Short Communication

Arboretum and Puerto Almendras viruses: two novel rhabdoviruses isolated from mosquitoes in Peru

Nikos Vasilakis,1,2,3 Fanny Castro-Llanos,4 Steven G. Widen,5 Patricia V. Aguilar,1,2,3 Hilda Guzman,1 Carolina Guevara,4 Roberto Fernandez,4 Albert J. Augoste,1,3 Thomas G. Wood,5 Vsevolod Popov,1,3 Kirk Mundal,4 Elodie Ghedin,6 Tadeusz J. Kochel,7 Edward C. Holmes,8 Peter J. Walker9 and Robert B. Tesh1,2,3

Correspondence
Nikos Vasilakis
nivasila@utmb.edu

Received 29 August 2013
Accepted 13 January 2014

1Center for Biodefense and Emerging Infectious Diseases and Department of Pathology, University of Texas Medical Branch, Galveston, TX, USA
2Center for Tropical Diseases, University of Texas Medical Branch, Galveston, TX, USA
3Institute for Human Infections and Immunity, University of Texas Medical Branch, Galveston, TX, USA
4US Naval Medical Research Unit no. 6, Lima, Peru
5Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA
6Center for Vaccine Research, Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA, USA
7Naval Medical Research Center, Silver Spring, MD, USA
8Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Biological Sciences and Sydney Medical School, University of Sydney, Sydney, NSW 2006, Australia
9CSIRO Animal, Food and Health Sciences, Australian Animal Health Laboratory, Geelong, VIC 3220, Australia

Arboretum virus (ABTV) and Puerto Almendras virus (PTAMV) are two mosquito-associated rhabdoviruses isolated from pools of Psorophora albigena and Ochlerotatus fulvus mosquitoes, respectively, collected in the Department of Loreto, Peru, in 2009. Initial tests suggested that both viruses were novel rhabdoviruses and this was confirmed by complete genome sequencing. Analysis of their 11 482 nt (ABTV) and 11 876 (PTAMV) genomes indicates that they encode the five canonical rhabdovirus structural proteins (N, P, M, G and L) with an additional gene (U1) encoding a small hydrophobic protein. Evolutionary analysis of the L protein indicates that ABTV and PTAMV are novel and phylogenetically distinct rhabdoviruses that cannot be classified as members of any of the eight currently recognized genera within the family Rhabdoviridae, highlighting the vast diversity of this virus family.

Rhabdoviridae is a diverse family of non-segmented, negative-sense ssRNA viruses that infect a wide range of vertebrates, invertebrates and plants (Dietzgen et al., 2012). Here we report the characterization of two novel rhabdoviruses, collected in Puerto Almendras (3° 50’ S, 73° 22’ W) in the Loreto Department of north-eastern Peru. Arboretum virus (ABTV; strain LO-121) and Puerto Almendras virus (PTAMV; strain LO-39) were isolated from pools of female Ochlerotatus fulvus and Psorophora albigena mosquitoes, collected by light trap on 22 February and 25 March 2009, respectively, in the grounds of the botanical garden located in the town. After collection, the mosquitoes were transported on dry ice to the US Naval Medical Research Unit no. 6 in Lima, where they were initially processed, and an aliquot was forwarded to the University of Texas Medical Branch for virus isolation. Here, we demonstrate that ABTV
and PTAMV are novel rhabdoviruses with a distinctive morphology and genome organization, and are phylogenetically divergent from known rhabdoviruses.

On initial inoculation of ABTV and PTAMV into flask cultures of C6/36 (Aedes albopictus) (Igarashi, 1978) cells maintained at 28 °C, viral cytopathic effect (CPE) was observed in 4–5 days. The infected C6/36 cells detached from the plastic surface and displayed clumping; some cells were enlarged or bloated in appearance (Fig. 1a, b). In contrast, neither virus produced detectable CPE in Vero E6 (African green monkey kidney) or BHK-21 (baby hamster kidney) cell cultures maintained at 37 °C. Likewise, they did not produce illness or death when inoculated intracranially in newborn (1–2 day old) mice. A reverse transcription PCR (RT-PCR) assay was used to assess further whether the mammalian cell lines could support replication of the viruses in the absence of CPE. MacVector version 12.7.3 was used to design specific primers Lo-121FW, 5'-AGGCGAACCAACTTACCC-3', and Lo-121REV, 5'-TCGTGCGGATACTGTAGTG-3' (targeting nucleotides 644–663 and 941–923, respectively, in the ABTV genome), and Lo-39FW, 5'-CGGTGCTCATAATCGGTTTC-3', and Lo-39REV, 5'-TGTATTGTGGTGCTCCTGTC-3' (targeting nucleotides 285–305 and 661–642, respectively, in the PTAMV genome) (Rastogi, 2000). Cell-free supernatants of the Vero and BHK-21 cells infected with ABTV or PTAMV were collected at 1, 2, 3, 4, 6 and 8 days post-infection (p.i.) and viral genetic load was evaluated by RT-PCR assay. Viral RNA was extracted using a QIAamp viral RNA mini kit (Qiagen), and amplicons were generated using the Titan one step RT-PCR kit (Roche) and visualized in 1.5 % TAE (Tris-acetate EDTA) agarose gels (Fig. S1, available in the online Supplementary Material). This assay demonstrated that ABTV and PTAMV were detected in the supernatants of each vertebrate cell line, although ABTV was only detected at 1 day p.i. in BHK-21 cells (Fig. S1). Each cell line is permissive for replication of a wide range of other arthropod-transmitted viruses.

In ultrathin sections of infected C6/36 cells, PTAMV virions were found predominantly within large intracytoplasmic vacuoles (Fig. 1c, d). They displayed typical bullet-shaped morphology with particles 40–55 nm in diameter and up to 190 nm long. ABTV virions formed at the cell surface (Fig. 1e). They displayed a thin rod-like shape and were

Fig. 1. Cytopathology and ultrastructure of the viruses ABTV and PTAMV. (a) Control (non-infected) C6/36 cells. Magnification, ×200. (b) CPE seen in C6/36 cells 4 days after ABTV infection; cells are clumping together and detaching; some cells are enlarged and bloated in appearance (arrows). Magnification, ×200. (c, d) PTAMV virions inside intracytoplasmic vacuoles. Bars, 100 nm. N, Nucleus. (e) ABTV virions budding from the cell surface. Bar, 100 nm.
40–45 nm in diameter and up to 460 nm long. Sections were processed as described previously (Vasilakis et al., 2013), and were cut on a Leica EM UC7 ultramicrotome (Leica Microsystems), stained with lead citrate and examined in a Philips 201 transmission electron microscope at 60 kV.

To further characterize ABTV and PTAMV, the complete genome sequences were determined. Briefly, infected C6/36 cells were grown to 90% confluence in T25 culture flasks, and virus harvest, isolation and processing of viral RNA (vRNA) for next generation sequencing (NGS) were as described previously (Vasilakis et al., 2013). The complete genome sequences were assigned GenBank accession numbers KC994644 (ABTV) and KF534749 (PTAMV); accession numbers for the genome sequences of rhabdoviruses used in the phylogenetic analyses are listed in Table S1.

The (−) ssRNA ABTV and PTAMV genomes have a similar organization, each containing long leader (l) and trailer (t) regions and six long ORFs arranged in the order 3′-l-N-P-M-G-U1-L-t-5′ (Fig. 2). The PTAMV genome (11 876 nt) is larger than the ABTV genome (11 482 nt), primarily due to longer intergenic regions and a longer P ORF. Each ORF is bounded by sequences similar to the transcription initiation (TI) and transcription termination-polyadenylation (TTP) sequences of other animal rhabdoviruses. The putative TI sequence is conserved (UUGUU); the TTP sequences feature the consensus motif ACU[U], which occurs in some genes as other variations of the dinucleotide preceding the septa-uridine run. Unlike some other rhabdoviruses, both contain no small alternative or overlapping ORFs; each of the six genes appears to be transcribed as a monocistronic mRNA.

The N ORFs each encode 434 aa mildly acidic polypeptides (ABTV, 49.7 kDa; PTAMV, 49.9 kDa) that share 38.6% overall amino acid sequence identity. CLUSTAL_X alignment of their N proteins with that of Vesicular stomatitis Indiana virus (VSIV) indicated preservation of several known conserved motifs, as well as each of the eight basic residues located in the RNA-binding cavity that are known to coordinate binding to viral genomic RNA in the ribonucleoprotein (RNP) (Green et al., 2006; Luo et al., 2007) (Fig. S2).
The P ORFs encode acidic 280 aa (32.0 kDa) and 320 aa (36.5 kDa) polypeptides, in ABTV and PTAMV, respectively, which, although displaying no significant sequence identity, are similar to other rhabdovirus P proteins in size and net charge. A pairwise alignment of the ABTV and PTAMV P proteins indicated a relatively high level of sequence identity in the N-terminal and C-terminal regions, and that three indels (9, 16 and 10 aa) in the N-terminal accounted for half the size difference between the proteins (data not shown). Rhabdovirus P proteins generally share low sequence identity and are intrinsically disordered with a modular organization comprising structured domains alternating with disordered regions (Karlin et al., 2003).

The M ORFs each encode 159 aa basic proteins in ABTV and PTAMV (19.3 kDa and 19.1 kDa, respectively), which share 49.1% pairwise amino acid identity, but show no match to other known rhabdovirus M proteins in a pBLAST search. However, like other rhabdovirus M proteins, each contains a sequence in the N-terminal region that resembles a late budding domain motif that mediates the interaction of M proteins with cellular proteins during the budding process (Freed, 2002). This motif occurs as PPxS in ABTV and PSxS in PTAMV, rather than PPxY and/or P(S/T)AP, which are characteristic of lyssaviruses, vesiculoviruses and ephemerviruses (Harty et al., 1999; Irie et al., 2004; Wirblitch et al., 2008). The PPxS motif has also been identified recently in Kolente rhabdovirus (Ghedin et al., 2013).

The G ORFs encode 476 aa and 472 aa polypeptides, in ABTV and PTAMV, respectively. Each is predicted to contain six potential N-glycosylation sites in the ectodomain (NetNGlyc 1.0 server), four at similar positions. On this basis, the mature G proteins would each have a molecular mass of ~71 kDa following cleavage at the predicted signal peptidase sites (SignalP). A pairwise alignment with the VSIV G protein indicated preservation of each of the highly conserved cysteine residues (C1–C211) that are characteristic of rhabdovirus G proteins, forming six disulfide bridges (Fig. S3a). As reported for some other rhabdovirus G proteins, two additional cysteine residues also occur in the downstream lateral domain of ABTV and PTAMV, and these are likely to form a seventh disulfide bridge (Walker & Kongsuwan, 1999).

The U1 ORFs encode, in ABTV and PTAMV, respectively, 73 aa (8.9 kDa) and 80 aa (9.5 kDa) small hydrophobic proteins, each with a predicted transmembrane domain, short N-terminal ectodomain and long C-terminal endodomain (49 aa and 44 aa) featuring a high proportion of basic residues (K, R, H) and large aromatic residues (F, Y, W) (Fig. S3b). This structural arrangement is typical of class IA viroporins such as the influenza A virus M2 protein and human immunodeficiency virus 1 Vpu, in which the basic and aromatic residues interact with the membrane and participate in ion channel formation (Nieva et al., 2012). Small hydrophobic proteins are encoded in the genomes of many other animal rhabdoviruses, either between the M and G genes (Tupaia rhabdovirus, Durham virus and Oak Vale rhabdovirus) or in the region downstream of the G gene (ephemeroviruses, tibroviruses and the Hart Park group) (Walker et al., 2011). Those located after the G gene generally have the structural features of class IA viroporins, and the viroporin-like properties have recently been confirmed experimentally for the bovine ephemeral fever virus x1 protein (Joubert et al., 2013).

The L ORFs encode 2062 aa (241 kDa) and 2060 aa (240 kDa) polypeptides, in ABTV and PTAMV, respectively, which share overall 36.1% amino acid sequence identity and contain all the structural motifs of the rhabdovirus RNA-dependent RNA polymerase (L protein).

Lastly, we generated a phylogenetic tree to reveal the relationship of ABTV and PTAMV to other members of the diverse family Rhabdoviridae. The ABTV and PTAMV L protein sequences were compared with those of 46 other rhabdoviruses (members of the genera Cytorhabdovirus, Novirhabdovirus and Nucleorhabdovirus were excluded because their excessive divergence reduced phylogenetic resolution). The GenBank accession numbers for the genome sequences of select rhabdoviruses used in the phylogenetic analyses are listed in Table S1. Protein sequences were aligned using Muscle (Edgar, 2004), and ambiguously aligned regions were removed using the Gblocks program (Talavera & Castresana, 2007), resulting in a sequence alignment of 1091 aa residues. The phylogenetic relationships were determined using the maximum-likelihood (ML) method available in PhyML 3.0 (Guindon et al., 2010), employing the WAG + Γ model of amino acid substitution, and subtree pruning and regrafting (SPR) branch-swapping. The phylogenetic robustness of each node was determined using 1000 bootstrap replicates.

The ML tree of 48 rhabdovirus L protein sequences indicates that ABTV and PTAMV cluster together as a distinct lineage within the major groups of animal rhabdoviruses, which include members of the genera Vesiculovirus and Ephemerovirus (Fig. 3). While both cluster in this phylogeny between Moussa virus (MOUV) and the other animal rhabdoviruses, the long branch lengths and low bootstrap support indicate that these viruses are divergent, experiencing a long evolutionary separation from other rhabdoviruses described to date. Although fine-scale resolution of the evolutionary history of ABTV and PTAMV is not possible based on these data, their divergent nature strongly suggests that they represent novel virus species.

Collectively the data indicate that ABTV and PTAMV are two novel rhabdoviruses isolated from different mosquito species living in close proximity, but little is known about their arthropod and vertebrate host range and epidemiology. Both viruses failed to produce detectable CPE in vertebrate cell lines. Similarly, both viruses failed to produce illness and/or death in newborn mice, suggesting that they may infect vertebrate hosts in the absence of clinical signs of illness, as is also true of Tibrogargan and coastal plains rhabdoviruses (Cybinski et al., 1980; Gubala et al., 2011). However, we acknowledge that cultured cells, particularly...
Vero cells, which are deficient in the IFN response (Mosca & Pitha, 1986), are imperfect models to establish host susceptibility. Phylogenetically, ABTV and PTAMV sit between MOUV and a diverse array of rhabdoviruses isolated from mammals (Kurz et al., 1986), birds (Allison et al., 2011) and insects (Quan et al., 2011; Vasilakis et al., 2013), highlighting the wide diversity of rhabdovirus hosts and the potential for human exposure. Indeed, both *P. albigena* (Watts et al., 1998) and *O. fulvus* (Turell et al., 2008) mosquitoes have been incriminated as competent vectors of several arboviruses of human importance (such as *Venezuelan equine encephalitis virus*, *Eastern equine encephalitis virus* and others) and both have been observed to feed on humans (Jones et al., 2004). However, since both ABTV and PTAMV have so far only been detected once in mosquito pools, a comprehensive and accurate assessment of their full geographical range, prevalence and host range remains to be determined through comprehensive surveillance studies.

The ABTV and PTAMV genomes have similar size and organization, encoding the five canonical structural proteins and a small hydrophobic protein (U1) in a single additional ORF located between the G and L genes. The structural characteristics of U1 suggest that it may function as a viroporin. ORFs encoding structurally similar proteins have been reported in the region between the G and L genes in several other rhabdoviruses, including ephemeroviruses, tibroviruses and members of the Hart Park serogroup, but these viruses have somewhat more complex genome organizations than ABTV and PTAMV, with multiple ORFs encoding additional accessory proteins (Walker et al., 2011).

Furthermore, the L protein phylogeny indicates that ABTV and PTAMV are relatively distantly related to all other known rhabdoviruses. The lack of available data about the host range and prevalence of these viruses underscores the necessity of further studies to decipher this diverse and complex family of viruses.

**Fig. 3.** ML phylogenetic tree of 48 rhabdovirus L protein sequences. ABTV and PTAMV are shaded and bootstrap support values (>70%) are shown for key nodes. All horizontal branch lengths are drawn to a scale of amino acid substitutions per site, and the tree is rooted in the position observed in a broader analysis of the *Rhabdoviridae*. *Genera not yet formally approved by the International Committee on Taxonomy of Viruses.*
Acknowledgements

We thank Anibal Huaynay for fieldwork assistance. This work was supported in part by a grant from the Institute for Human Infections and Immunity, University of Texas Medical Branch (N. V.), and NIH contract HHSN272201000040I/HHSN272200004/D04 (R. B. T., N. V.). E.C.H. is supported by an NHMRC Australia Fellowship. A. J. A. is supported by the J. W. McLaughlin Fellowship. This study was supported by the United States Department of Defense Global Emerging Infections Surveillance and Response System, a Division of the Armed Forces Health Surveillance Center, work unit no. 847705.8200.B0016. The sponsors had no role in this study other than providing funding. The opinions expressed in this work are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense or US Government. None of the authors has a financial or personal conflict of interest related to this study. K. M. and T. J. K. are military service members, and F. C., C. G. and R. F. are employees of the US Government. This work was prepared as part of their official duties.

References


