomine sandlies. Science **175:**1477–1479.
66. **Tomonaga, K., et al.** 2000. Identification of alternative splicing and negative splicing activity of a nonsegmented negative-strand RNA virus, Borna disease virus. Proc. Natl. Acad. Sci. U. S. A. **97:**12788–12793.
67. **Tomonaga, K., T. Kobayashi, and K. Ikta.** 2002. Molecular and cellular biology of Borna disease virus infection. Microbes Infect. **4:**491–500.
68. **Tordo, N., et al.** 2005. *Rhabdoviridae*, p. 623–644. *In* C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (ed.), Virus taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier, London, United Kingdom.
69. **Tsai, C.-W., et al.** 2005. Complete genome sequence and in planta subcellular localization of maize fine streak virus proteins. J. Virol. **79:**5304–5314.
70. **Tsuda, Y., and K. S. Kim.** 2008. Sudden autumnal appearance of adult *Culex tritaeniorhynchus* (Diptera: Culicidae) at a park in urban Tokyo: first field evidence for prediapause migration. J. Med. Entomol. **45:**610–616.
71. **Turell, M. J., and J. L. Hardy.** 1980. Carbon dioxide sensitivity of mosquitoes infected with California encephalitis virus. Science **209:**1029–1030.
72. **Yanai, H., et al.** 2006. A methionine-rich domain mediates CRM1-dependent nuclear export activity of Borna disease virus phosphoprotein. J. Virol. **80:**1121–1129.
73. **Vazeille, M. C., and L. Rosen.** 1987. Cultivation in mosquito cells of sigma virus, the aetiologic agent of hereditary CO_2 sensitivity in *Drosophila melanogaster.* Ann. Inst. Pasteur. Virol. **138:**451–460.
74. **Vazeille, M. C., L. Rosen, and J. C. Guillon.** 1988. An orbivirus of mosquitoes which induces CO_2 sensitivity in mosquitoes and is lethal for rabbits. J. Virol. **62:**3484–3487.
75. **Vazeille-Falcoz, M., H. Ohayon, P. Gounon, and L. Rosen.** 1992. Unusual morphology of a virus which produces carbon dioxide sensitivity in mosquitoes. Virus Res. **24:**235–247.
76. **Veerassamy, S., A. Smith, and E. R. Tillier.** 2003. A transition probability model for amino acid substitutions from Blocks. J. Comput. Biol. **10:**997–1010.
77. **Vidy, A., J. E-l, Bougrini, M. K. Chelbi-Alix, and D. Blondel.** 2007. The nucleocytoplasmic rabies virus P protein counteracts interferon signaling by inhibiting both nuclear accumulation and DNA binding of STAT1. J. Virol. **81:**4255–4263.
78. **Wada, Y., et al.** 1973. Ecology of vector mosquitoes of Japanese encephalitis, especially of *Culex tritaeniorhynchus summorosus*. 5. Overwintering of *Culex tritaeniorhynchus summorosus* and *Anopheles sinensis.* Trop. Med. **15:**11–22.
79. **Wada, Y., et al.** 1975. Ecology of Japanese encephalitis in Japan. II. The population of vector mosquitoes and the epidemic of Japanese encephalitis. Trop. Med. **17:**111–127.
80. **Walker, P. J., and K. Kongsuwan.** 1999. Deduced structural model for animal rhabdovirus glycoproteins. J. Gen. Virol. **80:**1211–12200.
81. **Walker, P. J., Y. Wang, J. A. Cowley, S. M. McWilliam, and C. J. N. Prehaud.** 1994. Structural and antigenic analysis of the nucleoprotein of bovine ephemeral fever rhabdovirus. J. Gen. Virol. **75:**1889–1899.
82. **Welsch, S., et al.** 2010. Electron tomography reveals the steps in filovirus budding. PLoS Pathog. **6:**e1000875.
83. **Whelan, S. P., J. N. Barr, and G. W. Wertz.** 2004. Transcription and replication of nonsegmented negative-strand RNA viruses. Curr. Top. Microbiol. Immunol. **283:**61–119.
84. **Wirblich, C., et al.** 2008. PPEY motif within the rabies virus (RV) matrix protein is essential for efficient virion release and RV pathogenicity. J. Virol. **82:**9730–9738.
85. **Zamora, M., and S. K. Samal.** 1992. Gene junction sequences of bovine respiratory syncytial virus. Virus Res. **24:**115–121.