JOURNAL OF VIROLOGY, Jan. 1994, p. 158–169
0022-538X/94/$04.00+0

Vol. 68, No. 1

Evolution of Alphaviruses in the Eastern Equine Encephalomyelitis Complex

SCOTT C. WEAVER,[1,2]* AMY HAGENBAUGH,[1] LIZ ANNE BELLEW,[1] LAURENT GOUSSET,[1] VARUNI MALLAMPALLI,[2] JOHN J. HOLLAND,[1] AND THOMAS W. SCOTT[2]

Department of Biology, University of California, San Diego, La Jolla, California 92093,[1] and Department of Entomology, University of Maryland, College Park, Maryland 20742[2]

Received 19 July 1993/Accepted 29 September 1993

Evolution of viruses in the eastern equine encephalomyelitis (EEE) complex was studied by analyzing RNA sequences and oligonucleotide fingerprints from isolates representing the North and South American antigenic varieties. By using homologous sequences of Venezuelan equine encephalomyelitis virus as an outgroup, phylogenetic trees revealed three main EEE virus monophyletic groups. A North American variety group included all isolates from North America and the Caribbean. One South American variety group included isolates from the Amazon basin in Brazil and Peru, while the other included strains from Argentina, Guyana, Ecuador, Panama, Trinidad, and Venezuela. No evidence of heterologous recombination was obtained when three separate regions of the EEE virus genome were analyzed independently. Estimates of the overall rate of EEE virus evolution (nucleotide substitution) were 1.6×10^{-4} substitution per nucleotide per year for the North American group and 4.3×10^{-4} for the Argentina-Panama South American group. Evolutionary rate estimates for the North American group increased over 10-fold (from about 2×10^{-5} to 4×10^{-4}) concurrent with divergence of two monophyletic groups during the early 1970s. The North and South American antigenic varieties diverged roughly 1,000 years ago, while the two main South American groups diverged about 450 years ago. Analysis of multiple strains isolated from an upstate New York transmission focus during the same years suggested that, in certain locations, EEE virus may be relatively isolated for short time periods.

The eastern equine encephalomyelitis (EEE) complex is a group of antigenically related viruses in the family *Togaviridae*, genus *Alphavirus* (3). Alphaviruses have positive (plus-strand) or messenger sense unsegmented RNA genomes of 11 to 12 kb. Infected cells also contain a subgenomic 26S RNA species, which is identical to the 3′ third of the genome; 26S RNA encodes three structural proteins designated capsid, E1, and E2, while the genomic RNA also encodes the four nonstructural proteins, nsP1 to nsP4 (30).

EEE virus is the only species in the EEE complex, which is one of seven antigenically defined alphavirus complexes. Two antigenic varieties of the EEE virus species have been described; all strains isolated in North America and most from the Caribbean belong to the North American variety, while isolates from Central and South America compose the South American variety (3, 4, 22, 25). An antigenic subtype of the North American variety, represented by a single human isolate from Mississippi, was described recently (2).

In North America, EEE virus is transmitted among songbirds in freshwater swamps by the mosquito vector, *Culiseta melanura*. Transmission occurs along the Atlantic coast from New Hampshire to Florida and along the Gulf coast to Texas. Enzootic foci also occur at inland locations including upstate New York, Michigan, Wisconsin, and Ontario. Epizootics, involving different mosquito vectors which transmit EEE virus to mammalian hosts, occur periodically with severe morbidity and high mortality rates in humans, horses, and gamebirds (22, 27).

In Central and South America, the epidemiology of EEE virus is poorly understood and human and equine disease is reported less frequently. Although arthropod vectors have not been identified definitively, mosquitoes belonging to the subgenus (*Melanoconion*) of the genus *Culex* probably transmit EEE virus among small mammals and/or birds in enzootic foci (27).

Recent work has indicated that EEE virus is genetically conserved in North America. RNA sequences from the 26S region of strains isolated 52 years apart yielded an estimate of 1.4×10^{-4} substitution per nucleotide per year, an evolutionary rate lower than that of many other non-arthropod-borne RNA viruses (38). A tree generated from sequence data of 13 isolates grouped them by year rather than the location of isolation, suggesting that EEE virus evolves as a single population with frequent exchange of viruses among transmission foci in North America (38). This conclusion was also supported by trees generated from complete 26S sequences and genomic RNA fingerprints from a total of 10 North American isolates, including the Mississippi North American antigenic subtype (36).

Evolutionary relationships among EEE viruses within the two antigenic varieties remain unknown. To examine these relationships, and to estimate the time frame for diversification of viruses in the EEE complex, we obtained RNA sequences from 16 South and Central American isolates. Using homologous nucleotide sequences of the Venezuelan equine encephalomyelitis virus (VEE virus) Trinidad donkey strain (21) as an outgroup (a related taxon used to root the tree and determine ancestral relationships), phylogenetic analysis indicated that the EEE complex comprises a monophyletic group (a group of viruses descended from a common ancestor) which includes three main groups. The North and South American varieties diverged ca. 1,000 years ago, while a single group in North America diverged into two distinct monophyletic groups around 1974. Divergence of the two North American groups

* Corresponding author. Mailing address: Department of Biology, 0116, The University of California, San Diego, La Jolla, CA 92093-0116. Phone: (619) 534-2520. Fax: (619) 534-7108. Electronic mail address: sweaver@ucsd.edu.

was accompanied by a 10-fold rise in the rate of nucleotide substitution.

MATERIALS AND METHODS

Viruses. The EEE virus strains that we analyzed are listed in Tables 1 and 2. Virus stocks were prepared on BHK-21 cell culture monolayers at 37°C, with multiplicities of infection of 0.1 to 1.0 PFU per cell. Extraction of genomic RNA was performed as described previously (38).

RNA sequencing. Primer-extension dideoxynucleotide sequencing of genomic viral RNA was performed as described by Fichot and Girard (13). DNA oligonucleotide primers described previously (38) were annealed to viral RNA by heating to 65°C, followed by gradual cooling. An additional DNA primer of sequence 5′-CTGCAAAGTGTCATCTG-3′, complementary to South American variety EEE virus RNA at positions 2289 to 2305 (numbering of Weaver et al. [36]), was used to sequence a region in the E2 glycoprotein of South American variety isolates.

Sequence analysis. For all North American variety isolates, two regions of the 26S region described previously (38), totaling 1,350 nucleotides, were sequenced; these regions include the C-terminal half of the E2 envelope glycoprotein, most of 6K, the C-terminal end of the E1 glycoprotein, and most of the 3′ untranslated region of the EEE virus genome. For South American variety isolates, three regions were sequenced: the C-terminal end of nsP4, 40 to 217 nucleotides upstream of the 26S region; nucleotides 2145 to 2266 encoding the C-terminal region of the E2 envelope glycoprotein; and nucleotides 3809 to 4110 (333 nucleotides including gaps) within the noncoding region at the 3′ end of the genome (numbering by Weaver et al. [36]).

Nucleotide sequences were aligned by using the PILEUP program of the Genetics Computer Group (6). Gaps or insertions greater than 4 nucleotides in length, found in more than one isolate, were treated as single characters. Phylogenetic trees were obtained with the Phylogenetic Analysis Using Parsimony (PAUP) program by using the heuristic algorithm (32), the DNAML maximum likelihood program (12), the PAPA3 nearest-neighbor method (7), and the FITCH distance-matrix method (12). For PAUP, characters were initially unordered, but a rescaled consistency index was used for a posteriori successive weighting to identify equally parsimonious trees that were supported best by the informative characters (10). This method adjusts the weight of characters (nucleotides) on the basis of their fit to the most-parsimonious trees (i.e., reduced the effect of nucleotides which undergo reversion). Confidence values were determined for tree groupings by the character resampling bootstrap analysis method (11). The DNAML and FITCH programs were implemented with the assumption of a 5:1 ratio of transitions-transversions, determined empirically from North American EEE virus sequence data (36, 38).

Nucleotide sequence accession numbers. All sequences were deposited with the GenBank library under accession numbers U01552 to U01656.

RESULTS

EEE virus complex. Aligned nucleotide sequences, totaling 633 bases, for 16 South American variety EEE virus isolates, along with North American variety sequences VA33 (the oldest isolate), FL82 (a recent isolate), and the VEE virus outgroup sequence, are shown in Fig. 1 (the virus codes are defined in Tables 1 and 2). The VEE sequence for the 3′ untranslated region was not included because of previously described uncertainties in the proper alignment of untranslated regions among different alphaviruses (35). Using unordered characters, the PAUP program produced 64 equally parsimonious trees from these sequences; these trees differed only in some relationships within the Argentina-Panama group (see below). When transversions were weighted five times that of transitions, 26 equally parsimonious trees were found. These trees differed only in the relationships depicted among the AR36, AR38, AR59, BG60, VE80, and VE81 isolates.

The maximum likelihood program produced a tree with topology identical to one of the two most parsimonious trees generated with PAUP; this PAUP tree is shown in Fig. 2. All trees revealed three main monophyletic groups of EEE virus, supported by bootstrap confidence values of 100%; one group included two Brazil isolates (BR56 and BR76) and a 1970 Peru isolate (henceforth referred to as the Brazil-Peru group). Both Brazil isolates are from the Belem area, while the 1970 Peru strain (PE70) was isolated in Iquitos, Peru, in the Amazon basin. Sequence data from more-recent and geographically diverse isolates are needed to determine whether more than one group exists and whether the Brazil-Peru group(s) is confined to the Amazon basin.

A second main group included isolates from Argentina, Panama, Ecuador, Trinidad, Venezuela, and Guyana (henceforth referred to as the Argentina-Panama group; Fig. 2). Groupings of many isolates were based on the geographical location of isolation, suggesting that independent evolution occurred for several years at some locations. For example, strains isolated within Panama or Venezuela tended to form separate groups in the trees (Fig. 2).

A third EEE virus group included two representatives of the North American variety (Fig. 2). Viruses in the North American group differed from the South American variety viruses by 26 to 29% in their nucleotide sequences; the two South American variety groups differed by 18 to 20%. The nsP4 amino acid sequences differed by about 4% (Brazil-Peru group versus Argentina-Panama group), 11% (Brazil-Peru group versus North American group), and 12% (Argentina-Panama versus North American group); amino acid sequences in the E2-6K-E1 region differed by about 3% (Brazil-Peru group versus Argentina-Panama group and Brazil-Peru group versus North American group) to 5% (Argentina-Panama versus North American group). Overall, 84% of the nucleotide differences among isolates were synonymous substitutions, primarily in the third position of codons.

To evaluate the possibility that heterologous recombination played a role in evolution of the EEE complex, we conducted separate phylogenetic analyses of nucleotide sequences within each of the three genome regions that we sampled. The trees from these analyses (data not shown) were similar to that generated by using the combined sequence regions, providing no evidence of recombination within the 26S portion of the EEE virus genome.

North American variety. To examine in greater detail the evolution of the North American group, we sequenced two maximally variable regions within the 26S portion of the EEE virus genome for 21 additional isolates. Nucleotides within these regions differing among isolates are listed in Fig. 3, along with those from 13 previously published sequences (36, 38). Initially, we included homologous outgroup sequences of all South American variety EEE isolates (Table 2), and the VEE Trinidad Donkey strain (21), to obtain a rooted North American tree. However, only 19 homologous nucleotides were available from the South American variety EEE sequences, and these provided inconsistent rooting. Most homologous

TABLE 1. North American variety EEE virus strains used in phylogenetic analyses

Code	Strain	Location	Date isolated (yr or mo-day-yr)	Host	Passage[a]
VA33[b]	Ten Broeck	Virginia	1933	Equine	sm12, v1
NJ45[b]	Hammon	New Jersey	1945	Pheasant	sm8, v1
LA47[c]	Decuir	Louisiana	1947	Human	p1
DR49[b]	Sanchez	Dominican Republic	1949	Horse	sm1, v2
LA50	Arth-167	Louisiana	1950	Mosquito	gp2, ch2
GA53	664	Georgia	1953	Horse	p1
MA56[b]	JW	Massachusetts	1956	Human	sm1, v1
NJ59[b]	NJ-1959	New Jersey	1959	Mosquito	sm3, v1
AL60	AL09-19-1045	Alabama	1960	Bird	m1
NJ60[c]	NJ/60	New Jersey	1960	Mosquito	p6, sm1
FL61	08-3-0855	Florida	1961	Bird	m1, sm3
ON61	32-14-989	Ontario	1961	Bird	p1, sm2
WI61	984	Wisconsin	1961	Horse	p1, sm1
GA62	WX2-1A	Georgia	1962	Mosquito	sm1
JA62	7761	Jamaica	1962	Horse	sm2, v1
FL64	B64-5282.01	Florida	1964	Mosquito	sm3
LA66	W-27933	Louisiana	1966	Bird	sm2
NJ68	A61-1K	New Jersey	1968	Bird	ch2, m4, sm1
NY69	69-7836	New York	1969	Pheasant	sm2
NY71A	71-3473	New York	8-28-71	Horse	sm2
NY71B	71-20685	New York	1971	Mosquito	sm1
NY74	74-33620	New York	1974	Mosquito	sm2
NY76	76-21743	Oswego County, N.Y.	1976	Pheasant	m1
MA77[b]	ME77132	Carver Swamp, Mass.	8-77	Mosquito	mq1, C6/36-1
DR78	10365	Dominican Republic	1978	Horse	sm2
MA78	78-3372	Raynham, Mass.	9-78	Mosquito	C6/36-1
MD79[b]	MP-9	Maryland	1979	Mosquito	v1, C6/36-1
MA80[b]	MA1833	Halifax, Mass.	8-80	Mosquito	p1, v2
MI80[b]	R-35108	Three Rivers, Mich.	1980	Horse	C6/36-1, v1
NJ80	M80-600	New Jersey	1980	Mosquito	ch2, sm1
WI80	WiAn5000-003	Wisconsin	1980	Horse	de2, sm1, v1
FL82[b]	82V-2137	Florida	1982	Mosquito	Unknown
GA82	82-8409	Georgia	1982	Dog	sm1
MA82	396	Massachusetts	1982	Mosquito	v2
MI82	83V-2264	Michigan	1982	Mosquito	de1
NJ82[b]	DV260-82	New Jersey	1982	Bird	v1
OH82	R-39922	Ohio	1982	Horse	v1, sm1
CT83[c]	M-210-83	Connecticut	9-83	Mosquito	v1
MS83[c]	4789	Mississippi	9-20-83	Human	rd2, sm3
MD84[c]	40-84	Maryland	8-16-84	Bird	Unpassaged
FL85	VO82085	Florida	1985	Horse	v2
MD85A	187-85	Maryland	8-29-85	Mosquito	Unpassaged
MD85B[b]	215-85	Maryland	9-4-85	Mosquito	Unpassaged
MD85C	219-85	Maryland	9-4-85	Mosquito	Unpassaged
MD85D	318-85	Maryland	10-20-85	Mosquito	Unpassaged
MD85E	323-85	Maryland	10-20-85	Mosquito	Unpassaged
GA86	86-20924	Georgia	1986	Dog	Unpassaged
MD87	953-87	Maryland	8-21-87	Mosquito	Unpassaged
MD88[b]	2061-88	Maryland	8-26-88	Mosquito	Unpassaged
NY88A	88-21529	Oswego County, N.Y.	8-22-88	Bird	Unpassaged
NY88B	88-21610	Oswego County, N.Y.	8-30-88	Bird	Unpassaged
NC89	R53442	North Carolina	8-2-89	Horse	Unpassaged
TN89	89-42687	Tennessee	1989	Horse	v2
MS89	89-45989	Mississippi	1989	Horse	v1
MI89	89-47595	Michigan	1989	Horse	v1
CT90	Williams	Connecticut	10-90	Horse	v1
MD90A	3067-90	Maryland	10-10-90	Mosquito	Unpassaged
MD90B	3071-90	Maryland	10-11-90	Mosquito	Unpassaged
NY90A	90-25078	Onandaga County, N.Y.	8-22-90	Pheasant	Unpassaged
NY90B	90-29015	Onandaga County, N.Y.	9-4-90	Pheasant	Unpassaged
NY90C	90-29060	Oneida County, N.Y.	10-3-90	Horse	Unpassaged
NY90D	90-31174	Oswego County, N.Y.	7-30-90	Mosquito	Unpassaged
NY90E	90-31315	Oswego County, N.Y.	8-9-90	Mosquito	Unpassaged
NY90F	90-31951	Oswego County, N.Y.	10-16-90	Mosquito	Unpassaged
NY90G	90-33222	Suffolk County, N.Y.	8-28-90	Mosquito	Unpassaged
NY90H	90-29059	Orange County, N.Y.	9-14-90	Horse	Unpassaged
RI90	Ar5-9690	Rhode Island	9-6-90	Mosquito	v1
FL91	FL91-4679	Florida	1991	Mosquito	v3
NY91A	91-29002	Oswego County, N.Y.	7-23-91	Horse	Unpassaged

Continued on following page

TABLE 1—*Continued.*

Code	Strain	Location	Date isolated (yr or mo-day-yr)	Host	Passage[a]
NY91B	91-30535	Onandaga County, N.Y.	6-25-91	Mosquito	Unpassaged
NY91C	91-31668	Onandaga County, N.Y.	8-13-91	Mosquito	Unpassaged
NY91D	91-32277	Onandaga County, N.Y.	9-25-91	Mosquito	Unpassaged
OH91	R55788	Ohio	1991	Horse	p2, v1

[a] C6/36, *Aedes albopictus* mosquito cells; ch, chicken; de, duck embryo cell; gp, guinea pig; m, mouse; mq, mosquito; p, unknown passage; rd, human rhabdomyosarcoma; sm, suckling mouse; v, Vero cell.
[b] Sequences from Weaver et al. (38).
[c] Sequences from Weaver et al. (36).

nucleotides were available for VEE virus (only the 3′ untranslated nucleotides were omitted as described above), but they also resulted in inconsistent rooting and the generation of 3,080 equally parsimonious trees (data not shown) with 363 mutations (total length). The majority of the roots in these trees were similar to that implied previously by midpoint rooting (38), which can be invalid for trees depicting a monophyletic group sampled at different times. Many of the VEE-rooted trees also predicted that genotypes like those represented by the strains isolated between 1933 and 1971 and between 1979 and 1991 (1933-to-1971 and 1979-to-1991 groups, respectively) were cocirculating from 1933 through 1971. However, the lack of sequences similar to those of the 1979-to-1991 isolates, among the 21 EEE strains isolated from 1933 to 1971 (Table 3; see below), makes this kind of sampling error seem unlikely. Further, constraining the VA33 isolate to be basal within the North American group added only three extra steps (total of 366 mutations) to these trees. The length distribution of 10,000 trees generated at random was not highly skewed, indicating that multiple substitutions of nucleotides probably obscured many phylogenetically informative substitutions (phylogenetic "signal"; see references 14 and 17). This suggested that reversions, combined with the small number of informative nucleotides available for the North American viruses, resulted in loss of resolution of relationships with outgroup sequences.

We therefore generated unrooted North American trees and arranged these trees with the sequence of the oldest North American isolate (VA33) at the proximal position because previous results implied a single North American EEE virus monophyletic group (36, 38). Analysis using the PAUP program resulted in 264 trees of equal (minimal) branch length, with 100 different branching patterns, by using unordered characters (transversion weighting resulted in even more [1,192] equally parsimonious trees); a posteriori successive character reweighting reduced this number to 24. One of the 24 trees, depicting groupings found in the majority of the 24 equally parsimonious trees, is shown in Fig. 4. Overall, the virus strains studied, including those from the Caribbean and midwestern North America, were associated by time and not location of isolation. Exact relationships among most of the 1933-to-1969 isolates could not be determined because of the small number of substitutions that occurred during this period. Within this older group of isolates, there was no consistent pattern of grouping with respect to time or space.

Distance matrix methods can sometimes provide greater accuracy and resolution than parsimony when few phylogenetically informative characters are available, because information from characters present in only one taxa are ignored by parsimony (5, 23, 29). Therefore, we also analyzed the North American EEE virus sequences, using the FITCH and PAPA3 programs. The tree generated with the PAPA3 program is shown in Fig. 5. This tree provided slightly greater resolution of relationships among the older isolates, as well as within the NY74A-MD79-NJ82-TN89-MD90A group. The tree constructed by using the FITCH program was nearly identical in topology; only the position of the FL82 isolate within the MI89-MS89-MD85B-MD88-NC89-FL82 group differed.

All of these trees indicated that, beginning around 1974, the North American EEE virus group apparently diverged into distinct monophyletic groups (Fig. 4 and 5). Group A included viruses from New York, Maryland, New Jersey, and Tennes-

TABLE 2. South American variety EEE virus strains used in phylogenetic analyses

Code	Strain	Location	Year isolated	Host	Passage[a]
AR36	AfgLL	Argentina	1936	Horse	p3
AR38	ArgB	Argentina	1938	Horse	p5
BR56	BeAn-5122	Brazil	1956	Monkey	sm2
PA58	GML207963	Panama	1958	Horse	sm5, v2
AR59	ArgM	Argentina	1959	Horse	p5
TR59	Tr24443	Trinidad	1959	Mosquito	m6, sm1
BG60	Tr25714	Guyana	1960	Horse	p5
PA62	GML900188	Panama	1962	Equine	sm2, v1
PE70	70U1104	Peru	1970	Hamster	v1
EC74	75V-1496	Ecuador	1974	Mosquito	v1, sm2
BR76	76V-25343	Brazil	1976	Mosquito	sm1
VE76	El Delerio	Venezuela	1976	Horse	sm7
VE80	IVICPan57151	Venezuela	1980	Hamster	v1, sm3
VE81	Pan66058-60	Venezuela	1981	Mosquito	v1, sm1
PA84	GML903866	Panama	1984	Chicken	v4
PA86	MARU435731	Panama	1986	Equine	v2

[a] p, unknown passage; sm, suckling mouse; v, Vero cell.

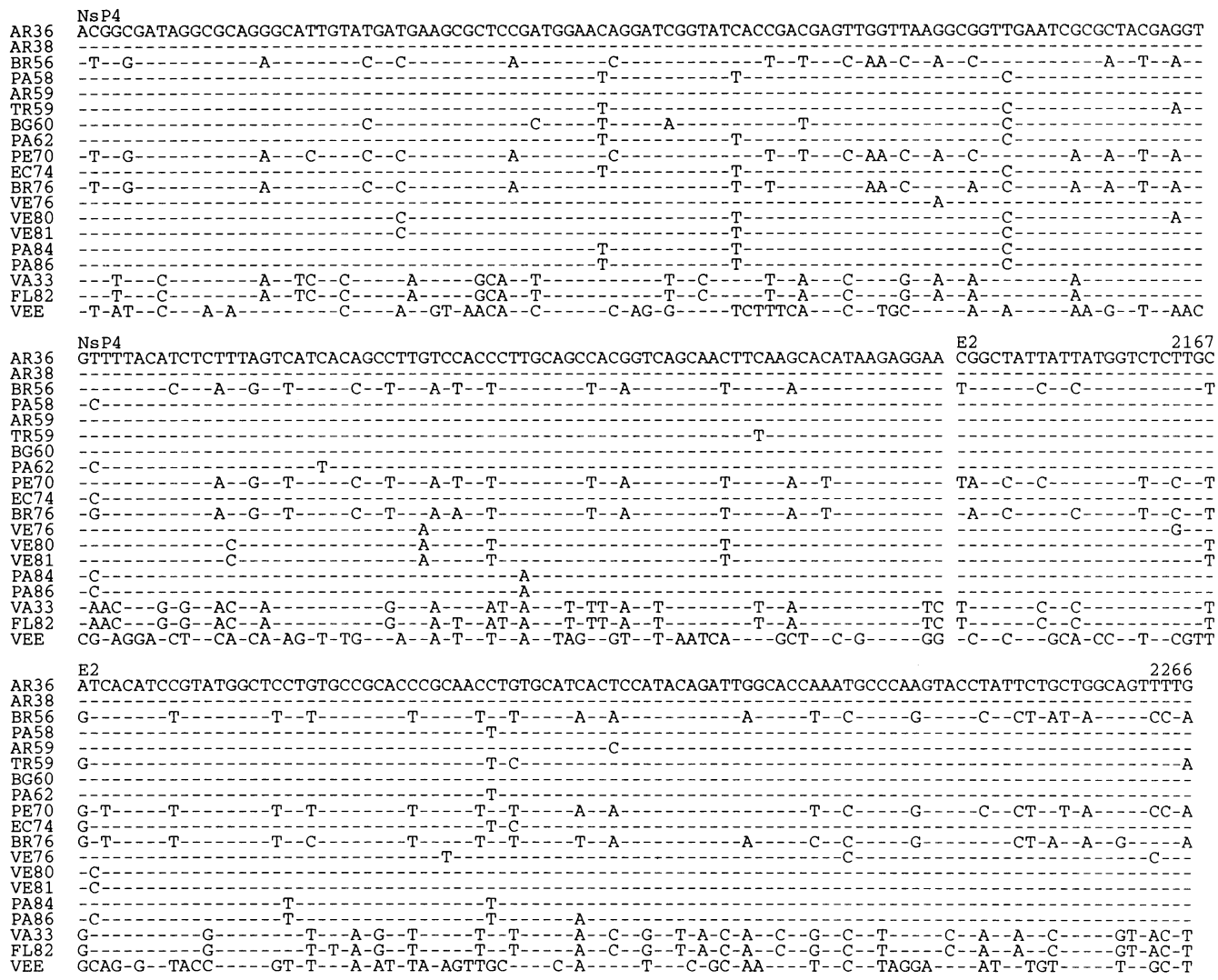

FIG. 1. Aligned sequences used for phylogenetic analysis of the EEE complex. Codes for virus isolates are defined in Table 2. Phylogenetically informative gaps used in the analysis are labeled above the sequence, beginning at the 5′ ends; — indicates that the nucleotide is identical to that of the uppermost sequence. Nucleotide numbers shown follow the numbering by Weaver et al. (26).

see, while group B contained representatives from Wisconsin, Michigan, Mississippi, Maryland, North Carolina, Connecticut, Rhode Island, and New York. Although our sample size was relatively small for viruses isolated prior to 1974, these data suggest that groups A and B shared a common ancestor (Fig. 4 and 5, node C) and were not present during the 1933-to-1969 period. Bootstrap analysis assigned confidence values of 81 and 98%, respectively, for groups A and B representing distinct monophyletic groups (Fig. 4). These values indicated that two or more groups were present during the 1974-to-1991 period. Typically, bootstrap values underestimate the correct probability of monophyletic groups. For example, empirical tests by Hillis and Bull (18) have shown that bootstrap values of 70% or higher correspond to actual probabilities of greater than 95% that a group is real. Therefore, the probability that EEE groups A and B are real may be higher than these bootstrap values indicate.

Within group B, two groupings (MD85B-MD88 group [89% confidence] and a CT90-RI90-NY91D group [76% confidence] suggested that regionally independent evolution may have occurred for a few years in temperate North America. The two Mississippi-Michigan groups (Fig. 4 and 5) are consistent with virus dispersal via birds traveling northward or southward through the Mississippi and Ohio River valleys.

To confirm the existence of the distinct North American groups, we sequenced the entire 26S region of a representative isolate from group A (MD90A) and constructed a phylogenetic tree including nine previously published (36, 38) 26S RNA sequences. When available homologous nucleotides for representatives of the two South American variety groups (BR56 and AR36) were used as outgroups, eight equally parsimonious trees were generated. These trees contained topological inconsistencies like those described above, placing the root for the North American variety strains in different positions. We therefore constructed unrooted trees of complete 26S sequences. A single most-parsimonious unrooted tree was obtained with the PAUP program (Fig. 6). This tree was consistent with a single EEE virus group circulating from 1933 to

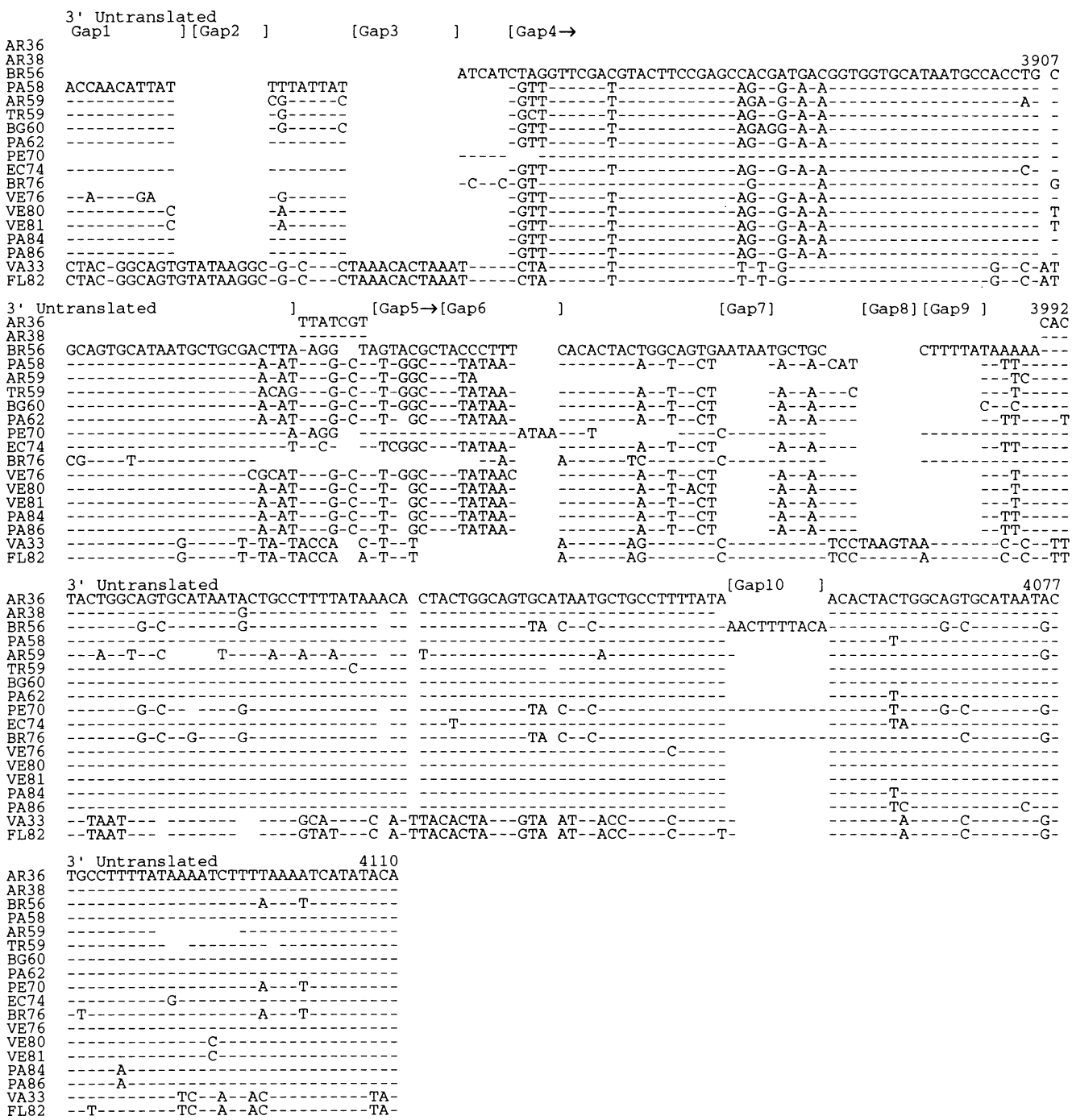

FIG. 1—*Continued.*

1977, followed by divergence of two monophyletic groups (MD90A representing group A; WI80, MS83, FL82, MD85B, and CT90 representing group B). The maximum likelihood, PAPA3, and FITCH programs produced topologically identical trees.

To confirm the existence of groups A and B, we also subjected previously described RNA fingerprint data (34, 36) and the fingerprint of the TN89 strain (33a) to phylogenetic analysis. Again, the VA33 isolate was included as the outgroup because homologous T_1-resistant oligonucleotides can be identified among only very closely related (ca. 10% or less nucleotide divergence) RNA genomes (19), precluding comparison with South American EEE strains. The PAUP program yielded 100 equally parsimonious trees; this number was reduced to 16 after successive character reweighting. A consensus tree depicting groupings found in all 16 trees is shown in Fig. 7. Isolates MD90A and B, as well as MD84A and TN89, formed a monophyletic group consistent with group A sequences (Table 3), while the other group included strains isolated from 1984 to 1991 in Connecticut, Florida, Maryland,

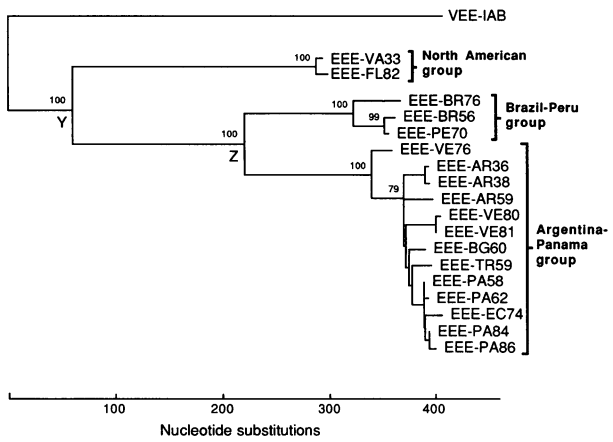

FIG. 2. One of three phylogenetic trees, obtained after successive reweighting, depicting relationships among members of EEE complex from nucleotide sequences listed in Fig. 1. Homologous sequences of the Trinidad donkey strain of VEE virus were used as an outgroup to root the tree. Nodes Y and Z represent hypothetical ancestral EEE viruses. Numbers indicate bootstrap confidence values for monophyletic groups defined by adjacent nodes (hypothetical ancestors). The virus codes are defined in Tables 1 and 2. The scale below the tree indicates the numbers of nucleotide substitutions represented in branch lengths.

TABLE 3. North American EEE virus groups isolated from 1933 to 1991

Group	Isolates
1933–1978	AL60, DR49, DR78, FL61, FL64, GA53, GA62, JA62, LA47, LA50, LA66, MA56, MA77, MA78, NJ45, NJ59, NJ60, NJ68, NY69, NY71A, NY71B, ON61, VA33, WI61
A	GA82, MA80, MD79, MD84(plaque 1), MD90A, MD90B, NJ82, NY74, NY76, TN89
B	CT83, CT90, FL82, FL85, GA86, MA82, MD84(plaque 4), MD85A–MD85D, MD87, MD88, MI80, MI82, MI89, MS83, MS90, NC89, NJ80, NY90A–NY90F, NY91A–NY91D, OH82, OH91, RI90, WI80

New York, Ohio, and Rhode Island; this group was completely consistent with group B sequences (Table 3). The MD84A isolate, believed to represent a dual infection of a sentinel quail (34), contained genotypes (plaque clones) representing both groups (A and B).

In an attempt to delineate further the temporal and spatial distribution of North American groups A and B, we obtained sequences from two smaller 26S regions, within the 1,360-base region, for an additional 30 strains (Table 3) isolated from 1950 to 1991. These sequences included most of the nucleotides which distinguished the 1974-to-1991 groups (nucleotides 2062, 2107, 3577, 3604, and 3619). All sequences fell

```
                1620 1633 1639 1640 1646 1650 1672 1678 1681 1684 1704 1723 1735 1771 1774 1783 1790 1804 1822 1825 1837 1870 1871 1912 1915 1916 1928 1960 1972 1978 1996 1997 2009 2022 2025 2062 2082 2104 2107 2113 2125 2152 2170 2177 2249 2266 2283 2314 2317 2366 2380 2441 3544 3553 3568 3577 3589 3594 3604 3619 3622 3648 3649 3665 3679 3694 3727 3730 3773 3779 3791 3792 3812 3823 3843 3855 3890 3897 3918 3935 3938 3949 3979 3980 3981 3985 3986 4009 4014 4028 4038 4043 4048
VA33   ACTGGGACCGCACTTAGTCCAGACGTGACCGTCCCTCAATCCGATTTCCCACTGTATCCCTTCACTAACTCTCTTCACTGTTTACTCTTCC
NJ45   ----------------------T--------------------------T----------T----------
LA47   ------------------------------------------------------------------------------XX---------
DR49   --------------------T---------------T-----------T-------------T----------T----------
GA53   ----AC----------------A-------------------------------------------------T----------
MA56   ---------A---C----T-----------------T------------T---A----------T----------T----C----------
NJ59   ----------------A-------------------------------------------------------T----------
NJ60   ----------------A--------------------------T-----T-------------T----------T----------
ON61   ------------------------------------------------------------T-------T----------
FL61   ------------T-----------------A-----------------------------G--------T----------
JA62   ------------T-------------------------------------------C-------------T--C-----T------
FL64   ------------------------------------------------------------------------T----------
LA66   -------------C----T---------T-------T------------T--------------T----------
NJ68   ----A---------C-------------------------------------------------------T----------
NY69   ----------------------------A----------T--------------T----------C---T----------T
NY74   --------------------G---------------G---------T-T----------C--TT----------T-A--------T----
MA77   ------------------------------------------------------------------------T----------T-
DR78   -------------C----------------------T------------T-T---G------T----------
MD79   ---A----------------G----------------GG--------TTT----------C---T----------TGA-----G--T----
WI80   --------T-----------G--------------CT-G------C-T------C-CT--C---T----------CT-A--------T----
MI80   --------------------G-----T-----T--C--G--T---C-T------C-CT--C---T----------T-A--------T----
FL82   -T------------------G---A--T-------C--G------C-T------C-CTT-C---T----C-----T-A--------T----
NJ82   --------------------G----------------GG--------T-T----------C---T-----G----T-A--------T----
MS83   ----------------G-TG-----------T--C--G------C-T------C-CT--C---T-------T--T-A--------T----
MD85B  -T-----T---G--------G--------------C--G------C-T------C-CTT-C---TC--T------T-A--------T----
MD88   -T---------G--------G-T------------C--G---C--CCT--G---C-CTT-C---TC--T------T-A--------T----
NC89   -T---------G--------G--------------C--G------C-T------C-CTT-C---T----------T-A--------T----
TN89   --C----------------G-----------T----GGA-------T-T----------C---T--------C-T-A----C---T----
MI89   -T---------G--------G-----------T-AC--G------C-T------C-CT--C---T----------T-A--------T----
MS89   -T---------G--------G--------------C--G------C-T------C-CT--C---T----------T-A--------T----
CT90   -T--------------------G--------------C--G-T----C-T------C-CT--C---T----------T-A--------T----
MD90A  ------G-------------G--------T-------GG-----C--T-T----------C---T--G-------T-A-A-C----T----
RI90   GT--------T---------G----G---------C--G-T----C-T----A-C-CT--C---T----------T-A--------T---T
NY91D  ----------------G----G------------T-C--G-T----C-T----A-C-CT--C---T----------T-A--------T---T
```

FIG. 3. Nucleotides within the 1,360-base maximally variable regions, differing among isolates, used for phylogenetic analysis of the North American EEE virus group. Numbers above the nucleotides indicate 26S genome positions according to the numbering by Weaver et al. (26). —, same nucleotide as VA33; X, deletion. Sequence data for VA33, NJ45, DR49, MA56, NJ59, MA77, MD79, WI80, MI80, FL82, NJ82, MD85B, and MD88 are from Weaver et al. (29); MS83 data are from Weaver et al. (26).

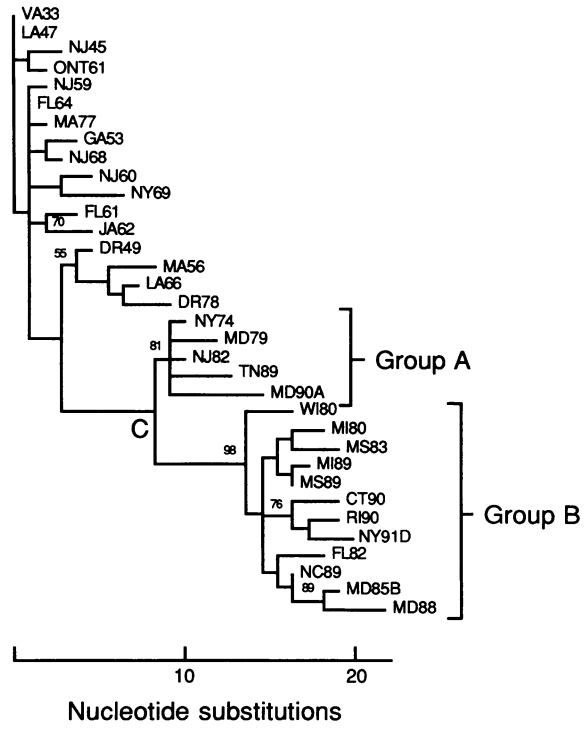

FIG. 4. One of 24 phylogenetic trees, obtained after successive reweighting, depicting relationships among North American variety EEE virus isolates from sequences described in Fig. 3. Numbers indicate bootstrap confidence values for monophyletic groups defined by adjacent nodes (hypothetical ancestors). The virus codes are defined in Table 1. The scale below the tree indicates the numbers of nucleotide substitutions represented in branch lengths. Node C represents the hypothetical ancestor of groups A and B.

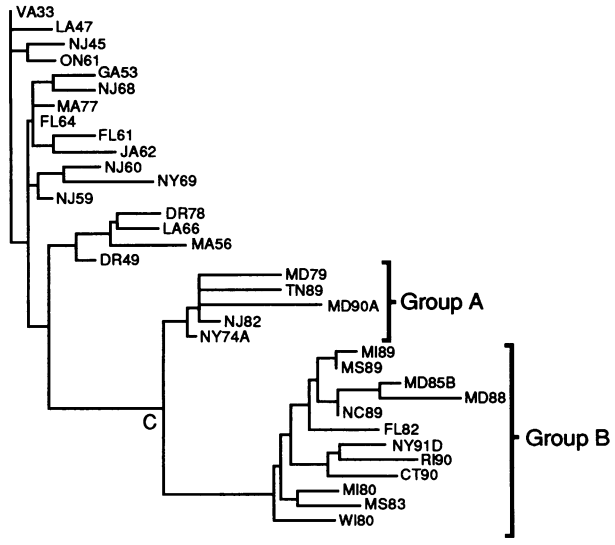

FIG. 5. Phylogenetic tree of North American variety EEE viruses obtained from sequences described in Fig. 3, by using the PAPA3 program. The virus codes are defined in Table 1. Node C represents the hypothetical ancestor of groups A and B.

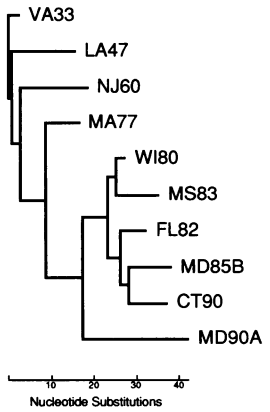

FIG. 6. The most-parsimonious phylogenetic tree for North American EEE virus strains generated from complete 26S nucleotide sequences. The virus codes are defined in Table 1.

into one of three patterns which were completely consistent with sequences of isolates placed by PAUP into groups A and B or the 1933-to-1978 group (see above); one group, assigned to group A, had the following nucleotides: 2062-T, 2107-G, 2113-G, 3577-T, 3604-T, 3619C, and 3649-C (Table 3); those isolates assigned to group B had the following nucleotides: 2062-C, 2107-A, 2113-G, 3577-T, 3604-C, 3619-T, and 3649-C; strains AL60, GA62, and NY71 were assigned to the 1933-to-1978 group on the basis of the following nucleotides: 2062-T, 2107-A, 2113-A, 3577-T, 3604-T, 3619-C, and 3649-T (Table 3). Group A included isolates from Georgia, Massachusetts, Maryland, New Jersey, New York, and Tennessee; group B included isolates from Connecticut, Florida, Georgia, Massachusetts, Maryland, Michigan, Mississippi, North Carolina, Ohio, Rhode Island, and Wisconsin. One strain, a 1984 Maryland unpassaged isolate (MD84) from a sentinel bobwhite quail, contained virions representing both groups (Table 3). This isolate is described elsewhere in greater detail (34).

The order of divergence of the isolates (VA33, LA47, NJ60, and MA77, followed by the 1980-to-1990 isolates) in the RNA fingerprint trees (Fig. 7), and the 26S tree (Fig. 6), seems unlikely to have occurred by chance alone. The lack of genomes belonging to groups A and B among 19 EEE strains isolated from 1933 to 1969 also seems unlikely to reflect sampling error, although this possibility cannot be ruled out. Most information therefore supports the concept of a single overall North American group from 1933 to 1970, followed by divergence of groups A and B during the early 1970s. However, additional EEE strains, or possibly additional sequence data, are needed to determine with greater certainty whether distinct EEE virus groups (possibly ancestors of groups A and B) were cocirculating from 1933 to 1969.

All 1990-to-1991 isolates from the upstate New York focus (NY90A, NY90B, NY90B to NY90F, and NY91A to NY91D) (Fig. 7, Upstate New York 1990-91 group) formed an exclusive monophyletic group in all 36 most-parsimonious trees generated from fingerprint data; the two 1988 isolates from upstate New York (NY88A and NY88B) also formed an exclusive group. Most of the Maryland isolates, as well as the downstate New York isolate NY90C, and a minority genome (plaque 6; see reference 34) within the NY88A isolate formed a separate group [Fig. 7, NY88A(p6)]. This suggested relative isolation of the upstate transmission focus from EEE virus immigration.

Rates of evolution. Rates of EEE virus evolution could not

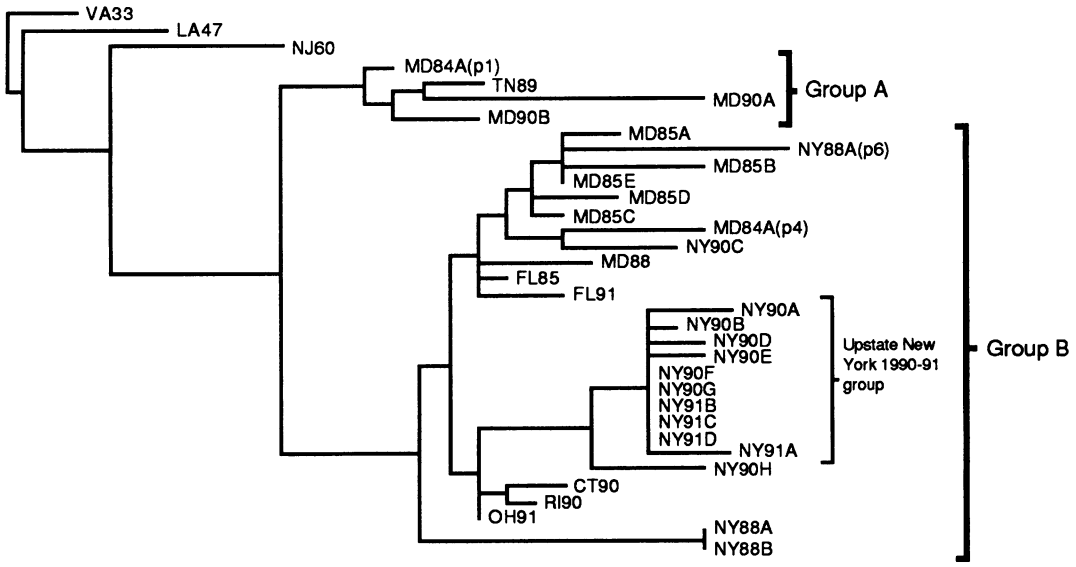

FIG. 7. Consensus phylogenetic tree showing groupings found in all 16 equally parsimonious trees (after successive weighting) generated from EEE virus oligonucleotide fingerprint data. The virus codes are defined in Table 1.

be estimated by using tree branch lengths because the year of occurrence of hypothetical ancestors (nodes in the trees) could not be determined accurately. Rates of evolution were therefore estimated as the slope of a linear regression for the number of nucleotide differences by the year of virus isolation (1). For the North American variety, data from the 1,360-base maximally variable regions were analyzed with respect to the oldest (VA33) isolate. The result is shown in Fig. 8A. Regression analyses gave estimated evolutionary rates (slopes) of 0.002%/year for isolates in the 1933-to-1978 group, and 0.027 and 0.041%/year for monophyletic groups A and B, respectively. Although this method suffers from pseudoreplication of mutations accumulating in ancestral viruses (internal branches in the tree), these data suggest that the rate of evolution in these regions of the genome increased about 10-fold concurrent with divergence of groups A and B. The relatively long branches preceding divergence of groups A and B (Fig. 4 and 5) also suggest a period of relatively rapid change preceding the divergence event. However, we cannot rule out the possibility that we failed to detect earlier divergence of groups A and B (prior to the 1970s) because of sampling limitations. Even if rooting the North American tree with the oldest isolate is incorrect (for example, if different progenitors of groups A and B were present from 1933 to 1969), the shorter distances among isolates within the 1933-to-1969 group, compared with longer distances among isolates within groups A and B, suggest that EEE viruses have evolved more rapidly since the 1970s.

The complete 26S data were also analyzed in the same manner. Regression yielded an average 26S evolutionary rate for 1947 to 1990 of 0.016%/year or 1.6×10^{-4} substitution per nucleotide per year; this is similar to the previous estimate of 1.4×10^{-4} obtained from direct comparison of VA33 and MD85 isolates (38). The rate of evolution of the South American variety of EEE virus was also estimated by using sequences from the Argentina-Panama group. The VE76 isolate was excluded because PAUP trees indicated that it diverged from this group earlier than all other isolates (Fig. 2). Figure 8B shows results of the regression analysis which yielded an estimated evolutionary rate (slope) of 0.043%/year or 4.3×10^{-4} substitution per nucleotide per year. Although the number of isolates used in this analysis was relatively small, there was no apparent change in this rate from 1938 to 1986. Regression data for the South American variety viruses yielded 95% confidence limits of 0.016 to 0.069%/year. Although these confidence limits are not entirely valid because of pseudoreplication in the regression (see above), they provide an estimate of the uncertainty in the evolutionary rate estimate.

We estimated times of divergence events during evolution of the EEE complex using the above evolutionary rate estimates of 1.6×10^{-4} and 4.3×10^{-4} substitutions per nucleotide per year for the North and South American varieties, respectively, and nucleotide substitutions (branch lengths) separating all virus isolates in the North American and Argentina-Panama groups from hypothetical common ancestors (Fig. 2, nodes Y and Z). This method yielded estimates of 410 to 800 years since the divergence of the main North and South American EEE virus groups; the estimate for the divergence of the two main (Argentina-Panama and Brazil-Peru) South American groups was 225 to 300 years ago. However, these estimates may include error resulting from sequential substitutions at nucleotide sites. To compensate, the phylogenetic tree was redrawn with branch lengths reflecting transversions only (data not shown). Since North American 26S sequence data indicated that EEE virus transitions occur roughly five times as often as transversions (38), the number of transversions was multiplied by 6 to yield an estimate of the total number of nucleotide substitutions between viruses and common ancestors (nodes A and B in the tree). Estimates for divergence of the North and South American groups were 1,290 to 1,390 years ago by using the average 26S North American EEE virus evolution rate of 0.016%/year (see above) and 1,035 to 1,040 years ago by using the average Argentina-Panama rate of 0.043%/year. The 95% confidence values from the South American regression analysis yielded estimates of 380 to 1,650 years ago. Divergence of the two main South American variety groups occurred an estimated 450 to 480 years ago, with 95% confidence limits of 165 to 770 years ago.

Estimates of divergence of the North and South American

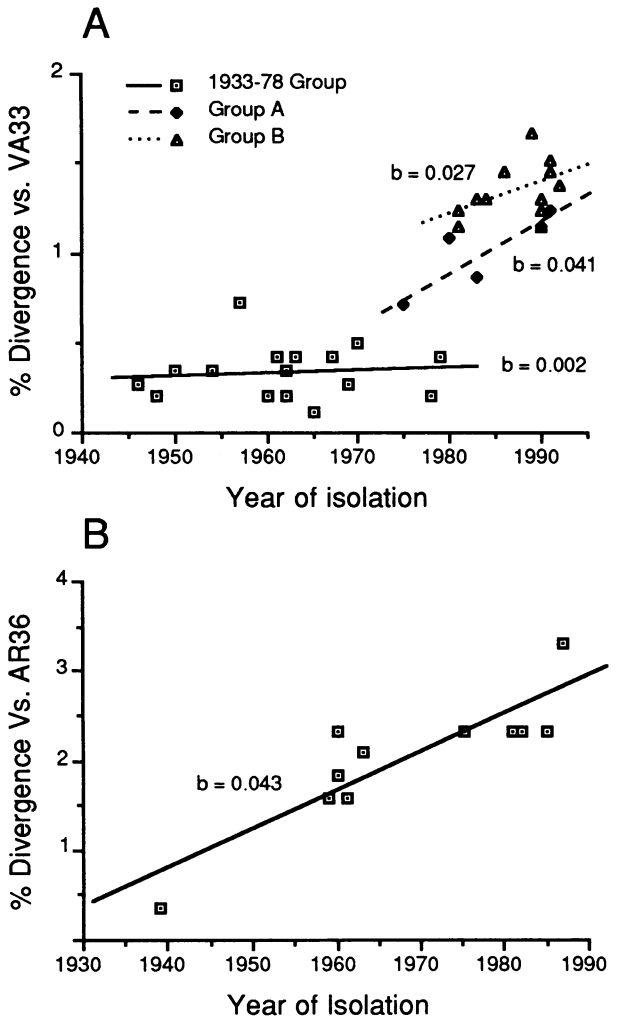

FIG. 8. Rates of nucleotide substitution of EEE virus. The slopes (b), printed adjacent to the lines, indicate the rates of evolution expressed as percent divergence per year. (A) Rate for 1,360-base maximally variable sequence regions of North American variety isolates. The ordinate shows percent divergence with respect to the VA33 isolate. Linear regression was used to plot average evolutionary rates for the 1933 to 1978 single monophyletic group, as well as the two 1974 to 1991 groups (A and B). (B) Rate of nucleotide substitution for South American variety EEE virus strains in the Argentina-Panama group. The ordinate shows percent divergence with respect to the AR36 isolate.

variety groups were also made, independently of phylogenetic trees, by using the one parameter formula of Gojobori et al. (15) and nucleotide differences between the oldest (AR36, BR56, and VA33) isolates. These estimates were about 800 years, by using the Argentina-Panama group evolutionary rate of 0.043%/year (295 to 1,280 years by using 95% confidence regression values) and about 2,200 years by using the North American rate of 0.016%/year.

DISCUSSION

EEE complex evolution. Our phylogenetic trees indicated that an ancestral EEE virus first diverged into North and South American groups about 800 to 2,200 years ago (Fig. 2, node Y); the South American group later diverged into distinct Peru-Brazil and Argentina-Panama groups roughly 450 years ago (Fig. 2, node Z). Divergence of these two main South American groups into smaller monophyletic groups may have occurred more recently, while divergence of North American groups A and B probably occurred during the early 1970s. The time estimates for the early divergence events may include a large error factor because North American EEE virus data indicate that the rate of evolution may vary by 10-fold or more. The estimate of 800 to 1,000 years may be more accurate than 1,400 to 2,200 years because the evolutionary rate of the South American group appears to be more uniform than that of the North American group, and the South American rate is similar to that estimated for other alphaviruses (37).

While overall relationships among viruses within the Argentina-Panama group in most PAUP trees and the maximum likelihood tree were based on time of isolation (Fig. 2; note AR36-38 strains diverge earliest, while the PA84 and PA86 isolates are terminal), some smaller groupings were based on the location of isolation (Fig. 2). This pattern suggested that EEE virus evolution can be regionally independent for several years. However, the overall time dependence of groupings suggested either (i) regional extinction of EEE virus lineages, followed by virus reintroduction, or (ii) occasional elimination of regionally divergent viruses following competition with an immigrant genotype. Because their transmission requires a complex series of events which must consistently occur during proper time intervals and at appropriate locations, arthropod-borne viruses are expected to exist in conditions of disequilibrium and therefore may frequently undergo extinction (28, 37).

The exclusivity of the NY90 and NY91 (upstate) EEE virus group in the tree generated from fingerprint data (Fig. 7) suggests that this focus may be relatively isolated from virus immigration via infected birds or mosquitoes. There are at least two obvious explanations for the exclusivity of this group: (i) the simplest explanation is that EEE virus overwinters in the upstate New York focus, which is relatively isolated from immigration of other EEE genomes, and/or (ii) annual reintroduction occurs in a highly specific pattern from a subtropical focus of continuous transmission that we did not sample. A mechanism for alphavirus overwintering in temperate climates (such as transovarial transmission) has not been demonstrated (26). Evaluation of the second hypothesis requires additional 1990-to-1991 isolates from subtropical areas of continuous transmission, which are not now available.

Several mechanisms of diversification for mosquito-borne viruses have been described. Eldridge (9) proposed that bunyaviruses in the California serogroup have diversified in concert with their mosquito vectors. Indirectly, our findings do not support codiversification of EEE complex viruses and their mosquito vectors because higher eucaryotes require much longer time frames for diversification than RNA viruses (31). Because (i) VEE and EEE viruses appear to have descended from a common ancestor in tropical America (35) and (ii) enzootic VEE viruses and South American EEE viruses are transmitted among small mammals by *Culex* (*Melanoconion*) mosquitoes, the most parsimonious evolutionary scenario would include an ancestral EEE virus utilizing these hosts in South and/or Central America. Therefore, introduction of an ancestor of the extant group of EEE viruses into North America probably involved host switching to *C. melanura* and songbirds. Colonization by geographically disjointed founder populations, followed by adaptive radiation and/or genetic drift, may be an important mechanism of evolution by New World alphaviruses (28, 33).

Rates of evolution. Our estimates of evolutionary rates for

North and South American EEE viruses are ca. 10-fold lower than those for many non-vector-borne RNA viruses (37). Several factors which may restrain alphavirus evolution have been reviewed previously (37, 38). Our South American EEE virus estimate of 4.3×10^{-4} substitution per nucleotide per year is very close to the estimate of 5×10^{-4}, derived from RNA fingerprint data, for the ID South American subtype of VEE virus (37). The difference in rates between North and South American EEE virus varieties (roughly threefold) may be at least partially due to temperature differences in habitats of transmission. If numbers of genome replication cycles are affected by lower ambient temperatures in North America versus South America, to which poikilothermic mosquito vectors are exposed, evolutionary rates of EEE virus may be correspondingly affected.

The 26S genome region of EEE viruses in the North American variety appears to have undergone a variable rate of evolution (nucleotide substitution) from 1945 to 1991. This finding contrasts with the constant, higher rate of evolution in the nonstructural genes of influenza viruses (1). The neutral theory of molecular evolution predicts that genomes primarily acquire neutral or synonymous nucleotide substitutions in a clock-like manner over time (20). Our results indicate that some RNA viruses may not follow the molecular clock pattern of evolution when viewed over a relatively short time frame.

The factor(s) responsible for the apparent increase in evolutionary rate of North American EEE virus during the 1970s is not known. One possibility is that virus dispersal and/or population size has been altered by changes in mosquito and/or vertebrate host populations. Populations of most passerine birds that breed in forests of eastern North America and migrate to the neotropics declined from 1978 to 1987 after a period of stability or increasing abundance (24). Many of these birds serve as hosts for EEE virus (27), suggesting that reductions in their populations may have affected dispersal of viruses among transmission foci and/or amounts of virus circulating in North America.

Another possibility is that divergence into two groups during the 1970s affected rates of evolution. This hypothesis bears superficial resemblance to the punctuated equilibrium theory of evolution, which predicts that organisms (eucaryotic) undergo rapid evolutionary change during peripatric speciation events (8). This mechanism involves founder effects in small, peripheral, isolated populations causing rapid genetic drift, resulting in reproductive isolation from the parent population. Later, the new form expands into the range of the parent, resulting in sympatric species. Because haploid alphaviruses cannot be considered sexual in this sense (recombination has only been detected in one alphavirus [16]), the same reproductive isolation mechanisms which lead to eucaryotic speciation cannot apply. However, spatial and temporal isolation may be required to prevent competitive exclusion of incipient or parental alphavirus groups. These concepts are discussed in greater detail elsewhere (33, 37). The dual infection of an avian EEE virus host (MD84) by members of both North American groups indicates that these genotypes are sometimes sympatric and may compete in nature. However, North American groups A and B could be spatially isolated in some North American regions not sampled by our collection of isolates. One possibility is that obscure ecological changes may have altered patterns of virus movement during the 1970s (e.g., dispersal patterns of birds or mosquitoes), resulting in isolation of groups A and B, and possibly affecting rates of evolution. More complete information regarding virus dispersal and the distribution of groups A and B is needed to evaluate this hypothesis.

Other possible mechanisms for an increase in the EEE virus evolutionary rate, coincident with divergence of distinct monophyletic groups include (i) reduction in the constraining effect of selective pressure; (ii) increased opportunities for founder effects and genetic drift, associated with a reduction in virus population sizes (39); (iii) changes in the environment or virus hosts which favored different viruses; and (iv) changes in the efficiency of dispersal of viruses in different groups. Theoretically, a weakly selected trait is more constrained by selection if the population is relatively large (39). Although most of the nucleotide substitutions accumulating in groups A and B were synonymous, selection could also act on primary RNA structure via codon usage preferences, RNA secondary structure, or packaging requirements. More information on the role of primary RNA sequence in the fitness of EEE virus genomes is needed to evaluate these hypotheses.

ACKNOWLEDGMENTS

We thank Charles Calisher, Margaret Grayson, John Howard, Nick Karabatsos, Robert Shope, and Robert Tesh for contributing virus isolates for this study. Estelle Bussey provided excellent technical assistance and Charles Mitter, David Swofford, and Brian Wiegmann made helpful suggestions concerning the manuscript.

This research was supported by National Institutes of Health grants AI26787, AI14627, and AI22119 and the Maryland Agricultural Experiment Station.

REFERENCES

1. **Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch.** 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. Science **232**:980–982.
2. **Calisher, C. H., N. Karabatsos, J. P. Foster, M. Pallansch, and J. T. Roehrig.** 1990. Identification of an antigenic subtype of eastern equine encephalitis virus isolated from a human. J. Clin. Microbiol. **28**:373–374.
3. **Calisher, C. H., R. E. Shope, W. Brandt, J. Casals, N. Karabatsos, F. A. Murphy, R. B. Tesh, and M. E. Wiebe.** 1980. Proposed antigenic classification of registered arboviruses. Intervirology **14**:229–232.
4. **Casals, J.** 1964. Antigenic variants of eastern equine encephalitis virus. J. Exp. Med. **119**:547–565.
5. **Cornish-Bowden, A.** 1983. Phenetic methods of classification use information that is disregarded by minimum-length methods. J. Theor. Biol. **101**:317–319.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
7. **Doolittle, R. F., and D.-F. Feng.** 1990. Nearest neighbor procedure for relating progressively aligned amino acid sequences. Methods Enzymol. **183**:659–669.
8. **Eldredge, N., and S. J. Gould.** 1972. Punctuated equilibria: an alternative to phyletic gradualism, p. 82–115. *In* T. J. M. Schopf (ed.), Models in paleobiology. Freeman, Cooper and Co., New York.
9. **Eldridge, B.** 1990. Evolutionary relationships among California serogroup viruses (*Bunyaviridae*) and *Aedes* mosquitoes (*Diptera*: *Culicidae*). J. Med. Entomol. **27**:738–749.
10. **Farris, J. S.** 1988. Hennig86, version 1.5. Distributed by the author, Port Jefferson, N.Y.
11. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.
12. **Felsenstein, J.** 1990. PHYLIP, manual version 3.3. University Herbarium, University of California, Berkeley, Calif.
13. **Fichot, O., and M. Girard.** 1990. An improved method for sequencing of RNA templates. Nucleic Acids Res. **18**:6162.
14. **Fitch, W. M.** 1984. Cladistic and other methods: problems, pitfalls, and potentials, p. 221–252. *In* T. Duncan and T. F. Stuessy (ed.), Cladistics. Columbia University Press, New York.

15. **Gojobori, T., E. N. Moriyama, and M. Kimura.** 1990. Statistical methods for estimating sequence divergence. Methods Enzymol. **183:**531–550.
16. **Hahn, C. S., S. Lustig, E. G. Strauss, and J. H. Strauss.** 1988. Western equine encephalitis virus is a recombinant virus. Proc. Natl. Acad. Sci. USA **85:**5997–6001.
17. **Hillis, D. M.** 1991. Discriminating between phylogenetic signal and random noise in DNA sequences, p. 278–294. *In* M. M. Miyamoto and J. Cracraft (ed.), Phylogenetic analysis of DNA sequences. Oxford University Press, Oxford.
18. **Hillis, D. M., and J. J. Bull.** 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. **42:**182–192.
19. **Kew, O. M., B. K. Nottay, and J. F. Obijeski.** 1984. Applications of oligonucleotide fingerprinting to the identification of viruses. Methods Enzymol. **8:**41–84.
20. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, London.
21. **Kinney, R. M., B. J. B. Johnson, J. B. Welch, K. R. Tsuchiya, and D. W. Trent.** 1989. Full-length nucleotide sequences of the virulent Trinidad donkey derivative, strain TC83. Virology **170:**19–31.
22. **Morris, C. D.** 1988. Eastern equine encephalomyelitis, p. 1–20. *In* T. P. Monath (ed.), The arboviruses: epidemiology and ecology, vol. III. CRC Press, Boca Raton, Fla.
23. **Nei, M.** 1991. Relative efficiencies of different tree-making methods for molecular data, p. 90–128. *In* M. M. Miyamoto and J. Cracraft (ed.), Phylogenetic analysis of DNA sequences. Oxford University Press, Oxford.
24. **Robbins, C. S., J. R. Sauer, R. S. Greenberg, and S. Droege.** 1989. Population declines in North American birds that migrate to the neotropics. Proc. Natl. Acad. Sci. USA **86:**7658–7662.
25. **Roehrig, J. T., A. R. Hunt, G.-J. Chang, B. Sheik, R. A. Bolin, T. F. Tsai, and D. W. Trent.** 1990. Identification of monoclonal antibodies capable of differentiating antigenic varieties of eastern equine encephalitis viruses. Am. J. Trop. Med. Hyg. **42:**394–398.
26. **Rosen, L.** 1987. Overwintering mechanisms of mosquito-borne arboviruses in temperate climates. Am. J. Trop. Med. Hyg. **37:**695–765.
27. **Scott, T. W., and S. C. Weaver.** 1989. Eastern equine encephalomyelitis virus: epidemiology and evolution of mosquito transmission. Adv. Virus Res. **37:**277–328.
28. **Scott, T. W., S. C. Weaver, and V. Mallampalli.** Evolution of mosquito-borne viruses. *In* S. S. Morse (ed.), Molecular evolution of viruses, in press. Raven Press, New York.
29. **Sourdis, J., and M. Nei.** 1988. Relative efficiencies of the maximum parsimony and distance-matrix methods in obtaining the correct phylogenetic tree. Mol. Biol. Evol. **5:**298–311.
30. **Strauss, E. G., and J. H. Strauss.** 1986. Structure and replication of the alphavirus genome, p. 35–90. *In* S. Schlesinger and M. Schlesinger (ed.), The togaviruses and flaviviruses. Plenum Press, New York.
31. **Strauss, J. H., and E. G. Strauss.** 1988. Evolution of RNA viruses. Annu. Rev. Microbiol. **42:**657–683.
32. **Swofford, D. L.** 1991. PAUP: phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign, Ill.
33. **Weaver, S. C.** Evolution of alphaviruses. *In* C. H. Calisher, A. J. Gibbs, and F. Garcia-Arenal (ed.), Molecular evolution of viruses, in press. Cambridge University Press, Cambridge.

33a. **Weaver, S. C.** Unpublished data.

34. **Weaver, S. C., L. A. Bellew, L. A. Gousset, P. A. Repik, T. W. Scott, and J. J. Holland.** 1993. Diversity within natural populations of eastern equine encephalomyelitis virus. Virology **195:**700–709.
35. **Weaver, S. C., L. A. Bellew, and R. Rico-Hesse.** 1992. Phylogenetic analysis of alphaviruses in the Venezuelan equine encephalitis complex and identification of the source of epizootic viruses. Virology **191:**282–290.
36. **Weaver, S. C., A. Hagenbaugh, L. A. Bellew, and C. H. Calisher.** 1992. Genetic characterization of an antigenic subtype of eastern equine encephalomyelitis virus. Arch. Virol. **127:**305–314.
37. **Weaver, S. C., R. Rico-Hesse, and T. W. Scott.** 1992. Genetic diversity and slow rates of evolution in New World alphaviruses. Curr. Top. Microbiol. Immunol. **176:**99–117.
38. **Weaver, S. C., T. W. Scott, and R. Rico-Hesse.** 1991. Molecular evolution of eastern equine encephalitis virus in North America. Virology **182:**774–784.
39. **Wright, S.** 1937. The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. USA **23:**307–320.