Dengue antigenic relationships predict evolutionary dynamics

Sidney Bell[1,2], Leah Katzelnick[3,4] & Trevor Bedford[1]

[1]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, [2]Molecular and Cell Biology Program, University of Washington, Seattle, WA, USA, [3]Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, Berkeley, CA, USA, [4]Department of Biology, University of Florida, Gainesville, FL, USA

October 1, 2018

Abstract

Dengue virus (DENV) exists as four genetically distinct serotypes, each of which is also antigenically distinct: the immune response to primary infection can be either cross-protective or associated with severe disease upon heterotypic secondary infection. Each serotype is historically assumed to be antigenically uniform. Recent analyses suggest that antigenic heterogeneity may exist within each serotype, but its source, extent and impact remain unclear. Here, we construct a phylogeny-based model to directly map antigenic change to underlying genetic divergence. We report moderate antigenic diversity within each serotype, and identify 12 antigenically distinct clades. We also quantify the impact of this antigenic heterogeneity on real-world DENV population dynamics. We find that antigenic fitness mediates fluctuations in DENV clade frequencies, although this appears to be driven by coarser serotype-level antigenic differences. These results provide a more nuanced understanding of dengue antigenic evolution, with important ramifications for vaccine design and epidemic preparedness.

Author Summary

Dengue virus (DENV), the causative agent of dengue hemorrhagic fever, exists as four genetically distinct serotypes, DENV1-4. These serotypes are antigenically distinct: symptomatic reinfection with a homotypic virus is very rare, while reinfection with a heterotypic virus is associated with severe disease. Until recently, it has been assumed that viruses within each serotype are antigenically uniform. However, specific genotypes within each serotype have been anecdotally associated with varying severity of patient outcomes and epidemic magnitude. One hypothesis is that each serotype contains overlooked, meaningful antigenic diversity. While antigenic cartography suggests that heterogeneity may exist within each serotype, its source, extent and impact is unclear. Here, we analyze a previously published neutralization titer dataset to quantify and phylogenetically characterize the extent of DENV intraserotype antigenic diversity. We map antigenic changes to specific branches of the virus phylogeny, and interpolate across the tree to estimate the antigenic distance between pairs of viruses based on their relative positions in the phylogeny. We find that DENV antigenic divergence is tightly coupled to DENV genetic divergence, and is likely a gradual, ongoing process. We report modest but significant antigenic diversity within each serotype of DENV, and identify a minimum of 12 distinct antigenic clades. Within each serotype, these antigenic clades are separated by antigenic distances comparable to those associated with differing vaccine efficacy in the recent CYD-TDV vaccine trial. This suggests that antigenic interactions between specific viral genotypes may contribute to secondary case outcomes and vaccine efficacy. To understand the impact of this antigenic heterogeneity on real-world DENV population dynamics, we also quantify the extent to which population immunity–accumulated through recent circulation of antigenically similar genotypes–determines the success and decline of DENV clades in a hyperendemic population. We find that antigenic fitness is a key determinant of DENV population turnover, although this appears to be driven by coarser serotype-level antigenic differences. By leveraging both molecular data and real-world population dynamics, these results provide a more nuanced understanding of dengue antigenic evolution, with important ramifications for improving vaccine design and epidemic preparedness.

Introduction

Dengue virus (DENV) is a mosquito-borne flavivirus which consists of four genetically distinct clades, canonically thought of as serotypes (Lanciotti et al., 1997). DENV circulates primarily in South America and Southeast Asia, infecting 400 million people annually. While most of these infections are asymptomatic or mild, ∼1–3% of cases progress to severe dengue hemorrhagic fever, causing approximately 10–15,000 deaths each year (Bhatt et al., 2013). Unlike most infectious diseases, DENV secondary infections are more likely than primary infections to cause severe disease, with estimates of relative risk of severe dengue of ∼24 (Mizumoto et al., 2014). Primary DENV infection is more often mild and is thought to generate lifelong homotypic immunity and temporary heterotypic immunity, which typically wanes over six months to two years (Katzelnick et al., 2016; Sabin, 1952). Subsequent heterotypic secondary infection induces broad cross-protection, and symptomatic tertiary and quaternary cases are rare (Gibbons et al., 2007; Olkowski et al., 2013). However, a small subset of secondary infections are enhanced by nonneutralizing, cross-reactive antibodies, resulting in severe disease (Halstead, 1979; Katzelnick et al., 2017; Sangkawibha et al., 1984). Thus, the antigenic relationships between dengue viruses — describing whether the immune response generated after primary infection results in protection or enhancement of secondary infection — are key drivers of DENV case outcomes and epidemic patterns.

While each serotype is clearly genetically and antigenically distinct, it is not clear how sub-serotype clades of DENV interact antigenically. Each DENV serotype consists of broad genetic diversity, including canonical clades termed 'genotypes' (Figure 1) (Rico-Hesse, 1990; Twiddy et al., 2003). Specific genotypes have been associated with characteristically mild or severe disease, and heterogeneous neutralization titers suggest that the immune response to some genotypes is more cross-protective than others (Gentry et al., 1982; Russell and Nisalak, 1967). Until recently, it has been assumed that these intraserotype differences are minimally important compared to interserotype differences. However, empirical evidence has demonstrated that these genotype-specific differences can drive case outcomes and epidemic severity (reviewed in Holmes and Twiddy, 2003). For example, analysis of a longitudinal cohort study demonstrated that specific combinations of primary infection serotype and secondary infection genotype can mediate individual case outcomes (OhAinle et al., 2011). On a population scale, the DENV1-immune population of Iquitos, Peru, experienced either entirely asymptomatic or very severe epidemic seasons in response to two different genotypes of DENV2 (Kochel et al., 2002).

One explanation for these and similar observations is that overlooked intraserotype antigenic variation contributes to these genotype-specific case outcomes and epidemic patterns. Recent efforts to antigenically characterize diverse DENV viruses suggests that each serotype may contain antigenic heterogeneity, but the source and impact of this heterogeneity is not clear (Katzelnick et al., 2015). Here, we take a phylogenetic approach to characterize the evolutionary basis for observed antigenic heterogeneity among DENV clades. We also quantify the impact of within- and between-serotype antigenic variation on real-world DENV population dynamics.

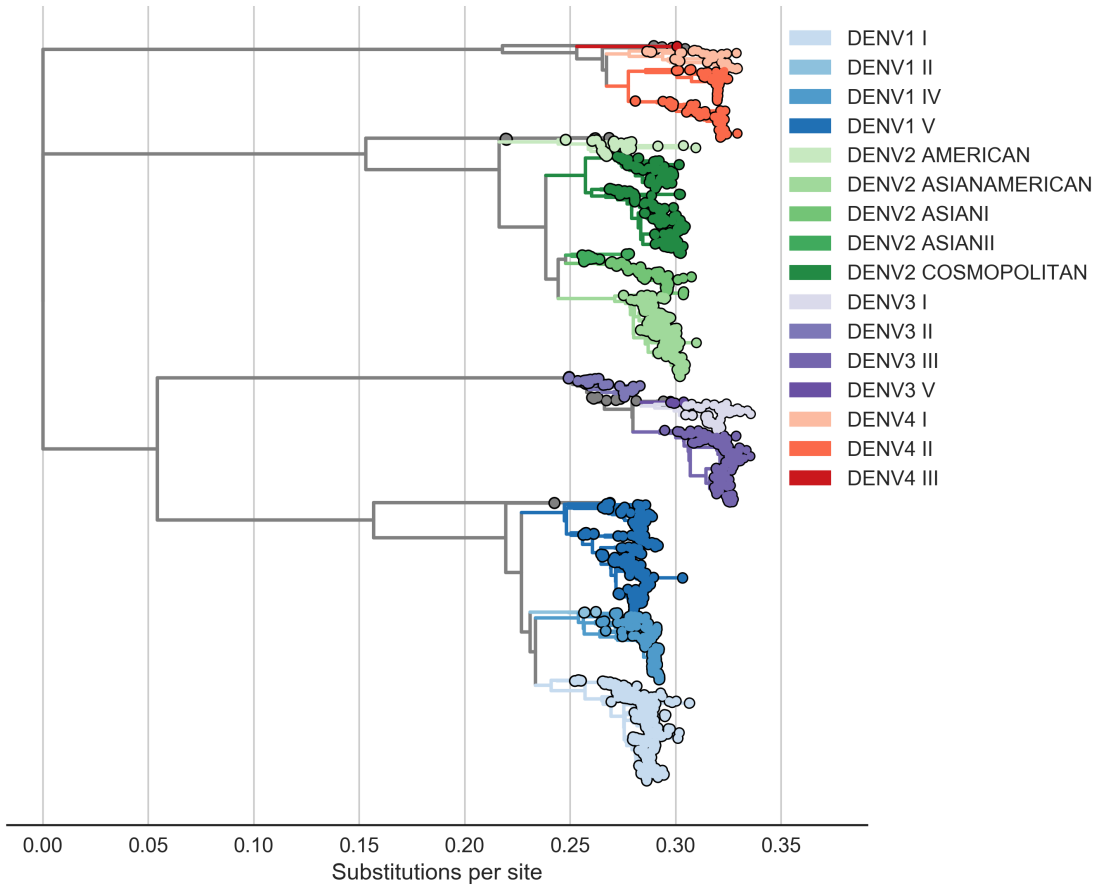

Figure 1. Phylogeny of dengue viral sequences. Maximum likelihood phylogeny of dengue virus genomes, colored by canonical genotype assignment.

Results

Dengue neutralization titer data

Antigenic distance between a pair of viruses i and j is experimentally quantified using neutralization titer, which measures how well serum drawn after infection with virus i is able to neutralize virus j in vitro(Russell and Nisalak, 1967). To measure the pairwise antigenic distances for a panel of diverse DENV viruses (Figure S1), Katzelnick et al. infected naive non-human primates (NHP) with each virus, drew sera at three months post-infection, and then titered this sera against a panel of test viruses (Katzelnick et al., 2015). To compare patterns of cross-protection in NHP and humans, they also drew sera from 31 study participants six weeks after inoculation with a monovalent component of the NIH dengue vaccine candidate. This sera was also titered against a broad panel of DENV viruses. As originally reported, we find generally consistent patterns of neutralization between the NHP and human sera data; see (Katzelnick et al., 2015) for a detailed comparison. In total, our dataset consists of 454 NHP sera titrations spanning the breadth of DENV diversity,

and 728 human sera titrations providing deep coverage of a small subset of viruses.

To standardize these measurements, we first take the $\log_2$ of each value, such that one titer unit corresponds to one, two-fold drop in neutralization. We then define antigenic distance between autologous virus-sera pairs (i.e., virus i and serum i) as zero. Normalized antigenic distance between i and j are thus calculated as $D_{ij} = \log_2(T_{ii}) - \log_2(T_{ij})$, such that a higher value of D_{ij} indicates that serum i is less effective at neutralizing virus j, implying greater antigenic distance between viruses i and j.

The full dataset of standardized titer values is shown in Figure 2. Here, we see that homotypic virus-serum pairs are more closely related antigenically than heterotypic pairs. However, we also observe large variance around this trend, both within and between serotypes. This suggests that treating each serotype as antigenically uniform overlooks potentially important antigenic heterogeneity across viruses within each serotype.

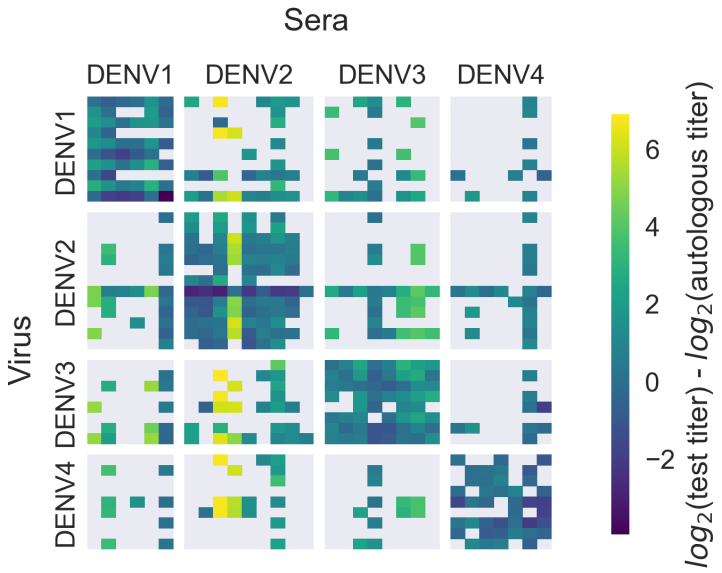

Figure 2. Normalized antigenic distance between pairs of dengue viruses and sera. Aggregated neutralization titers from Katzelnick et al. are standardized such that the distance between autologous virus-serum pairs is 0, and each titer unit corresponds to one, two-fold change in PRNT50 value. Light gray areas represent missing data. Higher values correspond to greater antigenic distance.

Dengue antigenic evolution corresponds to phylogenetic divergence

Titer measurements are prone to noise, and there is a limited amount of available titer data. If the antigenic heterogeneity observed in the raw data is truly the result of an underlying evolutionary process, we expect that changes in antigenic phenotype correspond to underlying changes in viral genotype. Figure 3 shows the relationship between genetic and antigenic distance between each pair of viruses in our dataset. There are two groups of comparisons — between serotype and within serotype — however, even within serotypes there is significant genetic diversity and a correlation between increased genetic distance

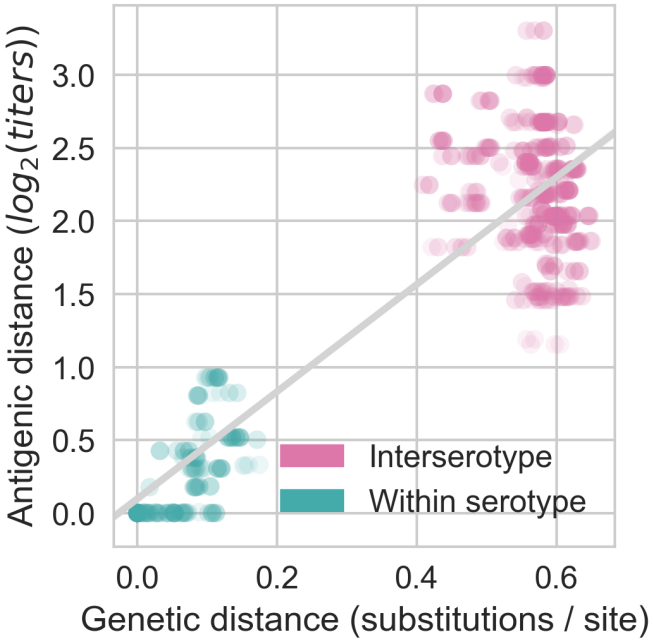

Figure 3. Normalized antigenic distance vs. genetic distance between pairs of dengue viruses. Antigenic distances are the same as from Figure 2; genetic distances are the patristic distances between viral genomes on a maximum likelihood phylogeny.

and increased antigenic distance. The relationship between genetic distance and antigenic distance is consistent within and between serotypes, where increasing genetic divergence generally corresponds to increased antigenic distance.

Antigenic evolution occurs within serotypes

To fully map the relationship between DENV genetic and antigenic evolution, we adapt a phylogeny-based model originally developed for influenza (Neher et al., 2016). Conceptually, this model predicts titer values through three steps. First, we build a phylogeny of dengue virus sequences to establish the genetic relationships between viruses (Figures 1, S2). Next, we infer how much antigenic change has occurred along each branch of the phylogeny by mapping titer changes to individual branches. This assigns each branch b an antigenic distance d_b. With this in hand, we estimate the antigenic distance between all pairs of viruses by tracing the path between them in the phylogeny, summing branch-specific distances d_b as we go (Methods, Eq. 2).

To learn these values of d_b, we first split our dataset into training (random 90% of measurements) and test data (the remaining 10% of values). We take the training data and fit d_b for each branch in the tree, subject to regularization as follows (also detailed in Methods, Eq. 3). Parsimoniously, we expect that antigenic change is more likely to occur through larger changes on a few branches than through small changes on many branches; correspondingly, our prior expectation of values of d_b is exponentially distributed

such that most values of $d_b = 0$. This is analogous to lasso regression to identify a few parameters with positive weights and set other parameters to 0. Additionally, some viruses have greater binding avidity, and some sera are more potent than others (Figure S3); these 'row' and 'column' effects, respectively, are normally distributed and are taken into account when estimating titers. The model uses convex optimization to learn the values of d_b that minimize the sum of squared errors (SSE) between observed and predicted titers in the training data.

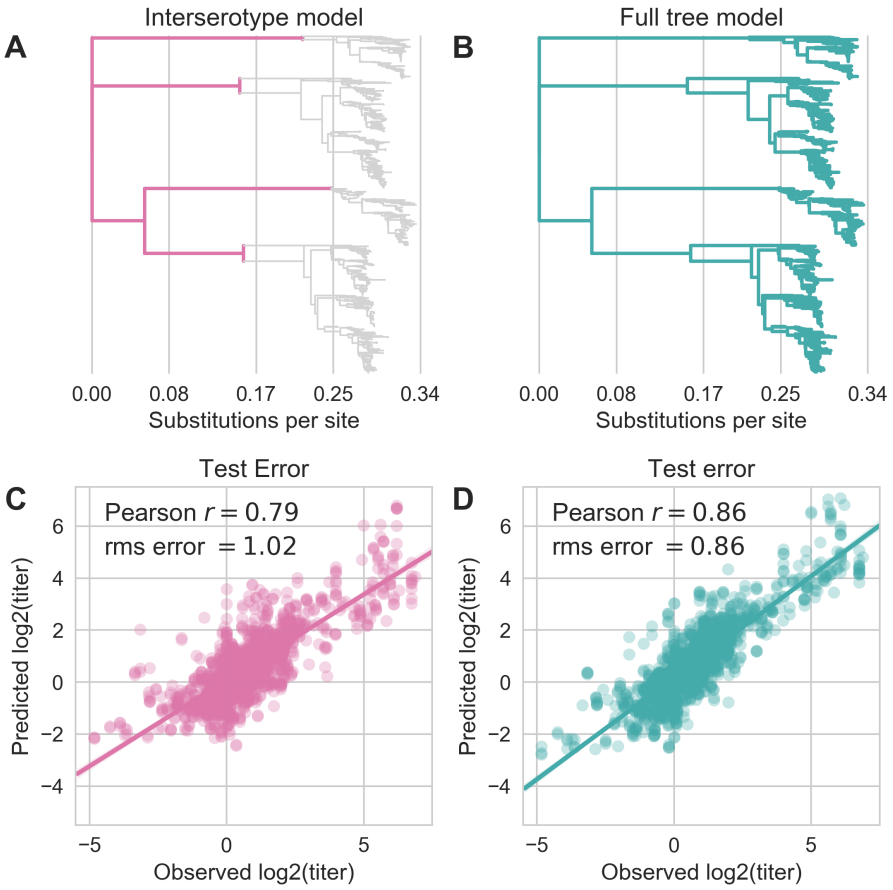

Figure 4. Titer model formulations and performance. **A** The 'interserotype model' only allows branches that lie between serotypes to contribute to antigenic evolution. All other branches are assigned $d_b = 0$. **B** The 'full tree model' allows any branch in the phylogeny to contribute to antigenic evolution ($d_b >= 0$). **C,D** Predictive performance of each model on the test dataset (aggregated from 10-fold cross-validation). Shading represents the 95% confidence interval.

This model formulation is an effective tool for estimating antigenic relationships between viruses based on their relative positions in the phylogeny. We can use variations of this model to explicitly test whether the observed antigenic phenotypes are better explained by the hypothesis that dengue serotypes are antigenically uniform ('interserotype model'), or by the hypothesis that serotypes are antigenically diverse ('full tree model'). In the interserotype model, we set $d_b = 0$ for all branches in the tree that do not lie between serotypes (Figure 4A). Alternatively, the 'full tree model' allows any branch in the phylogeny

to contribute to antigenic evolution (Figure 4B). For each model, we learn model parameters from the training data, and then use those parameters to predict test data values. We assess model performance by comparing the predicted test titer values to the actual values. Model performance indicates how well the hypothesis embedded in the model explains the observed data.

We find that serotype-level characterization alone explains observed antigenic phenotypes to a reasonable degree. On average, this interserotype model predicts titers within 1.02 $\log_2$ titer units of the true value (root mean squared error, RMSE), and explains 62% of the observed variation in neutralization titers overall (Figure 4).

However, we find that accounting for within-serotype antigenic evolution substantially improves our ability to explain dengue antigenic phenotypes. The full tree model is able to predict test titers within 0.86 $\log_2$ titer units of the true value (RMSE approaching the level of error intrinsic to the assay), and explains 74% of the observed variation in neutralization titers overall (Figure 4). Importantly, all reported error metrics refer to performance on test data, so this difference in model performance is not due to the number of free parameters. The full tree model performance is comparable to the model error from a cartography-based characterization of the same dataset (RMSE 0.65–0.8 $\log_2$ titer units), and to the error observed when this model was used to characterize an influenza dataset (RMSE of 0.5 $\log_2$ titer units) (Katzelnick et al., 2015; Neher et al., 2016). From this, we conclude that there is antigenic evolution within each serotype of DENV, and that this is driven by underlying genetic divergence.

Collapsing putatively antigenically uniform clades, we observe at least 12 distinct antigenic phenotypes of DENV (Figure 5) and reject the null hypothesis that all DENV viruses within a serotype are antigenically uniform. The titer dataset spans the breadth of canonical DENV genotypes, but in most cases lacks the resolution to detect within-genotype antigenic diversity. We thus expect that these results represent a lower-bound on the true extent of DENV intraserotype antigenic diversity.

Antigenic novelty predicts serotype success

From the titer model, we observe strong evidence that homotypic genotypes of DENV vary in their ability to escape antibody neutralization (Figures 4, S5). However, antibody neutralization is only one of many factors that shape epidemic patterns. We investigate whether the observed antigenic diversity influences dengue population dynamics in the real world. The size of the viral population (i.e., prevalence, commonly analyzed using SIR models as reviewed in (Lourenço et al., 2018)) is determined by many complex factors, and reliable values for population prevalence are largely unavailable. Contrastingly, the composition of the viral population (i.e., the relative frequency of each viral clade currently circulating) can be estimated over time by examining historical sequence data (Lee et al., 2018; Neher et al., 2016), and is primarily driven by viral fitness (Bedford et al., 2011). In meaningfully antigenically diverse viral populations, antigenic novelty (relative to standing population immunity) contributes to viral fitness: as a given virus i circulates in a population, the proportion of the population that is susceptible to infection with i–and other viruses antigenically similar to i–decreases over time as more people acquire

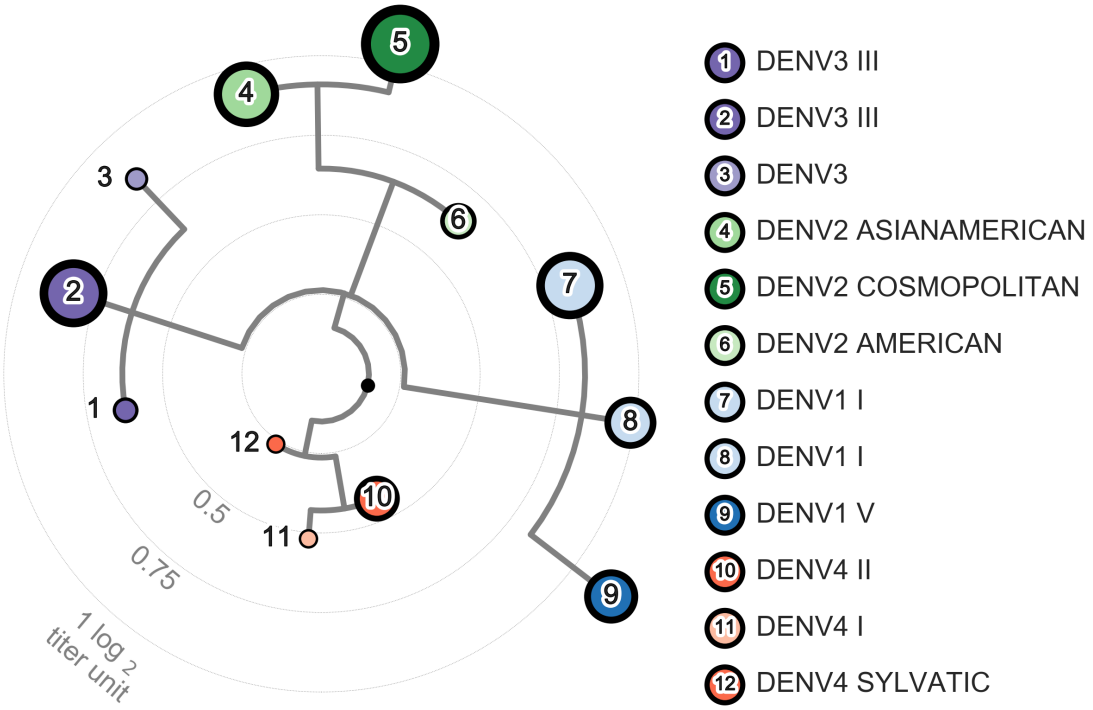

Figure 5. Tree of dengue antigenic phenotypes. We infer the topology from a maximum likelihood phylogeny of DENV genomes. Branch lengths are scaled to reflect d_b, the relative contributions of each branch to antigenic divergence, as inferred by the 'full tree' model of DENV evolution. Antigenically uniform clades are collapsed, and node diameter reflects the size of the collapsed clade. Figure S4 offers an additional view of the same topology.

immunity (Bedford et al., 2012; Łuksza and Lässig, 2014). Antigenically novel viruses that are able to escape this population immunity are better able to infect hosts and sustain transmission chains, making them fitter than the previously circulating viruses (Bedford et al., 2012; Zhang et al., 2005). Thus, if antigenic novelty constitutes a fitness advantage for DENV, then we would expect greater antigenic distance from recently circulating viruses to correlate with higher growth rates.

To test this hypothesis, we examine the composition of the dengue virus population in Southeast Asia from 1970 to 2015. We estimate the relative population frequency of each DENV serotype at three month intervals, $x_i(t)$ (Figure 6A), based on available sequence data (Methods, Eq. 4).

Fitter virus clades increase in frequency over time, such that $x_i(t+dt) > x_i(t)$. It follows that these clades have a growth rate–defined as the fold-change in frequency over time–greater than one: $\frac{x_i(t+dt)}{x_i(t)} > 1$. To isolate the extent to which antigenic fitness contributes to clade success and decline, we extend work by Łuksza and Lässig* (Łuksza and Lässig, 2014) to build a simple model that attempts to predict clade growth rates based on two variables: the antigenic fitness of the clade at time t, and a time-invariant free parameter representing the intrinsic fitness of the serotype the clade belongs to. We estimate the antigenic fitness

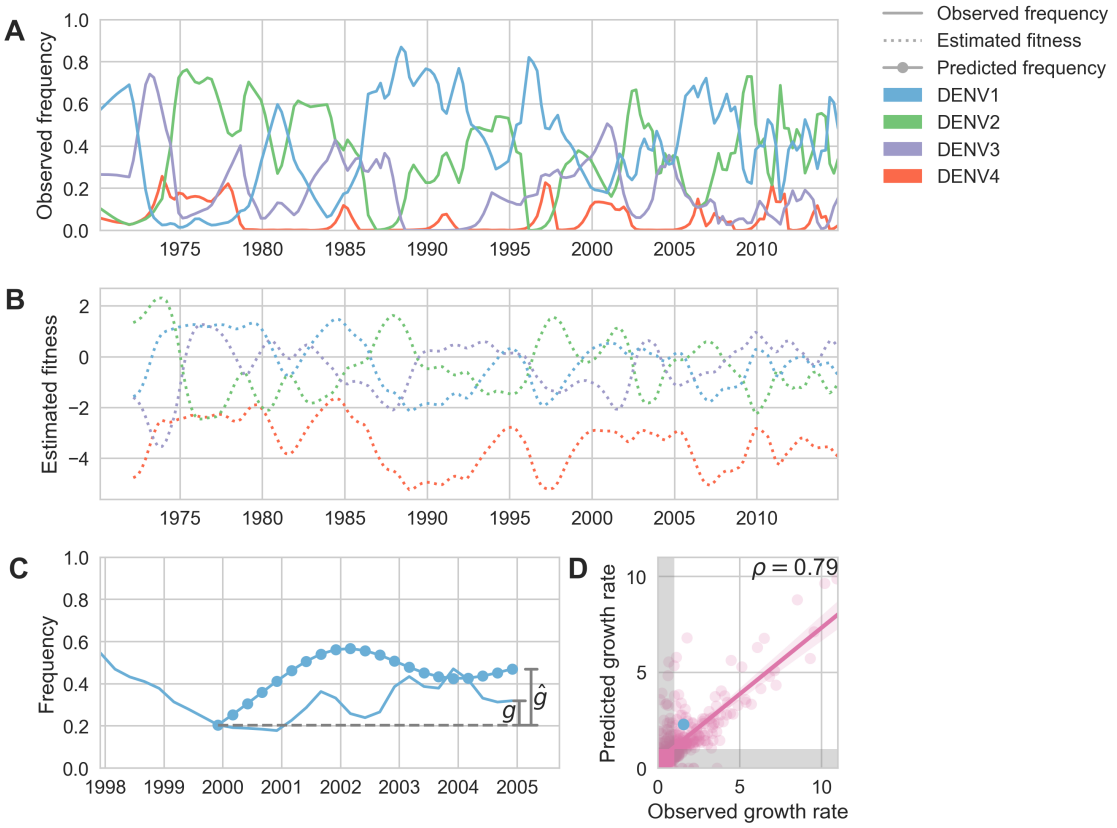

Figure 6. Antigenic novelty predicts serotype success. A The relative frequency of each serotype, x_i, in Southeast Asia is estimated every three months based on available sequence data. **B** We calculate antigenic fitness for each serotype over time as its frequency-weighted antigenic distance from recently circulating viruses. We then add this to a time-invariant intrinsic fitness value to calculate total fitness (shown here, arbitrary units). **C** illustrates how the model predicts clade growth rates. At each timepoint t, we blind the model to all empirical data from timepoints later than t and predict each serotype's future trajectory based on its initial frequency, time-invariant intrinsic fitness, and antigenic fitness at time t (Methods, Eq. 10). We predict forward in three-month increments for a total prediction period of $dt = 5$ years. At each increment, we use the predicted stepwise frequency change to adjust our estimates of antigenic fitness on a rolling basis (Methods, Eq. 14). Predicted growth rates are calculated as $\hat{g} = \frac{\hat{x_i}(t+dt)}{x_i(t)}$ and compared to empirically observed growth rates, $g = \frac{x_i(t+dt)}{x_i(t)}$ in **D**. The example illustrated in **C** is also shown in **D** as the blue point. Serotype growth versus decline is accurate (i.e., the predicted and actual growth rates are both > 1 or both < 1, all points outside the gray area) for 80% of predictions.

of clade i at time t as a function of its antigenic distance from each viral clade j that has circulated in the same population over the previous two years, weighted by the relative frequency of j and adjusted for waning population immunity (Figure 6B; Methods, Eq. 7). Growth rates are estimated based on a five year sliding window (Figure 6C).

This simple model explains 62% of the observed variation in serotype growth rates, and predicts serotype growth vs decline correctly for 80% of predictions (Figure 6D). This strongly suggests that antigenic fitness is a major driver of serotype population dynamics.

This also demonstrates that this model captures key components of dengue population dynamics; examining the formulation of this model in more detail can yield insights into how antigenic relationships influence DENV population composition. The fitness model includes eight free parameters that are optimized such that the model most accurately reproduces the observed fluctuations in DENV population composition (reduction in prediction error relative to a null model, see Methods). We find that serotype fluctuations are most consistent with a model wherein population immunity wanes linearly over time, with the probability of protection dropping by about 48% per year for the first two years after primary infection. We also find that these dynamics are best explained by intrinsic fitness that moderately varies by by serotype (Table S1).

Antigenic novelty also partially predicts genotype success

To estimate how well antigenic fitness predicts genotype dynamics, we used the same model to predict genotype success and decline. As before, fitness of genotype i is based on the intrinsic fitness of the serotype i belongs to, and the antigenic distance between i and each other genotype, j, that has recently circulated (Figure 7B). Importantly, we can calculate antigenic distance between i and j at the serotype level (i.e., the antigenic distances computed from the 'interserotype model' as illustrated in Figure 4A) or at the genotype level (i.e., the antigenic distances computed by the 'full tree model' as illustrated in Figure 4B, which incorporates the observed within-serotype heterogeneity). If within-serotype antigenic heterogeneity contributes to genotype fitness, then we would expect estimates of antigenic fitness based on the 'full tree model' to better predict genotype growth rates.

We find that antigenic fitness contributes to genotype turnover, although it explains less of the observed variation than for serotypes. When antigenic distance is estimated from the 'interserotype model', we find that our model of antigenic fitness explains approximately 31% of the observed variation in genotype growth rates, and correctly predicts genotype growth vs. decline 68% of the time (Figure 7C). Perhaps surprisingly, more precise estimates of antigenic distance between genotypes from the 'full tree model' does not improve our predictions of genotype success ($R^2 = 0.28$, 70% accuracy; Figure 7D). This suggests that although we find strong evidence that genotypes vary in their ability to escape neutralizing antibodies (Figures 4, S5), these differences are subtle enough that they do not impact broad-scale regional dynamics over time.

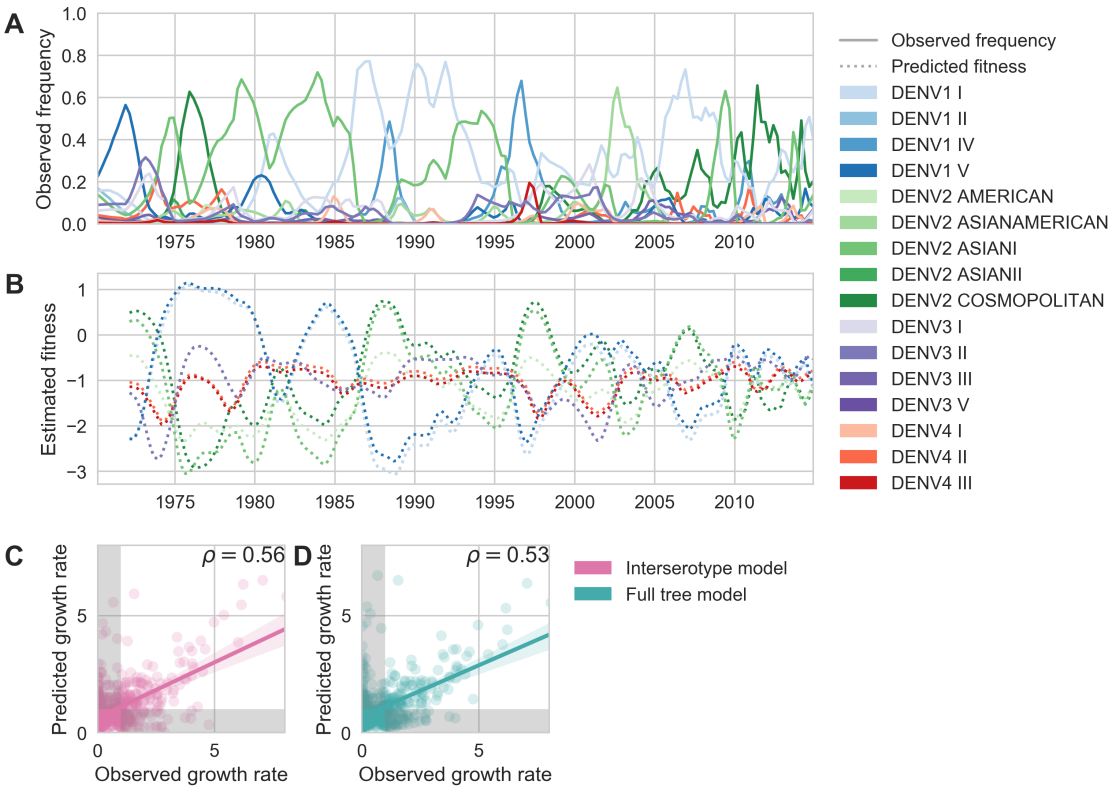

Figure 7. Relative frequencies and fitness of dengue genotypes, 1970-2015. A Relative frequencies of each canonical dengue genotype across Southeast Asia, estimated from available sequence data. **B** Antigenic fitness is calculated for each genotype as its frequency-weighted antigenic distance from recently circulating genotypes. We then add this to a time-invariant, serotype-specific intrinsic fitness value to calculate total fitness (shown here, arbitrary units). We assess antigenic distance using either the 'full tree model' or the 'interserotype model' of antigenic relationships (Figure 4). In this panel, we show total fitness over time, incorporating estimates of antigenic fitness derived from the 'full tree' model. **C, D** Fitness estimates were used to predict clade growth rates over 5 years, compounding immunity every three months based on predicted frequency changes (Methods Eq. 14). Here, we compare observed vs predicted growth rates for both formulations of the fitness model (using fitness derived from either the 'interserotype' or 'full tree' model estimates of antigenic distance). Growth versus decline was accurate (predicted and actual growth rates both > 1 or both < 1, points outside the gray shaded area) for 68% – 70% of predictions, respectively.

Discussion

Within-serotype antigenic heterogeneity

We show that mapping antigenic change to individual branches and interpolating across the DENV phylogeny is able to explain a large majority of the observed variation in antigenic phenotypes, as measured by neutralization titers. This demonstrates that DENV antigenic divergence is closely coupled to genetic divergence. We also find that accounting for within-serotype antigenic evolution is necessary to explain the observed variation in antigenic phenotypes. This supports and expands upon previous reports (Katzelnick et al.,

2015) that the null hypothesis of antigenically uniform serotypes is inconsistent with observed patterns of cross-protection and susceptibility. We observe at least 12 distinct antigenic phenotypes that are both genetically and antigenically distinct. Analysis of the recent CYD-TDV vaccine trial shows different vaccine efficacy against genotypes I and II of DENV4 (Juraska et al., 2018); the antigenic distance between these two genotypes is comparable to the antigenic distance separating each of the 12 antigenic phenotypes we report (Figures 5, S5). This suggests that these phenotypes may be sufficiently distinct to have important impacts on secondary case outcomes and vaccine efficacy.

Overall, we expect that these antigenic phenotypes represent a lower-bound on the extent, magnitude, and nature of antigenic heterogeneity with DENV. Our current titer dataset spans the breadth of DENV diversity, but due to small sample size, it lacks the resolution to detect most sub-genotype antigenic variation or to identify which specific mutations correspond to changes in antigenic phenotype. The appearance of the deep antigenic divergence of the four serotypes, and the more recent antigenic divergences within each serotype, suggest that DENV antigenic evolution is likely an ongoing, though gradual, process. We therefore expect that future studies with richer datasets will find additional antigenic variation within each genotype. This dataset also contains many left-censored titer values, where we know two viruses are at least T titer units apart, but do not know exactly how far apart. If we knew the true value of these censored titers, many of them would indicate larger antigenic distances than the reported values, T, which are used to train the model. Thus, it is likely that our model systematically underestimates the magnitude of titer distances. Finally, antibody neutralization and escape (as measured by PRNT titers) is only one component of the immune response to DENV. Although analysis of a longitudinal cohort study shows that these neutralization titers correlate with protection from severe secondary infection, it is unclear how PRNT titers correspond to antibody-dependent enhancement (Katzelnick et al., 2016). It is also important to note that DENV case outcomes are partially mediated by interactions with innate and T-cell immunity, the effects of which are not captured in neutralization titers (Green et al., 2014). Overall, while richer datasets and the development of more holistic assays will be required in order to fully characterize the extent of DENV antigenic diversity, it is clear that the four-serotype model is insufficient to explain DENV antigenic evolution.

Viral clade dynamics

We use these inferred antigenic relationships to directly quantify the impact of antigenic fitness on DENV population composition. To do so, we measure serotype frequencies across Southeast Asia over time and construct a model to estimate how they will fluctuate (Methods, Eq. 6–19). This model places a fitness value on each serotype that derives from a constant intrinsic component alongside a time-dependent antigenic component. Antigenic fitness declines with population immunity, which is accumulated via the recent circulation of antigenically similar viruses. We find that antigenic fitness is able to explain most of the observed variation in serotype growth and decline (Figure 6). This demonstrates that antigenic fitness is a strong determinant of DENV serotype dynamics in a real-world, hyperendemic population.

We similarly use this model to quantify the effect of within-serotype antigenic variation on the success and decline of canonical DENV genotypes (Figure 7). As above, genotype antigenic fitness declines with population immunity. Here, we estimate population immunity based on antigenic distance from recently circulating genotypes, using distances derived from the 'interserotype model' or the 'full tree model' of DENV antigenic evolution. We then directly compare how strongly these coarser serotype-level versus specific genotype-level antigenic relationships impact DENV population dynamics. Overall, we find that antigenic fitness explains a moderate portion of the observed variation in genotype growth and decline. Surprisingly, however, we find that incorporating within-serotype antigenic differences does not improve our predictions (Figure 7C,D). This suggests that although genotypes are antigenically diverse, these differences do not appear to influence large-scale regional dynamics over time.

However, this observation is subject to caveats imposed by the the available data and model assumptions. Our estimates of antigenic fitness are informed by the antigenic distances inferred by the titer tree model; thus, as above, we are unable to account for nuanced antigenic differences between sub-genotype clades of DENV due to limited titer data. We estimate DENV population composition over time based on available sequence data, pooled across all of Southeast Asia (Methods, Eq. 4). As the vast majority of cases of DENV are asymptomatic, sequenced viruses likely represent a biased sample of more severe cases from urban centers where patients are more likely to seek and access care. We also assume that Southeast Asia represents a closed viral population with homogeneous mixing. However, increasing globalization likely results in some amount of viral importation that is not accounted for in this model (Allicock et al., 2012). Finally, although Southeast Asia experiences hyperendemic DENV circulation, the majority of DENV transmissions are hyper-local (Salje et al., 2017), and viral populations across this broad region may not mix homogeneously each season. Thus, it is possible that these sub-serotype antigenic differences impact finer-scale population dynamics, but we lack the requisite data to examine this hypothesis.

Conclusions

We find that within-serotype antigenic evolution is necessary to explain the observed patterns of cross-immunity and susceptibility after primary DENV infection. These within-serotype differences are large enough that we believe they likely impact secondary case outcomes and small scale population dynamics. We also find that population immunity is a strong determinant of the composition of the DENV population across Southeast Asia, although this is putatively driven by coarser, serotype-level antigenic differences. As richer datasets become available, future studies that similarly combine viral genomics, functional antigenic characterization, and population modeling have great potential to improve our understanding of how DENV evolves antigenically and moves through populations.

Model sharing and extensions

We have provided all code, configuration files and datasets at 'github.com/blab/dengue-antigenic-dynamics', and wholeheartedly encourage other groups to adapt and extend this framework for further investigation of DENV antigenic evolution and population dynamics.

Methods

Data

Sequences

We downloaded all dengue virus sequences available from the Los Alamos National Lab Hemorrhagic Fever Virus Database as of March 7, 2018, that contained the full coding sequence of E (total N=12,645) (Kuiken et al., 2011). We discarded sequences which were putative recombinants, duplicates, lab strains, or which lacked an annotated sampling location and/or sampling date. We then randomly subsampled up to 8 viruses per region, per month, preferentially including records with available titer data and longer sequences. Our final dataset consists of 2,563 viral sequences (Figure S2). We used the annotated reference dataset from (Pyke et al., 2016) to assign sequences to canonical genotypes.

Titers

Antigenic distance between pairs of viruses i and j is experimentally measured using a neutralization titer, which measures how well serum drawn after infection with virus i is able to neutralize virus j in vitro (Russell and Nisalak, 1967). Briefly, two-fold serial dilutions of serum i are incubated with a fixed concentration of virus j. Titers represent the lowest serum concentration able to neutralize 50% of virus, and are reported as the inverse dilution. We used two publicly available plaque reduction neutralization titer (PRNT50) datasets generated by Katzelnick et al. in (Katzelnick et al., 2015). The primary dataset was generated by infecting each of 36 non-human primates with a unique strain of DENV. NHP sera was drawn after 12 weeks and titered against the panel of DENV viruses. The secondary dataset was generated by vaccinating 31 human trial participants with a monovalent component of the NIH DENV vaccine. Sera was drawn after 6 weeks and titered against the same panel of DENV viruses. As discussed in Katzelnick et al., these two datasets show similar patterns of antigenic relationships between DENV viruses. In total, our dataset includes 47 virus strains, 36 serum strains, and 1182 measurements.

Titer Model

We compute standardized antigenic distance between virus i and serum j (denoted D_{ij}) from measured titers relative to autologous titers (denoted T_{ii} and T_{ij}, respectively), such that

$$D_{ij} = \log_2(T_{ii}) - \log_2(T_{ij}) \tag{1}$$

To predict unmeasured titers, we employ the 'tree model' from Neher et al. and implemented in Nextstrain, which assumes that antigenic evolution is driven by underlying genetic evolution (Hadfield et al., 2018; Neher et al., 2016). We use RAxML version 8.2.11 and a GTRCAT nucleotide substitution model to build a maximum likelihood phylogeny of dengue viral genome sequences (Stamatakis, 2014). Observed titer drops are mapped to branches in the viral phylogeny after correcting for overall virus avidity, v_i, and serum

potency, p_j ('row' and 'column' effects, respectively):

$$\hat{D}_{ij} \approx D_{ij} = \sum_{b \in path(i,j)} d_b + v_i + p_j \tag{2}$$

where d_b is the titer drop assigned to each branch, b, in the phylogeny. We randomly withhold 10% of titer measurements as a test set. We use the remaining 90% of titer measurements as a training set to learn values for virus avidity, serum potency, and branch effects. As in Neher et al., we formulate this as a convex optimization problem and solve for these parameter values to minimize the cost function:

$$C = \sum_{i,j} (\hat{D}_{ij} - D_{ij})^2 + \lambda \sum_b d_b + \gamma \sum_i v_i^2 + \delta \sum_j p_j^2 \tag{3}$$

Respectively, these terms represent the squared training error; an L1 regularization term on branch effects, such that most values of $d_b = 0$; and L2 regularization terms on virus avidities and serum potencies, such that they are normally distributed. These parameter values are then used to predict the antigenic distance between all pairs of viruses, i and j, in the phylogeny. We assess performance by comparing predicted to known titer values in our test data set, and present test error (aggregated from 10-fold cross-validation) throughout the manuscript.

Viral Clade Dynamics

Empirical Clade Frequencies

As discussed in Neher et al and Lee et al, we estimate empirical clade frequencies from 1970 to present based on observed relative abundance of each clade in the 'slice' of the phylogeny corresponding to each quarterly timepoint (Lee et al., 2018; Neher et al., 2016).

Briefly, the frequency trajectory of each clade in the phylogeny is modeled according to a Brownian motion diffusion process discretized to three-month intervals. Relative to a simple Brownian motion, the expectation includes an 'inertia' term that adds velocity to the diffusion and the variance includes a term $x(1-x)$ to scale variance according to frequency following a Wright-Fisher population genetic process. This results in the following diffusion process:

$$x(t+dt) = \mathcal{N}\left(x(t) + \epsilon dx,\ dt\sigma^2 x(t)(1-x(t))\right) \tag{4}$$

with 'volatility' parameter σ^2. The term dx is the increment in the previous timestep, so that $dx = x(t) - x(t-dt)$. We used $\epsilon = 0.7$ and $\sigma = 2.0$ to maximize fit to empirical trajectory behavior.

We also include an Bernoulli observation model for clade presence / absence among sampled viruses at timestep t. This observation model follows

$$f(x,t) = \Pi_{v \in V} x(t)\ \Pi_{v \notin V}(1 - x(t)) \tag{5}$$

where $v \in V$ represents the set of viruses that belong to the clade and $v \notin V$ represents the set of viruses that do not belong to the clade. Each frequency trajectory is estimated

by simultaneously maximizing the likelihood of the process model and the likelihood of the observation model via adjusting frequency trajectory $\vec{x} = (x_1, ...x_n)$.

Population Immunity

For antigenically diverse pathogens, antigenic novelty represents a fitness advantage (Lipsitch and O'Hagan, 2007). This means that viruses that are antigenically distinct from previously-circulating viruses are able to access more susceptible hosts, allowing the antigenically novel lineage to expand. We adapt a simple deterministic model from Łuksza and Lässig to directly quantify dengue antigenic novelty and its impact on viral fitness (Łuksza and Lässig, 2014). We quantify population immunity to virus i at time t, $P_i(t)$, as a function of which clades have recently circulated in the past N years, and how antigenically similar each of these clades is to virus i:

$$P_i(t) = \sum_{n=1}^{n=N} \left(w(n) \sum_j \Big(x_j(t-n)\, C(D_{ij}) \Big) \right) \tag{6}$$

Where D_{ij} is the antigenic distance between i and each non-overlapping clade j, n is the number of years since exposure, and $x_j(t-n)$ is the relative frequency of j at year $t-n$. Waning immunity is modeled as a non-negative linear function of time:

$$w(n) = max(-\gamma n + 1, 0) \tag{7}$$

The relationship between antigenic distance and the probability of protection, C, is also assumed to be linear and non-negative, such that:

$$C(D_{ij}) = max(-\sigma D_{ij} + 1, 0) \tag{8}$$

We model the effects of population immunity, $P_i(t)$, on viral antigenic fitness, $f_i(t)$, as:

$$f_i(t) = f_0 - \beta P_i(t) \tag{9}$$

where β and f_0 are fit parameters representing the slope of the linear relationship between immunity and fitness, and the intrinsic relative fitness of each serotype, respectively.

Frequency Predictions

Similar to the model implemented in Łuksza and Lässig*, we estimate predicted clade frequencies at time $t + dt$ as

$$\hat{x_i}(t + dt) = \frac{x_i(t) e^{f_i(t) dt}}{\sum_i x_i(t) e^{f_i(t) dt}} \tag{10}$$

for short-term predictions (where $dt < 1$ year).

For long-term predictions, we must account for immunity accrued at each intermediate timepoint between t and dt. We divide the interval between t and dt into a total of U 3 month timepoints, $[t + u, t + 2u, ...t + U]$, such that $t + U = dt$. We then compound immunity based on predicted clade frequencies at each intermediate timepoint:

$$\hat{x_i}(t + u) = x_i(t) e^{f_i(t) u} \tag{11}$$

$$\hat{x_i}(t+2u) = \hat{x_i}(t+u)e^{f_i(t+u)u} \tag{12}$$

...

$$\hat{x_i}(t+U) = x_i(t)e^{f_i(t)u}e^{f_i(t+u)u}e^{f_i(t+2u)u}...e^{f_i(t+U)u} \tag{13}$$

$$\hat{x_i}(t+dt) = \hat{x_i}(t+U) = x_i(t)e^{\sum_u f_i(t+u)u} \tag{14}$$

We then calculate clade growth rates, defined as the fold-change in relative clade frequency between time t and time $t+dt$:

$$\frac{\hat{x_i}(t+dt)}{x_i(t)} \tag{15}$$

Null model and model performance

To quantify the impact of antigenic fitness on DENV clade success, we compare our antigenically-informed model to a null model wherein all viruses have equal antigenic fitness at all timepoints:

$$f_i^{null}(t) = 0 \tag{16}$$

$$\hat{x_i}^{null}(t+dt) = x_i(t)e^0 = x_i(t) \tag{17}$$

For both the null model and the antigenically-informed model, we can assess predictive power as the sum of squared error between predicted and empirical clade growth rates:

$$SSE = \sum_{i,t}\left(\frac{\hat{x_i}(t+dt)}{x_i(t)} - \frac{\hat{x_i}^{null}(t+dt)}{x_i(t)}\right)^2 \tag{18}$$

We can then estimate how much more error is present in the null model than the antigenically-informed model:

$$\Delta SSE = SSE^{null} - SSE^{model} \tag{19}$$

Our frequency prediction model has a total of 8 free parameters:

β	Slope of linear relationship between population immunity and viral fitness
γ	Slope of linear relationship between titers and probability of protection
σ	Proportion of titers waning each year since primary infection
f_{s0}	Relative intrinsic fitness of each serotype ($f_0 = 0$ for DENV4)
N	Number of years of previous immunity that contribute to antigenic fitness
dt	Number of years in the future to predict clade frequencies

For each dataset, we jointly fit these parameters to maximize ΔSSE (Table S1).

Data availability

Sequence and titer data, as well as all code used for analyses and figure generation, is publicly available at github.com/blab/dengue-antigenic-dynamics.

Acknowledgements

We would like to thank Richard Neher, Molly OhAinle, David Shaw, Paul Edlefsen, Michal Juraska, and all members of the Bedford Lab for useful discussion and advice. SB is a Graduate Research Fellow and is supported by NSF DGE-1256082. TB is a Pew Biomedical Scholar and is supported by NIH R35 GM119774-01. LK is supported by NIH awards R01AI114703-01 and P01AI106695.

References

Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, Suchard MA, Foster JE, Rambaut A, Carrington CV. Phylogeography and population dynamics of dengue viruses in the Americas. Molecular biology and evolution. 2012; 29(6):1533–1543.

Bedford T, Cobey S, Pascual M. Strength and tempo of selection revealed in viral gene genealogies. BMC Evolutionary Biology. 2011; 11(1):220.

Bedford T, Rambaut A, Pascual M. Canalization of the evolutionary trajectory of the human influenza virus. BMC Biol. 2012; 10(1):38.

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, et al. The global distribution and burden of dengue. Nature. 2013; 496(7446):504.

Gentry M, Henchal E, McCown J, Brandt W, Dalrymple J. Identification of distinct antigenic determinants on dengue-2 virus using monoclonal antibodies. The American journal of tropical medicine and hygiene. 1982; 31(3):548–555.

Gibbons RV, Kalanarooj S, Jarman RG, Nisalak A, Vaughn DW, Endy TP, Mammen Jr MP, Srikiatkhachorn A. Analysis of repeat hospital admissions for dengue to estimate the frequency of third or fourth dengue infections resulting in admissions and dengue hemorrhagic fever, and serotype sequences. The American journal of tropical medicine and hygiene. 2007; 77(5):910–913.

Green AM, Beatty PR, Hadjilaou A, Harris E. Innate immunity to dengue virus infection and subversion of antiviral responses. Journal of molecular biology. 2014; 426(6):1148–1160.

Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA. Nextstrain: Real-time tracking of pathogen evolution. Bioinformatics. 2018; .

Halstead SB. In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody. Journal of Infectious Diseases. 1979; 140(4):527–533.

Juraska M, Magaret CA, Shao J, Carpp LN, Fiore-Gartland AJ, Benkeser D, Girerd-Chambaz Y, Langevin E, Frago C, Guy B, et al. Viral genetic diversity and protective efficacy of a tetravalent dengue vaccine in two phase 3 trials. Proceedings of the National Academy of Sciences. 2018; p. 201714250.

Katzelnick LC, Fonville JM, Gromowski GD, Arriaga JB, Green A, James SL, Lau L, Montoya M, Wang C, VanBlargan LA, et al. Dengue viruses cluster antigenically but not as discrete serotypes. Science. 2015; 349(6254):1338–1343.

Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, Balmaseda A, Harris E. Antibody-dependent enhancement of severe dengue disease in humans. Science. 2017; 358(6365):929–932.

Katzelnick LC, Montoya M, Gresh L, Balmaseda A, Harris E. Neutralizing antibody titers against dengue virus correlate with protection from symptomatic infection in a longitudinal cohort. Proceedings of the National Academy of Sciences. 2016; 113(3):728–733.

Kochel TJ, Watts DM, Halstead SB, Hayes CG, Espinoza A, Felices V, Caceda R, Bautista CT, Montoya Y, Douglas S, et al. Effect of dengue-1 antibodies on American dengue-2 viral infection and dengue haemorrhagic fever. The Lancet. 2002; 360(9329):310–312.

Kuiken C, Thurmond J, Dimitrijevic M, Yoon H. The LANL hemorrhagic fever virus database, a new platform for analyzing biothreat viruses. Nucleic acids research. 2011; 40(D1):D587–D592.

Lanciotti RS, Gubler DJ, Trent DW. Molecular evolution and phylogeny of dengue-4 viruses. Journal of General Virology. 1997; 78(9):2279–2284.

Lee JM, Huddleston J, Doud MB, Hooper KA, Wu NC, Bedford T, Bloom JD. Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. Proceedings of the National Academy of Sciences. 2018; .

Lipsitch M, O'Hagan JJ. Patterns of antigenic diversity and the mechanisms that maintain them. Journal of the Royal Society Interface. 2007; 4(16):787–802.

Lourenço J, Tennant W, Faria NR, Walker A, Gupta S, Recker M. Challenges in dengue research: A computational perspective. Evolutionary applications. 2018; 11(4):516–533.

Łuksza M, Lässig M. A predictive fitness model for influenza. Nature. 2014; 507(7490):57.

Mizumoto K, Ejima K, Yamamoto T, Nishiura H. On the risk of severe dengue during secondary infection: a systematic review coupled with mathematical modeling. Journal of vector borne diseases. 2014; 51(3):153.

Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. Proceedings of the National Academy of Sciences. 2016; 113(12):E1701–E1709.

OhAinle M, Balmaseda A, Macalalad AR, Tellez Y, Zody MC, Saborío S, Nuñez A, Lennon NJ, Birren BW, Gordon A, et al. Dynamics of dengue disease severity determined by the interplay between viral genetics and serotype-specific immunity. Science translational medicine. 2011; 3(114):114ra128–114ra128.

Olkowski S, Forshey BM, Morrison AC, Rocha C, Vilcarromero S, Halsey ES, Kochel TJ, Scott TW, Stoddard ST. Reduced risk of disease during postsecondary dengue virus infections. The Journal of infectious diseases. 2013; 208(6):1026–1033.

Pyke AT, Moore PR, Taylor CT, Hall-Mendelin S, Cameron JN, Hewitson GR, Pukallus DS, Huang B, Warrilow D, Van Den Hurk AF. Highly divergent dengue virus type 1 genotype sets a new distance record. Scientific reports. 2016; 6:22356.

Rico-Hesse R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. Virology. 1990; 174(2):479–493.

Russell PK, Nisalak A. Dengue virus identification by the plaque reduction neutralization test. The Journal of Immunology. 1967; 99(2):291–296.

Sabin AB. Research on dengue during World War II1. The American journal of tropical medicine and hygiene. 1952; 1(1):30–50.

Salje H, Lessler J, Berry IM, Melendrez MC, Endy T, Kalayanarooj S, Atchareeya A, Chanama S, Sangkijporn S, Klungthong C, et al. Dengue diversity across spatial and temporal scales: local structure and the effect of host population size. Science. 2017; 355(6331):1302–1306.

Sangkawibha N, Rojanasuphot S, Ahandrik S, Viriyapongse S, Jatanasen S, Salitul V, Phanthumachinda B, Halstead SB. Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong, Thailand. American journal of epidemiology. 1984; 120(5):653–669.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30(9):1312–1313.

Twiddy SS, Holmes EC, Rambaut A. Inferring the rate and time-scale of dengue virus evolution. Molecular biology and evolution. 2003; 20(1):122–129.

Zhang C, Mammen MP, Chinnawirotpisan P, Klungthong C, Rodpradit P, Monkongdee P, Nimmannitya S, Kalayanarooj S, Holmes EC. Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. Journal of virology. 2005; 79(24):15123–15130.

Supplement

Table S1. Fit parameter values for fitness model

Genetic resolution	Antigenic resolution	Metric	Metric value	β	γ	σ	DENV1 f_0	DENV2 f_0	DENV3 f_0	DENV4 f_0
Serotype	Interserotype	Δ SSE	15.02	2.57	0.57	0.86	4.57	3.43	2.14	0.00
Serotype	Interserotype	Pearson R^2	0.63	2.57	0.57	0.86	3.43	2.29	0.71	0.00
Genotype	Interserotype	Δ SSE	14.83	2.57	0.57	0.86	5.71	4.57	3.57	0.00
Genotype	Interserotype	Pearson R^2	0.36	2.57	0.57	0.86	5.71	5.71	2.86	0.00
Genotype	Full tree	Δ SSE	14.22	1.71	0.57	0.43	1.40	0.80	0.40	0.00
Genotype	Full tree	Pearson R^2	0.33	1.29	0.57	0.43	1.40	1.60	0.40	0.00

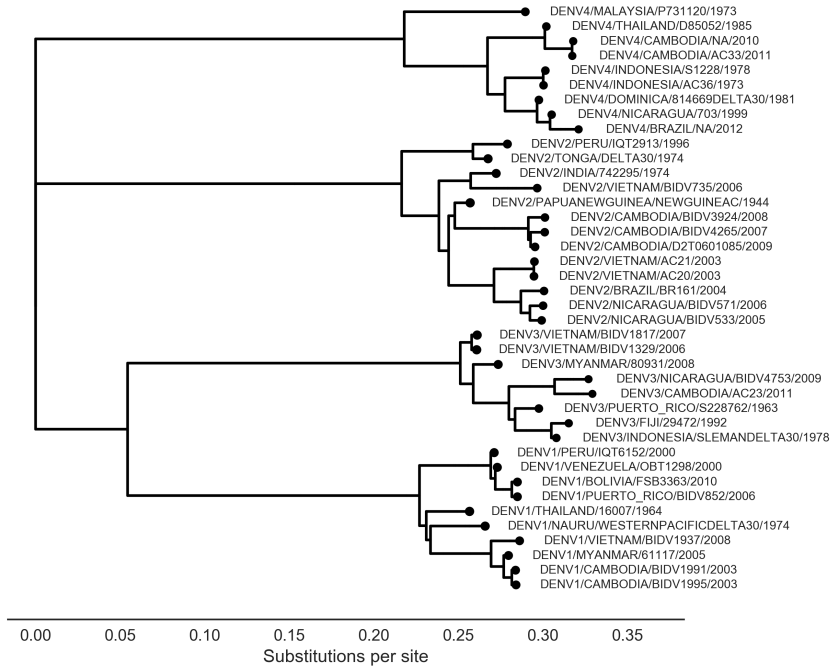

Figure S1. Tree of DENV viruses in titer dataset Maximum likelihood phylogeny of all DENV viruses that were included in the titer dataset.

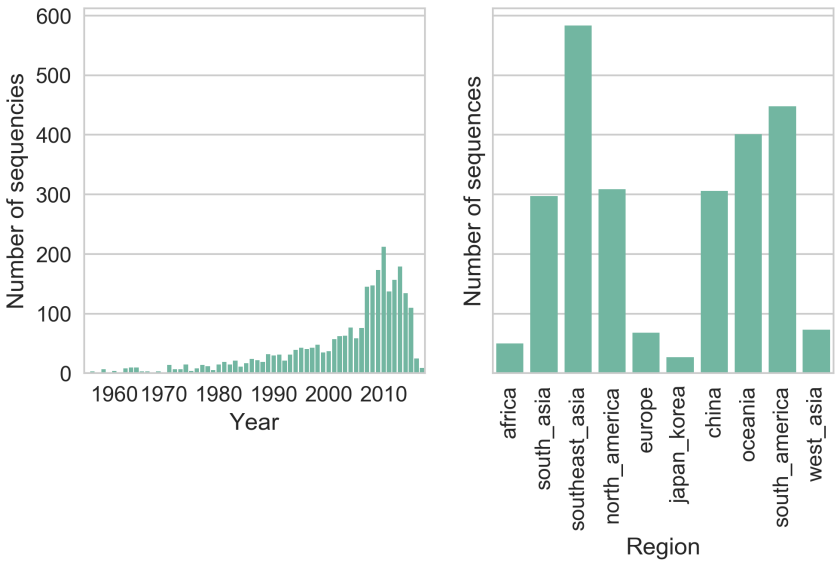

Figure S2. Sequence dataset distribution Temporal and geographic distribution of sequences included in our dataset after subsampling.

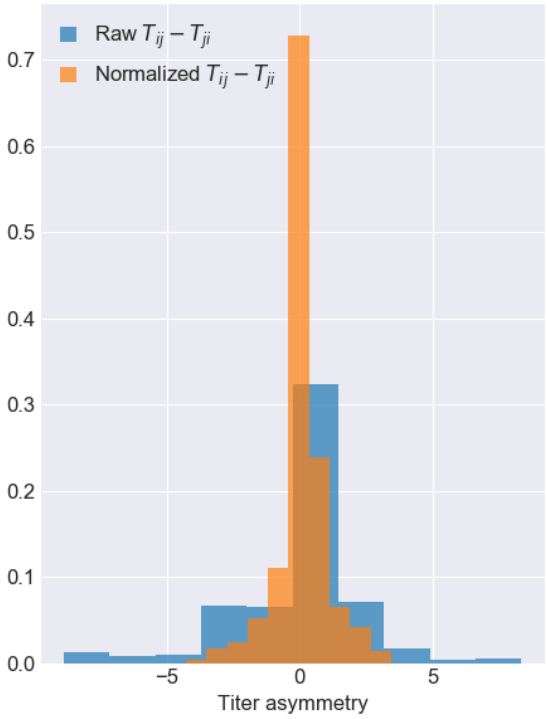

Figure S3. Titer value symmetry Some viruses have greater avidity overall, and some sera are more potent overall. We normalize for these row and column effects (v_a and p_b, respectively) in the titer model. Once overall virus avidity and serum potency are accounted for, titers are roughly symmetric (i.e., $D_{ij} \approx D_{ji}$).

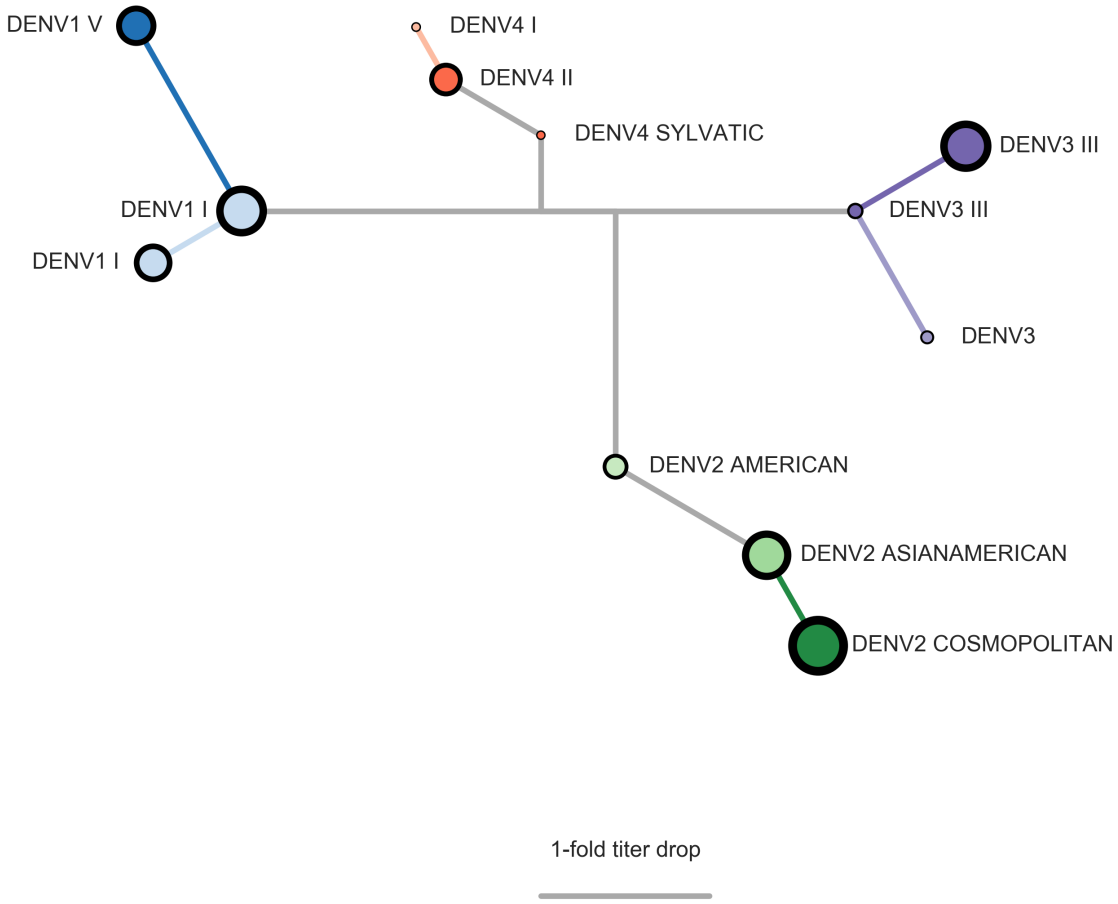

Figure S4. Tree of dengue antigenic phenotypes (alternate view) As in Figure 5, the topology is inferred from a maximum likelihood phylogeny of DENV sequences. Branch lengths are scaled to represent the d_b assigned to each branch by the 'full tree' model of antigenic evolution.

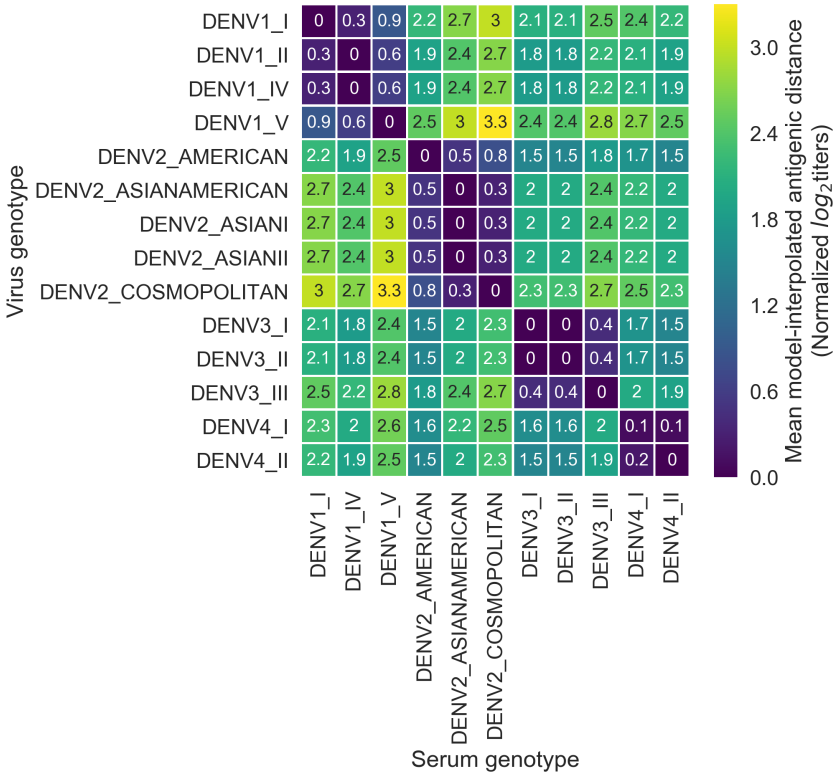

Figure S5. Titer distance by genotype Values represent the mean interpolated antigenic distance between canonical dengue genotypes (in standardized $\log_2$ titer units).