JVI Journals.ASM.org

Genome-Scale Phylogeny of the Alphavirus Genus Suggests a Marine Origin

N. L. Forrester,[a] G. Palacios,[b]* R. B. Tesh,[a] N. Savji,[b]* H. Guzman,[a] M. Sherman,[c] S. C. Weaver,[a] and W. I. Lipkin[b]

Institute for Human Infections and Immunity, Center for Biodefense and Emerging Infectious Diseases and Department of Pathology, University of Texas Medical Branch, Galveston, Texas, USA[a]; Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, New York, USA[b]; and W. M. Keck Center for Virus Imaging, University of Texas Medical Branch, Galveston, Texas, USA[c]

The genus Alphavirus comprises a diverse group of viruses, including some that cause severe disease. Using full-length sequences of all known alphaviruses, we produced a robust and comprehensive phylogeny of the Alphavirus genus, presenting a more complete evolutionary history of these viruses compared to previous studies based on partial sequences. Our phylogeny suggests the origin of the alphaviruses occurred in the southern oceans and spread equally through the Old and New World. Since lice appear to be involved in aquatic alphavirus transmission, it is possible that we are missing a louse-borne branch of the alphaviruses. Complete genome sequencing of all members of the genus also revealed conserved residues forming the structural basis of the E1 and E2 protein dimers.

Many medically important viruses are arboviruses (arthropod-borne viruses). The typical life cycle of an arbovirus involves a vertebrate host, such as a bird, rodent, amphibian, reptile, nonhuman primate, or human, and a hematophagous arthropod vector, such as a mosquito, biting fly, or tick. Therefore, maintenance of arbovirus fitness to infect both the vertebrate host and arthropod vector is required, leading to complex evolutionary constraints.

The alphaviruses are a diverse group of small, spherical, enveloped viruses with single-stranded, positive-sense, RNA genomes and have been isolated from all continents except Antarctica (see Table 1). They belong to the family *Togaviridae*, and include 29 recognized species (80). Their genomes contain two open reading frames (ORFs): one flanked by a 5′ cap and an untranslated region that encodes the nonstructural proteins and one controlled by a subgenomic promoter that encodes the structural proteins (71). The four nonstructural proteins produced, nsP1 to nsP4, are involved in RNA replication and modification and in proteolytic cleavage. A leaky opal stop codon near the 3′ end of the nsP3 gene is present in the genomes of most but not all alphaviruses (42, 51), such that two products, P123 and P1234, are produced during translation (63, 71). The second polyprotein encodes the structural proteins, including the capsid protein, two major envelope proteins (E2 and E1), and two smaller structural proteins not usually found in virions (23, 71).

Alphaviruses are transmitted by mosquitoes with two exceptions: salmon pancreatic disease virus (SPDV) and its subtype sleeping disease virus (SDV), which infect salmon and trout, causing mortality in farmed fish (82, 83), and Southern elephant seal virus (SESV). For both of these viruses the presence of the virus within lice *Lepeophtheirus salmonus* for SPDV and *Lepidohthirus macrorhini* for SESV (40) suggests an arthropod-borne cycle, but the vector has yet to be incriminated.

Many of the remaining pathogenic alphaviruses cause acute, febrile illness in humans and/or domestic animals that culminates either in encephalitis or arthralgia/arthritis. However, some alphaviruses that circulate enzootically are not known to cause disease. Most of these were first isolated during mosquito surveillance, and for many the transmission cycle remains enigmatic. These include Trocara virus (TROV) and Aura virus (AURAV) (80). Among the New World encephalitic alphaviruses, the western equine encephalitis (WEE) complex arose from a rare recombination event among arboviruses resulting in the virulent important human and veterinary pathogen, WEE virus (WEEV) (30, 81), as well as other viruses not incriminated in human disease. Among the Old World arthralgic alphaviruses of the Semliki Forest complex, the recently emerged Chikungunya virus (CHIKV) is the most important, causing disease in millions of people in Africa, Asia, and parts of Europe (22, 78). It is the only alphavirus to emerge into an urban or peridomestic cycle, where the virus is transmitted by anthrophilic mosquitoes from human-to-human with no involvement of wild animals as amplification or reservoir hosts. Among this group of viruses in the Semliki Forest complex, some such as Una virus (UNAV) and Getah virus (GETV), cause little or no human disease but do cause disease in horses (15, 24).

Previous attempts to understand the evolutionary history of the alphaviruses relied on partial E1 gene sequences (57) or a partial set of complete genomes (45). To better understand the evolution of the alphaviruses, we conducted a more comprehensive phylogenetic analysis using complete genomic sequences for all known members of the genus. We first sequenced the eight missing genomes: Bebaru virus (BEBV), Buggy Creek virus (BCRV), Ndumu virus (NDUV), Sindbis virus strain Babanki (SINV-Babanki), Southern elephant seal virus (SESV), Trocara virus (TROV), Una virus (UNAV), and Whataroa virus (WHAV).

Received 1 July 2011 **Accepted** 12 December 2011

Published ahead of print 21 December 2011

Address correspondence to N. L. Forrester, naforres@utmb.edu.

* Present address: G. Palacios, School of Medicine, New York University, New York, New York, USA; and N. Savji, U.S. Army Medical Institute for Infectious Diseases, Fort Detrick, Maryland, USA.

N. L. Forrester and G. Palacios contributed equally to this article.

Supplemental material for this article may be found at http://jvi.asm.org/.

doi:10.1128/JVI.05591-11

We then conducted detailed phylogenetic analyses to estimate the origins and patterns of evolution of the alphaviruses. In addition, given that recent studies have resolved the structural proteins of the alphaviruses to atomic resolution (67, 79), we used these to visualize the conserved residues of the structural proteins of the alphaviruses.

MATERIALS AND METHODS

Viruses. Viruses were obtained from the World Reference Center for Emerging Viruses and Arboviruses at the University of Texas Medical Branch. Table 1 provides the names, strain numbers, sources, dates, and locality of isolation, as well as the GenBank accession numbers, for the eight newly sequenced genomes described in the present study.

Genome sequencing. RNA was extracted from virus stocks using TRIzol LS (Invitrogen, Carlsbad, CA) and treated with DNase I (DNA-Free; Ambion, Austin, TX). cDNA was generated using the Superscript II system (Invitrogen) using random hexamers linked to an arbitrary 17-mer primer sequence (53). Resulting cDNA was treated with RNase H and then randomly amplified by PCR with a 9:1 mixture of primer corresponding to the 17-mer sequence and the random hexamer linked 17-mer primer (53). Products greater than 70 bp were selected by column chromatography (MinElute; Qiagen, Hilden, Germany) and ligated to specific adapters for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT) without fragmentation (13, 47, 52). Software programs accessible through the analysis applications at the GreenePortal website (http://tako.cpmc.columbia.edu/tools/) were used for removal of primer sequences, redundancy filtering, and sequence assembly. Sequence gaps were completed by reverse transcription-PCR amplification using primers based on pyrosequencing data. Amplification products were size fractionated on 1% agarose gels, purified (MinElute), and directly sequenced in both directions with ABI Prism BigDye Terminator 1.1 cycle sequencing kits on ABI Prism 3700 DNA analyzers (Perkin-Elmer Applied Biosystems, Foster City, CA). The terminal sequences for each virus were amplified using the Clontech Smarter RACE kit (Clontech, Mountain View, CA). Sequences of the genomes were verified by Sanger dideoxy sequencing using primers designed from the draft sequence to create products of 1,000 bp with 500-bp overlaps.

Phylogenetic analysis. The remaining genomic alphavirus sequences were downloaded from GenBank (see Table 1) and aligned in SeaView (27) using the MUSCLE algorithm (20). Sequences were aligned as deduced amino acids (aa) from ORFs and then returned to nucleotide sequences for most analyses. The two ORFs were concatenated, and the C terminus of the nsP3 and the N terminus of the capsid sequences, which do not produce reliable alignments due to numerous insertions and deletions and extensive sequence divergence, were removed to increase the reliability of the analysis. After manual adjustments, the complete alignment was split into nonstructural and structural protein ORFs.

Phylogenetic analyses were undertaken using PAUP* version 4.0,10b (72). The optimal evolutionary model for each data set was estimated from 56 models implemented using Modeltest version 3.06 (56). An optimal maximum-likelihood (ML) tree was then estimated using the appropriate model and a heuristic search with tree-bisection-reconstruction branch swapping and 10 replicates, estimating variable parameters from the data, where necessary. Bootstrap replicates were calculated for each data set under the same models, but using GARLI (86) due to computational constraints. A neighbor-joining tree was also generated with PAUP* utilizing the p-distance algorithm. Bayesian analysis was undertaken using MrBayes v3.1 (33, 59), and data sets were run for two million generations (structural ORF) or four million generations (nonstructural ORF and full-length data sets) until they reached congruence. The model used was the GTR+I+G model.

Maximum-parsimony and neighbor-joining trees were also generated with PAUP* using deduced amino acid sequences with the default settings and 10 replicates. Homoplasy indices were calculated for the maximum-parsimony trees using the default settings in PAUP*. The trees generated were compared by using the Kishino-Hasegawa (KH) test implemented in PAUP* to determine the most probable topology of the trees.

Recombination analysis was carried out using RAT (21) utilizing a subset of sequences, including the WEEV group of viruses (WEEV, WEEVAg80, BCRV, FMV, and HJV), the SINV group of viruses (SINV, SINV-Babanki, and SINV-Ock/Eds), and the EEEV group of viruses (NAEEEV, SAEEEV LinII, SAEEEV LinIII, and SAEEEV LinIV).

Analysis of conserved residues. Amino acid sequence alignments were constructed as described above, and conserved residues not previously described as functionally important were identified. The E1-E2 heterodimer structure (chains F and G) of CHIKV was extracted from PDB 2XFB and fitted into a cryo-electron microscopy (cryo-EM) map of WEEV (67) using Chimera (55). The same program was used to prepare Fig. 4.

GenBank accession numbers. The complete genomic sequencing of the eight additional alphaviruses described here are available from GenBank under accession numbers HM147984, HM147985, HM147986, HM147988, HM147989, HM147990, HM147991, and HM147993.

RESULTS

The complete genomic sequencing of the eight additional alphaviruses described above allowed us to generate the phylogenetic history of the Alphavirus genus using Bayesian, ML, and distance methods at the nucleotide and amino acid levels. Three different trees (from different data sets) were produced using different alignments: (i) full-length genomes, excluding the C terminus of nsP3 and the N terminus of the capsid genes and also excluding the recombinant WEE-like viruses, which would have confounded the result; (ii) the nonstructural polyprotein (excluding the C terminus of nsP3); and (iii) the E2-6K-E1 region to reflect the recombination event in the WEE complex. Because the methods that produce the phylogenetic trees rely on determining the similarity among the sequences, including a virus whose genes represent sisters with those of two distinct viruses from two different groups creates error within the analysis. Thus, to avoid introducing known error into the analysis, the recombinant viruses known to be present in the WEE complex were removed from the full-length analysis. Instead, we partitioned the genome into the nonstructural and structural genomes so that the recombinant viruses could be included in the analysis without confounding the results.

Full-length genome trees. The full-length genome trees showed strong groupings for all four subgroups previously classified as antigenic complexes based on serological cross-reactivity (see Table 1): the Semliki Forest complex and the Sindbis virus-like members of the WEE complex, as well as the VEE and EEE complexes, in all analyses. When likelihood-ratio tests were performed, the Bayesian analysis and the ML analysis produced the most robust trees ($P < 0.001$) (For the full results of the KH analysis, see Table S1 in the supplemental material.) The only difference was the placement of SESV, which had no bootstrap or posterior probability support in either topology. The Bayesian midpoint rooted tree is shown in Fig. 1. The long branch length leading to SPDV suggests that it is an appropriate outgroup for the terrestrial vertebrate alphaviruses. The only other virus in this analysis without a mosquito vector, but which may be transmitted by a louse (*Lepidophthirus macrorhini*), SESV (40) was outside the major groups of the phylogeny, along with SPDV.

BFV was placed at the basal position among the Old World arthralgic viruses, which include the BF, NDU, MID, and SF complexes (see Table 1), with NDUV and MIDV branching off in

TABLE 1 Alphavirus genus, the nomenclature, host and vector, and geographic range[a]

Virus	Antigenic complex	Subtype	Abbreviation	First isolation	Host range	Mosquito vector	Human disease	Geographic range	Source or reference	Accession no.
Salmon pancreatic disease			SPDV		Salmon (*Salmo salar*)	Unknown		Aquatic	83	AJ316246
Southern elephant seal			SESV	Australia, 2001	Seal (*Mirounga leonine*)	*Lepidohthirus macrolini*		Aquatic	40	This study
Barmah Forest	BF		BFV	Australia, 1974		*Culex* and *Aedes* spp.	Yes	Australia	14	U73745
Middelburg	MID		MIDV	South Africa, 1957		*Aedes* spp.	Yes	Africa	36	EF536323
Ndumu	NDU		NDUV	South Africa, 1959		*Aedes* spp.	Yes	Africa	37	This study
Sagiyama	SF		SAGV	Japan, 1956		*Culex* and *Aedes* spp.		Japan	65, 66	AB032553
Getah	SF		GETV	Malaysia, 1955	Cattle and horses	*Culex* and *Aedes* spp.		Australasia, Asia	17, 18	AY702913
Ross River	SF		RRV	Australia, 1959	Rodents	*Culex* and *Aedes* spp.	Yes	Australasia	17, 25	GQ433358
Bebaru	SF		BEBV	Malaysia, 1956		*Culex* spp.		Malaysia	75	This study
Semliki Forest	SF		SFV	Uganda, 1942				Africa	69, 70	X04129
Mayaro	SF		MAYV	Trinidad, 1954		*Haemogogus* spp.	Yes	South America and Caribbean	1, 9, 11	This study
Una	SF		UNAV	Brazil, 1959		*Psorophora* and *Aedes* spp.		South America	10	This study
Chikungunya	SF		CHIKV	Tanzania, 1953	Humans, nonhuman primates	*Aedes aegypti* and *Aedes albopictus*	Yes	Africa, Indian Ocean, Asia	46, 58, 60	AF369024
O'nyong nyong	SF		ONNV	Uganda, 1959	Humans	*Anopheles* spp.	Yes	Africa	84	AF079456
Venezuelan equine encephalitis	VEE	IAB	VEEV	Venezuela, 1938	Rodents, horses, humans	*Aedes*, *Culex*, and *Psorophora* spp.	Yes	South and Central America	38, 39	AF069903
	VEE	IC								L04653
	VEE	ID								L00930
	VEE	IE								AY823299
78V-3531	VEE	IF							7	AF075257
Everglades	VEE	II	EVEV	Florida, USA, 1963	Birds	*Culex* (melanoconion) spp.		South Florida, USA	4, 12	AF075251
Tonate	VEE	IIIA	TONV	French Guiana, 1973				South and Central America	16	AF075254
Mucambo	VEE	IIIB	MUCV	Brazil, 1954		*Culex* spp.		South America and Carribbean	68	AF075253
71D1252	VEE	IIIC							64	AF075255
Pixuna	VEE	IV	PIXV	Brazil, 1961				South America	68	AF075256
Cabassou	VEE	V	CABV	French Guiana, 1968		*Culex portesi*		French Guiana	16, 35	AF075259
Rio Negro (Ag80-663)	VEE	VI	RNV			*Culex* spp.		South America	50	AF075258
78V-3531	VEE	IF							7	AF075257
Eastern equine encephalitis	EEV	Lin I (NA)	EEEV	Maryland, USA, 1933	Birds, rodents	*Culex melanura*	Yes	North and South America	48, 74	EF151502
	EEE	Lin II (SA)				*Culex* and *Aedes* spp.				DQ241303
	EEE	Lin III (SA)								DQ241304
	EEE	Lin IV (SA)								EF151503
Aura	WEE		AURAV	Brazil, 1959		*Culex* (melanoconion) spp.		South America	10	AF126284
Buggy Creek	WEE		BCRV	Oklahoma, USA, 1980	Cliff sparrow (*Petrochelidon pyrrhonota*)	*Oeciacus vicarious*		North America	32, 44	This study
Fort Morgan	WEE		FMV	Colorado, USA, 1973	Birds	*Oeciacus vicarious*		North America	8	GQ281603
Highlands J	WEE		HJV	Florida, USA, 1960	Birds	*Culex melanura*		North America	31	GU167952
Sindbis	WEE		SINDV	Egypt, 1952	Birds	*Culex* and *Aedes* spp.		Global	34, 73	AF429428
	WEE	Ockelsbo								M69205
	WEE	Babanki								This study
Trocara	WEE		TROV	Brazil, 1984		*Culex serratus*		Brazil	76	This study
Western equine encephalitis	WEE	NA	WEEV	California, USA, 1930	Birds	*Aedes* and *Culex* spp.	Yes	North and South America	49	AF214040
	WEE	SA			Rodents					GQ287646
Whataroa	WEE		WHAV	New Zealand, 1962	Birds	*Culiseta* and *Culex* spp.		New Zealand	61	This study

[a] The table also includes the original source reference(s) and the accession numbers used in this study. For some viruses, the vertebrate host or the mosquito vector has not been incriminated, and therefore these cells are blank. Viruses have been ordered via antigenic complex.

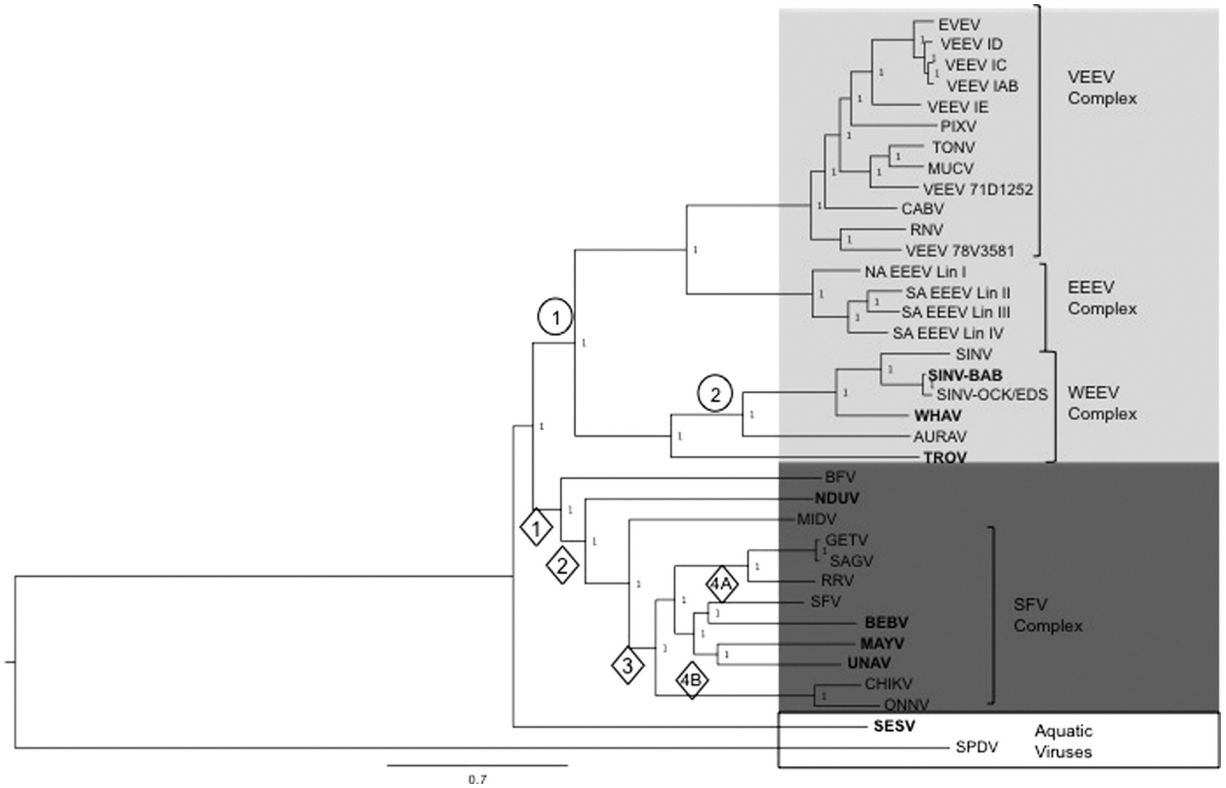

FIG 1 Phylogenetic tree produced using Bayesian methods and rooted using the midpoint. The tree includes representatives from all species of the alphaviruses, except the WEEV complex, using the full-genome alignment of both ORFs, excluding portions of the nsP3 and capsid that do not produce significant alignments due to frequent indels. Viruses sequenced for the present study are indicated in boldface. The light gray shading indicates viruses classified as New World alphaviruses, while dark gray shading indicates those classified as Old World viruses, and the open box signifies the aquatic alphaviruses. It should be noted that the Old and New World designation refers to the geographical placement of the majority of the viruses within the group, although representatives of the New World alphaviruses are found in the Old World and vice versa. Posterior probabilities are shown on major branches. The recombinant WEEV complex alphaviruses were excluded to prevent bias. (Likelihood scores for both Bayesian and ML trees were −ln L 277174.42319).

subsequent order. Within the Semliki Forest complex, RRV, GETV, and SAGV comprised a clade, as did SFV, BEBV, MAYV, and UNAV. SFV and BEBV consistently grouped together, as did MAYV and UNAV, presumably reflecting their common geographic distribution and divergence from an ancestor introduced into the tropics. The remaining groupings were expected; the VEE complex was monophyletic, and the EEEV and SINV groups formed clades as expected. All trees supported two well-defined monophyletic groups: the SINV and VEEV/EEEV clades (the encephalitic New World group) as one group and the SF, MID, NDU, and BF complexes (the arthralgic Old World group) as the other. As expected, based on antigenic (6) and previously determined genetic similarities (57), WHAV fell within the SINV-like clade of nonrecombinant viruses in the WEE antigenic complex, and TROCV was basal to WHATV.

Nonstructural protein ORF trees. As with the genomic trees, the two methods giving the best topology as identified by the KH test were the Bayesian and the ML analyses. Figure 2B shows the midpoint rooted Bayesian analysis. Both generated trees with the same overall topology. However, in both analyses, support for some of the groups was weak. This lack of resolution was also seen in trees constructed using amino acid sequences (data not shown). Similarly to the full-length trees, the placement of SESV was not well supported (<50% bootstrap and 0.86 posterior probability); it was basal to the SFV group, rather than to the entire terrestrial alphavirus clade. Interestingly, although MAYV and UNAV grouped together in the analyses of the full-length genome, UNAV grouped with SFV and BEBV in nonstructural ORF trees, albeit with low bootstrap values under ML conditions. In fact, the MAYV-UNAV-SFV-BEBV clade was the only one well supported by ML, with a bootstrap value of 77. The Bayesian analysis did not support the UNAV and SFV grouping but gave strong support to BEBV-UNAV-SFV groupings.

The nonstructural tree still supported two well-defined monophyletic groups: the encephalitic group and the arthralgic group. As expected, the WEEV-like recombinant group appeared as a sister to the EEEV group, and this topology was supported by strong bootstrap and posterior probability values. However, the topologies of the WEE and SF complexes were different from those depicted in the full-genome trees, although not well supported by bootstrap or posterior values.

Envelope protein gene trees. Topologically similar trees were generated using the E2-6K-E1 genome region. Based on the KH likelihood test, the best trees were again those generated using Bayesian and ML methods, with nearly identical branching patterns. The Bayesian tree is shown in Fig. 2A. The only difference between these two trees was that the Bayesian tree showed a polytomy comprised of the SFV/EEE/VEE/WEE complexes and SESV, whereas the ML tree placed SESV as basal to the SF complex. However, neither topology showed high posterior or bootstrap

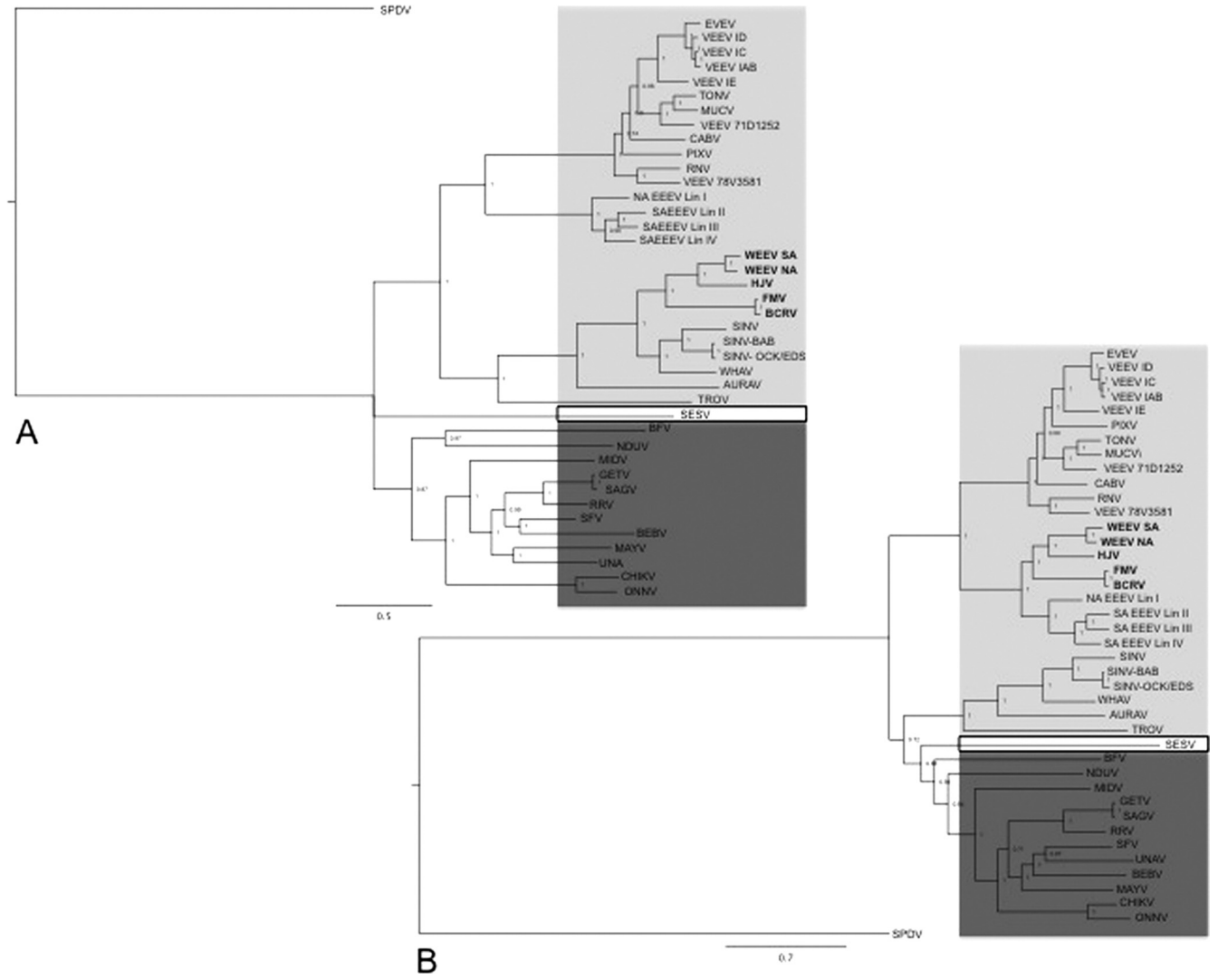

FIG 2 Phylogenetic trees produced using Bayesian methods, the trees were rooted using the midpoint. The trees include representatives from all species of the alphaviruses with the structural proteins comprising E2, 6K, and E1 proteins (the likelihood scores for the Bayesian and ML trees were −ln L 87879.75390 and −ln L 87872.05667, respectively) (A) and the nonstructural proteins excluding regions of the nsP3 (the likelihood score for both Bayesian and ML trees was −ln L 198542.90397) (B). The recombinant alphaviruses are highlighted in boldface.

support, and therefore the placement of SESV should be considered unresolved. Interestingly, the ML tree could not resolve MIDV, NDUV, and BFV, while the Bayesian tree showed high confidence for the NDUV-BFV grouping but low support for the placement of viruses basal to them. MAYV and UNAV grouped together, as did SFV and BEBV in the full-length trees. SFV and BEBV showed a closer relationship to RRV, GETV, and SAGV, a grouping that was unique to the envelope protein tree. The Bayesian tree had high support for this topology, whereas the ML tree showed weaker bootstrap support for the SFV-BEBV grouping, and none at all for the grouping of SFV and BEBV with RRV.

The WEEV-like recombinant viruses grouped, including FMV, BCV, and HJV, with the SINV-like group within the WEE complex, an observation consistent with the previously described origins of the envelope proteins of the recombinant ancestor from a SINV-like ancestor (81). This grouping had strong support both from posterior probabilities and bootstrap values generated by the Bayesian and ML analyses, respectively. Recombination analysis of the coding regions gave approximately the same breakpoint for Fort Morgan, Highlands J, Buggy Creek, and Western equine encephalitis viruses. We were only able to place the recombination event in the capsid protein as it was not possible to generate a robust alignment of the 3′ untranslated region. The breakpoint occurred somewhere between amino acids (aa) 293 and 328 of the WEEV structural protein. This corresponds to the C terminus of the E3 and the N terminus of the E2. However, there was a discrepancy of ~40 aa between the viruses making up the recombinant clade, with HJV breakpoint occurring between aa 293 and 300 and BCRV occurring between aa 321 and 328 of the WEEV structural proteins. FMV, WEEV NA, and WEEV SA fell somewhere between these two extremes.

Conserved envelope protein residues. To visualize conserved amino acid residues in the E1 and E2 glycoproteins, we used as a template a three-dimensional map of WEEV at 13-Å resolution

obtained from a cryo-EM reconstruction (67). Part of the map containing a trimeric E1-E2 spike at the 3-fold axis was computationally extracted from the whole map of WEEV, and the Chikungunya virus E1-E2 X-ray structure (PDB code 2XFB, chains F and G [79]) was fitted into it using Chimera (55). The overall fit of the E1-E2 heterodimer into the spike in the map is shown in Fig. 4B and C. Conserved residues in E1 are shown in green, and those conserved in the E2 sequence are shown in cyan in Fig. 4D and E. Strikingly, most of the conserved residues in both the E1 and the E2 proteins are positioned close to one another in adjacent beta sheets (e.g., Tyr-46 and Ala-121, Thr-48 and Ala-119, etc. [Fig. 4E]) and alpha-helices (Gly 239 and Trp 243) in domain II of E1, suggesting their evolutionary co-conservation to maintain the protein fold. Trp-89 is obviously a very important residue in the fusion loop of E1; it is inserted into a cleft in E2 and interacts with the domain B of E2 (79). The same pattern continues in E2 (e.g., Glu-35 and Gln-49 in domain A, etc.). Some of the conserved residues participate in the E1-E2 interactions both within the heterodimer and within the spike.

DISCUSSION

We generated the complete sequences of eight alphaviruses for which full-genome sequences were previously unavailable. Our completion of the full-length sequences of the Alphavirus genus allows the generation of a robust, comprehensive phylogenetic analysis of the entire genus. In particular, the phylogenetic placement of UNAV was of interest since previous analyses had given conflicting results as to the placement of UNAV within the SF complex. This analysis also included SESV an ecologically novel alphavirus, which had previously not been included in any full-length analysis. The full-length sequences allowed us to rule out the possibility of additional recombination events (other than the ancestor of WEEV) in other genome regions and/or involving other alphavirus species.

Our phylogenetic analysis based on genomic sequences increased the accuracy and reliability of phylogenetic trees depicting the evolution of the genus, compared to previous studies based on partial genomic sequences (57) or incomplete representation where only 20 out of 29 recognized species were included (45). The latter study was also inappropriate in its use of RUBV, which has no detectable sequence homology to the alphaviruses (19), as an outgroup to root trees (45). Although RUBV is a sister virus to the alphaviruses based on its overall genome organization and virion morphology (80), the extensive divergence between the genera of the *Togaviridae* means that there remains no significant sequence homology remaining between the genera. Also, an apparent rearrangement of the nsP2 and nsP3-like genes leaves the RUBV and alphavirus nonstructural ORFs without the same order of genes sharing functions (19).

In previous phylogenies, MIDV has been placed within the SFV complex (57). Our data suggest that the Middelburg complex (and thus MIDV) sit basal to the SFV complex, a finding more consistent with serological relationships. Our analysis also placed UNAV/MAYV and SFV/BEBV as sisters and members of a strongly supported clade. The sister grouping of MAYV and UNAV is the most parsimonious considering their geographical distributions, unique within the SF complex, in the New World. This grouping is different from the one observed in previous analysis using partial E1 envelope glycoprotein sequences (57), where BEBV and SFV were grouped with RRV, GETV, and SAGV. Interestingly, we also observed some inconsistencies in the internal branching of this clade among different tree topologies, depending on the genome area utilized. A recombination analysis did not show evidence of any such process to explain these inconsistencies (data not shown) and additional analysis of maximum-parsimony trees did not show significant differences in the homoplasy indices (see Materials and Methods).

Based on our analysis, we propose a revised evolutionary history for the alphaviruses. A New or Old World origin for the alphaviruses, with the alphaviruses emerging as encephalitides in the New World and moving to the Old World, or emerging as arthalgic viruses in the Old World and transitioning to the New World, had been proposed without clear evidence for either; both require numerous reintroductions to give the extant geographical distribution of the alphaviruses (57). However, an alternative explanation for the origin of the alphaviruses could involve a Pacific emergence from marine to terrestrial vertebrate hosts and to mosquito vectors. Although we hypothesize a Pacific emergence based on the extant geographical locations of viruses such as BFV and VEEV/EEEV, given the range of the alphaviruses, this could have occurred in any ocean. After emergence into terrestrial hosts, subsequent movements both east and west would result in the Old and New World ancestors of the mosquito-borne viruses. This scenario would also require subsequent reintroductions between the hemispheres (Fig. 3). The presence of the aquatic alphaviruses at basal positions in our trees when defined by midpoint rooting suggests that these viruses may be ancestral. We recognize that the correct rooting of our trees is not certain and could be influenced by variable evolutionary rates among alphavirus lineages, as proposed previously (3, 78), as well as by the highly diverse hosts and environments in which alphaviruses circulate. The placement of the aquatic virus SESV, which was isolated from the seal louse, *L. macrorhini*, also reinforces this hypothesis. Although we could not identify a robust placement for this virus within the alphavirus phylogeny, it clearly diverged from the mosquito-borne viruses in the distant past. We recognize that further study of the aquatic ecosystems, and identification of additional alphaviruses is needed to more conclusively determine the origins of the alphaviruses. A comprehensive study of aquatic invertebrates would undoubtedly reveal many new viruses, although in the absence of disease these investigations are unlikely to occur. The identification of SPDV as a pathogen of farmed fish was the major reason for its discovery. The retention by at least some of the New World alphaviruses of the ability to replicate in fish cells and at lower temperatures, such as those found in aquatic habitats (54, 85) also supports an ancestral aquatic habitat. Further experiments to determine the ability of the nonaquatic alphaviruses to infect fish would enable this to be verified more fully, still it is additional evidence that the alphaviruses secondarily acquired their ability to infect warm-blooded vertebrates and mosquito vectors.

A notable evolutionary trait of alphaviruses is their ability to move across continents and colonize new areas. All hypothetical scenarios for the origin of alphaviruses require repeated movement across the globe to explain the distributions observed today. Assuming that most of the global movement of ancestral alphaviruses occurred before the age of frequent human transoceanic travel, it seems likely that zoonotic hosts were responsible for the alphavirus movement depicted in our phylogenies. Birds are the most obviously mobile hosts. However, many of the alphaviruses that are found closer to the root of the tree, such as AURAV and

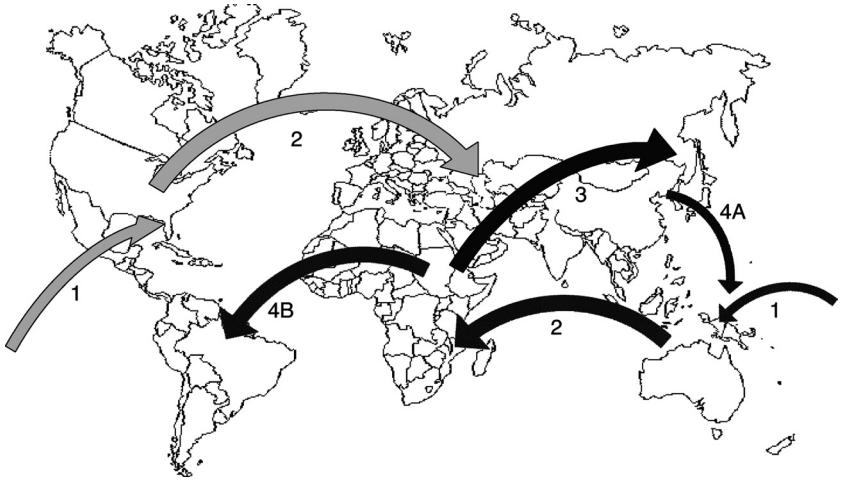

FIG 3 Diagram showing a hypothetical origin of the alphaviruses. New World alphaviruses are indicated by gray arrows: arrow 1, introduction from Oceania to the New World; and arrow 2, secondary introduction to the Old World. Old World viruses are indicated by black arrows: arrow 1, introduction from Oceania to Australasia; arrow 2, secondary introduction into Southern Africa; arrow 3, tertiary introduction to Northern Africa and Eurasia; arrow 4A, secondary introduction of RRV to Australasia; and arrow 4B, secondary introduction of MAYV and UNAV to the New World.

TROCV, do not have a vertebrate host identified. Thus, it is difficult to determine whether birds were solely responsible for this transfer of viruses over large distances, it is possible that arthropod vectors could be an alternative vehicle for transfer.

The presence of aquatic viruses in our alphavirus phylogenies and the limited surveillance for nonhuman pathogens suggests that there may be many undiscovered alphaviruses transmitted by lice. Recent work has shown that the number of viruses within the oceans is far larger than only recently could have been imagined (41). It is therefore probable that many alphaviruses like the louse-borne SESV remain to be discovered.

The alphaviruses and the flaviviruses, many of which share vector-borne transmission, nevertheless have major differences in their evolutionary histories. In general, the flaviviruses exhibit clear evolutionary associations with particular groups (26), with a few exceptions, and can be subdivided into the tick-borne (29) and mosquito-borne (28) groups. In contrast, the alphaviruses do not show obvious vector-virus relationships. Instead, within each alphavirus group are viruses transmitted primarily by *Aedes* and *Culex* species, as well as many other mosquito genera. The WEEV-like group is particularly diverse in its vector usage, with WEEV transmitted by *Culex tarsalis*, while the closely related BCRV and FMV are transmitted by the nest bug, *Oeciacus vicarius*. Moreover, many alphaviruses use multiple, taxonomically diverse mosquito vectors for transmission. This suggests that the alphaviruses are more promiscuous in their ability to adapt to new vectors and hosts than the flaviviruses. It is likely that this promiscuity has facilitated the ability of alphaviruses to traverse oceans and continents. We suspect that numerous geographic introductions and reintroductions of these alphaviruses are undetected by our phylogenetic methods due to incomplete sampling, as well as extinctions of ancestral lineages. The propensity of alphaviruses to spread and change their host range underscores their potential as emerging and reemerging pathogens. Individual mutations that mediate important host range changes have been linked to of the reemergence of CHIKV (77) and VEEV (2, 5).

One of the main differences between the alphaviruses and the flaviviruses is the presence of a subgenomic promoter within the alphaviruses. Whereas the flaviviruses exist as a single polyprotein the alphaviruses replicate using two polyproteins. It is possible that this difference in genome increases the ability of the alphaviruses to change host range and vector with greater frequency. The structural proteins are under the control of a subgenomic promoter, which increases the number of copies that is produced. Although some Alphaviruses package this sgRNA into the virion, this does not occur in all alphaviruses (62). Moreover, the sgRNA is not involved in replication and therefore cannot be responsible for the added plasticity. However, it is possible that the shorter polyproteins results in less defective mRNAs produced than the larger single polyprotein of the flaviviruses, resulting in the potential for more mutations to be incorporated.

Our alignments of the structural proteins allowed the identification of numerous conserved residues of no known function within the envelope proteins of the alphaviruses. These conserved residues are paired throughout the dimer indicating they are structurally conserved to maintain the folds and stability of the envelope dimer. Interestingly, we identified a conserved residue in the fusion loop, *viz.*, TRP89 (Fig. 4). The recent resolution of Alphavirus structure at both neutral and low pH demonstrates the importance of the fusion loop and its structure (43, 79). The fusion loop becomes exposed at low pH within endosomes, and we suggest that the interaction of this TRP89 with the domain B of the E2 becomes disrupted and allows the conformation transformation that results in virus fusion and entry into the cell. Further mutagenesis studies are required to determine how important this residue is and whether it interacts with domain B of the E2 protein.

In summary, we have produced a comprehensive alphavirus phylogeny using complete genomic sequences from all of the known members of the genus. This phylogeny has resolved some previous issues such as the placing of MIDV and NDUV, and the grouping of SFV, BEBV, UNAV, and MAYV. It also allows us to propose an alternative hypothesis for the aquatic origins of the genus. Improved understanding of the underlying relationships

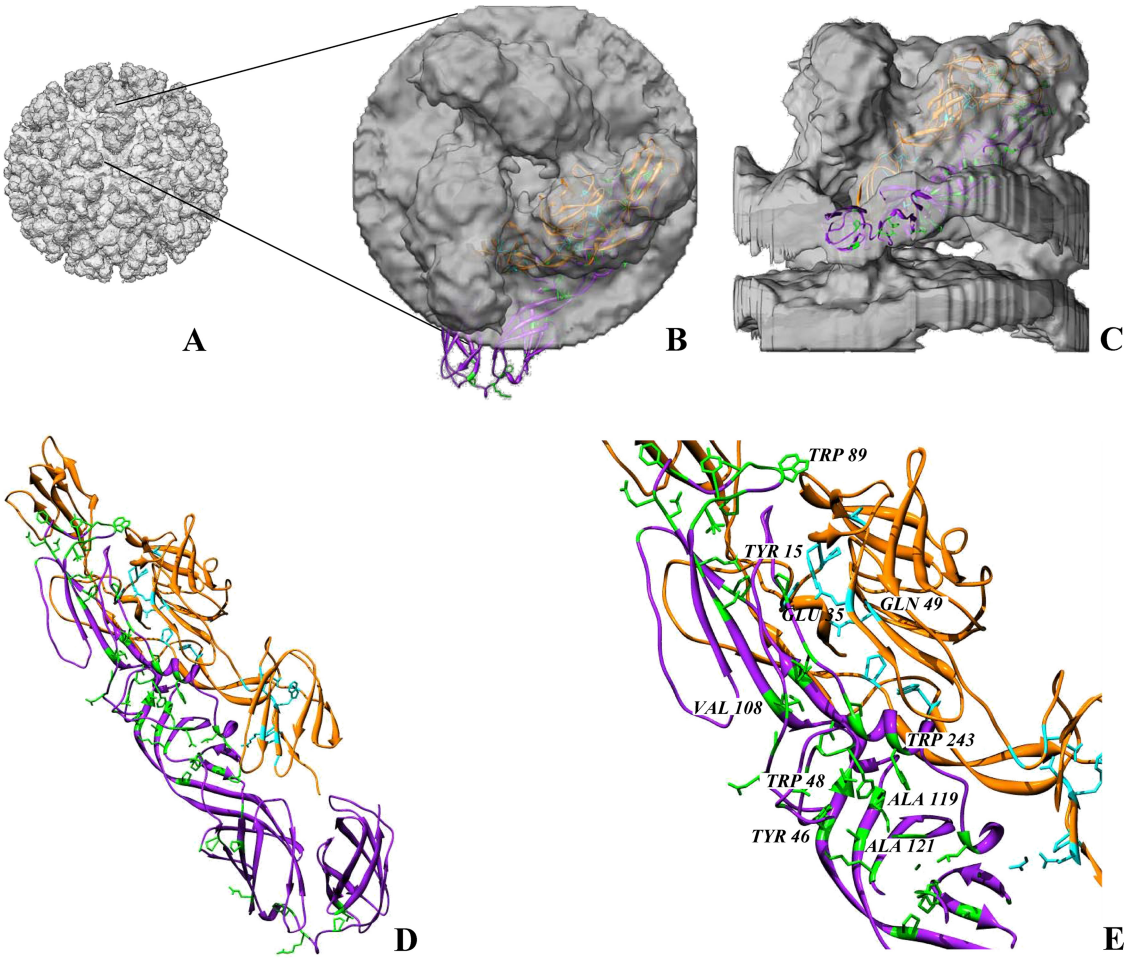

FIG 4 Chikungunya E1-E2 heterodimer (PDB code 2XFB, chains F and G [79]) fitted into cryo-EM map of WEEV (67). (A) Three-dimensional WEEV cryo-EM map showing E1-E2 spikes on the surface of the virus. (B) A spike from the map with trimer of E1-E2 heterodimers fitted into the density. Part of the E1 is sticking out from the density owing to restricted size of the spike density cut out from the WEEV map. (C) Side view of the spike with a single E1-E2 heterodimer fitted into cryo-EM density. (D) E1-E2 structure (79) rotated by 180° relative to the orientation in panel C with the conserved residues shown in green (E1) and in cyan (E2). (E) Close-up view of a fragment of the E1-E2 heterodimer as in panel D, with some of the conserved residues labeled. See the details in the text.

among alphaviruses may facilitate the identification of potential threats prior to the emergence of new arboviral diseases.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health (NIH) grant AI069145 and NIH contract HHSN27220100004OI/HHSN27200004/D04 and by the John S. Dunn Foundation. G.P., N.S., and W.I.L. were supported by AI57158 (Northeast Biodefense Center-Lipkin), AI079231, AI070411, USAID PREDICT, and the U.S. Department of Defense.

REFERENCES

1. **Anderson CR, Downs WG, Wattley GH, Ahin NW, Reese AA.** 1957. Mayaro virus: a new human disease agent. II. Isolation from blood of patients in Trinidad, B. W. I. Am. J. Trop. Med. Hyg. **6**:1012–1016.
2. **Anishchenko M, et al.** 2006. Venezuelan encephalitis emergence mediated by a phylogenetically predicted viral mutation. Proc. Natl. Acad. Sci. U. S. A. **103**:4994–4999.
3. **Arrigo NC, Adams AP, Weaver SC.** 2010. Evolutionary patterns of eastern equine encephalitis virus in North versus South America suggest ecological differences and taxonomic revision. J. Virol. **84**:1014–1025.
4. **Bigler WJ, Ventura AK, Lewis AL, Wellings FM, Ehrenkanz NJ.** 1974. Venezuelan equine encephalomyelitis in Florida: endemic virus circulation in native rodent populations of Everglades hammocks. Am. J. Trop. Med. Hyg. **23**:513–521.
5. **Brault AC, et al.** 2004. Venezuelan equine encephalitis emergence: enhanced vector infection from a single amino acid substitution in the envelope glycoprotein. Proc. Natl. Acad. Sci. U. S. A. **101**:11344–11349.
6. **Calisher CH, Karabatsos N, Lazuick JS, Monath TP, Wolff KL.** 1998. Reevaluation of the western equine encephalitis antigenic complex of alphaviruses (family *Togaviridae*) as determined by neutralization tests. Am. J. Trop. Med. Hyg. **38**:447–452.
7. **Calisher CH, et al.** 1982. Identification of a new Venezuelan equine encephalitis virus from Brazil. Am. J. Trop. Med. Hyg. **31**:1260–1272.
8. **Calisher CH, et al.** 1980. Characterization of Fort Morgan virus, an alphavirus of the Western equine encephalitis virus complex in an unusual ecosystem. Am. J. Trop. Med. Hyg. **29**:1428–1440.
9. **Casals J, Whitman L.** 1957. Mayaro virus: a new human disease agent. I. Relationship to other arbor viruses. Am. J. Trop. Med. Hyg. **6**:1004–1011.
10. **Causay OR, Casals J, Shope RE, Udomsakdi S.** 1963. Aura and Una, two new group A arthropod-borne viruses. Am. J. Trop. Med. Hyg. **12**:777–781.
11. **Causay OR, Maroja OM.** 1957. Mayaro virus: a new human disease agent. III. Investigation of an epidemic of acute febrile illness on the river Guama in Para, Brazil, and isolation of Mayaro virus as a causative agent. Am. J. Trop. Med. Hyg. **6**:1017–1023.

12. **Chamberlain RW, Sudia WD, Coleman PH, Work TH.** 1964. Venezuelan equine encephalitis virus from South Florida. Science **145**:272–274.
13. **Cox-Foster DL, et al.** 2007. A metagenomic survey of microbes in honey bee colony collapse disorder. Science **318**:283–287.
14. **Dalgarno L, et al.** 1984. Characterization of Barmah forest virus: an alphavirus with some unusual properties. Virology **133**:416–426.
15. **Diaz LA, Diaz Mdel P, Almiron WR, Contigiani MS.** 2007. Infection by UNA virus (Alphavirus; Togaviridae) and risk factor analysis in black howler monkeys (*Alouatta caraya*) from Paraguay and Argentina. Trans. R. Soc. Trop. Med. Hyg. **101**:1039–1041.
16. **Digoutte JP, Girault G.** 1976. The protective properties in mice of Tonate virus and two strains of Cabassou virus against neurovirulent everglades Venezuelan encephalitis virus. Ann. Microbiol. **127B**:429–437. (In French.)
17. **Doherty RL, Carley JG, Mackerras MJ, Marks EN.** 1963. Studies of arthropod-borne virus infections in Queensland. III. Isolation and characterization of virus strains from wild-caught mosquitoes in North Queensland. Aust. J. Exp. Biol. Med. Sci. **41**:17–39.
18. **Doherty RL, Gorman BM, Whitehead RH, Carley JG.** 1966. Studies of arthropod-borne virus infections in Queensland. V. Survey of antibodies to group A arboviruses in man and other animals. Aust. J. Exp. Biol. Med. Sci. **44**:365–377.
19. **Dominguez G, Wang CY, Frey TK.** 1990. Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution. Virology **177**:225–238.
20. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.
21. **Etherington GJ, Dicks J, Roberts IN.** 2005. Recombination Analysis Tool (RAT): a program for the high-throughput detection of recombination. Bioinformatics **21**:278–281.
22. **Flahault A, et al.** 2007. Chikungunya, La Reunion and Mayotte, 2005–2006: an epidemic without a story? Sante Publique **19**:S165–S195. (In French.)
23. **Frolov I, Schlesinger S.** 1996. Translation of Sindbis virus mRNA: analysis of sequences downstream of the initiating AUG codon that enhance translation. J. Virol. **70**:1182–1190.
24. **Fukunaga Y, Kumanomido T, Kamada M.** 2000. Getah virus as an equine pathogen. Vet. Clin. N. Am. Equine Practice **16**:605–617.
25. **Gard G, Marshall ID, Woodroofe GM.** 1973. Annually recurrent epidemic polyarthritis and Ross River virus activity in a coastal area of New South Wales. II. Mosquitoes, viruses, and wildlife. Am. J. Trop. Med. Hyg. **22**:551–560.
26. **Gould EA, de Lamballerie X, Zanotto PM, Holmes EC.** 2003. Origins, evolution and vector/host co-adaptations within the genus *Flavivirus*. Adv. Virus Res. **59**:277–314.
27. **Gouy M, Guindon S, Gascuel O.** 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. **27**:221–224.
28. **Grard G, et al.** 2010. Genomics and evolution of *Aedes*-borne flavivirus. J. Gen. Virol. **91**:87–94.
29. **Grard G, et al.** 2007. Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenic determinants, and taxonomy. Virology **361**:80–92.
30. **Hahn CS, Lustig S, Strauss EG, Strauss JH.** 1988. Western equine encephalitis virus is a recombinant virus. Proc. Natl. Acad. Sci. U. S. A. **85**:5997–6001.
31. **Henderson JR, Karabatsos N, Bourke AT, Wallis RC, Taylor RM.** 1962. A survey for arthropod-borne viruses in south-central Florida. Am. J. Trop. Med. Hyg. **11**:800–810.
32. **Hopla CE, Francy DB, Calisher CH, Lazuick JS.** 1993. Relationship of cliff swallows, ectoparasites, and an alphavirus in west-central Oklahoma. J. Med. Entomol. **30**:267–272.
33. **Huelsenbeck JP, Ronquist F.** 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics **17**:754–755.
34. **Jost H, et al.** 2010. Isolation and phylogenetic analysis of Sindbis viruses from mosquitoes in Germany. J. Clin. Microbiol. **48**:1900–1903.
35. **Kinney RM, Pfeffer M, Tsuchiya KR, Chang GJ, Roehrig JT.** 1998. Nucleotide sequences of the 26S mRNA's of the viruses defining the Venezuelan equine encephalitis antigenic complex. Am. J. Trop. Med. Hyg. **59**:952–964.
36. **Kokernot RH, De Meillon B, Paterson HE, Heymann CS, Smithburn KC.** 1957. Middelburg virus; a hitherto unknown agent isolated form *Aedes* mosquitoes during an epizootic in sheep in the eastern Cape Province. S. Afr. J. Med. Sci. **22**:145–153.
37. **Kokernot RH, McIntosh BM, Worth CB.** 1961. Ndumu virus, a hitherto unknown agent, isolated from culicine mosquitoes collected in northern Natal. Union of South Africa. Am. J. Trop. Med. Hyg. **10**:383–386.
38. **Kubes V, Rios FA.** 1939. The causative agent of infectious equine encephalomyelitis in Venezuela. Science **90**:20–21.
39. **Kubes V, Rios FA.** 1939. Equine encephalomyelitis in Venezuela: advance data concerning the causative agent. Can. J. Comp. Med. **3**:43–44.
40. **La Linn M, et al.** 2001. Arbovirus of marine mammals: a new alphavirus isolated form the elephant seal louse, *Lepidohthirus macrorhini*. J. Virol. **75**:4103–4109.
41. **Lang AS, Rise ML, Culley AI, Steward GF.** 2009. RNA viruses in the sea. FEMS Microbiol. Rev. **33**:295–323.
42. **Li GP, Rice CM.** 1989. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. J. Virol. **63**:1326–1337.
43. **Li L, Jose J, Xiang Y, Kuhn RJ, Rossman MG.** 2010. Structural changes of envelope proteins during alphavirus fusion. Nature **468**:705–708.
44. **Loye JE, Hopla CE.** 1983. Ectoparasites and microorganisms associated with the cliff swallow in west-central Oklahoma. II. Life history patterns. Bull. Soc. Vector Ecol. **8**:79–84.
45. **Luers AJ, Adams SD, Smalley JV, Campanella JJ.** 2005. A phylogenomic study of the genus Alphavirus employing whole genome comparison. Comparative Functional Genomics **6**:217–227.
46. **Lumsden WH.** 1955. An epidemic of virus disease in Southern province, Tanganyika Territory, in 1952–53. II. General description and epidemiology. Trans. R. Soc. Trop. Med. Hyg. **49**:33–57.
47. **Margulies M, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
48. **Merrill MH, Lacailladie CW, Jr, Broeck CT.** 1934. Mosquito transmission of equine encephalomyelitis. Science **1934**:251–252.
49. **Meyer KF, Haring CM, Howitt B.** 1931. The etiology of epizootic encephalomyelitis of horses in the San Joaquin valley, 1930. Science **74**:227–228.
50. **Mitchell CJ, et al.** 1987. Arbovirus isolations form mosquitoes collected during and after the 1982–1983 epizootic of western equine encephalitis in Argentina. Am. J. Trop. Med. Hyg. **36**:107–113.
51. **Myles KM, Kelly CLH, Ledermann JP, Powers AM.** 2006. Effects of an opal termination codon preceding the nsP4 gene sequence in the O'Nyong-Nyong virus genome on *Anopheles gambiae* infectivity. J. Virol. **80**:4992–4997.
52. **Palacios G, et al.** 2008. A new arenavirus in a cluster of fatal transplant-associated diseases. N. Engl. J. Med. **358**:991–998.
53. **Palacios G, et al.** 2007. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. Emerg. Infect. Dis. **13**:73–81.
54. **Peleg J, Pecht M.** 1978. Adaptation of an *Aedes aegypti* mosquito cell line to growth at 15°C and its response to infection by Sindbis virus. J. Gen. Virol. **38**:231–239.
55. **Pettersen EF, et al.** 2004. UCSF Chimera: a visualization system for exploratory research and analysis. J. Comput. Chem. **25**:1605–1612.
56. **Posada DC, Crandall KA.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817–818.
57. **Powers AM, et al.** 2001. Evolutionary relationship and systematics of the alphaviruses. J. Virol. **75**:10118–10131.
58. **Robinson MC.** 1955. An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952–53. I. Clinical features. Trans. R. Soc. Trop. Med. Hyg. **49**:28–32.
59. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.
60. **Ross RW.** 1956. A laboratory technique for studying the insect transmission of animal viruses, employing a bat-wing membrane, demonstrated with two African viruses. J. Hyg. (Lond.) **54**:192–200.
61. **Ross RW, Miles JA, Austin FJ, Maguire T.** 1964. Investigations into the ecology of a group A arbovirus in Westland, New Zealand. Aust. J. Exp. Biol. Med. Sci. **42**:689–702.
62. **Rumenapf T, et al.** 1995. Aura alphavirus subgenomic RNA is packaged into virions of two sizes. J. Virol. **69**:1741–1746.
63. **Sawicki DL, Perri S, Polo JM, Sawicki SG.** 2006. Role for nsP2 proteins in the cessation of alphavirus minus-strand synthesis by host cells. J. Virol. **80**:360–371.
64. **Scherer WF, Chin J.** 1983. An unusual strain of Venezuelan encephalitis

existing sympatrically with subtype I-D strains in a Peruvian rain forest. Am. J. Trop. Med. Hyg. **32**:871–876.
65. **Scherer WF, Funkenbusch M, Buescher EL, Izumi T.** 1962. Sagiyama virus, a new group A arthropod-borne virus from Japan. I. Isolation, immunologic classification, and ecologic observations. Am. J. Trop. Med. Hyg. **11**:255–268.
66. **Scherer WF, Izumi T, McCown J, Hardy JL.** 1962. Sagiyama virus. II. Some biologic, physical, chemical and immunologic properties. Am. J. Trop. Med. Hyg. **11**:269–282.
67. **Sherman MB, Weaver SC.** 2010. Structure of the recombinant alphavirus western equine encephalitis virus revealed by cryoelectron microscopy. J. Virol. **84**:9775–9782.
68. **Shope RE, Causay OR, De Andrade AH.** 1964. The Venezuelan equine encephalomyelitis complex of group A arthropod-borne viruses, including Mucambo and Pixuna from the Amazon region of Brazil. Am. J. Trop. Med. Hyg. **13**:723–727.
69. **Smithburn KC, Haddow AJ.** 1944. Semliki Forest virus: I. Isolation and pathogenic properties. J. Immunol. **49**:141–157.
70. **Smithburn KC, Mahafferty AF, Haddow AJ.** 1944. Semliki Forest virus: II. Immunological studies with specific antiviral sera and sera from humans and wild animals. J. Immunol. **49**:159–173.
71. **Strauss JH, Strauss EG.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. **58**:491–562.
72. **Swofford D.** 2000. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer Associates, Sunderland, MA.
73. **Taylor RM, Hurlbut HS.** 1953. The isolate of coxsackie-like viruses from mosquitoes. J. Egypt. Med. Assoc. **36**:489–494.
74. **Tenbroeck C, Hurst EW, Traub E.** 1935. Epidemiology of equine encephalomyelitis in the eastern United States. J. Exp. Med. **62**:677–685.
75. **Tesh RB, Gajdusek DC, Garruto RM, Cross JH, Rosen L.** 1975. The distribution and prevalence of group A arbovirus neutralizing antibodies among human populations in Southeast Asia and the Pacific Islands. Am. J. Trop. Med. Hyg. **24**:664–675.
76. **Travassos da Rosa AP, et al.** 2001. Trocara virus: a newly recognized alphavirus (*Togaviridae*) isolated from mosquitoes in the Amazon Basin. Am. J. Trop. Med. Hyg. **64**:93–97.
77. **Tsetsarkin KA, Vanlandingham DL, McGee CE, Higgs S.** 2007. A single mutation in Chikungunya virus affects vector specificity and epidemic potential. PLoS Pathog. **3**:e201.
78. **Volk SM, et al.** 2010. Genome-scale phylogenetic analyses of Chikungunya virus reveal independent emergences of recent epidemic and various evolutionary rates. J. Virol. **84**:6497–6504.
79. **Voss JE, et al.** 2010. Glycoprotein organisation of Chikungunya virus particles revealed by X-ray crystallography. Nature **468**:709–712.
80. **Weaver SC, et al.** 2005. *Togaviridae*, p 999–1008. *In* Fauquet CM, et al. (ed), Virus taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses. Elsevier/Academic Press, London, England.
81. **Weaver SC, et al.** 1997. Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. J. Virol. **71**: 613–623.
82. **Weston J, et al.** 2002. Comparison of two aquatic alphaviruses, Salmon pancreas disease virus and Sleeping disease virus, by using genome sequence analysis, monoclonal reactivity, and cross-infection. J. Virol. **76**: 6155–6163.
83. **Weston JH, Welsh MD, McLoughlin MF, Todd D.** 1999. Salmon pancreas disease virus, an alphavirus infecting farmed Atlantic salmon, *Salmo salar* L. Virology **256**:188–195.
84. **Williams MC, Woodall JP.** 1961. O'nyong nyong fever: an epidemic virus disease in East Africa. II. Isolation and some properties of the virus. Trans. R. Soc. Trop. Med. Hyg. **55**:135–141.
85. **Wolf K, Mann JA.** 1980. Poikilotherm vertebrate cell lines and viruses: a current listing for fishes. In Vitro **16**:168–179.
86. **Zwickl DJ.** 2006. Genetic algorithm approaches for the phylogenetics analysis of large biological sequence datasets under the maximum-likelihood criterion. PhD dissertation. University of Texas at Austin, Austin, TX.