



Published in final edited form as:

*Virus Res.* 2011 January ; 155(1): 112–122. doi:10.1016/j.virusres.2010.09.007.

## Characterization of Durham virus, a novel rhabdovirus that encodes both a C and SH protein

A. B. Allison<sup>a,f</sup>, G. Palacios<sup>b</sup>, A. Travassos da Rosa<sup>c</sup>, V. L. Popov<sup>c</sup>, L. Lu<sup>c</sup>, S. Y. Xiao<sup>c,d</sup>, K. DeToy<sup>b</sup>, T. Briese<sup>b</sup>, W. Ian Lipkin<sup>b</sup>, M. K. Keel<sup>a</sup>, D. E. Stallknecht<sup>a</sup>, G. R. Bishop<sup>e</sup>, and R. B. Tesh<sup>c</sup>

<sup>a</sup> Southeastern Cooperative Wildlife Disease Study, College of Veterinary Medicine, University of Georgia, Athens, GA 30602

<sup>b</sup> Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY 10032

<sup>c</sup> Department of Pathology, Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch, Galveston, TX 77555

<sup>d</sup> Department of Pathology, University of Chicago Medical Center, Chicago, IL 60637

<sup>e</sup> Mississippi River Field Institute, National Audubon Society, 1208 Washington St., Vicksburg, MS 39183

### Abstract

The family *Rhabdoviridae* is a diverse group of non-segmented, negative-sense RNA viruses that are distributed worldwide and infect a wide range of hosts including vertebrates, invertebrates, and plants. Of the 114 currently recognized vertebrate rhabdoviruses, relatively few have been well characterized at both the antigenic and genetic level; hence, the phylogenetic relationships between many of the vertebrate rhabdoviruses remain unknown. The present report describes a novel rhabdovirus isolated from the brain of a moribund American coot (*Fulica americana*) that exhibited neurological signs when found in Durham County, North Carolina, in 2005. Antigenic characterization of the virus revealed that it was serologically unrelated to 68 other known vertebrate rhabdoviruses. Genomic sequencing of the virus indicated that it shared the highest identity to Tupaia rhabdovirus (TUPV), and as only previously observed in TUPV, the genome encoded a putative C protein in an overlapping open reading frame (ORF) of the phosphoprotein gene and a small hydrophobic protein located in a novel ORF between the matrix and glycoprotein genes. Phylogenetic analysis of partial amino acid sequences of the nucleoprotein and polymerase proteins indicated that, in addition to TUPV, the virus was most closely related to avian and small mammal rhabdoviruses from Africa and North America. In this report, we present the morphological, pathological, antigenic, and genetic characterization of the new virus, tentatively named Durham virus (DURV), and discuss its potential evolutionary relationship to other vertebrate rhabdoviruses.

<sup>f</sup>Corresponding author: allisona@uga.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Keywords

rhabdovirus; Durham virus; vesiculovirus; C protein; small hydrophobic protein; overprinting

---

## Introduction

The *Rhabdoviridae* is one of the four families classified within the order *Mononegavirales* (Pringle and Easton, 1997). The rhabdovirus genome, like those of paramyxoviruses, filoviruses, and *Borna disease virus* (BDV), is a non-segmented, single strand of negative-sense RNA, although the genome of Orchid fleck virus (OFV) has been proposed to be bipartite (Kondo *et al.*, 2006). Six genera are currently recognized within the *Rhabdoviridae*, four of which, *Vesiculovirus*, *Lyssavirus*, *Ephemerovirus*, and *Novirhabdovirus*, infect vertebrates, while members of *Cytorhabdovirus* and *Nucleorhabdovirus* infect plants (Tordo *et al.*, 2005). Many of the vertebrate and plant rhabdoviruses replicate in, and are transmitted by, arthropod vectors, although some viruses (e.g., lyssaviruses) are spread by direct contact without any apparent arthropod component.

Of the presently recognized 114 vertebrate rhabdoviruses, 52 are classified as members (or tentative members) of one of the four existing genera (Tordo *et al.*, 2005). Of the 62 remaining unclassified vertebrate rhabdoviruses, 20 are placed within six serogroups based on antigenic cross-reactivity, while 42 remain unassigned to any existing serogroup. Although the antigenic relationships for many vertebrate rhabdoviruses have been determined through serological studies (Tesh *et al.*, 1983; Calisher *et al.*, 1989), the phylogenetic relationships between many of these viruses remain unknown due to the lack of available sequence data. Additionally, as many of the unclassified vertebrate rhabdoviruses are represented by only a single or few isolates, the normal host associations, transmission cycles, and geographical distributions of these viruses remain obscure.

The prototypical rhabdovirus genome consists of five genes [nucleoprotein (N), phosphoprotein (P), matrix (M), glycoprotein (G), polymerase (L)] that are sequentially transcribed in decreasing molar abundance according to their placement in the genome (i.e., 3'-N>P>M>G>L-5') (Abraham and Banerjee, 1976). The genomic 3' and 5' ends contain the untranslated leader and trailer sequences, respectively, which exhibit some degree of terminal complementarity (Wertz *et al.*, 1994; Rose and Whitt, 2001). Each of the five individual genes is flanked by transcription initiation and termination/polyadenylation signals, which may be conserved among members of the same genus (Fu, 2005). Between each transcription unit (gene and associated flanking signals) is a nontranscribed intergenic region that usually contains a single or dinucleotide sequence [e.g., G or GG in Tupaia rhabdovirus (TUPV)] (Springfeld *et al.*, 2005).

In addition to the five structural proteins (N-P-M-G-L) normally encoded in the genome, a number of additional proteins, either encoded in the same or overlapping open reading frame (ORF) within an existing gene, or in a novel ORF, have been identified in numerous rhabdoviruses. For instance, Wongabel virus (WONV), a putative member of the Hart Park serogroup, has recently been demonstrated to encode three novel genes (U1, U2, U3) from three consecutive ORFs located between the P and M genes, in addition to two genes (U4, U5) in overlapping ORFs of the N and G genes, respectively (Bourhy *et al.*, 2005; Gubala *et al.*, 2008). Other gene products in addition to N-P-M-G-L have also been described for vesiculoviruses (Spiropoulou and Nichol, 1993; Springfeld *et al.*, 2005), ephemeroviruses (Walker *et al.*, 1992; Wang *et al.*, 1994), lyssaviruses (Chenik *et al.*, 1995), novirhabdoviruses (Kurath *et al.*, 1985; Alonso *et al.*, 2004), nucleorhabdoviruses (Scholthof *et al.*, 1994; Huang *et al.*, 2003), and cytorhabdoviruses (Tanno *et al.*, 2000;

Dietzgen *et al.*, 2006). As the genomic sequences of a number of both vertebrate and plant rhabdoviruses have become available in recent years, it appears that many members of the family do not conform to the prototypical N-P-M-G-L genomic organization. Currently, the function of most of these novel genes remains unknown (Fu, 2005).

In November 2005, a novel rhabdovirus was isolated from the brain of a moribund American coot (*Fulica americana*) found in Durham County, North Carolina. The virus was provisionally named Durham virus (DURV), after the county in North Carolina where the bird was originally recovered. Comparative analysis of the full-length genome of DURV to other rhabdoviruses indicated that it was most closely related to TUPV. Moreover, as previously only documented with TUPV, the genome of DURV encodes a putative C protein in a second overlapping ORF in the P gene, and also a unique small hydrophobic (SH) protein located between the M and G genes (Springfeld *et al.*, 2005). In this report, we present the morphological, pathological, antigenic, and genetic characterization of DURV and discuss the potential evolutionary relationships of DURV to other vertebrate rhabdoviruses.

## Methods

### Case history

On 08 November 2005, a moribund American coot (*Fulica americana*) was found by a private wildlife rehabilitator in Bahama, Durham County, North Carolina (NC), and was brought to the Piedmont Wildlife Center, Chapel Hill, NC, on the following day for supportive care. At the time of admittance, the bird was ataxic and was unable to stand and died before a physical examination could be completed. The bird was then shipped to the Southeastern Cooperative Wildlife Disease Study at the University of Georgia for diagnostic evaluation (i.e., necropsy and pathological, virological, bacteriological, and toxicological testing).

### Virus isolation

For virus isolation, samples of brain and liver (~0.5cm<sup>3</sup>) were mechanically homogenized in 650µl of medium [1X minimum essential medium (MEM), 2.2 g/l NaHCO<sub>3</sub>, 20% fetal bovine serum (FBS), and 4X antibiotic/antimycotic solution (400 units/ml penicillin, 400 µg/ml streptomycin, 1 µg/ml amphotericin B) (Sigma, St. Louis, MO)]. Homogenized tissues were centrifuged (6700 × g for 10 min) to pellet debris, and an aliquot (100 µl) of clarified supernatant was used to inoculate 2-day-old cultures of Vero cells in a 12-well plate format. For initial characterization, stock virus was inoculated into three 75 cm<sup>2</sup> flasks of Vero cells at an m.o.i. of 0.1 and virus was precipitated on day 4 post-inoculation with polyethylene glycol (PEG) as described by Killington *et al.* (1996). cDNA synthesis from extracted RNA was carried out using random decamers, ImProm-II reverse transcriptase (Promega, Madison, WI), and GoTaq Flexi DNA polymerase (Promega), according to the manufacturer's instructions. Amplicons were excised and purified from agarose using a QIAquick Gel Extraction Kit (Qiagen), cloned using a PCR Cloning<sup>Plus</sup> Kit (Qiagen), and subsequently purified using a QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. Sequencing of clones was performed using an Applied Biosystems Inc. 3100 Genetic Analyzer (Foster City, California, USA).

### Transmission electron microscopy

Infected monolayers of Vero cells were fixed in a mixture of 2.5% formaldehyde and 0.1% glutaraldehyde containing 0.03% trinitrophenol and 0.03% CaCl<sub>2</sub> in 0.05 M cacodylate buffer. Cells were scraped off the plastic, pelleted in buffer, post-fixed in 1% OsO<sub>4</sub> in 0.1 M cacodylate buffer, stained en bloc with 2% aqueous uranyl acetate, dehydrated in ethanol,

and embedded in Poly/Bed 812 (Polysciences, Warrington, PA). Ultrathin sections were cut on a Reichert-Leica Ultracut S ultramicrotome, stained with 0.4% lead citrate, and examined in a Philips 201 electron microscope at 60 kV.

### Experimental infection of newborn mice

Two-day old newborn mice (ICR strain, Harlan Sprague-Dawley, Indianapolis, IN) were inoculated intracerebrally with approximately  $10^5$  PFU of a stock of DURV prepared from infected Vero cells. When the mice appeared severely ill, they were euthanized with CO<sub>2</sub> gas and a necropsy was performed. Samples of lung, liver, spleen, kidney and brain were removed and placed in 10% buffered formalin for fixation.

### Histologic and immunohistochemical examinations

Tissues from the coot (brain) and experimentally-infected mice (lung, liver, spleen, kidney, brain) were fixed in 10% formalin for 24 hr and then transferred to 70% ethanol for storage and subsequent embedding in paraffin and sectioning. Mouse and coot tissue sections (3–5  $\mu$ m) were prepared for immunohistochemistry (IHC) or stained with hematoxylin and eosin (H&E). IHC for DURV antigen was performed as described previously (Xiao *et al.*, 2001a; Xiao *et al.*, 2001b). DURV mouse hyperimmune ascitic fluid was used as the primary antibody at a dilution of 1:100 and incubated overnight at 4 °C. An ISO-IHC immunostain kit (Inno-Genex, San Ramon, CA) was used to detect bound primary antibody and to prevent nonspecific binding between species (Xiao *et al.*, 2001a; Xiao *et al.*, 2001b).

### Antigens and immune reagents

Antigens used in complement fixation (CF) tests and for immunizing animals were infected newborn mouse brains prepared by the sucrose-acetone extraction method (Clarke and Casals, 1958). Specific hyperimmune mouse ascitic fluids were prepared against each of the 68 vertebrate rhabdoviruses listed in Supplementary Table 1. The immunization schedule consisted of four intraperitoneal injections given at weekly intervals. Immunogens consisted of 10% suspensions of homogenized infected mouse brain mixed with equal volumes of Freund's adjuvant just prior to inoculation. Sarcoma 180 cells were also given intraperitoneally with the final immunization in order to induce ascites formation. Five adult hamsters were also inoculated with DURV antigen prepared from a frozen harvest of infected baby hamster kidney (BHK) cells grown in MEM with 10% hamster serum. The immunization schedule was the same, except that sarcoma 180 cells were not used. All animal work was carried out under an animal protocol approved by the University of Texas Medical Branch.

### Serologic testing

Complement fixation tests were done according to a microtechnique described previously (Tesh *et al.*, 1983), using 2 full units of guinea-pig complement. Titers were recorded as the highest dilutions giving 3+ or 4+ fixation of complement on a scale of 0 to 4+.

### In vitro host range and growth kinetics

The host range and replicative capacity of DURV was assessed in the following mammalian, avian, reptilian, fish, and mosquito cell lines: Vero, CPAE, QNR/K2, PDE, VH-2, TH-1, FHM, and C6/36. Cells were infected with approximately  $10^3$  PFU of DURV in a 12-well plate format and wells were harvested daily for 8 days. All titrations were performed on Vero cells overlaid with 1% gum tragacanth/1X MEM supplemented with 3% FBS and 1X antibiotic/antimycotic solution (Sigma). Four additional rhabdoviruses, 1) Farmington virus (FARV) [CT AN 114 Clone B, unidentified bird spp., Connecticut, 1969], 2) Flanders virus (FLAV) [WV 382-02, sparrow spp., West Virginia, 2002], 3) Klamath virus (KLAV)

[M-1056, montane vole (*Microtus montanus*), Oregon, 1962], and 4) *Vesicular stomatitis Indiana virus* (VSINV) [97-25323, horse (*Equus caballus*), New Mexico, 1997] were also analyzed for comparison.

### Genomic sequencing

Preparation of genomic material for pyrosequencing was performed as described previously (Palacios *et al.*, 2008). DURV RNA was extracted from infected Vero cell supernatant using TRIzol LS (Invitrogen, Carlsbad, CA). Total RNA extracts were treated with DNase I (DNA-free kit, Ambion, Austin, TX) and cDNA was subjected to a modified degenerate oligonucleotide-primed PCR (DOP-PCR) procedure (Palacios *et al.*, 2007). Products >70 base pairs (bp) were selected by column purification (MinElute, Qiagen) and ligated to specific linkers for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT) without fragmentation of the cDNA (Palacios *et al.*, 2008). Removal of primer sequences, redundancy filtering, and sequence assembly were performed with software programs accessible through the analysis applications at the GreenePortal website (<http://tako.cpmc.columbia.edu/Tools/>).

Sequence gaps between the aligned fragments were filled in by specific RT-PCR amplification with primers designed on the pyrosequence data. Specific primer sequences are available upon request. Amplification products were sequenced using ABI PRISM Big Dye Terminator v1.1 Cycle Sequencing kits on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems Inc.). Terminal sequences were obtained by RACE (SMART RACE cDNA Amplification Kit, Clontech, Mountain View, CA). Overlapping primer sets based on the draft genome were designed to facilitate sequence validation by classical dideoxy sequencing. The accumulated data revealed the complete DURV genome (GenBank accession number FJ952155).

### Protein analysis

All molecular visualizations of protein structures were carried out using the Visual Molecular Dynamics (VMD) program (<http://www.ks.uiuc.edu/Research/vmd>) running in an OpenGL 32-bit Windows XP Professional format (Humphrey *et al.*, 1996). The predicted molecular weights, isoelectric points, and grand average hydrophobicity scores of the DURV proteins were determined using the ProtParam tool on the ExpASY server (<http://www.expasy.ch/tools/protparam.html>) (Kyte and Doolittle, 1982; Gasteiger *et al.*, 2003). Phosphorylation sites of the P protein were predicted using the NetPhos 2.0 server (<http://www.cbs.dtu.dk/services/NetPhos/>) (Blom *et al.*, 1999). The putative transmembrane topology of the SH protein was determined using Phobius (<http://phobius.sbc.su.se/>), SOSUI ([http://bp.nuap.nagoya-u.ac.jp/sosui/cgi-bin/adv\\_sosui.cgi](http://bp.nuap.nagoya-u.ac.jp/sosui/cgi-bin/adv_sosui.cgi)), and TMHMM (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) servers. Hydrophobicity plot analysis of the SH protein was performed using the ProtScale program on the ExpASY server (<http://www.expasy.ch/tools/protscale.html>). N-glycosylation sites of the DURV G protein were predicted using the NetNGlyc 1.0 server (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

### Phylogenetic analysis

Evolutionary relationships of DURV with representative rhabdoviruses were derived by construction of phylogenetic trees generated with CLUSTALW alignments of partial amino acid (aa) sequences of the N, G, and L proteins using the neighbor-joining method in the MEGA program (Tamura *et al.*, 2007). Bootstrap values were determined using 2000 replicates. The trees were calculated using Poisson correction and evolutionary distances were represented as the number of amino acid substitutions per site. N, G, and L sequences of 212, 334, and 156 aa in length, respectively, were used in generating phylogenies with the

cognate regions of other rhabdovirus sequences available in GenBank (Bourhy *et al.*, 2005; Kuzmin *et al.*, 2006).

## Results

### Virus isolation

Cytopathic effects were observed in Vero cells inoculated with brain homogenate from the American coot and a diethyl ether-sensitive and 5-bromo-2'-deoxyuridine-resistant virus (DURV) was isolated. Subsequent RT-PCR, using random decamers, of RNA extracted from PEG-precipitated virus yielded a 640 bp clone that, when translated, exhibited identity to a single sequence by BLASTp analysis: TUPV G protein (NCBI accession number AAX47601). Primer walking from this clone resulted in a 2,850 bp contig covering portions of the G and L genes that, similar to the initial clone, shared the highest amino acid identity (45%) to TUPV.

### Pathological examination of bird tissues

Microscopic examination of H&E-stained sections of the brain showed features of both meningitis and encephalitis. Specifically, there was vascular congestion involving both the meninges (Fig. 1A) and the brain parenchyma. In addition, inflammatory cellular infiltration was evident in the meninges, particularly in perivascular areas. Neuronal degeneration was seen in most areas of the parenchyma, with vacuolation and inflammation (Fig. 1B). IHC of brain tissue showed perinuclear staining with viral antigen in some neurons (Fig. 1C). The other tissues examined did not have any significant lesions.

### Ultrastructural studies

In ultrathin sections of infected Vero cells, areas of massive virion formation could be observed in the cytoplasm (Fig. 2A). Rod-like virions ~30 nm in diameter and 140–160 nm long were observed to be budding from the limiting membranes into the expanded membrane-limited compartments formed by rough endoplasmic reticulum (Fig. 2B). Infected cells did not show significant cytopathy.

### Antigenic relationships

In CF tests, the hyperimmune DURV mouse and hamster antisera were positive with the homologous (DURV) mouse brain antigen at dilutions of 512/≥2 and ≥2048/≥2, respectively. The DURV mouse brain antigen was negative in CF tests with antisera against each of the 68 rhabdoviruses listed in Supplementary Table 1, as well as with an antiserum against *Newcastle disease virus* (NDV) (*Avulavirus: Paramyxoviridae*).

### Pathogenicity of DURV in mice and hamsters

Two-day old ICR mice inoculated intracerebrally, with approximately  $10^4$  TCID<sub>50</sub> of DURV, became ill on approximately the sixth day and most were dead by the ninth day after inoculation. Intraperitoneal inoculation of adult mice and hamsters with a crude 10% suspension of brains from the moribund and dead infant mice did not produce detectable illness, but both species had high levels of CF antibodies when tested approximately one month after infection.

### Histopathology in newborn mice

Examination of stained sections of the brain showed that the overall pathologic process was that of multifocal necrosis of different stages. There was no specific preferential distribution and lesions were noted in the cerebral cortex, subcortical nuclei, hippocampus, and cerebellum. Earlier stage lesions showed degeneration or necrosis of neurons, followed by

neutrophilic infiltration in later stage lesions and eventually loss of the neuropil, leading to vacuoles containing heavy mixed inflammatory cellular infiltration predominately composed of neutrophils. In addition, there were multiple foci of vascular reaction characterized by a prominence of endothelial cells, with intravascular and perivascular lymphocytic infiltration. Representative photomicrographs are shown in Fig. 1D-F. Examination of liver, spleen, and kidneys did not reveal significant abnormalities.

### In vitro growth characteristics

*In vitro* growth curve analysis demonstrated that quail glial cells of the neuroretina (QNR/K2) and monkey kidney (Vero), cattle (CPAE), and bat (Tb 1 Lu) cells all supported replication of DURV, but duck embryo (PDE), fish (FHM), mosquito (C6/36), and reptilian (VH-2, TH-1) cell lines were either refractory or relatively non-permissive to infection (Table 1). DURV replicated to lower titers in permissive cell lines (Vero, CPAE, Tb 1 Lu, QNR/K2) than either VSINV, FARV, or KLAV; however, the extent to which the titer of the initial inoculum and passage history of each virus affects its ability to replicate on the different cell lines tested is unknown, and therefore, direct comparisons between viruses of different passage histories may be biased. In contrast, DURV replicated much more efficiently than FLAV [same passage history (Vero P2) as DURV], with FLAV being restricted in both its host cell range and replicative capacity (Table 1), as previously noted by Whitney (1964).

### Genomic and protein analysis

The genome of DURV is 11,265 nt long, encoding 3,784 aa, and is schematically represented as 3'-I-N-P/C-M-SH-G-L-t-5' (Fig. 3A). The putative transcription start and stop/polyadenylation sequences were KUKY and NBACUUUUUUU (NBACU<sub>7</sub>), respectively. The deduced intergenic region between each transcription unit was GA, similar to *Vesicular stomatitis New Jersey virus* (VSNJV) (Stillman and Whitt, 1998). All transcription initiation, intergenic, and transcription termination/polyadenylation sequences, along with the 3' leader and 5' trailer sequences, are shown in Fig. 3B-C.

The DURV N ORF is 1,293 nt long, encoding 430 aa. Pairwise comparisons with other selected rhabdoviruses indicated that the DURV N protein shared the highest aa identity to TUPV (57%) (Table 3). Kolongo virus (KOLV), an African bird-associated rhabdovirus, shared the second closest amino acid identity (34%) (not shown). The region of the N protein reported to be the RNA binding motif conserved among the rhabdoviruses (Kouznetzoff *et al.*, 1998), was <sup>285</sup>-GISKNSPYSS-<sup>294</sup> in DURV, with N<sup>289</sup> being unique.

The DURV P ORF is 1,014 nt long, encoding 337 aa, and contains 25 potential phosphorylation sites. The DURV P protein shared the lowest identity of all of the five major proteins when compared against other rhabdoviruses, with a maximum identity of 16% to TUPV (Table 3). A putative dynein light chain binding motif [(K/R)XTQT] (Jacob *et al.*, 2000;Lo *et al.*, 2001) could not be identified in the DURV P gene.

Similar to that first described in VSNJV (Spiropoulou and Nichol, 1993), the DURV P gene encodes an additional protein, designated as the C protein, from a second overlapping ORF located near the 5' end of the P gene. The putative start codon for the C protein lies 40 nt downstream of the initiator AUG codon for the P gene. Position -3 from the start codon of the P gene is a C, consistent with Kozak's rules for leaky scanning through an upstream AUG codon (Kozak, 1989). The predicted DURV C protein is 136 aa in length, which was 85 aa shorter than the TUPV homolog (221 aa), but significantly larger than the C' proteins reported for VSINV (67 aa) and VSVNJV (65 aa) or the putative C proteins reported for additional rhabdoviruses (Spiropoulou and Nichol, 1993; Schütze *et al.*, 1999; Marriott,

2005; Pauszek *et al.*, 2008; Tao *et al.*, 2008). BLASTp analysis of the C protein of DURV revealed that it shared homologies, covering very similar regions of the DURV C protein (encompassing aa positions 15–75) (Fig. 3E), to two sequences of interest. Using the non-redundant NCBI sequence database, DURV C residues 16–75 shared a 22–23% identity and a 43–45% homology to the dynein heavy chain proteins of two mosquito species found in the southeastern United States, *Culex quinquefasciatus* (NCBI accession number CPIJ002912) and *Aedes aegypti* (NCBI accession number AAEL014313). Additionally, BLASTp analysis using the NCBI protein data bank (PDB) demonstrated that the DURV C protein (aa positions 15–71) had the highest alignment score (i.e., 26% identity and 36% homology) to the fusion domain of the G protein of VSINV (aa positions 55–103; PDB ID 2CMZ) (Fig. 4). No other viral sequences aligned with DURV C in the PDB BLAST.

The M ORF of DURV is 582 nt long, encoding 193 aa. Similar to that observed with the P protein, the DURV M protein exhibited little sequence homology to other rhabdoviruses, with a maximum aa identity of 20% to TUPV (Table 3). The DURV M protein did not contain the highly conserved PPXY motif involved in viral budding reported for other rhabdoviruses (Harty *et al.*, 1999).

As recognized previously only in TUPV, the DURV genome contains an additional ORF of 234 nt between the M and G genes that encodes a putative protein of 77 aa (Springfeld *et al.*, 2005). Based on the nomenclature of Springfield *et al.* (2005), we have also denoted this gene product as the small hydrophobic (SH) protein. The putative transmembrane topology of the DURV SH protein was predicted to contain two transmembrane helices spanning residues 5–26 and 32–48, separated by a short extracellular domain (aa positions 27–31). Residues 1–4 and 49–77 were predicted to be cytoplasmic, with no putative signal sequence associated with the N-terminus. The C-terminal region of the DURV SH protein stretching from residues 50–75 shows homology to the  $\alpha 1$  and  $\alpha 2$  regions of the sterile alpha motif (SAM) domains of the Eph family of receptor tyrosine kinases by BLASTp analysis of the PDB (Stapleton *et al.*, 1999). Specifically, residues 57–63 (RLTGDWL) exhibited a high identity to a portion of the  $\alpha 1$  motif in the SAM domain of EphA3 (RTTGDWL) (PDB ID 3FYD2).

The DURV G ORF is 1,521 nt long, encoding 506 aa. Topological analysis using the Phobius server predicted an N-terminal signal peptide (<sup>1</sup>-MWILLHVSFVASQVII-<sup>17</sup>), followed by an ectodomain (aa 18–475), transmembrane domain (<sup>476</sup>-ILLASITLIALITSTLLLCCVC-<sup>498</sup>), and a short cytoplasmic tail (<sup>499</sup>-KKRQHRSV-<sup>506</sup>), consistent with the overall topology of other rhabdovirus G proteins (Coll, 1995; Walker and Kongsuwan, 1999). DURV, as reported for TUPV (Springfeld *et al.*, 2005), contained only 10 cysteine residues in the G protein and was missing residues C<sub>VIII</sub> and C<sub>X</sub>. DURV G had two potential N-glycosylation sites (NXS/T) at aa positions <sup>317</sup>-NST-<sup>320</sup> and <sup>404</sup>-NNT-<sup>407</sup>. Similar to VSINV and other vesiculoviruses, the DURV G protein contains the aa residues <sup>82</sup>-WY-<sup>83</sup> and <sup>126</sup>-YA-<sup>127</sup> in the proposed two non-contiguous fusion loops that are responsible for G-mediated fusion (Roche *et al.*, 2006; Roche *et al.*, 2007; Sun *et al.*, 2008).

The DURV L ORF is 6,318 nt long, encoding 2,105 aa. The N-terminal LNSPL motif found in all animal rhabdoviruses, other than TUPV and *snakehead rhabdovirus* (SHRV) (Kuzmin *et al.*, 2008), is also different in DURV, with an aspartic acid substitution for asparagine (<sup>41</sup>-LDSPL-<sup>45</sup>). Although the start of the polyadenylation sequence directly follows (N, M) or precedes (P, SH, G) the stop codon for other DURV genes, the first stop codon in the L gene was 68 nt upstream of the start of the putative polyadenylation site (NBACU<sub>7</sub>). However, a second in-frame stop codon (AUC) starts two nt upstream of the start of the polyadenylation signal (Fig. 3B), which may suggest that the first stop codon could possibly be read-through, culminating in an L protein containing an additional 22 aa.



## Phylogenetic analysis

Neighbor-joining phylogenies of the N, G, and L proteins of DURV were constructed against the cognate regions of selected representative rhabdoviruses. Phylogenetic analysis of the partial N aa sequence of DURV indicated that it was most closely related to TUPV and, as previously noted for TUPV by Kuzmin *et al.* (2006), both viruses clustered with KOLV and Sandjimba virus (SJA), two rhabdoviruses isolated from birds from the Central African Republic (Fig. 5). The G protein phylogeny demonstrated that DURV and TUPV formed a monophyletic clade completely separate from all other vertebrate rhabdoviruses (not shown). Phylogenetic analysis of a portion of the L protein indicated that, similar to the N and G phylogenies, DURV was most closely related to TUPV. Additionally, both DURV and TUPV clustered with KLAV to form a distinct group (89% bootstrap support), which was separate from other rhabdoviruses (Fig. 5).

## Discussion

Genomic sequencing of DURV revealed it shared the highest amino acid identity to, and an identical genomic organization with, TUPV, a virus originally isolated from a Northern tree shrew (*Tupaia belangeri*) imported from Thailand (Kurz *et al.*, 1986). As tree shrews are geographically confined to Southeast Asia (Nowak, 1999), the recognition of a rhabdovirus isolated from a bird in the southeastern United States that was most closely related to TUPV was an intriguing finding and raised questions concerning the evolutionary history of these two viruses. Although DURV and TUPV are most closely related to one another based on currently available rhabdovirus sequences, they only share a 42% amino acid identity over their entire genomes, suggesting they are distantly related and have undergone considerable genetic divergence due to contrasting transmission cycles and/or geographical ranges. Additionally, unlike that reported for TUPV (Kurz *et al.*, 1986), DURV was not host-restricted either *in vitro* (Table 1) or *in vivo* (Fig. 1). Nevertheless, DURV and TUPV share a number of common and unique genetic features (as outlined below) which may facilitate inferring evolutionary histories of the *Rhabdoviridae* as a whole, when additional sequence of other vertebrate rhabdoviruses (e.g., bird or small mammal-associated viruses) becomes available.

Like TUPV and most members of the genus *Vesiculovirus*, the DURV genome encodes a putative C protein from a second overlapping reading frame in the P gene (Spiropoulou and Nichol, 1993; Kretzschmar *et al.*, 1996; Springfield *et al.*, 2005; Pauszek *et al.*, 2008). The DURV C protein was not conserved among related viruses, as it exhibited an 8 and 14% amino acid identity to the C proteins of TUPV and VSINV, respectively (Table 3). For VSNJV and VSINV, the C protein has been shown to be a small, highly basic, nonstructural protein that is translated in major (C) and minor (C') forms (Spiropoulou and Nichol, 1991; Kretzschmar *et al.*, 1996). Currently, the precise function of the C protein of the vesiculoviruses remains unknown, although it has been speculated that it may enhance transcriptional activity or be involved in host pathogenicity or insect transmission (Kretzschmar *et al.*, 1996; Peluso *et al.*, 1996).

Protein BLAST analysis of the C protein of DURV, using the NCBI non-redundant sequence database, indicated that it shared identity and homology to the ciliary dynein heavy chain proteins of *Culex* and *Aedes* species of mosquitoes. Previously, it has been demonstrated that the P protein of RABV has a dynein binding motif and that the interaction of the P protein with dynein, rather than promoting retrograde axonal transport as previously suggested (Raux *et al.*, 2000), enhances viral transcription (Tan *et al.*, 2007). By analogy, this could suggest that the C protein, at least in DURV, may potentially act as a competitive inhibitor or functional analog of dynein by binding to and interacting with the P protein during viral infection, thereby modulating transcriptional activity of the polymerase

complex. Additionally, BLAST analysis of the DURV C protein against the NCBI PDB database revealed it had the highest alignment score with the G protein of VSINV (PDB ID 2CMZ), with the area of homology mapping to the fusion domain in the G protein (Fig. 4). That the DURV C protein shared the greatest identity/homology to the VSINV G protein of all the structural sequences in the PDB (~56,000 entries) was intriguing and possibly suggested that this relationship represented a functional commonality, such that the DURV C protein may be fusogenic *in vivo*, or that the homology between the C and G proteins could be indicative of gene duplication in an ancestral virus. However, the evolution of the C gene and its biological role during viral infection are only speculative until other related rhabdoviruses containing bicistronic P genes are sequenced, in addition to performing functional studies to investigate the biological properties of the DURV C protein.

Similar to only TUPV, the DURV genome encodes a small hydrophobic (SH) protein in an ORF located between the M and G genes (Springfeld *et al.*, 2005). Although TUPV and DURV are currently the only two rhabdoviruses known to encode an SH protein, SH proteins have been described, and the functional roles investigated, in other related negative-sense RNA viruses such as pneumoviruses and rubulaviruses within the family *Paramyxoviridae*. In *parainfluenza virus 5* (PIV5) [*simian virus 5*], expression of the SH protein has been demonstrated to inhibit tumor necrosis factor-mediated apoptosis, thereby suggesting the protein may have a role in abrogating viral clearance (Lin *et al.*, 2003; Wilson *et al.*, 2006). Although the SH proteins of paramyxoviruses such as PIV5, *respiratory syncytial virus* (RSV), and *mumps virus* (MuV) share little sequence homology and are predicted to adopt different topologies (i.e., both type I and type II transmembrane proteins), it has been suggested that they all may function similarly to inhibit programmed cell death (He *et al.*, 2001; Lin *et al.*, 2003; Fuentes *et al.*, 2007). Hence, although the DURV SH protein is not predicted to adopt a type I transmembrane topology as TUPV (i.e., it lacks a predicted signal peptide), the 1) two transmembrane domains flanking a short extracellular domain, coupled with a long cytoplasmic C terminus, 2) high level of amino acid conservation in the N-terminal half of the protein, along with the conserved leucine residues (Fig. 3D), and 3) identical genomic placement (Fig. 3A), strongly suggests that the SH protein plays a similar, albeit unknown, biological function in both viruses.

Comparison of the G protein of DURV with other rhabdoviruses revealed that both DURV and TUPV share the same cysteine configuration, in that they are missing the 8<sup>th</sup> and 10<sup>th</sup> cysteine residues as determined by Walker and Kongsuwan (1999). Although the cysteine configuration in the G protein for rhabdoviruses is well conserved, some viruses deviate from the 12 cysteine residue configuration observed in the vesiculoviruses and ephemeroiruses, and those that do, are phylogenetically related. For example, viruses within the genus *Lyssavirus* were demonstrated to be missing the 3<sup>rd</sup> and 5<sup>th</sup> residues, while viruses within the genus *Novirhabdovirus* were missing the 2<sup>nd</sup> and 4<sup>th</sup> residues (Walker and Kongsuwan, 1999), possibly implying that other rhabdoviruses closely related to DURV and TUPV may also share the same cysteine configuration. Secondary structure predictions for each of four genera analyzed (*Lyssavirus*, *Vesiculovirus*, *Empherovirus*, and *Novirhabdovirus*) indicated disulfide bridges between the 8<sup>th</sup>-11<sup>th</sup> and 9<sup>th</sup>-10<sup>th</sup> cysteine residues (Walker and Kongsuwan, 1999), suggesting that the disulfide bridge connections in the G protein of DURV and TUPV, and the subsequent structural motif of the G protein, may be unique.

With VSINV, crystallographic studies have determined the structural organization of the pre- and post-fusion forms of the G protein and mapped the location of the two non-contiguous fusion loops that are involved in mediating fusion to the host cell membrane (72-WY-73 and 116-YA-117) (Roche *et al.*, 2006; Roche *et al.*, 2007). Among rhabdoviruses for which G protein sequences are available, the fusion loop amino acid configuration of

VSINV is shared only by other vesiculoviruses such as *Vesicular stomatitis Alagoas virus* (VSAV), *Cocal virus* (COCV), *Chandipura virus* (CHPV), *Isfahan virus* (ISFV), and *Piry virus* (PIRYV). Although the bipartite fusion loops contain hydrophobic residues that are conserved among vertebrate and plant rhabdoviruses (Roche *et al.*, 2006; Sun *et al.*, 2008), DURV currently appears to be the only other rhabdovirus, other than the aforementioned vesiculoviruses, to contain the same amino acid configuration (<sup>82</sup>-WY-<sup>83</sup> and <sup>126</sup>-YA-<sup>127</sup>) in the exposed ends of the fusion loops.

Kuzmin *et al.* (2006) recently reported on the phylogenetic relationships of seven unclassified rhabdoviruses, including KOLV and SJAV. Neighbor-joining analysis of the N gene revealed that KOLV and SJAV were demonstrated to consistently group with TUPV, possibly suggesting a phylogenetic relationship between African avian rhabdoviruses and TUPV. The fact that DURV is closely related genetically to TUPV, and that both viruses share the same unique genomic organization, reiterates that notion that TUPV may be linked phylogenetically to avian rhabdoviruses. Phylogenetic analysis of the DURV N protein demonstrated a monophyletic origin with TUPV, with both viruses additionally clustering with KOLV and SJAV to form a distinct clade (Fig. 5). As suggested by Kuzmin *et al.* (2006), the phylogenetic relationship between TUPV, KOLV, SJAV, and now DURV, suggest that these related rhabdoviruses are widely distributed on multiple continents (i.e., Africa, Asia, North America) in numerous host species (i.e., small mammals and birds).

Previously, TUPV has been demonstrated to be antigenically related to known vesiculoviruses [i.e., VSAV and *Maraba virus* (MARAV)] (Calisher *et al.*, 1989), leading to its current classification as a tentative member of the *Vesiculovirus* genus (Tordo *et al.*, 2005; Lyles and Rupprecht, 2007). Additionally, TUPV has been shown to be antigenically related to KLAV (Calisher *et al.*, 1989), another tentative vesiculovirus originally isolated from a montane vole (*Microtus montanus*) in Oregon and subsequently from Northern red-backed voles (*Clethrionomys rutilus*) and tundra voles (*Microtus oeconomus*) in Alaska (Johnson, 1965; De and Banerjee, 1999). Phylogenetic analysis of a portion of the L protein of KLAV indicated that it formed a monophyletic clade with TUPV, with both viruses additionally grouping with DURV (Fig. 5).

Although DURV did not exhibit cross-reactivity to members of *Vesiculovirus* genus by CF (Supplementary Table 1), the fact that DURV encodes of putative C protein (which additionally shares homology to the fusion domain of the G protein of VSINV), in addition to the fusion loop motif identity and phylogenetic relationship to TUPV and KLAV, suggests that DURV is related to the vesiculoviruses. However, based on the phylogenetic analysis presented here (Fig. 5), a direct relationship between DURV and classified members of the *Vesiculovirus* genus (e.g., VSINV, VSNJV, CHPV) could not be demonstrated, implying that until additional unclassified (or currently unknown) vertebrate rhabdoviruses are sequenced, the evolutionary pathway of these relationships will remain obscure. Additionally, it would be of interest to see if other related avian (e.g., KOLV and SJAV) or small mammal (e.g., KLAV) rhabdoviruses also share some of the unique genetic features (i.e., C and SH proteins, G protein cysteine configuration, G fusion loop motif) observed in DURV. Further genetic characterization of these and other morphologically or serologically-confirmed, yet unclassified, rhabdoviruses will likely provide additional insight into the phylogenetic relationships among the many vertebrate rhabdoviruses that are, at present, arbitrarily grouped as 'dimarhabdoviruses' (Bourhy *et al.*, 2005; Kuzmin *et al.*, 2009).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to thank I. K. Jordan for review of the manuscript. A. B. A., M. K. K., and D. E. S. were supported through sponsorship from SCWDS member state fish and wildlife agencies; through the Federal Aid to Wildlife Restoration Act (50 Stat. 917) and Grant Agreement 06ERAG0005, Biological Resources Division, U.S. Geological Survey, U.S. Department of the Interior; and through Cooperative Agreements 0596130032CA and 0696130032CA, Veterinary Services, Animal and Plant Health Inspection Service, U.S. Department of Agriculture. R. B. T., A. T. R., V. L. P. and S. Y. X. were supported by NIH contract NO1-AI25489. G. P., K. D., T. B., and W. I. L. were supported by NIH awards AI051292 and AI57158 (Northeast Biodefense Center - W. I. L.), the U. S. Department of Defense, and Google.org.

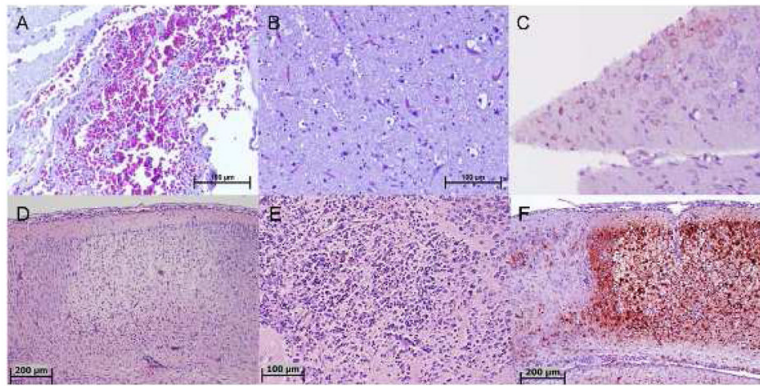
## References

- Abraham G, Banerjee AK. Sequential transcription of the genes of vesicular stomatitis virus. *Proc Natl Acad Sci U S A* 1976;73:1504–1508. [PubMed: 179088]
- Alonso M, Kim CH, Johnson MC, Pressley M, Leong J. The NV gene of snakehead rhabdovirus (SHRV) is not required for pathogenesis, and a heterologous glycoprotein can be incorporated into the SHRV envelope. *J Virol* 2004;78:5875–5882. [PubMed: 15140985]
- Blom N, Gammeltoft S, Brunak S. Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 1999;294:1351–1362. [PubMed: 10600390]
- Bourhy H, Cowley JA, Larrous F, Holmes EC, Walker PG. Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. *J Gen Virol* 2005;86:2849–2858. [PubMed: 16186241]
- Calisher CH, Karabatsos N, Zeller H, Digoutte JP, Tesh RB, Shope RE, Travassos da Rosa AP, St George TD. Antigenic relationships among rhabdoviruses from vertebrates and hematophagous arthropods. *Intervirology* 1989;30:241–257. [PubMed: 2676900]
- Chenik M, Chebli K, Blondel D. Translation initiation at alternate in-frame AUG codons in the rabies virus phosphoprotein mRNA is mediated by a ribosomal leaky scanning mechanism. *J Virol* 1995;69:707–712. [PubMed: 7815533]
- Clarke DH, Casals J. Techniques for hemagglutination and hemagglutination-inhibition with arthropod-borne viruses. *Am J Trop Med Hyg* 1958;7:561–573. [PubMed: 13571577]
- Coll JM. The glycoprotein G of rhabdoviruses. *Arch Virol* 1995;140:827–851. [PubMed: 7605197]
- De, BP.; Banerjee, AK. Ungrouped mammalian, bird and fish rhabdoviruses. In: Granoff, A.; Webster, RG., editors. *Encyclopedia of Virology*. 2. Elsevier Academic Press; San Diego, California: 1999. p. 1541-1544.
- Dietzgen RG, Callaghan B, Wetzel T, Dale JL. Completion of the genome sequence of Lettuce necrotic yellows virus, type species of the genus *Cytorhabdovirus*. *Virus Res* 2006;118:16–22. [PubMed: 16313992]
- Fu ZF. Genetic comparison of the rhabdoviruses from animals and plants. *Curr Top Microbiol Immunol* 2005;292:1–24. [PubMed: 15981465]
- Fuentes S, Tran KC, Luthra P, Teng MN, He B. Function of the respiratory syncytial virus small hydrophobic protein. *J Virol* 2007;81:8361–8366. [PubMed: 17494063]
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003;31:3784–3788. [PubMed: 12824418]
- Gubala AJ, Proll DF, Barnard RT, Cowled CJ, Crameri SG, Hyatt AD, Boyle DB. Genomic characterisation of Wongabel virus reveals novel genes within the Rhabdoviridae. *Virology* 2008;376:13–23. [PubMed: 18436275]
- Harty RN, Paragas J, Sudol M, Palese P. A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with the WW domains of cellular proteins: Implications for viral budding. *J Virol* 1998;73:2921–2929. [PubMed: 10074141]
- He B, Lin GY, Durbin JE, Durbin RK, Lamb RA. The SH integral membrane protein of the paramyxovirus simian virus 5 is required to block apoptosis in MDBK cells. *J Virol* 2001;75:4068–4079. [PubMed: 11287556]

- Huang Y, Zhao H, Luo Z, Chen X, Fang RX. Novel structure of the genome of rice yellow stunt virus: identification of the gene 6-encoded virion protein. *J Gen Virol* 2003;84:2259–2264. [PubMed: 12867659]
- Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. *J Mol Graphics* 1996;14:33–38.
- Jacob Y, Badrane H, Ceccaldi PE, Tordo N. Cytoplasmic dynein LC8 interacts with lyssavirus phosphoprotein. *J Virol* 2000;74:10217–10222. [PubMed: 11024152]
- Johnson HN. Diseases derived from wildlife. *Proc Vert Pest Cont Conf* 1965;2:138–142.
- Killington, RA.; Stokes, A.; Hierholzer, JC. Virus purification. In: Mahy, BWJ.; Kangro, HO., editors. *Virology Methods Manual*. Academic Press; New York, New York: 1996. p. 73-74.
- Kondo H, Maeda T, Shirako Y, Tamada T. Orchid fleck virus is a rhabdovirus with an unusual bipartite genome. *J Gen Virol* 2006;87:2413–2421. [PubMed: 16847138]
- Kozak M. The scanning model for translation: an update. *J Cell Biol* 1989;108:229–241. [PubMed: 2645293]
- Kouznetzoff A, Buckle M, Tordo N. Identification of a region of the rabies virus N protein involved in direct binding to the viral RNA. *J Gen Virol* 1998;79:1005–1113. [PubMed: 9603315]
- Kretzschmar E, Peluso R, Schnell MJ, Whitt MA, Rose JK. Normal replication of vesicular stomatitis virus without C proteins. *Virology* 1996;216:309–316. [PubMed: 8607260]
- Kurath G, Leong JC. Characterization of infectious hematopoietic necrosis virus mRNA species reveals a nonvirion rhabdovirus protein. *J Virol* 1985;53:462–468. [PubMed: 4038520]
- Kurz W, Gelderblom H, Flügel RM, Darai G. Isolation and characterization of a tupaia rhabdovirus. *Intervirology* 1986;25:88–96. [PubMed: 3013798]
- Kuzmin IV, Hughes GJ, Rupprecht CE. Phylogenetic relationships of seven previously unclassified viruses within the family Rhabdoviridae using partial nucleoprotein gene sequences. *J Gen Virol* 2006;87:2323–2331. [PubMed: 16847128]
- Kuzmin IV, Wu X, Tordo N, Rupprecht CE. Complete genomes of Aravan, Khujand, Irkut and West Caucasian bat viruses, with special attention to the polymerase gene and non-coding regions. *Virus Res* 2008;136:81–90. [PubMed: 18514350]
- Kuzmin IV, Novella IS, Dietzgen RG, Padhi A, Rupprecht CE. The rhabdoviruses: biodiversity, phylogenetics, and evolution. *Infect Genet Evol* 2009;9:541–553. [PubMed: 19460320]
- Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982;157:105–132. [PubMed: 7108955]
- Levinson W, Oppermann H, Rubenstein P, Jackson J. Host range restriction of vesicular stomatitis virus on duck embryo cells. *Virology* 1978;85:612–616. [PubMed: 208241]
- Lin Y, Bright AC, Rothermel TA, He B. Induction of apoptosis by paramyxovirus simian virus 5 lacking a small hydrophobic gene. *J Virol* 2003;77:3371–3383. [PubMed: 12610112]
- Lo KWH, Naisbitt S, Fan JS, Sheng M, Zhang M. The 8-kDa dynein light chain binds to its targets via a conserved (K/R)XTQT motif. *J Biol Chem* 2001;276:14059–14066. [PubMed: 11148209]
- Lyles, DS.; Rupprecht, CE. Rhabdoviridae. In: Knipe, DM.; Howley, PM.; Griffin, DE.; Martin, MA.; Lamb, RA.; Roizman, B.; Straus, SE., editors. *Fields Virology*. 5. Vol. 1. Lippincott Williams and Wilkins; Philadelphia, Pennsylvania: 2007. p. 1363-1408.
- Marriott AC. Complete genome sequences of Chandipura and Isfahan vesiculoviruses. *Arch Virol* 2005;150:671–680. [PubMed: 15614433]
- Nowak, RM., editor. *Walker's Mammals of the World*. 6. Vol. 1. John Hopkins University Press; Baltimore, Maryland: 1999. Scandentia: Tupaiidae: Tree Shrews; p. 245-252.
- Palacios G, Quan PL, Jabado OJ, Conlan S, Hirschberg DL, Liu Y, Zhai J, Renwick N, Hui J, Hegyi H, Grolla A, Strong JE, Towner JS, Geisbert TW, Jahrling PB, Buchen-Osmond C, Ellerbrok H, Sanchez-Seco MP, Lussier Y, Formenty P, Nichol MS, Feldmann H, Briese T, Lipkin WI. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg Infect Dis* 2007;13:73–81. [PubMed: 17370518]
- Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, Conlan S, Quan PL, Hui J, Marshall J, Simons JF, Egholm M, Paddock CD, Shieh WJ, Goldsmith CS, Zaki SR, Catton M, Lipkin WI. A new

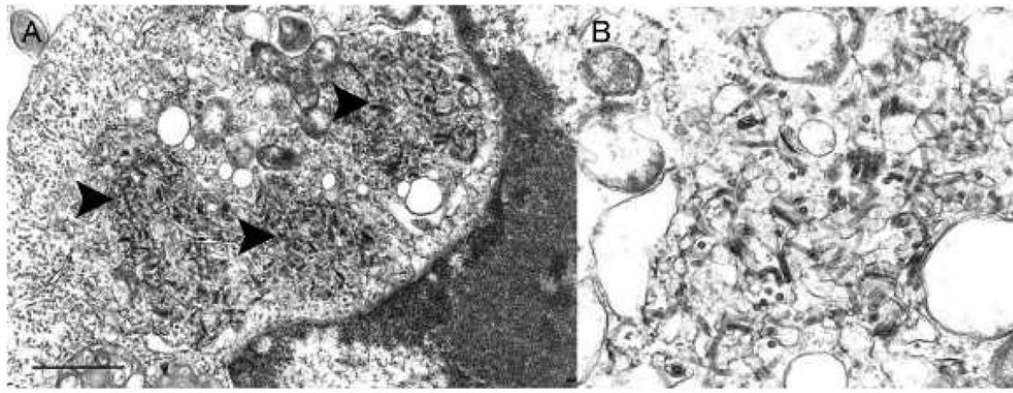
- arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med* 2008;358:991–998. [PubMed: 18256387]
- Pauszek SJ, Allende R, Rodriguez LL. Characterization of the full-length genomic sequences of vesicular stomatitis Cocal and Alagoas viruses. *Arch Virol* 2008;153:1353–1357. [PubMed: 18512001]
- Peluso RW, Richardson JC, Talon J, Lock M. 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* 1996;218:335–342. [PubMed: 8610460]
- Pringle CR, Easton AJ. Monopartite negative strand RNA genomes. *Semin Virol* 1997;8:49–57.
- Raux H, Flamand A, Blondel D. Interaction of the rabies virus P protein with the LC8 dynein light chain. *J Virol* 2000;21:10212–10216. [PubMed: 11024151]
- Roche S, Bressanelli S, Rey FA, Gaudin Y. Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. *Science* 2006;313:187–191. [PubMed: 16840692]
- Roche S, Rey FA, Gaudin Y, Bressanelli S. Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. *Science* 2007;315:843–848. [PubMed: 17289996]
- Rose, JK.; Whitt, MA. Rhabdoviridae: the viruses and their replication. In: Knipe, DM.; Howley, PM.; Griffin, DE.; Martin, MA.; Lamb, RA.; Roizman, B.; Straus, SE., editors. *Fields Virology*. 4. Vol. 1. Lippincott Williams and Wilkins; Philadelphia, Pennsylvania: 2001. p. 1221-1244.
- Scholthof KB, Hillman BI, Modrell B, Heaton LA, Jackson AO. Characterization and detection of sc4: a sixth gene encoded by sonchus yellow net virus. *Virology* 1994;204:279–288. [PubMed: 8091658]
- Schütze H, Mundt E, Mettenleiter TC. Complete genomic sequence of viral hemorrhagic septicemia virus, a fish rhabdovirus. *Virus Genes* 1999;19:59–65. [PubMed: 10499451]
- Spiropoulou CF, Nichol ST. A small highly basic protein is encoded in overlapping frame within the P gene of vesicular stomatitis virus. *J Virol* 1993;67:3103–3110. [PubMed: 8388490]
- Springfeld C, Darai G, Cattaneo R. Characterization of the tupaia rhabdovirus genome reveals a long open reading frame overlapping with P and a novel gene encoding a small hydrophobic protein. *J Virol* 2005;79:6781–6790. [PubMed: 15890917]
- Stapleton D, Balan I, Pawson T, Sicher F. The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat Struct Biol* 1999;6:44–49. [PubMed: 9886291]
- Stillman EA, Whitt MA. The length and sequence composition of vesicular stomatitis virus intergenic regions affect mRNA levels and the site of transcript initiation. *J Virol* 1998;72:5565–5572. [PubMed: 9621014]
- Sun X, Belouzard S, Whittaker GR. Molecular architecture of the bipartite fusion loops of vesicular stomatitis virus glycoprotein G, a class III viral fusion protein. *J Biol Chem* 2008;283:6418–6427. [PubMed: 18165228]
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–1599. [PubMed: 17488738]
- Tan GS, Preuss MA, Williams JC, Schnell MJ. The dynein light chain 8 binding motif of rabies virus phosphoprotein promotes efficient viral transcription. *Proc Natl Acad Sci U S A* 2007;104:7229–7234. [PubMed: 17438267]
- Tao JJ, Zhou GZ, Gui JF, Zhang QY. Genomic sequence of mandarin fish rhabdovirus with an unusual small non-transcriptional ORF. *Virus Res* 2008;132:86–96. [PubMed: 18068257]
- Tanno F, Nakatsu A, Toriyama S, Kojima M. Complete nucleotide sequence of Northern cereal mosaic virus and its genome organization. *Arch Virol* 2000;145:1373–1384. [PubMed: 10963343]
- Tesh RB, Travassos da Rosa AP, Travassos da Rosa JS. Antigenic relationship among rhabdoviruses infecting terrestrial vertebrates. *J Gen Virol* 1983;64:169–76. [PubMed: 6337233]
- Tordo, N.; Benmansour, A.; Calisher, C.; Dietzgen, RG.; Fang, R-X.; Jackson, AO.; Kurath, G.; Nadin-Davis, S.; Tesh, RB.; Walker, PJ. Family Rhabdoviridae. In: Fauquet, CM.; Mayo, MA.; Maniloff, J.; Desselberger, U.; Ball, LA., editors. *Virus Taxonomy: Eighth report of the International Committee on Taxonomy of Viruses*. Elsevier Academic Press; San Diego, California: 2005. p. 623-644.

- Walker PJ, Byrne KA, Riding GA, Cowley JA, Wang Y, McWilliam S. The genome of bovine ephemeral fever rhabdovirus contains two related glycoprotein genes. *Virology* 1992;191:49–61. [PubMed: 1413521]
- Walker PJ, Kongsuwan K. Deduced structural model for animal rhabdovirus glycoproteins. *J Gen Virol* 1999;80:1211–1220. [PubMed: 10355768]
- Wang Y, McWilliam SM, Cowley JA, Walker PJ. Complex genome organization in the GNS-L intergenic region of Adelaide River rhabdovirus. *Virology* 1994;203:63–72. [PubMed: 8030285]
- Wertz GW, Whelan S, LeGrone A, Ball LA. Extent of terminal complementarity modulates the balance between transcription and replication of vesicular stomatitis virus. *Proc Natl Acad Sci USA* 1994;91:8587–8591. [PubMed: 8078927]
- Whitney E. Flanders strain, an arbovirus newly isolated from mosquitoes and birds of New York state. *Am J Trop Med Hyg* 1964;13:123–131. [PubMed: 14106044]
- Wilson RL, Fuentes SM, Wang P, Taddeo EC, Klatt A, Henderson AJ, He B. Function of small hydrophobic proteins of paramyxovirus. *J Virol* 2006;80:1700–1709. [PubMed: 16439527]
- Xiao SY, Guzman H, Zhang H, Travassos da Rosa APA, Tesh RB. West Nile virus infection in the golden hamster (*Mesocricetus auratus*): A model for West Nile encephalitis. *Emerg Infect Dis* 2001a;7:714–721. [PubMed: 11585537]
- Xiao SY, Zhang H, Guzman H, Tesh RB. Experimental yellow fever virus infection in the golden hamster (*Mesocricetus auratus*). 2 Pathology. *J Infect Dis* 2001b;183:1437–1444. [PubMed: 11319680]

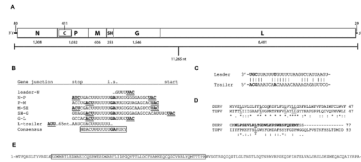


**Fig. 1.** Pathology of DURV in avian and mammalian brain sections. **(A-C)**: Sections of coot brain naturally-infected with DURV. **(A)** Severe congestion in the meninges (H&E); **(B)** Diffuse neuronal degeneration (accompanied by prominent vacuolation), mononuclear inflammatory cellular infiltration, and prominent capillary vessels (H&E); **(C)** Immunohistochemical staining demonstrating neurons with perinuclear DURV antigen. **(D-F)**: Sections of mouse brain experimentally-infected with DURV. **(D)** Neurophil and neurons of the cerebral cortex; inflammatory cellular infiltrate has started from the bottom edge of the lesions. A prominent reactive blood vessel is visible (bottom) and marked meningitis is also present (top) (H&E); **(E)** A necrotic focus in the deeper subcortical nuclei showing loss of neurons and infiltration with many neutrophils and macrophages (H&E); **(F)** Immunohistochemical staining of same area of brain as shown in (D), demonstrating necrosis of the central cortex, with strong staining for DURV antigen. Scattered neurons in the adjacent area (far left) are also positive for viral antigen.

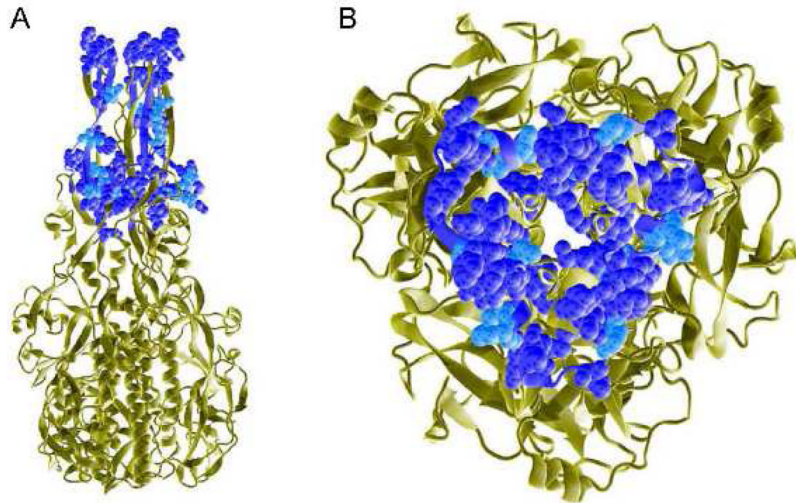




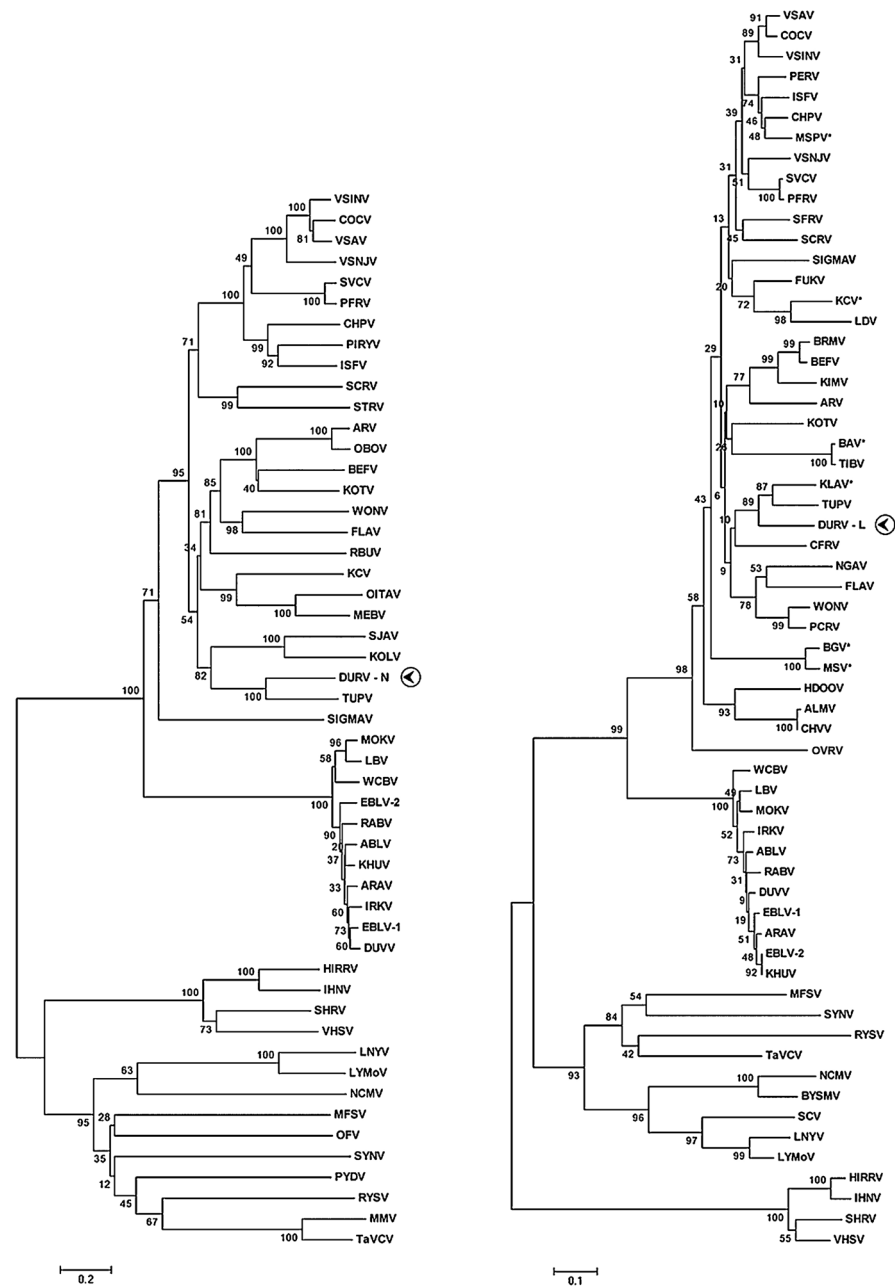
**Fig. 2.** Ultrastructure of DURV in ultrathin sections of Vero cells. **(A)** Portion of a cell cytoplasm with three areas of virus formation (arrowheads); Bar = 1  $\mu\text{m}$ . **(B)** Detail of a virus formation area, demonstrating virions budding into enlarged endoplasmic reticulum cisterns; Bar = 100 nm.



**Fig. 3.** DURV genome and regions of interest. **(A)** Schematic organization of the DURV genome. Including the 3' leader and 5' trailer sequences, the DURV genome was 11,265 nucleotides (nt) in length. The length of each gene, in nt, encoding the N, P, M, SH, G, and L proteins, is listed below the schematic. The length of the C gene (encoded in an overlapping reading frame of the P gene), along with the leader and trailer sequences, are noted above the genome schematic. **(B)** Transcription initiation, intergenic, and transcription termination/polyadenylation sequences. The start and stop codon for each gene is underlined and in bold. The intergenic sequence (i.s.) between each transcription unit, where applicable, is in bold. The consensus sequence for the transcription termination, intergenic (in bold), and transcription initiation regions is boxed. **(C)** Complementarity of the 3' leader and 5' trailer sequences, showing a 65% (19/29nt) inverse identity. Nucleotides that are conserved among the vertebrate rhabdoviruses (positions 1–3; 10) are in bold. **(D)** Amino acid alignment of the SH proteins of DURV and TUPV. Asterisks denote identity; colons and periods represent conserved and semi-conserved amino acid substitutions, respectively. Conserved leucine residues are underlined. The C-terminal region of the SH protein of DURV that shares homology to the  $\alpha 1$  and  $\alpha 2$  regions of the SAM domain of the EphA3 protein tyrosine kinase receptor is shown in bold. **(E)** Amino acid sequence of the DURV C protein. The region of the C protein that shows homology to the fusion domain of the VSINV G protein, as well as to the dynein heavy chains of *Culex quinquefasciatus* and *Aedes aegypti*, is boxed.



**Fig. 4.** DURV C protein identity and homology to the VSINV G protein. The ribbon schematic of the post-fusion homotrimeric form of the G protein of VSINV (PDB 2CMZ) is shown in tan. The region of the VSINV G protein (amino acids 55–103) which aligns with the DURV C protein (amino acids 15–71) is rendered as van der Waals space-filling and corresponds to domain IV, the fusion domain of the VSINV G protein. The 15 amino acid residues of the DURV C protein which are identical to the VSINV G protein are shown in dark blue, while the five amino acid substitutions that are conserved are in light blue. Panel **A** shows a side view of the VSINV G protein with the distal ends of the fusion loops pointing upwards, while panel **B** is an enlarged 90° downward rotation of the fusion loops.



**Figure 5.**

Evolutionary relationships of DURV with representative rhabdoviruses generated by neighbor-joining phylogenies of the N and L proteins. DURV sequences and the corresponding protein used in the phylogeny are indicated with a circled arrowhead. Asterisks beside ICTV abbreviations in the L tree denote new sequences. Bootstrap values were determined using 2000 replicates and are listed at each node. Branch lengths are drawn to scale. The trees were calculated using Poisson correction and evolutionary distances are represented as the number of amino acid substitutions per site. Gaps in the alignments were analyzed by complete deletion. Amino acid sequences used to construct the N and L trees were: ABLV, Australian bat lyssavirus (N: AAD01267; L: NP\_478343); ALMV, Almpiwar virus (L: AAZ43273); ARAV, Aravan virus (N: Q6X1D8; L: ABV03822); ARV, Adelaide

River virus (N: Q65111; L: AAG10421); BAV, Bivens Arm virus (L: GU085726); BEFV, Bovine ephemeral fever virus (N: NP\_065398; L: NP\_065409); BGV, Bahia Grande virus (L: HQ207195); BRMV, Berrimah virus (L: AAZ43265); BYSMV, Barley yellow striate mosaic virus (L: ACT21686); CFRV, China fish rhabdovirus (L: AAX86686); COCV, Cocal virus (N: ACB47434; L: ACB47438); CHPV, Chandipura virus (N: P11211; L: P13179); CHVV, Charleville virus (L: AAZ43300); DURV, Durham virus (FJ952155); DUVV, Duvenhage virus (N: Q66453; L: ABZ81216); EBLV-1, European bat lyssavirus (N: AAX62875; L: ABZ81181); EBLV-2, European bat lyssavirus 2 (N: YP\_001285393; L: ABZ81191); FLAV, Flanders virus (N: AAN73283; L: AAN73288); FUKV, Fukuoka virus (L: AAZ43279); HDOOV, Humpty Doo virus (L: AAZ43271); HIRRV, Hirame rhabdovirus (N: ACO87995; L: NP\_919035); IHNV, Infectious hematopoietic necrosis virus (N: Q08449; L: CAA52076); IRKV, Irkut virus (N: Q5VKP6; L: ABV03823); ISFV, Isfahan virus (N: Q5K2K7; L: Q5K2K3); KCV, Kern Canyon virus (N: ABE69215; L: HQ207197); KHUV, Khujand virus (N: Q6X1D4; L: ABV03824); KIMV, Kimberley virus (L: AAZ43266); KLAV, Klamath virus (L: GU085725); KOLV, Kolongo virus (N: ABE69214); KOTV, Kotonkon virus (N: ABE69213; L: AAZ43267); LBV, Lagos bat virus (N: ABF56214; L: ABZ81171); LDV, Le Dantec virus (L: AAZ43278); LNYV, Lettuce necrotic yellows virus (N: YP\_425087; L: YP\_425092); LYMoV, Lettuce yellow mottle virus (N: YP\_002308371; L: YP\_002308376); MEBV, Mount Elgon bat virus (N: ABE69217); MFSV, Maize fine streak virus (N: YP\_052843; L: YP\_052849); MMV, Maize mosaic virus (N: YP\_052850; L: YP\_052855); MOKV, Mokola virus (N: YP\_142350; L: ABZ81211); MSPV, Malpais Springs virus (L: GU085727); MSV, Muir Springs virus (L: HQ207196); NCMV, Northern cereal mosaic virus (N: NP\_057954; L: NP\_597914); NGAV, Ngaingan virus (L: AAZ43277); OBOV, Obodhiang virus (N: ABE69212); OFV, Orchid fleck virus (N: BAH97109; L: YP\_001294929); OITAV, Oita virus (N: BAD13431); OVRV, Oak Vale virus (L: AAZ43298); PCRV, Parry Creek virus (L: AAZ43275); PERV, Perinet virus (L: AAZ43280); PFRV, Pike fry rhabdovirus (N: ACP27998; L: ACP28002); PIRYV, Piry virus (N: P26037); PYDV, Potato yellow dwarf virus (N: ABW35154); RABV, Rabies virus (N: ACN51666; L: Q66T60); RBUV, Rochambeau virus (N: ABE69218); RYSV, Rice yellow stunt virus (N: NP\_620496; L: NP\_620502); SCR, Siniperca chuatsi rhabdovirus (N: YP\_802937; L: YP\_802942); SCV, Strawberry crinkle virus (L: AAP03645); SFRV, Starry flounder rhabdovirus (L: AAS02285); SHR, Snakehead rhabdovirus (N: NP\_050580; L: NP\_050585); SIGMAV, Sigma virus (N: ACV67011; L: ACU65438); SJAV, Sandjimba virus (N: ABE69216); STRV, Sea trout rhabdovirus (N: AAL35756); SVCV, Spring viremia of carp virus (N: ABW24033; L: Q91DR9); SYN, Sonchus yellow net virus (N: P10550; L: NP\_042286); TaV, Taro vein chlorosis virus (N: YP\_224078; L: YP\_224083); TIBV, Tibrogargan virus (L: AAZ43274); TUPV, Tupaia rhabdovirus (N: YP\_238528; L: YP\_238534); VHSV, Viral hemorrhagic septicemia virus (N: P24378; L: CAB40833); VSINV, Vesicular stomatitis Indiana virus (N: P11212; L: NP\_041716); VSAV, Vesicular stomatitis Alagoas virus (N: ACB47439; L: ACB47443); VSNJV, Vesicular stomatitis New Jersey virus (N: P04881; L: P16379); WCBV, West Caucasian bat virus (N: Q5VKP2; L: ABV03821); WONV, Wongabel virus (N: YP\_002333271; L: AAZ43276).

**Table 1**

In vitro host range and replicative capacity of DURV and other selected rhabdoviruses.

	DURV	FARV	FLAV	KLAV	VSINV
Vero	6.68 (3) <sup>a</sup>	9.15 (2)	5.73 (2)	7.63 (4)	9.20 (2)
CPAE	5.67 (2)	7.60 (3)	nd	6.67 (6)	8.06 (2)
Tb 1 Lu	5.24 (3)	7.62 (3)	nd	4.26 (4)	7.27 (5)
PDE	nd <sup>b</sup>	7.63 (2)	nd	4.20 (3)	6.42 (5) <sup>c</sup>
QNR/K2	6.72 (4)	8.33 (3)	nd	nd	7.55 (2)
C6/36	nd	nd	3.30 (4)	nd	nd
FHM	nd	6.94 (3)	nd	nd	nd
TH-1	nd	6.89 (3)	nd	nd	8.63 (3)
VH-2	nd	nd	nd	nd	7.58 (5)

<sup>a</sup>Maximum titers (log<sub>10</sub> PFU/mL) reached over the 8-day growth period are given, followed by the day at which the maximum titer was reached in parentheses.

<sup>b</sup>nd, not detected above input level; i.e., maximum titer recovered never exceeded the initial inoculum/mL.

<sup>c</sup>PDE cells were essentially non-permissive to VSINV infection, as previously reported by Levinson et al. (1978); titer given at day 5 represents the single well of eight in which a titer above the initial inoculum was recorded. Nucleotide sequencing of the G protein from the initial inoculum versus virus harvested on day 5 did not disclose any nucleotide substitutions, suggesting adaptation of VSINV to PDE cells was not at the level of receptor binding.

**Table 2**

Length of the DURV ORFs and associated untranslated regions in antigenomic orientation and predicted length, molecular weight, isoelectric point, and hydrophobic index of the putative proteins.

DURV gene	5'UTR (nt)	ORF (nt)	3'UTR (nt)	Gene (nt)	Protein (aa)	MW (kDa)	pI (pH)	Hydrop. index
N	4	1293	11	1308	430	48.4	8.48	-0.231
P	12	1014	6	1032	337	36.6	4.70	-0.521
C	-	411	-	411	136	15.7	4.99	-0.476
M	13	582	11	606	193	21.9	6.73	-0.302
SH	13	234	6	253	77	9.2	4.77	0.790
G	19	1521	6	1546	506	57.1	6.38	-0.303
L	4	6318	79	6401	2105	239.9	8.74	-0.145

**Table 3**

Pairwise amino acid identity of DURV proteins to other selected rhabdoviruses [DURV, Durham virus; TUPV, Tupaia rhabdovirus (tentative vesiculovirus); FLAV, Flanders virus (unclassified, Hart Park serogroup); SVCV, Spring viremia of carp virus (tentative vesiculovirus); VSINV, Vesicular stomatitis Indiana virus (genus Vesiculovirus); BEFV, Bovine ephemeral fever virus (genus Ephemerovirus); RABV, Rabies virus (genus Lyssavirus); IHNV, Infectious hematopoietic necrosis virus (genus Novirhabdovirus); LNYV, Lettuce necrosis yellows virus (genus Cytorhabdovirus); SYN, Sonchus yellow net virus (genus Nucleorhabdovirus)].

	DURV	TUPV	FLAV	SVCV	VSINV	BEFV	RABV	IHNV	LNYV	SYNV
N	57	32	32	32	32	31	17	5	2	5
P	16	4	9	8	8	7	7	7	6	6
C	8	-	-	14	-	-	-	-	-	-
M	20	15	8	11	15	5	4	12	2	2
SH	25	-	-	-	-	-	-	-	-	-
G	25	19	18	18	14	14	12	15	8	2
L	51	38	39	41	37	29	12	14	14	14