

# Isolation of a novel species of flavivirus and a new strain of *Culex flavivirus* (*Flaviviridae*) from a natural mosquito population in Uganda

Shelley Cook,<sup>1</sup> Gregory Moureau,<sup>2</sup> Ralph E. Harbach,<sup>1</sup> Louis Mukwaya,<sup>3</sup> Kim Goodger,<sup>1</sup> Fred Ssenfuka,<sup>3</sup> Ernest Gould,<sup>2,4</sup> Edward C. Holmes<sup>5</sup> and Xavier de Lamballerie<sup>2</sup>

## Correspondence

Shelley Cook  
s.cook@nhm.ac.uk

<sup>1</sup>Natural History Museum, Cromwell Road, London SW7 5BD, UK

<sup>2</sup>Unité des Virus Emergents UMR190 'Emergence des Pathologies Virales', Université de la Méditerranée et Institut de Recherche pour le Développement, Marseille, France

<sup>3</sup>Mosquito Research Programme, Uganda Virus Research Institute, PO Box 49, Entebbe, Uganda

<sup>4</sup>Centre for Ecology and Hydrology Oxford, Mansfield Road, Oxford OX1 3SR, UK

<sup>5</sup>Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA

The genus *Flavivirus*, which contains approximately 70 single-stranded, positive-sense RNA viruses, represents a unique model for studying the evolution of vector-borne disease, as it includes viruses that are mosquito-borne, tick-borne or have no known vector. Both theoretical work and field studies suggest the existence of a large number of undiscovered flaviviruses. Recently, the first isolation of cell fusing agent virus (CFAV) was reported from a natural mosquito population in Puerto Rico, and sequences related to CFAV have been discovered in mosquitoes from Thailand. CFAV had previously been isolated from a mosquito cell line in 1975 and represented the only known 'insect-only' flavivirus, appearing to replicate in insect cells alone. A second member of the 'insect-only' group, Kamiti River virus (KRV), was isolated from Kenyan mosquitoes in 2003. A third tentative member of the 'insect-only' group, *Culex flavivirus* (CxFV), was first isolated in 2007 from Japan and further strains have subsequently been reported from the Americas. We report the discovery, isolation and characterization of two novel 'insect-only' flaviviruses from Entebbe, Uganda: a novel lineage tentatively designated Nakiwogo virus (NAKV) and a new strain of CxFV. The individual mosquitoes from which these strains were isolated, identified retrospectively by using a reference molecular phylogeny generated using voucher specimens from the region, were *Mansonia africana nigerrima* and *Culex quinquefasciatus*, respectively. This represents the first isolation, to our knowledge, of a novel insect-only flavivirus from a *Mansonia* species and the first isolation of a strain of CxFV from Africa.

Received 8 June 2009

Accepted 30 July 2009

## INTRODUCTION

The genus *Flavivirus* represents a unique model for studying the evolution of vector-borne disease because it includes viruses that are: (i) arthropod-borne, infecting a range of vertebrate hosts through mosquito or tick bites;

The GenBank/EMBL/DDBJ accession numbers for the COI gene sequences from mosquito species obtained during this study are GQ165759–GQ165807, that for *Culex flavivirus* Uganda is GQ165808, that for Nakiwogo virus is GQ165809 and that for cell fusing agent virus strain Rio Piedras is GQ165810.

Three supplementary tables listing primer sequences used, flaviviral abbreviation codes and accession numbers, and mosquito accession numbers and Barcode of Life Datasystem (BOLD) reference numbers are available with the online version of this paper.

(ii) presumed to be limited to vertebrates alone; and (iii) appear to be restricted to insects alone ('insect-only' flaviviruses). In addition, the genus includes pathogens of significant importance to human health, including dengue virus and yellow fever virus (YFV). Cell fusing agent virus (CFAV), first isolated from a *Stegomyia aegypti* (as *Aedes aegypti*) cell line [Stollar & Thomas, 1975; nomenclature for the aedine mosquitoes follows Reinert (2000) and Reinert *et al.* (2004, 2006, 2008)] and characterized by Cammisa-Parks *et al.* (1992), previously represented the only known 'insect-only' flavivirus, appearing to replicate uniquely in insect cells (Kuno, 2007). In recent years, Cook *et al.* (2006) reported the first isolation of CFAV, from a natural mosquito population in Puerto Rico. The virus was found in several different species and in mosquitoes of both sexes.

Two strains of a second member of the 'insect-only' group, designated Kamiti River virus (KRV), were isolated from *Neomelanimonion mcintoshii* mosquitoes from Kenya [Crabtree *et al.*, 2003; Sang *et al.*, 2003; these authors stated that KRV was isolated from '*Ae. macintoshii*' from flooded dambos in Central Province, Kenya. However, *Aedes (Ochlerotatus) macintoshii* (Marks) [*Ochlerotatus macintoshii* in the classification of Reinert *et al.* (2008)] is an Australian mosquito, and the species in this case is actually the Rift Valley fever vector *Neomelanimonion mcintoshii* (Huang) (originally named *Aedes mcintoshii*), whose larvae are frequently found in dambos}. Taken together with theoretical work that suggested the existence of over 2000 unknown mosquito-borne flaviviruses (Pybus *et al.*, 2002), as well as the discovery of DNA sequences related to flaviviruses in the genomes of *St. aegypti* and *Stegomyia albopicta*, which probably resulted from integration events following infection of each mosquito species by a virus (or viruses) related to the CFAV group, it was proposed that further 'insect-only' flaviviruses were likely to be discovered in natural mosquito populations (Crochu *et al.*, 2004). Indeed, sequences related to CFAV been discovered have more recently in *St. aegypti* mosquitoes from Thailand (Kihara *et al.*, 2007). A third member of the 'insect-only' group, *Culex flavivirus* (CxFV), was first isolated in 2007 from *Culex tritaeniorhynchus* and *Culex quinquefasciatus* in Japan and Indonesia (Hoshino *et al.*, 2007). Further strains have subsequently been isolated from *Cx. quinquefasciatus* in Guatemala (Morales-Betoulle *et al.*, 2008) and Mexico (Farfan-Ale *et al.*, 2009), and from *Cx. quinquefasciatus* and *Culex restuans* in Texas, USA, and Trinidad (Kim *et al.*, 2009). Finally, flavivirus RNA has also recently been discovered in phlebotomine sandflies from Algeria (Moureau *et al.*, 2009).

We now report the discovery, isolation and characterization of two novel flaviviruses of the 'insect-only' group from the Entebbe area of Uganda: a novel lineage, tentatively designated Nakiwogo virus (NAKV), and a new strain of CxFV. The individual mosquitoes from which these strains were isolated, identified retrospectively from a backbone phylogeny generated using the 'barcode' region of voucher specimens from the area, were specimens of *Mansonia africana nigerrima* and *Cx. quinquefasciatus*, respectively. To date, insect-only flaviviruses have been isolated from *Culex*, *Stegomyia* and *Ochlerotatus* mosquitoes, in common with the majority of the mosquito-borne flaviviruses. This work represents, to our knowledge, the first isolation of a novel insect-only flavivirus from a *Mansonia* species and the first isolation of a strain of CxFV from Africa. In addition, we include in our analyses, for the first time, the full genome sequence from CFAV Rio Piedras, isolated previously from Puerto Rico (Cook *et al.*, 2006).

## METHODS

**Trapping protocol.** In total, 419 individual mosquitoes were collected in February 2008 for virus screening. In addition, 204

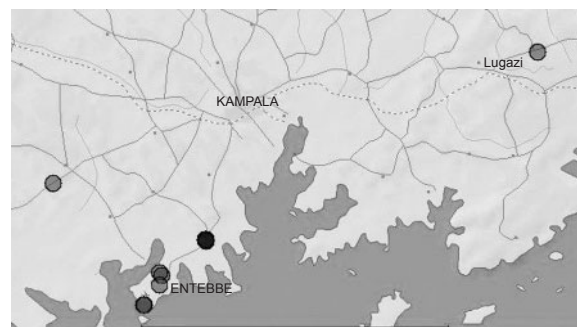
mosquitoes representing species found at sampling locations were retained as voucher specimens for the development of molecular identification protocols. Mosquitoes were sampled from a variety of locations (Fig. 1), using different methods to maximize species diversity. Centers for Disease Control (CDC) fan-augmented light traps were supplemented with dry ice for approximately 8 h trapping periods between dusk and dawn and placed at various heights in urban, rural and forest environments. Modified backpack aspirators and hand-held aspirators were used to sample resting mosquitoes. Specimens were placed on dry ice immediately upon collection at the sampling sites. Upon arrival at the laboratory, samples were sorted on a chill table according to trap, location and sex and placed in individual wells in microtitre plates before storage at  $-80^{\circ}\text{C}$ .

**Sample preparation, nucleic acid extraction and RT-PCR screening.** Whole mosquitoes were homogenized individually as described previously (Cook *et al.*, 2006). Samples of 65  $\mu\text{l}$  from five individual mosquitoes were pooled for RNA extraction. Total nucleic acid (NA) extraction was conducted via the MagNAPure LC System using a Culture Cells kit and standard manufacturer's protocol (Roche Diagnostics), or using a Biorobot EZ1, with Virus Mini kit v2.0 (Qiagen).

A 5  $\mu\text{l}$  aliquot of each pooled RNA extraction was used in a one-step real-time QuantiTect SYBR-Green RT-PCR assay, according to Moureau *et al.* (2007), with primers PF1S and PF2R-bis (see Supplementary Table S1, available in JGV Online). Second-round PCR on 2.5  $\mu\text{l}$  cDNA, using primers PF3S and PF2R-bis, used the following conditions: 94  $^{\circ}\text{C}$  for 2 min; 40 cycles of 94  $^{\circ}\text{C}$  for 30 s, 50  $^{\circ}\text{C}$  for 45 s and 72  $^{\circ}\text{C}$  for 1 min; followed by a final extension of 72  $^{\circ}\text{C}$  for 7 min and cooling to 20  $^{\circ}\text{C}$  for 2 min. To avoid any potential contamination, positive cDNA samples from first-round RT-PCR were manipulated separately from all other samples for second-round PCR. Negative controls comprising both NA extractions from non-infected mosquitoes and reactions in which water replaced RNA were performed through all stages. Positive controls comprised Rio Bravo, Japanese encephalitis and Montana myotis leukoencephalitis viruses (RBV, JEV and MMLV, respectively).

Second-round PCR products were analysed visually by electrophoresis through ethidium bromide-stained 2% agarose gels under UV light. Products were purified by using a QIAquick Spin PCR Purification kit (Qiagen). Amplicons were sequenced on both strands with an ABI 377 automated sequencer (Applied Biosystems).

For those samples that gave positive results using the primer set PF1S-2Rbis-3S on RNA extractions from pooled mosquitoes, NA were re-extracted from the original individual mosquito homogenates by using a Biorobot EZ1, with Virus Mini kit v2.0 according to the manufacturer's instructions (Qiagen). One-step SYBR green first-



**Fig. 1.** Mosquito collection sites in Uganda. Bar, 50 km.

round RT-PCR was followed by second-round PCR using the PF1–2bis–3 system as described for the pooled samples. Second-round PCR products were analysed visually, purified and sequenced as described above.

#### Further investigation of individual samples and virus isolation.

Individual mosquito homogenates were investigated further based on (i) congruent positive sequences in both pooled sample and individual mosquitos and (ii) the resulting 157 nt NS5 sequence clustering with other known ‘insect-only’ flaviviruses.

For the individual mosquito homogenates chosen for virus isolation work, 100 µl homogenate was centrifuged at 13 000 r.p.m. (16 100 g) for 20 min. Subsequently, the supernatant was diluted in a 1 ml final volume of L15 medium without fetal bovine serum (FBS), but enriched with antibiotics (ml<sup>-1</sup>: 100 IU penicillin G, 100 mg streptomycin, 100 mg kanamycin, 2.5 µg amphotericin B) and inoculated onto C6/36 cells in single 12.5 cm<sup>2</sup> flasks. After incubation at room temperature for 1 h, 4 ml fresh 3% FBS L15 medium was added. The flask was incubated at 28 °C. A negative-control flask was handled under identical conditions. Flasks were examined daily for the presence of cytopathic effect [i.e. cell fusion and/or syncytium formation similar to that seen in the prototype CFAV, as described by Stollar & Thomas (1975)] and 400 µl of each supernatant medium was extracted and tested by real-time RT-PCR. A series of four passages on C6/36 cells was carried out for each positive isolate.

The strategy for sequencing of the entire open reading frame (ORF) followed Grard *et al.* (2007). Initial RT-PCR amplifications targeted regions encoding the E, NS3 and NS5 proteins using degenerate primers (Crochu *et al.*, 2004; Gaunt & Gould, 2005; Gaunt *et al.*, 2001; Moureau *et al.*, 2007). Specific primers were then designed to complete gaps via long-range RT-PCR (cMaster RT<sub>plus</sub> PCR system; Eppendorf) and amplicons were sequenced using the long PCR product sequencing (LoPPS) method, a shotgun-based approach applied to long PCR amplification products (Emonet *et al.*, 2006, 2007). Primer sequences are available from the corresponding author upon request.

**Transmission electron microscopy.** Tissue-culture samples showing cytopathic effect were prepared for electron-microscopic examination. Negative-stained electron-microscopic specimens were prepared by drying culture supernatant medium, mixed 1:1 with 2.5% paraformaldehyde, onto Formvar/carbon-coated grids and staining with 2% methylamine tungstate.

**Investigation for the presence of virus-specific DNA.** It has been previously shown that DNA forms of CFAV are produced during infection of a number of different mosquito cell types (Cook *et al.*, 2006). Therefore, for flaviviral isolates, total NA were extracted from C6/36 cultures of each virus at passage 5, day 6 post-infection by using a Biorobot EZ-1. Both supernatant medium and sedimented cells were tested using the primer pairs detailed in Supplementary Table S1. First, sedimented cells and supernatant medium were tested by using standard PCR with no RT step. Second, the same primer pairs were used for a one-step RT-PCR (Promega). Third, NA from sedimented cells were treated with recombinant DNase I according to the manufacturer’s instructions (Roche). These DNase-treated samples were then tested in a classic PCR with no RT step.

**Phylogenetic analyses of viral sequences.** Sequencher v4.8 (Gene Codes) was used to combine reverse and forward viral sequences. Sequences were then compared with those of all other members of the genus *Flavivirus* available to date. Supplementary Table S2 (available in JGV Online) lists the flaviviral sequences that were used in analyses. Datasets were prepared by using Se-Al (available at <http://tree.bio.ed.ac.uk/software/seal/>) for (i) nucleotide data for the region encoding the NS5 protein, (ii) amino acid alignment of the complete ORF and

(iii) nucleotide and amino acid sequences for the region encoding the E protein for the insect-only flaviviruses, as this currently comprises the greatest number of species available for this group. All nucleotide analyses were repeated both including and excluding third-codon positions (as these were likely to experience site saturation).

For the NS5 region, the dataset was primarily limited to those flaviviruses for which ORF data were also available, to allow comparison between phylogenies. Nucleotides were aligned by using MUSCLE (Edgar, 2004) with manual adjustment to maintain correct reading frame. The model of nucleotide substitution and parameter values were selected via MODELTEST (Posada & Crandall, 1998) and used to estimate maximum-likelihood (ML) phylogenetic trees in PAUP (Swofford, 2000). Values for the substitution matrix, base composition, gamma distribution of among-site rate variation ( $\Gamma$ ) and the proportion of invariant sites (I) are available from the authors on request. A bootstrap-resampling analysis was conducted using 1000 replicate neighbour-joining (NJ) trees based on the ML substitution matrix. Phylogenetic analyses were also performed using the Bayesian method available in MrBayes v3.1.2 (Huelsenbeck & Ronquist, 2001) with a minimum of 10 million generations and a burn-in of 10%. Stationarity was assessed at effective sample size >400 using Tracer v1.4.1 (part of the BEAST package; Drummond & Rambaut, 2007). For the full ORF dataset, amino acids were aligned by using MUSCLE and phylogenetic analyses were performed in MrBayes, with a minimum of 10 million generations and a burn-in of 10%. For the E region, nucleotides were aligned based on alignment of amino acid sequences in MUSCLE. ML phylogenetic analyses conducted by using MODELTEST and PAUP were identical to those carried out for the NS5 region. In addition, Bayesian analyses were conducted on the amino acid dataset in MrBayes as described above, after removing 17 sites of ambiguous alignment using GBlocks (Talavera & Castresana, 2007). Final datasets comprised 60 sequences (2514 sites) for the NS5 region (corresponding to nucleotide positions 7790–10270 of YFV; GenBank accession no. X03700), 58 sequences (3801 sites) for the ORF and 27 sequences (417 sites) for the E region (corresponding to amino acid positions 1184–1531 of YFV; GenBank accession no. U17067).

#### Extraction, amplification and sequencing of voucher mosquito DNA.

Voucher mosquito specimens had previously been identified using morphology (Edwards, 1941; Service, 1990). A 130 µl volume of Chelex 100 Resin 15% (w/v) (Bio-Rad) was added to 20 µl mosquito homogenate. This was incubated at 95 °C for 20 min and centrifuged at 13 000 r.p.m. (16 100 g) for 15 min. Further analysis was conducted using 50–60 µl supernatant and, of this extraction, 1.0–2.5 µl was used in PCRs. DNA was also extracted, amplified and sequenced from those individual mosquito homogenates found to be positive for novel flavivirus strains.

The COI gene was amplified by using primers UEA3 (5'-TATRGCWTTYCCWCGAATAAATAA-3') (Lunt *et al.*, 1996) and Fly10 (5'-ASTGCACTAATCTGCCATATTAG-3') (Sallum *et al.*, 2002) according to Cook *et al.* (2006) (hereafter referred to as the ‘Fly’ region). An additional overlapping region, the ‘barcode’ section of the COI gene, was amplified using the primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAATCA-3') [both from Folmer *et al.* (1994)] with a primer concentration of 10 µM and reaction conditions of 5 min at 95 °C; 40 cycles of 30 s at 95 °C, 30 s at 48 °C and 45 s at 72 °C; followed by a final extension time of 5 min at 72 °C. COI sequences obtained during this work have been deposited in GenBank with accession numbers GQ165759–GQ165807.

**Phylogenetic analysis of mosquito sequences.** Sequencher v4.8 was used to combine reverse and forward sequences from each mosquito and final datasets were compiled using Se-Al. These

sequences were added to those currently available from GenBank and the Barcode of Life Datasystem (BOLD) (Ratnasingham & Hebert, 2007). For all voucher specimens, COI sequences from a minimum of five field-collected individuals per species were included in analyses where possible. In the case of sequences from public databases, five sequences per species (preferably from different geographical locations) were again included where available. Sequences from flavivirus-positive mosquito samples were also included. Datasets for each region comprised 1152 bp from 112 individuals and 636 bp from 338 individuals for the 'Fly' and 'barcode' regions, respectively. Analyses were conducted separately for each region, due to the fact that taxonomic coverage of mosquito species in public databases varies for each. Following confirmation that all species groups of five were monophyletic, analyses were then repeated using single representatives per species for ease of presentation, except in the case of clades including flavivirus-positive mosquitoes, in which case sequences from multiple vouchers or public data for that species were included if possible. These 'single representative per species' datasets comprised 138 and 46 individuals, respectively, for the 'barcode' and 'Fly' regions.

In each case, nucleotide sequences were aligned by using MUSCLE (Edgar, 2004) with manual adjustment to maintain correct reading frame, the model of nucleotide substitution and parameter values were selected via MODELTEST (Posada & Crandall, 1998) and used to estimate ML phylogenetic trees in PAUP (Swofford, 2000). Bootstrap-resampling analyses were conducted by using 1000 replicate NJ trees based on the ML substitution matrix. For sample H3E6 in the barcode region, the ML phylogeny of *Culex* and *Culiseta* sequences only was also estimated for ease of presentation, and is shown in Fig. 6 (see Supplementary Table S3, available in JGV Online). Trees were rooted by using *Chagasia* (Fly region) or *Culiseta* (barcode subset).

## RESULTS

### Flavivirus screening

Positive and negative controls carried through all protocols showed that the system was both sensitive to the presence of flaviviruses and free from contamination. From 419 mosquitoes assayed using semi-nested RT-PCR with the PF1S–2R–3S system, four pools were PCR-positive, with the resultant sequences clustering with known 'insect-only' flaviviruses.

Upon further testing of the 20 individual samples that contributed to the four positive pools, three individual mosquito homogenates gave congruent positive sequences. No congruent sequences were amplified from individual mosquitoes contributing to the fourth positive pool and virus isolation from both the pool and the relevant homogenates was unsuccessful. Of the remaining three potential flavivirus strains, two individual sequences were identical and originated from mosquitoes in the same trap. Isolation was attempted using only one of these two homogenates. Therefore, in total, two flavivirus strains were taken forward for further characterization and isolation work. Nevertheless, retrospective identification was conducted for all three individual mosquitoes for which flavivirus-positive results were obtained via RT-PCR.

### Virus isolation and characterization

The first novel strain (H4A1) was isolated from a female mosquito, UG134-26, collected in a CDC light trap supplemented with dry ice on 24 February 2008 near Entebbe (latitude 0.07555867, longitude 32.4574212; Fig. 1). The second strain (H3E6) was isolated from a female mosquito, UG134-14, from the same trap. Notably, the third virus-positive sample, H4D1, which was identical at the nucleotide level to H4A1 for the PF3S/PF2R-bis amplicon, was isolated from another female mosquito, UG134-26, also captured in the same trap.

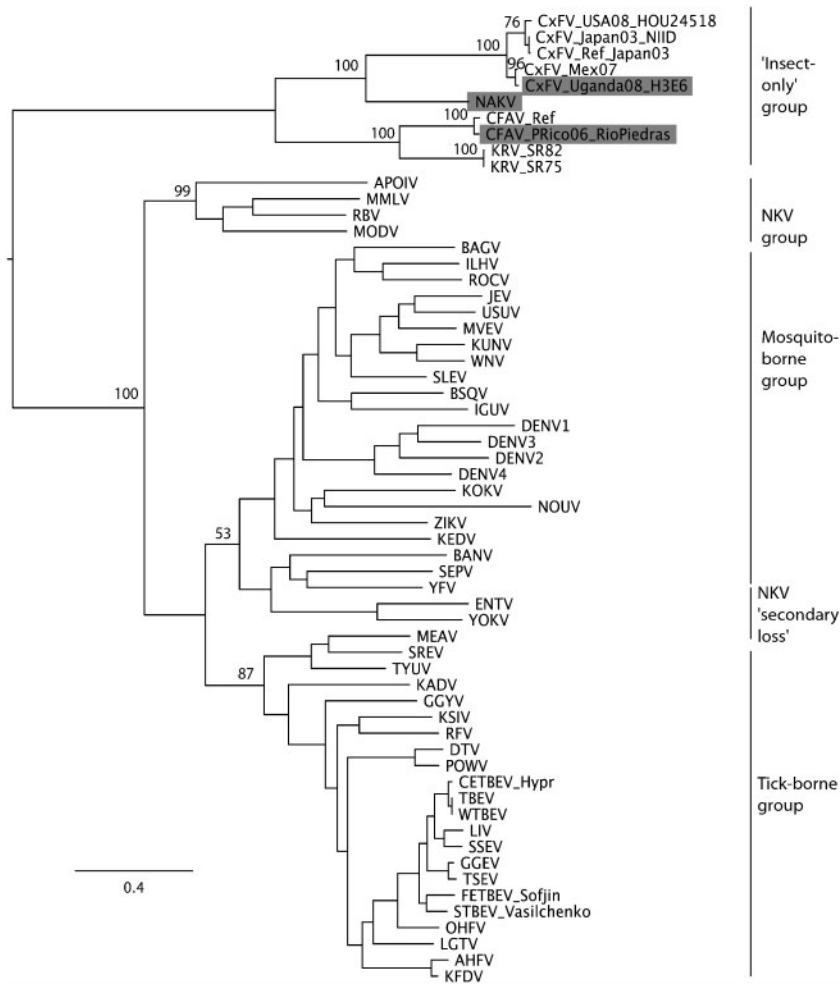
The ORFs of isolate H3E6, designated CxFV Uganda, and isolate H4A1, tentatively designated Nakiwogo virus (NAKV, named after the location of the trap), are 10 092 and 10 122 nt long, respectively (GenBank accession numbers GQ165808 and GQ165809). The topology of the insect-only group, as well as its position relative to the other members of the genus, was similar in both the NS5 and ORF phylogenies (Figs 2 and 3, respectively). CxFV Uganda appears to be related most closely to CxFV from Mexico, with an uncorrected p-distance of 9.7% in NS5. NAKV appears to form a sister group to the CxFV group, with an uncorrected p-distance between NAKV and CxFV Uganda of 40.5%. In contrast, in the E region phylogeny shown in Fig. 4, CFAV is related more closely to CxFV than is NAKV.

As shown in Fig. 5, flavivirus-like particles were visible in infected C6/36 culture for both CxFV Uganda and NAKV. In common with other members of the genus, the spherical virions are enveloped with clear projections and approximately 40 nm in diameter. Cytopathic effect in C6/36 cells infected with NAKV was moderate and similar to that seen in the prototype CFAV (Stollar & Thomas, 1975), with the formation of large syncytia. In contrast, CxFV Uganda was associated with a reduction in cell density and modification of cell shape from circular to triangular.

Results from the investigation of DNA forms are shown in Table 1. For both H4A1 and H3E6, results were consistent with RNA forms of each flavivirus being present in both culture supernatant and pelleted cells. DNA forms were also present, but in sedimented cells only.

### Phylogeny and identification of flavivirus-positive mosquitoes

Alignment of COI sequences from voucher and flavivirus-positive mosquitoes was unambiguous. In all cases, groups of five individuals from a given species from both the field and public databases were monophyletic, indicating, in common with other studies, that this gene is suitable for species identification (Hebert *et al.*, 2003). Further analyses, using single representatives from each species except in the case of species related to the flavivirus-positive mosquitoes (in which case, multiple representatives are included where possible), are shown in Fig. 6. Sample H3E6, from which CxFV Uganda was isolated,



**Fig. 2.** Maximum-likelihood (ML) phylogenetic tree of the novel CxFV Uganda and NAKV strains and other members of the genus *Flavivirus* for the NS5 nucleotide dataset. Bootstrap values are shown for the main clades only. See Supplementary Table S2 for virus abbreviations and GenBank accession numbers. Identical topologies were obtained using both PAUP (including and excluding third-codon positions) and MrBayes for the insect-only flaviviruses. All horizontal branch lengths are drawn to scale; bar, 0.4 substitutions per site. The tree is midpoint-rooted for purposes of clarity only. Those flavivirus sequences that are published for the first time in the current study, namely NAKV, CxFV strain Uganda and CFAV strain Rio Piedras, are highlighted in grey.

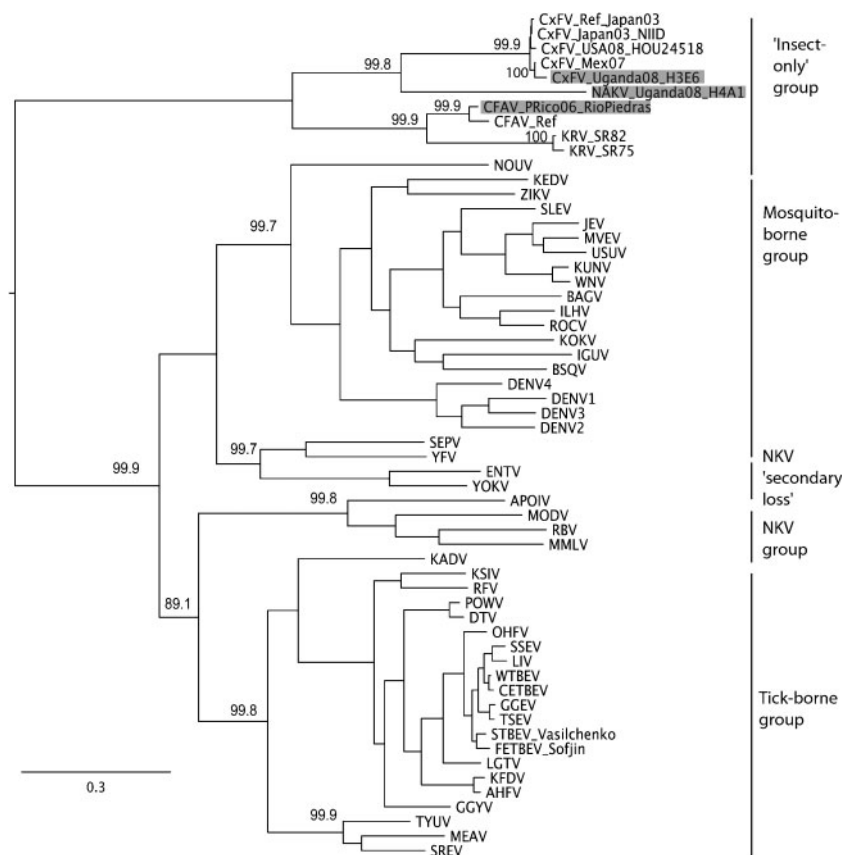
grouped with sequences from field-collected *Cx. quinquefasciatus* voucher specimens in both phylogenies. Mosquitoes comprising samples H4A1 and H4D1 grouped with a field-collected voucher specimen of *Ma. africana nigerrima* for the 'Fly' region. Primers LCO1490 and HCO2198 did not amplify the barcode region for the mosquitoes that comprised samples H4A1 and H4D1 and, for ease of presentation, we show the *Culex* subset phylogeny, including sample H3E6, for this region of the COI gene.

## DISCUSSION

We have described the first isolation, to our knowledge, of a novel flavivirus, NAKV, from a field-collected *Mansonia* mosquito. We have also isolated a novel strain of the flavivirus CxFV, designated here CxFV Uganda, from adult *Cx. quinquefasciatus* mosquitoes. Kuno *et al.* (1998) suggested that pairwise sequence identity of >84% correlates well with among-species antigenic characteristics for the flaviviruses. As such, and considering the >30% sequence divergence of NAKV from all other members of the group, it is possible that NAKV comprises a novel

species within the genus, although all such definitions are essentially arbitrary. The approximately 10% difference between the reference CxFV ('Japan03') and the Ugandan strain is comparable to that seen between other strains of CxFV.

Both the NS5 and ORF ML analyses indicate that CxFV Uganda is related most closely to CxFV from Mexico, with strong support for this clade. The larger E region dataset suggests that CxFV isolates from the Caribbean and Guatemala also fall into this group. A second CxFV lineage contains isolates from Japan and North America, which fall as sister groups. Previous studies had also identified these two main clades, namely Trinidad/Guatemala and Asia/Texas, as potential subtypes or genotypes of CxFV (Kim *et al.*, 2009). The addition of CxFV Uganda from Africa, which falls into the first of these groups, supports the hypothesis that CxFV may have been introduced multiple times into the Americas, (i) from Africa to the Caribbean, potentially via routes similar to the introduction of other mosquito-borne flaviviruses such as YFV (Gould *et al.*, 2003) and (ii) from Asia to North America (Lounibos, 2002). As such, the distribution of CxFV may reflect that of the *Culex pipiens* complex, which is cosmopolitan and



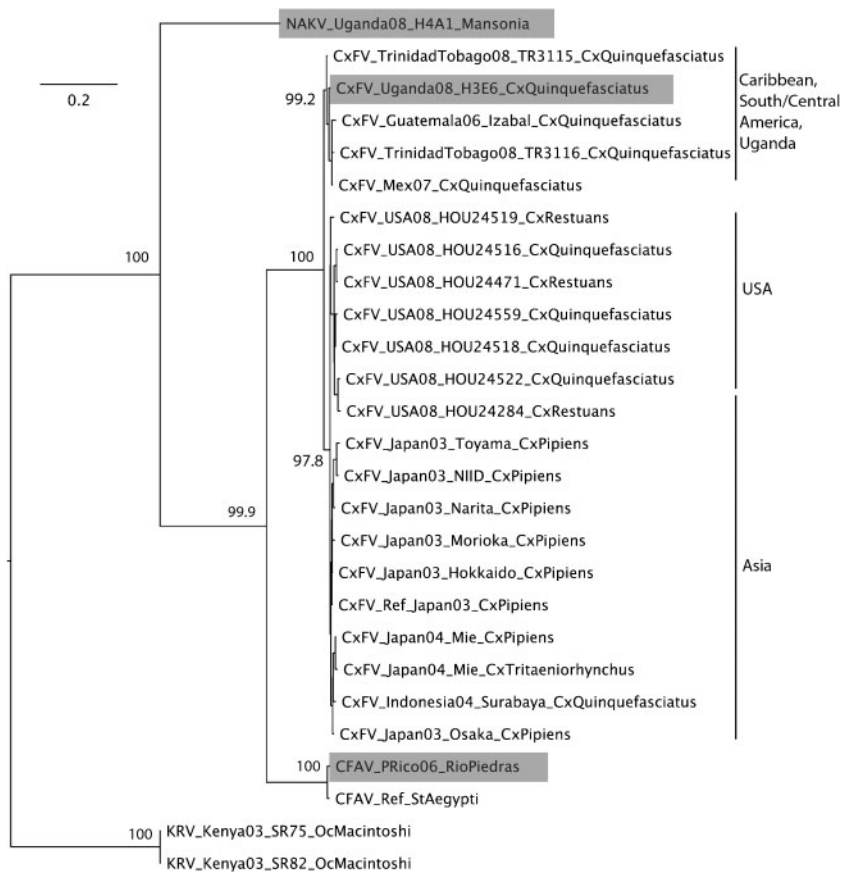
**Fig. 3.** Bayesian phylogeny of the ORF amino acid dataset of the genus *Flavivirus*, including novel strain CxV Uganda and NAKV. Posterior probabilities (%) are shown for the main clades only. See Supplementary Table S2 for virus abbreviations and GenBank accession numbers. All horizontal branch lengths are drawn to scale; bar, 0.3 substitutions per site. The tree is midpoint-rooted for purposes of clarity only. Those flavivirus sequences that are published for the first time in the current study, namely NAKV, CxV strain Uganda and CFAV strain Rio Piedras, are highlighted in grey.

ubiquitous, with multiple introductions of this mosquito species into the New World. Notably, although members of the *Cx. pipiens* complex differ in physiology, ecology and behaviour, sibling species are very difficult to distinguish from one another via morphological methods. Until molecular protocols for the accurate identification of *Culex* species across the range of the complex are fully developed, there is the potential for misidentification of the vectors of CxFV and, as such, it is difficult to draw conclusions regarding the relative prevalence and distribution of CxFV in *Cx. pipiens* and *Cx. quinquefasciatus*. As shown in Fig. 6, although the *Cx. quinquefasciatus* clade is monophyletic and the molecular identification of the mosquito species from which CxFV Uganda was isolated is robust, resolution of the *Cx. pipiens* plus *Cx. quinquefasciatus* clade is not high and the development of a molecular backbone phylogeny for these species in Uganda using an additional gene, for example ITS2 (Severini *et al.*, 1996), is recommended for future work.

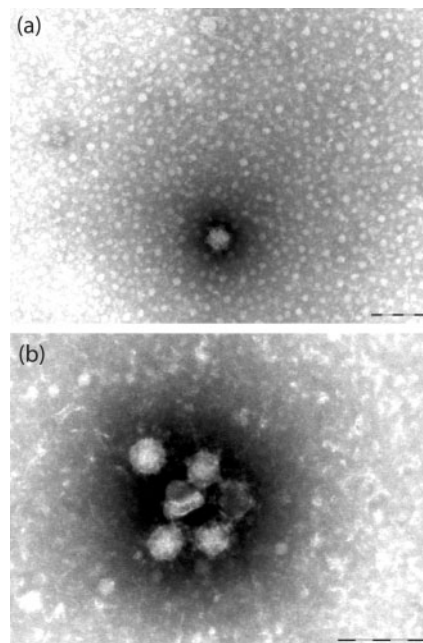
NAKV tentatively forms a novel species of flavivirus that appears to be related most closely to the CxFV group in the NS5 and ORF phylogenies. CFAV and KRV then form a sister group to the NAKV/CxFV lineage. In contrast, in the E region phylogeny, CFAV is related more closely to CxFV than NAKV. Although this is compatible with potential interspecies recombination, interpretation of the E region tree is not unambiguous, in particular the position of the

root, because the high genetic distances involved mean that other members of the genus *Flavivirus* were not incorporated into this analysis. It is therefore clear that additional data, particularly comprising ORFs for all available insect-only flaviviruses, would be required in future to test this more fully via formal tests.

Notably, both CxV Uganda and the two samples of NAKV (i.e. homogenates H4A1 and H4D1, which were both positive via RT-PCR, with identical sequences) were isolated from the same trap, but from different mosquito species, namely *Cx. quinquefasciatus* and *Ma. africana nigerrima*, respectively. Both of these mosquito species were trapped in other locations and are present across Uganda; hence, although sampling was not designed to investigate virus distribution, this suggests that the presence of insect-only flaviviruses is patchy and potentially linked to a common source of infection. This would be in agreement with the hypothesis that insect-only flaviviruses can be transmitted vertically (Cook *et al.*, 2006). Indeed, evidence exists that West Nile virus can be transmitted vertically by *Culex* and *Stegomyia* mosquitoes (as *Aedes albopictus* and *Ae. aegypti*) (Baqar *et al.*, 1993). Although CxFV Uganda and NAKV were both isolated from female mosquitoes, these were collected by using CDC light traps supplemented with dry ice, therefore biasing sampling towards host-seeking females. Hence, the possibility that the viruses may also have been present in



**Fig. 4.** Bayesian phylogeny of the E region insect-only amino acid dataset from the genus *Flavivirus*, including CxFV and NAKV. Country and year of isolation, strain and mosquito species are shown for each virus. Posterior probabilities (%) are shown for the main clades only. See Supplementary Table S2 for virus abbreviations and GenBank accession numbers. The topology of the E region nucleotide phylogeny, both including and excluding third-codon positions, was identical to that obtained by using amino acid data. All horizontal branch lengths are drawn to scale; bar, 0.2 substitutions per site. The tree is rooted by using KRV for the purpose of clarity only (midpoint rooting produces a similar topology). Those flavivirus sequences that are published for the first time in the current study, namely NAKV, CxFV strain Uganda and CFAV strain Rio Piedras, are highlighted in grey.

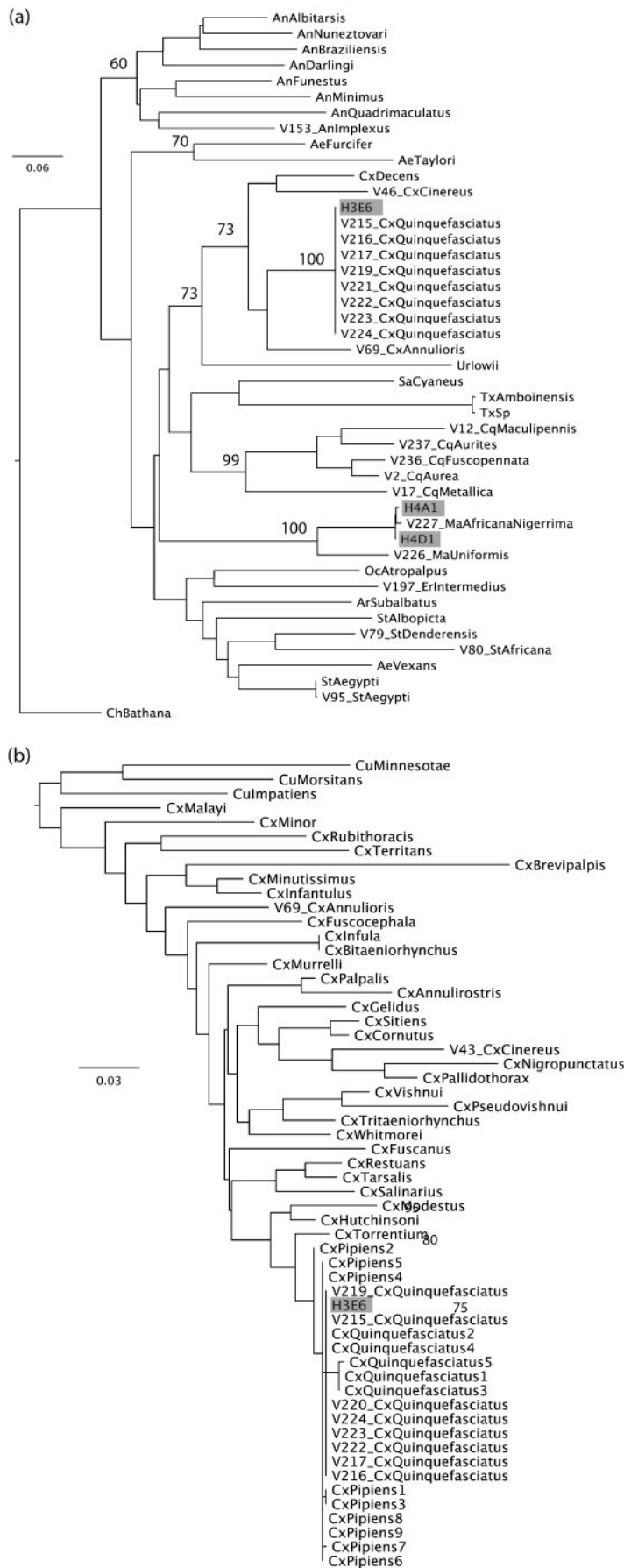


**Fig. 5.** Electron micrographs of C6/36 cells infected with (a) CxFV Uganda and (b) NAKV. Bars, 100 nm.

male mosquitoes, as described for CFAV in Puerto Rico (Cook *et al.*, 2006), cannot be discounted.

The development of a molecular ‘backbone’ phylogeny for the mosquitoes present in a given region of study, as conducted here, is essential for the screening of potential viral vectors, because many aedine and *Culex* species can only be identified by dissection of male genitalia. Without the development of a molecular system, particularly in species-rich tropical sampling areas, identification of entire adults via hand lens or microscope over a chill block in the field prior to viral screening may result in misidentification. This study has shown that primers LCO1490 and HCO2198 did not amplify the barcode region for *Ma. africana nigerrima*. Barcode primers have, to date, primarily been used for potential vectors of malaria, i.e. species of *Anopheles*. Future work will aim to tailor the system to all taxa of the family Culicidae.

Overall, considering that (i) many undiscovered insect-only flaviviruses are likely to exist in nature, (ii) flavivirus RNA has recently been discovered in phlebotomine sandflies (Moureau *et al.*, 2009) and (iii) the transmission dynamics of the insect-only flaviviruses remain unclear, but may involve vertical transmission, it is clear that this group requires further investigation. Whilst there is no evidence that they cause disease in humans, the high prevalence of the flaviviruses in general raises a number of



**Fig. 6.** ML phylogenetic tree of mosquitoes plus virus-positive samples from Uganda for the COI gene. Bootstrap values for main clades of >50% only are shown for clarity. All horizontal branch lengths are drawn to scale; bars, 0.06 (a) or 0.03 (b) substitutions per site. V, Field-collected voucher specimen; H3E6, mosquito from which CxFV Uganda was isolated; H4A1 and H4D1, mosquitoes from which NAKV was isolated. All other sequences are from public databases as detailed in Supplementary Table S3. (a) 'Fly' region of COI, rooted using *Chagasia*. (b) 'Barcode' region of COI, *Culex* subset, rooted using *Culiseta*.



**Table 1.** Results of investigation of DNA forms in NAKV and CxJV C6/36 cell culture

Sample	Primers	No RT classic PCR		RT-PCR one-step		No RT classic PCR	
		Supernatant	Cells	Supernatant	Cells	Cells +DNase	Cells no DNase
H4A1	NS3 verif S/R	–	+	+	+	–	+
	NS5 verifB S/R	–	+	+	+	–	+
H3E6	ENV verif S1/R1	–	+	+	+	–	+
	NS3 verif S/R	–	+	+	+	–	+
	NS5 verifB S/R	–	+	+	+	–	+

issues for further consideration. First, future work should aim to investigate the effect of insect-only flaviviruses on different mosquito cell lines and on all life stages of vector species to clarify potential mechanisms of infection and vertical transmission. Second, the possibility of competitive exclusion amongst co-circulating flaviviruses within the vector population, including the effect of infection by insect-only flaviviruses on vector competence for other flaviviruses, also requires investigation via experimental-infection studies. Finally, given that insect-only flaviviruses appear to infect both mosquitoes (Culicidae) and phlebotomine sandflies (Psychodidae), it is advisable to screen potential vector species of both families from any given locality via an integrated approach.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr Jonathan Kayondo (Uganda Virus Research Institute, Entebbe, Uganda) for his help in fieldwork and mosquito collection and Professor Alan Clements for helpful comments on the manuscript. The work of S. C. is funded by a Sir Henry Wellcome Postdoctoral Fellowship from the Wellcome Trust (grant 039833).

## REFERENCES

- Baqar, S., Hayes, C. G., Murphy, J. R. & Watts, D. M. (1993). Vertical transmission of West Nile virus by *Culex* and *Aedes* species mosquitoes. *Am J Trop Med Hyg* **48**, 757–762.
- Cammisa-Parks, H., Cisar, L. A., Kane, A. & Stollar, V. (1992). The complete nucleotide sequence of cell fusing agent (CFA): homology between the nonstructural proteins encoded by CFA and the nonstructural proteins encoded by arthropod-borne flaviviruses. *Virology* **189**, 511–524.
- Cook, S., Bennett, S. N., Holmes, E. C., De Chesse, R., Moureau, G. & de Lamballerie, X. (2006). Isolation of a new strain of the flavivirus cell fusing agent virus in a natural mosquito population from Puerto Rico. *J Gen Virol* **87**, 735–748.
- Crabtree, M. B., Sang, R. C., Stollar, V., Dunster, L. M. & Miller, B. R. (2003). Genetic and phenotypic characterization of the newly described insect flavivirus, Kamiti River virus. *Arch Virol* **148**, 1095–1118.
- Crochu, S., Cook, S., Attoui, H., Charrel, R., De Cheese, R., Belhouchet, M., Lemasson, J.-J., de Micco, P. & de Lamballerie, X. (2004). Sequence of flavivirus-related RNA viruses persist in DNA form in the genome of *Aedes* spp. mosquitoes. *J Gen Virol* **85**, 1971–1980.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Edwards, F. W. (1941). *Mosquitoes of the Ethiopian Region III – Culicine Adults and Pupae*. London: British Museum.

Emonet, S., Grard, G., Brisbarre, N., Moureau, G., Temmam, S., Charrel, R. & de Lamballerie, X. (2006). LoPPS: a long PCR product sequencing method for rapid characterisation of long amplicons. *Biochem Biophys Res Commun* **344**, 1080–1085.

Emonet, S. F., Grard, G., Brisbarre, N. M., Moureau, G. N., Temmam, S., Charrel, R. N. & de Lamballerie, X. (2007). Long PCR product sequencing (LoPPS): a shotgun-based approach to sequence long PCR products. *Nat Protoc* **2**, 340–346.

Farfan-Ale, J. A., Loroño-Pino, M. A., Garcia-Rejon, J. E., Hovav, E., Powers, A. M., Lin, M., Dorman, K. S., Platt, K. B., Bartholomay, L. C. & other authors (2009). Detection of RNA from a novel West Nile-like virus and high prevalence of an insect-specific flavivirus in mosquitoes in the Yucatan Peninsula of Mexico. *Am J Trop Med Hyg* **80**, 85–95.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3**, 294–297.

Gaunt, M. W. & Gould, E. (2005). Rapid subgroup identification of the flaviviruses using degenerate primer E gene RT-PCR and site specific restriction enzyme analysis. *J Virol Methods* **128**, 113–117.

Gaunt, M. W., Sall, A. A., de Lamballerie, X., Falconar, A. K. I., Dzihivanian, T. I. & Gould, E. A. (2001). Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J Gen Virol* **82**, 1867–1876.

Gould, E. A., de Lamballerie, X., Zanotto, P. M. & Holmes, E. C. (2003). Origins, evolution, and vector/host coadaptations within the genus *Flavivirus*. *Adv Virus Res* **59**, 277–314.

Grard, G., Moureau, G., Charrel, R. N., Lemasson, J.-J., Gonzalez, J.-P., Gallian, P., Gritsun, T., Holmes, E. C., Gould, E. A. & de Lamballerie, X. (2007). Genetic characterisation of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. *Virology* **361**, 80–92.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**, 313–321.

Hoshino, K., Isawa, H., Tsuda, Y., Kazuhiko, Y., Sasaki, T., Yuda, M., Takasaki, T., Kobayashi, M. & Sawabe, K. (2007). Genetic characterization of a new insect flavivirus isolated from *Culex pipiens* mosquito in Japan. *Virology* **359**, 405–414.

Huelsenbeck, J. P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755.

- Kihara, Y., Satho, T., Eshita, Y., Sakai, K., Kotaki, A., Takasaki, T., Rongsriyam, Y., Komalamisra, N., Srisawat, R. & other authors (2007). Rapid determination of viral RNA sequences in mosquitoes collected in the field. *J Virol Methods* **146**, 372–374.
- Kim, D. Y., Guzman, H., Bueno, R., Jr, Dennett, J. A., Auguste, A. J., Carrington, C. V. F., Popov, V. L., Weaver, S. C., Beasley, D. W. C. & Tesh, R. B. (2009). Characterization of *Culex flavivirus* (Flaviviridae) strains isolated from mosquitoes in the United States and Trinidad. *Virology* **386**, 154–159.
- Kuno, G. (2007). Host range specificity of flaviviruses: correlation with *in vitro* replication. *J Med Entomol* **44**, 93–101.
- Kuno, G., Chang, G.-J. J., Tsuchiya, K. R., Karabatsos, N. & Cropp, C. B. (1998). Phylogeny of the genus *Flavivirus*. *J Virol* **72**, 73–83.
- Lounibos, L. P. (2002). Invasions by insect vectors of human disease. *Annu Rev Entomol* **47**, 233–266.
- Lunt, D. H., Zhang, D.-X., Szymura, J. M. & Hewitt, G. M. (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Mol Biol* **5**, 153–165.
- Morales-Betoulle, M. E., Monzón Pineda, M. L., Sosa, S. M., Panella, N., López, M. R., Cordón-Rosales, C., Komar, N., Powers, A. & Johnson, B. W. (2008). *Culex flavivirus* isolates from mosquitoes in Guatemala. *J Med Entomol* **45**, 1187–1190.
- Moureaux, G., Temmam, S., Gonzalez, J. P., Charrel, R. N., Grard, G. & de Lamballerie, X. (2007). A real-time RT-PCR method for the universal detection and identification of flaviviruses. *Vector Borne Zoonotic Dis* **7**, 467–478.
- Moureaux, G., Ninove, L., Izri, A., Cook, S., de Lamballerie, X. & Charrel, R. N. (2009). Flavivirus RNA in phlebotomine sandflies. *Vector Borne Zoonotic Dis* (in press). doi:10.1089/vbz.2008.0216
- Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.
- Pybus, O. G., Rambaut, A., Holmes, E. & Harvey, P. H. (2002). New inferences from tree shape: numbers of missing taxa and population growth rates. *Syst Biol* **51**, 881–888.
- Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: the Barcode of Life Data System. *Mol Ecol Notes* **7**, 355–364.
- Reinert, J. F. (2000). New classification for the composite genus *Aedes* (Diptera: Culicidae: Aedini), elevation of subgenus *Ochlerotatus* to generic rank, reclassification of the other subgenera, and notes on certain subgenera and species. *J Am Mosq Control Assoc* **16**, 175–188.
- Reinert, J. F., Harbach, R. E. & Kitching, I. J. (2004). Phylogeny and classification of *Aedini* (Diptera: Culicidae), based on morphological characters of all life stages. *Zool J Linn Soc* **142**, 289–368.
- Reinert, J. F., Harbach, R. E. & Kitching, I. J. (2006). Phylogeny and classification of *Finlaya* and allied taxa (Diptera: Culicidae: Aedini) based on morphological data from all life stages. *Zool J Linn Soc* **148**, 1–101.
- Reinert, J. F., Harbach, R. E. & Kitching, I. J. (2008). Phylogeny and classification of *Ochlerotatus* and allied taxa (Diptera: Culicidae: Aedini) based on morphological data from all life stages. *Zool J Linn Soc* **153**, 29–114.
- Sallum, M. A. M., Schultz, T. R., Foster, P. G., Aronstein, K., Wirtz, R. A. & Wilkerson, R. C. (2002). Phylogeny of *Anophelinae* (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences. *Syst Entomol* **27**, 361–382.
- Sang, R. C., Gichogo, A., Gachoya, J., Dunster, M. D., Ofula, V., Hunt, A. R., Crabtree, M. B., Miller, B. R. & Dunster, L. M. (2003). Isolation of a new flavivirus related to cell fusing agent virus (CFAV) from field-collected flood-water *Aedes* mosquitoes sampled from a dambo in central Kenya. *Arch Virol* **148**, 1085–1093.
- Service, M. W. (1990). *Handbook to the Afrotropical Toxorhynchitine and Culicine Mosquitoes, excepting Aedes and Culex*. London: British Museum.
- Severini, C., Silvestrini, F., Mancini, P., La Rosa, G. & Marinucci, M. (1996). Sequence and secondary structure of the rDNA second internal transcribed spacer in the sibling species *Culex pipiens* L. and *Cx. quinquefasciatus* Say (Diptera: Culicidae). *Insect Mol Biol* **5**, 181–186.
- Stollar, V. & Thomas, V. (1975). An agent in the *Aedes aegypti* cell line (Peleg) which causes fusion of *Aedes albopictus* cells. *Virology* **64**, 367–377.
- Swofford, D. L. (2000). PAUP\*: phylogenetic analysis using parsimony (\*and other methods), v4. Sunderland, MA: Sinauer Associates.
- Talavera, G. & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56**, 564–577.