Ledantevirus: A Proposed New Genus in the Rhabdoviridae Has a Strong Ecological Association with Bats

Kim R. Blasdell, Hilda Guzman, Steven G. Widen, Cadhla Firth, Thomas G. Wood, Edward C. Holmes, Robert B. Tesh, Nikos Vasilakis, and Peter J. Walker

CSIRO Biosecurity, Australian Animal Health Laboratory, Geelong, Victoria, Australia; Center for Biodefense and Emerging Infectious Diseases and Department of Pathology, Center for Tropical Diseases, and Institute for Human Infections and Immunity, The University of Texas Medical Branch, Galveston, Texas; Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, Texas; Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Biological Sciences and Sydney Medical School, Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia

Abstract. The Le Dantec serogroup of rhabdoviruses comprises Le Dantec virus from a human with encephalitis and Keuriliga virus from rodents, each isolated in Senegal. The Kern Canyon serogroup comprises a loosely connected set of rhabdoviruses many of which have been isolated from bats, including Kern Canyon virus from California, Nkolbisson virus from Cameroon, Central African Republic, and Cote d’Ivoire, Kolente virus from Guinea, Mount Elgon bat and Fikirini viruses from Kenya, and Oita virus from Japan. Fukuoka virus isolated from mosquitoes, mites, and cattle in Japan, Barur virus from a rodent in India and Nishimuro virus from pigs in Japan have also been linked genetically or serologically to this group. Here, we analyze the genome sequences and phylogenetic relationships of this set of viruses. We show that they form three subgroups within a monophyletic group, which we propose should constitute the new genus Ledantevirus.

INTRODUCTION

Le Dantec virus (LDV) was isolated in May 1965 at the Le Dantec University Hospital in Dakar, Senegal, from the blood of a 10-year-old girl who had an acute febrile illness with signs of hepatosplenomegaly and fever. The virus was subsequently identified morphologically as a rhabdovirus and found to cross-react solidly in complement-fixation tests with Keuriliga virus (KEUV), which was isolated in 1968 from the liver of a gerbil (Tatera kempi) trapped in a millet and peanut plantation in Saboya, Senegal. Surveys of human sera in Senegal detected no other evidence of either LDV or KEUV infection but KEUV antibody was detected in 1% of gerbils and other rodent species tested. However, LDV antibody was detected in a 47-year-old male in Wales who reported fever, headache, and delirium after being bitten in 1969 by an insect while unloading peanuts from a ship that had come from Nigeria. The patient subsequently developed neurological symptoms diagnosed as Parkinson’s disease but a clear causal relationship with LDV was never established for either of the two human cases.

Although initial studies indicated that KEUV cross-reacts weakly in complement-fixation tests with several vesiculoviruses, phylogenetic analyses using partial L protein (RdRp) sequences suggest that LDV is more closely related to KEUV than to other rodent species. Genome sequences for Fukuoka virus (FUKV), which has been assigned to the Kern Canyon serogroup of rhabdoviruses, were obtained by 50 base pair (bp) paired-end sequencing on the HiSeq2000 Illumina platform. Briefly, confluent monolayers of BHK-21 cells (T25 flask) were infected with the respective viruses and harvested 5–7 days later when extensive cytopathic effect (CPE) were present. Viral RNA was prepared and processed for sequencing as described previously.

MATERIALS AND METHODS

We obtained the complete genome sequence for FUKV and near complete genome sequences for NKOV, BARV, KCV, LDV, KEUV, MEBV, and OITAV, lacking only near-terminal sequences of the 3′- and 5′-UTRs. These sequences were obtained by 50 base pair (bp) paired-end sequencing on the HiSeq2000 Illumina platform. Briefly, confluent monolayers of BHK-21 cells (T25 flask) were infected with the respective viruses and harvested 5–7 days later when extensive cytopathic effect (CPE) were present. Viral RNA was prepared and processed for sequencing as described previously.

*Address correspondence to Kim R. Blasdell, CSIRO Biosecurity, Australian Animal Health Laboratory, 5 Portarlington Road, Geelong, Victoria, Australia. E-mail: Kim.Blasdell@csiro.au
Between 0.4% and 70% of the sequence reads in the samples mapped to the corresponding viral contigs (data available on request). The published genome sequences of FKRV, KOLEV, and NISV were also included in our data set (Supplemental Table 1). Transcription profiling was performed using a previously described method and either primer LDV_U1 (CTAGTCATACCTGTCTAAC) or KCV_U1 (GATG TCTGTCAGGTGTTCC) for the U1 ORFs of LDV and KCV, respectively.22

RESULTS

The genomes ranged in size from 10,863 nt for FUKV (the smallest rhabdovirus genome yet reported) to >11,528 nt for KCV, which is intermediate in length to those of vesiculoviruses and lyssaviruses. The genome organizations are shown in Figure 1A. Each genome contains open reading frames (ORFs) encoding the five canonical structural proteins in the order 3'-N-P-M-G-L-5' (in negative polarity). The length

Figure 1. (A) Schematic representation of ledatevirus genome organizations showing the locations of all open reading frames (ORFs) > 180 nt, with dark gray denoting U1 ORFs, black Fukuoka virus (FUKV) Mx ORF, and pale gray with dashed outline other possible ORFs. (B) Alignment of the U1 proteins of Keuraliba virus (KEUV), Le Dantec virus (LDV), and Kern Canyon virus (KCV). (C) Sequence of the FUKV Mx protein and its predicted membrane topology (TMHMM; http://web.expasy.org).
of each ORF was observed to vary little between viruses with the exception of the P ORF, which ranges in size from 765 nt (BARV, FUKV, NISV) to 1,125 nt (OITAV). Major variations in the length of the P ORF are not unusual within rhabdoviruses. Additional short ORFs (U1) flanked by partially conserved transcription initiation (TI) and transcription termination/polyadenylation (TTT) consensus sequences are situated between the G and L genes in LDV (195 nt), KEUV (195 nt) and KCV (234 nt). Transcription profiling of LDV and KCV U1 proteins established that these genes are transcribed (Supplemental Figure 1). In some viruses, potential alternative ORFs (> 180 nt) are also present within the N (LDV, NISV), P (KEUV), M (FUKV), G (BARV, KEUV, FKRV), and L genes (BARV, OITAV, FKRV, KOLEV) but it is not known if these are expressed. Intergenic regions (IGRs) are generally very short (0–4 nt) with the exception of the [U1-L] IGR in KCV (16 nt), and in FKRV, KOLEV, MEBV and OITAV the M gene overlaps the P gene by 12–16 nt.

In pairwise comparisons of the genomes in our data set, nucleotide sequence identities ranged from 44.8% (FKRV and LDV) to 79.5% (BARV and FUKV) (Table 1). Three distinct subgroups were identified: subgroup A (KCV, LDV and KEUV) with 55.0–69.2% sequence identities; subgroup B (FKRV, KOLEV, MEBV and OITAV) with 50.0–62.0% sequence identities; and subgroup C (BARV, FUKV, NISV and NKOV) with 59.2–79.5% sequence identities. The assignment of these subgroups is supported by pairwise comparisons of amino acid (aa) sequence identities of the most highly conserved proteins (Supplemental Table 2). The BARV and FUKV consistently showed the highest pairwise aa sequence identities for all individual proteins (N = 96.4%; P = 81.1%; M = 95.7%; G = 84.4%; L = 92.2%).

The viruses assigned to subgroup A (LDV, KEUV and KCV) were the only three viruses to contain an additional gene (U1). The PSI-Blast and HHblits homology searches revealed no significant similarity between the U1 proteins and any previously identified protein sequences, and none of the proteins displayed any striking structural characteristics. The U1 proteins of LDV (7.4 kDa) and KEUV (7.3 kDa) were each predicted to be mildly acidic and a pairwise alignment indicated they are closely related (56.3% aa identity). The KCV U1 protein (8.8 kDa) is mildly basic and did not align convincingly with any of those of LDV and KEUV (Figure 1B). Of the alternative ORFs in the structural protein genes, only the 10.4 kDa protein encoded in the alternative ORF in the FUKV M gene (Mx protein) displayed remarkable structural characteristics with a predicted double-membrane spanning topology and a highly basic central ectodomain (Figure 1C). Like many other rhabdovirus accessory proteins, the functions of the U1 and Mx proteins remain unknown.

Maximum likelihood phylogenetic trees were generated from multiple sequence alignments of the deduced aa sequences of the N and L proteins for all viruses in the data set along with select animal rhabdoviruses. Amino acid sequences were aligned using MUSCLE23 and ambiguously aligned regions were removed using the Gblocks program for the L protein only.24 This resulted in alignments of 1,071 aa and 62 taxa for the L protein and 449 aa and 27 taxa for the N protein. Trees were estimated assuming the WAG+Γ model of aa substitution in the program PhyML 3.0, utilizing subtree pruning and regrafting (SPR) branch-swapping.25 The phylogenetic robustness of each node was determined using 1,000 bootstrap replicates and nearest-neighbor branch-swapping. Analysis of each protein indicated that all viruses in our primary data set clustered within the larger dimarhabdovirus (“dipteran-mammal associated rhabdovirus”) supergroup in a strongly supported monophyletic group in the L protein phylogeny (bootstrap support [BSP] = 100), hereafter referred to as the ledantevirus clade (Figure 2).6 The ledantevirus clade comprised three subgroups that corresponded to those identified above based on pair-wise identities of genome nucleotide sequences. Each subgroup was strongly supported in both the L and N protein analyses (BSP = 100 in both N and L trees); however, the relationships between these subgroups were not well resolved (BSP = 79 in the L tree, BSP < 70 in the N tree) (Figure 2).

**DISCUSSION**

The Rhabdoviridae currently contains 11 approved genera (Vesiculovirus, Lyssavirus, Ephemerovirus, Tibrovirus, Novirhabdovirus, Perhabdovirus, Signavirus, Sprivirus, Tupaiviruses, Cytorhabdovirus, and Nucleorhabdovirus) and four unassigned species.26-28 The data presented here provide a strong case for the creation of a new genus Ledantevirus to which the 11 viruses included in this study can be assigned. Phylogenies of the N and L proteins clearly show the monophyletic nature of this group. The full genome nt sequence identities between the members vary from 44.8% to 79.5% and L protein aa identities vary from 46.7% to 92.2%, falling well within the range seen for other genera in the

**Table 1**

<table>
<thead>
<tr>
<th></th>
<th>NKOV</th>
<th>NISHV</th>
<th>BARV</th>
<th>FUKAV</th>
<th>OITV</th>
<th>MEBV</th>
<th>FIKV</th>
<th>KOLEV</th>
<th>KCV</th>
<th>LEDV</th>
<th>KEUV</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKOV</td>
<td>59.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NISHV</td>
<td>59.3</td>
<td>75.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BARV</td>
<td>59.8</td>
<td>76.2</td>
<td>79.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUKAV</td>
<td>45.4</td>
<td>46.8</td>
<td>46.5</td>
<td>47.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OITV</td>
<td>47.2</td>
<td>49</td>
<td>49.3</td>
<td>49.3</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEBV</td>
<td>46.6</td>
<td>47.2</td>
<td>46.7</td>
<td>47</td>
<td>53</td>
<td>50.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIKV</td>
<td>46.7</td>
<td>47.8</td>
<td>47.4</td>
<td>47.8</td>
<td>53.3</td>
<td>50.8</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOLEV</td>
<td>47.6</td>
<td>48.9</td>
<td>48.8</td>
<td>49.3</td>
<td>45.4</td>
<td>46.8</td>
<td>45.1</td>
<td>45.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCV</td>
<td>48.2</td>
<td>49.2</td>
<td>49.4</td>
<td>49.4</td>
<td>45.8</td>
<td>47.6</td>
<td>44.8</td>
<td>46.1</td>
<td>55.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEDV</td>
<td>48.1</td>
<td>48.9</td>
<td>48.6</td>
<td>49.7</td>
<td>45.7</td>
<td>47.3</td>
<td>45.8</td>
<td>45.8</td>
<td>55</td>
<td>69.2</td>
<td></td>
</tr>
<tr>
<td>KEUV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>69.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Rhabdoviridae. The genome organizations of the viruses are also similar, comprising five ORFs encoding the structural proteins, concise intergenic regions and a small additional ORF between the G and L in only three of the viruses. This relatively simple genome organization is similar to that of vesiculoviruses.

Based on the N and L phylogenies and full genome nt sequence pairwise comparisons, the ledanteviruses can be further subdivided into three subgroups. This is partially supported by the complement fixation tests reported previously, although strong serological cross-reactions were generally only observed between closely related viruses (BARV and FUKV; KEUV, and LEDV); NISV and FKRV were not included in this study. Although this previous study also included Gossas virus, which had previously been identified as a rhabdovirus, subsequent sequence analysis of the source material indicated that the virus used was actually NKOV and that Gossas virus is not a rhabdovirus (data not shown). Some differences in genome organization were also observed between the three subgroups, with an additional ORF between the G and L ORFs in all subgroup A viruses, whereas subgroup B viruses contained a considerably longer P ORF than that found in the other subgroups. Based on the isolation data available for each virus, there are indications that there may also be differences in the ecology of viruses associated with each of these subgroups.

Figure 2. Maximum likelihood phylogenetic trees of the ledanteviruses (boxed) and select members of the Rhabdoviridae based on L (a) and N (b) protein alignments. Sequences generated in this study are shown in bold, and subgroups A, B, and C are indicated above the respective nodes. Bootstrap support values > 70% are indicated below each node; for clarity these values are given for major clades in the L tree only.
Species demarcation criteria for the *Rhabdoviridae* vary among different genera, complicated by differences in interspecies genetic diversities that are likely to be associated with different rates of evolution and/or periods of speciation. However, intra-species sequence diversities have been reported previously for several rhabdoviruses. Diversity analysis of the P proteins of 77 isolates of rabies virus from China identified a minimum aa identity of 85%, whereas intra-species (intra-genotype) P protein variation among 128 lyssaviruses isolated identified aa identities of ≥ 73.5% and inter-species identities of ≥ 65.9%.31,32 Full genome analysis of nine virus isolates of vesicular stomatitis New Jersey virus from throughout the known geographic range of the virus identified minimum aa identities of ≥ 91.5% for all proteins except P for which the minimum identity was 82.8%.33 Analysis of the hememeroviruses bovine ephemeral fever virus (BEFV) and Kimberley virus showed minimum intra-species aa sequence identities of 79.1% for the partial P protein and 94.9% for the partial G protein, whereas maximum inter-species identities between BEFV and the closely related Berrimah virus were 44.9% and 90.8% for the partial P and G proteins, respectively.34 In the current study, aa sequence identities were < 81% for L and N, < 70.0% for M and G, and ≤ 57% for all viruses except BARV, FUKV, and NISV. Amino acid sequence identities between BARV and FUKV were > 90% for the N, M, and L proteins and > 80% when compared with NISV. Amino acid sequence identities of P and G proteins were lower but were still 81.1% and 84.4%, respectively, between BARV and FUKV, and > 70% for NISV. Although identities for most proteins are similar to those observed previously for intra-species diversity of other rhabdovirus species, sequence identity of the G protein is much lower. Sequences analysis of further isolates would be required to ascertain the intra-species sequence diversity for the ledanteviruses and to establish if BARV, FUKV, and NISV should be considered as separate species or genotypes of a single species.

Several of the genera within the *Rhabdoviridae* show associations with a dominant group of vertebrate hosts. Ephemeraloviruses and tibroviruses appear to be hosted by cattle, with many members of each genus isolated either from cattle and/or from mosquitoes or biting midges that feed on cattle.35,36 Lyssaviruses circulate in bats.37,38 Several viruses suspected of being vectored by arthropods. Further studies are needed to assess the genetic diversity, ecology, and pathogenicity of these viruses and to better define the risks they may pose to human and animal health.

Received September 25, 2014. Accepted for publication October 30, 2014. Published online December 8, 2014.

Note: Supplemental figure and tables appear at www.ajtmh.org.

References


