Ecuador Paraiso Escondido virus, a new flavivirus isolated from New World sand flies in Ecuador, is the first representative of a novel clade in the genus flavivirus

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Ecuador Paraiso Escondido Virus, a New Flavivirus Isolated from New World Sand Flies in Ecuador, Is the First Representative of a Novel Clade in the Genus *Flavivirus*

Cigdem Alkan,[a,b] Sonia Zapata,[c] Laurence Bichaud,[a,b] Grégory Moureau,[a,b] Philippe Lemey,[d] (ORCID) Andrew E. Firth,[e] Tamara S. Gritsun,[f] Ernest A. Gould,[a,b] Xavier de Lamballerie,[a,b] Jérôme Depaquit,[g] Rémi N. Charrel[a,b]

Aix Marseille Université, IRD French Institute of Research for Development, EHESP French School of Public Health, EPV UMR_D 190 Emergence des Pathologies Virales, Marseille, France[a]; IHU Mediterranée Infection, APHM Public Hospitals of Marseille, Marseille, France[b]; Instituto de Microbiología, Universidad San Francisco de Quito, Quito, Ecuador[c]; Department of Microbiology and Immunology, Rega Institute, KU Leuven, Leuven, Belgium[d]; Division of Virology, Department of Pathology, University of Cambridge, Cambridge, United Kingdom[e]; School of Biological Sciences, University of Reading, Whiteknights, Reading, United Kingdom[f]; Université de Reims Champagne-Ardenne, ANSES, EA 4688 USC VECPAR Transmission Vectorielle et Épidémiosurveillance de Maladies Parasitaires, SFR Capsanté, Reims, France[g]

ABSTRACT

A new flavivirus, Ecuador Paraiso Escondido virus (EPEV), named after the village where it was discovered, was isolated from sand flies (*Psathyromyia abonnenci*, formerly *Lutzomyia abonnenci*) that are unique to the New World. This represents the first sand fly-borne flavivirus identified in the New World. EPEV exhibited a typical flavivirus genome organization. Nevertheless, the maximum pairwise amino acid sequence identity with currently recognized flaviviruses was 52.8%. Phylogenetic analysis of the complete coding sequence showed that EPEV represents a distinct clade which diverged from a lineage that was ancestral to the nonvectored flaviviruses Entebbe bat virus, Yokose virus, and Sokoluk virus and also the *Aedes*-associated mosquito-borne flaviviruses, which include yellow fever virus, Sepik virus, Saboya virus, and others. EPEV replicated in C6/36 mosquito cells, yielding high infectious titers, but failed to reproduce either in vertebrate cell lines (Vero, BHK, SW13, and XTC cells) or in suckling mouse brains. This surprising result, which appears to eliminate an association with vertebrate hosts in the life cycle of EPEV, is discussed in the context of the evolutionary origins of EPEV in the New World.

IMPORTANCE

The flaviviruses are rarely (if ever) vectored by sand fly species, at least in the Old World. We have identified the first representative of a sand fly-associated flavivirus, Ecuador Paraiso Escondido virus (EPEV), in the New World. EPEV constitutes a novel clade according to current knowledge of the flaviviruses. Phylogenetic analysis of the virus genome showed that EPEV roots the *Aedes*-associated mosquito-borne flaviviruses, including yellow fever virus. In light of this new discovery, the New World origin of EPEV is discussed together with that of the other flaviviruses.

The genus *Flavivirus* currently includes 86 viruses, of which 73 are grouped into 53 species ([1](#)). More than 40 of these flaviviruses are known to be pathogenic for humans and other vertebrates, in which they cause a variety of clinical diseases from mild febrile illness to severe encephalitis and/or hemorrhagic fever. Flaviviruses are small enveloped viruses with positive-sense single-stranded RNA genomes consisting of an open reading frame (ORF) flanked by 5′ and 3′ noncoding regions (NCR). The polyprotein encodes three structural proteins (the capsid [C], membrane [M], and envelope [E] proteins) and seven nonstructural (NS) proteins (NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5) ([2, 3](#)).

Flaviviruses have extensive geographic distributions and diverse arthropod vectors, and many of them infect vertebrate hosts ([4](#)). Among the arthropod-borne flaviviruses there is a correlation between phylogenetic relationships and virus-vector-host interactions ([5–8](#)). On the basis of virus neutralization studies and, separately, the association of arthropod vectors with vertebrates, 4 major groups of flaviviruses are recognized: the tick-borne flaviviruses (TBFVs), the mosquito-borne flaviviruses (MBFVs), no-known-vector flaviviruses (NKVs), and no-known-vertebrate-host flaviviruses ([5, 6, 9, 10](#)).

The mosquito- and tick-borne borne flaviviruses contain important animal and human pathogens, including yellow fever virus (YFV), dengue virus (DENV), West Nile virus (WNV), St. Louis encephalitis virus (SLEV), Japanese encephalitis virus (JEV), and tick-borne encephalitis virus (TBEV), which, in total, annually cause millions of human infections worldwide. Subsequently, on the basis of phylogenetic analysis of a relatively limited number of viral envelope gene sequences, the mosquito-borne flaviviruses were subdivided into the *Aedes*- or *Culex*-associated

viruses. Some of the *Aedes*-transmitted viruses induce hemorrhagic fevers in humans and primates, whereas many of the *Culex*-transmitted viruses are associated with encephalitic infection in avian species (5). In addition to NKVs, Sokoluk virus (SOKV), Entebbe bat virus (ENTV), and Yokose virus (YOKV), which share ancestral roots with the *Aedes*-associated MBFVs, were isolated from vertebrates but not from arthropods. They form a clade normally referred to as the "Entebbe bat virus group," which is closely related to the YFV and Edge Hill virus (EHV) clades. It was previously suggested that these viruses may have lost their association with arthropods during their evolution and divergence (11).

In addition to these three groups of flaviviruses which infect vertebrates, additional groups have been isolated only from mosquitoes or sand flies and under experimental conditions appear to infect and replicate only in insect cell lines. Accordingly, they have tentatively been defined as insect-specific flaviviruses, until appropriate taxonomic criteria are devised and approved. The cell-fusing-agent virus (CFAV) was the first of many insect-specific flaviviruses (classical insect-specific flaviviruses [clSFs]) to be identified. CFAV was first isolated from an *Aedes aegypti* cell line in 1975 (12), and its genomic sequence was characterized in 1992 (13). CFAV and a subsequently identified heterogeneous group of related clSFs form a distinct lineage in flavivirus phylogenies. These viruses have subsequently been isolated from a wide range of mosquito species in many countries throughout the world (14–22). An additional separate group of flaviviruses that do not appear to infect vertebrate cells currently consists of nine viruses: Lammi virus (LAMV) (23), Ilomantsi virus (ILOV) (24), Marisma mosquito virus (MMV) (19), Donggang virus (DONV) (unpublished data; GenBank accession number NC_016997), Chaoyang virus (CHAOV) (25, 26), Nounane virus (NOUV) (27), Barkedji virus (BJV) (28), Nhumirim virus (NHUV) (29), and Nanay virus (NANV) (30). These nine viruses form a distinct clade that sits within the MBFV group of viruses. Moreover, flavivirus-like genomic sequences integrated within the genomes of *Aedes* mosquitoes (21, 31) have also been identified. Finally, three viruses with very highly divergent genetic lineages, *viz.*, Tamana bat virus (TABV), Ngoye virus (NGOV), and Mogiana tick virus (MGTV), are also considered to be flaviviruses because they share similar genome organizations with the recognized flaviviruses (1, 32).

The only representative of a sand fly-borne flavivirus, Saboya virus (SABV), sits in the group of mosquito-borne flaviviruses primarily associated with *Aedes* mosquitoes in Africa (5, 33). Flavivirus RNA has also been discovered in phlebotomine sand flies from Algeria (34) and Portugal (unpublished data; GenBank accession number HM563684). However, these sequences align closely with those of the clSFs.

We report here on the detection, isolation, complete genome sequence, and phylogenetic assignment of a novel sand fly-borne flavivirus in *Psathyromyia abonnenci* (*Pa. abonnenci*) sand flies. We propose the name Ecuador Paraiso Escondido virus (EPEV), based on the village where the sand flies were collected. We also discuss the possible significance of the discovery of a New World (NW) sand fly-associated virus that shares a common ancestral lineage with nonvectored Old World (OW) flaviviruses.

MATERIALS AND METHODS

Trapping of sand flies. Sand flies were trapped during March 2011 in the locality of Paraiso Escondido (00°85′03″N, 79°17′49″W), Pichincha Province, Ecuador. CDC miniature light traps were placed from dusk to dawn. Individual sand flies were identified morphologically using a microscope (35, 36). The identified sand flies were pooled on the basis of species and sex with up to 50 individuals per pool and placed in 1.5-ml tubes for storage at −80°C.

Detection of virus. Pools of sand flies were ground in 600 μl of Eagle minimal essential medium (supplemented with 7% fetal bovine serum, 1% penicillin-streptomycin, and 1% L-glutamine [200 mM]) in the presence of 3-mm tungsten beads using an MM300 mixer mill (Qiagen, Courtaboeuf, France) as described previously (37). A 200-μl aliquot was used for viral nucleic acid (NA) extraction with a BioRobot EZ1-XL Advanced virus extraction minikit (Qiagen) and eluted in a 90-μl volume. Five microliters of this solution was used for SYBR green reverse transcription (RT)-PCR and seminested PCR assays with the primers and by the protocol described previously (34, 38). Bands of the expected size were purified (Amicon Ultra centrifugal filters; Millipore) and directly sequenced.

Isolation of virus. A 100-μl volume of the PCR-positive sand fly pool was inoculated onto C6/36 cell monolayers in 25-cm² tissue culture flasks after being mixed with 900 μl of L15 medium enriched with 1% penicillin-streptomycin, 1% L-glutamine (200 mM), 5% kanamycin, 3% amphotericin B (Fungizone), and 5% tryptose phosphate broth solution. After incubation at room temperature for 1 h, 5 ml fresh medium containing 5% fetal bovine serum (FBS) was added. The flasks were incubated at 28°C and examined daily for the presence of a cytopathic effect (CPE).

***In vitro* culture of EPEV.** Cell lines of different vertebrate species, including human (SW13), hamster (BHK), monkey (Vero), and amphibian (XTC), were inoculated with the supernatant medium of EPEV-infected C6/36 cells obtained at passage 6. Two flasks were inoculated for each cell line and incubated at either 32°C or 37°C. The flasks were examined daily for the presence of a CPE. A 100-μl volume of the PCR-positive sand fly homogenate was also inoculated onto Vero cells. In the absence of a CPE, the cells were harvested after 7 days, and nucleic acids were purified. Regardless of the absence of a CPE, 5 serial passages were performed, and each was tested by real-time RT-PCR (38) for the presence of EPEV RNA.

Mouse brain inoculation. A total of 15 μl of undiluted EPEV-containing supernatant medium (passage 4) or 15 μl of EPEV-containing supernatant medium (passage 4) diluted 1:10 with minimal essential medium was injected intracerebrally into 2-day-old newborn OF1 mice. The baby mice were observed for 14 days and then euthanized. Nucleic acids were purified from the brain tissues and used for the detection of EPEV RNA by a specific real-time RT-PCR assay (38). Additional mice were injected with supernatant medium containing a pool of infected brain tissue from the previously infected mice. They were observed for 14 days and then euthanized, and nucleic acids were purified from the brain tissues and used for detection of EPEV RNA by a specific real-time RT-PCR assay (38). Veterinary Services of the Ministry of Agriculture has approved animal experimentation under the number A1301309.

Complete genome sequencing. The EPEV strain (passage 6 in C6/36 cells) and the original homogenate of the EPEV-positive sand fly pool were used independently for complete genome characterization through next-generation sequencing (NGS). Briefly, 140 μl of each sample was incubated at 37°C for 7 h in 30 U of Benzonase endonuclease (catalog number 70664-3; Novagen) to eliminate cellular DNA and RNA and preserve encapsidated viral particles. The encapsidated viral particles were then processed for RNA extraction using a BioRobot EZ1-XL Advanced viral RNA minikit (Qiagen) without an RNA carrier. Random amplification was performed using a tagged random primer for RT and using tag-specific and random primers for PCR amplification (Applied Biosystems). The PCR products were purified (Amicon Ultra centrifugal filters; Millipore), and quantification was done using a Qubit fluorometer. Two hundred nanograms of the sample was processed for sequencing using an Ion PGM sequencer (Life Technologies SAS, Saint Aubin, France) (39).

Analyses of the sequencing data were performed using the CLC Genomics Workbench program (v6.5; Qiagen).

Ultracentrifugation and genome circularization. The clarified supernatant medium of EPEV-infected C6/36 cells was incubated at 4°C overnight in a polyethylene glycol 6000 (Sigma-Aldrich) solution, to concentrate the virions. The concentrate was centrifuged at 3,000 rpm for 30 min, and the pelleted sediments were resuspended and ultracentrifuged at 30,000 rpm for 3 h using a Fiberlite F14-6 × 250y fixed-angle rotor with a Thermo Scientific Heraeus Multifuge X3R ultracentrifuge. The phase 4 fraction of the centrifuge tube was collected to provide the sample that was then used for viral RNA extraction. A 400-μl volume was used for viral RNA extraction as mentioned above and treated with 10 U/μl tobacco acid pyrophosphatase (Epicentre); circularization was performed with 5 U/μl T4 RNA ligase (Ambion) at 10°C overnight. Specific primers were designed to perform RT-PCR and to amplify the extremities using an Access RT-PCR one-step kit (Promega). The positive samples were gel purified using a QIAquick gel extraction kit (Qiagen), and both strands were sequenced.

Phylogenetic analysis. Flavivirus sequences available in the GenBank database were collected to obtain a data set including a representative of at least one sequence for each species of the genus *Flavivirus* available as an amino acid sequence of the complete ORF. Alignments of the amino acid sequences of the complete ORF were generated using both the Clustal W2 (40, 41) and MUSCLE (42) programs, available at the EMBL server (http://www.ebi.ac.uk/Tools/), and refined manually for comparison.

The effect of mass removal of regions of ambiguous alignment by use of the GBlocks algorithm (43) was also investigated. Phylogenetic trees were reconstructed using Markov chain Monte Carlo (MCMC) analysis implemented in the MrBayes software program (v3.1.2) (44). The analysis was performed using the WAG substitution model with a gamma-distributed rate variation among sites and default priors. Five independent Markov chains were run for 10 million generations, with the first 10% of samples being discarded as burn-in. Stationarity was confirmed on the basis of effective sample sizes of >400 using the Tracer program (v1.4.1) (45). A maximum clade credibility tree was summarized using the Tree-Annotator program, which annotates all nodes with posterior probability support values.

Genetic distances. The amino acid distances among the representatives of the flaviviruses were calculated for the complete ORF and NS5 protein by use of the p-distance method in MEGA software (v5) (46). Nucleotide distances were also calculated for the 1-kb region in the NS5 protein reported by Kuno et al. (4).

Comparison of cleavage sites, glycosylation sites, cysteine residues, and conserved motifs of EPEV and known flaviviruses. Putative cleavage sites from amino acid sequence alignments that included EPEV and other flaviviruses were deduced and compared according to the proteolytic processing cascade pattern previously described for the flavivirus ORFs (47). Predicted glycosylation and conserved cysteine residue sites were determined for EPEV and compared with those of known flaviviruses. Conserved enzymatic patterns of EPEV were also analyzed and compared with those of known flaviviruses.

Search for ribosomal −1 frameshifting. Possible ribosomal frameshifting sites on the EPEV genome were investigated using the methods described previously (48–50).

Nucleotide sequence accession numbers. The complete genomic sequence of EPEV has been submitted to the GenBank database and may be found under accession number KJ152564.

RESULTS

Trapping of sand flies and detection of the virus. One thousand three hundred sand flies (1,150 females and 150 males) were collected in Paraiso Escondido village, and the following species were identified morphologically: *Nyssomyia trapidoi*, *Psychodopygus panamensis*, *Psathyromyia aragoi*, *Pa. abonnenci*, *Lutzomyia hartmanni*, *Trichophoromyia reburra*, and a *Trichophoromyia* sp.

Twenty-six pools were tested by PCR for the presence of flavivirus RNA. One pool of 30 female *Pa. abonnenci* sand flies contained flavivirus RNA when the pan-flavivirus RT-PCR (38) was used, and on the basis of quantitative real-time PCR, the pool contained >10^{12} genome copies. Importantly, this EPEV-positive pool consisted exclusively of females that were neither gravid nor engorged, thus precluding the possibility that a blood meal from a vertebrate host might be the source of EPEV in the tested sample. On the basis of the fact that all the other sand fly pools were negative and the positive pool contained nonengorged sand flies, it seems reasonable to assume that a single sand fly might have been positive for EPEV RNA. Thus, the evidence strongly supports the concept that EPEV must have infected and reproduced efficiently in the positive sand fly. However, we cannot rule out the possibility that EPEV might have been detected in mosquitoes if they had been collected from the same area where the sand flies were trapped.

Virus isolation, mouse brain inoculation, and investigation of host range *in vitro*. C6/36 cells infected with the sand fly homogenate from the PCR-positive pool showed a cytopathic effect within 3 days and 1 day at the 4th passage and 6th passage, respectively (Fig. 1). Amplification of the virus using the pan-flavivirus RT-PCR assay (38) for the first three passages was also confirmed, with threshold cycle values being 18.32, 16.21, and 19.01, respectively. In order to test whether or not EPEV could infect human and/or other vertebrate cells in culture, Vero, BHK, SW13, and XTC cell cultures were inoculated with passage 6 EPEV and incubated at 32°C and 37°C. None showed signs of a CPE during 5 serial passages. Moreover, no CPE was detected following five serial passages of Vero cells inoculated with the original EPEV-positive sand fly homogenate. Viral RNA detection tests using the pan-flavivirus RT-PCR assay also produced negative results, indicating that EPEV did not replicate in any of the vertebrate cell lines tested. Finally, newborn mice inoculated with the original EPEV-positive sand fly homogenate showed no signs of disease, and EPEV RNA could not be detected in the brains of these newborn mice. Thus, under the conditions employed in these experiments, EPEV did not infect or replicate in vertebrate cells or newborn mice.

EPEV genome strategy. (i) Complete genome sequencing and cyclization. Initially, an almost complete sequence of C6/36 cell-derived virus was obtained by NGS. Unfortunately, the sand fly homogenate-derived virus gave very poor sequence readouts after NGS. Reads with minimum lengths of 30 nucleotides and with a minimum quality of 99% per base were trimmed using the CLC Genomic Workbench program (v6.5) and mapped to reference sequences previously obtained by Sanger sequencing. Parameters were set to ensure that each accepted read mapped to the reference sequence for at least 50% of its length and had a minimum of 80% identity to the reference sequence.

Thus, the reported sequence was derived from NGS of the C6/36 cell-derived virus and Sanger sequencing of the gaps. Missing regions, including segments of the capsid protein, the Pr protein, and the 5′ and 3′ noncoding regions, were successfully amplified after cyclization of the genome and sequencing of the concatenated 3′ and 5′ termini. This resulted in determination of the complete genomic sequence of EPEV (GenBank accession number KJ152564), which was 10,761 nucleotides (nt) in length and which had an ORF of 10,323 nt (positions 120 to 10,442). The single complete ORF was predicted to encode a polyprotein of
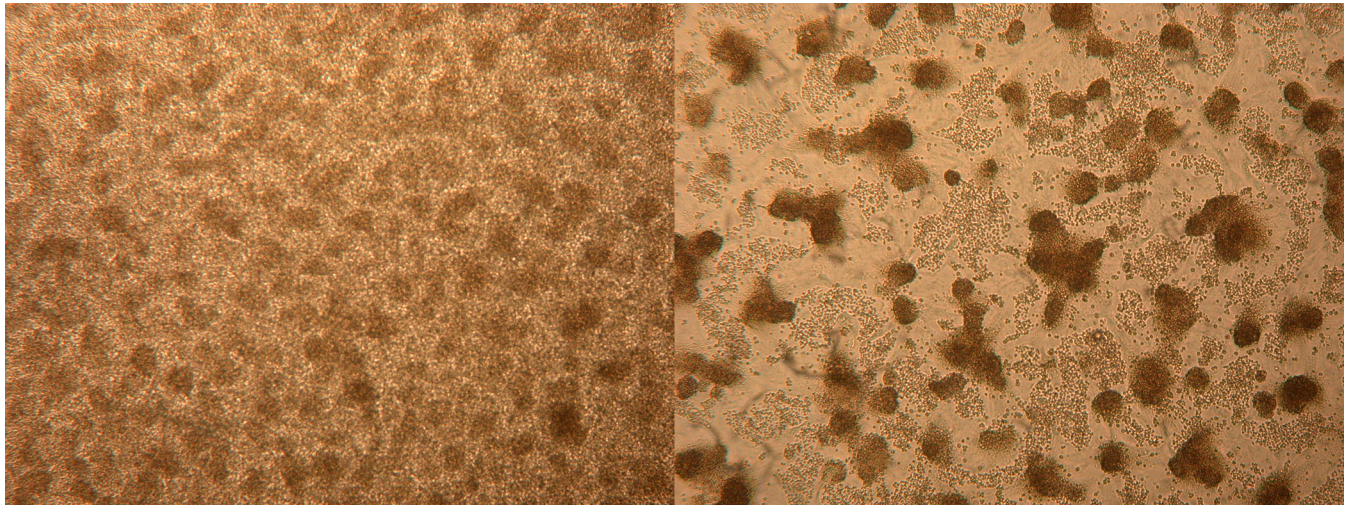
FIG 1 C6/36 cell cultures at the 6th passage after 1 day. (Left) Noninfected cells; (right) cells infected with EPEV.

3,441 amino acids (aa). The lengths and positions of the genes or genomic regions are shown in Table 1.

(ii) Terminal dinucleotides. At each terminus of the genome, 2 nucleotides which are conserved among members of the entire *Flavivirus* genus (51–53), i.e., 5′-AG and CT-3′, were detected. Mutagenesis studies have previously shown that CT-3′ is probably a recognition site for the replication complex to initiate RNA synthesis (54, 55).

(iii) 5′ NCR and 3′ NCR. The 5′ and 3′ NCRs are thought to function as recognition sites for virus translation, replication, and possibly, assembly (56). The lengths of the EPEV 5′ and 3′ NCRs were 119 and 316 nt, respectively. On the basis of comparative alignment of the nucleotide sequences of all flaviviruses, the EPEV 316-nt 3′ NCR is significantly shorter than that of any of the previously studied flaviviruses and lacks a number of conserved sequences found in the mosquito-borne flaviviruses. The only conserved sequence detected, conserved sequence 1 (CS1), was 5′-CATATTGACACCAGGGAAAAGAC-3′, which is located near the

3′ terminus of the 3′ NCR. This is conserved in all mosquito-borne flaviviruses. CS1 is complementary to a conserved sequence within the N-terminal coding region of the capsid protein. This complementarity results in a long-range RNA interaction that promotes cyclization of the flavivirus genome, forming a panhandle-like RNA structure required for viral RNA synthesis (57). In the EPEV genome, the first 12 nucleotides of CS1 are complementary to a region in the capsid protein. Conserved sequence 2 (CS2) is also found in all mosquito-borne flaviviruses. The nucleotides in a short stretch prior to CS1 in EPEV showed very poor homology with the nucleotides at the equivalent positions of other flaviviruses, and the degenerate reverse primer 3′UTR-MOS (5′-GGT CTCCWMTAACCTCTAG-3′), which anneals to CS2 (58), was unable to anneal to this related region in the EPEV 3′ untranslated region.

The highly conserved pentanucleotide motif 5′-CACAG-3′ at the terminus of the 3′ NCR was identified. This pentanucleotide is conserved among almost all mosquito- and tick-borne flaviviruses (the exception is 5′-CACCG-3′ in Murray Valley encephalitis virus [MVEV]), and it has been reported to be an element essential for viral RNA synthesis, whereas its functional role is not known (59–61). It is more variable in the NKV group.

Comparison of the EPEV polyproteins with the polyproteins of all available flaviviruses. The predicted pentapeptide cleavage sites and conserved domains of EPEV were compared with the corresponding sites of all available flavivirus polyprotein sequences generated using amino acid sequence alignments.

(i) Cleavage sites. The predicted cleavage sites for EPEV are shown in Table 2. A comparison of these cleavage sites with those of other flaviviruses is shown in Table S1 in the supplemental material. The similarities detected are consistent with the phylogenetic position of EPEV, which is explained below in "Phylogenetic analysis" and "Genetic distances."

(a) Structural genes. The mature virion capsid protein (VirC) is cleaved from the nascent capsid protein (AnchC) after a dibasic amino acid sequence at positions 101 and 102 and before a C-terminal hydrophobic domain (CTHD) by the viral serine protease (VSP) in the cytoplasm of the host cell. For EPEV the proposed residues Arg-Arg were also identified in other flaviviruses. How-

TABLE 1 Genome organization of EPEV

Protein type	Gene(s) or region	Gene size (no. of nucleotides)	Genome position	Protein size (no. of amino acids)
Structural	5′ NCR	119	1–119	
	VirC	306	120–425	102
	AnchC	51	426–476	17
	prM	333	477–809	111
	M	234	810–1043	78
	E	1,476	1044–2519	492
Nonstructural	NS1	1,059	2520–3578	353
	NS2A	690	3579–4268	230
	NS2B	393	4269–4661	131
	NS3	1,872	4662–6533	624
	NS4A	378	6534–6911	126
	2K	69	6912–6980	23
	NS4B	750	6981–7730	250
	NS5	2,712	7731–10442	904
	3′ NCR	316	10446–10761	

TABLE 2 Putative polyprotein cleavage sites of EPEV polyprotein

Cleavage site	Protease	Amino acid sequence
VirC/AnchC	VSP	KKKRR/NGGTA
AnchC/prM	HS	MMIVC/YVNAR
Pr/M	Furin	HRHRR/RSPRS or RRSPR/SIAMP
M/E	HS	GPAYG/THCLT
E/NS1	HS	TGVGA/DVGCS
NS1/NS2A	?	SLVSA/GNGMV
NS2A/NS2B	VSP	WRKTR/SWPIS
NS2B/NS3	VSP	RSKRR/AGILW
NS3/NS4A	VSP	AEGRR/NWTGL
NS4A/2K	VSP	EPGCQR/SAQDN
2K/NS4B	HS	AGVAA/NELGW
NS4B/NS5	VSP	KSFRR/GKAQA

ever, Asn at position −1 has not been reported for VirC/AnchC but it was reported for other cleavage sites.

The CTHD protein is cleaved from the prM polyprotein after Cys (at position 119 for EPEV) by a host signalase (HS) in the lumen of the endoplasmic reticulum (ER). Cys at position −1 was not reported for flaviviruses, but it is consistent with the (−3, −1) rule of von Heijne (62).

The prM protein is the glycosylated precursor of the M protein. Pr is cleaved from the M protein by the host cell furin or a similar enzyme (63, 64) in the lumen of the ER. A specific pattern of Arg-X-(Arg/Lys)-Arg was reported for this cleavage site (65). For EPEV, we propose the signature His-Arg-(Arg/Arg)-Ser after the amino acid at position 230 or Ser-Pro-Arg/Ser-Ile after the amino acid at position 234.

The M protein is cleaved from the envelope protein after Gly at position 308 by an HS in the lumen of the ER, which is consistent with the rule of von Heijne (62).

The envelope protein is cleaved from the nonstructural NS1 protein after Ala at position 799 by a signalase (HS) in the lumen of the ER, which is consistent with the rule of von Heijne (62).

(b) Nonstructural genes. The NS1/NS2A cleavage site occurs after Ala at position 1153, which is consistent with the rule of von Heijne (62).

The NS2A/NS2B cleavage site occurs after Arg at position 1384, which is cleaved by the viral serine protease, but Thr at position −2 is not consistent with the dibasic requirement at positions −1 and −2. Experimental evidence is required to determine whether the proposed cleavage site is functional or not.

The NS2B/NS3 cleavage site occurs after two Arg residues at positions 1514 and 1515, which are cleaved by the viral serine protease.

The NS3/NS4A cleavage site occurs after a double Arg at positions 2138 and 2139, which are cleaved by the viral serine protease.

The NS4A/2K cleavage site occurs after Gln and Arg at positions 2264 and 2265, respectively, which are cleaved by the viral serine protease.

The 2K/NS4B cleavage site occurs after Ala at position 2288, which is cleaved by a host signalase. This is consistent with the rule of von Heijne (62).

The NS4B/NS5 cleavage site occurs after a double Arg at positions 2537 and 2538, which are cleaved by the viral serine protease.

(ii) Glycosylation sites and conserved cysteine residues. The Pr gene region contains 1 possible N-linked glycosylation site and 6 Cys residues which are conserved among other flaviviruses ex-

cept the clSFs, the E-gene region contains 2 possible N-linked glycosylation sites and 12 Cys residues which are conserved in all other flaviviruses except the clSFs, and the NS1 protein contains 2 possible N-linked glycosylation sites and 12 Cys residues which are highly conserved. Cysteine residues are required for stabilizing disulfide bridges in all mosquito-borne viruses (3).

(iii) Fusion peptide. A sequence homologous to the fusion peptide, a 14-aa motif thought to be involved in virus fusion with cellular membranes (66), is present at positions 406 to 419 in the NS1 protein. It conforms with the Asp-Arg-Gly-Trp-X-X-(Gly/His)-Cys-X-X-Phe-Gly-Lys-Gly motif observed for all mosquito-borne flaviviruses (67, 68).

(iv) Conserved enzymatic patterns. The N-terminal sequence of the flavivirus NS3 protein contains conserved regions (boxes 1 to 4) that have significant similarity to serine proteases and belong to the trypsin superfamily (69, 70). This protease activity was reported to be required for the polyprotein processing of flaviviruses. On the other hand, the C-terminal sequence contains conserved regions (motifs I, IA, and II to VI) which are similar to RNA helicases of the DEAD family (71). These motifs are shown in Fig. 2 for EPEV. Compared with the sequences of other flaviviruses, these patterns are highly conserved in the EPEV genome. The EPEV NS5 protein contains the conserved sequences (motifs A to D) associated with RNA-dependent RNA polymerase activity (72) and also the Gly-Asp-Asp sequence of motif C, which is the active site of the polymerase. In the N-terminal domain, two conserved motifs, *viz.*, motifs 1 and 2, are similar to methyltransferases (Fig. 2).

Comparative alignment of the amino acid sequences of EPEV and all available flaviviruses. With the ultimate objective of generating a phylogenetic tree, the nucleotide and amino acid sequences of the entire genome of EPEV were aligned with the corresponding sequences of all available flaviviruses. Despite the fact that EPEV was genetically significantly different from the other flaviviruses, the results of the alignment clearly identified EPEV to be a flavivirus. Nevertheless, the EPEV genome was distinct from the genomes of all other recognized flaviviruses in that the capsid protein contained a unique sequence of 17 amino acids from positions 141 to 157. Use of the BLAST search engine revealed no equivalent sequence in any known virus or cellular sequences. However, fragments from this sequence of up to 10 amino acids matched fragments of sequences found in a variety of bacterial enzymes.

We then prepared a truncated flavivirus alignment containing the amino acid sequences of only the capsid and prM, and we attempted to align manually a 26-aa fragment that contained the 17-aa fragment of the unique EPEV sequence. This truncated alignment was manually adjusted as described previously when the untranslated regions of the flaviviruses were analyzed (73–76). Using this method, residual amino acids within the 26-aa fragment aligned with amino acids in the upstream region of the prM protein, particularly with those of YFV, Sepik virus (SEPV), and Wesselsbron virus (WSLV), i.e., the most closely related viruses (Fig. 3).

Phylogenetic analysis. The topology of the phylogenetic tree based on the genomic sequence is in agreement with that in previously published analyses (11, 23, 33, 58, 77, 78) and clearly demonstrates the segregation of the major clusters consisting of the mosquito-borne, tick-borne, and no-known-vector flaviviruses with 100% support. The MBFVs showed a divergence of the two

In NS3

1. Serine protease

	Box 1	Box 2	Box 3	Box 4
EPEV	GVLHTMWHVTRG	DLVSYGGSW	DFPSGTSGSP	GNVLGLYGNGILVGEWDYVS
YFV	GVFHTMWHVTRG //	DLVAYGGSW //	DYPSGTSGSP //	GEVIGLYGNGILVGDNSFVS

2. Helicase / NTPase

	Motif I	Motif IA	Motif II
EPEV	DFHPGAGKT	RTLVLAPTRVVAAEMA	PNWEVIIDEAHFTDPRSIAVRG
YFV	DFHPGAGKT //	RTLVLAPTRVVLSEMK //	VNWEVIIMDEAHFLDPASIAARG

	Motif III	Motif IV	Motif V	Motif VI
EPEV	MSATPPG	TLWFVPSIRSGNEIAACL	FILATDIAEMGAN	SAAQRRGRVGR
YFV	MTATPPG //	TAWFLPSIRAANVMAASL //	FILATDIAEMGAN //	SAAQRRGRIGR

In NS5

3. Methyltransferase

	Motif 1	Motif 2
EPEV	VVDLGCGRGGW	DTVLCDIGES
YFV	VIDLGCGRGGW //	DTLLCDIGES

4. RNA-dependent RNA polymerase

	Motif A	Motif B	Motif C	Motif D
EPEV	YADDTAGWDT	GQVVTYALNTFTN	SGDDQV	LIHLNAMAKTRKDIDEW
YFV	YADDTAGWDT //	GQVVTYALNTITN //	SGDDCV //	LSHLNAMSKVRKDISEW

FIG 2 Conserved motifs in the polyprotein of EPEV. The numbers 1 to 4 indicate the conserved enzymatic motifs in the proteins encoded by the NS3 and NS5 genes. Slashes, gaps; residues in yellow, nonconserved positions.
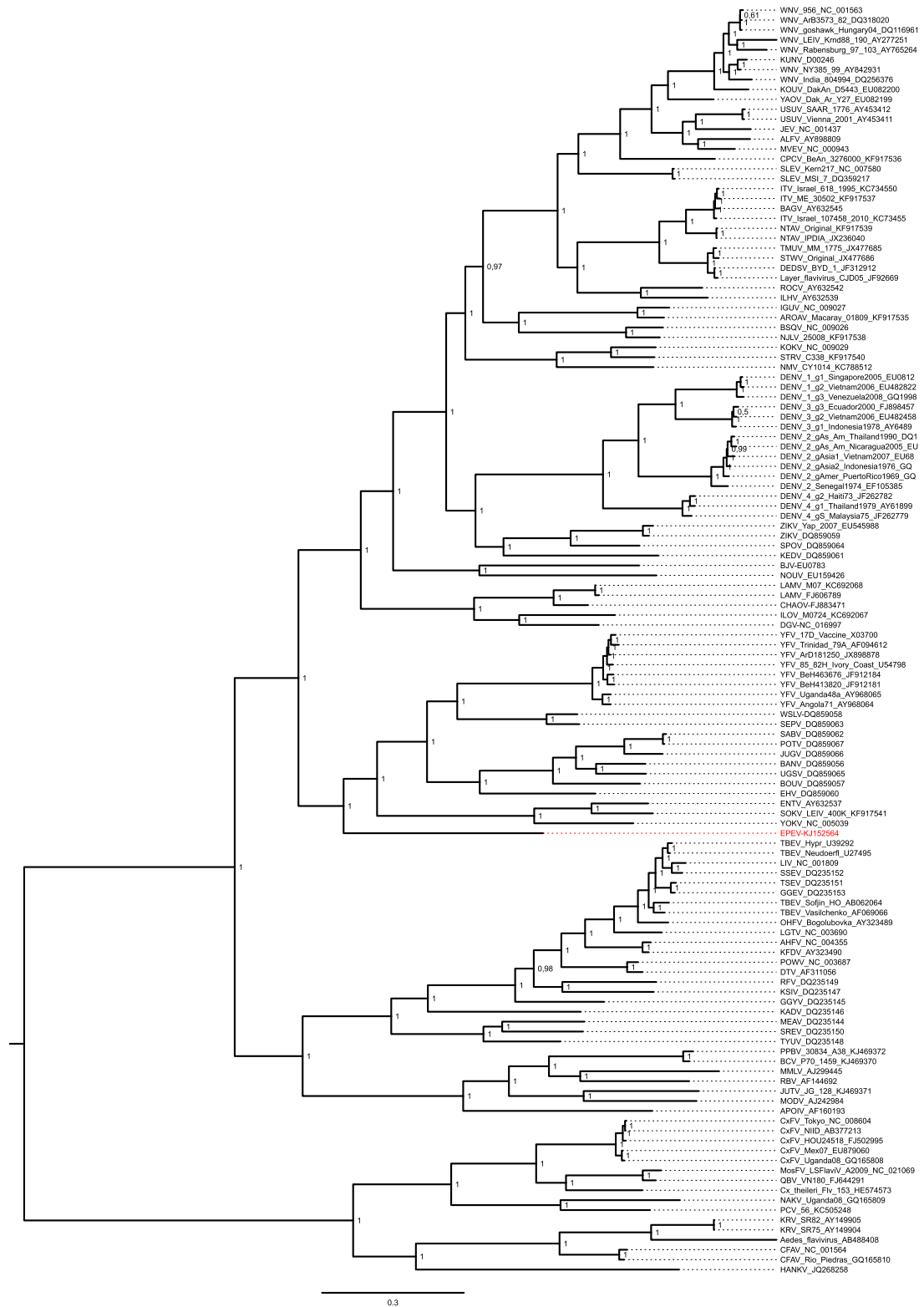
major clades, i.e., the viruses primarily associated with either *Aedes* or *Culex* mosquitoes (as delineated in Fig. 3). Interestingly, the most recent common ancestor for EPEV was shared with the nonvectored ENTV clade and the larger MBFV clade that includes SEPV, YFV, and several other viruses. It is also important to note that, apart from EPEV, in these three clades that share an ancestor, only YFV is found in the NW, where it is believed to have been introduced during the slave trade. Nevertheless, EPEV diverged from these other lineages and was clearly distinct from these other flavivirus species, the closest of which was WSLV, with which it had a genome sequence homology of 52.8%.

In view of the discovery of a unique sequence fragment in the EPEV capsid region and the knowledge that some flavivirus genes possess regions with ambiguous, highly variable sequences, it was decided to conduct two phylogenetic analyses. In the first, a tree was constructed in which all ambiguous/highly variable sequence regions, including the unique EPEV capsid sequence, were retained in the alignment. In the second tree, these ambiguous regions were removed using GBlocks trimming. Despite this cautious approach, the overall topology, branching patterns, and support for the trees were closely similar (data not shown). Moreover, the relative position of EPEV remained the same in both trees with 100% support. Independent phylogenetic analyses based on the complete ORF and, separately, the E-gene sequence also resulted in a similar position of EPEV, in which it appeared outside the ENTV and YFV clusters (data not shown). However, in trees based on the NS3 and NS5 gene sequences, EPEV was positioned between the ENTV and YFV clusters (data not shown).

Genetic distances. The pairwise distances of the amino acid sequences between EPEV and other flaviviruses are shown in Ta-

ble S2 in the supplemental material. More than 15 years ago, in a pioneer study, it was proposed that virus species and clades within the genus *Flavivirus* should be based on nucleotide distances using a 1-kb region of the NS5 gene (4), and cutoff values were set at

69% and 84% for discriminating clades and species, respectively. Using these values, EPEV, which is at best 68% similar to the 13 most closely related viruses (SOKV, ENTV, YOKV, SEPV, WSLV, YFV, EHV, Bouboui virus [BOUV], Banzi virus [BANV],

Uganda S virus, Jugra virus [JUGV], Potiskum virus [POTV], and SABV), would represent a new species. The maximum 68.0% pairwise nucleotide sequence identity of EPEV with these viruses was <69.0%, which was the cutoff value that grouped viruses in the same clade, and <84%, which was the cutoff to consider two viruses to be the same species. Moreover, the maximum pairwise amino acid sequence identities of the complete ORF and NS5 gene (52.8% and 65.4%, respectively) between EPEV and the YFV-related clade were lower than the minimum pairwise amino acid identities of the complete ORF and the NS5 gene (61% and 69.4%, respectively) between the clades defined at that time (4). In a separate analysis, the cutoff value based on complete ORF sequences was recorded to be <55% amino acid sequence identity for a number of arthropod-borne flaviviruses (77), which is also higher than the maximum identity of 52.8% with EPEV.

A search for ribosomal −1 frameshift sites. Many RNA virus genomes contain specific sequences that direct a proportion of ribosomes into an alternative reading frame. Where functionally utilized, this is known as programmed ribosomal frameshifting (PRF). The most common type of PRF involves −1 tandem slippage of the P- and A-site tRNAs on a slippery heptanucleotide sequence with the consensus motif X_XXY_YYZ, where X_XX represents any three identical nucleotides (with some exceptions, such as U_CC and G_GA); Y_YY represents A_AA or U_UU; Z represents A, C, or U; and underscores separate zero-frame codons. For efficient −1 PRF to occur, an extra stimulatory element is required, and this normally takes the form of a downstream RNA stem-loop or pseudoknot structure separated from the heptanucleotide shift site by a spacer sequence of 5 to 9 nt (79).

Ribosomal −1 frameshifting in the gene expression of JEV serogroup viruses (33, 48, 80) and the clSFs (i.e., CFAV, *Culex* flavivirus [CxFV], and relatives [49]) has been described previously. On the basis of comparative genomic analysis, −1 PRF has also been predicted to be utilized by viruses in the CHAOV-LAMV-DONV-ILOV and WSLV SEPV flavivirus clades (33, 49). These frameshift sites occur within the region of the genome encoding NS2A-NS2B, with the site being conserved within a clade but not obviously conserved between clades, suggesting that −1 PRF in the flaviviruses may have evolved independently on several occasions. Analysis of EPEV for potential −1 PRF sites was inconclusive. On the one hand, the genome contains several slippery heptanucleotide motifs, a few of which have potential downstream RNA structures. However, predictions based on a single sequence are inadequate, as such motifs have a high probability of occurring in random sequences. None of the potential −1 PRF sites identified aligned with the previously predicted −1 PRF site in the related WSLV SEPV clade. Moreover, the high degree of divergence

of EPEV from all other sequenced flavivirus species precluded robust comparative genomic analyses to identify (or tentatively rule out) the occurrence of purifying selection (i.e., evidence for functionality) at other sites with a potential −1 PRF.

DISCUSSION

During the past 10 years, thousands of sand flies (*Phlebotomus* and *Sergentomyia* spp.) have been collected in Europe, Africa, and the Middle East and analyzed for the presence of flavivirus RNA. We also systematically inoculated sand fly extracts onto Vero cells. None of these analyses produced evidence of infection by EPEV or any other flavivirus, suggesting that the well-known flaviviruses are rarely (if at all) vectored by sand fly species, at least in these regions of the OW. In the current study, we analyzed pools of sand flies collected in Ecuador for the presence of flavivirus-related RNA. One pool of *Pa. abonnenci* sand flies yielded a PCR product, the sequence of which suggested a genetic relationship with viruses in the genus *Flavivirus*. This PCR-positive pool of sand flies was also inoculated onto C6/36 cells, and a cytopathic virus was subsequently isolated. We have named that virus Ecuador Paraiso Escondido virus (EPEV). However, attempts to isolate the virus using Vero cells failed, implying that EPEV is another example of an insect-specific flavivirus. Thus, EPEV appears to be unique, in that it was isolated in Ecuador, i.e., the NW, from *Pa. abonnenci* sand flies. Detailed comparative analysis of the entire viral RNA sequence of EPEV revealed a genomic organization very similar to that of viruses in the genus *Flavivirus* (2, 3). However, alignment of the corresponding regions of EPEV and other flaviviruses revealed in EPEV a unique insert of 51 nucleotides (17 amino acids) between amino acid positions 141 and 157 (as determined from the 5′-terminal starting codon), i.e., near the boundary between the viral capsid and prM genes. The lack of hydrophobic amino acids in this region indicates that this unique sequence is more representative of the newly formed N-terminal segment of the prM protein than the hydrophobic transmembrane (TM) domain of immature C protein. The TM domain is proteolytically cleaved to generate mature C and prM proteins following coordinated signalase- and NS2B/3pro-mediated cleavage (81, 82). By manually adjusting the alignment as described previously (73–76), residual amino acids within the region unique to EPEV aligned with amino acids in the upstream region of the prM protein, particularly but not exclusively with those of YFV, SEPV, and WSLV. Importantly, no similar nucleotide sequences were identified among other virus and cellular sequences when the BLAST search engine was used. This unique EPEV sequence therefore appears to contain ancestral elements of the flaviviral prM sequence which have presumably

FIG 3 Bayesian phylogeny of the ORFs of the members of the genus *Flavivirus*. Only posterior probabilities are included. The tree is rooted at the midpoint. Strain names and GenBank accession numbers are given after the names of the viruses, which are abbreviated as follows: KUNV, Kunjin virus; KOUV, Koutango virus; YAOV, Yaounde virus; USUV, Usutu virus; ALFV, Alfuy virus; CPCV, Cacipacore virus; ITV, Israel turkey meningoencephalitis virus; BAGV, Bagaza virus; NTAV, Ntaya virus; TMUV, Tembusu virus; STWV, Sitiawan virus; DEDSV, duck egg drop syndrome virus; ROCV, Rocio virus; ILHV, Ilheus virus; IGUV, Iguape virus; AROAV, Aroa virus; BSQV, Bussuquara virus; NJLV, Naranjal virus; KOKV, Kokobera virus; STRV, Stratford virus; NMV, New Mapoon virus; DENV_1 to DENV_4, dengue virus serotypes 1 to 4, respectively; ZIKV, Zika virus; SPOV, Spondweni virus; KEDV, Kedougou virus; DGV, Donggang virus; UGSV, Uganda S virus; LIV, Louping ill virus; SSEV, Spanish sheep encephalitis virus; TSEV, Turkish sheep encephalitis virus; GGEV, Greek goat encephalitis virus; OHFV, Omsk hemorrhagic fever virus; LGTV, Langat virus; AHFV, Alkhurma virus; KFDV, Kyasanur Forest disease virus; RFV, Royal Farm virus; KSIV, Karshi virus; GGYV, Gadgets Gully virus; KADV, Kadam virus; MEAV, Meaban virus; SREV, Saumarez Reef virus; TYUV, Tyuleniy virus; PPBV, Phnom Penh bat virus; BCV, Batu Cave virus; MMLV, Montana myotis leukoencephalitis virus; RBV, Rio Bravo virus; JUTV, Jutiapa virus; MODV, Modoc virus; MosFV, mosquito flavivirus virus; QBV, Quang Ninh virus; Cv_theileri, Culex theileri flavivirus; NAKV, Nakiwogo virus; PCV, Palm Creek virus; KRV, Kamiti River virus; HANKV, Hanko virus. The definitions of the other abbreviations for viruses are provided in the text.

survived during the insertion and deletion process of RNA evolution, as described previously (73–76).

The unique single polyprotein with conserved cleavage sites, the 5′-to-3′ orientation of structural and nonstructural proteins, and the highly conserved amino acid sequence domains in the virus genome-encoded enzymes justify the inclusion of EPEV in the genus *Flavivirus*. The cysteine residues, 6 in the prM, 12 in the envelope, and 12 in the NS1, proteins which play an important role in protein folding through disulfide bridges, are all conserved among other flaviviruses. There are also N-glycosylation motifs which were speculated to be very important for viral replication, virulence, and maturation of viral particles (83–86) in the prM, envelope, and NS1 proteins, some of which are conserved in EPEV. EPEV has the shortest 3′ NCR, with 316 nt, among the flaviviruses. It contains only conserved sequence 1, which is found in all MBFVs, whereas conserved sequence 2 was highly variable. This conserved sequence 2 was reported to have a minor effect on viral RNA synthesis but seems to decrease viral pathogenicity (87–90).

The presence of a given virus in an arthropod does not confirm that the corresponding vector species plays a role in the natural cycle of the virus. However, in the case of the EPEV-positive sand fly pool, the *Pa. abonnenci* females were neither gravid nor engorged, and an extremely high viral load was detected. This is compelling evidence that EPEV replicates in this sand fly species. Thus, *Pa. abonnenci* and possibly other NW sand flies may play a direct role in the natural infectious life cycle of EPEV.

EPEV is not the first flavivirus isolated from sand flies, although few viruses have been found in phlebotomine flies until now. Saboya virus was isolated from sand flies in Senegal (91), although it was originally isolated from the gerbil species *Tatem itempi* in Saboya Village in Senegal (92) and was later recovered from rodents of the *Mastomys*, *Aivicantis niliaticus*, and *Mus musculus* species (93, 94) and chiropters in the Republic of Guinea (95). It was also isolated from *Anopheles* mosquitoes and ticks (77). In addition, two flavivirus sequences were detected in *Phlebotomus perniciosus* sand flies in Algeria. These formed a monophyletic group more closely related to *Culex*-associated insect-only flaviviruses (34). Flavivirus RNA has also been discovered in phlebotomine sand flies from Portugal (unpublished data; GenBank accession number HM563684).

The phylogenetic analysis showed that EPEV diverges from the common ancestral lineage of *Aedes*-associated mosquito-borne flaviviruses that include YFV and also the NKV-like flaviviruses (YOKV, ENTV, SOKV). EPEV is more distantly related to the other group of *Aedes*-borne flaviviruses that include dengue viruses and the *Culex*-associated group that includes WNV, SLEV, and Zika virus. Interestingly, EPEV is also only distantly related to the clSFs (CFAV, Kamiti River virus, Hanko virus, and CxFV) and dual-host-affiliated ISFs (dISFs) (NOUV, LAMV, CHAOV, Donggang virus (DGV) and ILOV) and the NKVs that include bat- and rodent-associated flaviviruses, such as Modoc virus and Montana myotis leukoencephalitis virus. EPEV appeared alone as a lineage which was separated from a major phylogenetic clade that included the YFV, EHV, and ENTV groups. The phylogenetic analysis showed that EPEV appears at the root of these groups with 100% bootstrap support. However, the phylogenetic position among the mosquito-borne viruses of EPEV conflicts with its lack of an ability to infect and replicate in vertebrate cell cultures or newborn mice. The insect-specific flaviviruses NOUV and LAMV

also appear to infect and replicate only in insect cell lines (23, 27), although they are phylogenetically grouped with mosquito-borne flaviviruses.

These data therefore support the identification of the first representative of a novel group of NW sand fly-borne flaviviruses. When the origins of EPEV and the other flaviviruses are considered, most of the earliest phylogenetic lineages of the mosquito-borne viruses, tick-borne viruses, and NKVs are found in Europe, Asia, Australia, and, in particular, Africa. On the basis of the close relationships of the NW and OW MBFVs, the current opinion is that these viruses were introduced to the NW at various times before, during, or after the period of slave trading (33). The only representative of the tick-borne flaviviruses in the NW is Powassan virus (POWV) and its relatively recent descendant variant, deer tick virus (DTV). On the basis of historical, ecological, anthropological, and molecular epidemiological evidence, it was proposed that POWV was introduced into the NW between 15,000 and 11,000 years ago, possibly during the tick and mammalian migrations, when the Bering land bridge connected Asia and North America (78, 96). Among the NKVs, the OW virus Apoi virus (APOIV), which roots all other NKVs, is the only rodent-associated virus related to the NKV group, whereas the bat-associated NKVs are found in both the OW and the NW. This supports the idea of a single dispersion into the NW, most likely through bats (11, 33). It has recently been suggested that the introduction of viruses into the NW occurred several times during two evolutionary time periods (33). It seems highly probable that the most recent introductions occurred from the period of increasing slave and commercial trading across the Atlantic Ocean from the 16th to the 19th centuries. The most convincing evidence of such dispersions includes the mosquito-borne viruses YFV, DENV, and, in 1999, WNV. For each of these viruses there are historical records and robust phylogenetic data (97–102). In the case of the potentially earlier introductions to the NW, which include Aroa virus, Ilheus virus, SLEV, Cacipacore virus, and the NKVs, their dispersion appears to have occurred thousands of years before slave and commercial trading (33). Whether or not these introductions to the NW occurred during the period when the Bering land bridge connected Asia and North America cannot be determined. However, it seems unlikely, since there are no recognized OW sand flies in the NW. On the other hand, it is recognized that estimation of the time of occurrence of such assumed introductions is highly dependent on the choice of dates used for calibration (33, 78). As analytical methods improve and more viruses are identified and characterized, it should become possible to provide more robust estimates for these virus dispersion patterns. The genus *Flavivirus* comprises both vectored and nonvectored viruses, raising the possibility that arthropod-mediated transmission is either a derived trait or a secondary loss. The former traditional hypothesis is supported extensively (4, 5, 10, 77, 103, 104), but the alternative hypothesis should not be discarded at this stage of our comprehension. The insect-specific flavivirus group consisting of NOUV, BJV, DONV, ILOV, CHAOV, LAMV, and the newly discovered NANV from Peru clusters phylogenetically with the MBFVs, but none of these insect-specific viruses have been found to replicate in mammalian cells (23, 24, 104). In contrast, the other nonvectored viruses, ENTV and YOKV, replicate in mosquito cells. This reflects the need for more in-depth field investigations to identify the possible arthropod vectors of these viruses. For example, the isolation of

SOKV, a close relative of YOKV, from *Argasidae* ticks, birds, and bats and serological evidence of human infections by YOKV were recently reported (105).

Our phylogenetic analysis places EPEV at the root of the YOKV, SOKV, and ENTV clade, which itself roots the *Aedes*-associated mosquito-borne flaviviruses (EHV, BOUV, Uganda S virus, BANV, JUGV, POTV, SABV, WSLV, SEPV, and YFV). In contrast to EPEV, all of these viruses have been found only in the OW. This finding requires a cautious interpretation before embarking on a definitive hypothesis. As previously elaborated (106), sampling issues should be taken into consideration in generating these apparently conflicting results. The situation with EPEV is further complicated when one considers that it was isolated from a pool of NW sand flies. However, it replicates in mosquito cells but does not appear to replicate in mammalian cells, and it is ancestral to the ENTV and YFV clades. Since EPEV roots clades of OW viruses, is closely related to YFV, and has also been shown to replicate to high titers in mosquito cells, we can speculate that the vectors for the transmission cycle of EPEV in the OW may include mosquitoes. In this case, the introduction of EPEV might then be attributable to the slave trade, in common with several other flaviviruses (33). This argument is further supported by the compelling evidence that the alphavirus chikungunya virus (CHIKV) caused many disease outbreaks in humans about 200 years ago in the Americas (107). It therefore seems highly likely that CHIKV was also introduced into the NW during the period of slave trading across the Atlantic Ocean. In this case, however, the introduced CHIKV does not appear to have become permanently established in NW reservoir hosts. The alternative possibility, that EPEV was introduced into the NW via infected sand flies from the OW, seems unlikely, since OW sand fly species have never been found in the NW. Clearly, mosquito trapping and screening are required in the region where the EPEV-positive sand fly pool was collected to discover whether or not EPEV circulates among the indigenous mosquito species. If this proves to be the case, it would support the hypothesis of an early divergence of the vector-borne viruses during the evolution of the flaviviruses. Moreover, the insect-specific flaviviruses (lSFs) also replicate only in mosquito cells, and the number of viruses in this group has been rapidly increasing. Most flaviviruses are associated with mosquitoes and vertebrates, but the ISF group is associated only with mosquitoes, appearing at the root of the MBFV, NKV, and TBFV branches. The highly divergent TABV, isolated in 1973 from an insectivorous bat in Trinidad (108), is now included as a tentative flavivirus (1). Recently, MGTV was isolated from ticks in Brazil (32). It appears to be highly divergent but associated with the genus *Flavivirus* on the basis of a partial NS3 and NS5 analysis. TABV and MGTV are highly divergent from each other, but they were both isolated in the NW. Nevertheless, they are related, albeit distantly, to the other flaviviruses. Thus, together with EPEV, TABV, and MGTV in the NW, viruses in the genus *Flavivirus* appear to have evolved following their introduction into the NW from the OW. However, the wide variety of different flaviviruses and their attributes are potentially biased by the small number of representatives of different lineages that have currently been identified (24).

At this stage, we cannot rule out the possibility that EPEV may be vectored by mosquitoes. However, the results in this study suggest that EPEV represents a new flavivirus species and could be the first recognized representative of a novel clade including other yet-to-be-discovered sand fly-borne flaviviruses. There is support

for this proposal: (i) EPEV has been isolated from *Pa. abonnenci*, a species of NW sand flies that is known to be present in Bolivia, Brazil, Colombia, French Guiana, Panama, Peru, Suriname, Venezuela, and Ecuador (CIPA: Computer-Aided Identification of Phlebotomine Sand flies of the Americas, http://cipa.snv.jussieu.fr [109]), (ii) *Pa. abonnenci* is not listed among the species of medical interest (110), and (iii) the absence of viral replication in various vertebrate cells and in the brains of newborn mice suggests that EPEV could be an insect-only flavivirus rather than an arbovirus.

Further studies are required to understand if there are other sand fly species infected with EPEV in the region where the *Pa. abonnenci* samples were collected. Alternatively, if mosquitoes are the natural vectors, it would seem sensible to look first in the region of Ecuador where EPEV was discovered.

REFERENCES

1. **Pletnev A, Gould E, Heinz FX, Meyers G, Thiel HJ, Bukh J, Stiasny K, Collett MS, Becher P, Simmonds P, Rice CM, Monath TP.** 2011. Flaviviridae, p 1003–1020. In King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy. Elsevier, Oxford, United Kingdom.
2. **Rice CM, Lenches EM, Eddy SR, Shin SJ, Sheets RL, Strauss JH.** 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science 229:726–733. http://dx.doi.org/10.1126/science.4023707.
3. **Chambers TJ, Hahn CS, Galler R, Rice CM.** 1990. Flavivirus genome organization, expression, and replication. Annu Rev Microbiol 44:649–688. http://dx.doi.org/10.1146/annurev.mi.44.100190.003245.
4. **Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB.** 1998. Phylogeny of the genus Flavivirus. J Virol 72:73–83.
5. **Gaunt MW, Sall AA, Lamballerie X, Falconar AK, Dzhivanian TI, Gould EA.** 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J Gen Virol 82:1867–1876. http://dx.doi.org/10.1099/0022-1317-82-8-1867.
6. **Gould EA, Solomon T.** 2008. Pathogenic flaviviruses. Lancet 371:500–509. http://dx.doi.org/10.1016/S0140-6736(08)60238-X.
7. **Grard G, Moureau G, Charrel RN, Lemasson JJ, Gonzalez JP, Gallian P, Gritsun TS, Holmes EC, Gould EA, de Lamballerie X.** 2007. Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. Virology 361:80–92. http://dx.doi.org/10.1016/j.virol.2006.09.015.
8. **Zanotto PM, Gao GF, Gritsun T, Marin MS, Jiang WR, Venugopal K, Reid HW, Gould EA.** 1995. An arbovirus cline across the northern hemisphere. Virology 210:152–159. http://dx.doi.org/10.1006/viro.1995.1326.
9. **Porterfield JS.** 1980. Antigenic characteristics and classification of Togaviridae, p 13–46. *In* Schlesinger RW (ed), The togaviruses. Academic Press, New York, NY.
10. **Gould EA, de Lamballerie X, Zanotto PM, Holmes EC.** 2003. Origins, evolution, and vector/host coadaptations within the genus Flavivirus. Adv Virus Res 59:277–314. http://dx.doi.org/10.1016/S0065-3527(03)59008-X.
11. **Cook S, Holmes EC.** 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (family: Flaviviridae) and the evolution of vector transmission. Arch Virol 151:309–325. http://dx.doi.org/10.1007/s00705-005-0626-6.
12. **Stollar V, Thomas VL.** 1975. An agent in the Aedes aegypti cell line (Peleg) which causes fusion of Aedes albopictus cells. Virology 64:367–377. http://dx.doi.org/10.1016/0042-6822(75)90113-0.
13. **Cammisa-Parks H, Cisar LA, 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. http://dx.doi.org/10.1016/0042-6822(92)90575-A.
14. **Cook S, Moureau G, Kitchen A, Gould EA, de Lambellerie X, Holmes EC, Harbach RE.** 2012. Molecular evolution of the insect-specific flaviviruses. J Gen Virol 93(Pt 2):223–234. http://dx.doi.org/10.1099/vir.0.036525-0.
15. **Huanyu W, Haiyan W, Shihong F, Guifang L, Hong L, Xiaoyan G, Lizhi S, Rayner S, Aiqiang X, Guodong L.** 2012. Isolation and identification of a distinct strain of Culex Flavivirus from mosquitoes

collected in mainland China. Virol J **9**:73. http://dx.doi.org/10.1186/1743-422X-9-73.

16. **Zuo S, Zhao Q, Guo X, Zhou H, Cao W, Zhang J.** 2014. Detection of Quang Binh virus from mosquitoes in China. Virus Res **180**:31–38. http://dx.doi.org/10.1016/j.virusres.2013.12.005.

17. **Parreira R, Cook S, Lopes A, de Matos AP, de Almeida AP, Piedade J, Esteves A.** 2012. Genetic characterization of an insect-specific flavivirus isolated from *Culex theileri* mosquitoes collected in southern Portugal. Virus Res **167**:152–161. http://dx.doi.org/10.1016/j.virusres.2012.04.010.

18. **Hobson-Peters J, Yam AW, Lu JW, Setoh YX, May FJ, Kurucz N, Walsh S, Prow NA, Davis SS, Weir R, Melville L, Hunt N, Webb RI, Blitvich BJ, Whelan P, Hall RA.** 2013. A new insect-specific flavivirus from northern Australia suppresses replication of West Nile virus and Murray Valley encephalitis virus in co-infected mosquito cells. PLoS One **8**:e56534. http://dx.doi.org/10.1371/journal.pone.0056534.

19. **Vázquez A, Sánchez-Seco MP, Palacios G, Molero F, Reyes N, Ruiz S, Aranda C, Marqués E, Escosa R, Moreno J, Figuerola J, Tenorio A.** 2012. Novel flaviviruses detected in different species of mosquitoes in Spain. Vector Borne Zoonotic Dis **12**:223–229. http://dx.doi.org/10.1089/vbz.2011.0687.

20. **Grisenti M, Vázquez A, Herrero L, Cuevas L, Perez-Pastrana E, Arnoldi D, Rosá R, Capelli G, Tenorio A, Sánchez-Seco MP, Rizzoli A.** 2015. Wide detection of Aedes flavivirus in north-eastern Italy—a European hotspot of emerging mosquito-borne diseases. J Gen Virol **96**(Pt 2):420–430. http://dx.doi.org/10.1099/vir.0.069625-0.

21. **Rizzo F, Cerutti F, Ballardini M, Mosca A, Vitale N, Radaelli MC, Desiato R, Prearo M, Pautasso A, Casalone C, Acutis P, Peletto S, Mandola ML.** 2014. Molecular characterization of flaviviruses from field-collected mosquitoes in northwestern Italy, 2011-2012. Parasit Vectors **7**:395. http://dx.doi.org/10.1186/1756-3305-7-395.

22. **Goenaga S, Fabbri CM, García JB, Rondán JC, Gardenal N, Calderón GE, Enria DA, Levis SM.** 2014. New strains of Culex flavivirus isolated in Argentina. J Med Entomol **51**:900–906. http://dx.doi.org/10.1603/ME13172.

23. **Huhtamo E, Putkuri N, Kurkela S, Manni T, Vaheri A, Vapalahti O, Uzcátegui NY.** 2009. Characterization of a novel flavivirus from mosquitoes in northern Europe that is related to mosquito-borne flaviviruses of the tropics. J Virol **83**:9532–9540. http://dx.doi.org/10.1128/JVI.00529-09.

24. **Huhtamo E, Cook S, Moureau G, Uzcátegui NY, Sironen T, Kuivanen S, Putkuri N, Kurkela S, Harbach RE, Firth AE, Vapalahti O, Gould EA, de Lamballerie X.** 2014. Novel flaviviruses from mosquitoes: mosquito-specific evolutionary lineages within the phylogenetic group of mosquito-borne flaviviruses. Virology **464-465**:320–329. http://dx.doi.org/10.1016/j.virol.2014.07.015.

25. **Wang Z, An S, Wang Y, Han Y, Guo Y.** 2009. A new virus of flavivirus: Chaoyang virus isolated in Liaoning Province. Chin J Pub Health **25**:769–772. (In Chinese.).

26. **Lee JS, Grubaugh ND, Kondig JP, Turell MJ, Kim HC, Klein TA, O'Guinn ML.** 2013. Isolation and genomic characterization of Chaoyang virus strain ROK144 from Aedes vexans nipponii from the Republic of Korea. Virology **435**:220–224. http://dx.doi.org/10.1016/j.virol.2012.10.020.

27. **Junglen S, Kopp A, Kurth A, Pauli G, Ellerbrok H, Leendertz FH.** 2009. A new flavivirus and a new vector: characterization of a novel flavivirus isolated from Uranotaenia mosquitoes from a tropical rainforest. J Virol **83**:4462–4468. http://dx.doi.org/10.1128/JVI.00014-09.

28. **Kolodziejek J, Pachler K, Bin H, Mendelson E, Shulman L, Orshan L, Nowotny N.** 2013. Barkedji virus, a novel mosquito-borne flavivirus identified in Culex perexiguus mosquitoes, Israel, 2011. J Gen Virol **94**(Pt 11):2449–2457. http://dx.doi.org/10.1099/vir.0.056200-0.

29. **Pauvolid-Corrêa A, Solberg O, Couto-Lima D, Kenney J, Serra-Freire N, Brault A, Nogueira R, Langevin S, Komar N.** 2015. Nhumirim virus, a novel flavivirus isolated from mosquitoes from the Pantanal, Brazil. Arch Virol **160**:21–27. http://dx.doi.org/10.1007/s00705-014-2219-8.

30. **Evangelista J, Cruz C, Guevara C, Astete H, Carey C, Kochel TJ, Morrison AC, Williams M, Halsey ES, Forshey BM.** 2013. Characterization of a novel flavivirus isolated from Culex (Melanoconion) ocossa mosquitoes from Iquitos, Peru. J Gen Virol **94**(Pt 6):1266–1272. http://dx.doi.org/10.1099/vir.0.050575-0.

31. **Crochu S, Cook S, Attoui H, Charrel RN, De Chesse R, Belhouchet M, Lemasson JJ, de Micco P, de Lamballerie X.** 2004. Sequences of flavi-

virus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. mosquitoes. J Gen Virol **85**:1971–1980. http://dx.doi.org/10.1099/vir.0.79850-0.

32. **Maruyama SR, Castro-Jorge LA, Ribeiro JM, Gardinassi LG, Garcia GR, Brandão LG, Rodrigues AR, Okada MI, Abrão EP, Ferreira BR, da Fonseca BA, de Miranda-Santos IK.** 2014. Characterisation of divergent flavivirus NS3 and NS5 protein sequences detected in Rhipicephalus microplus ticks from Brazil. Mem Inst Oswaldo Cruz **109**:38–50. http://dx.doi.org/10.1590/0074-0276130166.

33. **Moureau G, Cook S, Lemey P, Nougairede A, Forrester NL, Khasnatinov M, Charrel RN, Firth AE, Gould EA, de Lamballerie X.** 2015. New insights into flavivirus evolution, taxonomy and biogeographic history, extended by analysis of canonical and alternative coding sequences. PLoS One **10**:e0117849. http://dx.doi.org/10.1371/journal.pone.0117849.

34. **Moureau G, Ninove L, Izri A, Cook S, De Lamballerie X, Charrel RN.** 2010. Flavivirus RNA in phlebotomine sand flies. Vector Borne Zoonotic Dis **10**:195–197. http://dx.doi.org/10.1089/vbz.2008.0216.

35. **Galati EAB.** 2011. Phlebotominae (Diptera, Psychodidae) classificação, morfologia, terminologia e identificação de adultos. Apostila Disciplina HEP 5752 Bioecologia e Identificação de Phlebotominae, vol I. Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São, São Paulo, SP, Brazil.

36. **Young DG, Duncan MA.** 1994. Guide to the identification and geographic distribution of Lutzomyia sand flies in Mexico, the West Indies, Central and South America (Diptera: Psychodidae). Mem Ann Entomol Institut **54**:1–88.

37. **Charrel RN, Moureau G, Temmam S, Izri A, Marty P, Parola P, da Rosa AT, Tesh RB, de Lamballerie X.** 2009. Massilia virus, a novel phlebovirus (Bunyaviridae) isolated from sand flies in the Mediterranean. Vector Borne Zoonotic Dis **9**:519–530. http://dx.doi.org/10.1089/vbz.2008.0131.

38. **Moureau G, Temmam S, Gonzalez JP, Charrel RN, 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**:467–477. http://dx.doi.org/10.1089/vbz.2007.0206.

39. **Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J.** 2011. An integrated semiconductor device enabling non-optical genome sequencing. Nature **475**:348–352. http://dx.doi.org/10.1038/nature10242.

40. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948. http://dx.doi.org/10.1093/bioinformatics/btm404.

41. **Thompson JD, Higgins DG, Gibson TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673–4680. http://dx.doi.org/10.1093/nar/22.22.4673.

42. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**:1792–1797. http://dx.doi.org/10.1093/nar/gkh340.

43. **Talavera G, Castresana J.** 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol **56**:564–577. http://dx.doi.org/10.1080/10635150701472164.

44. **Huelsenbeck JP, Ronquist F.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754–755. http://dx.doi.org/10.1093/bioinformatics/17.8.754.

45. **Drummond AJ, Rambaut A.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7**:214. http://dx.doi.org/10.1186/1471-2148-7-214.

46. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28**:2731–2739. http://dx.doi.org/10.1093/molbev/msr121.

47. **Rice CM, Strauss JH.** 1990. Production of flavivirus polypeptides by proteolytic processing. Semin Virol **1:**357–367.

48. **Firth AE, Atkins JF.** 2009. A conserved predicted pseudoknot in the NS2A-encoding sequence of West Nile and Japanese encephalitis flaviviruses suggests NS1′ may derive from ribosomal frameshifting. Virol J **6:**14. http://dx.doi.org/10.1186/1743-422X-6-14.

49. **Firth AE, Blitvich BJ, Wills NM, Miller CL, Atkins JF.** 2010. Evidence for ribosomal frameshifting and a novel overlapping gene in the genomes of insect-specific flaviviruses. Virology **399:**153–166. http://dx.doi.org/10.1016/j.virol.2009.12.033.

50. **Firth AE.** 2014. Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. Nucleic Acids Res **42:**12425–12439. http://dx.doi.org/10.1093/nar/gku981.

51. **Mandl CW, Holzmann H, Kunz C, Heinz FX.** 1993. Complete genomic sequence of Powassan virus: evaluation of genetic elements in tick-borne versus mosquito-borne flaviviruses. Virology **194:**173–184. http://dx.doi.org/10.1006/viro.1993.1247.

52. **Brinton MA, Fernandez AV, Dispoto JH.** 1986. The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology **153:**113–121. http://dx.doi.org/10.1016/0042-6822(86)90012-7.

53. **Wengler G, Castle E.** 1986. Analysis of structural properties which possibly are characteristic for the 3′-terminal sequence of the genome RNA of flaviviruses. J Gen Virol **67**(Pt 6)**:**1183–1188. http://dx.doi.org/10.1099/0022-1317-67-6-1183.

54. **Nomaguchi M, Ackermann M, Yon C, You S, Padmanabhan R.** 2003. De novo synthesis of negative-strand RNA by dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. J Virol **77:**8831–8842. http://dx.doi.org/10.1128/JVI.77.16.8831-8842.2003.

55. **Khromykh AA, Kondratieva N, Sgro JY, Palmenberg A, Westaway EG.** 2003. Significance in replication of the terminal nucleotides of the flavivirus genome. J Virol **77:**10623–10629. http://dx.doi.org/10.1128/JVI.77.19.10623-10629.2003.

56. **Gritsun TS, Gould EA.** 2007. Origin and evolution of 3′UTR of flaviviruses: long direct repeats as a basis for the formation of secondary structures and their significance for virus transmission. Adv Virus Res **69:**203–248.

57. **Hahn CS, Hahn YS, Rice CM, Lee E, Dalgarno L, Strauss EG, Strauss JH.** 1987. Conserved elements in the 3′ untranslated region of flavivirus RNAs and potential cyclization sequences. J Mol Biol **198:**33–41. http://dx.doi.org/10.1016/0022-2836(87)90455-4.

58. **Grard G, Moureau G, Charrel RN, Holmes EC, Gould EA, de Lamballerie X.** 2010. Genomics and evolution of Aedes-borne flaviviruses. J Gen Virol **91**(Pt 1)**:**87–94. http://dx.doi.org/10.1099/vir.0.014506-0.

59. **Silva PA, Molenkamp R, Dalebout TJ, Charlier N, Neyts JH, Spaan WJ, Bredenbeek PJ.** 2007. Conservation of the pentanucleotide motif at the top of the yellow fever virus 17D 3′ stem-loop structure is not required for replication. J Gen Virol **88**(Pt 6)**:**1738–1747. http://dx.doi.org/10.1099/vir.0.82811-0.

60. **Elghonemy S, Davis WG, Brinton MA.** 2005. The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology **331:**238–246. http://dx.doi.org/10.1016/j.virol.2004.11.008.

61. **Tilgner M, Deas TS, Shi PY.** 2005. The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology **331:**375–386. http://dx.doi.org/10.1016/j.virol.2004.07.022.

62. **Von Heijne G.** 1984. How signal sequences maintain cleavage specificity. J Mol Biol **173:**243–251. http://dx.doi.org/10.1016/0022-2836(84)90192-X.

63. **Stadler K, Allison SL, Schalich J, Heinz FX.** 1997. Proteolytic activation of tick-borne encephalitis virus by furin. J Virol **71:**8475–8481.

64. **Steiner DF, Smeekens SP, Ohagi S, Chan SJ.** 1992. The new enzymology of precursor processing endoproteases. J Biol Chem **267:**23435–23438.

65. **Rice CM.** 1996. Flaviviridae: the viruses and their replication, p 931–959. *In* Fields BN, Knipe DM, Howley PM (ed), Fields virology, 3rd ed. Lippincott-Raven, Philadelphia, PA.

66. **Roehrig JT, Hunt AR, Johnson AJ, Hawkes RA.** 1989. Synthetic peptides derived from the deduced amino acid sequence of the E-glycoprotein of Murray Valley encephalitis virus elicit antiviral antibody. Virology **171:**49–60. http://dx.doi.org/10.1016/0042-6822(89)90509-6.

67. **De Lamballerie X, Crouchu S, Billoir F, Neyts J, de Micco P, Holmes EC, Gould EA.** 2002. Genome sequence analysis of Tamana bat virus and its relationship with the genus Flavivirus. J Gen Virol **83:**2443–2454. http://dx.doi.org/10.1099/0022-1317-83-10-2443.

68. **Medeiros DB, Nunes MR, Vasconcelos PF, Chang GJ, Kuno G.** 2007. Complete genome characterization of Rocio virus (Flavivirus: Flaviviridae), a Brazilian flavivirus isolated from a fatal case of encephalitis during an epidemic in Sao Paulo state. J Gen Virol **88**(Pt 8)**:**2237–2246. http://dx.doi.org/10.1099/vir.0.82883-0.

69. **Bazan JF, Fletterick RJ.** 1989. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. Virology **171:**637–639. http://dx.doi.org/10.1016/0042-6822(89)90639-9.

70. **Gorbalenya AE, Donchenko AP, Koonin EV, Blinov VM.** 1989. N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. Nucleic Acids Res **17:**3889–3897. http://dx.doi.org/10.1093/nar/17.10.3889.

71. **Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM.** 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res **17:**4713–4730. http://dx.doi.org/10.1093/nar/17.12.4713.

72. **Tan BH, Fu J, Sugrue RJ, Yap EH, Chan YC, Tan YH.** 1996. Recombinant dengue type 1 virus NS5 protein expressed in Escherichia coli exhibits RNA-dependent RNA polymerase activity. Virology **216:**317–325. http://dx.doi.org/10.1006/viro.1996.0067.

73. **Gritsun DJ, Jones IM, Gould EA, Gritsun TS.** 2014. Molecular archaeology of Flaviviridae untranslated regions: duplicated RNA structures in the replication enhancer of flaviviruses and pestiviruses emerged via convergent evolution. PLoS One **9:**e92056. http://dx.doi.org/10.1371/journal.pone.0092056.

74. **Gritsun TS, Gould EA.** 2007. Direct repeats in the flavivirus 3′ untranslated region; a strategy for survival in the environment? Virology **358:**258–265. http://dx.doi.org/10.1016/j.virol.2006.09.033.

75. **Gritsun TS, Gould EA.** 2006. The 3′ untranslated region of tick-borne flaviviruses originated by the duplication of long repeat sequences within the open reading frame. Virology **354:**217–223. http://dx.doi.org/10.1016/j.virol.2006.03.052.

76. **Gritsun TS, Gould EA.** 2006. Direct repeats in the 3′ untranslated regions of mosquito-borne flaviviruses: possible implications for virus transmission. J Gen Virol **87**(Pt 11)**:**3297–3305. http://dx.doi.org/10.1099/vir.0.82235-0.

77. **Billoir F, de Chesse R, Tolou H, de Micco P, Gould EA, de Lamballerie X.** 2000. Phylogeny of the genus flavivirus using complete coding sequences of arthropod-borne viruses and viruses with no known vector. J Gen Virol **81**(Pt 9)**:**2339. http://dx.doi.org/10.1099/0022-1317-81-9-2339.

78. **Pettersson JH, Fiz-Palacios O.** 2014. Dating the origin of the genus Flavivirus in the light of Beringian biogeography. J Gen Virol **95**(Pt 9)**:**1969–1982. http://dx.doi.org/10.1099/vir.0.065227-0.

79. **Firth AE, Brierley I.** 2012. Non-canonical translation in RNA viruses. J Gen Virol **93**(Pt 7)**:**1385–1409. http://dx.doi.org/10.1099/vir.0.042499-0.

80. **Melian EB, Hinzman E, Nagasaki T, Firth AE, Wills NM, Nouwens AS, Blitvich BJ, Leung J, Funk A, Atkins JF, Hall R, Khromykh AA.** 2010. NS1′ of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. J Virol **84:**1641–1647. http://dx.doi.org/10.1128/JVI.01979-09.

81. **Yamshchikov VF, Compans RW.** 1993. Regulation of the late events in flavivirus protein processing and maturation. Virology **192:**38–51. http://dx.doi.org/10.1006/viro.1993.1006.

82. **Stocks CE, Lobigs M.** 1998. Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. J Virol **72:**2141–2149.

83. **Li J, Bhuvanakantham R, Howe J, Ng ML.** 2006. The glycosylation site in the envelope protein of West Nile virus (Sarafend) plays an important role in replication and maturation processes. J Gen Virol **87:**613–622. http://dx.doi.org/10.1099/vir.0.81320-0.

84. **Beasley DW, Whiteman MC, Zhang S, Huang CY-H, Schneider BS, Smith DR, Gromowski GD, Higgs S, Kinney RM, Barrett ADT.** 2005. Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. J Virol **79:**8339–8347. http://dx.doi.org/10.1128/JVI.79.13.8339-8347.2005.

85. **Crabtree MB, Kinney RM, Miller BR.** 2005. Deglycosylation of the NS1 protein of dengue 2 virus, strain 16681: construction and characteriza-

tion of mutant viruses. Arch Virol **150:**771–786. http://dx.doi.org/10.1007/s00705-004-0430-8.

86. **Goto A, Yoshii K, Obara M, Ueki T, Mizutani T, Kariwa H, Takashima I.** 2005. Role of the N-linked glycans of the prM and E envelope proteins in tick-borne encephalitis virus particle secretion. Vaccine **23:**3043–3052. http://dx.doi.org/10.1016/j.vaccine.2004.11.068.

87. **Blaney JE, Jr, Sathe NS, Goddard L, Hanson CT, Romero TA, Hanley KA, Murphy BR, Whitehead SS.** 2008. Dengue virus type 3 vaccine candidates generated by introduction of deletions in the 3′ untranslated region (3′-UTR) or by exchange of the DENV-3 3′-UTR with that of DENV-4. Vaccine **26:**817–828. http://dx.doi.org/10.1016/j.vaccine.2007.11.082.

88. **Alvarez DE, De Lella Ezcurra AL, Fucito S, Gamarnik AV.** 2005. Role of RNA structures present at the 3′UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology **339:**200–212. http://dx.doi.org/10.1016/j.virol.2005.06.009.

89. **Bredenbeek PJ, Kooi EA, Lindenbach B, Huijkman N, Rice CM, Spaan WJ.** 2003. A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication. J Gen Virol **84:**1261–1268. http://dx.doi.org/10.1099/vir.0.18860-0.

90. **Lo MK, Tilgner M, Bernard KA, Shi PY.** 2003. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3′ untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. J Virol **77:**10004–10014. http://dx.doi.org/10.1128/JVI.77.18.10004-10014.2003.

91. **Fontenille D, Traore-Lamizana M, Trouillet J, Leclerc A, Mondo M, Ba Y, Digoutte JP, Zeller HG.** 1994. First isolations of arboviruses from phlebotomine sand flies in West Africa. Am J Trop Med Hyg **50:**570–574.

92. **Karabatsos N.** 1985. International catalogue of arthropod-borne viruses, 3rd ed, p 137–152. American Society for Tropical Medicine and Hygiene, San Antonio, TX.

93. **Traoré-Lamizana M, Fontenille D, Diallo M, Bâ Y, Zeller HG, Mondo M, Adam F, Thonon J, Maïga A.** 2001. Arbovirus surveillance from 1990 to 1995 in the Barkedji area (Ferlo) of Senegal, a possible natural focus of Rift Valley fever virus. J Med Entomol **38:**480–492. http://dx.doi.org/10.1603/0022-2585-38.4.480.

94. **Ba Y, Trouillet J, Thonnon J, Fontenille D.** 1999. Phlebotomus of Senegal: survey of the fauna in the region of Kedougou. Isolation of arbovirus. Bull Soc Pathol Exot **92:**131–135. (In French.)

95. **Konstantinov OK, Diallo SM, Inapogi AP, Ba A, Kamara SK.** 2006. The mammals of Guinea as reservoirs and carriers of arboviruses. Med Parazitol (Mosk) **2006:**34–39. (In Russian.)

96. **Heinze DM, Gould EA, Forrester NL.** 2012. Revisiting the clinal concept of evolution and dispersal for the tick-borne flaviviruses by using phylogenetic and biogeographic analyses. J Virol **86:**8663–8671. http://dx.doi.org/10.1128/JVI.01013-12.

97. **Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, Suchard MA, Foster JE, Rambaut A, Carrington CV.** 2012. Phylogeography and population dynamics of dengue viruses in the Americas. Mol Biol Evol **29:**1533–1543. http://dx.doi.org/10.1093/molbev/msr320.

98. **Bryant JE, Holmes EC, Barrett AD.** 2007. Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. PLoS Pathog **3:**e75. http://dx.doi.org/10.1371/journal.ppat.0030075.

99. **Hanley KA, Monath TP, Weaver SC, Rossi SL, Richman RL, Vasilakis N.** 2013. Fever versus fever: the role of host and vector susceptibility and interspecific competition in shaping the current and future distributions of the sylvatic cycles of dengue virus and yellow fever virus. Infect Genet Evol **19:**292–311. http://dx.doi.org/10.1016/j.meegid.2013.03.008.

100. **Mann BR, McMullen AR, Swetnam DM, Barrett AD.** 2013. Molecular epidemiology and evolution of West Nile virus in North America. Int J Environ Res Public Health **10:**5111–5129. http://dx.doi.org/10.3390/ijerph10105111.

101. **May FJ, Davis CT, Tesh RB, Barrett AD.** 2011. Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. J Virol **85:**2964–2974. http://dx.doi.org/10.1128/JVI.01963-10.

102. **Reisen WK.** 2013. Ecology of West Nile virus in North America. Viruses **5:**2079–2105. http://dx.doi.org/10.3390/v5092079.

103. **Jenkins GM, Pagel M, Gould EA, de A Zanotto PM, Holmes EC.** 2001. Evolution of base composition and codon usage bias in the genus Flavivirus. J Mol Evol **52:**383–390.

104. **Varelas-Wesley I, Calisher CH.** 1982. Antigenic relationships of flaviviruses with undetermined arthropod-borne status. Am J Trop Med Hyg **31:**1273–1284.

105. **L'vov DK, Al'khovskiĭ SV, Shchelkanov MIu, Shchetinin AM, Deriabin PG, Gitel'man AK, Samokhvalov EI, Botikov AG.** 2014. Taxonomy of the Sokuluk virus (SOKV) (Flaviviridae, Flavivirus, Entebbe bat virus group) isolated from bats (Vespertilio pipistrellus Schreber, 1774), ticks (Argasidae Koch, 1844), and birds in Kyrgyzstan. Vopr Virusol **59:**30–34. (In Russian.)

106. **Twiddy SS, Holmes EC.** 2003. The extent of homologous recombination in members of the genus Flavivirus. J Gen Virol **84**(Pt 2)**:**429–440. http://dx.doi.org/10.1099/vir.0.18660-0.

107. **Halstead SB.** 2015. Reappearance of chikungunya, formerly called dengue, in the Americas. Emerg Infect Dis **21:**557–561. http://dx.doi.org/10.3201/eid2104.141723.

108. **Price JL.** 1978. Isolation of Rio Bravo and a hitherto undescribed agent, Tamana bat virus, from insectivorous bats in Trinidad, with serological evidence of infection in bats and man. Am J Trop Med Hyg **27**(Pt 1)**:**153–161.

109. **Floch H, Chassignet R.** 1947. Phlébotomes de la Guyane Française. XXI. P. abonnenci n. sp. Publication 157. Institut Pasteur de la Guyane et du Territoire de l'Inini, Cayenne, French Guiana.

110. **Jones LA, Cohnstaedt LW, Beati L, Terán R, León R, Munstermann LE.** 2010. New records of phlebotomine sand flies (Diptera: Psychodidae) from Ecuador. Proc Entomol Soc Wash **112:**47–53. http://dx.doi.org/10.4289/0013-8797-112.1.47.