

## Genomic Epidemiology of a Dengue Virus Epidemic in Urban Singapore<sup>▽†</sup>

Mark J. Schreiber,<sup>1\*</sup> Edward C. Holmes,<sup>2,3</sup> Swee Hoe Ong,<sup>1,4,5</sup> Harold S. H. Soh,<sup>6</sup> Wei Liu,<sup>1</sup> Lukas Tanner,<sup>1</sup> Pauline P. K. Aw,<sup>4</sup> Hwee Cheng Tan,<sup>7</sup> Lee Ching Ng,<sup>7</sup> Yee Sin Leo,<sup>8</sup> Jenny G. H. Low,<sup>8</sup> Adrian Ong,<sup>8</sup> Eng Eong Ooi,<sup>9</sup> Subhash G. Vasudevan,<sup>1‡</sup> and Martin L. Hibberd<sup>4</sup>

*Novartis Institute for Tropical Diseases, 10 Biopolis Road, Chromos 05-01, Singapore 138670<sup>1</sup>; Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, University Park, Pennsylvania 16802<sup>2</sup>; Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892<sup>3</sup>; Genome Institute of Singapore, A\*STAR, 60 Biopolis Street, Genome 02-01, Singapore 138672<sup>4</sup>; Swiss Tropical Institute, Socinstrasse 57, P.O. Box CH-4002, Basel, Switzerland<sup>5</sup>; Institute for High Performance Computing, A\*STAR, 1 Science Park Road, 01-01 The Capricorn, Singapore Science Park II, Singapore 117528<sup>6</sup>; National Environment Agency, 40 Scotts Road, 13-00, Singapore 228231<sup>7</sup>; Tan Tock Seng Hospital, 11 Jalan Tan Tock Seng, Singapore 308433<sup>8</sup>; and DSO National Laboratories, 27 Medical Drive, 09-01, Singapore 117510<sup>9</sup>*

Received 26 November 2008/Accepted 3 February 2009

Dengue is one of the most important emerging diseases of humans, with no preventative vaccines or antiviral cures available at present. Although one-third of the world's population live at risk of infection, little is known about the pattern and dynamics of dengue virus (DENV) within outbreak situations. By exploiting genomic data from an intensively studied major outbreak, we are able to describe the molecular epidemiology of DENV at a uniquely fine-scaled temporal and spatial resolution. Two DENV serotypes (DENV-1 and DENV-3), and multiple component genotypes, spread concurrently and with similar epidemiological and evolutionary profiles during the initial outbreak phase of a major dengue epidemic that took place in Singapore during 2005. Although DENV-1 and DENV-3 differed in viremia and clinical outcome, there was no evidence for adaptive evolution before, during, or after the outbreak, indicating that ecological or immunological rather than virological factors were the key determinants of epidemic dynamics.

The phylogenetic analysis of gene sequence data is commonly used to determine the origins of disease outbreaks, particularly those caused by rapidly evolving RNA viruses (1, 2, 7, 9, 16, 26). Historically, such molecular epidemiological studies have usually utilized a small number of genes and have largely concentrated on determining the origins of disease outbreaks and retracing their pathways of spread. As such, these studies have rarely been able to shed light on the precise spatial and temporal dynamics of viral transmission. Recently, whole-genome sequencing of viruses has been utilized to provide greater phylogenetic resolution on outbreak dynamics (12, 14, 15, 20) and is likely to become the benchmark in the near future.

One disease where genomic sequence may be especially informative is dengue fever, an acute febrile disease caused by a mosquito-borne RNA virus (dengue virus [DENV]; single-strand positive-sense, family *Flaviviridae*) and the most common vector-borne viral infection of humans; some 100 million dengue cases are reported on an annual basis, with epidemics especially common in Southeast Asia (10, 11). The potential

expansion of the viable geographic range for *Aedes aegypti* mosquitoes following global warming coupled with the current lack of an effective vaccine or antiviral therapies make understanding the epidemic dynamics of this important emerging virus a key priority.

Dengue has the ability to cause major outbreaks in urban settings, often with high levels of morbidity. The Singapore dengue outbreak of 2005 was the largest of its kind in this locality, with a case rate of 335 per 100,000 population (17). The 2005 outbreak was also notable in that it was characterized by the cocirculation of two of the four viral serotypes—DENV-1 and DENV-3—combined with a low level of DENV-2 transmission. The resurgence of dengue in Singapore is particularly striking given that an aggressive vector control program started in 1970 has resulted in a very low household index of mosquito breeding sites (21). Lessons learned in Singapore, with its long history of commitment to dengue control, may therefore be vital to the overall global effort in dengue prevention.

By undertaking a detailed analysis of whole-genome data sampled from a major outbreak of DENV in Singapore during 2005—the first of its kind—we demonstrate how a synthesis of comparative genomics and fine-scale spatial and temporal epidemiological data provides unprecedented power to reveal the origins, causes, and dynamics of this important emerging virus in a densely populated urban environment.

### MATERIALS AND METHODS

\* Corresponding author. Mailing address: Novartis Institute for Tropical Diseases, 10 Biopolis Road, Chromos 05-01, Singapore 138670. Phone: (65) 672-2973. Fax: (65) 672-2910. E-mail: mark.schreiber@novartis.com.

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

‡ Present address: Emerging Infectious Diseases Program, Duke-NUS Graduate Medical School, 2 Jalan Bukit Merah, Singapore 169547.

▽ Published ahead of print on 11 February 2009.

consenting patients presenting with fever of  $\geq 38^\circ\text{C}$  for less than 72 h. Portions (1 ml) of serum from samples confirmed as dengue positive by reverse transcription-PCR (RT-PCR) were inoculated onto the *Aedes albopictus* mosquito (C6/36) cell line (ATCC CRL-1660). Cells were incubated at  $37^\circ\text{C}$  for up to 10 days or until 75% of the cell monolayer showed cytopathic effects. Isolation of the virus was confirmed and serotyped by indirect immunofluorescence using DENV group-specific and DENV serotype-specific monoclonal antibodies.

**Molecular analysis.** Viruses were propagated by two passages in C6/36 mosquito cell culture. Virus titer was measured by using a plaque assay. Virus titers of at least  $10^6$  were found to be required for optimal success in subsequent RT-PCR steps. Viral RNA was extracted from the culture supernatant by using a Qiagen QiaAmp kit and extraction protocol.

The extracted RNA was copied to cDNA using an RT reaction, followed by PCR amplification. The virus was amplified as five overlapping fragments. Unless specified below the same conditions were used for all five fragments. RT of fragments 1, 3, and 4 was performed in a single tube. RT of fragments 2 and 5 was performed in separate tubes. Samples were kept on ice and pipetting was carried out using RNase free filter tips. PCR primers used in the RT reaction are detailed in Table S1 in the supplemental material and reaction conditions are presented in Table S2 in the supplemental material. The five RT fragments were subsequently amplified by PCR. The PCR primers that were used to amplify the RT fragments are detailed in Table S3 in the supplemental material. PCR of the RT fragments was carried out in a thermal cycler using the program shown in Table S4 in the supplemental material.

Gel electrophoresis was used for visualization of the PCR products as well as gel extraction and purification of products. The products were separated in 1% agarose Tris-borate-EDTA gels after visualization of ethidium bromide-stained bands under UV light. DNA was extracted from bands excised from agarose gels by using a Qiagen QiaQuick extraction kit.

Finally, gel-purified fragments were sequenced by using an Applied Biosystems BigDye ddNTP capillary sequencer. Viral genome sequences generated in the present study are deposited in GenBank with accession numbers EU081177 to EU081281. All genome sequences, their GenBank accession numbers, and their standard strain names (27) used in the analyses in the present study are detailed in Table S5 in the supplemental material.

**Evolutionary analysis.** To reveal the origins of the Singapore viruses we conducted a phylogenetic analysis of the complete coding region of the genome sequences of all those viruses sequenced here, as well as those already available in GenBank. This resulted in data sets of the following sizes: DENV-1, 145 taxa, 10,176 nucleotides (nt); DENV-2, 116 taxa, 10,173 nt; and DENV-3, 122 taxa, 10,173 nt. To determine the best-fit model of nucleotide substitution we used MODELTEST (24). In all cases, the most general GTR+I+ $\Gamma_4$  model, where GTR is generalized time reversible, I is proportion of invariable sites, and  $\Gamma$  is the shape parameter of the gamma distribution, was favored. Maximum-likelihood (ML) trees were then inferred under this model using PAUP\* (30), with tree bisection reconnection branch-swapping in each case. Finally, a neighbor-joining bootstrap analysis (1,000 replications), but using the ML substitution model, was used to determine the robustness of key nodes on each phylogeny.

To determine the population dynamics of DENV-1 and DENV-3 during the 2005 dengue outbreak in Singapore, we analyzed isolates that clearly diversified during the course of the epidemic. For DENV-1, this meant our analysis was confined to 53 genome sequences from Singapore (genotype I), while 42 genomes (genotype III) were used in the case of DENV-3. There were insufficient sequences for an analysis of DENV-2. Demographic and evolutionary parameters for both serotypes were estimated by using the Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in the BEAST package (4). Because of the typically complex population dynamics we utilized the Bayesian skyline plot as a coalescent prior. This provides a piecewise graphical depiction of changes in the effective number of infections ( $N_e\tau$ ), where  $N_e$  is the effective population size and  $\tau$  is the generation time. We also utilized both strict and relaxed (uncorrelated lognormal) molecular clocks. The GTR+ $\Gamma_4$  model of nucleotide substitution was used in all cases, with the invariant-sites parameter (I) excluded since it tended to overfit to the data. Uncertainty in the data for each estimated parameter is reflected in values of the 95% high probability density (HPD), with all MCMC chains run for sufficient time (50 million steps, with 10% removed as burn-in) to ensure statistical convergence.

To determine the nature of selection pressures acting on each gene of DENV-1 and DENV-3 sampled from Singapore, we computed the mean ratio of nonsynonymous ( $d_N$ ) to synonymous ( $d_S$ ) substitutions per site ( $d_N/d_S$ ) using the single likelihood ancestor counting method available through the DATAMONKEY web interface (23) and assuming the GTR model of nucleotide substitution and an input neighbor-joining tree. This analysis also allowed us to compute the tree length ( $T_L$ ) in substitutions per site for each gene. In addition, we used the

CODEML program within the PAML package (33) to estimate an overall  $d_N/d_S$  for the entire coding region of both serotypes (the “one-ratio” model). This was compared to the case in which a separate  $d_N/d_S$  value was estimated for the external and internal branches of each data set (the “two-ratio” model).

Finally, to determine the strength of spatial structure in both DENV-1 and DENV-3, we first obtained the physical address of each viral isolate and produced clusters according to their geographical proximity by K-means clustering. For DENV-1, the physical address was available for 48 isolates, which were then placed into one of six different spatial groups (with a single-letter character state code identifying each group). In the case of DENV-3, address information was available for 42 isolates, and these were separated into four spatial groups. Although there are a variety of methods by which the extent of spatial structure in phylogenetic data can be determined, particularly utilizing parsimony character state mapping (28), we used a Bayesian MCMC approach (22), thereby accounting for any error in the underlying phylogeny. This analysis was based on the trees output from the previous BEAST analysis (with a new BEAST analysis conducted on the 48-sequence DENV-1 data set), using 1,000 replications and with 10% of trees removed as burn-in. From these trees we computed the mean values, credible intervals, and significance of the parsimony score (PS) and association index (AI) statistics of the strength of geographical clustering (22).

## RESULTS AND DISCUSSION

Fortunately, the 2005 outbreak coincided with the launch of the longitudinal EDEN (early dengue infection and outcome) study in the central Ang Mo Kio district of Singapore (17). Of 133 RT-PCR dengue-positive patients enrolled during the EDEN study collected between April and November 2005, serotyping determined 66 (48.9%) to be DENV-1, 62 (46.6%) to be DENV-3, and 5 to be DENV-2 (3.8%). No cases due to DENV-4 were observed (17), and one patient was found to be coinfecte with serotypes 1 and 3. The detection of large numbers of DENV-3 in the Ang Mo Kio area was unusual as DENV-1 was the predominant serotype in the rest of Singapore. We were able to isolate 112 (84.2%) viruses, corresponding to 57 DENV-1, 50 DENV-3, and 5 DENV-2 isolates. Complete genome sequences were obtained for 54 DENV-1, 44 DENV-3, and 4 DENV-2 viruses. The remaining samples, although shown to be dengue positive by RT-PCR, did not yield sufficient viral RNA for genome sequencing.

To determine the origins of the viral isolates responsible for the 2005 dengue epidemic in Singapore, we conducted a phylogenetic analysis of the complete genomes of the viruses sampled here combined with representative DENV isolates taken from GenBank. To assist in this comparison, we also completed whole-genome sequences of some historical DENV samples from Singapore. All but one of the DENV-1 genomes from this epidemic were classified as genotype I, which commonly circulates in Southeast Asia (Fig. 1). The single outlier belongs to genotype III, which is predominantly found in Latin America and West Africa (8), although a genotype III virus was previously sampled in Singapore in 1993. The closest relatives of the 53 genotype I Singapore DENV-1 isolates are the Chinese isolates DENV-1/CN/Fj231/2004 and DENV-1/CN/ZH1067/XXX, suggesting that frequent transfer of DENV may occur between Singapore and China, and DENV-1/JP/20-Feb/XXX from Japan. Importantly, since three historical Singapore DENV-1 samples isolated in 2003 fell at the base of the 2005 cluster, it is possible that this particular lineage of DENV-1 genotype I has been circulating continuously in Singapore since at least 2003.

The four DENV-2 genomes form part of the “cosmopolitan” genotype (Fig. 2), which has a wide distribution in tropical and

# DENV-1



FIG. 1. Phylogenetic relationships of 145 complete genomes DENV-1 sampled globally determined by using a ML method. Isolates sampled from Singapore are shown in red, and individual genotypes are shown. Bootstrap values (>80%) are shown next to key nodes, and all horizontal branch lengths are drawn to scale.

## DENV-2

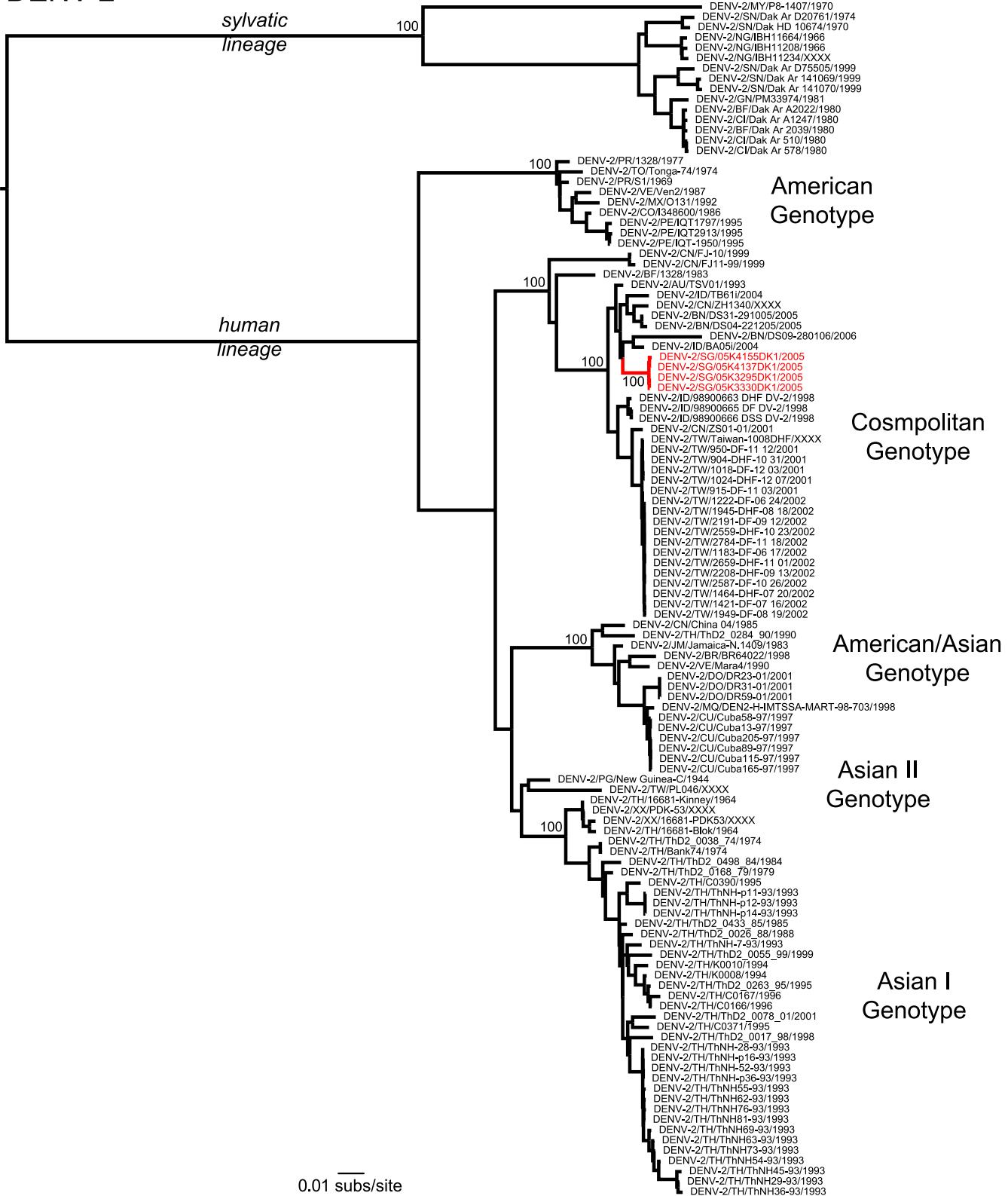


FIG. 2. Phylogenetic relationships of 116 complete genomes DENV-2 sampled globally determined by using a ML method. Isolates sampled from Singapore are shown in red, and individual genotypes are shown. Bootstrap values ( $>80\%$ ) are shown next to key nodes, and all horizontal branch lengths are drawn to scale.

## DENV-3

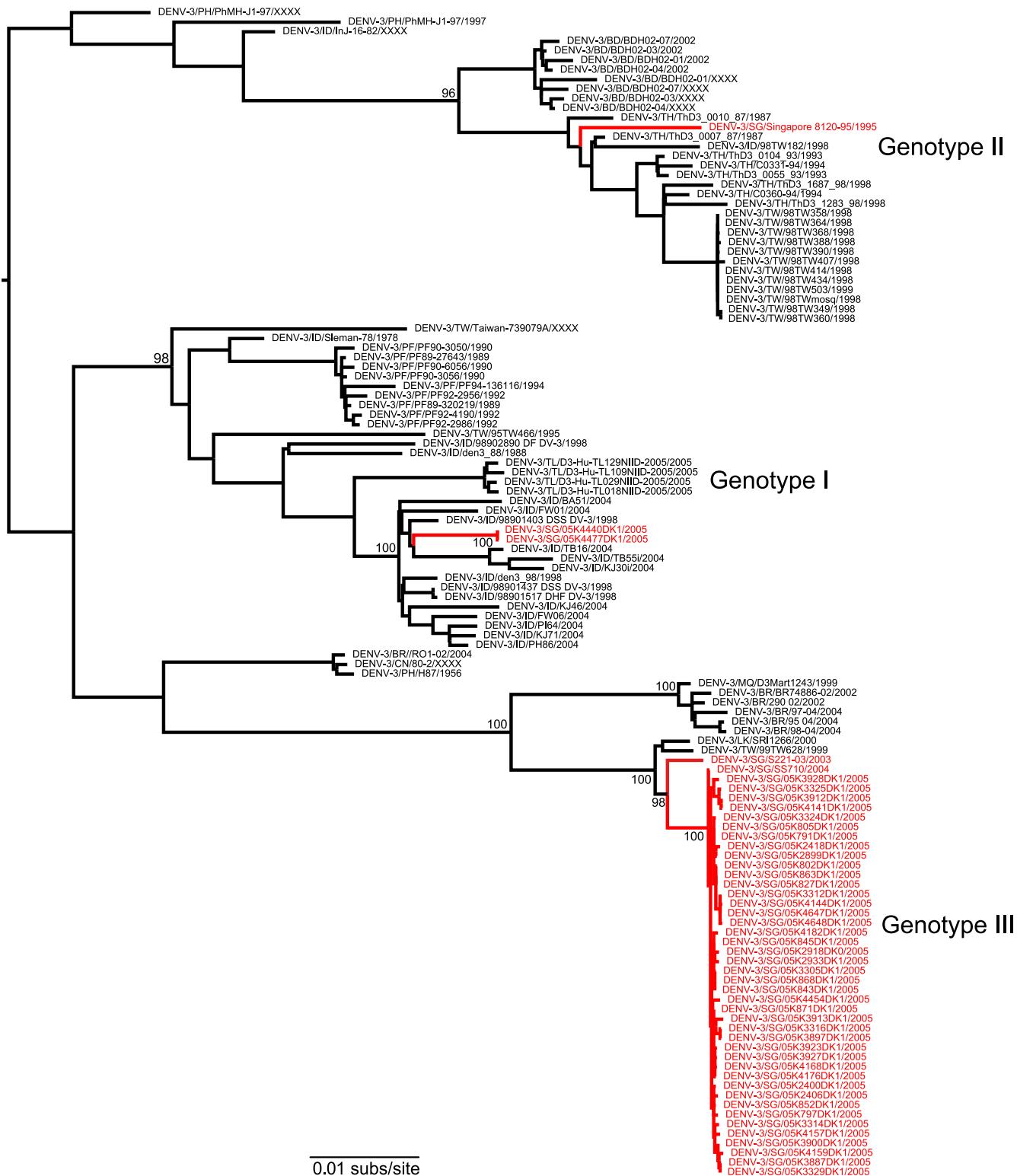


FIG. 3. Phylogenetic relationships of 122 complete genomes DENV-3 sampled globally determined by using a ML method. Isolates sampled from Singapore are shown in red, and individual genotypes are shown. Bootstrap values (>80%) are shown next to key nodes, and all horizontal branch lengths are drawn to scale.

**A**

A Q L S R E K T N G M W D H F Y C I P V - ?

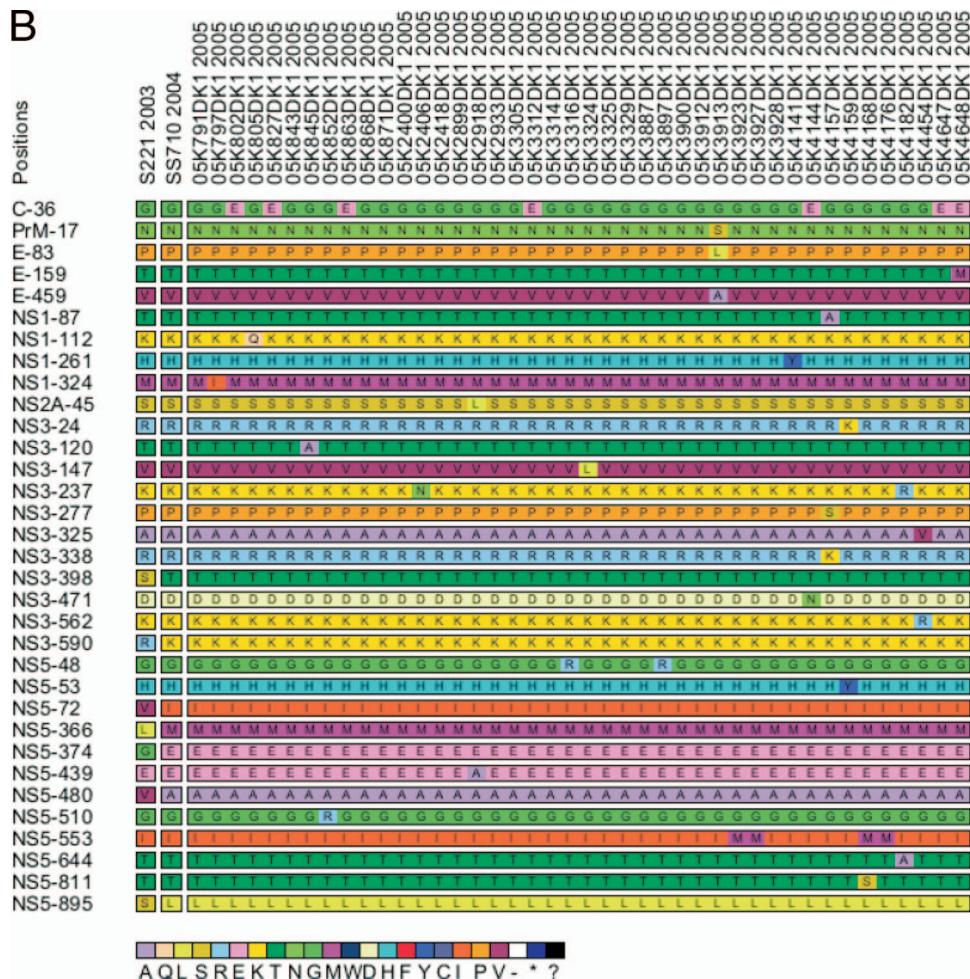


FIG. 4. Observed amino acid changes in DENV-1 (A) and DENV-3 (B). Viral isolates are plotted on the x axis. The aligned polyproteins of each virus were compared to count the number and distribution of amino acid changes. Positions that were not completely conserved are shown with the individual protein name and the amino acid position within the protein on the y axis. The color of each square indicates the type of amino acid residue found in isolate x at position y.

subtropical localities (32). Close relatives of these strains include DENV-2/ID/BA05i/2004 and DENV-2/ID/TB16i/2004, which were isolated during a dengue fever epidemic in Jakarta in 2004 (29), as well as three strains from Brunei Darussalam, one from China, and an older isolate from Queensland, Australia (DENV-2/AU/TSV01/1993) (13), possibly introduced from the nearby Indonesian islands. Since a number of these viruses were isolated between 2004 and 2006, it seems likely that this lineage was relatively common in this geographical area at the time of the outbreak.

The majority of DENV-3 genomes fell into genotype III (Fig. 3). This genotype was originally associated with the Indian subcontinent until the mid-1990s, when it was introduced into Latin America and the Caribbean (19). Of more importance from the perspective of this outbreak was that an isolate from genotype III was first detected in Singapore in 2003 (DENV-3/SG/S221/2003) and which fell basal to the 2005 outbreak viruses in our phylogenetic analysis. Such a phylogenetic pattern is compatible with the *in situ* evolution of this lineage in Singapore since at least 2003. Hence, as is also likely the case with DENV-1, the 2005 outbreak of DENV-3 may also be due

to the amplification of a preexisting viral lineage rather than the invasion of an “exotic” DENV strain. To further test this hypothesis, we obtained the additional genome sequence of a Singaporean DENV-3 genotype I isolate (DENV-3/SG/SS710/2004) sampled in 2004. As expected under the hypothesis of *in situ* evolution, this isolate occupies an intermediate position between the 2003 and 2005 strains. Finally, two of the 44 DENV-3 isolates from the 2005 outbreak in Singapore fall into genotype I, which is endemic in the Malay archipelago. This observation provides an additional point of similarity between the DENV-1 and DENV-3 components of the 2005 DENV outbreak in Singapore: that individual epidemic serotypes can be composed of multiple viral genotypes.

To determine the evolutionary processes that enabled the emergence of multiple DENV genotypes in a single outbreak we examined each of the amino acid changes associated with these viruses. Remarkably, no amino acid changes were completely fixed on the branches leading to the DENV-1 or DENV-3 isolates sampled in Singapore during 2005 (Fig. 4A). In DENV-1 residue 76 of NS4A was observed to be K in all samples from 2003 and R in many 2005 samples. However, the

TABLE 1. Phylogenetic and evolutionary patterns among the proteins of DENV-1 and DENV-3 sampled from Singapore during 2005<sup>a</sup>

Protein	Length(s) (bp) <sup>b</sup>	DENV-1			DENV-3		
		IS	T <sub>L</sub> <sup>b</sup>	d <sub>N</sub> /d <sub>S</sub>	IS	T <sub>L</sub>	d <sub>N</sub> /d <sub>S</sub>
Capsid	342	5	0.026	0.359	3	0.023	0.064
Membrane	498	9	0.034	0.332	1	0.008	0.175
Envelope	1,479 and 1,485*	11	0.031	0.206	9	0.013	0.095
NS1	1,056	7	0.025	0.042	4	0.011	0.204
NS2A	654	6	0.029	0.318	3	0.009	0.067
NS2B	390	8	0.043	0.029	1	0.010	0
NS3	1,857	16	0.024	0.059	4	0.014	0.265
NS4A	450	6	0.029	0.037	2	0.009	0
NS4B	744 and 747*	6	0.031	0.019	3	0.009	0
NS5	2,679 and 2,700*	19	0.022	0.157	11	0.011	0.131
Genome (coding)	10,158 (10,170)	96	NA <sup>c</sup>	NA	41	NA	NA

<sup>a</sup> IS, number of parsimony informative sites; T<sub>L</sub>, tree length expressed as the number of substitutions per site.<sup>b</sup> \*, Lengths for DENV-1 and DENV-3, respectively.<sup>c</sup> NA, not applicable. That is, no significant evidence for positive selection was observed in any gene in either DENV-1 or DENV-3.

penetration of the mutation was incomplete and is quite conservative in nature and therefore unlikely to be significant. Similarly, no amino acid variants in DENV-1 observed during 2005 were found in the two genomes isolated and sequenced from the 2006 nonepidemic year, indicating that no mutations that occurred in 2005 became fixed. Some substitutions were observed in DENV-3 between the 2003 and 2004 nonepidemic strains, although all are conservative, except for the change from serine to lysine at residue 895 in NS5. Position 895 is not conserved in the four serotypes and is usually S, P, or E. In the recently solved structure of the DENV-3 RNA-dependent RNA polymerase (34), position 895 is near the C terminus (position 900) and does not appear to be functionally significant. More notably, there are no fixations between the 2004 isolate and the 2005 epidemic isolates. Although a number of nonconservative substitutions were observed within the 2005 isolates of DENV-3, that this serotype was not detected in 2006 indicates that none were capable of perpetuating the clade (Fig. 4B). Finally, there was no evidence for positive selection in any gene of the Singapore viruses, with a relatively low ratio of nonsynonymous to synonymous substitutions per site (d<sub>N</sub>/d<sub>S</sub>) in all genes and no evidence for site-specific positive selection (Table 1; which also gives a variety of other gene-specific measures of genetic diversity).

To infer the epidemiological dynamics of DENV-1 and DENV-3 during the Singapore outbreak, we used a Bayesian coalescent approach (3, 5) incorporating data on the exact day of viral sampling. For both serotypes, we estimated the changing patterns of relative genetic diversity through time as reflected in the effective number of infections (N<sub>e</sub>τ) using a Bayesian skyline plot and assuming a relaxed (uncorrelated lognormal) molecular clock (although very similar results were observed under a strict molecular clock; these results are available from the authors on request). Similar epidemic profiles were observed in both viruses, comprising a rapid growth phase followed by a constant population size, although the mean age of the common ancestor was significantly greater in DENV-1 (1,740 days; 95% HPD = 741 to 3,222 days) compared to DENV-3 (298 days; 95% HPD = 225 to 387 days), indicating that already diverse lineages of DENV-1 were present in Singapore at the outset of the 2005 outbreak (Fig. 5). Similarly, mean estimates of peak N<sub>e</sub>τ were smaller in DENV-3 (1,632

days; 95% HPD = 56 to 4,334) than DENV-1 (7,211 days; 95% HPD = 637 to 23130), although with overlapping HPD values. Interestingly, DENV-3 was not widely reported in Singapore during the outbreak and appeared to be mainly contained to the sampling area, which supports this result. The substitution dynamics of both serotypes were also similar, with mean evolutionary rates of  $1 \times 10^{-3}$  substitutions/site/year (95% HPD,  $0.4 \times 10^{-3}$  to  $1.6 \times 10^{-3}$  substitutions/site/year) and  $1.3 \times 10^{-3}$  substitutions/site/year (95% HPD,  $8.7 \times 10^{-4}$  to  $1.8 \times 10^{-3}$  substitutions/site/year) for DENV-1 and DENV-3, respectively, and equivalent to those estimated previously for DENV (32).

As an additional analysis of evolutionary dynamics, we determined the genetic distance (under the ML substitution model) for each pair of DENV-1 and DENV-3 sequences and compared these values to time intervals of sampling (based on day of fever onset) (Fig. 6). Interestingly, it appears that genetic distances are often higher between isolates separated by shorter periods of time but then decline between pairs sampled over a longer time period (although this analysis does not take into account phylogenetic structure). This may in part be due to the presence of transient deleterious mutations in samples that are only separated by short time periods (such that genetic distances strongly reflect the background mutation rate), which are later purged by purifying selection, so that longer-term genetic distances are more indicative of the population substitution rate (12). This is supported by the observation that d<sub>N</sub>/d<sub>S</sub> is higher on external (0.130 and 0.157) than internal (0.081 and 0.030) branches of the complete coding region phylogenies for both DENV-1 and DENV-3, respectively, as expected if most nonsynonymous polymorphisms are deleterious (25).

Interestingly, our analysis of DENV spatial dynamics in Singapore revealed significant population substructure (i.e., the existence of distinct spatial clusters) in both DENV-1 and DENV-3. This spatial dynamic was especially strong in the case of DENV-3 ( $P < 0.001$  for both the PS and AI statistics) compared to DENV-1 ( $P = 0.008$  and  $0.037$  for the PS and AI statistics, respectively). Hence, although these viruses were sampled from a relatively restricted region within Singapore, the movement of hosts and/or vectors is sufficiently limited that spatial structure is present in the data.

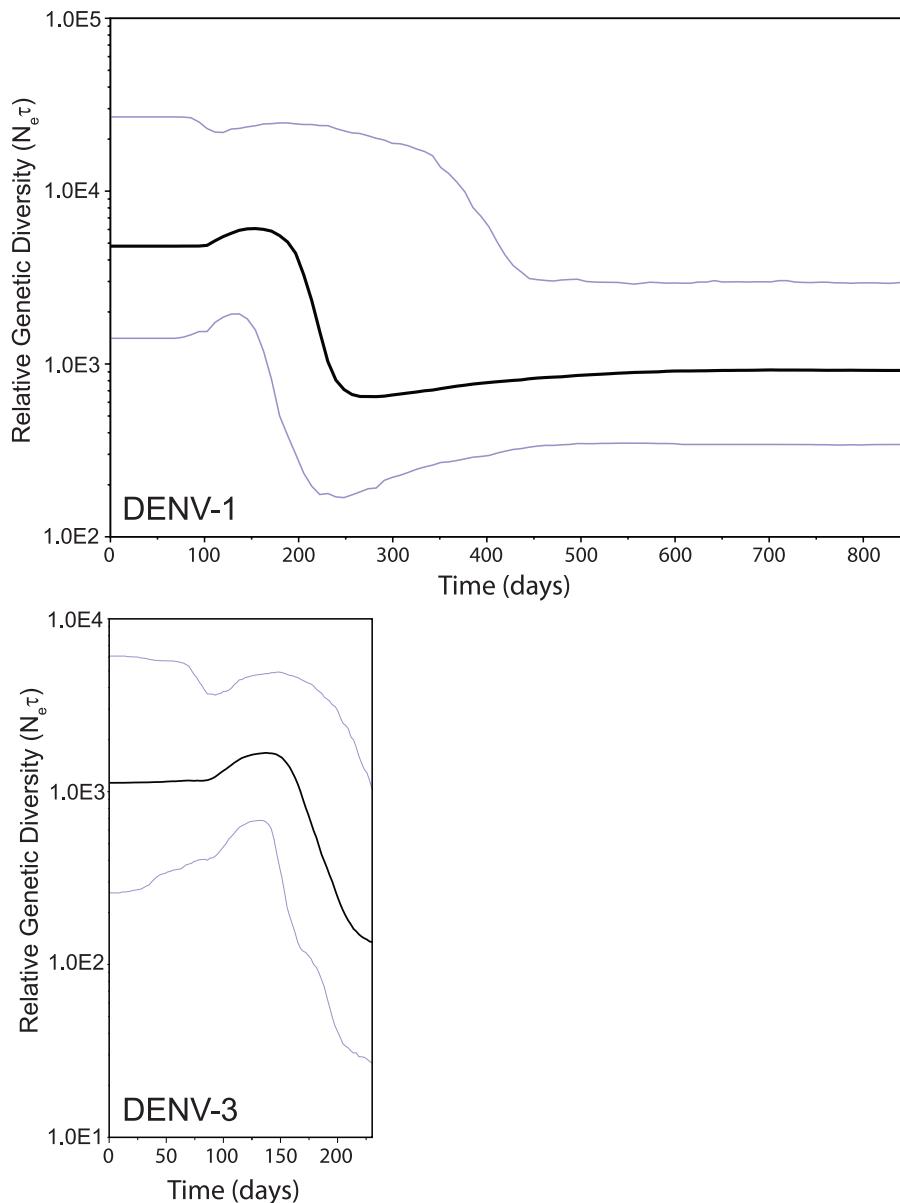


FIG. 5. Population dynamics of DENV-1 and DENV-3 in Singapore during 2005 depicted using Bayesian skyline plots. The plots show changes in relative genetic diversity, depicted as the effective number of infections ( $N_e\tau$ ), through time. The black line represents the mean estimate of  $N_e\tau$ , while the 95% HPD intervals are shown in blue. Time is shown as the number of days from the most recent sample. To aid interpretation, DENV-1 and DENV-3 have been shown on the same time axis.

Finally, the viremia from each patient's sera was estimated from a crossover threshold ( $C_t$ ) calculated using quantitative RT-PCR at 1 to 3 days and 4 to 7 days after fever onset. A low  $C_t$  value of the RT-PCR, indicating high viremia levels, in the first sampling has been previously shown to be predictive of severe thrombocytopenia in our cohort (31). The  $C_t$  value for DENV-1 was significantly lower ( $P = 0.002$ ) than for DENV-3 at both the first (17.07 versus 19.76) and the second (26.87 versus 29.57) serum samplings, indicating a higher viremia level for DENV-1. DENV-1 also resulted in a significantly ( $P = 0.021$ ) higher ratio of hospitalizations among the sampled population compared to DENV-3 (0.74 versus 0.50) and may

also be reflected in the higher values of  $N_e\tau$  for DENV-1 than DENV-3.

Overall, these results suggest that ecological and/or immunological factors, rather than aspects of viral evolution, were central in shaping the dynamics of this dengue outbreak. Most notably, two different serotypes, and multiple cocirculating genotypes, emerged simultaneously, experienced similar epidemiological dynamics, and seemingly spread without the aid of positive selection, accumulating no amino acid fixations. Similarly, it is clear that viral evolution did not succeed in extending the epidemic; despite numerous mutations (Table 1), the number of dengue cases declined rapidly in 2006, suggesting

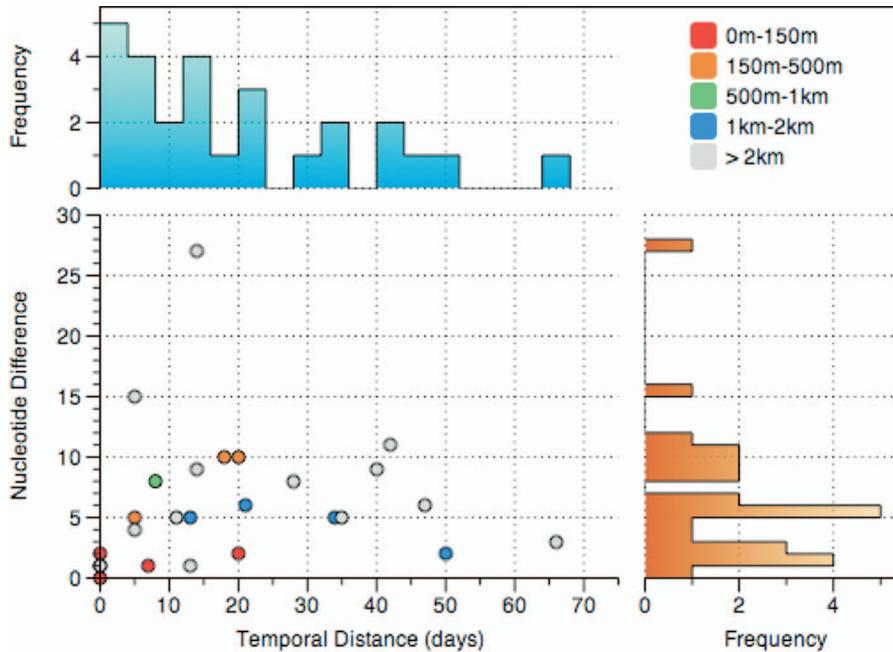


FIG. 6. Distribution of pairwise nucleotide, temporal, and geographic distances among complete coding sequences of DENV-1 and DENV-3 combined. Histograms show the frequency of temporal and nucleotide distances on the x and y axes, respectively. Each pair is represented as a point, and the color of the point indicates the physical distance between the home address of the patients from whom the viruses were isolated.

that few, if any, of these mutations provided any selective advantage in the face of rising immunity. However, both DENV-1 and DENV-3 appear to possess an inherent robustness that allows them to persist at a low level of infection at times when ecological and immunological conditions do not favor an outbreak. This is supported by the observation that the DENV-1 and DENV-3 lineages that characterized this outbreak were found in Singapore as early as 2003 but did not result in an outbreak until 2005. However, a wider sampling of viral isolates from neighboring geographic areas is needed to fully test this analysis. Furthermore, dengue epidemics in Singapore follow a regular 6- to 7-year periodic cycle, which is difficult to explain by patterns of viral evolution or by vector population density and might be more attributable to changing levels of herd immunity (6). In addition, in Colombia it has been observed that epidemic years are correlated with an increase in the infection rate of mosquitoes and not the total number of mosquitoes per household (18), indicating surveillance of mosquito populations may also be important in understanding epidemics.

Together, our results have important implications for the future study and control of DENV epidemics. In particular, the epidemic surveillance of viral genome sequences in this case would not have been sufficient to predict the 2005 outbreak. Hence, incumbent strains with apparently inherent epidemic potential are required but apparently not sufficient to spark an outbreak. Concurrent surveillance of viral isolates, mosquito vectors (including the proportion of mosquitoes infected with DENV), and periodic surveys of seroprevalence rates of the population may therefore provide the additional required predictive information. The chance discovery of the DENV-3 out-

break also highlights the value of comprehensive city-wide fever surveys in detecting rare events.

#### ACKNOWLEDGMENTS

We thank the staff involved in undertaking the EDEN study, together with Lai Yee Ling from the Environmental Health Institute, Fu Xiuju from the Institute for High Performance Computing, and Leng Marcel from the National University of Singapore. We thank Edison Liu and Duane Gubler for useful comments on the manuscript.

M.J.S., W.L., and S.G.V. are employed by the Novartis Institute for Tropical Diseases, which is funded by Novartis AG for the purpose of developing anti-dengue therapies. S.H.O. is a Novartis-funded Ph.D. candidate at the University of Basel. L.T. was enrolled in the Joint M.Sc. Programme in Infectious Diseases organized in conjunction with the National University of Singapore, the Novartis Institute of Tropical Diseases, the Swiss Tropical Institute, and the University of Basel.

This study was in part funded by NIH grant number GM080533-01 to E.C.H., a Singapore BioMedical Research grant to A.O., the Singapore Agency for Science Technology and Research, the Singapore Tissue Network, and the Novartis Institute for Tropical Diseases.

#### REFERENCES

- Bennett, S. N., E. C. Holmes, M. Chirivella, D. M. Rodriguez, M. Beltran, V. Vorndam, D. J. Gubler, and W. O. McMillan. 2006. Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope accompanies clade reintroduction. *J. Gen. Virol.* 87:885–893.
- de Oliveira, T., O. G. Pybus, A. Rambaut, M. Salemi, S. Cassol, M. Ciccozzi, G. Rezza, G. C. Gattinara, R. D'Arrigo, M. Amicosante, L. Perrin, V. Colizzi, and C. F. Perno. 2006. Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* 444:836–837.
- Drummond, A. J., S. Y. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22:1185–1192.
- Egger, J. R., E. E. Ooi, D. W. Kelly, M. E. Woolhouse, C. R. Davies, and P. G. Coleman. 2008. Reconstructing historical changes in the force of infection of

- dengue fever in Singapore: implications for surveillance and control. *Bull. W. H. O.* **86**:187–196.
7. Gilbert, M. T., A. Rambaut, G. Wlasiuk, T. J. Spira, A. E. Pitchenik, and M. Worobey. 2007. The emergence of HIV/AIDS in the Americas and beyond. *Proc. Natl. Acad. Sci. USA* **104**:18566–18570.
  8. Goncalvez, A. P., A. A. Escalante, F. H. Pujol, J. E. Ludert, D. Tovar, R. A. Salas, and F. Liprandi. 2002. Diversity and evolution of the envelope gene of dengue virus type 1. *Virology* **303**:110–119.
  9. Guan, Y., B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L. Cheung, S. W. Luo, P. H. Li, L. J. Zhang, Y. J. Guan, K. M. Butt, K. L. Wong, K. W. Chan, W. Lim, K. F. Shortridge, K. Y. Yuen, J. S. Peiris, and L. L. Poon. 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**:276–278.
  10. Gubler, D. J. 2002. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol.* **10**:100–103.
  11. Halstead, S. B. 2007. Dengue. *Lancet* **370**:1644–1652.
  12. Jarman, R. G., E. C. Holmes, P. Rodpradit, C. Klungthong, R. V. Gibbons, A. Nisalak, A. L. Rothman, D. H. Libraty, F. A. Ennis, M. P. Mammen, Jr., and T. P. Endy. 2008. Microevolution of Dengue viruses circulating among primary school children in Kamphaeng Phet, Thailand. *J. Virol.* **82**:5494–5500.
  13. Johansson, M., A. J. Brooks, D. A. Jans, and S. G. Vasudevan. 2001. A small region of the dengue virus-encoded RNA-dependent RNA polymerase, NS5, confers interaction with both the nuclear transport receptor importin-beta and the viral helicase, NS3. *J. Gen. Virol.* **82**:735–745.
  14. King, C. C., D. Y. Chao, L. J. Chien, G. J. Chang, T. H. Lin, Y. C. Wu, and J. H. Huang. 2008. Comparative analysis of full genomic sequences among different genotypes of dengue virus type 3. *J. Virol.* **82**:563.
  15. Klungthong, C., C. Zhang, M. P. Mammen, Jr., S. Ubol, and E. C. Holmes. 2004. The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand. *Virology* **329**:168–179.
  16. Lanciotti, R. S., J. T. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. E. Volpe, M. B. Crabtree, J. H. Scherret, R. A. Hall, J. S. MacKenzie, C. B. Cropp, B. Panigrahy, E. Ostlund, B. Schmitt, M. Malkinson, C. Banet, J. Weissman, N. Komar, H. M. Savage, W. Stone, T. McNamara, and D. J. Gubler. 1999. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* **286**:2333–2337.
  17. Low, J. G., E. E. Ooi, T. Tolvenstam, Y. S. Leo, M. L. Hibberd, L. C. Ng, Y. L. Lai, G. S. Yap, C. S. Li, S. G. Vasudevan, and A. Ong. 2006. Early dengue infection and outcome study (EDEN): study design and preliminary findings. *Ann. Acad. Med. Singapore* **35**:783–789.
  18. Mendez, F., M. Barreto, J. F. Arias, G. Rengifo, J. Munoz, M. E. Burbano, and B. Parra. 2006. Human and mosquito infections by dengue viruses during and after epidemics in a dengue-endemic region of Colombia. *Am. J. Trop. Med. Hyg.* **74**:678–683.
  19. Messer, W. B., D. J. Gubler, E. Harris, K. Sivananthan, and A. M. de Silva. 2003. Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg. Infect. Dis.* **9**:800–809.
  20. Ong, S. H., J. T. Yip, Y. L. Chen, W. Liu, S. Harun, E. Lystianingsih, B. Heriyanto, C. G. Beckett, W. P. Mitchell, M. L. Hibberd, A. Suwandoni, S. G. Vasudevan, and M. J. Schreiber. 2008. Periodic re-emergence of endemic strains with strong epidemic potential—a proposed explanation for the 2004 Indonesian dengue epidemic. *Infect. Genet. Evol.* **8**:191–204.
  21. Ooi, E. E., K. T. Goh, and D. J. Gubler. 2006. Dengue prevention and 35 years of vector control in Singapore. *Emerg. Infect. Dis.* **12**:887–893.
  22. Parker, J., A. Rambaut, and O. G. Pybus. 2008. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. *Infect. Genet. Evol.* **8**:239–246.
  23. Pond, S. L., and S. D. Frost. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**:2531–2533.
  24. Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**:817–818.
  25. Pybus, O. G., A. Rambaut, R. Belschaw, R. P. Freckleton, A. J. Drummond, and E. C. Holmes. 2007. Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. *Mol. Biol. Evol.* **24**:845–852.
  26. Rico-Hesse, R., L. M. Harrison, A. Nisalak, D. W. Vaughn, S. Kalayanarooj, S. Green, A. L. Rothman, and F. A. Ennis. 1998. Molecular evolution of dengue type 2 virus in Thailand. *Am. J. Trop. Med. Hyg.* **58**:96–101.
  27. Schreiber, M. J., S. H. Ong, R. C. Holland, M. L. Hibberd, S. G. Vasudevan, W. P. Mitchell, and E. C. Holmes. 2007. DengueInfo: a web portal to dengue information resources. *Infect. Genet. Evol.* **7**:540–541.
  28. Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* **123**:603–613.
  29. Suwandoni, A., H. Kosasih, Nurhayati, R. Kusriastuti, S. Harun, C. Ma'roef, S. Wuryadi, B. Herianto, D. Yuwono, K. R. Porter, C. G. Beckett, and P. J. Blair. 2006. Four dengue virus serotypes found circulating during an outbreak of dengue fever and dengue haemorrhagic fever in Jakarta, Indonesia, during 2004. *Trans. R. Soc. Trop. Med. Hyg.* **100**:855–862.
  30. Swofford, D. L. 2003. PAUP\*: phylogenetic analysis using parsimony (\* and other methods), version 4. Sinauer Associates, Sunderland, MA.
  31. Tanner, L., M. Schreiber, J. G. Low, A. Ong, T. Tolvenstam, Y. L. Lai, L. C. Ng, Y. S. Leo, L. Thi Puong, S. G. Vasudevan, C. P. Simmons, M. L. Hibberd, and E. E. Ooi. 2008. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. *PLoS Negl. Trop. Dis.* **2**:e196.
  32. Twiddy, S. S., E. C. Holmes, and A. Rambaut. 2003. Inferring the rate and time-scale of dengue virus evolution. *Mol. Biol. Evol.* **20**:122–129.
  33. Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**:1586–1591.
  34. Yap, T. L., T. Xu, Y. L. Chen, H. Malet, M. P. Egloff, B. Canard, S. G. Vasudevan, and J. Lescar. 2007. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J. Virol.* **81**:4753–4765.