Genomic and phylogenetic characterization of viruses included in the Manzanilla and Oropouche species complexes of the genus Orthobunyavirus, family Bunyaviridae

Jason T. Ladner,† Nazir Savji,†† Loreen Lofts,1 Amelia Travassos da Rosa,3 Michael R. Wiley,1 Marie C. Gestole,1 Gail E. Rosen,2 Hilda Guzman,3 Pedro F. C. Vasconcelos,4 Marcio R. T. Nunes,5 Tadeusz J. Kochel,6 W. Ian Lipkin,2 Robert B. Tesh3 and Gustavo Palacios1

1Center for Genomic Sciences, United States Army Medical Institute for Infectious Disease, Frederick, MD, USA
2Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY, USA
3Center for Biodefense and Emerging Infectious Diseases, Department of Pathology, University of Texas Medical Branch, Galveston, TX, USA
4Department of Arbovirology and Hemorrhagic Fevers, Instituto Evandro Chagas, Ananindeua, Brazil
5Center for Technological Innovation, Instituto Evandro Chagas, Ananindeua, Brazil
6Virology Department, Naval Medical Research Unit Six, Lima, Peru

A thorough characterization of the genetic diversity of viruses present in vector and vertebrate host populations is essential for the early detection of and response to emerging pathogenic viruses, yet genetic characterization of many important viral groups remains incomplete. The Simbu serogroup of the genus Orthobunyavirus, family Bunyaviridae, is an example. The Simbu serogroup currently consists of a highly diverse group of related arboviruses that infect both humans and economically important livestock species. Here, we report complete genome sequences for 11 viruses within this group, with a focus on the large and poorly characterized Manzanilla and Oropouche species complexes. Phylogenetic and pairwise divergence analyses indicated the presence of high levels of genetic diversity within these two species complexes, on a par with that seen among the five other species complexes in the Simbu serogroup. Based on previously reported divergence thresholds between species, the data suggested that these two complexes should actually be divided into at least five species. Together these five species formed a distinct phylogenetic clade apart from the rest of the Simbu serogroup. Pairwise sequence divergences among viruses of this clade and viruses in other Simbu serogroup species complexes were similar to levels of divergence among the other orthobunyavirus serogroups. The genetic data also suggested relatively high levels of natural reassortment, with three potential reassortment events present, including two well-supported events involving viruses known to infect humans.
INTRODUCTION

Globalization of travel and trade, climate change and ever-growing human population sizes are all contributing to an increase in the emergence of pathogenic viruses (Lipkin, 2013). Many of these viruses are coming from well-characterized and expected groups (e.g. influenza, flaviviruses), whereas others belong to groups that have largely been ignored by past surveillance programmes. An example of the latter are members of the genus Phlebovirus, which, with the notable exceptions of Rift Valley fever and Toscana viruses, were generally thought to be of little current public health importance until the recent emergence of severe fever with thrombocytopenia syndrome virus (SFTSV) and Heartland virus (HRTV) (McMullan et al., 2012; Yu et al., 2011). Another, even larger, neglected group is the orthobunyaviruses. Due to their abundance and diversity, many orthobunyaviruses have yet to be fully sequenced and there are probably many others that have yet to be detected. However, a thorough understanding of the sequence diversity of such circulating viruses is a critical part of surveillance and preparedness for future disease outbreaks.

The genus Orthobunyavirus is the largest in the family Bunyaviridae with over 170 named viruses corresponding to 18 different serogroups and 48 species complexes (Elliott & Blakgör, 2011; Nichol et al., 2005). Although the term ‘serogroup’ is not currently utilized by the International Committee on Taxonomy of Viruses (ICTV), the concept of serogroups has played an important historical role in viral taxonomy (Calisher & Karabatsos, 1988); the classification of arthropod-borne viruses (arboviruses) was initially based on antigenic relationships revealed by serological tests (Casals, 1957). In general, genetic-based classifications are starting to supplant antigenic classifications; however, due to the lack of genetic information for many named viruses in the family Bunyaviridae, most current taxonomic assignments are still based on serological criteria (Nichol et al., 2005; Plyusnin et al., 2012). Thus, in this report, in the interests of continuity and clarity, the term ‘serogroup’ will continue to be used for a group of serologically related viruses and the term ‘species complex’ will be used for ICTV-defined (Nichol et al., 2005) groups of closely related, differently named viruses whose exact taxonomic status remains uncertain because of slight antigenic variation or differences in host range, vector species, geographical distribution and/or pathogenic potential from the designated type species. The purpose of the present report was to explore genetic diversity within the Simbu serogroup of the genus Orthobunyavirus, a diverse and geographically widespread group that includes important human and livestock pathogens (Kinney & Calisher, 1981; Saeed et al., 2001a).

The Simbu serogroup currently comprises 22 officially recognized viruses that have been grouped into seven different species complexes (Akabane, Manzanilla, Oropouche, Sathuperi, Simbu, Shamonda and Shuni; Nichol et al., 2005), as well as several other recently described viruses that have yet to be officially assigned to a species (Aguilar et al., 2011; Figueiredo & Da Rosa, 1988; Goller et al., 2012; Plyusnin et al., 2012; Saeed et al., 2001b). Full genomes have recently been obtained for 12 viruses within the Simbu serogroup (Goller et al., 2012); however, these genome sequences are not equally distributed among the seven species complexes. For example, only one representative, Oropouche virus (OROV), has been fully sequenced from the Oropouche species complex, and no complete sequences are available from the Manzanilla species complex (Kinney & Calisher, 1981). Yet, these are two of the largest species complexes within the Simbu serogroup, and the Oropouche species complex is the only one with members that are known to cause human disease. The lack of complete genomes for all members of these species complexes impacts diagnostic capacity (e.g. the ability to identify conserved/divergent regions for primer design in PCR applications and for recognition with sequence-based diagnostic methods); it also prevents the recognition of reassortants.

Here, we utilized high-throughput sequencing technologies to improve our understanding of the diversity and evolutionary history of these two species complexes by obtaining full genome sequences for 11 previously uncharacterized viruses, including the three remaining members of the Oropouche species complex, four of the five members of the Manzanilla species complex and four other unassigned viruses that have demonstrated genetic and/or antigenic similarities to one of these two species complexes. In order to compare the sequences with previous taxonomic characterizations, serological comparisons were conducted among these 11 uncharacterized viruses and the other fully sequenced members of these two species complexes.

RESULTS AND DISCUSSION

Genome sequences

Genome sequences for 11 viruses (Table 1) were obtained through de novo assembly from either 454 (Roche) or Illumina sequences. Each sequenced orthobunyavirus genome consisted of three distinct RNA segments, and the sizes and organization of the ORFs were generally consistent with what has been described previously for the genus (Plyusnin et al., 2012) (Table S1, available in the online Supplementary Material). The 3’ terminal sequence was obtained for 18 segments (nine different viruses) and the 5’ terminal sequence was obtained for seven segments (five different viruses). In all cases, the 10 most terminal nucleotides were identical to those that have been reported previously for the genus (Plyusnin et al., 2012). Segments without sufficient coverage to assemble the ends were completed using primers targeting these conserved terminal sequences. The large (L) genome segment of members of the genus Orthobunyavirus contains a single ORF that encodes an RNA polymerase; in our sequences, this ORF ranged in size from 6756 to 6783 bases. The
### Table 1. Virus isolates sequenced in this study

<table>
<thead>
<tr>
<th>Virus Strain</th>
<th>Source</th>
<th>Locality</th>
<th>Year</th>
<th>Species*</th>
<th>GenBank accession nos</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANV TRVL 3587</td>
<td>Alouatta seniculus (red howler monkey)</td>
<td>Trinidad</td>
<td>1954</td>
<td>A. seniculus</td>
<td>KF697148–KF697150</td>
<td>Anderson et al. (1960)</td>
</tr>
<tr>
<td>INGV SA An 4165</td>
<td>Alouatta seniculus (red howler monkey)</td>
<td>South Africa</td>
<td>1959</td>
<td>A. seniculus</td>
<td>KF697139–KF697141</td>
<td>McIntosh et al. (1965)</td>
</tr>
<tr>
<td>MERV AV 782</td>
<td>Progne subis (purple martin)</td>
<td>USA</td>
<td>1964</td>
<td>P. subis</td>
<td>KF697151–KF697153</td>
<td>Calisher et al. (1969)</td>
</tr>
<tr>
<td>Cat Que virus VN 04-2108</td>
<td>DCulex sp. (mosquitoes)</td>
<td>Vietnam</td>
<td>2004</td>
<td>Culex sp.</td>
<td>JQ675598–JQ675600</td>
<td>Bryant et al. (2005)</td>
</tr>
<tr>
<td>BUTV BFS 5002</td>
<td>Culicoides sp. (biting midges)</td>
<td>USA</td>
<td>1964</td>
<td>C. sp.</td>
<td>KF697160–KF697162</td>
<td>Reeves et al. (1970)</td>
</tr>
<tr>
<td>FPV Aus Ch 16129</td>
<td>Culex sp. (mosquitoes)</td>
<td>Australia</td>
<td>1974</td>
<td>Culex sp.</td>
<td>KF697136–KF697138</td>
<td>Doherty et al. (1979)</td>
</tr>
<tr>
<td>UTIV Be An 84785</td>
<td>Bradypus tridactylus (pale-throated sloth)</td>
<td>Brazil</td>
<td>1965</td>
<td>B. tridactylus</td>
<td>KF697154–KF697156</td>
<td>Shope et al. (1983)</td>
</tr>
<tr>
<td>UVV Pan An 48878</td>
<td>Bradypus variegates (brown-throated sloth)</td>
<td>Panama</td>
<td>1975</td>
<td>B. variegates</td>
<td>KF697157–KF697159</td>
<td>Seymour et al. (1983)</td>
</tr>
<tr>
<td>JATV BeAn 423380</td>
<td>Nasua nasua (South American coati)</td>
<td>Brazil</td>
<td>1984</td>
<td>N. nasua</td>
<td>JQ675601–JQ675603</td>
<td>Figueiredo &amp; Da Rosa (1988)</td>
</tr>
<tr>
<td>IQTV IQT9924</td>
<td>Homo sapiens (human)</td>
<td>Peru</td>
<td>1999</td>
<td>H. sapiens</td>
<td>KF697142–KF697144</td>
<td>Aguilar et al. (2011)</td>
</tr>
<tr>
<td>FMD1303</td>
<td>Homo sapiens (human)</td>
<td>Peru</td>
<td>2007</td>
<td>H. sapiens</td>
<td>KF697145–KF697147</td>
<td>NA</td>
</tr>
</tbody>
</table>

MANV, Manzanilla virus; INGV, Ingwavuma virus; MERV, Mermet virus; BUTV, Buttonwillow virus; FPV, Facey’s Paddock virus; UTIV, Utinga virus; UVV, Utive virus; JATV, Jatobal virus; IQTV, Iquitos virus.

*D = Previously described as an isolate of OYAV. ‘Cat Que virus’ is an unofficial name proposed here to refer to isolate VN 04-2108. ‘Madre de Dios virus’ is an unofficial name proposed here to refer to isolate FMD1303. NA, Not applicable.

---

**Genus-level phylogenetic relationships**

Phylogenetic analyses of the three genome segments confirmed that the 11 sequenced viruses were all genetically closely related to the previously sequenced Simbu serogroup viruses. The Simbu serogroup formed a monophyletic clade in both the M- and S-segment trees (bootstrap supportive 66.9 and 100, respectively; Figs S1 and S2), and in the L-segment tree the serogroup was paraphyletic, also including Leanyer virus (LEAV; Fig. 1) (Savji et al., 2011). Uncertainty in placing LEAV was the main reason for a low bootstrap in the M-segment tree; without LEAV, the Simbu serogroup clade was supported in 90% of bootstrap analyses. Our extended genetic sampling of the Oropouche and Manzanilla species complexes illuminated a deep evolutionary divide within the Simbu serogroup; one phylogenetic clade (clade A) included the Manzanilla and Oropouche species complexes and a second clade (clade B) included the other five Simbu species complexes (Figs 1, S1 and S2) (Kinney & Calisher, 1981). These monophyletic clades were well supported across all three genomic segments with bootstrap support ranging from 97.5 to 100%, and there was no evidence of reassortment between these clades.

In 1981, Kinney and Calisher used a combination of serological analyses to provide finer levels of classification within the Simbu serogroup (Kinney & Calisher, 1981); the two genetic clades seen in our analysis are consistent with the serocomplexes they identified. Clade B corresponds completely to their original Simbu serocomplex, whereas the Oropouche and Manzanilla species complexes (clade A) each correspond to unique serocomplexes. Pairwise genetic similarities at the amino acid level between clades A and B range from 97.5 to 100%.
**Fig. 1.** Phylogenetic tree of members of the genus *Orthobunyavirus* based on the protein-coding portion of the L segment. The tree was built using translated amino acid sequences in MEGA v5.1 (Tamura et al., 2011) using the neighbour-joining algorithm and a p-distance matrix. The tree is unrooted and the node labels represent percentage bootstrap support values after 1000 resampling events. Filled circles indicate the genomes that were sequenced in this study. Species designations (left brackets) are based on the genetic data presented in this paper. Clade labels on the far right correspond to serogroups.
and B ranged from 56.6 to 61.5 %, 32.4 to 39 % and 59.7 to 71.9 % for the L, M and S segments, respectively, whilst similarities within the clades ranged from 65.9 to 99.7 %, 34.4 to 89.1 % and 69.8 to 100 %, respectively (Fig. 2). In our previous evaluation of LEAV (Savji et al., 2011), minimum percentage amino acid similarities were proposed as criteria for inclusion of viruses within the same group, where the term 'group' refers to the taxonomic division currently occupied by serogroups. Clear distinctions were found between intra- and intergroup genetic similarities at the L and S segments, and cut-offs of 59 and 60 %, respectively, were proposed (Savji et al., 2011). With currently available data, we saw similarly clear distinctions in comparisons of L-segment sequences, even when Simbu serogroup clades A and B were treated separately. In fact, ~77 % (92/120) of the pairwise comparisons between these two clades fell at or below the proposed 59 % similarity cut-off, and 90.8 % (109/120) of the pairwise comparisons were below a similarity cut-off of 60 %, whilst all intraclade comparisons were well above these cut-offs. In general, we found levels of intra- and interserogroup genetic similarity to be less distinct when comparing the S segments, and no pairwise comparisons between clades A and B met the previously proposed cut-off for different serogroups. However, 76.9 % (120/156) of pairwise comparisons between clades A and B exhibited <69 % similarity, whilst all comparisons within the two clades were above this threshold. Furthermore, this degree of S segment similarity was on a par with pairwise comparisons between members of the Wyeomyia and Bunyamwera serogroups (Fig. 2); the

Fig. 2. Pairwise genetic similarities (1–amino acid p-distance) among viruses within and between serogroups of orthobunyaviruses based on the L segment (a) and S segment (b). For the two Simbu serogroup clades, an extra category is presented that includes only the pairwise similarities between these two groups; this is a subset of the intergroup distances for both clade A and clade B. See Tables S2 and S3 for the list of sequences used in these analyses.
former was not included in our previous analysis (Savji et al., 2011). Therefore, we argue that the level of evolutionary divergence between these two Simbu serogroup clades is more consistent with levels of divergence seen among the other orthobunyavirus serogroups than that seen within serogroups.

One of the hallmarks of the Simbu serogroup is its extensive geographical distribution, and this has been suggested to be one of the major factors behind the high level of genetic diversity found within this group (Saeed et al., 2005). Even when considered independently, the two Simbu serogroup clades, both still exhibited extensive geographical distributions (Fig. S3), which may explain the high levels of genetic diversity within each of these two groups compared with many other orthobunyavirus serogroups (Fig. 2). However, clade A (Manzanilla and Oropouche species complex viruses) was unique in being found within the Americas. To date, 77% (10/13) of Manzanilla and Oropouche species viruses have been isolated from North and/or South America (including Inini virus; Saeed et al., 2001), but no clade B viruses have been found yet in this region. However, clade A viruses are not restricted only to the Americas, as isolates of some representatives have been obtained from Australia, South Africa and Vietnam (Bryant et al., 2005; Doherty et al., 1979; McIntosh et al., 1965), and serological evidence exists for the presence of Ingvavuma virus (INGV) in Nigeria, the Central African Republic, India, Thailand and Cyprus (http://www.cdc.gov/arbocat/). These differences in distribution between the two genetic clades may relate to differences in vector and/or host range, which can facilitate or restrict the geographical spread of viruses. More effort on the surveillance and identification of arthropod vectors of viruses in this group is necessary to understand better the forces that have shaped and maintained these distinct distributions.

**Species-level phylogenetic relationships**

Within clade A, two species complexes have been proposed (Nichol et al., 2005); however, the genetic data presented here suggested the presence of at least five different lineages, which should probably be considered distinct species (Figs 1, S1 and S2). Based on a limited number of available sequences, it has been observed that distinct *Orthobunyavirus* species tend to differ by at least 10% when comparing NP amino acid sequences (S segment). Utilizing this criterion, clade A should be divided into five different species with Buttonwillow virus (BUTV) forming its own species apart from the other Manzanilla species complex viruses and with the Oropouche species complex being split into three distinct species. Together, UTIV and UVV formed one of the new species within the current Oropouche species complex; the name Utinga is suggested for this species, as this was the first of its members to be described (Shope et al., 1967). Facey’s Paddock virus (FPV), which was a phylogenetic outlier and highly divergent from all of the other clade A viruses (minimum NP divergence=26%), represents the other new species previously attributed to the Oropouche species complex.

These species divisions were phylogenetically consistent in both the S- and L-segment trees with between-species amino acid divergences of at least 14.1 and 20.9%, respectively, and maximum within-species divergences of 7.3 and 17.3%, respectively. The M-segment phylogeny was also generally consistent with these species, with the exception of the Oropouche species, which was complicated by several potential reassortment events (see below). Excluding Oropouche species viruses, all M-segment, amino acid divergences within species were ≤21.4%, whilst all divergences between species were ≥32.8%. These species were also consistent with serological comparisons (Tables 2 and 3), and in certain cases corresponded with available phenotypic information. For example, mosquitoes are the primary vectors for most of the Manzanilla species complex viruses, whereas BUTV has only been associated with *Culicoides* midges (http://www.cdc.gov/arbocat/). Throughout the rest of the paper, we utilize these five species designations without the use of the term ‘complex’, in order to distinguish them from the species identified previously by the ICTV (Nichol et al., 2005).

**Unassigned viruses**

Four of the viruses sequenced here have not been officially assigned to a species complex by the ICTV. However, using a combination of genetic (Figs 1, S1 and S2) and serological data (Tables 2 and 3), it was clear that one of these viruses (VN04-2108) belonged to the Manzanilla species, whilst the other three [Iquitos (IQTV), Jatobal (IATV) and FMD1303] belonged to the Oropouche species. Virus strain VN 04-2108 was originally reported to be Oya virus (OYAV), based on indirect immunofluorescence assays and its high nucleotide similarity to OYAV, based on a partial S-segment sequence (Bryant et al., 2005). Our S-genome segment was identical to the sequence from the original characterization of VN 04-2108; in fact, this sequence is characteristic of all the ‘Oya’ isolates obtained in Vietnam during that study (Bryant et al., 2005). However, when compared now with all of our newly available sequence data, VN 04-2108 exhibited similar levels of divergence to four named viruses in the Manzanilla species complex: the original OYAV isolate (4.1% amino acid divergence, 9.6% nucleotide divergence), *Manzanilla virus* (MANV, 2.5% amino acid divergence, 13.1% nucleotide divergence), INGV (5.7% amino acid divergence, 11.2% nucleotide divergence) and Mermet virus (MERV, 4.1% amino acid divergence, 14.5% nucleotide divergence). The prototype OYAV virus was isolated from a sick pig in Malaysia during a Nipah virus outbreak and was not available for this study, so all of these divergences were based only on the published portion of the S segment available for the original OYAV isolate (GenBank accession no. AB075611). No genus-wide framework has been proposed for genetically determining which orthobunyaviruses should be uniquely named;
Table 2. Complement fixation (CF) results

Values are displayed as levels of dilution for antibody/antigen (Φ, undiluted). Within the family Bunyaviridae, a CF test generally detects NP antibodies, a marker for the S-segment RNA. In this study, some of the homologous CF titres of four dose hyperimmune ascitic fluid were high (512–1024), probably explaining the more extensive, low-titre heterologous relationship obtained. MDDV, Madre de Dios virus (an unofficial name proposed here to refer to isolate FMD1303); CQV, Cat Que virus (an unofficial name proposed here to refer to isolate VN 04-2108); OROV, Oropouche virus; see Table 1 for other abbreviations. Bold text indicates results from CF tests with antigen and antibody generated from the same virus.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>OROV</th>
<th>IQTV</th>
<th>MDDV</th>
<th>JATV</th>
<th>UTIV</th>
<th>UVV</th>
<th>BUTV</th>
<th>CQV</th>
<th>INGV</th>
<th>MERV</th>
<th>MANV</th>
<th>FPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>OROV</td>
<td>512/≥32</td>
<td>512/≥8</td>
<td>512/≥8</td>
<td>128/≥8</td>
<td>8/≥32</td>
<td>32/≥32</td>
<td>0</td>
<td>32/≥8</td>
<td>0</td>
<td>16/≥8</td>
<td>0</td>
<td>16/≥8</td>
</tr>
<tr>
<td>IQTV</td>
<td>512/≥Φ</td>
<td>512/≥Φ</td>
<td>1024/≥Φ</td>
<td>256/≥Φ</td>
<td>0</td>
<td>64/≥Φ</td>
<td>0</td>
<td>64/≥Φ</td>
<td>8/≥Φ</td>
<td>32/≥Φ</td>
<td>0</td>
<td>16/≥Φ</td>
</tr>
<tr>
<td>MDDV</td>
<td>512/≥Φ</td>
<td>512/≥Φ</td>
<td>1024/≥Φ</td>
<td>256/≥Φ</td>
<td>16/≥Φ</td>
<td>64/≥Φ</td>
<td>0</td>
<td>64/≥Φ</td>
<td>8/≥Φ</td>
<td>0</td>
<td>0</td>
<td>32/≥Φ</td>
</tr>
<tr>
<td>JATV</td>
<td>512/≥Φ</td>
<td>512/≥Φ</td>
<td>1024/≥Φ</td>
<td>256/≥Φ</td>
<td>0</td>
<td>32/≥Φ</td>
<td>0</td>
<td>64/≥Φ</td>
<td>0</td>
<td>16/≥Φ</td>
<td>0</td>
<td>8/≥8</td>
</tr>
<tr>
<td>UTIV</td>
<td>32/≥32</td>
<td>64/≥8</td>
<td>32/≥8</td>
<td>16/≥8</td>
<td>32/≥32</td>
<td>256/≥32</td>
<td>0</td>
<td>32/≥8</td>
<td>0</td>
<td>8/≥8</td>
<td>0</td>
<td>32/≥8</td>
</tr>
<tr>
<td>UVV</td>
<td>32/≥32</td>
<td>64/≥8</td>
<td>64/≥8</td>
<td>16/≥8</td>
<td>32/≥32</td>
<td>256/≥32</td>
<td>0</td>
<td>32/≥8</td>
<td>0</td>
<td>8/≥8</td>
<td>0</td>
<td>16/≥8</td>
</tr>
<tr>
<td>BUTV</td>
<td>64/≥8</td>
<td>64/≥8</td>
<td>64/≥8</td>
<td>32/≥8</td>
<td>0</td>
<td>8/≥Φ</td>
<td>32/≥8</td>
<td>32/≥8</td>
<td>8/≥8</td>
<td>8/≥8</td>
<td>0</td>
<td>8/≥8</td>
</tr>
<tr>
<td>CQV</td>
<td>64/≥Φ</td>
<td>64/≥Φ</td>
<td>64/≥Φ</td>
<td>32/≥Φ</td>
<td>0</td>
<td>8/≥Φ</td>
<td>0</td>
<td>512/≥Φ</td>
<td>64/≥Φ</td>
<td>512/≥Φ</td>
<td>16/≥Φ</td>
<td>0</td>
</tr>
<tr>
<td>INGV</td>
<td>32/≥8</td>
<td>64/≥8</td>
<td>64/≥8</td>
<td>32/≥8</td>
<td>0</td>
<td>16/≥8</td>
<td>0</td>
<td>512/≥Φ</td>
<td>64/≥Φ</td>
<td>512/≥Φ</td>
<td>16/≥8</td>
<td>0</td>
</tr>
<tr>
<td>MERV</td>
<td>8/≥8</td>
<td>32/≥8</td>
<td>64/≥8</td>
<td>16/≥8</td>
<td>0</td>
<td>0</td>
<td>256/≥8</td>
<td>32/≥8</td>
<td>512/≥Φ</td>
<td>8/≥8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MANV</td>
<td>64/≥Φ</td>
<td>64/≥Φ</td>
<td>32/≥Φ</td>
<td>32/≥8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>512/≥Φ</td>
<td>64/≥Φ</td>
<td>512/≥Φ</td>
<td>128/≥Φ</td>
<td>0</td>
</tr>
<tr>
<td>FPV</td>
<td>64/≥Φ</td>
<td>32/≥Φ</td>
<td>64/≥Φ</td>
<td>0</td>
<td>8/≥Φ</td>
<td>32/≥Φ</td>
<td>0</td>
<td>32/≥Φ</td>
<td>8/≥Φ</td>
<td>0</td>
<td>0</td>
<td>512/≥2</td>
</tr>
<tr>
<td>Normal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

However, based on the levels of genetic divergence among currently named viruses in this group, VN 04-2108 should probably be given its own unique name, in which case, Cat Que virus (CQV) is suggested, as this is the name of the community in Vietnam where the infected mosquitoes were collected (Bryant et al., 2005). Alternatively, given the overall genetic and phylogenetic similarity of VN 04-2108, MANV, MERV and INGV across all three genome segments (Figs 1, S1 and S2), it may be more prudent to simplify the nomenclature by referring to all four simply as distinct strains of a single named virus, utilizing the name of the species to which they all belong (e.g. Manzanilla virus VN 04-2108, Manzanilla virus AV 782). UTIV and UVV are also strong candidates for synonymization under a single virus name, as they exhibit even higher levels of sequence similarity than those within the Manzanilla species. High levels of genetic similarity are also seen among named members of the Oropouche species at particular segments; however, relationships are complicated by patterns of reassortment (see below).

FMD1303 is a previously uncharacterized orthobunyavirus that was isolated from a febrile human in the Madre de Dios region of Peru. Serologically, FMD1303 is broadly reactive with OROV, IQTV and JATV, which suggests that it is a member of the Oropouche species, and this is consistent with the available genetic data. The S segment of FMD1303 is identical at the amino acid level to both OROV and IQTV. This is consistent with what is known about the epidemiology of FMD1303, as OROV and IQTV are the only other two Simbu serogroup viruses that have been shown to cause disease in humans (Aguilar et al., 2011; Anderson et al., 1961). However, based on the nucleotide sequence of the S segment, FMD1303 falls outside the range of diversity currently described for OROV/IQTV (Fig. S4). The L-segment phylogeny is also consistent with this finding.

Table 3. Haemagglutination inhibition results

Haemagglutinating antigen preparation was unsuccessful for BUTV, MANV, UVV, JATV, IQTV, FPV and MDDV. See Tables 1 and 2 for abbreviations. Bold text indicates results from tests with antigen and antibody generated from the same virus.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>OROV</th>
<th>UTIV</th>
<th>CQV</th>
<th>INGV</th>
<th>MERV</th>
</tr>
</thead>
<tbody>
<tr>
<td>OROV</td>
<td>1:10</td>
<td>1:16</td>
<td>1:80</td>
<td>1:20</td>
<td>1:40</td>
</tr>
<tr>
<td>IQTV</td>
<td>1:320</td>
<td>1:80</td>
<td>1:80</td>
<td>1:40</td>
<td>1:40</td>
</tr>
<tr>
<td>MDDV</td>
<td>1:320</td>
<td>1:80</td>
<td>1:80</td>
<td>1:40</td>
<td>1:40</td>
</tr>
<tr>
<td>JATV</td>
<td>0</td>
<td>0</td>
<td>1:40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UTIV</td>
<td>1:20</td>
<td>1:20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UVV</td>
<td>1:80</td>
<td>1:80</td>
<td>1:40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BUTV</td>
<td>0</td>
<td>0</td>
<td>1:20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CQV</td>
<td>1:160</td>
<td>1:80</td>
<td>1:2560</td>
<td>1:640</td>
<td>1:320</td>
</tr>
<tr>
<td>INGV</td>
<td>1:80</td>
<td>1:40</td>
<td>1:80</td>
<td>1:160</td>
<td>1:160</td>
</tr>
<tr>
<td>MANV</td>
<td>0</td>
<td>0</td>
<td>1:40</td>
<td>1:40</td>
<td>1:40</td>
</tr>
<tr>
<td>MERV</td>
<td>1:40</td>
<td>1:40</td>
<td>1:160</td>
<td>1:160</td>
<td>1:640</td>
</tr>
<tr>
<td>FPV</td>
<td>1:20</td>
<td>1:20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*1 unit is the highest antigen dilution in which complete or almost complete haemagglutination (HA) occurred. Tests were run with 4 units/0.025 ml.
although for this segment many fewer sequences are available for comparison, whilst the patterns on the M segment are complicated by reassortment (see below). Nonetheless, this virus expands the known diversity of Oropouche species viruses that infect humans. Due to its distinctiveness, we suggest that this virus be given its own name; we propose the name Madre de Dios virus (MDDV), based on the collection locality.

Partial genome sequences from JATV have previously demonstrated an association with the Oropouche species (Saeed et al., 2001a, b), and our genetic and serological analyses confirm this assignment. However, the complete genome sequence of JATV (GenBank accession nos JQ675601–JQ675603) differs from the partial sequences previously reported for the same strain (GenBank accession nos AF312380 and AF312382) (Saeed et al., 2001a, b). These results were confirmed at the Center for Technological Innovation, Genomic Core, Evandro Chagas Institute, by obtaining a second full-length sequence of a lower passage of JATV strain BeAn 423380 (original seed). Sequences of the original seed and the passed virus from the World Reference Center for Emerging Viruses and Arboviruses collection are identical. The largest discrepancy between the JATV sequences described here and those that were reported previously lies in the S segment where the two sequences are only 83.4 % identical at the nucleotide level (90.4 % amino acid identity), with polymorphisms present throughout the ~700 bases. The M-segment sequences, on the other hand, are essentially identical across the 570 bases covered by the partial sequence, except for several differences at the ends of the partial M sequence. We cannot account for these differences in the S-segment sequences, as we could not detect the previously published sequence in either of our JATV samples, but we are confident in the quality of the genome sequence reported here.

Reassortment

Reassortment in multi-segmented viruses is a form of genetic exchange that has the potential to provide many of the benefits of sexual exchange; for example, the rapid introduction of novel variation within a lineage, the uncoupling of beneficial mutations from detrimental changes and the ability to combine multiple beneficial mutations that originated in different lineages (Simon-Loriere & Holmes, 2011). In fact, reassortment has been implicated in multiple instances of host/vector range shifts and changes in pathogenicity (Briese et al., 2006; Idris et al., 2008; Le Nouëñ et al., 2006; Nelson & Holmes, 2007; Parrish & Kawoaka, 2005; Schrauwen et al., 2011). Laboratory experiments have demonstrated that reassortment is common between many bunyaviruses in vitro (Gentsch & Bishop, 1976; Gentsch et al., 1977, 1980; Irocugb & Pringle, 1981; Pringle & Irocugb, 1982; Reese et al., 2008), and a number of recent genomic analyses have suggested that reassortment also plays an important role in viral evolution in natural populations (Aguilar et al., 2011; Briese et al., 2006, 2013; Kobayashi et al., 2007; Nunes et al., 2005; Reese et al., 2008; Yanase et al., 2006, 2010). These examples include multiple viruses within the genus Orthobunyavirus. Given the potential evolutionary implications, it is important to monitor for the prevalence of viral reassortment events, especially in virus groups known to infect mammals.

To look for reassortment events, nucleotide-level phylogenetic analyses were conducted, which included only fully sequenced members of the five clade A species (Fig. 3). Phylogenetic discordance, representing potential reassortment, was evident among the phylogenetic trees built from the three different genome segments. More specifically, whilst the S and L trees exhibited nearly identical branching patterns, the M segment supported different relationships among several of the sequenced viruses. In total, there were three discrepancies between the M-segment tree and the S/L-segment trees. The two best-supported discrepancies (bootstraps ≥75.9 in all trees) involved viruses in the Oropouche species. The first involved OROV, IQtV and MDDV. OROV and IQtV were sister taxa in the S/L trees, whilst IQtV and MDDV were sister taxa in the M tree (Fig. 3). Support for the relationships of these three taxa was extremely high in all trees (≥99 bootstrap). Sliding-window analyses in RDP confirmed that these disparate patterns of divergence were consistent throughout each genomic segment, as expected with a reassortment event along one of these three lineages (Figs S5 and S6).

IQtV was recently described as a reassortant between OROV and an unknown Simbu serogroup virus based on partial sequences from the S and M segments (Aguilar et al., 2011). Our results are consistent with this finding; furthermore, we were able to identify MDDV as a potential source for the M segment of IQtV. Natural reassortment between MDDV and OROV is certainly plausible given their documented geographical distributions: both viruses have been isolated from the Madre de Dios region of Peru, and the S segment of IQtV is most similar to the S segments from clade II of OROV, which is the only clade, so far, that has been found in Peru (Aguilar et al., 2011; Saeed et al., 2000). Interestingly, the level of amino acid divergence between OROV and both IQtV and MDDV (41–42 %) was on par with levels of divergence seen between species. Whether this reassortment event has resulted in any changes in virulence, vectors or range has yet to be determined. It is also important to keep in mind that, with the current available data, it is impossible to know for certain which of these three viruses represents the true reassortant (Briese et al., 2013). The addition of more complete genome sequences from each of these viral lineages should clarify relationships.

The second potential case of reassortment involved JATV, the Utinga species (i.e. UTIV/UVV) and the lineage leading to the three human viruses (OROV, IQtV and MDDV). In the L and S trees, JATV formed a clade with the three human viruses, whereas in the M-segment tree JATV formed a well-supported clade with the Utinga species (Fig. 3).
3). The RDP analyses demonstrated that these discordant relationships were consistent throughout the entirety of each genome segment (Figs S5 and S6). Based on the high levels of sequence divergence among all three groups of viruses (>45%), this reassortment event is likely to have occurred many generations ago or to have involved parental viruses that have yet to be isolated and/or sequenced.

JATV has been reported previously to be a reassortant with an S segment from OROV and an M segment from an uncharacterized virus (Saeed et al., 2001b). Whilst our findings were generally consistent with the overall conclusion of reassortment, our S-segment sequence was not consistent with the previously reported S segment (Saeed et al., 2001b). The S segment in this previous publication fell within the OROV clade, whereas our S sequence was a phylogenetic outgroup to the S segments of OROV, IQTV and MDDV (Fig. 3). This placement demonstrated that this is an S/L- vs M-segment reassortment event. Also, based on the available data, it was again impossible to distinguish which of the viral lineages involved represents the true reassortant.

The third potential reassortment event involves Manzanilla species viruses. CQV and INGV were sister taxa in the L/S trees, whereas CQV formed a clade with MANV and MERV in the M-segment tree (Fig. 3). The bootstrap support for these different relationships, however, was lower than that seen in the other discrepancies and the RDP analyses demonstrated that the divergence signals were not consistent across the genome segments (Figs S5 and S6). Therefore, it is unclear whether this third discrepancy is due to reassortment or simply the result of ambiguity in the patterns of divergence.

**Conclusion**

The addition of 11 fully sequenced genomes for viruses in the Manzanilla and Oropouche species complexes has highlighted a deep evolutionary divide between these two species complexes and the rest of the Simbu serogroup. With sequence data from all three genome segments, we found compelling evidence to divide these two species complexes into five distinct species, and we were also able
to identify three potential reassortment events among viruses in these species. Two of these involved viruses that infect humans, and levels of sequence divergence on the reassorted segment were on a par with divergences seen between species. Future work is needed to determine whether any of these reassortments have affected virulence.

**METHODS**

**Virus isolates.** All virus stocks used in this study were obtained from the World Reference Center for Emerging Viruses and Arboviruses at the University of Texas Medical Branch (UTMB). The JATV original seed was provided by the World Health Organization Reference Centre for Arboviruses at the Department of Arbovirology and Hemorrhagic Fevers, Instituto Evandro Chagas, Brazilian Ministry of Health. Virus strain FMD1303 was originally isolated at the US Naval Medical Research Unit No. 6 (NAMRU-6) in Lima from a blood sample obtained from a febrile human in Madre de Dios Department, Peru, on 22 March 2007. The histories of the other isolates sequenced in this study have been published previously (see Table 1).

**Serological characterization.** All the sequenced viruses were compared with each other and with OROV for serological similarity. Methods used to prepare antigens for the complement fixation (CF) tests and for the preparation of immune ascitic fluids have been described previously (Beaty et al., 1989; Travassos da Rosa et al., 1983; Xu et al., 2007). Both antigens and antibodies were produced in mice. CF tests were performed by the microtitre technique (Beaty et al., 1989; Xu et al., 2007), using 2 U of guinea pig complement with overnight incubation of the antigen and antibody at 4 °C. CF titres were recorded as the highest dilutions giving 3+ or 4+ fixation of complement. Titres of 1:8 or greater were considered positive. Haemagglutination inhibition (HI) testing was performed in microtitre plates, as described previously (Travassos da Rosa et al., 1983). HI tests were performed with four haemagglutination units of virus at the optimal pH (5.75) against serial twofold antiserum dilutions, starting at 1:20. HI titres of 1:20 or greater were considered positive. CF and HI tests were performed at the UTMB, Galveston, TX, USA.

**Genome sequencing.** The BeAn 423380 (JATV) and VN 04-2108 (CQV) strains were sequenced and assembled at the Center for Infection and Immunity, Columbia University. The JATV original seed was sequenced and assembled at the Center for Technological Innovation, Genomic and Bioinformatic Cores, Evandro Chagas Institute, Brazil. For these strains, total RNA was first extracted from viral supernatant preserved in TRIzol LS and was amplified using sequence-independent single primer amplification, as described previously (Dijkeng et al., 2008). Amplicons were sheared to ~400 bp and used as starting material for Illumina TRUsq DNA libraries. Sequencing was performed on a HiSeq 2500. Primers were trimmed from the sequencing reads using Cutadapt (Martin, 2011), quality filtering was conducted with Prinseq-lite (Schmieder & Edwards, 2011) and then genomes were assembled using Ray Meta (Boisvert et al., 2012) in combination with custom scripts. When necessary, terminal sequences were completed through PCR and dideoxynucleotide sequencing using a universal orthobunyavirus primer targeting the conserved viral termini (5’-AGTAGTGRRC-3’) in combination with specific primers designed from the sequences generated from the de novo assembly. In addition, four genomes were confirmed with dideoxynucleotide sequencing (BUTV, FPV, UTIV and UVV). These included the six genome segments with the lowest levels of sequence coverage (× 30 to × 767). These sequences confirmed the high-quality of assemblies achieved through these methods.

**Phylogenetic analysis.** Separate phylogenetic analyses were conducted for each of the three genome segments using only the protein coding portions of the genome. Orthobunyavirus sequences from GenBank were included to provide a representative picture of the entire genus; many of the sequences included cover only a portion of the coding region. Sequences were aligned using the CLUSTAL algorithm, which was implemented at the amino acid level in MEGA v5.1 (Tamura et al., 2011) with additional manual editing to ensure the highest possible alignment quality. Neighbour-joining analyses using p-distance at the amino acid level were performed. The statistical significance of the tree topology was evaluated by 1000 replications of bootstrap resampling. Phylogenetic analyses were performed using MEGA v5.1 (Tamura et al., 2011).

**Reassortment analysis.** To identify potential reassortment events, the data were mined for evidence of phylogenetic discordance. For this analysis, additional phylogenetic trees were reconstructed, which included only fully sequenced members of the Oropouche and Manzanilla species complexes. These trees were reconstructed from the same alignments used above; however, to provide additional power, these trees were conducted using a maximum-likelihood framework at the nucleotide level [implemented in MEGA v5.1 (Tamura et al., 2011) with the Tamura–Nei substitution model, partial deletion, uniform rates among sites and 1000 bootstrap replications]. Potential reassortment events were then verified using the manual BOOTSCAN (Martin et al., 2005) and distance plot methods in RDP4 (Martin et al., 2010).

**Pairwise sequence analysis.** Pairwise sequence divergences were calculated among each of our 11 viruses and all of the other orthobunyaviruses with complete genome segment sequences using MEGA v5.1 with pairwise deletions (Tamura et al., 2011). For the comparisons of divergence within and between serogroups, only one representative of each named species was utilized. This downsampling was done to avoid bias due to intensive sampling of certain viruses.

**ACKNOWLEDGEMENTS**

We would like to thank Carolina Guevara and staff of the NAMRU-6 Virology Department, Lima, Peru, for providing MDDV (FMD1303) and Clayton Lima, Keley Nunes and Jedson Cardoso, Genomic and Bioinformatic Cores, Center for Technological Innovation, Evandro Chagas Institute, for sequencing and assembling the JATV original seed. The work at USAMRIID was funded by the Defense Threat Reduction Agency project no. 1881290. Work at UTMB was supported by National Institutes of Health contract HHSN27220100004OI/HHSN272200004/D04. Sequencing and assembling of the JATV original seed was supported by Evandro Chagas Institute internal grant, and
REFERENCES


Taxonomy of Viruses


