# Establishment of a minigenome system for Oropouche orthobunyavirus reveals the S genome segment to be significantly longer than previously reported

Gustavo Olszanski Acrani[1,2*], Natasha L. Tilston-Lunel[1,3*], Martin Spiegel[4], Manfred Weidmann[4, 5] Meik Dilcher[4], Daisy Elaine Andrade da Silva[6], Marcio R. T. Nunes[6] & Richard M. Elliott[1+]

*These authors contributed equally to this work.

[1]MRC-University of Glasgow Centre for Virus Research, 464 Bearsden Road, Glasgow G61 1QH, Scotland, UK.
[2]Department of Cell and Molecular Biology, University of Sao Paulo School of Medicine, 3900, Av Bandeirantes, Ribeirão Preto, SP 14049-900, Brazil
[3]Biomedical Sciences Research Complex, School of Biology, University of St Andrews, KY16 9ST, Scotland, UK.
[4]Department of Virology, University Medical Center Göttingen, Kreuzbergring 57, D-37075 Göttingen, Germany.
[5]Present address: Pathfoot Building, University of Stirling, Stirling FK9 4LA, Scotland, UK.
[6]Center for Technological Innovation, Instituto Evandro Chagas, Ananindeua, Brazil.

**Running title: Oropouche virus minigenome**

Animal - Negative-strand RNA Viruses

+ Corresponding author
email: richard.elliott@glasgow.ac.uk

Sequence accession numbers KP052850, KP052851, KP052852, KP026179, KP026180 and KP026181.

**Abstract**

Oropouche virus (OROV) is a medically important orthobunyavirus, which causes frequent outbreaks of a febrile illness in the Northern parts of Brazil. However, despite being the cause for an estimated half a million human infections since its first isolation in Trinidad, 1955, details of the molecular biology of this tripartite, negative-sense RNA virus remain limited. We have determined the complete nucleotide sequence of the Brazilian prototype strain of OROV, BeAn 19991, and found a number of differences compared to sequences in the database. Most notable were that the S segment contains an additional 204 nucleotides at the 3' end and that there is a critical nucleotide mismatch at position 9 within the base-paired terminal panhandle structure of each genome segment. In addition, we obtained the complete sequence of the Trinidadian prototype strain TRVL 9760 that showed similar characteristics to the BeAn 19991 strain. By using a T7 RNA polymerase-driven minigenome system, we demonstrated that cDNA clones of the BeAn 19991 L and S segments expressed functional proteins and also that the newly determined terminal untranslated sequences acted as functional promoters in the minigenome assay. By co-transfecting a cDNA to the viral glycoproteins, virus-like particles (VLP) were generated that packaged a minigenome and were capable of infecting naive cells.

## Introduction

Oropouche virus (OROV) is one of the most important arboviruses in Brazil, after Dengue virus and yellow fever virus, and was first isolated in 1955 from a febrile patient in Trinidad (Anderson *et al.*, 1961). Subsequently, the virus was isolated in Brazil in 1960 from the blood of a pale-throated three-toed sloth, *Bradypus tridactylus,* at a forest camp-site during construction of the Belem-Brasilia highway, just before the first documented epidemic in Brazil in 1961 (Pinheiro *et al.*, 1962). It is estimated that half a million OROV infections have occurred in more than 30 outbreaks since the virus became recognised, but it is probable that the actual numbers are much higher as cases may be masked by other febrile illnesses, such as Dengue or Mayaro fever, and diseases caused by other orthobunyaviruses such as Guama virus, that are prevalent in the region (reviewed in (Vasconcelos *et al.*, 2011). OROV has also been isolated from various mosquito species (e.g. *Coquillettidia venezuelensis, Ochlerotatus serratus*) but during epidemics, OROV is transmitted to humans by the biting midge *Culicoides paraensis* (Pinheiro *et al.*, 1981a; Pinheiro *et al.*, 1982; Pinheiro *et al.*, 1981b).

OROV belongs to the Simbu serogroup of the genus *Orthobunyavirus*, which includes a number of veterinary pathogens such as Akabane (AKAV), Aino, Shuni, Sabo and Douglas viruses, as well the newly emerged Schmallenberg virus (SBV) (Afonso *et al.*, 2014). OROV is currently the only known human pathogen in the serogroup and recent phylogenetic analysis (Ladner *et al.*, 2014b) places it in a clade separate to the other members. Like all bunyaviruses the OROV genome consists of three segments of single-stranded negative-sense RNA designated large (L), medium (M) and small (S). The L segment encodes the viral polymerase (L protein) and the M segment encodes the glycoproteins Gn and Gc, along with a non-structural protein called NSm. The S segment encodes the viral nucleocapsid protein (N) and a second non-structural protein, NSs, in over-lapping reading frames, though both proteins are translated from the same mRNA (Elliott, 2014; Plyusnin & Elliott, 2011). The terminal sequences at the 3’ and 5’ ends of each segment are complementary, allowing the formation of a

84 panhandle structure that is crucial for genome replication and transcription (Barr *et al.*,
85 2003; Barr & Wertz, 2004; Kohl *et al.*, 2004).
86
87 The epidemiology and genetic variation of OROV has been widely studied, and
88 phylogenetic analysis of numerous partial S segment sequences (mainly N ORF
89 sequences), together with more limited partial sequence data on the M and L
90 segments, suggests the existence of four genotypes (reviewed in (Vasconcelos *et al.*,
91 2011). However, much less is known about the general molecular biology of OROV or
92 virus-host interactions. To facilitate such investigations we intend to develop a reverse
93 genetics system for OROV, as has been reported for other orthobunyaviruses (Elliott,
94 2012) including two Simbu group viruses AKAV (Ogawa *et al.*, 2007) and SBV (Elliott
95 *et al.*, 2013; Varela *et al.*, 2013). When we produced cDNA clones of the OROV
96 genome segments, we noticed several discrepancies between the viral sequences we
97 obtained and the sequences in the database, notably that the S segment contains an
98 additional 204 nucleotides. The functionality of our cDNA clones was confirmed by
99 establishing minigenome (Blakqori *et al.*, 2003; Weber *et al.*, 2002) and virus-like
100 particle (VLP; (Shi *et al.*, 2007) systems. Our results highlight the importance of
101 obtaining complete and correct viral sequences, including direct confirmation of the
102 genome termini, in order to establish reverse genetic systems.
103
104 **Results**
105
106 **Cloning and sequence determination of the genome of Oropouche virus strain**
107 **BeAn 19991.** Total RNA was extracted from BHK-21 cells infected with OROV strain
108 BeAn 19991 (prototype Brazilian strain isolated from *B. tridactylus*) and reverse
109 transcribed using random primers. Segment-specific oligonucleotides, based on
110 available complete sequences in the database (L, NC_005776.1 (Aquino *et al.*, 2003);
111 M, NC_005775.1 (Wang *et al.*, 2001); and S, NC_005777.1; Aquino et al.,
112 unpublished), were used in PCR (Table 1). Full-length cDNAs were cloned into the T7
113 RNA polymerase transcription plasmid TVT7R(0,0) (Johnson *et al.*, 2000); the inserts
114 included an extra G residue at their 5' ends for efficient T7 transcription, and the
115 cDNAs were cloned such that T7 polymerase would transcribe antigenome-sense

RNAs, as previously described (Elliott *et al.*, 2013). Descriptions of the sequences in this paper are presented for the anti-genomic sense RNA, in the conventional 5’ – 3’ orientation.

*L segment.* The full-length L segment sequence that we obtained was 6852 nt in length, 6 nt longer than the database deposition NC_005776.1. Alignment of our sequence with that of NC_005776.1 revealed a number of differences in the region from nucleotide positions 2405 to 2450 and from 2592 to 2617, resulting in amino acid changes in the region from 798 to 812, and from 860 to 867 (Fig. 1). We verified the sequence of this region by RT-PCR amplification of a fragment from nt 2130 to 2980 using specific primers and viral RNA as template. Furthermore, alignment of our sequence with partial sequences of the L segments of OROV strains TRVL-9760, GML-444479 and IQT-1690 (accession numbers KC759122.1, KC759128.1 and KC759125.1 respectively) revealed that, apart from a few variations at the nucleotide level, the translated amino acid sequence for this region is conserved (Fig. 1). Therefore we consider the published sequence for BeAn 19991 L segment contains some errors in this region. In addition, we noted two other amino acid differences: L to F at position 415 and N to D at position 1021. Both of these have been confirmed by independent sequence analysis of our stock of virus, and the F residue at position 415 is also found in the L protein of other strains of OROV (TRVL-9760, GML-444479 and IQT-1690).

The terminal sequences of the L segment UTRs were determined by a 3’ RACE procedure on total infected cell RNA, using oligonucleotides designed to anneal to either the genomic or anti-genome strands. Position 9 of the 5’ UTR was determined as a C residue and the corresponding -9 position in the 3’ UTR as an A residue, resulting in the characteristic mismatch that has been observed in the predicted panhandle structure of other orthobunyavirus genome segments (Kohl *et al.*, 2004). This mismatch is not recorded in the published sequence. Additionally, position 18 at the 5’ end was determined to be a U rather than a C residue as in the published sequence (Fig. 2).

*M segment.* The full length M segment was determined to be 4385 in length, in agreement with the published sequence. There were a small number of nucleotide variations compared to the database entry (NC_005775.1), six of which resulted in amino acid differences: I274F, F587L, K614N, D750G, K981Q and G982S changes. The sequences encoding these residues were confirmed in independent cDNA clones of the M segment cDNA and also by specific RT-PCR amplification of appropriate regions of the viral RNA. Results from RACE analysis revealed two single nucleotide differences in the 5' UTR (C at position 9 and A at position 15) and one difference at the 3' end (U at position 15) compared to the database sequence. Thus the predicted panhandle has a C-A mismatch at position 9/-9 and a U-A pairing at position 15/-15 (Fig. 2)

*S segment.* The PCR reaction to amplify the S segment surprisingly generated 2 products of approximately 750 bp and 1000 bp in size (Fig. 3A). After cloning, the sequences of both products were determined. The nucleotide sequence of the smaller fragment was identical to the database entry NC_005777 (Aquino *et al*., unpublished) that is described as "Oropouche virus segment S, complete genome", but no strain designation is given. Saeed et al. (2000) reported the complete sequence of the TRVL 9760 strain of OROV also to be 754nt long, though the database entry (accession no. AF164531) only gives the coding sequence for this strain. In addition, the sequence of the N ORF of the BeAn 19991 was also reported by Saeed et al. (accession no. AF164531) and the amino acid sequence is identical to that that we obtained.

The larger fragment contained an additional 204 nucleotides after the apparent consensus 3' terminal sequence in the database entry (Fig. 3B).

The DNA products were extracted from the gel and used as templates in further PCR. The shorter template gave rise to a single, similarly sized amplicon, whereas the longer template again generated products approx 750 bp and 1000 bp in length (Fig. 3C). To investigate this observation further, we amplified the S segment of a clinical isolate of OROV (H759025 AMA2080; Tilston-Lunel *et al.,* in preparation) using the same primers and PCR conditions that were used for BeAn 19991, and again

observed two amplified DNA fragments (data not shown). The sequences of both of these amplicons largely matched that of the BeAn 19991 products (data not shown).

Inspection of the “long” sequence showed that nt 735 – 752 could allow annealing of the primer used in PCR (Fig. 3D). Thus, binding to the primer to this internal sequence in the S segment would result in a cDNA product with a terminus matching that of the orthobunyavirus consensus sequence, making it appear complete. Using 3’ RACE and RNA ligation methods we confirmed that the OROV S segment did indeed contain the additional 204 nt at the 3’ end (data not shown). Therefore the full-length OROV S segment is 958 nt in length.

The corrected sequences of the OROV strain BeAn 19991 genome have been deposited in the databases with accession numbers KP052850 (L), KP052851 (M), and KP052852 (S).

**Sequence determination of the Oropouche virus TRVL 9760 strain**. Determination of the complete sequence of another strain of OROV, the Trinidadian prototype TRVL 9760, was carried out independently from that of the BeAn 19991 strain. Total RNA was extracted from infected murine IFNAR-/- cells and reverse transcribed using random hexamer primers. Sequences comprising the L, M and N ORFs were amplified by RT-PCR using specific oligonucleotides based on the sequences available in the database (L, NC_005776.1; M, NC_005775.1; and S, NC_005777.1, as described above). While the N ORF sequence was completely amplified in one step, the L and the M ORF sequences were amplified as six (L) or three (M) overlapping fragments. The resulting cDNAs were inserted into the TA-vector pCRII, and their sequences were determined by Sanger sequencing. In comparison to the BeAn 19991 L ORF sequence NC_005776.1, the TRVL 9760 L ORF contained 151 nucleotide exchanges, 7 single nucleotide insertions and 1 single nucleotide deletion. While 134 of the 151 nucleotide exchanges were silent, the nucleotide insertions and deletions which were found from nucleotide positions 2405 to 2446 and from 2592 to 2617 lead to several amino acid exchanges and the insertion of two additional amino acids at position 799 and 810 (Figure 1). The majority of the amino acid substitutions caused by single

nucleotide exchanges were found in the N-terminal half of the L ORF (A136T, M145V, N210S, N273D, Q308K, S313N, I355V, F415L, D442N, T479A, I558M, T640A, S921N, L974I, S1021N) while only three exchanges were found in the C-terminal half (T1159I, E2056G, R2241K). When compared to the BeAn 19991 M ORF sequence NC_005775.1, the TRVL 9760 M ORF showed 100 nucleotide exchanges with 15 of them leading to amino acid substitutions (S12G, I13V, L67P, A244V, I274F, T463I, A609T, K615N, V732L, D750G, R801K, V846I, S849G, V1241I, M1363I). For the TRVL 9760 N ORF we detected 13 nucleotide exchanges in comparison to the BeAn 19991 N ORF sequence NC_005777.1, but none of these exchanges leads to an amino acid substitution. Three of these nucleotide exchanges also affect the overlapping NSs ORF and two of them lead to amino acid exchanges (K13R and N74S).

To determine the sequence of the complete L, M and S segments, pyrosequencing was performed. OROV genomic RNA isolated from supernatants of infected murine IFNAR -/- cells was converted to dsDNA by whole transcriptome amplification, which served as starting material for a shotgun library preparation. After pyrosequencing of the shotgun library *de novo* assembly with the obtained sequence reads was performed which resulted in sequences for the OROV L, M and the S ORFs identical to those obtained by Sanger sequencing. It was not, however, possible to determine the sequences of the non-coding regions by *de novo* assembly. Therefore, an additional reference mapping was performed using the OROV genomic segment sequences from the database as reference. With this approach we were able to map the obtained sequence reads to the complete L and the M segment sequences NC_005776.1 and NC_005775.1. In the case of the S segment, however, it was not possible to map the sequence reads to the 3’ end of the S segment sequence NC_005777.1 but mapping was possible for the 5’ non-coding end and the N ORF. We therefore performed another round of reference mapping using an S segment fragment comprising the 5’ end and the N ORF of NC_005777.1 as reference sequence. Using this approach the reference mapping resulted in an S segment sequence with 204 additional nucleotides at the 3’ end.

The complete sequences of the OROV strain TRVL 9760 genome segments have been deposited in the databases with accession numbers KP026179 (L), KP026180 (M) and KP026181 (S).

**Establishment of an OROV minigenome system.** Minigenome systems have been described for a number of orthobunyaviruses, and comprise a negative-sense genome analogue encoding a reporter gene that is packaged into ribonucleoprotein complex (RNP), transcribed and replicated by co-expressed viral N and L proteins, leading to measurable reporter activity (Elliott, 2012). After confirmation of the nucleotide sequences, the open reading frames (ORF) in each segment were amplified by PCR and subcloned into the pTM1 expression vector (Moss *et al.*, 1990). Minigenome constructs were created by replacing the viral ORF in each segment with the sequence for *Renilla* luciferase, and then inverting the insert in plasmid TVT7R (0,0) (Johnson *et al.*, 2000) so that T7 transcripts would be in the genomic sense (Weber *et al.*, 2001). We first used a minigenome based on the OROV M segment, as studies with BUNV showed the M segment minigenome to be the most active (Barr *et al.*, 2003). However, initial attempts using the M segment UTR sequences as reported in the database gave low activity over background. When we subsequently obtained the M segment terminal sequences by 3' RACE analysis and redesigned the minigenome accordingly, with the C-A mismatch at position 9/-9, high levels of luciferase activity were observed, indicating that, firstly, both N and L protein expressing constructs were functional and, secondly, that the M segment UTR sequences determined herein were active promoters. The amounts of transfected N- and L- expressing plasmids were titrated to determine the optimal amounts that gave maximum luciferase activity (data not shown) and the optimised amounts used in all further experiments.

The effects of nucleotide differences in M segment UTR on minigenome activity are compared in Figure 4A. The minigenome with UTR sequences as previously published (9C:G, 15C:G) showed low activity, whereas the minigenome with UTR sequences as determined in our work (9C:A, 15U:A) showed over 2,000-fold increased activity over background (cells where no L expressing plasmid was transfected). However, it was not just the mismatch at position 9/-9 that was critical for maximal activity, but also the

base-pairing at position 15/-15, as the minigenome with the position 9 C:A mismatch but C:G at position 15/-15 showed only 500-fold increase in activity. Introduction of the U:A pairing was not able to rescue activity when position 9/-9 was C:G, and other nucleotide combinations at position 15 were less active than U:A. Taken together, these results highlight the importance of certain residues within the M segment promoter.

The minigenome assay was also used to compare the short and long S segment UTR sequences (Fig. 4B). Minigenome constructs contained the same 5' UTR and either the 14nt (as previously published) or 218nt (as determined herein) long 3' UTR. The minigenome with the short UTR was inactive whereas the minigenome with the 218nt 3' UTR showed robust luciferase activity. Lastly, we compared L segment derived minigenomes, with either a C or U residue at position 18 in the 5' UTR. Both minigenomes gave similar high luciferase activity (Fig. 4C).

Together, these results confirmed that the N and L proteins were functional in a minigenome assay, and also that the UTR sequences as determined for the S, M and L segments were functional promoters, and that a base mismatch at position 9/-9 was critical for promoter activity.

**Virus-like particle production assay.** To investigate whether the glycoprotein gene was also functional, a VLP assay was developed. In addition to M segment minigenome, N and L expressing plasmids, cells were also transfected with a plasmid expressing the glycoprotein precursor. Luciferase activity was measured in these donor cells at 24 and 48 h post transfection (Fig. 5A), and it was noted that there was a significant increase in luciferase activity in cells additionally transfected with the glycoprotein cDNA at 48 h, suggesting spread of VLPs within the culture. The supernatants from transfected cells were harvested at 48 h post transfection and transferred onto naïve BHK cells; luciferase activity in these cells was measured 24 h later. High levels of luciferase activity were recorded in cells exposed to supernatants expressing the glycoproteins (column L + M in Fig 5B) compared to those exposed to supernatants from cells not transfected with the glycoprotein cDNA (column L). This

is a stringent assay relying only on transcription of the packaged minigenome in the VLP without the need for exogenously supplied viral N and L proteins. Incubation of the supernatant with antibodies to OROV before infection markedly reduced luciferase expression, whereas incubation with an irrelevant antiserum (anti-BUNV serum) had no effect (Fig. 5B). Taken together these results indicate that the OROV glycoprotein gene cDNA was functional in this VLP assay.

## Discussion

A crucial step in developing reverse genetic systems for RNA viruses is obtaining cDNA clones that are representative of the authentic viral genome sequence. As described above, we found a number of sequence differences in our clones derived from the BeAn 19991 strain compared to sequences in the database, including approx. 200 additional nucleotides at the 3' end of the S genome segment, an apparent frame shift in the L segment coding sequence and a critical mismatched nucleotide pair in the terminal panhandle sequence on each segment. These significant differences were confirmed when the complete sequence of the Trinidadian prototype strain TRVL 9760 was also determined.

Early studies comparing orthobunyavirus genome sequences indicated that the terminal 11 nucleotides of each segment exhibited a high degree of conservation, and hence consensus primers based on sequences of Bunyamwera and California serogroup viruses (Dunn *et al.*, 1994; Elliott, 1989a; b; Elliott *et al.*, 1991) have traditionally been used to amplify unknown bunyavirus genomes. However, the actual terminal sequences for the majority of sequences currently available in the database have not been verified directly, for example by RACE techniques. With regard to the orthobunyavirus "consensus sequence" there is a single nucleotide difference between the 3' and 5' complementary ends such that, using total infected cell RNA as template, mispriming by either primer could occur, or a single primer could bind to both genomic and antigenomic RNAs. Indeed, a single primer was used to amplify the OROV M segment (Aquino & Figueiredo, 2004) or the S segments of a range of orthobunyaviruses (Lambert & Lanciotti, 2008). The importance of the terminal

sequence has been investigated by minigenome assays for BUNV (Dunn *et al.*, 1995; Kohl *et al.*, 2003; Kohl *et al.*, 2004) (Barr *et al.*, 2003) (Barr & Wertz, 2004) and the mismatch at position 9/-9 was shown to be crucial for promoter activity (Barr & Wertz, 2005). As more diverse orthobunyavirus genomes have been sequenced, particularly using next generation sequencing methods (deep sequencing) that are not reliant on specific primers to amplify cDNA, it has become clear that there is more variation in the "bunyavirus consensus" than observed between Bunyamwera and California serogroup viruses (e.g.(Ladner *et al.*, 2014b)), highlighting the requirement for direct determination of the terminal sequences. In a similar vein, as the genomes of more phleboviruses (that constitute another genus in the *Bunyaviridae* family) have been sequenced, it is apparent that the termini also diverge from the "phlebovirus consensus" (Dilcher *et al.*, 2012a; Elliott & Brennan, 2014; Matsuno *et al.*, 2013).

A recent paper (Ladner *et al.*, 2014a) has suggested the standards that should be applied to viral genome sequence determination and we strongly support the recommendations proposed therein.

In 2000, Saeed and others reported the first nucleocapsid gene sequences of 28 strains of OROV, including the prototypic Trinidadian OROV isolate TRVL 9760 and the Brazilian isolate BeAn 19991 (Saeed *et al.*, 2000). They determined the complete S segment to be 754 bases and noted the unusually short length of the 3' UTR, just 14 bases after the translational stop codon, compared to other orthobunyavirus S segments. They employed various experimental procedures to verify the 3'UTR including chemical denaturation of the purified viral RNA with methylmercury hydroxide before RT-PCR (in case there was a secondary structure that impeded reverse transcription), and a 5' RACE procedure using both purified viral RNA and total cellular RNA as starting material (Saeed *et al.*, 2000). All approaches yielded that same short 3'UTR. Our results indicate that the true length of the S segment is actually 958 nt which was verified by independent experimental analyses, including deep-sequencing of the TRVL 9760 strain. Examination of the correct sequence reveals an internal region highly similar to the terminal sequence that could hybridise with the

primer, and in our studies resulted in two PCR products. The functionality of the longer 3'UTR determined in this study was demonstrated in the minigenome assay.

We further confirmed that the sequences of the BeAn 19991 N and L proteins were functional in driving reporter gene expression from minigenomes, and similarly that the determined UTR sequence for all three segments could be used to construct functional minigenomes. Lastly, by cotransfecting a cDNA that expressed the glycoprotein gene, we produced virus like particles that were capable of packaging a minigenome and to infect naïve cells. Together these data provide strong evidence that the cDNA clones reported in this paper are fully functional and pave the way to establishing a virus rescue system. The availability of such a system will play a crucial role in understanding the molecular biology of this important yet poorly characterised emerging viral zoonosis. The corrected sequences of the BeAn 19991 and TRVL9670 genome segments have been deposited in the database.

## Methods and Materials

### Cells and Virus

Vero-E6 and murine IFNAR -/- cells were grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% foetal calf serum (FCS). BHK-21 cells were grown in Glasgow minimal essential medium (GMEM, Invitrogen) supplemented with 10% newborn calf serum (NCS) and 10% tryptose phosphate broth (TPB, Invitrogen). BSR-T7/5 cells, which stably express T7 RNA polymerase (Buchholz *et al.*, 1999), were grown in GMEM supplemented with 10% FCS, 10% TPB and 1 mg/ml G418 (Geneticin; Invitrogen).

OROV strain BeAn 19991 was kindly donated by Prof. Luiz Tadeu Moraes Figueiredo, from the Ribeirao Preto School of Medicine, University of Sao Paulo, Brazil, and strain TRVL 9760 was kindly provided by Dr. Robert Shope from the University of Texas Medical Branch in Galveston, USA. A sample of total infected cell RNA obtained from the strain H759025 AMA2080 was provided by Dr. Pedro Vasconcelos, from the Department of Arboviruses and Hemorrhagic Fevers, Brazil.

All experiments with infectious viruses were conducted under CL3 laboratory conditions.

**Cloning of OROV cDNA**

OROV was grown in BHK-21 cells at 37°C and after 30 h both cells and supernatant were harvested, and RNA extracted using TRIzol reagent (Invitrogen). cDNAs to each segment were synthesised separately, using segment-specific primers for the L and M segments (OROLFg and OROMFg, Table 1), and random primers (Promega) for the S segment, together with M-MLV reverse transcriptase (Promega). Each cDNA preparation was used in a segment specific PCR using the appropriate primer pairs (OROMFg and OROMRg for the M segment, and OROSFg and OROSRg for the S segment; Table 1) and KOD Hot Start DNA polymerase (Merck), according to the manufacturer's protocol. The full-length PCR products were cloned into pGEM-T Easy (Promega). After selection of positive clones, the inserts were excised by digestion with BsmBI and ligated into BbsI-linearized plasmid TVT7R(0,0) (Johnson *et al.*, 2000). The L segment cDNA was amplified in two fragments using primer pairs (OROFLg and OROL1, and OROL2 and OROLRg; Table 1). The first primer pair amplified nt 1 to 3706, and the second pair nt 3537 to 6852, resulting in two PCR products with a 170 bp overlapping region containing a unique BsgI restriction site (position 3590 in the full length segment). PCR products were purified from an agarose gel and then cloned into pGEM-T Easy. The inserts were excised by digestion with restriction enzymes BsgI and BsmBI, and the full-length L segment was assembled by ligating both fragments with BbsI-linearized TVT7R(0,0). The cDNA inserts included an extra G residue at their 5′ ends for efficient T7 transcription, and the inserts were cloned such that T7 polymerase would transcribe antigenome-sense RNAs. The plasmids were named pTVTOROVL, pTVTOROVM and pTVTOROVS.

**Construction of protein-expressing and minigenome-expressing plasmids.**

The complete open reading frames (ORF) in the L and M segments were amplified by PCR using specific primers (pTM1 series in Table 1) and the pTVT7 transcription plasmids as templates, and subcloned into expression vector pTM1 (Moss *et al.*, 1990), under the control of the T7 promoter and encephalomyocarditis virus internal ribosome

entry site sequence (IRES). The constructs were called pTM1OROV-L and pTM1OROV-M. To generate a plasmid expressing only the N protein we introduced three point mutations (T68C, T113C and G116A) into pTVTOROVS, using primers OROdelNSsF and OROdelNSsR (Table 1), by QuikChange Site-directed Mutagenesis (Stratagene), prior to PCR amplification of the N ORF. These mutations changed the first and second methionine codons in the NSs ORF into threonine codons, and introduced an in-frame translation stop codon at codon 17; the coding sequence of the overlapping N ORF was unaffected. This plasmid was designated pTM1OROV-N.

The minigenome plasmids were created in three steps. First, the sequence encoding the coding sequence in each pTVT7 clone was deleted by excision PCR, leaving the UTRs intact. These linearised DNAs were then used in an In-Fusion reaction (In-Fusion HD Cloning, Clontech) with PCR-amplified DNA of the *Renilla* luciferase gene. The amplified luciferase gene contained 15 nt extensions homologous to the OROV L, M or S segment UTR sequences in the linearised pTVT7 construct. The UTR-luciferase-UTR sequence was then amplified by PCR using primers containing 15 nt extensions homologous to the T7 terminator (5' end) and T7 promoter (3' end). This amplified products were combined with TVT7R(0,0) DNA in an In-Fusion reaction to generate minigenome-expressing plasmids such that in T7 transcripts the *Renilla* luciferase was in the negative-sense. These constructs were designated pTVT7OROVSRen(-), pTVT7OROVMRen(-) and pTVT7OROVLRen(-).

**Sequencing OROV BeAn 19991 5' and 3' termini**

As total infected cell RNA contains both genomic and anti-genomic segments, 3' RACE analysis would be capable of generating both the 5' and 3' terminal sequences using strand specific primers. Briefly, RNA was polyadenylated (Ambion kit) for 1 hour at 37°C and then purified using the RNeasy minikit (Qiagen). The polyadenylated RNA was then used in a reverse transcription reaction with M-MLV reverse transcriptase (Promega) and oligo d(T) primer, followed by PCR using 3' PCR anchor primer (Roche) and the appropriate segment specific primer (OROVL_anti and OROVL_gen for the L segment, and OROVM_anti and OROVM_gen for the M segment; Table 1)

with KOD Hot Start DNA polymerase (Merck). Amplified products were purified on an agarose gel and their nucleotide sequence determined.

To confirm the S segment terminal sequences, total infected cell RNA was first denatured at 90°C for 3 min and then ligated using T4 RNA ligase (New England Biolabs) for 2 hours at 37°C. The reaction was heat inactivated at 65°C and purified using the RNeasy minikit (Quiagen). cDNA was synthesised using M-MVL reverse transcriptase (Promega) and oligonucleotide OROSlig1 (Table 1). PCR was then performed with KOD Hot Start DNA polymerase (Merck) and primers OROSlig1 and OROSlig2 (Table 1). The PCR product was purified on an agarose gel and its nucleotide sequence determined.

**Pyrosequencing of the OROV TRVL 9760 strain**

OROV TRVL 9760 was grown in IFNAR -/- cells at 37°C and after 48 h supernatant was harvested. (Preliminary results showed that IFNAR-/- cells gave the highest amounts of genomic RNA in the extracted supernatant compared to Vero-E6 or BHK-21 cells; unpublished observations). For removal of cell debris the supernatant was centrifuged at 700 x *g* for 10 min and at 2,800 x *g* for 5 min followed by filtration through a 0.2 µm sterile filter. To enrich viral particles 20 ml cleared supernatant was mixed with 1.48 ml 5M NaCl and 10.8 ml 30% PEG8000 in NTE (100 mM NaCl; 10 mM Tris, pH 6.5; 1 mM EDTA), incubated on a shaker for 30 min at 4 °C, and subsequently centrifuged at 6,000 x *g* for 60 min at 4 °C. The virus pellet was resuspended in 500 µl PBS. RNA extraction was performed using PeqGold Trifast (Peqlab, Erlangen Germany). To be able to cover the 3’ terminal parts of the OROV genome segments, 500 ng self-complementary FLAC adapters were ligated to 500 ng purified viral RNA as described (Dilcher *et al.*, 2012b). To achieve coverage of the 5’ terminal parts, a 5’-RACE RNA adapter (Ambion) was ligated to the viral RNA after the removal of two phosphate groups via RNA 5’-polyphosphatase. To remove unligated adapters a subsequent purification step was performed using the CleanAll DNA/RNA Clean-Up and Concentration-Kit (Norgen Biotek). The concentration of the adapter-ligated and purified ssRNA was determined by Qant-iT RiboGreen Assay (Invitrogen). 60 ng of the adapter-ligated viral RNA was amplified and converted to dsDNA using the

TransPlex Whole Transciptome Amplification kit (WTA2, Sigma-Aldrich). The newly synthesized dsDNA was purified using the QIAquick PCR Purification kit (Qiagen), and DNA fragments shorter than 350 bp were removed using Ampure-XP beads (Agencourt). 300 ng of the whole genome amplified dsDNA was used for Titanium Shotgun Rapid Library Preparation and pyrosequencing on a Genome Sequencer FLX (Roche) as described in the FLX Titanium Protocol (Roche) but omitting the DNA fragmentation by nebulization step. Assembly of the sequenced OROV genome segments was done by means of the Genome Sequencer FLX System Software Package version 2.3 (GS De Novo Assembler, GS Reference Mapper) in combination with the commercially available SeqMan Pro Software version 10.1.1 (DNASTAR, Lasergene).

**Minigenome and virus-like particle assays**

Subconfluent monolayers of BSR-T7/5 cells were transfected with 1 µg each pTM1OROV-L and pTM1OROV-N, 0.5 µg of a minigenome-expressing plasmid and 100ng pTM1-FF-Luc (Weber *et al.*, 2001). At 24 h post-transfection *Renilla* and firefly luciferase activities were measured using Dual-Luciferase Reporter Assay kit (Promega).

To generate VLPs, the M segment minigenome transfection mix was supplemented with 0.5 µg pTM1OROV-M. At 24 and 48 h post-transfection supernatants were harvested, clarified by centrifugation (4000 rpm for 5 mins at 4$^{o}$C), digested with benzonase, and used to infect BHK-21 cells. *Renilla* activity was measured after 24 h using the Renilla Reporter Assay kit (Promega). To neutralise the VLPs, samples were incubated with hyperimmune mouse ascetic fluid to OROV or with anti-BUNV rabbit antiserum for 1 h at room temperature before infecting BHK-21 cells.

**Acknowledgements**

This work was supported by a Wellcome Trust Senior Investigator Award to RME (099220), a Medical Research Council postgraduate studentship to NT-L, a FAPESP-Sao Paulo Research Foundation fellowship to GOA (2013/02798-0, CNPQ), and a

CAPES-National Council for the Improvement of Higher Education (Brasil) scholarship to DEAS (3851/10-9).

## References


**Afonso, A., Abrahantes, J. C., Conraths, F., Veldhuis, A., Elbers, A., Roberts, H., Van der Stede, Y., Meroc, E., Gache, K. & Richardson, J. (2014).** The Schmallenberg virus epidemic in Europe-2011-2013. *Prev Vet Med* **DOI: 10.1016/j.prevetmed.2014.02.012**, in press.

**Anderson, C., Spence, L., Downs, W. & Aitken, T. (1961).** Oropouche virus: a new human disease agent from Trinidad, West Indies. *Am J Trop Med Hyg* **10**, 574-578.

**Aquino, V. H. & Figueiredo, L. T. (2004).** Linear amplification followed by single primer polymerase chain reaction to amplify unknown DNA fragments: complete nucleotide sequence of Oropouche virus M RNA segment. *J Virol Methods* **115**, 51-57.

**Aquino, V. H., Moreli, M. L. & Moraes Figueiredo, L. T. (2003).** Analysis of Oropouche virus L protein amino acid sequence showed the presence of an additional conserved region that could harbour an important role for the polymerase activity. *Arch Virol* **148**, 19-28.

**Barr, J. N., Elliott, R. M., Dunn, E. F. & Wertz, G. W. (2003).** Segment-specific terminal sequences of Bunyamwera bunyavirus regulate genome replication. *Virology* **311**, 326-338.

**Barr, J. N. & Wertz, G. W. (2004).** Bunyamwera bunyavirus RNA synthesis requires cooperation of 3'- and 5'-terminal sequences. *J Virol* **78**, 1129-1138.

**Barr, J. N. & Wertz, G. W. (2005).** Role of the conserved nucleotide mismatch within 3'- and 5'-terminal regions of Bunyamwera virus in signaling transcription. *J Virol* **79**, 3586-3594.

**Blakqori, G., Kochs, G., Haller, O. & Weber, F. (2003).** Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids. *J Gen Virol* **84**, 1207-1214.

**Buchholz, U. J., Finke, S. & Conzelmann, K. K. (1999).** Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. *J Virol* **73**, 251-259.

**Dilcher, M., Alves, M. J., Finkeisen, D., Hufert, F. & Weidmann, M. (2012a).** Genetic characterization of Bhanja virus and Palma virus, two tick-borne phleboviruses. *Virus Genes* **45**, 311-315.

**Dilcher, M., Hasib, L., Lechner, M., Wieseke, N., Middendorf, M., Marz, M., Koch, A., Spiegel, M., Dobler, G., Hufert, F. T. & Weidmann, M. (2012b).** Genetic characterization of Tribec virus and Kemerovo virus, two tick-transmitted human-pathogenic Orbiviruses. *Virology* **423**, 68-76.

**Dunn, E. F., Pritlove, D. C. & Elliott, R. M. (1994).** The S RNA genome segments of Batai, Cache Valley, Guaroa, Kairi, Lumbo, Main Drain and Northway bunyaviruses: sequence determination and analysis. *J Gen Virol* **75** 597-608.

**Dunn, E. F., Pritlove, D. C., Jin, H. & Elliott, R. M. (1995).** Transcription of a recombinant bunyavirus RNA template by transiently expressed bunyavirus proteins. *Virology* **211**, 133-143.

**Elliott, R. M. (1989a).** Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae. *Virology* **173**, 426-436.

**Elliott, R. M. (1989b).** Nucleotide sequence analysis of the small (S) RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae. *J Gen Virol* **70**, 1281-1285.

**Elliott, R. M. (2012).** Bunyavirus reverse genetics and applications to studying interactions with host cells. In *Reverse Genetics of RNA Viruses: Applications and Perspectives*, pp. 200-223. Edited by A. Bridgen: John Wiley & Sons, Ltd.

**Elliott, R. M. (2014).** Orthobunyaviurses: recent genetic and structural insights. *Nature Reviews Microbiology* **12**, 673-685.

**Elliott, R. M., Blakqori, G., van Knippenberg, I. C., Koudriakova, E., Li, P., McLees, A., Shi, X. & Szemiel, A. M. (2013).** Establishment of a reverse genetics system for Schmallenberg virus, a newly emerged orthobunyavirus in Europe. *J Gen Virol* **94**, 851-859.

**Elliott, R. M. & Brennan, B. (2014).** Emerging phleboviruses. *Curr Opin Virol* **5**, 50-57.

**Elliott, R. M., Schmaljohn, C. S. & Collett, M. S. (1991).** Bunyaviridae genome structure and gene expression. *Curr Top Microbiol Immunol* **169**, 91-141.

**Johnson, K. N., Zeddam, J. L. & Ball, L. A. (2000).** Characterization and construction of functional cDNA clones of Pariacoto virus, the first Alphanodavirus isolated outside Australasia. *J Virol* **74**, 5123-5132.

**Kohl, A., Bridgen, A., Dunn, E., Barr, J. N. & Elliott, R. M. (2003).** Effects of a point mutation in the 3' end of the S genome segment of naturally occurring and engineered Bunyamwera viruses. *J Gen Virol* **84**, 789-793.

**Kohl, A., Dunn, E. F., Lowen, A. C. & Elliott, R. M. (2004).** Complementarity, sequence and structural elements within the 3' and 5' non-coding regions of the Bunyamwera orthobunyavirus S segment determine promoter strength. *J Gen Virol* **85**, 3269-3278.

**Ladner, J. T., Beitzel, B., Chain, P. S., Davenport, M. G., Donaldson, E. F., Frieman, M., Kugelman, J. R., Kuhn, J. H., O'Rear, J., Sabeti, P. C., Wentworth, D. E., Wiley, M. R., Yu, G. Y., Sozhamannan, S., Bradburne, C. & Palacios, G. (2014a).** Standards for sequencing viral genomes in the era of high-throughput sequencing. *mBio* **5**, e01360-01314.

**Ladner, J. T., Savji, N., Lofts, L., Travassos da Rosa, A., Wiley, M. R., Gestole, M. C., Rosen, G. E., Guzman, H., Vasconcelos, P. F., Nunes, M. R., Lipkin, W. I., Tesh, R. & Palacios, G. F. (2014b).** Genomic and phylogenetic characterization of viruses included in the Manzanilla and Oropouche species complexes of the genus Orthobunyavirus, family Bunyaviridae. *J Gen Virol* **95**, 1055-1066.

**Lambert, A. J. & Lanciotti, R. S. (2008).** Molecular characterization of medically important viruses of the genus Orthobunyavirus. *J Gen Virol* **89**, 2580-2585.

**Matsuno, K., Weisend, C., Travassos da Rosa, A. P., Anzick, S. L., Dahlstrom, E., Porcella, S. F., Dorward, D. W., Yu, X. J., Tesh, R. B. & Ebihara, H. (2013).** Characterization of the Bhanja serogroup viruses (Bunyaviridae): a novel species of the genus Phlebovirus and its relationship with other emerging tick-borne phleboviruses. *J Virol* **87**, 3719-3728.

**Moss, B., Elroy-Stein, O., Mizukami, T., Alexander, W. A. & Fuerst, T. R. (1990).** New mammalian expression vectors. *Nature* **348**, 91-92.

**Ogawa, Y., Sugiura, K., Kato, K., Tohya, Y. & Akashi, H. (2007).** Rescue of Akabane virus (family Bunyaviridae) entirely from cloned cDNAs by using RNA polymerase I. *J Gen Virol* **88**, 3385-3390.

**Pinheiro, F. P., Hoch, A. L., Gomes, M. L. & Roberts, D. R. (1981a).** Oropouche virus. IV. Laboratory transmission by Culicoides paraensis. *Am J Trop Med Hyg* **30**, 172-176.

**Pinheiro, F. P., Pinheiro, M., Bensabath, G., Causey, O. R. & Shope, R. E. (1962).** Epidemia de virus Oropouche em Belem. . *Revista do Servico Especial de Saude Publica* **12**, 15-23.

**Pinheiro, F. P., Travassos da Rosa, A. P., Gomes, M. L., LeDuc, J. W. & Hoch, A. L. (1982).** Transmission of Oropouche virus from man to hamster by the midge Culicoides paraensis. *Science* **215**, 1251-1253.

**Pinheiro, F. P., Travassos da Rosa, A. P., Travassos da Rosa, J. F., Ishak, R., Freitas, R. B., Gomes, M. L., LeDuc, J. W. & Oliva, O. F. (1981b).** Oropouche virus. I. A review of clinical, epidemiological, and ecological findings. *Am J Trop Med Hyg* **30**, 149-160.

**Plyusnin, A. & Elliott, R. M., (eds) (2011).** *Bunyaviridae. Molecular and Cellular Biology*. Norfolk: Caister Academic Press.

**Saeed, M. F., Wang, H., Nunes, M., Vasconcelos, P. F., Weaver, S. C., Shope, R. E., Watts, D. M., Tesh, R. B. & Barrett, A. D. (2000).** Nucleotide sequences and phylogeny of the nucleocapsid gene of Oropouche virus. *J Gen Virol* **81**, 743-748.

**Shi, X., Kohl, A., Li, P. & Elliott, R. M. (2007).** Role of the cytoplasmic tail domains of Bunyamwera orthobunyavirus glycoproteins Gn and Gc in virus assembly and morphogenesis. *J Virol* **81**, 10151-10160.

**Varela, M., Schnettler, E., Caporale, M., Murgia, C., Barry, G., McFarlane, M., McGregor, E., Piras, I. M., Shaw, A., Lamm, C., Janowicz, A., Beer, M., Glass, M., Herder, V., Hahn, K., Baumgartner, W., Kohl, A. & Palmarini, M. (2013).** Schmallenberg virus pathogenesis, tropism and interaction with the innate immune system of the host. *PLoS Pathog* **9**, e1003133.

**Vasconcelos, H. B., Nunes, M. R., Casseb, L. M., Carvalho, V. L., Pinto da Silva, E. V., Silva, M., Casseb, S. M. & Vasconcelos, P. F. (2011).** Molecular epidemiology of Oropouche virus, Brazil. *Emerg Infect Dis* **17**, 800-806.

**Wang, H., Beasley, D. W., Li, L., Holbrook, M. R. & Barrett, A. D. (2001).** Nucleotide sequence and deduced amino acid sequence of the medium RNA segment of Oropouche, a Simbu serogroup virus: comparison with the middle RNA of Bunyamwera and California serogroup viruses. *Virus Res* **73**, 153-162.

**Weber, F., Bridgen, A., Fazakerley, J. K., Streitenfeld, H., Randall, R. E. & Elliott, R. M. (2002).** Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon. *J Virol* **76**, 7949-7955.

**Weber, F., Dunn, E. F., Bridgen, A. & Elliott, R. M. (2001).** The Bunyamwera virus nonstructural protein NSs inhibits viral RNA synthesis in a minireplicon system. *Virology* **281**, 67-74.

## Figure Legends

**Figure 1.** Alignment of part of the OROV L segment highlighting the differences between the published sequence for BeAn 19991 strain (accession number NC_005776) and the sequence obtained in this study (new data), along with three published OROV sequences from different genotypes TRVL-9760 (KC759122.1), GML-444479 (KC759128.1) and IQT-1690 (KC759125.1). The nucleotide alignment is shown in the top panel and the amino acid alignment in the bottom panel. Alignments were performed using CLC Genomics Workbench 6.5.

**Figure 2.** Comparison of the published and the revised OROV UTR sequences shown as a panhandle structure. The terminal 11 conserved residues are separated by a vertical line. Differences are highlighted in red.

**Figure 3.** Analysis of the OROV S segment. (A) Agarose gel electrophoresis of the S segment RT-PCR product. (B) Schematic drawing of OROV S segment, comparing the published sequence of 754bp (upper drawing) to the newly determined 958bp sequence (lower drawing). Black boxes represents the N ORF, grey boxes the NSs ORF, and hatched boxes the UTRs. The sequence is presented in the anti-genomic 5' to 3' sense. Numbers indicate the nucleotide position in the sequence. (C) Agarose gel electrophoresis of reamplified DNA products using the 754bp and the 958bp PCR products as template. (D) Diagram showing the potential internal binding site (bold) in the OROV S segment. Numbers represent the nucleotide position. OROSRg Primer: primer sequence that was used in this paper to amplify the S segment.

**Figure 4.** Minigenome assay. (A) Comparison of M segment based minigenomes. BSR-T7/5 cells were transfected with 1 µg each pTM1OROV-L and pTM1OROV-N, 0.5 µg of M segment minigenome-expressing plasmid and 100 ng pTM1-FF-Luc; the background control lacked pTM1OROV-L. M segment minigenomes contained different nucleotides at position 9/-9 as indicated. Minigenome activity is expressed as fold induction over the background control. (B) Comparison of S segment minigenomes containing the published (14 nt) or newly-defined long (218 nt) 5' UTR.

(C). Comparison of L segment minigenomes containing a C or U residue at position 18 in the 3' UTR.

**Figure 5. Virus-like particle production assay.** BSR-T7/5 cells were transfected with 1 µg each pTM1OROV-L and pTM1OROV-N, 0.5 µg pTM1OROV-M, 0.5 µg of the M segment minigenome-expressing plasmid and 100ng pTM1-FF-Luc; control transfection mixes lacked pTM1OROV-L (No L) or pTM1OROV-M (+L). At 24 or 48 h post transfection, clarified supernatants were used to infect naive BHK-21 cells, and luciferase activity measured 24 h later. (A) Minigenome activity in transfected BSR-T7/5 cells at 24 or 48 h post transfection. (B) Minigenome activity in BHK-21 cells infected with supernatants from cells in graph A. VLPs were also incubated with anti-OROV antibodies (+ve) or irrelevant antibodies (-ve) before infection of cells as indicated.