Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 548–553, January 1996
Evolution

Population dynamics of flaviviruses revealed by molecular phylogenies

PAOLO M. DE A. ZANOTTO*, ERNEST A. GOULD*, GEORGE F. GAO*, PAUL H. HARVEY†, AND EDWARD C. HOLMES†‡

*National Environment Research Council Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford, OX1 3SR, United Kingdom; and †Wellcome Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, United Kingdom

Communicated by Robert M. May, University of Oxford, Oxford, United Kingdom, October 20, 1995 (received for review August 1, 1995)

ABSTRACT **The phylogeny of 123 complete envelope gene sequences was reconstructed in order to understand the evolution of tick- and mosquito-borne flaviviruses. An analysis of phylogenetic tree structure reveals a continual and asymmetric branching process in the tick-borne flaviviruses, compared with an explosive radiation in the last 200 years in viruses transmitted by mosquitoes. The distinction between these two viral groups probably reflects differences in modes of dispersal, propagation, and changes in the size of host populations. The most serious implication of this work is that growing human populations are being exposed to an expanding range of increasingly diverse viral strains.**

The genus *Flavivirus*, formerly known as the group B arboviruses, comprises more than 70 agents sharing common antigenic determinants and contains the first human virus to be isolated—yellow fever (YF) virus, which is also the prototype virus of the genus. Flaviviruses have spherical particles 40 to 50 nm in diameter with single-stranded positive-sense RNA genomes (1). They are responsible for considerable morbidity and mortality and may cause severe encephalitic, hemorrhagic, hepatic, and febrile illnesses in vertebrates, including humans.

Of particular importance for public health are the mosquito-borne dengue viruses. It is estimated that up to 100 million cases of dengue fever (DF) occur annually around the world, producing at least 250,000 cases of dengue hemorrhagic fever (DHF), with a 5% mortality rate (2). DF is primarily an urban disease of the tropics and the viruses which cause it are maintained in a cycle that involves humans and *Aedes* mosquitoes. The tick-borne encephalitis (TBE) complex, the most important arboviruses in Europe in terms of morbidity and mortality, consists of antigenically related viruses producing a variety of diseases. They are mainly found in the Northern Hemisphere, where they infect a wide range of vertebrate (mainly mammalian) species, including humans. Strains of TBE virus from far east Asia cause severe encephalitis among humans, with 40–60% case mortality (3), whereas western European strains generally cause <2% case mortality. Louping-ill (LI) virus is found in the United Kingdom and causes fatal encephalitis in sheep, grouse, and occasionally humans (4).

Previous analyses using both phylogenetic (5) and serological techniques (6) have indicated that there might be considerable differences in the evolutionary dynamics of tick- and mosquito-borne flaviviruses. For example, the tick-borne viruses appear to have dispersed westward across most of Eurasia and possibly North America, with the oldest lineages in the east and the most recent in the west, and may evolve at lower rates than mosquito-borne viruses (5, 7). Mosquito-borne viruses, on the other hand, show disjunct and cosmopolitan populations, sometimes with geographical isolation (8–10). Furthermore, comparisons of the envelope (E) genes of both vector groups, on which most phylogenetic studies are based, indicate that variability is not randomly distributed along the primary sequence but is characterized by distinct variable domains which differ between tick-borne and mosquito-borne viruses (3). This finding is of significance because the E glycoprotein, as well as being the target for neutralizing antibodies and T-cell responses, probably defines the tropism of the flaviviruses and may therefore partially determine virulence.

Flaviviruses therefore constitute informative natural models of comparative molecular evolution because they differ in host, vector, and associated disease. There is, however, clearly a need for a better understanding of their population dynamics, especially given the dramatic increase in the number of cases of DF, DHF, and dengue shock syndrome (DSS) (2, 11). Here we describe the population biology of the genus *Flavivirus* by analyzing the branching structure of phylogenetic trees reconstructed from E gene sequences.

MATERIALS AND METHODS

Compilation of Sequence Data. Complete E gene sequences were obtained from 26 tick-borne flaviviruses, including 24 isolates taken from across Eurasia and which mainly infect mammals, as well as 2 representatives of viruses associated with sea birds, the Tyuleniy (TYU) and Saumarez reef (SRE) viruses (5), and 96 mosquito-borne flaviviruses, including 66 dengue (DEN) viruses (representative of all four serotypes), 24 viruses of the Japanese encephalitis (JE) complex [including Kunjin (KUN) virus, Murray Valley encephalitis (MVE) virus, St. Louis encephalitis (SLE) virus, and West Nile (WN) virus], and 6 YF viruses. The E gene sequence from the *Aedes albopictus* cell fusion agent (CFA) virus (12) was included as an outgroup. A 1548-bp alignment was constructed by using the progressive alignment method in the MULTALIN program (13). All 123 sequences were obtained from GenBank, and a full list and alignment are available through the World-Wide Web at http://evolve.zps.ox.ac.uk//. To improve the phylogenetic analysis an additional 240-bp-long data set encompassing the junction of the E and NS1 genes from 40 representatives of DEN1 and 40 from DEN2 serotypes (10) was also included.

Phylogenetic Analysis. Phylogenetic trees were reconstructed for the entire E gene data set as well as separately for each of the tick- and mosquito-borne groups described above. Because of substitutional saturation, the third codon position was excluded in all these analyses. In each case the maximum

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DF, dengue fever; DHF, dengue hemorrhagic fever; DSS, dengue shock syndrome; E, envelope; LTTP, lineages-through-time plot; CFA, cell fusion agent; JE, Japanese encephalitis; KUN, Kunjin; LI, Louping-ill; MVE, Murray Valley encephalitis; SLE, St. Louis encephalitis; WN, West Nile; SRE, Saumarez reef; TBE, tick-borne encephalitis; TYU, Tyuleniy; YF, yellow fever.

‡To whom reprint requests should be addressed.

likelihood (FastDNAml; ref. 14), parsimony (PAUP; ref. 15), and neighbor-joining and Fitch–Margoliash methods (PHYLIP programs NEIGHBOR and FITCH, respectively; ref. 16) were used. In the case of the distance matrix programs, distances between sequences were estimated by using the six-parameter model of Felsenstein (program DNADIST) and by using the method of Nei and Gojobori (17) for nonsynonymous sites only (MEGA package; ref. 18). Optimum values of transition and transversion were estimated from the data by finding the ratios which produced trees of the highest likelihood. For some analyses of branching structure it is necessary to date nodes on the tree relative to each other. Consequently trees were also reconstructed by assuming a molecular clock using the ultrametric version of the Fitch–Margoliash method (program KITSCH). Finally, to assess the robustness of the relationships obtained, a bootstrap resampling analysis (1000 replications) was undertaken with the neighbor-joining method (PHYLIP programs SEQBOOT, DNADIST, NEIGHBOR, and CONSENSE).

Analysis of Branching Patterns. A simple graphical tool by which to visualize differences in population processes, caused by changes in the rate of lineage birth and death, is a plot of the logarithmic number of lineages in the tree against the time at which they appear. These "lineages-through-time" plots were constructed with the ENDEMIC–EPIDEMIC program (19). The shape of the plot obtained determines whether the population is more likely to have grown exponentially or to have maintained a constant size. If exponential ("epidemic") growth is indicated, a further y-axis transformation can be applied to determine whether the exponent has been constant, increasing, or decreasing through time (20).

Estimating Rates and Dates of Flavivirus Evolution. Because times of isolation are available for some sequences in the data set, it is possible to estimate rates of nucleotide substitution and subsequently the times of divergence of other nodes. There were 47 pairwise comparisons where this information was known, and for these the rates of nonsynonymous substitution per site per year (r_n) were estimated from the relation $r' = d_n/\Delta T$, where r' is an estimator of r_n (see below), d_n is the number of nonsynonymous substitutions per nonsynonymous site, and ΔT is the difference in times of isolation for the two taxa in question. Because lineages may have split before the time of sampling of the earlier taxon, time of isolation is likely to be an underestimate of the true time of divergence (T). Consequently, lower values of r' tend to be better estimators of r_n and the mean value of r' is a less reliable estimate than a single minimal r'. Moreover, because of multiple substitution, the reliability of estimations will decrease when deeper nodes are dated, although nonsynonymous rates tend to show more linearity in time than synonymous rates. However, this method is useful in that it allows independent estimations of r' and provides a minimum boundary for the times of sequence divergence.

Rate variation was also assessed by summing branch lengths, estimated by maximum-likelihood and parsimony methods, between pairs of taxa on both DNA and protein sequence trees (patristic distances). Twenty-two representatives of tick- and mosquito-borne viruses comprising single representatives of each mosquito-borne serotype as previously described (5) were included in this analysis. In the case of DNA sequences, patristic distances were estimated on first and second positions only. With the protein data, changes were weighted by use of a matrix based on physicochemical and mutational distances (21) and by one in which the number of amino acid changes is set by the minimum number of nucleotide substitutions determined by the genetic code (15, 16). Patristic distances were calculated for the TBE complex from the divergence of the TYU group and for the DEN and JE serotypes from the divergence of YF virus, for both unrooted trees and those rooted by CFA.

RESULTS

Global Phylogeny of the Flaviviruses. The maximum-likelihood tree for all 123 flaviviruses, rooted by CFA, is shown in Fig. 1. The numbers next to some of the nodes represent the percentage of supporting bootstrap replications. The deepest split is that between the tick- and mosquito-borne viruses. The mosquito-borne viruses then further divide into the YF, JE, and dengue viruses in that order. Within the dengue viruses, those assigned to DEN4 diverge first, followed by DEN2, and finally DEN1 and DEN3. All these major groupings were supported by high numbers of bootstrap replicates. This topology agrees with one obtained with fewer representatives of the entire genus (5) and with one obtained by using all codon positions (tree not shown).

Patterns of Cladogenesis in the Flaviviruses. One way in which to analyze the branching structure of phylogenetic trees is to determine the extent to which branches differ in their probability of producing daughter lineages. This can be done by use of the "relative cladogenesis statistic," P_k, which is the probability that a particular lineage, existing at time t, will have k tips (compared with the total number of tips) by time 0 (the present) (22). This allows the identification of branches that have higher-than-expected rates of cladogenesis. From Fig. 1 it is clear that the TBE group has a highly asymmetric (pectinate) topology characterized by a continuous stepwise branching process (chaining) throughout the evolutionary period. Of the 26 lineages in the phylogenetic tree, 14 produce significantly more daughters than expected under a null model of a uniform rate of cladogenesis ($P < 0.01$). In contrast, both the dengue and JE complex of viruses have much more balanced topologies with no more branches with higher rates of cladogenesis than would be expected by chance. Furthermore, and unlike the situation for the TBE group, most of the branching activity has occurred in the recent past.

Another way to analyze branching structure is to examine the distribution of lineage-splitting events through time, as this allows us to distinguish between different models of population growth and decline (20). However, this method is applicable only if there are no major differences in the rate of cladogenesis between branches, and consequently this method was applied only to the mosquito-borne viruses. Furthermore, it is necessary to use trees in which nodes are dated relative to each other (constructed with the PHYLIP program KITSCH). Although the assumption of rate constancy is controversial, the trees produced had branching patterns very similar to those seen in the maximum likelihood analysis (Fig. 1), with the majority of the branching activity occurring recently (trees not shown).

In an exponentially growing population it is expected that the semilogarithmic lineages-through-time plot (LTTP) for a full phylogeny will approximate to a straight line steepening toward the present (22, 23). However, because we have only a small sample of viruses, exponential population growth is expected to leave a different signature in the LTTP, one in which the line gently flattens toward the present (20, 23). This is caused by an underrepresentation of recently diverged nodes. A simple model of exponential population growth is clearly not sufficient to explain the LTTP of dengue viruses, which appears to have two distinct phases (Fig. 2*A*). The first phase is characterized by a gently rising line which implies a low rate of population growth. The second phase begins when the number of lineages rises sharply (a period of rapid cladogenesis), indicative of very rapid population expansion. This pattern is also apparent in the epidemic transformation of these data (Fig. 2*B*). It is also informative to examine the rapid expansion phase in more detail, and the LTTP for this period is shown in Fig. 2*C*. The slight curve over toward the present confirms that a model of exponential population growth is the most appropriate one, and the upward curvature in the epi-

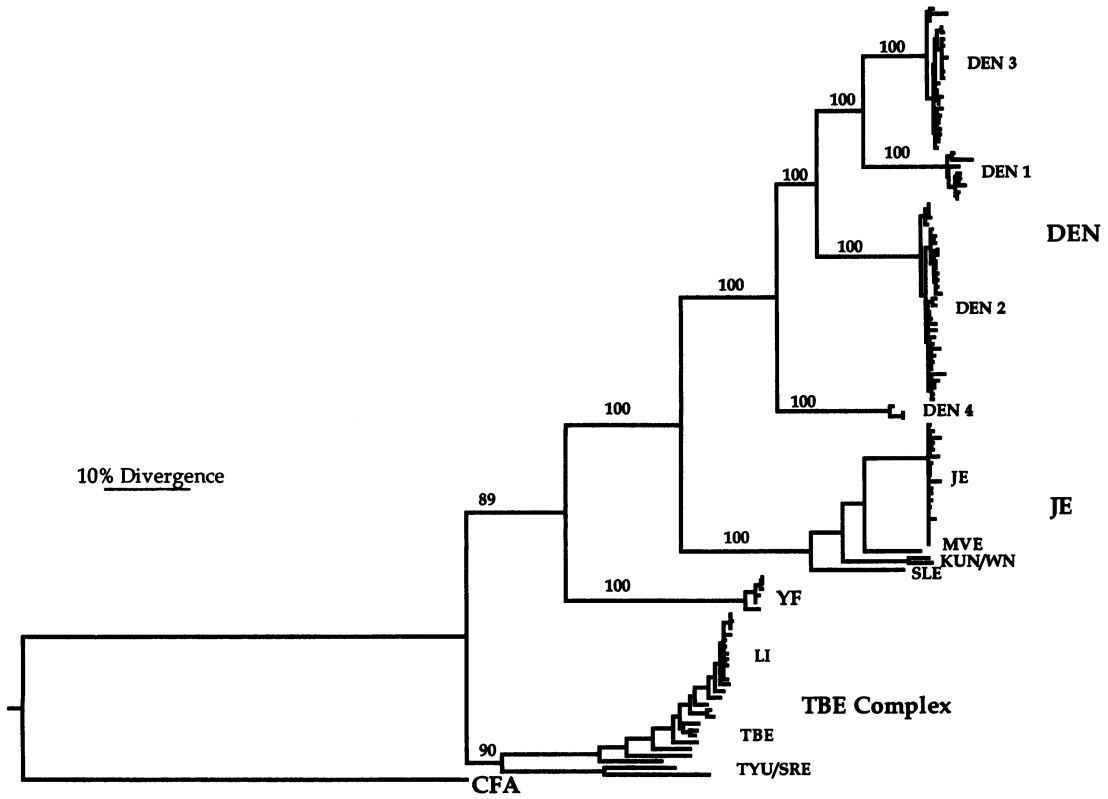

FIG. 1. Maximum likelihood tree for the E gene from 123 flaviviruses. The tree is rooted by the sequence from *Aedes albopictus* CFA virus. The numbers next to some branches represent the numbers of bootstrap replications supporting that branch. All horizontal branch lengths are drawn to scale.

demic transformation indicates that the growth exponent has been increasing through time (Fig. 2*D*). An equivalent analysis of the smaller number of JE viruses presents an identical picture of population growth (data not shown).

Estimating Rates of Nucleotide Substitution. A summary of r' values for tick- and mosquito-borne viruses is shown in Table 1. The lowest values obtained for rates of nonsynonymous substitution are 7.28×10^{-5} per site per year for the TBE-complex viruses and 7.5×10^{-5} per site per year for the mosquito-borne viruses. The observation of such similar rates is surprising given the short length of the branch leading to the tick-borne group in the maximum likelihood tree (Fig. 1), which implies a lower rate of substitution, and is probably explained by the fact that only two comparisons were made in the tick-borne group against 45 for the mosquito-borne viruses (Table 1). Therefore, rates of substitution were also estimated by calculating patristic distances between taxa on trees reconstructed for 22 representatives of the genus. This analysis, based on both DNA and protein sequences, indicates that the tick-borne viruses are evolving about 0.56 times as fast as the mosquito-borne viruses, which, if we accept the previously determined rate for the mosquito-borne viruses, means a rate of 4.12×10^{-5} nonsynonymous substitutions per site per year.

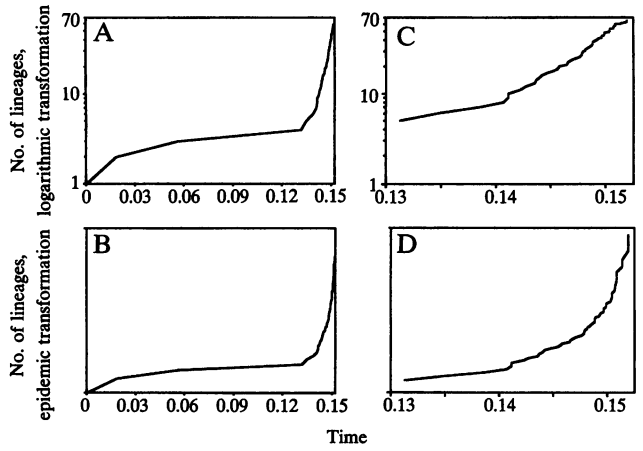

FIG. 2. (*A*) Lineages through time of 66 dengue viruses. The time at which each lineage division occurs on a constant-rate (KITSCH) phylogenetic tree is plotted with the time axis scaled as the number of nucleotide substitutions from the root of the tree to the tips. (*B*) Epidemic transformation of the same data. This transformation of the *y* axis determines whether the rate of population growth has been constant through time (straight line), increasing (upward curvature, as here), or decreasing (downward curvature). (*C*) Lineages through time of the most recent 62 nodes of dengue viruses. (*D*) Epidemic transformation of the same data.

Table 1. Independent estimates of minimum rates of nonsynonymous substitution per site per year (r') based on date of isolation

Group	r'	No. of taxa compared	Reference taxon
Mosquito-borne			
DEN1	4.42×10^{-4}	5	Japan 43
DEN2	7.50×10^{-5}	8	India 57
DEN2	2.52×10^{-4}	5	New Guinea 44
DEN2	1.62×10^{-4}	10	Trinidad 54
DEN3	2.92×10^{-4}	2	Puerto Rico 63
DEN3	6.84×10^{-4}	3	Puerto Rico 65
DEN3	3.04×10^{-4}	6	Philippines 56
DEN4	8.54×10^{-4}	2	Philippines 56
JE	2.60×10^{-4}	4	Vellore 58
Tick-borne			
TBE	1.22×10^{-4}	1 (FETBE)	FETBE/205
LI	7.28×10^{-5}	1 (LI/SB256)	LI/31

FETBE, Far Eastern tick-borne encephalitis.

Table 2. Times of divergence (in years before present) for selected nodes of mosquito-borne viruses dated by using rates of nonsynonymous substitution

	DEN1	DEN2	DEN3	DEN4	JE
DEN1	**170 ± 30**	1421 ± 96	895 ± 70	1708 ± 109	2772 ± 157
DEN2	1421 ± 96	**215 ± 34**	1451 ± 98	1690 ± 107	2935 ± 163
DEN3	895 ± 70	1451 ± 98	**178 ± 31**	1730 ± 111	2886 ± 161
DEN4	1708 ± 109	1690 ± 107	1730 ± 111	**119 ± 25**	2811 ± 158
JE	2772 ± 157	2935 ± 163	2886 ± 161	2811 ± 158	**130 ± 26**
YF	3284 ± 181	3237 ± 179	3294 ± 180	3573 ± 195	3123 ± 172
CFA	7890 ± 564	8081 ± 585	7787 ± 547	8153 ± 595	8172 ± 597

Dates for the recent cladogenesis in the dengue and JE groups are given in boldface type on the diagonal.

In spite of the biases in these estimations, the values obtained are in good agreement with those previously determined and support the notion that tick-borne viruses might be evolving at a lower rate than the mosquito-borne viruses due to differences in vector biology (3, 5, 7). Rates of synonymous substitution are about an order of magnitude higher (data not shown) and are in good agreement with the mutation rates estimated for other RNA viruses (24). However, synonymous sites are so saturated in this data set that estimations of their rate of substitution are unreliable.

Dating Flavivirus Evolution. By using the rates of substitution estimated for the tick- and mosquito-borne viruses separately (4.12×10^{-5} and 7.5×10^{-5} nonsynonymous substitutions per site per year, respectively), the dates of divergence of other parts of the tree can be estimated. Values for selected divergence events are shown in Tables 2 and 3. For the tick-borne viruses (Table 3) these data suggest a continual divergence over the last 2000 years, although the two bird-associated viruses, SRE and TYU, diverged somewhat earlier (≈4600 years ago). According to these estimates, LI virus arose ≈300 years ago. In contrast, although the major groups of mosquito-borne viruses arose thousands of years ago—for example, the separation of the dengue and JE groups from YF dates to over 3000 years ago—most of the lineages have diverged recently. In the dengue and JE serotypes the period of intense cladogenesis occurred during the last two centuries. These estimates are in accord with the earliest reports of epidemics caused by these diseases. The first recognized epidemic of DF was in 1779 during the Philadelphia epidemics (25), whereas a JE epidemic was described as early as 1871 (1).

DISCUSSION

Phylogenetic Relationships of the Flaviviruses. The phylogenetic analysis of the flaviviruses presented here complements and extends those undertaken earlier on smaller data sets (5, 7–10, 26). The phylogeny for the family, reconstructed by using first and second codon positions, is strongly supported by bootstrap resampling analysis. Moreover, rooting with the CFA sequence did not affect the relationships obtained or the estimates of branch lengths to a great extent.

Table 3. Times of divergence (in years before present) for selected nodes of tick-borne viruses dated by using rates of nonsynonymous substitution

	LI/31		LI/31
LI/I	168 ± 42	OHF	728 ± 91
LI/MA54	273 ± 55	LGT	1192 ± 119
SSE	268 ± 54	KFD	1695 ± 146
WTBE	431 ± 67	POW	1995 ± 159
TSE	449 ± 70	TYU	4466 ± 269
FETBE	595 ± 82	SRE	4910 ± 289

SSE, Spanish sheep encephalomyelitis; WTBE, Western European TBE; TSE, Turkish sheep encephalomyelitis; FETBE, Far Eastern TBE; OHF, Omsk hemorrhagic fever; LGT, Langat; KFD, Kyasanur Forest disease; POW, Powassan.

Cladogenetic Patterns in the Flaviviruses. The analysis of the branching structure of the flavivirus phylogenetic tree indicates that there is a major difference between the tick-borne and mosquito-borne viruses. Specifically, the dengue and JE serotypes show a two-phase pattern in which a slow growth phase is followed by one of rapid growth. In contrast, the tick-borne viruses are characterized by a continual and highly asymmetric branching process. The two-phase pattern of mosquito-borne viruses is caused by the lack of observable lineage splitting from the origin of the main serotypes to the point where they increase dramatically in rates of cladogenesis. Phylogenetic gaps are also apparent in the estimation of times of divergence; the tick-borne viruses have produced observable lineages throughout the last 2000 years, whereas in the mosquito-borne group few lineages which diverged between 200 and 3000 years ago have survived to the present. This interpretation is supported by serological studies which indicate that the tick-borne encephalitis viruses constitute a single antigenic complex but that there is lack of cross-neutralization among dengue serotypes, indicative of little sharing of epitopes and an absence of intermediate lineages (6). Furthermore, possible members of putative DEN5 and DEN6 were shown by sequence analysis to be DEN2 and DEN1, respectively (27).

Distinct Population Processes Among Flaviviruses. Tick-borne viruses exist in a forest cycle where they infect a wide range of vertebrates through vectors whose generation time can be measured in months and even years (28). Consequently, it is possible to view them as zoonoses in relatively stable transmission relationships with a set of hosts, a stability that is facilitated by the fact that tick-to-tick transmission by cofeeding on nonviremic vertebrates is also possible (29). Transmission networks are therefore sustained through long periods of time and the survival of ancient lineages becomes possible. The highly asymmetric branching pattern of the tree is compatible with a model of the gradual dispersal of these viruses across the Northern Hemisphere (7).

The simplest explanation for the branching pattern observed in the mosquito-borne viruses, and particularly dengue, is that they were low-prevalence endemic viruses until recently when the availability of new and susceptible hosts, coupled with increasing worldwide movement and mixing of vectors, viruses, and hosts, initiated a massive population growth. Evidence for this is that the rapid expansion in dengue population size, reflected in the increase in number of viral lineages, and the acceleration in the growth rate of the human host population (30) both occurred ≈200 years ago (Fig. 3). Mosquito lineages produced before this time did not survive, we suggest, because they would generally encounter a small population, cause a localized epidemic, and then die out because of a lack of susceptible hosts and vectors. Such nonsustainable dengue epidemics are seen in nature (31). It was not until the last 200 years or so that viral lineages would encounter a large enough host population to be able to maintain a sustainable transmission network. The movement of the viruses into new suscep-

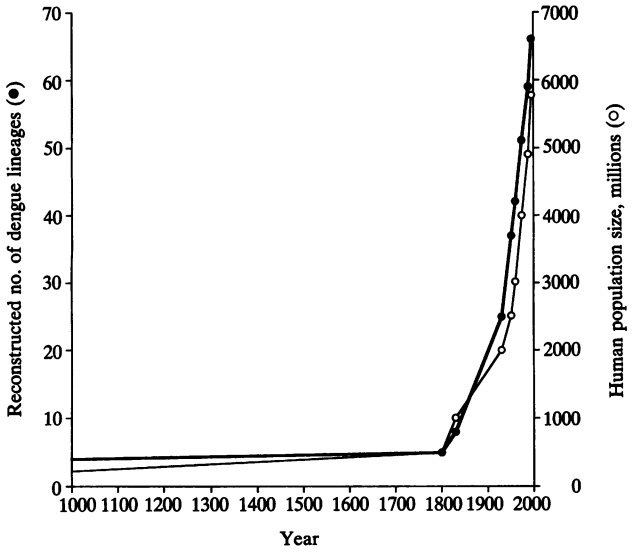

FIG. 3. Plot of the number of reconstructed lineages of dengue virus against the time at which they appear (data presented here) and the corresponding growth of the world's human population (data from ref. 30).

tible populations would also prune lineages through founder effects. The expansion of human populations would have less effect on the transmission of tick-borne flaviviruses, where humans are effectively dead-end hosts.

Two factors complicate the inferences made here: the small and biased nature of the sample of viral lineages and the influence of immune selection. There are a number of reasons for thinking that unrepresentative sampling does not determine the branching patterns observed. In growing populations such as the tick- and mosquito-borne flaviviruses, the effect of a small sample of sequences will be to disproportionately underestimate the number of *recent* nodes, whereas these two virus groups differ in the number of deep nodes. The difference in the number of deep nodes between the tick- and mosquito-borne flaviviruses is also observed in a phylogeny of 16 complete polyprotein sequences, indicating that it is not simply an artefact of looking at the E gene (data not shown). We have also reanalyzed the dengue virus data by constructing trees and LTTP on smaller, randomly chosen, samples of sequences and in samples which were controlled for time and place of origin. In all cases the two-phase LTTP was observed. Finally, the explosion of dengue viruses was also found in an independent analysis of the E and NS1 junction sequences from DEN1 and DEN2 (data not shown).

In theory, immune selection could lead to the branching pattern seen in the mosquito-borne viruses if an escape mutant emerges that out-competes other viruses and then generates a great deal of branching activity as its descendants propagate (a "selective sweep"). However, this scenario is unlikely because if immunological selection really is the driving force of genetic diversity, we would expect to see higher rates of nonsynonymous substitution than of synonymous substitution. When codons that have fixed amino acid replacements are compared, DEN1 shows average values for $d_n = 0.1550 \pm 0.0217$ and $d_s = 0.3443 \pm 0.0548$, indicating by a t test (18) that nonsynonymous rates in the E gene are significantly lower than synonymous rates ($P < 0.001$) even among closely related sequences. Therefore, most of the selection in the DEN1 E gene is purifying, probably caused by constraints imposed by protein structure and function. Furthermore, it is striking that although mosquito-borne viruses show different patterns of variation along their E genes (3) and propagate in different vectors (i.e., dengue in *Aedes* spp. and JE in *Culex* spp.), exactly the *same* branching patterns are seen. Although we accept that selection must be an important factor in shaping viral genetic diversity, these effects will probably be localized to particular regions of the genome [e.g., naturally arising antibody-escape mutant epitopes in the E gene of LI virus (32)] or possibly to epitopes on different gene products (e.g., C, prM, E, NS2a, and NS3 of dengue) challenged by the immune system (33).

Implications of the Dengue Virus Radiation. Most of the evidence on disease in humans prior to the turn of the century comes from anecdotal descriptions (34). In the case of dengue-like diseases, the etiology is common to several distinct agents and proper identification as a disease entity had to wait until the beginning of this century and retrospective serology (25). However, disease descriptions characteristic of dengue date back to 1779, and the virus may have already become a cosmopolitan agent, causing almost simultaneous epidemics in locations as far apart as Philadelphia, Jakarta, Bavaria, and possibly Cairo (25). Since that time dengue virus population dispersal has probably been enhanced by the expansion of breeding places for mosquitoes, the mass movement of non-immune human populations into areas infested with mosquitoes (35), and the increased urbanization and destitution which followed. More recently, efficient and rapid mass transport has favored the introduction and cocirculation of serotypes around the globe, followed by an increase in cases of DF, DHF, and DSS (2). Indeed, the first reported outbreaks of DHF occurred in Southeast Asia after World War II and DHF is now endemic to this region, being one of the main causes of hospitalization and death among children (33). As well as in Southeast Asia, the mixing of serotypes has been associated with a huge rise in the cases of DHF in the Americas starting in the mid-1970s (35).

Although historical and epidemiological data are compelling in showing the increase in numbers of dengue infections, our analysis presents fresh evidence on viral population dynamics and the increase in genetic diversity. In our view, the most serious implication of this work is that human populations are being exposed to an expanding range of increasingly diverse viral strains. Furthermore, our findings have implications for DHF and DSS, the most important human diseases caused by flaviviruses. There are at least three possible explanations for the occurrence of DHF and DSS: (*i*) increased viral load caused by enhanced infection of monocytes and macrophages in the presence of preexisting dengue antibody at subneutralizing levels (35, 36), (*ii*) hyperviremic variant strains (8), and (*iii*) exacerbated immunopathological response (33). Although these explanations are not exclusive and might be synergistic, reinfection is important, since secondary infection is responsible for a 157-fold increase of DSS and an 11-fold increase in DHF compared with primary infections in children between 1 and 14 years of age in Bangkok.[§] The "iceberg" (i.e., incidence) model is a more general view stating that a given percentage of dengue cases, at the tip of the iceberg, will suffer DHF or DSS (35). Given that cocirculation of heterologous serotypes of dengue might become more common, the rapid increase in viral diversity described here may lead to the emergence of viral lineages with increased transmission potential and pathogenicity.

Whatever the future prognosis of dengue fever in humans, it is evident that extensive surveys of flavivirus diversity and geographical dispersal patterns are essential for properly describing and understanding arbovirus epidemics and thereby providing the possibility of assisting in their control.

[§]Kurane, I., Gagnon, S. J., Green, S., Livingstone, P. G., Brinton, M. A., Zenf, L. L. & Lai, C. J., Fourth International Symposium on Positive Strand RNA Viruses, May 25–30, 1995, Utrecht, The Netherlands.

We thank Geoff Garnett, Sean Nee, Andy Purvis, and Andrew Rambaut from the Department of Zoology, University of Oxford;

Patricia A. Nuttall from the National Environment Research Council Institute of Virology and Environmental Microbiology, Oxford; Edwin Westaway from the Sir Albert Sakzewski Virus Research Center, Brisbane, Australia; and three anonymous referees. This work was supported by a scholarship from Conselho Nacional de Pesquisas of Brazil (P.M.d.A.Z.) and grants from the Biotechnology and Biological Sciences Research Council (P.H.H.), the Royal Society (E.C.H.), and the Wellcome Trust (E.C.H. and P.H.H.).

1. Monath, T. P. (1990) in *Virology*, eds. Fields, B. N., Knipe, D. M., Chanock, R. M., Melnick, J. L., Hirsch, M. S. & Monath, T. P. (Raven, New York), 2nd Ed., pp. 763–815.
2. Monath, T. P. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2395–2400.
3. Gritsun, T. S., Holmes, E. C. & Gould, E. A. (1995) *Virus Res.* **35,** 307–321.
4. Reid, H. W. (1988) in *The Arboviruses: Epidemiology and Ecology*, ed. Monath, T. P. (CRC, Boca Raton, FL), Vol. 3, pp. 117–135.
5. Marin, M. S., Zanotto, P. M. de A., Gritsun, T. S. & Gould, E. A. (1995) *Virology* **206,** 1133–1139.
6. Calisher, C. H., Karabatsos, N., Dalrymple, J. M., Shope, R. E., Porterfield, J. S., Westaway, E. G. & Brandt, W. E. (1989) *J. Gen. Virol.* **70,** 37–43.
7. Zanotto, P. M. de A., Gao, G. F., Gritsun, T., Marin, M. S., Jiang, W. R., Venugopal, K., Reid, H. W. & Gould, E. A. (1995) *Virology* **210,** 152–159.
8. Lanciotti, R. S., Lewis, J. G., Gubler, D. J. & Trent, D. W. (1994) *J. Virol.* **75,** 65–75.
9. Lewis, J. A., Chang, G. J., Lanciotti, R. S., Kinney, R. M., Mayer, L. W. & Trent, D. W. (1993) *Virology* **197,** 216–224.
10. Rico-Hesse, R. (1990) *Virology* **174,** 479–493.
11. Gubler, D. J. & Clark, G. G. (1995) *Emerging Infect. Dis.* **1,** 55–57.
12. Camisa-Parks, H., Cisar, L. A., Kane, A. & Stollar, V. (1992) *Virology* **189,** 511–524.
13. Corpet, F. (1988) *Nucleic Acids Res.* **16,** 10881–10890.
14. Olsen, G. J., Matsuda, H., Hagstrom, R. & Overbeek, R. (1994) *Comput. Appl. Biosci.* **10,** 41–48.
15. Swofford, D. L. (1993) Phylogenetic Analysis Using Parsimony (PAUP) (Univ. of Illinois, Champaign), Version 3.1.1.
16. Felsenstein, J. (1993) PHYLIP: Phylogeny Inference Package (Univ. of Washington, Seattle), Version 3.5c.
17. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
18. Kumar, S., Tamura, K. & Nei, M. (1993) MEGA: Molecular Evolutionary Genetics Analysis (Pennsylvania State Univ., University Park, PA), Version 1.01.
19. Harvey, P. H., Rambaut, A. & Nee, S. (1995) in *Genesis and Maintenance of Biodiversity*, eds. Hochberg, M. E., Clobert, J. & Barbault, R. (Oxford Univ. Press, Oxford), pp. 60–68.
20. Nee, S., Holmes, E. C., Rambaut, A. & Harvey, P. H. (1995) *Philos. Trans. R. Soc. London B* **349,** 25–31.
21. Taylor, R. W. & Jones, D. T. (1993) *J. Theoret. Biol.* **164,** 65–83.
22. Nee, S., May, R. M. & Harvey, P. H. (1994) *Philos. Trans. R. Soc. London B* **344,** 305–311.
23. Harvey, P. H., May, R. M. & Nee, S. (1994) *Evolution* **48,** 523–529.
24. Domingo, E. & Holland, J. J. (1994) in *The Evolutionary Biology of Viruses*, ed. Morse, S. S. (Raven, New York), pp. 161–184.
25. Theiler, M. & Downs, W. G. (1973) *An Account of the Rockefeller Foundation Virus Program (1951-1970)* (Yale Univ. Press, London).
26. Blok, J., McWilliam, S. M., Butler, H. C., Gibbs, A. J., Weiller, G., Herring, B. L., Hemsley, A. C., Aaskov, J. G., Yoksan, S. & Bhamarapravati, N. (1992) *Virology* **187,** 573–590.
27. Shiu, S. Y. W., Jiang, W. R., Porterfield, J. S. & Gould, E. A. (1992) *J. Gen. Virol.* **73,** 207–212.
28. Sonenshine, D. E. (1988) in *The Arboviruses: Epidemiology and Ecology*, ed. Monath, T. P. (CRC, Boca Raton, FL), Vol. 1, pp. 219–241.
29. Labuda, M., Nuttall, P. A., Kozuch, O., Eleckova, E., Williams, T., Zuffova, E. & Sabo, A. (1993) *Experientia* **49,** 802–805.
30. Johnson, S. P. (1994) *World Population—Turning the Tide: Three Decades of Progress* (Grahan & Trotman, London).
31. Groot, H., Morales, A., Romero, M., Vidales, H., Lesmes-Donaldson, C., Márquez, G., Escobar de Calvache, D. & Sáenz, M. O. (1979) in *Dengue in the Caribbean, 1977* (Pan Am. Health Organ, Washington, DC), Sci. Publ. No. 375, pp. 31–39.
32. Gao, G. F., Hussain, M. H., Reid, H. W. & Gould, E. A. (1994) *J. Gen. Virol.* **75,** 609–614.
33. Kurane, I. & Ennis, F. (1994) *Semin. Virol.* **5,** 443–448.
34. McKeown, T. (1988) *The Origins of Human Disease* (Blackwell, Oxford).
35. Gubler, D. J. (1988) in *The Arboviruses: Epidemiology and Ecology*, ed. Monath, T. P. (CRC, Boca Raton, FL), Vol. 2, pp. 233–260.
36. Halstead, S. B. (1981) in *Dengue Hemoragic Fever 1981*, Proceedings of the First International Center for Medical Research Seminar (Kobe Univ. School of Medicine, Kobe, Japan), pp. 115–124.