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


**A B S T R A C T**

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 (SH) protein located in a novel ORF between the matrix and glycoprotein genes. Phylogenetic analysis of partial amino acid sequences of the nucleoprotein and polymerase gene 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.

**1. 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
members of the same genus (Fu, 2005). Between each transcription/polyadenylation signals, which may be conserved among 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., 2005). 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), ephemerviruses (Walker et al., 1992; Wang et al., 1994), lyssaviruses (Chenik et al., 1995), novirhabdoviruses (Kurath and Leong, 1985; Alonso et al., 2004), and cytorhabdoviruses (Tao et al., 2000; Dietzgen et al., 2006). As the genomic sequences of a number of both vertebrate and invertebrate 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.

2. Methods

2.1. Case history

On 08 November 2005, a moribund American coot (F. 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).

2.2. Virus isolation

For virus isolation, samples of brain and liver (~0.5 cm³) were mechanically homogenized in 650 μL of medium [1× minimum essential medium (MEM), 2.2 g/L NaHCO₃, 20% fetal bovine serum (FBS), and 4× 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 cm² 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 CloningMix® 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, CA, USA).

2.3. Transmission electron microscopy

Infected monolayers of Vero cells were fixed in a mixture of 2.5% formaldehyde and 0.1% glutaraldehyde containing 0.03% tri- tolene phosphine and 0.03% CaCl₂ in 0.05 M cacodylate buffer. Cells were scraped off the plastic, pelleted in buffer, post-fixed in 1% OsO₄ 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.

2.4. Experimental infection of newborn mice

Two-day-old newborn mice (ICR strain, Harlan Sprague–Dawley, Indianapolis, IN) were inoculated intracerebrally with approximately 10⁵ PFU of a stock of DURV prepared from infected Vero cells. When the mice appeared severely ill, they were euthanized with CO₂ gas and a necropsy was performed. Samples of lung, liver, spleen, kidney and brain were removed and placed in 10% buffered formalin for fixation.

2.5. Histologic and immunohistochemical examinations

Tissues (lung, liver, spleen, kidney, brain) from the coot and experimentally infected mice were fixed in 10% formalin for 24 h and then transferred to 70% ethanol for storage and subsequent embedding in paraffin and sectioning. Mouse and coot tissue sections (3–5 μm) 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,b). 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,b).

2.6. 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 same, except that sarcoma 180 cells were not used. All animal experiments were carried out under an animal protocol approved by the University of Texas Medical Branch.

2.7. 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. Titters were recorded as the highest dilutions giving 3+ or 4+ fixation of complement on a scale of 0–4+

2.8. In vitro host range and growth kinetics

The host range and replicative capacity of DURV were 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/1× MEM supplemented with 3% FBS and 1× 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.

2.9. 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-primer 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).

2.10. 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/advsosui.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/protscal.html). N-Glycosylation sites of the DURV G protein were predicted using the NetN Glyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/).

2.11. 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 the Poisson correction method 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).

3. Results

3.1. Virus isolation

Cytopathic effects were observed in Vero cells inoculated with brain homogenate from the American coot and a diethyl ether-resistant 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).

3.2. 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.

### 3.3. 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 30–35 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.

### 3.4. Antigenic relationships

In CF tests, the hyperimmune DURV mouse and hamster antisera were positive with the homologous (DURV) mouse brain antigen at dilutions of $512/\geq 2$ and $\geq 2048/\geq 2$, 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).

### 3.5. Pathogenicity of DURV in mice and hamsters

Two-day-old ICR mice inoculated intracerebrally, with approximately $10^4$ TCID$_{50}$ 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 1 month after infection.

### 3.6. 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,
3.7. 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 in 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).

3.8. Genomic and protein analysis

The genome of DURV is 11,265 nt long, encoding 3784 aa, and is schematically represented as 3′-N-P/C–M–SH–G–L-t-5′ (Fig. 3A). The putative transcription start and stop/polyadenylation sequences are KUKY and NBACUUUUUUU (NBACU7), respectively. The deduced intergenic region between each transcription unit is GA, similar to Vesicular stomatitis New Jersey virus (VSNJ) (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 and C.

The DURV N ORF is 1293 nt long, encoding 430 aa (Table 2). Pair-wise 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 (Kozak et al., 1998), was 285-GISKNSPYSS-294 in DURV, with N289 being unique.

The DURV P ORF is 1,014 nt long, encoding 337 aa, and contains 25 phosphorylation sites. The DURV P protein shared...
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.

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

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. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

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, Vescular 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); SYNV, Sonchus yellow net virus (genus Nucleorhabdovirus)].

<table>
<thead>
<tr>
<th>DURV</th>
<th>TUPV</th>
<th>FLAV</th>
<th>SVCV</th>
<th>VSINV</th>
<th>BEFV</th>
<th>RABV</th>
<th>IHNV</th>
<th>LNYV</th>
<th>SYNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>57</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>31</td>
<td>17</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>P</td>
<td>16</td>
<td>4</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>–</td>
<td>–</td>
<td>14</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>M</td>
<td>20</td>
<td>15</td>
<td>8</td>
<td>11</td>
<td>15</td>
<td>5</td>
<td>4</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>SH</td>
<td>25</td>
<td>19</td>
<td>18</td>
<td>18</td>
<td>14</td>
<td>12</td>
<td>15</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>25</td>
<td>38</td>
<td>39</td>
<td>41</td>
<td>37</td>
<td>29</td>
<td>12</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>L</td>
<td>51</td>
<td>38</td>
<td>39</td>
<td>41</td>
<td>37</td>
<td>29</td>
<td>12</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

the lowest identity of all of the five major proteins when com-
pared 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 identi-
fied 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. Posi-
tion −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 sig-
ificantly 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 anal-
ysis of the C protein of DURV revealed that it shared homologies,
covering very similar regions of the DURV C protein (encompass-
ing aa positions 15–75) (Fig. 3E), to two sequences of interest.
Using the non-redundant NCBI sequence database, DURV 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 pro-
tein data bank (PDB) demonstrated that the DURV C protein (aa
positions 15–71) had the highest alignment score (i.e., 26% identi-
ty and 36% homology) to the fusion domain of the G protein of

The 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 6 411 136 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
Fig. 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 the Poisson correction method 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
The DURV G ORF is 1521 nt long, encoding 506 aa. Topological analysis using the Phobius server predicted an N-terminal signal peptide (1-MWILLHVSFVASQVII-17), followed by an ectodomain (aa 18-475), transmembrane domain (476-948), and a cytoplasmic tail (949-1096), consistent with the overall topology of other rhabdovirus G proteins (Coll, 1995; Walker and Kongsuwan, 1999). DURV, as reported for TUPV (Springfield et al., 2005), contained only 10 cysteine residues in G protein and was missing Cys110 and Cys112. DURV G had two potential N-glycosylation sites (NXXS/T) at positions 292-294 and 295-297. Similar to VSINV and other vesculoviruses, the DURV G protein contains the aa residues 210-WY121 and 126-YA127 in the proposed two non-continuous fusion loops that are responsible for G-mediated fusion (Roche et al., 2006, 2007; Sun et al., 2008).

4. Discussion

Genomic sequencing of DURV revealed it shared the highest amino acid identity to, and an identical genomic organization with, a rhabdovirus originally 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. Unlike that reported for TUPV (Kuczko 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; Springfeld 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 VSINV and VSINS, 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, 1993; 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 RAVV 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 8th and 10th 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 ephemerviruses, and those that do, are phylogenetically related. For example, viruses within the genus *Lyssavirus* were demonstrated to be missing the 3rd and 5th residues, while viruses within the genus *Novirhabdovirus* were missing the 2nd and 4th 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 8–11th and 9–10th 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 conserved.

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, 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 *Pry virus* (PIRV). 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 (82-WY-83 and 126-YA-127) 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 KLV (Calisher et al., 1989), another tentative vesiculovirus originally isolated from a montane vole (Microtus montanus) in Oregon and subsequently from Northern
References


