

# Revisiting the Clinal Concept of Evolution and Dispersal for the Tick-Borne Flaviviruses by Using Phylogenetic and Biogeographic Analyses

D. M. Heinze,<sup>a</sup> E. A. Gould,<sup>b</sup> and N. L. Forrester<sup>a</sup>

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,<sup>a</sup> and Unité des Virus Emergents, Université de la Méditerranée, Marseille, France<sup>b</sup>

**Tick-borne flaviviruses (TBF) are widely dispersed across Africa, Europe, Asia, Oceania, and North America, and some present a significant threat to human health. Seminal studies on tick-borne encephalitis viruses (TBEV), based on partial envelope gene sequences, predicted a westward clinal pattern of evolution and dispersal across northern Eurasia, terminating in the British Isles. We tested this hypothesis using all available full-length open reading frame (ORF) TBF sequences. Phylogenetic analysis was consistent with current reports. However, linear and nonlinear regression analysis of genetic versus geographic distance combined with BEAST analysis identified two separate clines, suggesting that TBEV spread both east and west from a central point. In addition, BEAST analysis suggested that TBF emerged and dispersed more than 16,000 years ago, significantly earlier than previously predicted. Thus, climatic and ecological changes may have played a greater role in TBF dispersal than humans.**

Flaviviruses are small, enveloped viruses containing a single-stranded, positive-sense RNA genome of approximately 11 kb (5). Based on phylogenetic analysis combined with virus, vector, and host relationships, the genus *Flavivirus* was originally divided into four groups: two mosquito-associated groups (either *Aedes* spp. or *Culex* spp.), one tick-associated group (*Ixodes* spp. and *Ornithodoros* spp.), and one group for which there is no known vector (19, 21). Recently, many insect-specific viruses that do not appear to infect vertebrates have been isolated and form a distinct group from the four previously proposed groups (7, 8), although some are in the mosquito-borne groups (29, 31).

Tick-borne flaviviruses (TBF) have been subdivided into mammalian and seabird-associated groups based on their specificity for vertebrate hosts and their associated ticks. Tick-borne flaviviruses known to infect humans include tick-borne encephalitis virus (TBEV), Omsk hemorrhagic fever virus (OHFV), Langat virus, Alkhurma hemorrhagic fever virus (AHFV), Kyasanur forest disease virus (KFDV), and Powassan virus (POWV). TBEV has the greatest effect on human health, with around 12,000 human cases per year and a case fatality rate ranging from 0.5 to 20% depending on location (24, 35). POWV is a rare but increasing public health risk in the United States, with around 30 to 40 clinical cases documented since its identification. This virus is also widespread in the Primorsky region of Russia, and although infection with POWV is rarely fatal, more than 50% of cases result in permanent neurologic sequelae (25).

TBEV is maintained in a cycle involving ixodid ticks and rodents. Ticks may become infected while feeding on a viremic vertebrate host or when cofeeding with infected ticks (33). After acquiring TBEV, a proportion of ticks remain persistently infected for the remainder of their life cycle. The virus is transmitted to subsequent vertebrate and invertebrate hosts via viremic and/or nonviremic transmission and also transovarially to tick progeny (9). Mathematical modeling of different transmission routes suggests that nonviremic transmission (tick to cofeeding tick) supported by transovarial transmission is sufficient to explain the observed prevalence and persistence of TBEV in natural foci; viremic transmission is of less impor-

tance due to the short period of viremia in rodent hosts (41). Large animals such as deer do not play a role in maintaining or amplifying virus, but they are very important as hosts for adult ticks (25). TBEV is transmitted to humans via the bite of infected ticks, although transmission through unpasteurized milk is also documented (13). In the case of other tick-borne flaviviruses, human infection through contact with infected animals is reported for Omsk hemorrhagic fever virus and Alkhurma hemorrhagic fever virus (45, 52); no natural human cases have been reported for Langat virus or Kadam virus (4).

POWV was isolated in 1958 from a fatal case of encephalitis in Powassan, Ontario (40). Molecular epizootiologic studies have shown that it exists in nature as two distinct lineages (14). Lineage I cycles between *Ixodes cookei* ticks and small- or medium-sized mammals such as groundhogs (*Marmota monax*), while lineage II cycles in *Ixodes scapularis* ticks and *Peromyscus leucopus* mice (38, 42). Lineage II has been named deer-tick virus (DTV) and shows distinct geographical clustering in New England and the midwest (42). POWV has also been isolated in the Primorsky region of Russia; the high level of similarity between Russian and North American isolates and their limited genetic variability resulted in estimates that POWV was introduced into Russia from North America during the past 100 years (36).

Based on differences in the nucleotide and amino acid sequence of the envelope (E) protein, three basic subtypes of TBEV have been delineated: TBEV-Europe (western or W-TBEV), TBEV-Siberia, and TBEV-far east (FE-TBEV) (15, 26, 53). Previously, Zanotto and colleagues showed that the rate of nonsynonymous nucleotide substitutions in the E protein gene had a linear

Received 24 April 2012 Accepted 22 May 2012

Published ahead of print 6 June 2012

Address correspondence to N. L. Forrester, [naforres@utmb.edu](mailto:naforres@utmb.edu).

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

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.01013-12

correlation with distance in kilometers from an arbitrarily defined focus, suggesting a continuous geographical cline of evolution from far-eastern Eurasia toward Europe (53). The present phylogenetic study sought to test the clinal hypothesis for the evolution and dispersal of the TBF using all available full-length open reading frame (ORF) TBF sequences to provide more comprehensive data and to assess whether or not the previous hypothesis of Zanotto et al. (53) was robust. Understanding these evolutionary and biogeographic characteristics is important in identifying the determinants of epidemiology and pathogenicity of these viruses. Such studies can also contribute to taxonomic decisions and predictions of their future effect on human health.

## MATERIALS AND METHODS

**Viral strains and genome sequences.** Full-length genome sequences of tick-borne flaviviruses were downloaded from the United States National Center for Biotechnology Information (NCBI). The resulting list of tick-borne flavivirus (TBF) sequences was culled based on the availability of reasonable data on the date and locality of isolation for each strain. Because of the wide geographical distribution represented in the resulting data set, localities were considered adequate if they were specific to a region (e.g., Primorsky region, Russia). In some instances, strains are named after readily identifiable cities; these were used as a location when no other source was available. Tyulenyi virus and Meaban virus were chosen as outgroup strains based on published phylogenies (35). Table 1 shows the strains of TBF and the related location data used in this study.

**Phylogenetic analysis. (i) Sequence alignment.** Open reading frames for each sequence were identified using MacVector (MacVector Inc., Cary, NC); this information was used to trim sequences to ORFs where necessary to ensure proper translation. The resulting sequences were imported into SeaView (22) and aligned as amino acids using the MUSCLE algorithm (16). Gap-only sites were deleted, and the resulting alignment was trimmed to eliminate untranslated regions (UTRs) using the sites tool.

**(ii) Neighbor-joining analysis.** A neighbor-joining analysis (46) was conducted using PAUP (Sinauer Associates, Inc., Sunderland, MA) with the following parameters: 1,000 bootstrap replicates, randomly broken ties, and general time-reversible distance measure with a gamma distribution shape parameter of 0.5. The output tree was a bootstrap 50% majority-rule consensus tree.

**(iii) Maximum likelihood analysis.** Model Test (43) was used to estimate the best-fit evolutionary model for the sequence alignment. This model was then entered into PAUP, and a maximum-likelihood tree was generated using the Bioportal at the University of Oslo (32). Bootstraps were run in Garli using the same parameters, again using the Bioportal for computational analysis.

**(iv) Kishino-Hasegawa test.** The test between the neighbor-joining tree and the maximum likelihood tree was implemented in PAUP\* using the default settings in the program.

**(v) Beast analysis.** Bayesian Markov chain Monte Carlo (MCMC), as implemented in the BEAST package, was used to estimate the time to the most recent common ancestor (tMRCA) for the tick-borne flaviviruses. Strains were selected on the basis of having a year of collection associated with the sequence. BEAST analysis was performed using a relaxed-clock model (uncorrelated lognormal) and the SRD06 nucleotide substitution model (51). All analyses were run in duplicate to determine convergence on the Bioportal at the University of Oslo. The Bayesian skyline coalescent tree prior was determined to produce the most robust results. The MCMC was run for 80 million iterations, and convergence was assessed using the TRACER program. A maximum clade credibility tree was produced using Tree Annotator (BEAST package) with the initial 10% of trees removed as burn-in.

**Phylogeographical analysis.** Synonymous and nonsynonymous substitutions per synonymous or nonsynonymous site were measured using

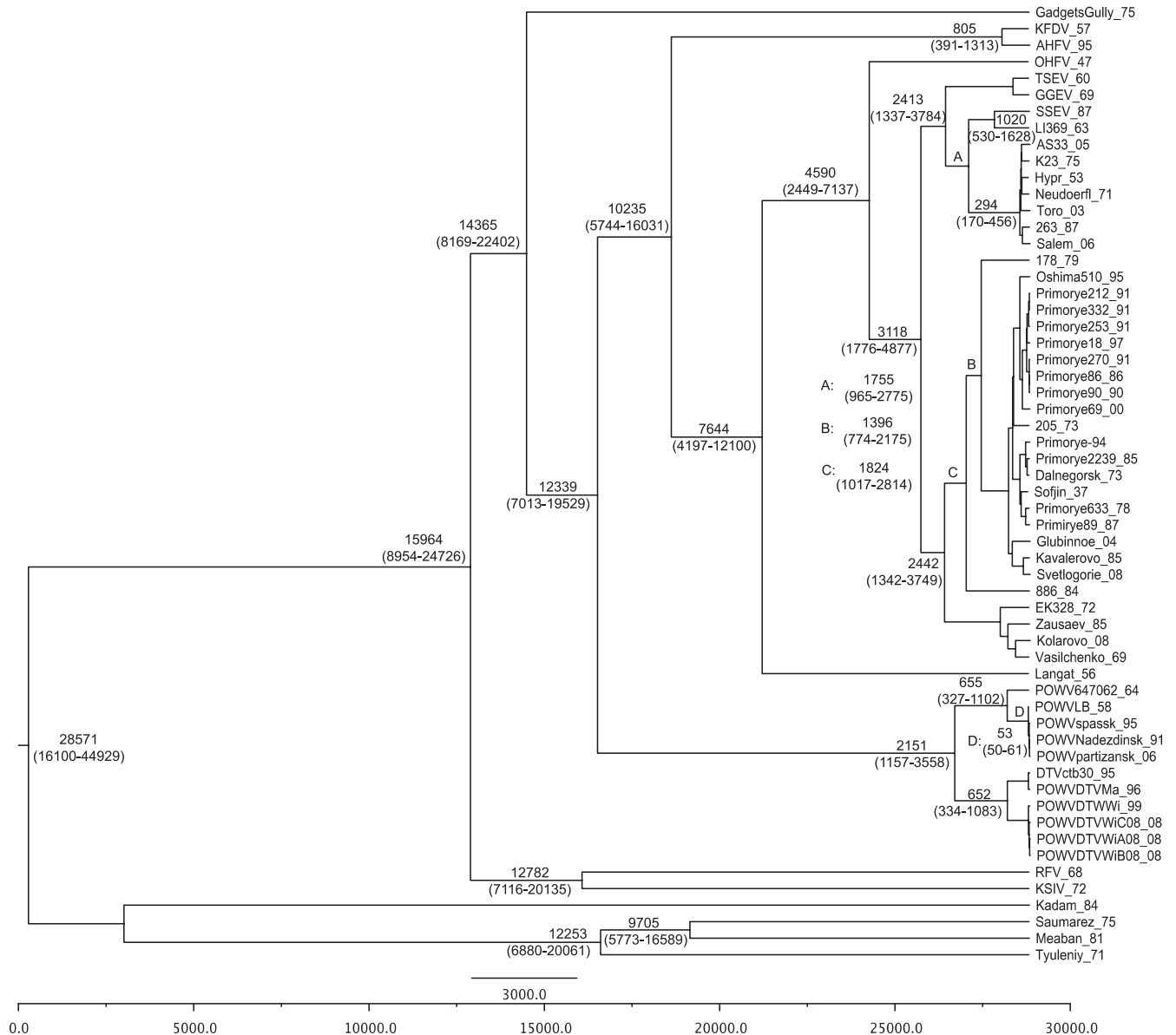
MEGA5 (49). This pairwise analysis was conducted using the Kumar model (40a). Pairwise comparisons between POWV LB (NC\_003687), OHFV Guriev (AB507800), or TBEV Sofjin-HO (AB062064) and all other sequences, excluding outgroup strains, were made and plotted versus the distance in kilometers between the point of isolation of the two viruses. These distances were calculated using Google Earth (<http://www.google.com/earth/index.html>). An assumption was made that distances should be measured across North America and Russia rather than the shortest path between isolation locations, because the shortest path often led across the Arctic, an unlikely route of dispersal. To accommodate this, a basic path between Ontario, Canada, and Primorsky Krai, Russia, was constructed, and all other distances (except for isolates in North America) were computed using this basic pathway. The TBEV Sofjin-HO data set was divided into TBEV-only and TBEV-FE-plus-TBF groups and regraphed to explore differences in dispersal. Linear and nonlinear regression was implemented in GraphPad Prism (GraphPad Software, Inc., La Jolla, CA).

## RESULTS AND DISCUSSION

**The age of the tick-borne flaviviruses and climate considerations.** To understand the temporal constraints on the origin and dispersal of the tick-borne flaviviruses, BEAST was used to estimate the time to most recent common ancestor (tMRCA) of each lineage. The resulting tree is shown in Fig. 1. The estimated time before the present for the basal node in our analysis (28,600 years; 95% HPD, 16,100 to 44,900) is surprising and suggests that the relationship between ticks and tick-borne flaviviruses has ancient roots. To our knowledge, this is the first time that full-length ORF sequences have been used to date more ancient lineages in the TBF tree. Our estimates of more recent events, such as the tMRCA for KFDV/AHFV and the TBEV group, are supported by previous work (12, 50), suggesting that our results are consistent with presently emerging temporal models of TBF evolution. Given the estimated length of time since the origin of the mammalian TBF group (16,000 to 45,000 years [16 to 45 kyr] ago), the climatic changes that have occurred over this period should be considered potential factors in the spread and distribution of these viruses. In the Northern Hemisphere, the last glacial maximum was approximately 20 kyr ago. After a rapid warming phase (termination), significant oscillations in global climate occurred between 18 and 11.5 kyr ago. Significant discharges of ice and melt water into the northern Atlantic Ocean caused two significant cold epochs (Heinrich stadial 1, from 18 to 15.5 kyr ago, and the Younger Dryas, from 13 to 11.5 kyr ago) punctuated by a period of warmer temperatures (Bølling-Allerød, from 15.5 to 13 kyr ago) (11). During this period of time, the Karshi virus (KSIV)/Royal Farm virus (RFV), Gadgets Gully virus (GGYV), and POWV lineages were predicted to have diverged from the remaining TBF. Present climatic conditions were attained 11 kyr ago at the beginning of the Holocene, although some climatic fluctuations have occurred since that time. Ice-core data suggest a widespread cooling event 8.2 kyr ago, associated with dry conditions across the Middle East and Asia (1) that is reflected in changes in pollen assemblages in sediments collected from lake Baikal (10). From the last glacial maximum to roughly 7 kyr ago, the sea level rose about 140 m. Sea level proxies, such as sedimentation rate and organic and foraminiferal carbon isotopes from cores collected in Singapore, suggest that sea level rise was relatively constant (~1.8 cm/yr) with a short cessation 7.8 to 7.4 kyr ago (2), while coral reef proxies in the Caribbean suggest there were some periods of rapid sea level increases (6.5 to 13.5 m over a century or two) at 14.2, 11.5, and 7.6

TABLE 1 Viral strains used in this analysis

Strain	Geographical origin	Yr of isolation	Source of isolate	Accession no.
Neudoerfl	Neudoerfl, Austria	1971	<i>Ixodes ricinus</i>	U27495.1
263	Temelin, Czech Republic	1987	<i>I. ricinus</i>	U27491.1
Salem	Bodensee, Germany	2006	<i>Macaca sylvanus</i>	FJ572210.1
AS33	Amberg, Germany	2005	<i>I. ricinus</i>	GQ266392.1
Hypr	Brno, Czech Republic	1953	Human blood	U39292.1
Toro-2003	Toro, Sweden	2003	<i>I. ricinus</i>	DQ401140.2
K23	Karlsruhe, Germany	1975	<i>I. ricinus</i>	AM600965.1
886-84	Irkutsk, Russia	1984	<i>Clethrionomys rufocanus</i>	EF469662.1
Kolarovo-2008	Kolarovo, Russia	2008	<i>I. pavlovskiyi</i>	FJ968751.1
Primorye-332	Mirnyy, Nadezhdinsky, and Primorsky, Russia	1991	Human blood	AY169390.3
Primorye-212	Vladivostok and Primorsky, Russia	1991	Human blood	EU816450.1
Primoye-253	Solovey-Klyuch, Nadezhdinsky, and Primorsky, Russia	1991	Human blood	EU816451.1
Primoye-270	Mirnyy, Nadezhdinsky, and Primorsky, Russia	1991	Human blood	EU816452
Kavalerovo	Kavalerovo, Russia	1985	Human	FJ402885.1
Dalnegorsk	Dalnegorsk, Russia	1973	Brain	FJ402886.1
Oshima 5-10	Kamiiso, Japan	1995	Dog blood	AB062063.2
178-79	Irkutsk, Russia	1979	<i>I. persulcatus</i>	EF469661.1
Zausaev	Tomsk, Russia	1985	Human brain	AF527415.1
Vasilchenko	Novosibirsk, Russia	1969		L40361.3
EK-328	Estonia	1972	<i>I. persulcatus</i>	DQ486861.1
205	Khabarovsk, Russia	1973	Unknown	DQ989336.1
Glubinnoe/2004	Glubinnoe, Primorsky, Russia	2004	Human brain	DQ862460.1
Svetlogorie	Svetlogor'e, Russia	2008	Human brain	GU121642.1
Sofjin-HO	Khabarovsk, Russia	1937	Human brain	AB062064.1
TSEV	Gebze area, northwestern Turkey	1960		DQ235151
LI 369/T2	Ayrshire, Scotland, United Kingdom	1963	<i>I. ricinus</i>	NC_001809.1
SSEV	Spain (Basque region)	1987	Sheep	DQ235152
GGEV	Vergina, Greece	1969	Goat brain	DQ235153
KFDV/KFD P 9605	Kyasanur Forest, Shimoga District, Karnataka, India	1957	Human blood	AY323490
AHJV	Jeddah, Saudi Arabia	1995	Human blood	NC_004355
OHFV/Guriev	Omsk region, Russia	1947	Human	AB507800
POWV/64-7062	New York	1964	Ticks on <i>Marmota</i> spp.	HM440563
POWV/DTVWiC08	Spooner, WI	2008	<i>I. scapularis</i>	HM440562
POWV/DTVWiB08	Spooner, WI	2008	<i>I. scapularis</i>	HM440561
POWV/DTVWiA08	Spooner, WI	2008	<i>I. scapularis</i>	HM440560
POWV/DTVMa96	Nantucket, MA	1996	<i>I. scapularis</i>	HM440559
POWV/DTVWi99	Chippewa Falls, WI	1999	<i>I. scapularis</i>	HM440558
POWV/Nadezdinsk-1991	Nadezdinsk region, Russia	1991	Human blood	EU670438
POWV/partizansk2006	Partizansk region, Russia	2006	Human blood	EU543649
POWV/LB	Powassan, Ontario, Canada	1958	Human brain	NC_003687
POWV/Spassk-9	Spassk region, Russia	1975	<i>Dermaecentor silvarum</i>	EU770575
DTV/ctb30	Connecticut	1995	<i>I. scapularis</i>	AF311056
KSIV	Beshkent, Uzbekistan	1972	<i>Ornithodoros papillipes</i>	NC_006947
Meaban virus	Meaban Island (French island)	1981	<i>O. maritimus</i>	DQ235144
Tyuleny virus	Three Arch Rocks National Wildlife Refuge, OR	1971	<i>I. uriae</i>	DQ235148
Royal Farm Virus	Kabul City, Afghanistan	1968	<i>Argas hermanni</i>	DQ235149
Langat TP21	Kuala Lumpur, Malaysia	1956	<i>I. granulatus</i>	NC_003690
Kadam virus	Kadam Mountain, Nakapiripirit, Karamoja, Uganda	1967	<i>Rhipicephalus pravus</i>	DQ235146
Gadgets Gully virus	Hurd Point, Macquarie Island	1975	<i>I. uriae</i>	DQ235145
Saumarez Reef virus	Australia	1977		DQ235150
Primorye-18	Russia	1997	Human blood	GQ228395.1
Primorye-86	Russia	1986	Human brain	EU816455.2
Primorye-90	Russia	1990	Human brain	FJ997899.1
Primorye-2239	Russia	1985	<i>I. persulcatus</i>	HM859895.1
Primorye-633	Russia	1978	Mouse	HM859894.1
Primorye-89	Russia	1987	Human brain	FJ906622.1
Primorye-94	Russia	1994	Human blood	EU816454.1
Primorye-69	Russia	2000	Human blood	EU816453.1



**FIG 1** Maximum clade credibility tree of the tick-borne flaviviruses. The node labels indicate the tMRCA, with the 95% HPD in parentheses. All nodes labeled with the tMRCA had a posterior probability of 1 (data not shown). Strains are labeled as strain name\_year of collection.

kyr ago (3). Under either scenario, the rise in sea level separated Siberia and Alaska 12.7 to 12.9 kyr ago (11,000  $^{14}\text{C}$  years ago, calibrated according to reference 44) (18) and Britain from mainland Europe roughly 8 to 6.5 kyr ago (6).

These climatic conditions likely limited northern dispersal of early lineages of mammalian TBF, such as KSIV/RFV, GGYV, and POWV, into areas free from glacial ice. Interestingly, the locations of isolation of AHFV/KFDV and Langat virus are still in southern Asia despite the predicted divergence of these lineages (roughly 10.2 and 7.6 kyr ago, respectively) well after modern climatic conditions were reached. However, this may be due to an incomplete understanding of the ranges of these viral lineages.

**Phylogenetic analysis.** Phylogenetic trees produced by neighbor-joining, maximum likelihood, and Bayesian methods shared

similar topology and were consistent with previous reports (23). The maximum likelihood and neighbor-joining trees showed equivalent likelihoods as defined using the Kishino-Hasegawa test ( $P < 0.0001$ ). In the interest of space, only the BEAST tree is shown (Fig. 1), while the maximum likelihood tree can be accessed in the supplemental material (see File S1 in the supplemental material). As expected, the POWV lineage showed a deep divergence from the TBF, branching between the GGYV and KFDV lineages with 98% bootstrap support. The POWV lineage then diverged into POWV and DTV groups (14). The POWV group contained lineages from the United States (POWV64\_7062), Canada (POWVLB), and Russia (POWV Nadezdinsk, Partizansk, and Spassk9). Based on tree topology, it appears the United States isolate represents the earliest lineage, followed by the Canadian isolate, while the Russian isolates represent the most recent lin-

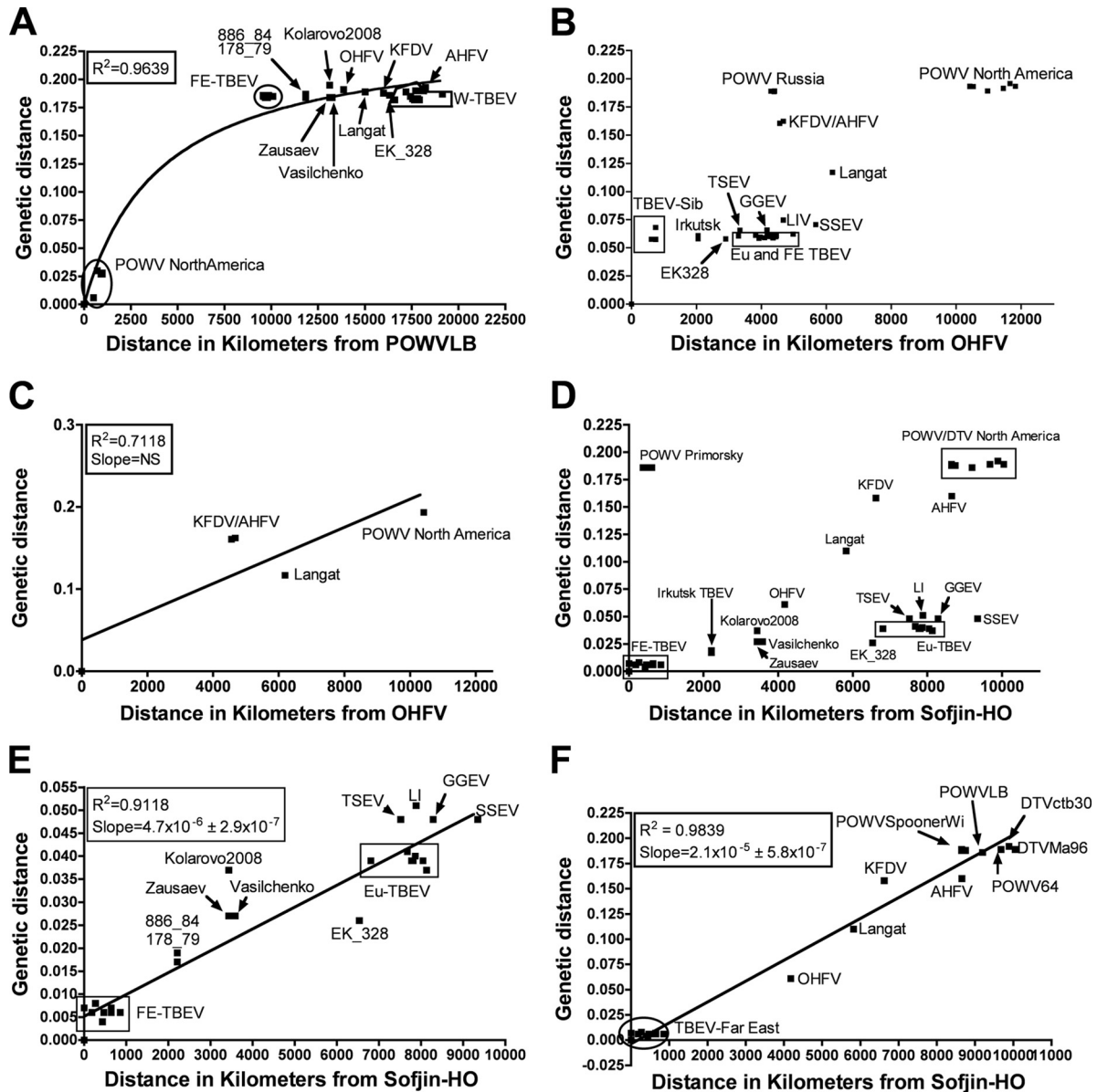


FIG 2 Genetic versus geographic distance between viral isolates was plotted for POWV LB (A), OHFV Guriev (B and C), and TBEV Sofjin-HO (D, E, and F). Linear and nonlinear regression lines are shown where calculated. For regression, Russian POWV isolates were removed from relevant analyses (A, C, and F).

age. These events have 100% bootstrap support in the maximum likelihood tree (see File S1). This topology suggests that the Russian isolates represent a recent introduction into Russia from Canada, as has been previously suggested (36).

The TBEV lineage showed a deep split between eastern and western TBEV. The eastern group contained the Siberian/Baltic lineages and FE-TBEV, while the western group contained the sheep encephalitic viruses and W-TBEV. Within the eastern TBEV group, the Siberian/Baltic lineage was the first to diverge, followed by isolates from Irkutsk and finally FE-TBEV. Within the western TBEV group, the Greek goat encephalitis virus (GGEV)/Turkish sheep encephalitis virus (TSEV) lineage was the first to diverge, followed by the Spanish sheep encephalitis virus (SSEV)/Louping ill virus (LIV) lineage, and lastly W-TBEV. All of these

nodes had 100% bootstrap support in the maximum likelihood tree and suggest a well-demarcated split between eastern and western TBEV groups. This has some implications for the clinal dispersal postulated for the TBEV, as discussed below.

**Geographic versus genetic distance.** Plotting genetic versus geographic distance from POWV LB to all sequences downstream of POWV showed rapid saturation of genetic distance. From the standpoint of POWV LB, all downstream sequences were roughly genetically equidistant from POWV LB, making this comparison uninformative as a measure of TBF dispersal, even though nonlinear regression gives a high  $R^2$  value (Fig. 2A). A similar comparison using OHFV Guriev showed rapid saturation of genetic distance for all members of the TBEV group. However, different genetic distances were found for members of the TBEV; linear re-

gression of these plots showed a moderate correlation between genetic and geographic distance with a nonsignificant slope (Fig. 2B and C). Thus, we conclude that outside the TBEV group, little information remains to allow speculation on TBF dispersal using this analysis method. In keeping with the work of Zanotto et al. (53), we also plotted genetic and geographic distance from TBEV Sofjin-HO, which occupied the most ancient lineage of the FE-TBEV viruses (Fig. 2D). This graph suggested that a correlation exists between genetic and geographic distance for the TBEV group and the TBF group. These two groupings (TBEV and TBF plus FE-TBEV) were plotted separately, and the correlation was assessed with linear regression (Fig. 2E and F). For linear regression analysis, the Russian POWV isolates were removed from the TBF-plus-FE-TBEV group, since they may represent recently introduced viruses. The slopes of both lines were significantly greater than zero, and  $R^2$  values were 0.9118 and 0.9839 for TBEV and TBF plus FE-TBEV, respectively. This suggests that genetic distance correlated with geographic distance for the TBF-plus-FE-TBEV group as well as the TBEV group. However, the correlation for the TBEV group suggests a route of dispersal from far eastern Russia toward Europe, whereas the TBF-plus-FE-TBEV analysis does not.

**POWV.** The origin and dispersal of GGYV and POWV is controversial, because their present distribution is far from any ancestral or descendant lineages. Our phylogenetic tree strongly suggested these viruses originated in the Old World in southern Asia. GGYV was isolated from *Ixodes uriae*, a seabird-associated tick, on Gadgets Gully Island in the southern ocean (48), and therefore it may have been widely disseminated by migratory seabirds. Limited sampling and the mobility of its vertebrate hosts likely explain the apparent anomaly of the location of isolation of GGYV far from Asia. The BEAST analysis (Fig. 1) predicted that the lineage that gave rise to POWV/DTV split from the TBF 12.3 (95% HPD, 7 to 19) kyr ago. While this is only an estimate, it is consistent with the possibility that this divergence event was due to movement across the Bering land bridge leading to POWV and DTV lineages circulating in the Americas. Reconstruction of climatic conditions in Beringia at this time suggests it was not colder than the present era and may even have encompassed the postglacial thermal maximum associated with a transition to a *Betula*-dominated tundra (17, 37). It is also interesting that this time period is roughly coincident with the second wave of human migrations into North America (47), providing additional evidence that climatic conditions were permissive for the dispersal of POWV. By 11 kyr ago, the Bering land bridge was flooded by the Bering Sea, isolating the ancestral POWV lineage from Old World TBF and limiting any further dispersal to the Americas. Perhaps the greatest detractor from this hypothesis is the lack of documented descendant lineages in North America. It is possible that with further sampling other viruses belonging to this lineage will be identified. Significant ecological disturbances, such as deforestation and unregulated hunting of deer and bison, occurred in North America after European colonization. These practices could have caused a drastic reduction in tick populations, leading to heavily localized distributions and potential extinction of lineages related to POWV/DTV. At a minimum, these changes would increase the difficulty of obtaining a wider sampling of potential POWV/DTV relatives.

Other scenarios are possible. The relatively close association between GGYV and POWV raises the possibility that early ancestors of POWV were also widely disseminated by seabirds. Lending some support to this speculation are data supporting the dispersal

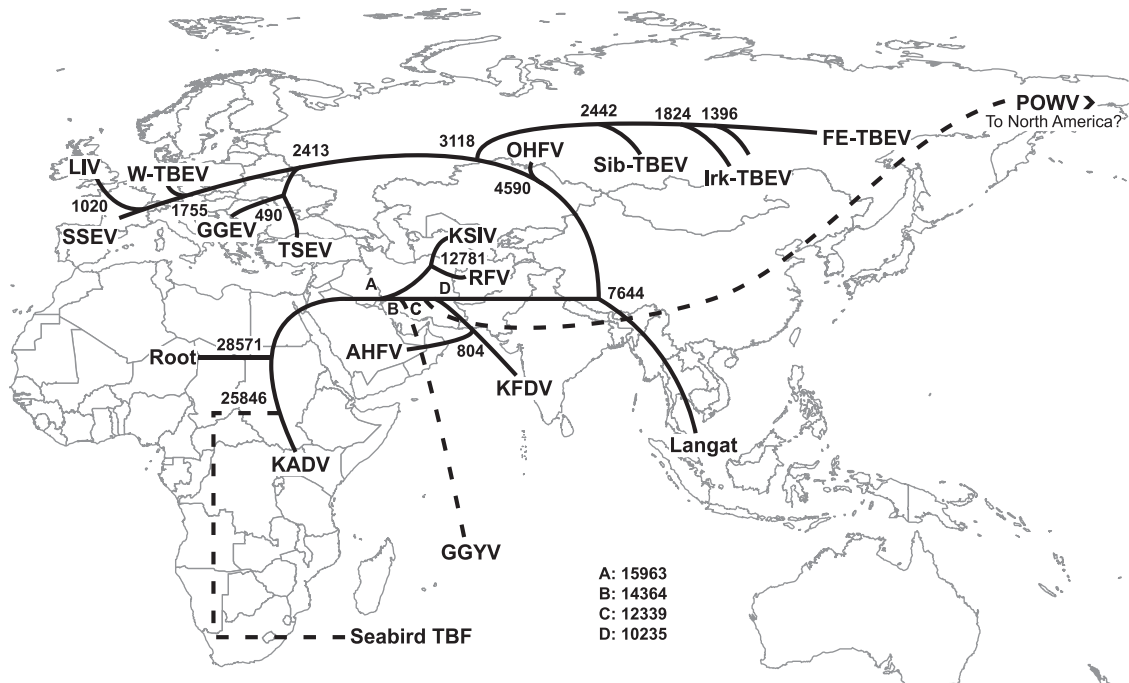
of *Borrelia* spp. by seabirds and associated *Ixodes uriae* organisms with evidence of admixture with terrestrial (mammal-*Ixodes ricinus*) cycles (20). However, if POWV were moving freely between the Americas and Siberia, we would expect several Russian lineages in the tree. In addition, if TBF movement by seabirds is common, then the apparent absence of TBEV in North America is difficult to explain. Another possibility is a very recent (100 to 300 years ago) introduction of POWV into North America, coincident with significant Russian immigration to Canada and the United States. This is very difficult to reconcile with our estimates and those of others (42) of the divergence of POWV and DTV. While it is possible this divergence occurred in Russia, it is more likely due to a New World adaptation of POWV from *Ixodes cookei*/*Ixodes marxi* cycles to *Ixodes scapularis* cycles or vice versa, since it is not clear which lineage is ancestral. Any recent introduction scenario also has to explain the lack of TBEV in North America.

The presence of POWV in the Primorsky region of Russia is difficult to explain. Based on the close relationship between the Russian POWV isolates, the original Canadian isolate, and the single Russian lineage in the tree, the most parsimonious explanation is a single introduction of POWV into Russia. Previous papers have speculated that importing American mink into Russia was responsible for this introduction (36). Given the Russian presence in Alaska until 1867, it is likely this introduction occurred during this time.

The DTV lineage has diverged to produce two subgroups, one in Wisconsin and the other in Connecticut/Massachusetts (42). While the significance of these foci is uncertain, it is interesting that the distribution of Lyme disease follows the same profile (<http://www.cdc.gov/lyme/stats/maps/map2010.html>). The phylogeography of *Borrelia burgdorferi* suggests a wide prehistoric distribution that was reduced to localized refugia in the 19th and early 20th century by deforestation and unregulated deer hunting (28). The earliest records of *I. scapularis* in North America come from the 1920s for Naushon Island, MA (34), and the 1960s for Wisconsin (30); these areas may represent refugial foci from which deer, ticks, and tick-borne disease are expanding into the United States. This may explain the DTV foci and suggests the virus will spread beyond these localized areas.

It is important to note that our model of POWV/DTV dispersal is based on our present estimates of the tMRCAs in this lineage and the present tree topology. It is possible that, with further sampling, other viruses belonging to this lineage would be identified, and these would enable a more accurate analysis of the origin of these viruses. However, given that DTV and POWV are the only two viruses identified in this lineage, a single introduction into the Americas is the most parsimonious explanation given the topology of the trees generated from the samples currently available. If POWV were being regularly transported between Russia and North America, we would expect the tree topology to reflect this rather than the distinct Russian and American POWV clades we observed. Thus, we postulate that either human involvement or significant ecological change (such as the opening or closing of the Bering land bridge) are the primary determinants of dispersal for this lineage of TBF.

**TBEV.** Northern expansion of the TBF began at least 7 kyr ago based on the predicted divergence of Langkat virus, the last Old World TBF with a location of isolation in southern Asia. According to our estimates, OHFV, a central Russian virus, diverged around 4.5 kyr ago. Thus, it seems reasonable to assert that between these divergence events, a significant northward expansion



**FIG 3** Possible model for TBF dispersal. The maximum clade credibility tree was manually plotted onto geographic space showing the most parsimonious route of dispersal for these viruses based on all complete genomic sequences. Approximate tMRCA dates are shown at all major nodes in the tree.

must have occurred. However, northern expansion before this time frame cannot be ruled out. According to our analysis, the most recent common ancestor of the TBEV group was present 3.1 (95% HPD, 1.8 to 4.9) kyr ago. This initial divergence separated the eastern from western TBEV groups. It was expected that the tMRCA of nodes within the eastern and western TBEV groups would support the clinal hypothesis of dispersal from far eastern Russia into Europe. However, the divergences of the Siberian/Baltic lineage from the eastern TBEV group and the GGEV/TSEV lineage from the western TBEV group appear almost simultaneously (2.4 [95% HPD, 1.3 to 3.7] and 2.4 [95% HPD, 1.3 to 3.8] kyr ago for eastern and western, respectively). After the divergence of the Siberian/Baltic lineages from the eastern TBEV group, viruses isolated in Irkutsk were predicted to diverge 1.8 (95% HPD, 1.4 to 3.8) and 1.4 (95% HPD, 1 to 2.8) kyr ago for TBEV strains 886\_84 and 178\_79, respectively. After the divergence of the GGEV/TSEV lineage from the western TBEV group, our analysis predicted the most recent common ancestor of the SSEV/LIV and W-TBEV groups at 1.8 (95% HPD, 1 to 2.8) kyr ago. These temporal considerations do not support a model of TBEV dispersal from eastern Eurasia moving westward but rather a dispersal beginning in central Russia moving eastwards for the eastern TBEV group and westwards for the western TBEV group. The divergence of the sheep encephalitic lineages may have been driven by contact between the forest-associated western TBEV and new hosts, such as sheep and goats brought near the forest to graze. LIV was likely introduced into the United Kingdom via transportation of animals to Ireland (39), while W-TBEV has been limited to mainland Europe due to its lack of association with livestock and the formation of the English Channel before the arrival of these lineages in Europe. These results support clinal aspects of the evolution of TBEV but do suggest a modification of the original hypothesis to

include a central Eurasian divergence of eastern and western TBEV lineages. This view is attractive, because it explains the correlation between genetic and geographic distance while remaining consistent with the phylogenetic tree and predicted times of divergence. Previous calculations postulated the divergence of Siberian and FE-TBEV 1.7 to 2 kyr ago (27), the arrival of LIV in the British Isles less than 800 years ago (39), and the divergence of TBEV-Oshima and Russian FE-TBEV 260 to 430 years ago (27). These numbers are in general support of our estimates of 2,424 (95% HPD, 1,342 to 3,749), 1,020 (95% HPD, 530 to 1,628), and 294 (95% HPD, 170 to 456) years ago for these three events.

**Conclusions.** The topology of our trees was consistent with previous reports (23). However, the inclusion of the tMRCA analysis suggests that TBF comprise a more ancient lineage than expected (Fig. 3). This suggests that TBF evolve even more slowly compared to mosquito-borne flaviviruses than previous estimates showed (54). This is likely due to differences in generation times and duration of persistent vector infection between ticks and mosquitoes and, possibly, the impact of humans as amplifying hosts for some mosquito-borne flaviviruses. Early TBF may have disseminated widely in southern Asia because of harsh climatic conditions in the north. Estimates of the tMRCA of the POWV/DTV lineage suggest that ancestral lineages had dispersed across the Bering land bridge before its closure at the beginning of the Holocene. Based on our estimates, Old World TBF lineages show clear northward movement between 7 and 5 kyr ago, culminating in a central Russian divergence of eastern and western TBEV groups. Dispersal within the TBEV group then progressed eastwards for the eastern TBEV group and westwards for the western TBEV group, producing the clinal pattern observed here and previously (53) but also refining its scope.

## ACKNOWLEDGMENTS

We thank Jonathan Auguste for his assistance with the maximum clade credibility analysis and Scott Weaver and Nikolaos Vasilakis for critical reading of the manuscript.

D.M.H. was supported by the Biodefense Training Program (NIH grant T32AI007526).

The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

There are no conflicts of interest in the publication of this work.

## REFERENCES

- Alley RB, et al. 1997. Holocene climatic instability: a prominent, widespread event 8200 yr ago. *Geology* 25:483–486.
- Bird MI, et al. 2010. Punctuated eustatic sea-level rise in the early mid-Holocene. *Geology* 38:803–806.
- Blanchon P, Shaw J. 1995. Reef drowning during the last deglaciation: evidence for catastrophic sea-level rise and ice-sheet collapse. *Geology* 23:4–8.
- Campbell MS, Pletnev AG. 2000. Infectious cDNA clones of Langat tick-borne flavivirus that differ from their parent in peripheral neurovirulence. *Virology* 269:225–237.
- Chambers TJ, Hahn CS, Galler R, Rice CM. 1990. Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* 44:649–688.
- Coles BJ. 2000. Doggerland: the cultural dynamics of a shifting coastline, p 393–401. *In* Pye K, Allen JRL (ed), *Coastal and estuarine environments: sedimentology, geomorphology and geoarchaeology*, vol 175. Geological Society of London, London, United Kingdom.
- Cook S, et al. 2012. Molecular evolution of the insect-specific flaviviruses. *J. Gen. Virol.* 93:223–234.
- Crabtree MB, Sang RC, Stollar V, Dunster LM, Miller BR. 2003. Genetic and phenotypic characterization of the newly described insect flavivirus, Kamiti River virus. *Arch. Virol.* 148:1095–1118.
- Danielova V, Holubova V, Pejnoch M, Daniel M. 2002. Potential significance of transovarial transmission in the circulation of tick-borne encephalitis virus. *Folia Parasitol.* 49:323–325.
- Demske D, et al. 2005. Late glacial and Holocene vegetation and regional climate variability evidenced in high-resolution pollen records from Lake Baikal. *Global Planet Change* 46:255–279.
- Denton GH, et al. 2010. The last glacial termination. *Science* 328:1652–1656.
- Dodd KA, et al. 2011. Ancient ancestry of KFDV and AHFV revealed by complete genome analyses of viruses isolated from ticks and mammalian hosts. *PLoS Negl. Trop. Dis.* 5:e1352. doi:10.1371/journal.pntd.0001352.
- Dumpis U, Crook D, Oksi J. 1999. Tick-borne encephalitis. *Clin. Infect. Dis.* 28:882–890.
- Ebel GD, Spielman A, Telford SR III. 2001. Phylogeny of North American Powassan virus. *J. Gen. Virol.* 82:1657–1665.
- Ecker M, Allison SL, Meixner T, Heinz FX. 1999. Sequence analysis and genetic classification of tick-borne encephalitis viruses from Europe and Asia. *J. Gen. Virol.* 80:179–185.
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi:10.1186/1471-2105-5-113.
- Edwards ME, Brubaker LB, Lozhkin AV, Anderson PM. 2005. Structurally novel biomes: a response to past warming in Beringia. *Ecology* 86:1696–1703.
- Elias SA, Short SK, Nelson CH, Birks HH. 1996. Life and times of the Bering land bridge. *Nature* 382:60–63.
- Gaunt MW, et al. 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J. Gen. Virol.* 82:1867–1876.
- Gomez-Diaz E, et al. 2011. Genetic structure of marine *Borrelia garinii* and population admixture with the terrestrial cycle of Lyme borreliosis. *Environ. Microbiol.* 13:2453–2467.
- Gould EA, de Lamballerie X, Zanotto PM, Holmes EC. 2003. Origins, evolution, and vector/host coadaptations within the genus *Flavivirus*. *Adv. Virus Res.* 59:277–314.
- 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.
- Grard G, et al. 2007. Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. *Virology* 361:80–92.
- Gritsun TS, Lashkevich VA, Gould EA. 2003. Tick-borne encephalitis. *Antiviral Res.* 57:129–146.
- Gritsun TS, Nuttall PA, Gould EA. 2003. Tick-borne flaviviruses. *Adv. Virus Res.* 61:317–371.
- Gritsun TS, et al. 1997. Complete sequence of two tick-borne flaviviruses isolated from Siberia and the UK: analysis and significance of the 5' and 3'-UTRs. *Virus Res.* 49:27–39.
- Hayasaka D, et al. 1999. Phylogenetic and virulence analysis of tick-borne encephalitis viruses from Japan and far-Eastern Russia. *J. Gen. Virol.* 80:3127–3135.
- Hoen AG, et al. 2009. Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. *Proc. Natl. Acad. Sci. U. S. A.* 106:15013–15018.
- Huhtamo E, et al. 2009. Characterization of a novel flavivirus from mosquitoes in northern Europe that is related to mosquito-borne flaviviruses of the tropics. *J. Virol.* 83:9532–9540.
- Jackson JO, DeFoliart GR. 1970. *Ixodes scapularis* Say in northern Wisconsin. *J. Med. Entomol.* 7:124–125.
- Junglen S, et al. 2009. A new flavivirus and a new vector: characterization of a novel flavivirus isolated from *Uranotaenia* mosquitoes from a tropical rain forest. *J. Virol.* 83:4462–4468.
- Kumar S, et al. 2009. AIR: a batch-oriented web program package for construction of supermatrices ready for phylogenetic analyses. *BMC Bioinformatics* 10:357. doi:10.1186/1471-2105-10-357.
- Labuda M, et al. 1993. Non-viraemic transmission of tick-borne encephalitis virus: a mechanism for arbovirus survival in nature. *Experientia* 49:802–805.
- Larrousse F, King AG, Wolbach SB. 1928. The overwintering in Massachusetts of *Ixodiphagus caucurtei*. *Science* 67:351–353.
- Lasala PR, Holbrook M. 2010. Tick-borne flaviviruses. *Clin. Lab. Med.* 30:221–235.
- Leonova GN, et al. 2009. Characterization of Powassan viruses from far eastern Russia. *Arch. Virol.* 154:811–820.
- Lozhkin AV, Anderson P, Eisner WR, Solomatkina TB. 2011. Late glacial and Holocene landscapes of central Beringia. *Quaternary Res.* 76:383–392.
- Main AJ, Carey AB, Downs WG. 1979. Powassan virus in *Ixodes cookei* and *Mustelidae* in New England. *J. Wildl. Dis.* 15:585–591.
- McGuire K, Holmes EC, Gao GF, Reid HW, Gould EA. 1998. Tracing the origins of louping ill virus by molecular phylogenetic analysis. *J. Gen. Virol.* 79:981–988.
- McLean DM, Donohue WL. 1959. Powassan virus: isolation of virus from a fatal case of encephalitis. *Can. Med. Assoc. J.* 80:708–711.
- Nei M, Kumar S. 2000. *Molecular evolution and phylogenetics*. Oxford University Press, New York, NY.
- Nonaka E, Ebel GD, Wearing HJ. 2010. Persistence of pathogens with short infectious periods in seasonal tick populations: the relative importance of three transmission routes. *PLoS One* 5:e11745. doi:10.1371/journal.pone.0011745.
- Pesko KN, Torres-Perez F, Hjelle BL, Ebel GD. 2010. Molecular epidemiology of Powassan virus in North America. *J. Gen. Virol.* 91:2698–2705.
- Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Reimer PJ, et al. 2009. Intcal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal Bp. *Radiocarbon* 51:1111–1150.
- Ruzek D, Yakimenko VV, Karan LS, Tkachev SE. 2010. Omsk haemorrhagic fever. *Lancet* 376:2104–2113.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Schurr TG. 2004. The peopling of the New World: perspectives from molecular anthropology. *Annu. Rev. Anthropol.* 33:551–583.
- St George TD, et al. 1985. The isolation of arboviruses including a new flavivirus and a new Bunyavirus from *Ixodes (Ceratiixodes) uriae* (Ixodoidea: Ixodidae) collected at Macquarie Island, Australia, 1975–1979. *Am. J. Trop. Med. Hyg.* 34:406–412.
- Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evo-



- lutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**:1596–1599.
50. Uzcategui NY, et al. 2011. Rate of evolution and molecular epidemiology of tick-borne encephalitis virus in Europe, including two isolations from the same focus 44 years apart. *J. Gen. Virol.* **93**:786–796.
  51. Volk SM, et al. 2010. Genome-scale phylogenetic analyses of chikungunya virus reveal independent emergences of recent epidemics and various evolutionary rates. *J. Virol.* **84**:6497–6504.
  52. Zaki AM. 1997. Isolation of a flavivirus related to the tick-borne encephalitis complex from human cases in Saudi Arabia. *Trans. R. Soc. Trop. Med. Hyg.* **91**:179–181.
  53. Zlotnik PM, et al. 1995. An arbovirus cline across the northern hemisphere. *Virology* **210**:152–159.
  54. Zlotnik PM, Gould EA, Gao GF, Harvey PH, Holmes EC. 1996. Population dynamics of flaviviruses revealed by molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* **93**:548–553.