

1 Dengue genetic divergence 2 generates within-serotype antigenic 3 variation, but serotypes dominate 4 evolutionary dynamics

5 **Sidney Bell^{1,2*}, Leah Katzelnick^{3,4}, Trevor Bedford¹**

*For correspondence:
sidneymb@uw.edu (SB)

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

11

12 **Abstract** Dengue virus (DENV) exists as four genetically distinct serotypes, each of which is
13 historically assumed to be antigenically uniform. However, recent analyses suggest that antigenic
14 heterogeneity may exist within each serotype, but its source, extent and impact remain unclear.
15 Here, we construct a sequence-based model to directly map antigenic change to underlying genetic
16 divergence. We identify 49 specific substitutions and four colinear substitution clusters that
17 robustly predict dengue antigenic relationships. We report moderate antigenic diversity within each
18 serotype, resulting in variation in genotype-specific patterns of heterotypic cross-neutralization. We
19 also quantify the impact of antigenic variation on real-world DENV population dynamics, and find
20 that serotype-level antigenic fitness is a dominant driver of dengue clade turnover. These results
21 provide a more nuanced understanding of the relationship between dengue genetic and antigenic
22 evolution, and quantify the effect of antigenic fitness on dengue evolutionary dynamics.

23

24 Author Summary

25 Dengue virus (DENV), the causative agent of dengue hemorrhagic fever, exists as four genetically
26 distinct serotypes, DENV1 to DENV4. These serotypes are antigenically distinct: symptomatic
27 reinfection with a homotypic virus is very rare, while reinfection with a heterotypic virus is some-
28 times associated with severe disease. Until recently, it has been assumed that viruses within each
29 serotype are antigenically uniform. However, specific genotypes within each serotype have been
30 anecdotally associated with varying severity of patient outcomes and epidemic magnitude. One
31 hypothesis is that each serotype contains overlooked, meaningful antigenic diversity. While anti-
32 genic cartography conducted on neutralization titers suggests that heterogeneity may exist within
33 each serotype, its source, extent and impact is unclear. Here, we analyze a previously published
34 titer dataset to quantify and characterize the extent of DENV intraserotype antigenic diversity. We
35 map antigenic changes to specific mutations in *E*, the dengue envelope protein, and interpolate
36 across the alignment to estimate the antigenic distance between pairs of viruses based on their
37 genetic differences. We identify 49 specific substitutions and four colinear substitution clusters
38 that contribute to dengue antigenic evolution. We find that DENV antigenic divergence is tightly
39 coupled to DENV genetic divergence, and is likely a gradual, ongoing process. We report modest
40 but significant antigenic diversity within each serotype of DENV, which may have important rami-
41 fications for vaccine design. To understand the impact of antigenic variation on real-world DENV
42 population dynamics, we also quantify the extent to which population immunity—accumulated
43 through recent circulation of antigenically similar genotypes—determines the success and decline
44 of DENV clades in a hyperendemic population. We find that serotype-level antigenic fitness is a
45 key determinant of DENV population turnover. By leveraging both molecular data and real-world
46 population dynamics, these results provide a more nuanced understanding of the relationship
47 between dengue genetic and antigenic evolution, and quantify the effect of antigenic fitness on
48 dengue evolutionary dynamics.

49 **Introduction**

50 Dengue virus (DENV) is a mosquito-borne flavivirus which consists of four genetically distinct clades,
 51 canonically thought of as serotypes (DENV1 – DENV4) (*Lanciotti et al., 1997*). DENV circulates
 52 primarily in South America and Southeast Asia, infecting 400 million people annually. Primary
 53 DENV infection is more often mild and is thought to generate lifelong homotypic immunity and
 54 temporary heterotypic immunity, which typically wanes over six months to two years (*Sabin, 1952*;
 55 *Reich et al., 2013; Katzelnick et al., 2016*). Subsequent heterotypic secondary infection induces
 56 broad cross-protection, and symptomatic tertiary and quaternary cases are rare (*Gibbons et al.,*
 57 *2007; Olkowski et al., 2013*). However, a small subset of secondary infections are enhanced by
 58 non-neutralizing, cross-reactive antibodies, resulting in severe disease via antibody dependent
 59 enhancement (ADE) (*Halstead, 1979; Katzelnick et al., 2017; Sangkawibha et al., 1984; Salje et al.,*
 60 *2018*). Approximately 1–3% of cases progress to severe dengue hemorrhagic fever, causing ~9,000
 61 deaths each year (*Bhatt et al., 2013; Stanaway et al., 2016*) and relative risk of severe dengue from
 62 secondary heterotypic infection relative to primary infection is estimated to be ~24 (*Mizumoto*
 63 *et al., 2014*). Thus, the antigenic relationships between dengue viruses — describing whether
 64 the immune response generated after primary infection results in protection or enhancement of
 65 secondary infection — are key drivers of DENV case outcomes and epidemic patterns.

66 While each serotype is clearly genetically and antigenically distinct, it is not clear how sub-
 67 serotype clades of DENV interact antigenically. Each DENV serotype consists of broad genetic
 68 diversity (Figure 1A), including canonical clades termed ‘genotypes’ (*Rico-Hesse, 1990; Twiddy et al.,*
 69 *2003*). Specific genotypes have been associated with characteristically mild or severe disease, and
 70 heterogeneous neutralization titers suggest that the immune response to some genotypes is more
 71 cross-protective than others (*Gentry et al., 1982; Russell and Nisalak, 1967*). Until recently, it has
 72 been assumed that these intraserotype differences are minimally important compared to inter-
 73 serotype differences. However, empirical evidence has demonstrated that these genotype-specific
 74 differences can drive case outcomes and epidemic severity (reviewed in *Holmes and Twiddy (2003)*).
 75 For example, analysis of a longitudinal cohort study demonstrated that specific combinations of
 76 primary infection serotype and secondary infection genotype can mediate individual case outcomes
 77 (*OhAinle et al., 2011*). On a population scale, the DENV1-immune population of Iquitos, Peru,
 78 experienced either entirely asymptomatic or very severe epidemic seasons in response to two
 79 different genotypes of DENV2 (*Kochel et al., 2002*).

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

87 **Results**

88 **Dengue neutralization titer data**

89 Antigenic distance between a pair of viruses i and j is experimentally quantified using neutralization
 90 titers, which measure how well serum drawn after infection with virus j is able to neutralize virus
 91 i *in vitro* (*Russell and Nisalak, 1967*). Throughout the following we refer to serum raised against
 92 virus j as serum j for brevity. To measure the pairwise antigenic distances for a panel of diverse
 93 DENV viruses (Figure 1), Katzelnick *et al.* infected naïve non-human primates (NHP) with each virus,
 94 drew sera at three months post-infection, and then titrated this sera against a panel of test viruses
 95 (*Katzelnick et al., 2015*). To compare patterns of cross-protection in NHP and humans, they also
 96 drew sera from 31 study participants six weeks after inoculation with a monovalent component of
 97 the NIH dengue vaccine candidate. This sera was also titrated against a broad panel of DENV viruses.

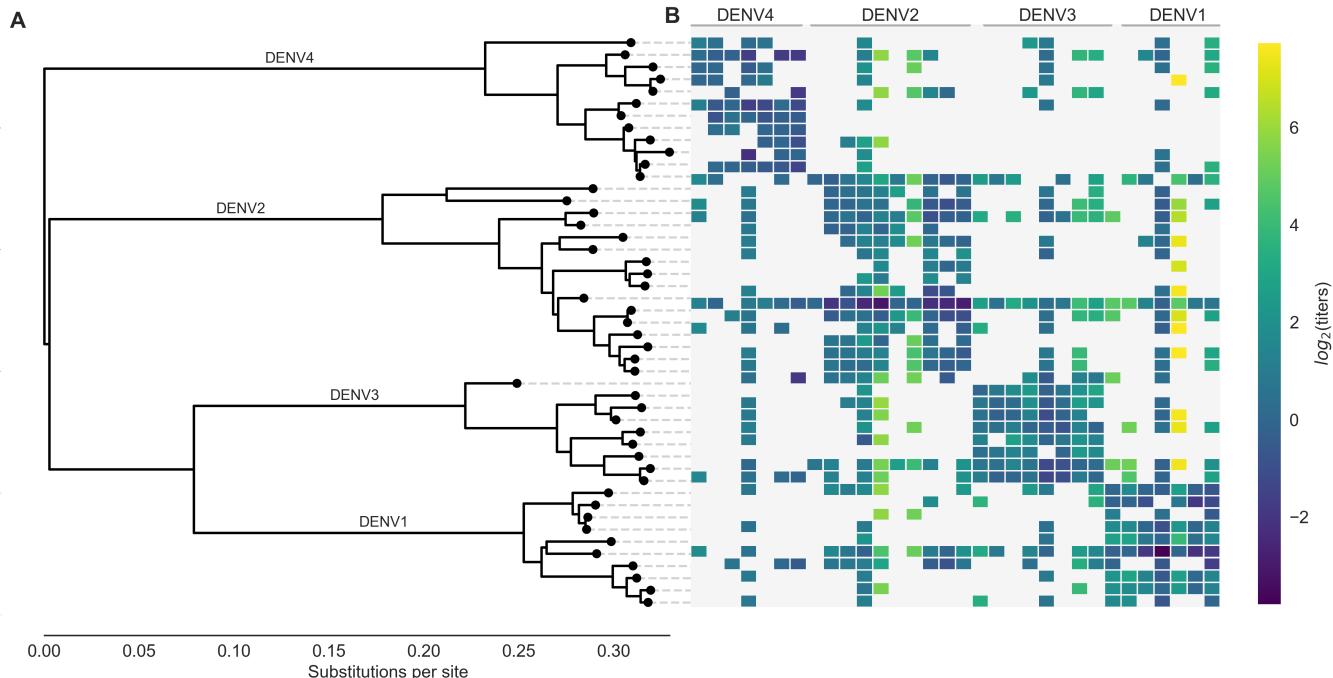


Figure 1. Phylogeny of dengue virus sequences and normalized antigenic distances. (A) Maximum likelihood phylogeny of the *E* (envelope) gene from titered dengue viruses. Notably, each of the four serotypes contains substantial genetic diversity. (B) Pairwise antigenic distances were estimated by Katzelnick *et al.* using plaque reduction neutralization titers (PRNT50, see Methods). Aggregated titer values are standardized such that the distance between autologous virus-serum pairs is 0, and each titer unit corresponds to a two-fold change in PRNT50 value. Light gray areas represent missing data. Larger values correspond to greater antigenic distance.

Figure 1–Figure supplement 1. 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., $D_{ij} \approx D_{ji}$).

98 As originally reported, we find generally consistent patterns of neutralization between the NHP
 99 and human sera data; see **Katzelnick et al. (2015)** for a detailed comparison. In total, our dataset
 100 consists of 454 NHP sera titers spanning the breadth of DENV diversity, and 728 human sera titers
 101 providing deep coverage of a small subset of viruses.

102 To normalize these measurements, we take the \log_2 of each value, such that one antigenic
 103 unit corresponds to a two-fold drop in neutralization, and we define antigenic distance between
 104 autologous serum-virus pairs (i.e., virus i and serum i) as zero. Normalized antigenic distance
 105 between virus i and serum j is thus calculated as $D_{ij} = \log_2(T_{ii}) - \log_2(T_{ij})$, such that a higher value of
 106 D_{ij} indicates that serum j is less effective at neutralizing virus i , implying greater antigenic distance
 107 between viruses i and j . For brevity, these normalized titer values are hereafter referred to simply
 108 as $\log_2(\text{titers})$.

109 The full dataset of standardized titer values is shown in Figure 1B. Here, we see that homotypic
 110 virus-serum pairs are more closely related antigenically than heterotypic pairs. However, we also
 111 observe large variance around this trend, both within and between serotypes. This suggests
 112 that treating each serotype as antigenically uniform potentially overlooks important antigenic
 113 heterogeneity across viruses within each serotype.

114 **Dengue antigenic evolution corresponds to genetic divergence**

115 Titer measurements are prone to noise, and there is a limited amount of available titer data. If the
 116 antigenic heterogeneity observed in the raw data is truly the result of an underlying evolutionary
 117 process, we expect that differences in antigenic phenotype correspond to underlying mutations
 118 in surface proteins. Dengue has two surface proteins, prM (membrane) and E (envelope). While
 119 previous studies have identified epitopes on both prM and E, it is believed that antibodies involved
 120 in ADE primarily target prM, while neutralizing antibodies primarily target E (**de Alwis et al., 2014**).
 121 The assay used to generate this titer dataset captures neutralization, but does not capture the
 122 effects of ADE; we thus focus our analysis on the E gene.

123 To fully map the relationship between DENV genetic and antigenic evolution, we adapt a
 124 substitution-based model originally developed for influenza (**Neher et al., 2016**). Conceptually,
 125 this model predicts titer values through three steps. First, we align E gene sequences from titered
 126 dengue viruses and catalog the amino acid mutations between each serum strain and test virus
 127 strain in our dataset. Next, we infer how much antigenic change is attributable to specific mutations
 128 by constructing a parsimonious model that links normalized antigenic distances to observed muta-
 129 tions. This assigns each mutation m an antigenic effect size, $d_m \geq 0$; forward and reverse mutations
 130 are assigned separate values of d_m . With this in hand, we estimate the asymmetrical antigenic dis-
 131 tance \hat{D}_{ij} between all pairs of sera and test viruses by summing over d_m for all mutations observed
 132 between the serum and the test virus (Methods, Eq. 2).

133 To learn these values of d_m , we first split our dataset into training (random 90% of measurements)
 134 and test data (the remaining 10% of values). We take the training data and fit d_m for each mutation
 135 that is observed two or more times, subject to regularization as follows (also detailed in Methods,
 136 Eq. 3). Parsimoniously, we expect that antigenic change is more likely to be incurred by a few
 137 key mutations than by many mutations; correspondingly, our prior expectation of values of d_m
 138 is exponentially distributed such that most values of $d_m = 0$. This is directly analogous to lasso
 139 regression to identify a few parameters with positive weights and set other parameters to 0
 140 (**Tibshirani, 1996**). Additionally, some viruses have greater binding avidity, and some sera are more
 141 potent than others (Figure 1—Figure Supplement 1); these ‘row’ and ‘column’ effects, respectively,
 142 are normally distributed and are taken into account when training the model. The model uses
 143 convex optimization to learn the values of d_m that minimize the sum of squared errors (SSE)
 144 between observed and predicted titers in the training data. We thus learn model parameters from
 145 the training data, and then use those parameters to predict test data values. We assess model
 146 performance by comparing the predicted test titer values to the actual values, aggregated across
 147 100-fold Monte Carlo cross validation.

Table 1. Antigenically relevant mutations. Each entry represents a mutation (or colinear cluster of mutations) inferred by the titer model to have a non-zero antigenic effect size d_m (shown in parentheses).

- I6V, S29G, F90Y, T176P, V197I, L475M (0.71)
- A19T (0.02)
- N83K (0.34)
- A88K (0.08)
- A88Q (1.39)
- Y90F (0.43)
- V91I (0.35)
- K93R (0.20)
- V114I (0.90)
- L122S (0.03)
- S122L (0.07)
- N124K (0.10)
- V129I (0.01)
- I129V (0.21)
- I129A (0.23)
- Y132I (0.12)
- Y132P, R233Q (0.34)
- I139V (0.05)
- D154E (0.03)
- K160V (0.03)
- E161T (0.22)
- A162I (0.19)
- I162A (0.29)
- I164V (0.01)
- T171S (0.23)
- V174E (0.46)
- E180T (0.43)
- N203E (0.04)
- N203K (0.08)
- D203N (0.14)
- E203N (1.09)
- E203D (1.34)
- I308V (0.18)
- G330D (0.22)
- I335V, N355T, P364V (0.18)
- L338E (0.43)
- T339I (0.65)
- V347A (0.28)
- I380V (0.45)
- V382A (0.04)
- V382I (0.37)
- N385K, V454I (0.12)
- D390S (0.12)
- N390S (0.18)
- V454T (0.25)
- V461F (0.01)
- F461I (0.03)
- I462V (0.10)
- L462I (0.16)
- V462L (2.11)
- T478S (0.10)
- S478M (0.43)
- V484I (0.39)

148 This model formulation is an effective tool for estimating antigenic relationships between
 149 viruses based on their genetic sequences. On average across cross-validation replicates, this
 150 model yields a root mean squared error (RMSE) of 0.75 when predicting titers relative to their true
 151 value (95% CI 0.74–0.77, RMSE), and explains 78% of the observed variation in neutralization titers
 152 overall (95% CI 0.77–0.79, Pearson R^2). This is comparable to the model error from a cartography-
 153 based characterization of the same dataset (RMSE 0.65–0.8 log₂ titer units) (**Katzelnick et al., 2015**).
 154 Prediction error was comparable between human and non-human primate sera, indicating that
 155 these genetic determinants of antigenic phenotypes are not host species-specific (Figure 2—Figure
 156 Supplement 1).

157 The 48 strains included in the titer dataset (as serum strains, test virus strains, or both) are
 158 25.7% divergent on average (amino acid differences in *E*). Pairwise comparisons of all serum strains
 159 and test viruses yields 1,534 unique mutations that are observed at least twice. Our parsimonious
 160 model attributes antigenic change to a total of 49 specific mutations and 4 colinear mutation
 161 clusters (each consisting of 2–6 co-occurring mutations) (Figure 2, Table 1). Each of these mutations
 162 confers 0.01–2.11 (median 0.19) log₂ titer units of antigenic change; 27 mutations or mutation
 163 clusters have $d_m \geq 0.2$. These mutations span all domains of *E*, and most occur both between and
 164 within serotypes (Figure 2).

165 **Each serotype of dengue contains moderate antigenic heterogeneity**

166 By linking antigenic change to specific mutations, we are able to estimate unmeasured antigenic dis-
 167 tances between any pair of viruses in the dataset based on their genetic differences. As an example,
 168 we estimated the antigenic distance between serum raised against each monovalent component
 169 of the NIH vaccine candidate and all other viruses in the dataset. As shown in Figure 3, vaccine-
 170 elicited antibodies result in strong homotypic neutralization, but heterotypic cross-neutralization
 171 varies widely between specific strains. This has important ramifications for vaccine design and trial
 172 evaluation.

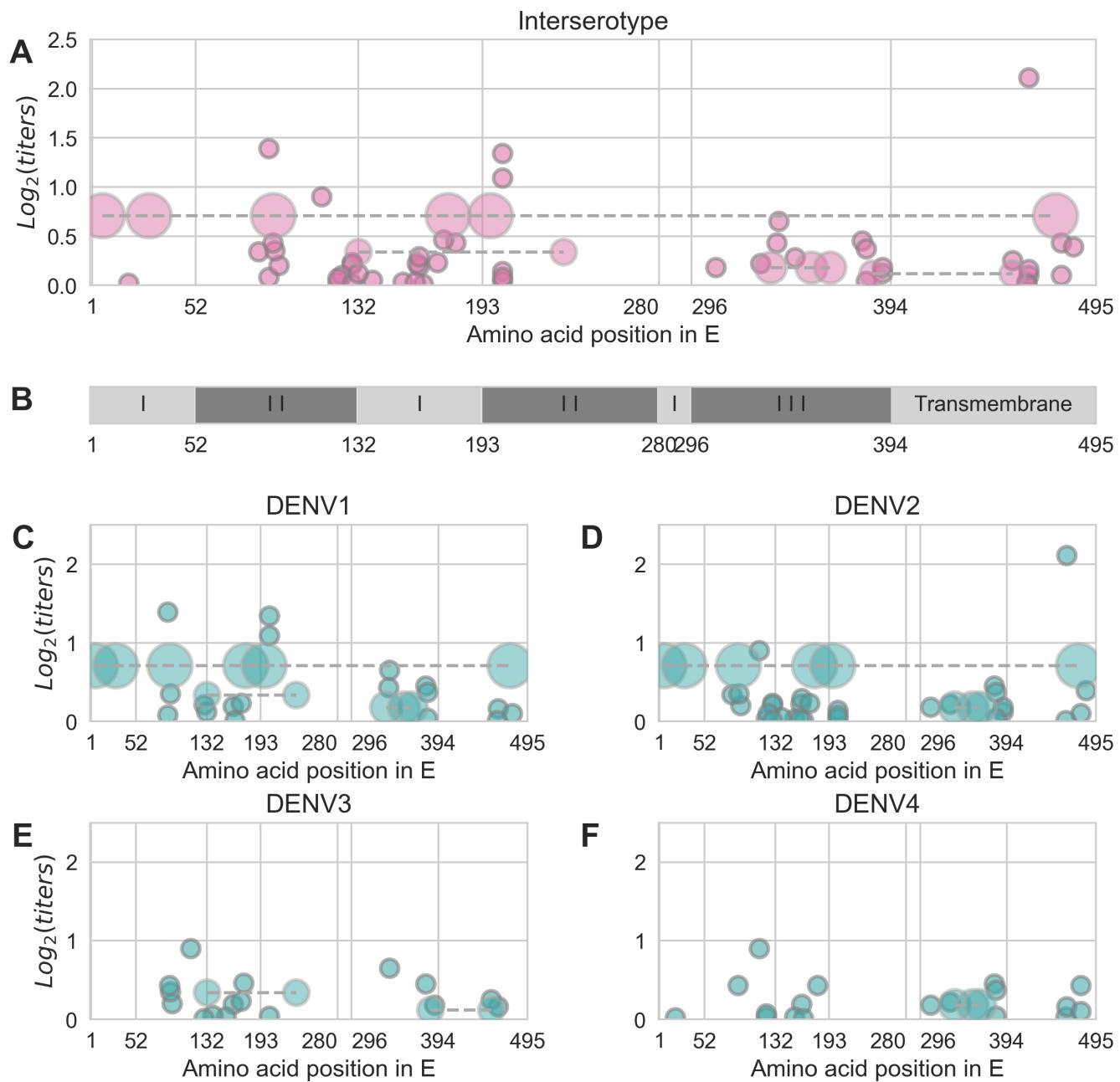


Figure 2. Distribution and effect size of antigenic mutations. Each point represents one antigenically relevant mutation or colinear mutation cluster. Clustered mutations are connected with dashed lines with point size proportionate to cluster size ($N=2-6$). The x axis indicates mutations' position in *E*, relative to each functional domain as noted in (B). The y axis indicates antigenic effect size.

Figure 2—Figure supplement 1. Titer prediction error by serum strain and species. Human sera was raised against four different virus strains (the monovalent vaccine components); non-human primate (NHP) sera was raised against many different virus strains. Here, we excluded NHP sera raised against the monovalent vaccine components, such that each normalized titer measurement is aggregated across individuals, but not across species. We report the out-of-sample titer prediction error for each serum strain (versus all available test viruses), aggregated across 100-fold Monte Carlo cross-validation.

Figure 2—Figure supplement 2. Genotype as site E 390 across dengue phylogeny. Dengue virus genotypes can be seen on Nextstrain ([Hadfield et al., 2018](#)). A live view of this figure is available at [nextstrain.org/dengue/](#).

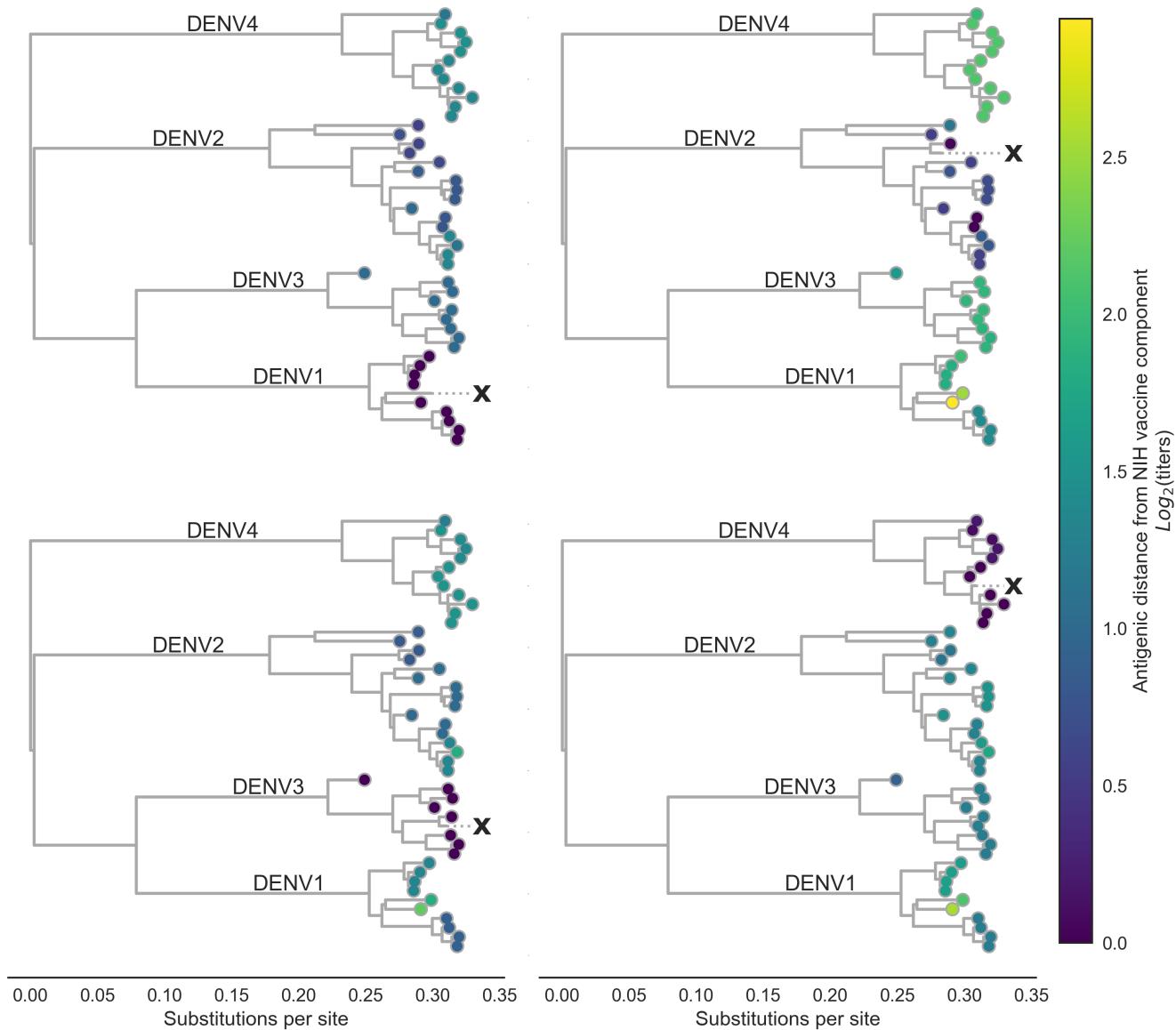


Figure 3. Antigenic distance from NIH vaccine strains. By assigning a discrete increment of antigenic change to each mutation, we can estimate the asymmetrical antigenic distance between any serum strain and test virus strain based on their genetic differences. Here, we show the estimated antigenic distance between serum raised against each monovalent component of the NIH vaccine candidate (indicated as 'X') and each test virus in the tree.

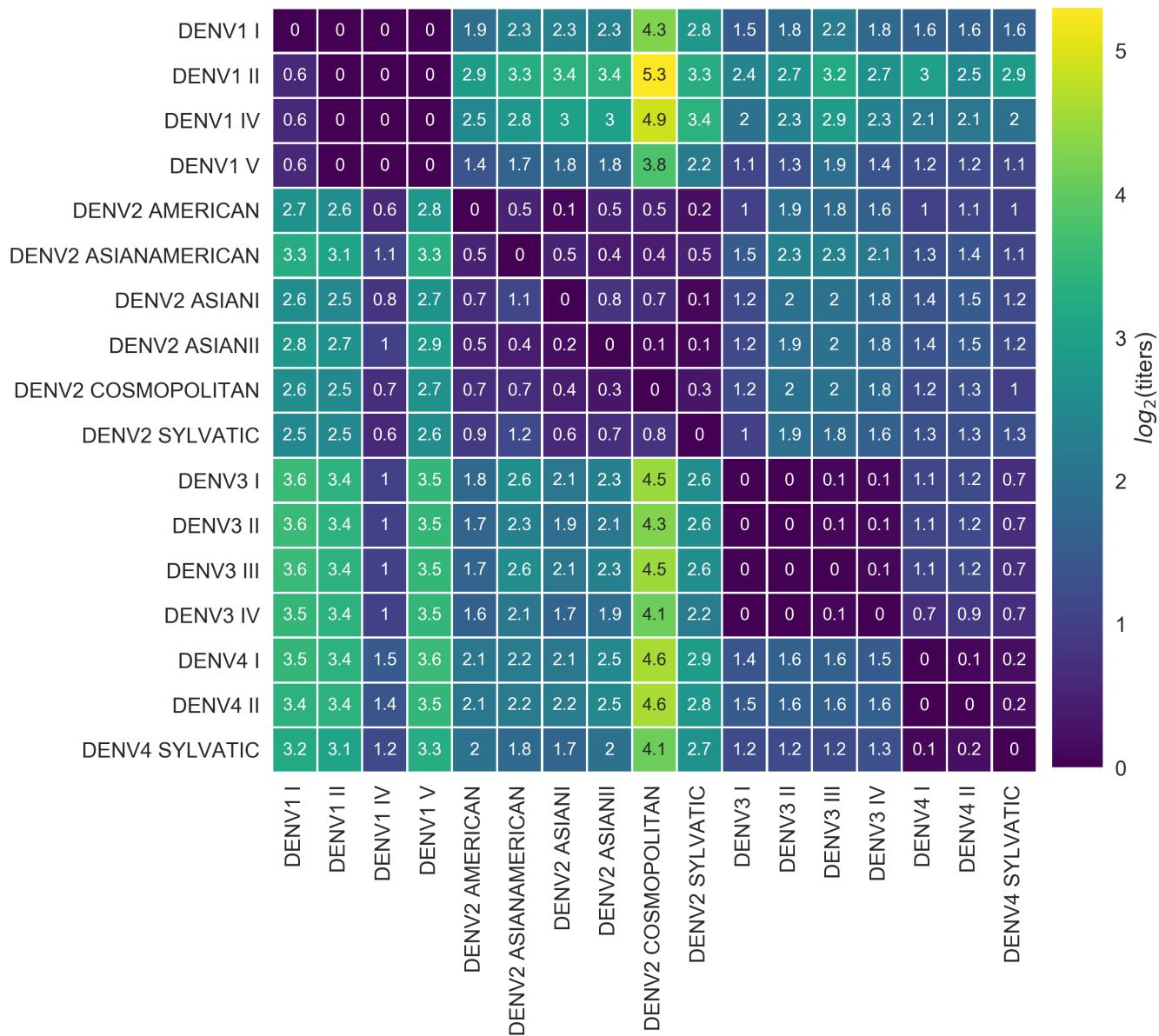


Figure 4. Titer distance by genotype. Values represent the mean interpolated antigenic distance between canonical dengue genotypes (in standardized \log_2 titer units).

173 We also observe antigenic heterogeneity at the genotype level. On average, heterotypic geno-
 174 types are separated by 6.9 antigenic mutations (or colinear mutation clusters) and $2.18 \log_2$ titers.
 175 Homotypic genotypes are separated by a mean of 1.9 antigenic mutations, conferring a total of 0.30
 176 \log_2 titers of antigenic distance (Figure 4). Notably, the titer dataset spans the breadth of canonical
 177 DENV genotypes, but in most cases lacks the resolution to detect within-genotype antigenic diversity.
 178 We thus expect that these results represent a lower-bound on the true extent of DENV intraserotype
 179 antigenic diversity.

180 In summary, we have identified a small number of antigenically relevant mutations that explain
 181 most of the observed antigenic heterogeneity in dengue, as indicated by neutralization titers. These
 182 mutations occur both between and within serotypes, suggesting that dengue antigenic evolution is
 183 an ongoing, though gradual, process. This results in strain-specific and genotype-level antigenic
 184 variation, although the scale of this variation is small compared to serotype-level differences. From
 185 this, we conclude that there is antigenic variation within each serotype of DENV, and that this is
 186 driven by underlying genetic divergence.

187 Antigenic novelty predicts serotype success

188 From the titer model, we find evidence that homotypic genotypes of DENV vary in their ability
 189 to escape antibody neutralization. However, antibody neutralization is only one of many factors
 190 that shape epidemic patterns. We investigate whether the observed antigenic diversity influences
 191 dengue population dynamics in the real world.

192 The size of the viral population (i.e., prevalence, commonly analyzed using SIR models as
 193 reviewed in *Lourenço et al. (2018)*) is determined by many complex factors, and reliable values for
 194 population prevalence are largely unavailable. Contrastingly, the composition of the viral population
 195 (i.e., the relative frequency of each viral clade currently circulating) can be estimated over time by
 196 examining historical sequence data (*Lee et al., 2018; Neher et al., 2016*), and is primarily driven by
 197 viral fitness (*Bedford et al., 2011*).

198 In meaningfully antigenically diverse viral populations, antigenic novelty (relative to standing
 199 population immunity) contributes to viral fitness: as a given virus i circulates in a population, the
 200 proportion of the population that is susceptible to infection with i -and other viruses antigenically
 201 similar to i -decreases over time as more people acquire immunity (*Bedford et al., 2012; Łuksza
 202 and Lässig, 2014*). Antigenically novel viruses that are able to escape this population immunity are
 203 better able to infect hosts and sustain transmission chains, making them fitter than the previously
 204 circulating viruses (*Zhang et al., 2005; Bedford et al., 2012; Gupta et al., 1998; Wearing and Rohani,
 205 2006; Lourenço and Recker, 2013*). Thus, if antigenic novelty constitutes a fitness advantage for
 206 DENV, then we would expect greater antigenic distance from recently circulating viruses to correlate
 207 with higher growth rates.

208 To test this hypothesis, we examine the composition of the dengue virus population in Southeast
 209 Asia from 1970 to 2015. We estimate the relative population frequency of each DENV serotype at
 210 three month intervals, $x_i(t)$ (Figure 5A), based on their observed relative abundance in the ‘slice’ of
 211 the phylogeny corresponding to each timepoint ($N=8,644$ viruses; see Methods, Eq. 4). While there
 212 is insufficient data to directly compare these estimated frequencies to regional case counts, we see
 213 good qualitative concordance between frequencies similarly estimated for Thailand and previously
 214 reported case counts from Bangkok (Figure 5—Figure Supplement 1).

215 Fitter virus clades increase in frequency over time, such that $x_i(t+dt) > x_i(t)$. It follows that these
 216 clades have a growth rate—defined as the fold-change in frequency over time—greater than one:
 217 $\frac{x_i(t+dt)}{x_i(t)} > 1$. To isolate the extent to which antigenic fitness contributes to clade success and decline,
 218 we extend work by *Łuksza and Lässig (2014)* to build a simple model that attempts to predict clade
 219 growth rates based on two variables: the antigenic fitness of the clade at time t , and a time-invariant
 220 free parameter representing the intrinsic fitness of the serotype the clade belongs to. We estimate
 221 the antigenic fitness of clade i at time t as a function of its antigenic distance from each viral clade j
 222 that has circulated in the same population over the previous two years, weighted by the relative

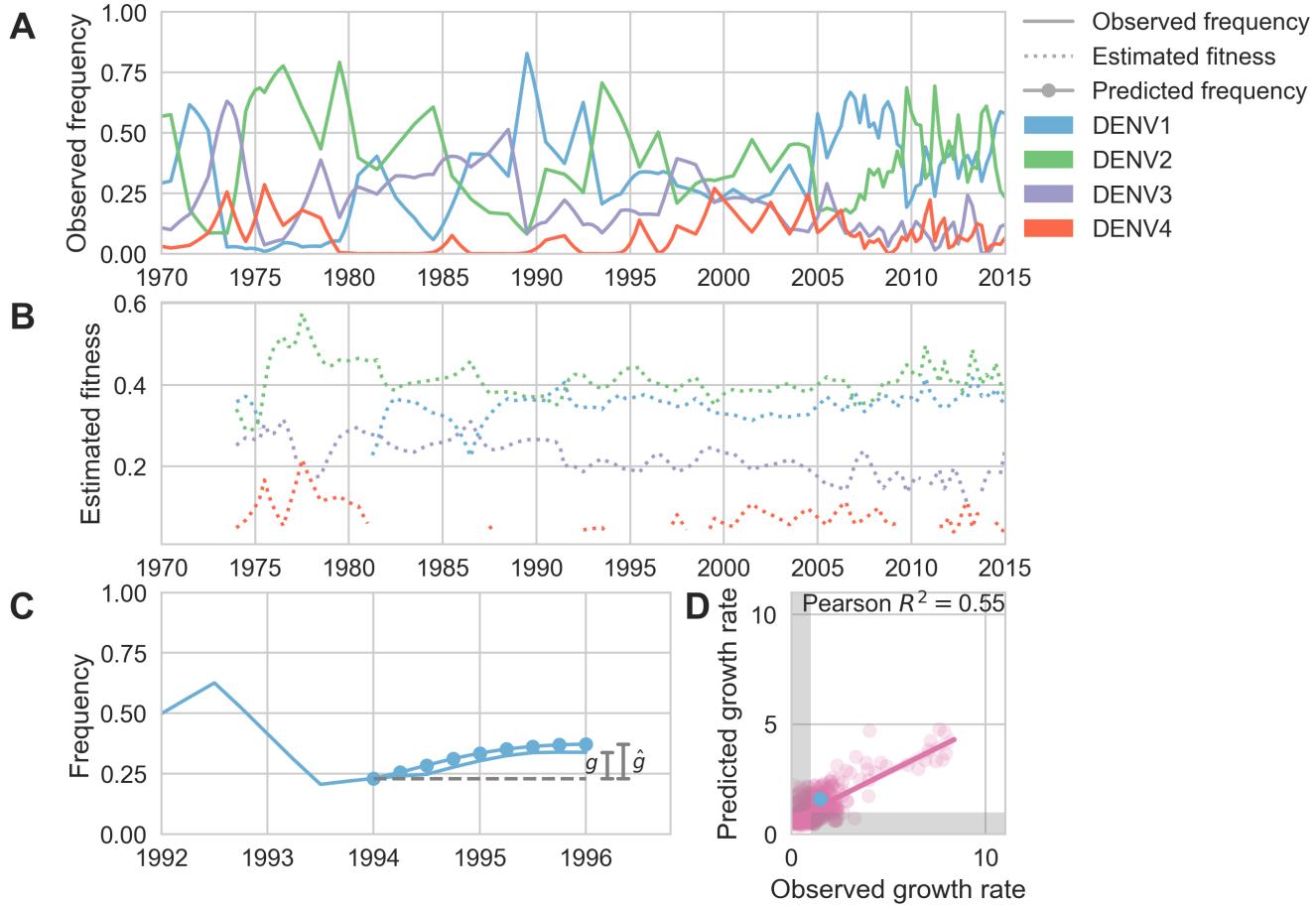


Figure 5. Antigenic novelty predicts serotype success. (A) The relative frequency of each serotype, x_i , in Southeast Asia estimated every three months based on available sequence data. (B) Total fitness of each serotype. We calculate antigenic fitness for each serotype over time as its frequency-weighted antigenic distance from recently circulating viruses. We then add this to a time-invariant intrinsic fitness value to calculate total fitness. (C) DENV1 frequencies between 1994 and 1996 alongside model projection. 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. 11). We predict forward in three-month increments for a total prediction period of $dt = 2$ years. At each increment, we use the current predicted frequency to adjust our estimates of antigenic fitness on a rolling basis (Methods, Eq. 15). (D) Predicted growth rates, $\hat{g} = \frac{\dot{x}_i(t+dt)}{x_i(t)}$, compared to empirically observed growth rates, $g = \frac{x_i(t+dt)}{x_i(t)}$. Predicted and empirical growth rate of the example illustrated in (C) is 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 66% of predictions.

Figure 5—Figure supplement 1. Case counts versus clade frequencies in Thailand. As described in the Methods, we estimate clade frequencies based on observed relative abundance in the ‘slice’ of the phylogeny at each quarterly timepoint. These frequency estimates are smoothed using a discretized Brownian motion diffusion process. Here, we compare estimated serotype frequencies across Thailand (all available high quality sequences) to case counts from a hospital in Bangkok between 1975–2010 (Reich et al., 2013). Biweekly case counts were aggregated into quarterly timepoints, but were not smoothed. While there are some instances where case counts and frequencies diverge (e.g., DENV4 in the early 1990s), the noisy nature of the unsmoothed case counts artificially deflates estimates of concordance.

Figure 5—Figure supplement 2. Simulated serotype frequencies. As described in the Methods, we seeded a simulation with two years of empirical frequencies and predicted forward to simulate the remainder of the timecourse. Here, we simulated under the model parameters described in Table 5.

Figure 5—Figure supplement 3. Simulated serotype frequencies (model parameters). As described in the Methods, we seeded a simulation with two years of empirical frequencies and predicted forward to simulate the remainder of the timecourse. Here, we simulated under the model parameters described in Table 4. This results in damped oscillations around the intrinsic fitness value for each serotype, but these intrinsic fitnesses alone are unable to predict observed clade dynamics (Table 2).

Table 2. Fitness model performance comparisons. Here we compare the performance of the antigenically-informed fitness models to model performance under two null formulations. In the ‘equal’ null model, all clades are assigned equal fitness (i.e., antigenic and intrinsic fitness are set to 0). In the ‘intrinsic’ null model formulation, only the serotype-specific, time-invariant intrinsic fitness values contribute to clade fitness (i.e., antigenic fitness is set to 0). For both formulations of generalized waning, all other parameters were set to the values reported in Table 4 (optimized for RMSE). Parameters for heterotypic waning were optimized separately.

Resolution	Fitness model	Waning	RMSE	Pearson R^2	Accuracy
Serotype	Interserotype	Generalized	0.105	0.547	0.660
Serotype	Equal fitness null	Generalized	0.130	0.000	0.480
Serotype	Intrinsic fitness null	Generalized	0.140	0.042	0.510
Genotype	Interserotype	Generalized	0.062	0.286	0.666
Genotype	Intergenotype	Generalized	0.062	0.254	0.610
Genotype	Equal fitness null	Generalized	0.070	0.000	0.440
Genotype	Intrinsic fitness null	Generalized	0.072	0.032	0.530
Serotype	Interserotype	Heterotypic	0.109	0.533	0.666
Genotype	Interserotype	Heterotypic	0.063	0.291	0.661
Genotype	Intergenotype	Heterotypic	0.063	0.203	0.599

frequency of j and adjusted for waning population immunity (Figure 5B; Methods, Eq. 7). Growth rates are estimated based on a two year sliding window (Figure 5C).

This simple model explains 54.7% of the observed variation in serotype growth rates, and predicts serotype growth vs. decline correctly for 66.0% of predictions (Figure 5D). This 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 six free parameters that are optimized such that the model most accurately reproduces the observed fluctuations in DENV population composition (minimizing the RMSE of frequency predictions, see Methods). We find that serotype fluctuations are consistent with a model wherein population immunity wanes linearly over time, with the probability of protection dropping by about 63% per year for the first two years after primary infection. This model assumes no fundamental difference between homotypic and heterotypic reinfection; rather, homotypic immunity is assumed to wane at the same rate as heterotypic immunity, but starts from a higher baseline of protection based on closer antigenic distances. We also find that these dynamics are best explained by intrinsic fitness that moderately varies by serotype (Table 4); we are not aware of any literature that directly addresses this observation via competition experiments. However, intrinsic fitness alone is unable to predict serotype dynamics (Table 2) and relative strength of antigenic fitness and intrinsic fitness are approximately matched in determining overall serotype fitness.

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 6B). For genotypes, we can calculate antigenic distance between i and j at either the serotype level or the genotype level. In the ‘interserotype model’, we treat each serotype as antigenically uniform, and assign the mean serotype-level antigenic distances to all pairs of constituent genotypes. In the ‘intergenotype model’, we incorporate the observed within-serotype heterogeneity, and use the mean genotype-level antigenic distances (as shown in Figure 4). If within-serotype antigenic heterogeneity contributes to genotype fitness, then we would expect estimates of antigenic fitness based on the ‘intergenotype model’ to better predict genotype

254 growth rates.

255 We find that antigenic fitness contributes to genotype turnover, although it explains less of
256 the observed variation than for serotypes. As for serotypes, intrinsic fitness alone was unable to
257 predict genotype turnover (Table 2). When antigenic distance is estimated from the ‘interserotype
258 model’, we find that our model of antigenic fitness explains approximately 28.6% of the observed
259 variation in genotype growth rates, and correctly predicts genotype growth vs. decline 66.6% of
260 the time (Figure 6C). Perhaps surprisingly, more precise estimates of antigenic distance between
261 genotypes from the ‘intergenotype model’ does not improve our predictions of genotype success
262 ($R^2 = 0.254$, 61.0% accuracy; Figure 6D, Table 2). This suggests that although we find strong evidence
263 that genotypes vary in their ability to escape neutralizing antibodies, these differences are subtle
264 enough that they do not impact broad-scale regional dynamics over time.

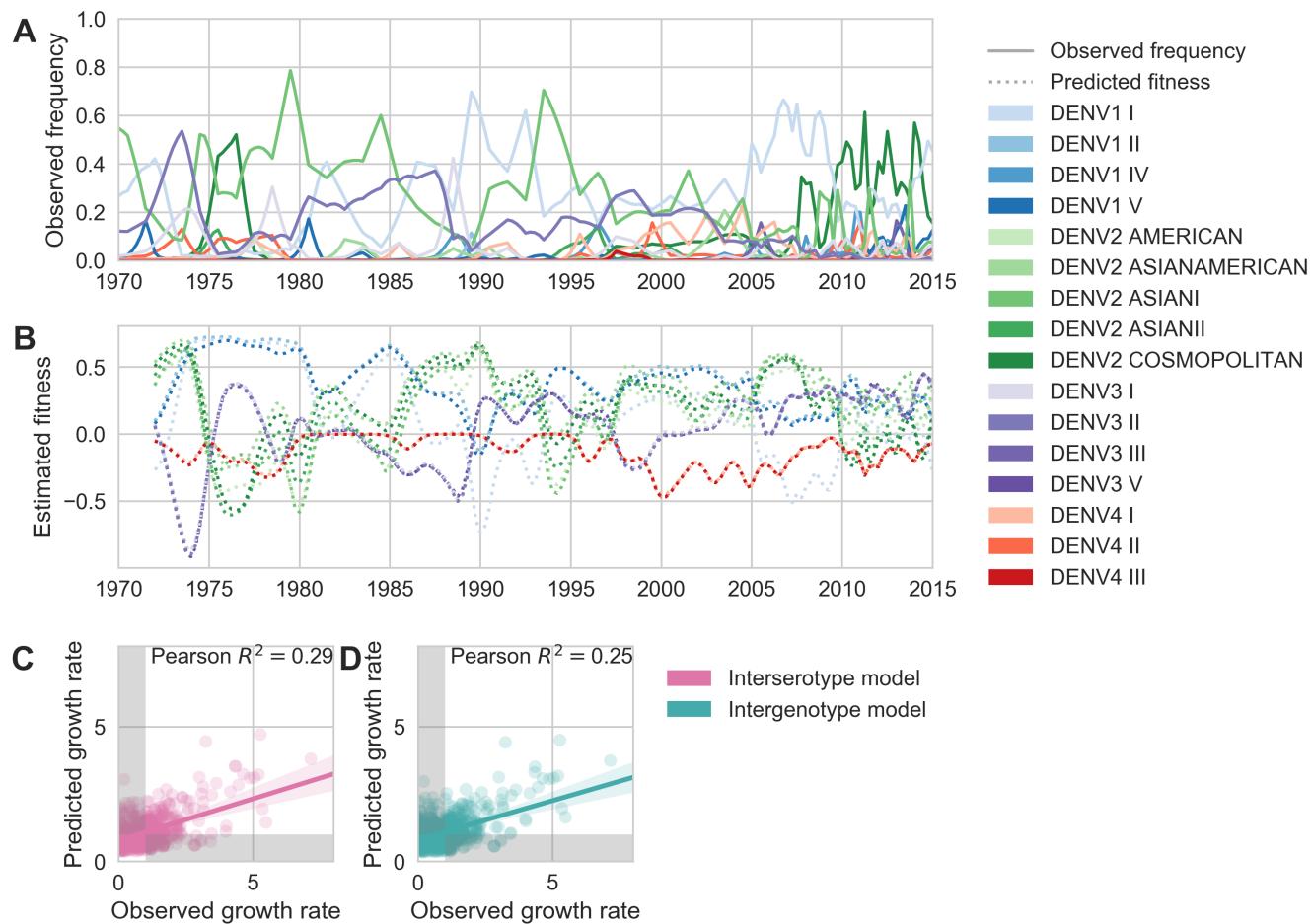


Figure 6. Antigenic novelty partially predicts genotype success **(A)** Relative frequencies of each canonical dengue genotype across Southeast Asia, estimated from available sequence data. **(B)** Antigenic fitness is calculated for each genotype as its frequency-weighted antigenic distance from recently 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 at either the ‘intergenotype’ or the ‘interserotype’ resolution. In this panel, we show total fitness over time, incorporating estimates of antigenic fitness derived from the ‘intergenotype’ model. **(C, D)** Fitness estimates were used to predict clade growth rates over 2 years, compounding immunity every three months based on predicted frequency changes (Methods Eq. 15). Here, we compare observed vs. predicted growth rates for both formulations of the fitness model (using fitness derived from either ‘interserotype’ or ‘intergenotype’ antigenic distances). Growth versus decline was accurate (predicted and actual growth rates both > 1 or both < 1 , points outside the gray shaded area) for 67% and 61% of predictions, respectively.

Table 3. Titer model performance comparisons. We compared performance across several different variations of the titer model. As described in *Neher et al. (2016)*, incremental antigenic change can be assigned to either amino acid substitutions ('Substitution' model) or to branches in the phylogeny ('Tree' model). For each of these models, we can constrain the model such that antigenic change is allowed to occur only between serotypes ('Interserotype') or between AND within serotypes ('Full'). For the substitution model, we constrain the interserotype model by reconstructing the amino acid sequence of the most recent common ancestor for each serotype and allowing the model to assign antigenic change only to mutations between these ancestral sequences. For the tree model, we constrain the interserotype model by allowing the model to assign antigenic change only to branches in the phylogeny that lie between serotypes. We also assess the impact of the virus avidity and serum potency terms, v_a and p_b . For all models and metrics, we report the mean and 95% confidence interval across 100-fold Monte Carlo cross validation with random 90%:10% training:test splits.

Model	Antigenic resolution	v_a and p_b	RMSE	Pearson R^2
Substitution	Full	Yes	0.75 (0.74–0.77)	0.78 (0.77–0.79)
Substitution	Full	No	1.13 (1.11–1.16)	0.50 (0.48–0.52)
Substitution	Interserotype	Yes	0.86 (0.85–0.88)	0.72 (0.70–0.73)
Substitution	Interserotype	No	0.86 (0.84–0.87)	0.71 (0.70–0.72)
Tree	Full	Yes	0.84 (0.83–0.86)	0.72 (0.71–0.73)
Tree	Full	No	1.40 (1.38–1.42)	0.24 (0.23–0.26)
Tree	Interserotype	Yes	0.87 (0.85–0.88)	0.70 (0.69–0.71)
Tree	Interserotype	No	0.86 (0.84–0.88)	0.72 (0.71–0.73)

265 Discussion

266 Within-serotype antigenic heterogeneity

267 We show that mapping antigenic change to specific mutations and interpolating across the DENV
 268 alignment is able to explain a large majority of the observed variation in antigenic phenotypes, as
 269 measured by neutralization titers. We identify 49 specific mutations and four colinear mutation
 270 clusters that contribute to antigenic variation, of which 27 mutations or mutation clusters have
 271 an antigenic impact of $0.20 \log_2$ titers or greater. These mutations span all major domains of
 272 *E*, and occur both within and between serotypes. This demonstrates that DENV antigenic diver-
 273 gence is closely coupled to genetic divergence. We use these mutations to infer unmeasured
 274 antigenic relationships between viruses, revealing substantial within-serotype antigenic variation.
 275 For comparison, we reconstructed the ancestral sequence of each serotype and constrained the
 276 model to only permit antigenic change to be attributed to these serotype-level differences. While
 277 this interserotype-only model predicts titers to a reasonable degree, we find that it has higher
 278 error (RMSE = 0.86) than the full model which accounts for within-serotype heterogeneity (RMSE =
 279 0.79; Table 3). This supports and expands upon previous reports (*Katzelnick et al., 2015; Forshey
 280 et al., 2016; Waggoner et al., 2016*) that the null hypothesis of antigenically uniform serotypes is
 281 inconsistent with observed patterns of cross-protection and susceptibility.

282 Consistent with the relatively long timescale of dengue evolution, we observe many sites in
 283 the dengue phylogeny to have mutated multiple times. These represent instances of parallelism,
 284 reversion and homoplasy. For example, we observe that site 390 is consistently S in DENV1, N in
 285 DENV3 and H in DENV4, while DENV2 genotypes show a mixture of D, N and S (Figure 2—Figure
 286 Supplement 2). We estimate an antigenic impact of $0.18 \log_2$ titers of the N390S mutation. Our
 287 model predicts that the parallel N390S mutations in DENV1 and DENV2 Cosmopolitan makes these
 288 viruses slightly more antigenically similar rather than more antigenically distinct. Along these lines,
 289 we compared the 'substitution' model to a similar model formulation (termed the 'tree' model),
 290 which assigns d_m values to individual branches in the phylogeny, rather than to individual mutations,
 291 so that each branch with a positive d_m value increases antigenic distance between strains (*Neher
 292 et al., 2016*). As expected from the high degree of homoplasy across the dengue phylogeny, we
 293 observe that the 'substitution' model outperforms the 'tree' model in predicting titers in validation
 294 datasets (Table 3).

295 To investigate the impact of this observed variation, we examine patterns of neutralization in
 296 response to vaccination with each monovalent component of the NIH vaccine candidate. Here, we
 297 see that each monovalent component elicits broad homotypic protection, but levels of heterotypic
 298 cross protection vary widely between heterotypic genotypes. This is consistent with previous
 299 reports of genotype-specific interactions between standing population immunity and subsequent
 300 heterotypic epidemics as modulating epidemic severity (*OhAinle et al., 2011; Kochel et al., 2002*).
 301 We hypothesize that this observed within-serotype variation primarily effects heterotypic secondary
 302 infection outcomes, rather than modulating homotypic immunity. Although we note that *Juraska*
 303 *et al. (2018)* demonstrate that vaccine efficacy decreases with increasing amino acid divergence of
 304 breakthrough infections from the vaccine insert.

305 Overall, we expect that these antigenic phenotypes represent a lower-bound on the extent,
 306 magnitude, and nature of antigenic heterogeneity with DENV. Our current titer dataset spans the
 307 breadth of DENV diversity, but due to small sample size, it lacks the resolution to detect most
 308 sub-genotype antigenic variation. The appearance of the deep antigenic divergence of the four
 309 serotypes, and the more recent antigenic divergences within each serotype, suggest that DENV
 310 antigenic evolution is likely an ongoing, though gradual, process. We therefore expect that future
 311 studies with richer datasets will find additional antigenic variation within each genotype. This
 312 dataset also contains many left-censored titer values, where we know two viruses are at least T
 313 titer units apart, but do not know exactly how far apart. If we knew the true value of these censored
 314 titers, many of them would indicate larger antigenic distances than the reported values, T , which
 315 are used to train the model. Thus, it is likely that our model systematically underestimates the
 316 magnitude of titer distances.

317 Finally, antibody neutralization and escape (as measured by PRNT titers) is only one component
 318 of the immune response to DENV. Although analysis of a longitudinal cohort study shows that these
 319 neutralization titers correlate with protection from severe secondary infection, it is unclear how
 320 PRNT titers correspond to antibody-dependent enhancement (*Katzelnick et al., 2016*). It is also
 321 important to note that DENV case outcomes are partially mediated by interactions with innate and
 322 T-cell immunity, the effects of which are not captured in neutralization titers (*Green et al., 2014*).
 323 Overall, while richer datasets and the development of more holistic assays will be required in order
 324 to fully characterize the extent of DENV antigenic diversity, it is clear that the four-serotype model is
 325 insufficient to explain DENV antigenic evolution.

326 **Viral clade dynamics**

327 We use these inferred antigenic relationships to directly quantify the impact of antigenic fitness on
 328 DENV population composition. To do so, we measure serotype frequencies across Southeast Asia
 329 over time and construct a model to estimate how they will fluctuate (Methods, Eq. 6–16). This model
 330 places a fitness value on each serotype that derives from a constant intrinsic component alongside
 331 a time-dependent antigenic component. Antigenic fitness declines with population immunity,
 332 which is accumulated via the recent circulation of antigenically similar viruses. Our primary model
 333 parameterization assumes that both heterotypic and homotypic immunity wane linearly over time
 334 at the same rate, with homotypic immunity starting from a higher baseline of protection based on
 335 closer antigenic distances. We compared this to a secondary model parameterization with only
 336 heterotypic waning (see Methods), under which we observe similar model performance (Table 2).

337 We find that antigenic fitness is able to explain much of the observed variation in serotype
 338 growth and decline (Figure 5). Forward simulations under the optimized parameter set display
 339 damped oscillations around the serotype-specific ‘set points’ determined by intrinsic fitnesses, but
 340 intrinsic fitness alone is unable to explain serotype fluctuations ($R^2 = 0.04$; Table 2, Figure 5—Figure
 341 Supplement 3). This demonstrates that although intrinsic fitness plays an important role in dictating
 342 long-term dynamics, wherein particular serotypes tend to circulate at low frequency (e.g., DENV4)
 343 and others at high frequency (e.g., DENV1 and DENV2), antigenic fitness plays out on shorter-term
 344 time scales, dictating circulation over several subsequent years.

345 We similarly use this model to quantify the effect of within-serotype antigenic variation on the
 346 success and decline of canonical DENV genotypes (Figure 6). As above, genotype antigenic fitness
 347 declines with population immunity. Here, we estimate population immunity based on antigenic
 348 distance from recently circulating genotypes, using distances that are either genotype-specific
 349 or based only on the serotype that each genotype belongs to. We then directly compare how
 350 strongly these coarser serotype-level versus specific genotype-level antigenic relationships impact
 351 DENV population dynamics. Overall, we find that antigenic fitness explains a moderate portion
 352 of the observed variation in genotype growth and decline. Surprisingly, however, we find that
 353 incorporating within-serotype antigenic differences does not improve our predictions (Figure 6C-D).
 354 We suggest two possible explanations for this observation.

355 First, it may be that although genotypes are antigenically diverse, these differences do not
 356 influence large-scale regional dynamics over time. We may then hypothesize that within-serotype
 357 antigenic heterogeneity mediates disease severity, but does not influence infection or onward
 358 transmission. This hypothesis is consistent with the findings of *Nagao and Koelle (2008)*, who
 359 demonstrated that dengue epidemiological dynamics are compatible with a model wherein im-
 360 munity confers protection against severe symptoms, but not asymptomatic infection. This is also
 361 consistent with *ten Bosch et al. (2018)*'s findings that asymptomatic dengue infections contribute
 362 to onward transmission.

363 Alternatively, this lack of signal could be methodologically explained if either (A) genotype-level
 364 frequency trajectories estimated from public data are overly noisy for this application or (B) our
 365 model of antigenic fitness based on PRNT assay data does not match reality, due to either PRNT
 366 assay data not well reflecting human immunity or due to our particular model formulation that
 367 parameterizes immunity from titer distances (Eq. 6-10). In the present analysis, we are not able to
 368 firmly resolve these disparate possibilities.

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

383 **Conclusions**

384 We find that within-serotype antigenic evolution helps explain observed patterns of cross-neutralization
 385 among dengue genotypes. We also find that serotype-level population immunity is a strong deter-
 386 minant of viral clade dynamics across Southeast Asia. As richer datasets become available, future
 387 studies that similarly combine viral genomics, functional antigenic characterization, and population
 388 modeling have great potential to improve our understanding of how DENV evolves antigenically
 389 and moves through populations.

390 **Model sharing and extensions**

391 We have provided all code, configuration files and datasets at github.com/blab/dengue-antigenic-
 392 [dynamics](https://github.com/blab/dengue-antigenic-dynamics), and wholeheartedly encourage other groups to adapt and extend this framework for
 393 further investigation of DENV antigenic evolution and population dynamics.

394 **Methods**

395 **Data**

396 **Titers**

397 Antigenic distance between pairs of viruses i and j is experimentally measured using a neutralization
 398 titer, which measures how well serum drawn after infection with virus i is able to neutralize virus
 399 j *in vitro* (**Russell and Nisalak, 1967**). Briefly, two-fold serial dilutions of serum i are incubated
 400 with a fixed concentration of virus j . Titers represent the lowest serum concentration able to
 401 neutralize 50% of virus, and are reported as the inverse dilution. We used two publicly available
 402 plaque reduction neutralization titer (PRNT50) datasets generated by Katzelnick *et al.* in (**Katzelnick**
 403 *et al., 2015*). The primary dataset was generated by infecting each of 36 non-human primates
 404 with a unique strain of DENV. NHP sera was drawn after 12 weeks and titered against the panel of
 405 DENV viruses. The secondary dataset was generated by vaccinating 31 human trial participants
 406 with a monovalent component of the NIH DENV vaccine. Sera was drawn after 6 weeks and titered
 407 against the same panel of DENV viruses. As discussed in Katzelnick *et al.*, these two datasets show
 408 similar patterns of antigenic relationships between DENV viruses. In total, our dataset includes
 409 1182 measurements across 48 virus strains: 36 of these were used to generate serum, and 47 were
 410 used as test viruses.

411 **Sequences**

412 For the titer model analysis, we used the full sequence of E (envelope) from the 48 strains in the
 413 titer dataset.

414 For the clade frequencies analysis, we downloaded all dengue genome sequences available
 415 from the Los Alamos National Lab Hemorrhagic Fever Virus Database as of March 7, 2018, that
 416 contained at least the full coding sequence of E (envelope) (total N=12,645) (**Kuiken et al., 2011**).
 417 We discarded sequences which were putative recombinants, duplicates, lab strains, or which lacked
 418 an annotated sampling location and/or sampling date. We selected all remaining virus strains that
 419 were annotated as a Southeast Asian isolate (total N = 8,644).

420 For both datasets, we used the annotated reference dataset from (**Pyke et al., 2016**) to assign
 421 sequences to canonical genotypes.

422 **Alignments and trees**

423 We used MAFFT v7.305b to align nucleotide E gene sequences for each strain before translating the
 424 aligned sequences (no frame-shift indels were present) (**Katoh and Standley, 2013**). All maximum
 425 likelihood phylogenies were constructed with IQ-TREE version 1.6.8 and the GTR+I+G15 nucleotide
 426 substitution model (**Nguyen et al., 2014**).

427 **Titer Model**

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

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

430 We then average normalized titers across individuals. To predict unmeasured titers, we employ the
 431 'substitution model' from (**Neher et al. (2016)**) and implemented in Nextstrain (**Hadfield et al., 2018**),
 432 which assumes that antigenic evolution is driven by underlying genetic evolution.

433 In the substitution model, observed titer drops are mapped to mutations between each serum
 434 and test virus strain after correcting for overall virus avidity, v_i , and serum potency, p_j ('row' and
 435 'column' effects, respectively), so that

$$\hat{D}_{ij} \approx D_{ij} = \sum_m d_m + v_i + p_j, \quad (2)$$

436 where d_m is the titer drop assigned to each mutation, m , between serum i and virus j , and m iterates
 437 over mutations. We randomly withhold 10% of titer measurements as a test set. We use the
 438 remaining 90% of titer measurements as a training set to learn values for virus avidity, serum
 439 potency, and mutation effects. As in **Neher et al. (2016)**, we formulate this as a convex optimization
 440 problem and solve for these parameter values to minimize the cost function

$$C = \sum_{i,j} (\hat{D}_{ij} - D_{ij})^2 + \lambda \sum_m d_m + \kappa \sum_i v_i^2 + \delta \sum_j p_j^2. \quad (3)$$

441 We used $\lambda = 3.0$, $\kappa = 0.6$, and $\delta = 1.2$ to minimize test error. Respectively, these terms represent
 442 the squared training error; an L1 regularization term on mutation effects, such that most values
 443 of $d_m = 0$; and L2 regularization terms on virus avidities and serum potencies, such that they are
 444 normally distributed. These parameter values are then used to predict the antigenic distance
 445 between all pairs of viruses, i and j . We assess performance by comparing predicted to known
 446 titer values in our test data set, and present test error (aggregated from 100-fold Monte Carlo
 447 cross-validation) throughout the manuscript.

448 Viral Clade Dynamics

449 Empirical Clade Frequencies

450 As discussed in **Neher et al. (2016)** and **Lee et al. (2018)**, we estimate empirical clade frequencies
 451 from 1970 to 2015 based on observed relative abundances of each clade in the ‘slice’ of the phylogeny
 452 corresponding to each quarterly timepoint.

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

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

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

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

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

463 where $v \in V$ represents the set of viruses that belong to the clade and $v \notin V$ represents the set of
 464 viruses that do not belong to the clade. Each frequency trajectory is estimated by simultaneously
 465 maximizing the likelihood of the process model and the likelihood of the observation model via
 466 adjusting frequency trajectory $\vec{x} = (x_1, \dots, x_n)$.

467 Population Immunity

468 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
 469 viruses are able to access more susceptible hosts, allowing the antigenically novel lineage to expand.
 470 We adapt a simple deterministic model from **Tuksza and Lässig (2014)** to directly quantify dengue
 471 antigenic novelty and its impact on viral fitness. We quantify population immunity to virus i at
 472 time t , $P_i(t)$, as a function of which clades have recently circulated in the past N years, and how
 473 antigenically similar each of these clades is to virus i , so that

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

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

$$w(n) = \max(1 - \gamma n, 0). \quad (7)$$

478 The relationship between antigenic distance and the probability of protection, C , is also assumed to
 479 be non-negative and linear with slope $-\sigma$, such that

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

480 In addition to this primary analysis, we conducted a secondary analysis with a different parameterization of immunity that removes waning of homotypic immunity while allowing waning of
 481 heterotypic immunity. In this case, we assume the relationship between antigenic distance and the
 482 probability of protection, C , to be 50% at antigenic distance $1/\sigma$ and to wane based on years since
 483 infection n modified by γ_{het} following

$$C(D_{ij}, n) = \exp(-\sigma(1/\gamma_{\text{het}})^n D_{ij}). \quad (9)$$

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

$$f_i(t) = f_0 - \beta P_i(t), \quad (10)$$

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

488 Frequency Predictions

489 Similar to the model implemented in [Łuksz and Lässig \(2014\)](#), we estimate predicted clade frequencies at time $t + dt$ as

$$\hat{x}_i(t + dt) = \frac{x_i(t)e^{f_i(t)dt}}{\sum_i x_i(t)e^{f_i(t)dt}} \quad (11)$$

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

492 We do not attempt to predict future frequencies for clades with $x_i(t) < 0.05$.

493 For long-term predictions, we must account for immunity accrued at each intermediate time-point between t and dt . We divide the interval between t and dt into a total of U 3 month timepoints,
 494 $[t + u, t + 2u, \dots, t + U]$, such that $t + U = dt$. We then compound immunity based on predicted clade
 495 frequencies at each intermediate timepoint following

$$\hat{x}_i(t + u) = x_i(t)e^{f_i(t)u} \quad (12)$$

497

$$\hat{x}_i(t + 2u) = \hat{x}_i(t + u)e^{f_i(t+u)u} \quad (13)$$

...

498

$$\hat{x}_i(t + U) = x_i(t)e^{f_i(t)u}e^{f_i(t+u)u}e^{f_i(t+2u)u}\dots e^{f_i(t+U)u} \quad (14)$$

499

$$\hat{x}_i(t + dt) = \hat{x}_i(t + U) = x_i(t)e^{\sum_u f_i(t+u)u} \quad (15)$$

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

$$\frac{\hat{x}_i(t + dt)}{x_i(t)}. \quad (16)$$

Table 4. Optimized fitness model parameters for primary analysis

Parameter	Value	Description
β	1.02	Slope of linear relationship between population immunity and viral fitness
γ	0.83	Proportion of titers waning each year since primary infection
σ	0.76	Slope of linear relationship between titers and probability of protection
$f_0^{(1)}$	0.74	Relative intrinsic fitness of DENV1
$f_0^{(2)}$	0.84	Relative intrinsic fitness of DENV2
$f_0^{(3)}$	0.50	Relative intrinsic fitness of DENV3
$f_0^{(4)}$	0.00	Relative intrinsic fitness of DENV4 (fixed)

502 Null models

503 To quantify the impact of antigenic fitness on DENV clade success, we compare our antigenically-
504 informed model to two null models.

505 Under the ‘equal fitness null’ model, all viruses have equal total fitness (antigenic and intrinsic
506 fitness) at all timepoints

$$f_i^{equal}(t) = 0 \quad (17)$$

507

$$\hat{x}_i^{equal}(t + dt) = x_i(t)e^0 = x_i(t). \quad (18)$$

508 Under the ‘intrinsic fitness null’ model, all viruses have equal antigenic fitness but serotype-
509 specific intrinsic fitness at all timepoints

$$f_i^{intrinsic}(t) = f_0 \quad (19)$$

510

$$\hat{x}_i^{intrinsic}(t + dt) = x_i(t)e^{f_0}. \quad (20)$$

511 Model performance assessment and parameter fitting

512 We assess predictive power as the root mean squared error between predicted and empirical clade
513 frequencies. To assess both the final frequency predictions and the predicted clade trajectories,
514 this RMSE includes error for each clade, for each starting timepoint t , and for each intermediate
515 predicted timepoint $t + u$.

516 Our frequency prediction model has a total of 6 free parameters. We jointly fit these parameters
517 to minimize RMSE of serotype frequency predictions via the Nelder-Mead algorithm as implemented
518 in SciPy v.1.0.0 (Table 4) ([Jones et al., 2001](#); [Gao and Han, 2012](#)). We use $N = 2$ years of previous
519 immunity that contribute to antigenic fitness and project $dt = 2$ years in the future when predicting
520 clade frequencies.

521 Simulations

522 To ensure the model machinery functions correctly, we seeded a forward simulation of clade
523 dynamics with two years of empirical frequencies and simulated predicted dynamics over the
524 remainder of the time course (Figure 5—Figure Supplement 2). We then fit model parameters as
525 described above, and obtained parameter values that well recover input values (Table 5).

526 Data and software availability

527 Sequence and titer data, as well as all code used for analyses and figure generation, is pub-
528 licly available at github.com/blab/dengue-antigenic-dynamics. Our work relies upon many open
529 source Python packages and software tools, including iPython ([Pérez and Granger, 2007](#)), Mat-
530plotlib ([Hunter, 2007](#)), Seaborn ([Waskom, 2017](#)), Pandas ([McKinney et al., 2010](#)), CVXOPT ([Andersen
et al., 2013](#)), NumPy ([Van Der Walt et al., 2011](#); [Gao and Han, 2012](#)), Biopython ([Cock et al., 2009](#)),
531 SciPy ([Jones et al., 2001](#)), Statsmodels ([Seabold and Perktold, 2010](#)), Nextstrain ([Hadfield et al.,](#)

Table 5. Parameter recovery against simulated data.

Parameter	Input value	Optimized value
β	3.25	3.10
γ	0.55	0.56
σ	2.35	2.57
$f_0^{(1)}$	0.70	0.72
$f_0^{(2)}$	0.85	0.78
$f_0^{(3)}$	0.40	0.41

533 2018), MAFFT (Katoh and Standley, 2013), and IQ-TREE (Nguyen et al., 2014). Package versions are
 534 documented in the GitHub repository.

535 Acknowledgements

536 We would like to thank Richard Neher, John Huddleston, Andrew Rambaut, Molly OhAinle, David
 537 Shaw, Paul Edlefsen, Michal Juraska, and all members of the Bedford Lab for useful discussion
 538 and advice. SB is a Graduate Research Fellow and is supported by NSF DGE-1256082. TB is a Pew
 539 Biomedical Scholar and is supported by NIH R35 GM119774-01. LK is supported by NIH awards
 540 R01AI114703-01 and P01AI106695. Our work depends on open data sharing and many open
 541 source software tools. We gratefully acknowledge the authors and developers who make our work
 542 possible.

543 References

- 544 **Allcock OM**, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, Suchard MA, Foster JE, Rambaut A,
 545 Carrington CV. Phylogeography and population dynamics of dengue viruses in the Americas. Molecular
 546 biology and evolution. 2012; 29(6):1533–1543.
- 547 **de Alwis R**, Williams KL, Schmid MA, Lai CY, Patel B, Smith SA, Crowe JE, Wang WK, Harris E, de Silva AM. Dengue
 548 viruses are enhanced by distinct populations of serotype cross-reactive antibodies in human immune sera.
 549 PLoS pathogens. 2014; 10(10):e1004386.
- 550 **Andersen M**, Dahl J, Vandenberghe L. CVXOPT: A Python package for convex optimization. abel ee ucla
 551 edu/cvxopt. 2013; .
- 552 **Bedford T**, Cobey S, Pascual M. Strength and tempo of selection revealed in viral gene genealogies. BMC
 553 Evolutionary Biology. 2011; 11(1):220.
- 554 **Bedford T**, Rambaut A, Pascual M. Canalization of the evolutionary trajectory of the human influenza virus.
 555 BMC Biol. 2012; 10(1):38.
- 556 **Bhatt S**, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O,
 557 et al. The global distribution and burden of dengue. Nature. 2013; 496(7446):504.
- 558 **ten Bosch Q**, Clapham HE, Lambrechts L, Duong V, Buchy P, Althouse BM, Lloyd AL, Waller LA, Morrison AC,
 559 Kitron U, et al. Contributions from the silent majority dominate dengue virus transmission. PLoS pathogens.
 560 2018; 14(5):e1006965.
- 561 **Cock PJ**, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B,
 562 et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics.
 563 Bioinformatics. 2009; 25(11):1422–1423.
- 564 **Forshey BM**, Reiner RC, Olkowski S, Morrison AC, Espinoza A, Long KC, Vilcarromero S, Casanova W, Wearing
 565 HJ, Halsey ES, et al. Incomplete protection against dengue virus type 2 re-infection in Peru. PLoS neglected
 566 tropical diseases. 2016; 10(2):e0004398.
- 567 **Gao F**, Han L. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. Computational
 568 Optimization and Applications. 2012; 51(1):259–277.

- 569 **Gentry M**, Henchal E, McCown J, Brandt W, Dalrymple J. Identification of distinct antigenic determinants on
570 dengue-2 virus using monoclonal antibodies. *The American journal of tropical medicine and hygiene*. 1982;
571 31(3):548–555.
- 572 **Gibbons RV**, Kalanarooj S, Jarman RG, Nisalak A, Vaughn DW, Endy TP, Mammen Jr MP, Sriiatkhachorn A.
573 Analysis of repeat hospital admissions for dengue to estimate the frequency of third or fourth dengue
574 infections resulting in admissions and dengue hemorrhagic fever, and serotype sequences. *The American
575 journal of tropical medicine and hygiene*. 2007; 77(5):910–913.
- 576 **Green AM**, Beatty PR, Hadjilaou A, Harris E. Innate immunity to dengue virus infection and subversion of
577 antiviral responses. *Journal of molecular biology*. 2014; 426(6):1148–1160.
- 578 **Gupta S**, Ferguson N, Anderson R. Chaos, persistence, and evolution of strain structure in antigenically diverse
579 infectious agents. *Science*. 1998; 280(5365):912–915.
- 580 **Hadfield J**, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA. Nextstrain:
581 Real-time tracking of pathogen evolution. *Bioinformatics*. 2018; 34:4121–4123.
- 582 **Halstead SB**. In vivo enhancement of dengue virus infection in rhesus monkeys by passively transferred
583 antibody. *Journal of Infectious Diseases*. 1979; 140(4):527–533.
- 584 **Holmes EC**, Twiddy SS. The origin, emergence and evolutionary genetics of dengue virus. *Infection, genetics
585 and evolution*. 2003; 3(1):19–28.
- 586 **Hunter JD**. Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*. 2007; 9(3):90–95. doi:
587 [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- 588 **Jones E**, Oliphant T, Peterson P, et al., SciPy: Open source scientific tools for Python; 2001. <http://www.scipy.org/>,
589 [Online; accessed <today>].
- 590 **Juraska M**, Magaret CA, Shao J, Carpp LN, Fiore-Gartland AJ, Benkeser D, Girerd-Chambaz Y, Langevin E, Frago
591 C, Guy B, et al. Viral genetic diversity and protective efficacy of a tetravalent dengue vaccine in two phase 3
592 trials. *Proceedings of the National Academy of Sciences*. 2018; p. 201714250.
- 593 **Katoh K**, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance
594 and usability. *Molecular biology and evolution*. 2013; 30(4):772–780.
- 595 **Katzelnick LC**, Fonville JM, Gromowski GD, Arriaga JB, Green A, James SL, Lau L, Montoya M, Wang C, VanBlargan
596 LA, et al. Dengue viruses cluster antigenically but not as discrete serotypes. *Science*. 2015; 349(6254):1338–
597 1343.
- 598 **Katzelnick LC**, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, Balmaseda A, Harris E. Antibody-dependent
599 enhancement of severe dengue disease in humans. *Science*. 2017; 358(6365):929–932.
- 600 **Katzelnick LC**, Montoya M, Gresh L, Balmaseda A, Harris E. Neutralizing antibody titers against dengue virus
601 correlate with protection from symptomatic infection in a longitudinal cohort. *Proceedings of the National
602 Academy of Sciences*. 2016; 113(3):728–733.
- 603 **Kochel TJ**, Watts DM, Halstead SB, Hayes CG, Espinoza A, Felices V, Caceda R, Bautista CT, Montoya Y, Douglas S,
604 et al. Effect of dengue-1 antibodies on American dengue-2 viral infection and dengue haemorrhagic fever.
605 *The Lancet*. 2002; 360(9329):310–312.
- 606 **Kuiken C**, Thurmond J, Dimitrijevic M, Yoon H. The LANL hemorrhagic fever virus database, a new platform for
607 analyzing biothreat viruses. *Nucleic acids research*. 2011; 40(D1):D587–D592.
- 608 **Lanciotti RS**, Gubler DJ, Trent DW. Molecular evolution and phylogeny of dengue-4 viruses. *Journal of General
609 Virology*. 1997; 78(9):2279–2284.
- 610 **Lee JM**, Huddleston J, Doud MB, Hooper KA, Wu NC, Bedford T, Bloom JD. Deep mutational scanning of
611 hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proceedings of the
612 National Academy of Sciences*. 2018; .
- 613 **Lipsitch M**, O'Hagan JJ. Patterns of antigenic diversity and the mechanisms that maintain them. *Journal of the
614 Royal Society Interface*. 2007; 4(16):787–802.
- 615 **Lourenço J**, Recker M. Natural, persistent oscillations in a spatial multi-strain disease system with application to
616 dengue. *PLoS computational biology*. 2013; 9(10):e1003308.

- 617 **Lourenço J**, Tennant W, Faria NR, Walker A, Gupta S, Recker M. Challenges in dengue research: A computational
618 perspective. *Evolutionary applications*. 2018; 11(4):516–533.
- 619 **Łuksza M**, Lässig M. A predictive fitness model for influenza. *Nature*. 2014; 507(7490):57.
- 620 **McKinney W**, et al. Data structures for statistical computing in python. In: *Proceedings of the 9th Python in*
621 *Science Conference*, vol. 445 Austin, TX; 2010. p. 51–56.
- 622 **Mizumoto K**, Ejima K, Yamamoto T, Nishiura H. On the risk of severe dengue during secondary infection: a
623 systematic review coupled with mathematical modeling. *Journal of vector borne diseases*. 2014; 51(3):153.
- 624 **Nagao Y**, Koelle K. Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic
625 fever. *Proceedings of the National Academy of Sciences*. 2008; 105(6):2238–2243.
- 626 **Neher RA**, Bedford T, Daniels RS, Russell CA, Shraiman BI. Prediction, dynamics, and visualization of anti-
627 genic phenotypes of seasonal influenza viruses. *Proceedings of the National Academy of Sciences*. 2016;
628 113(12):E1701–E1709.
- 629 **Nguyen LT**, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for
630 estimating maximum-likelihood phylogenies. *Molecular biology and evolution*. 2014; 32(1):268–274.
- 631 **OhAinle M**, Balmaseda A, Macalalad AR, Tellez Y, Zody MC, Saborío S, Nuñez A, Lennon NJ, Birren BW, Gordon
632 A, et al. Dynamics of dengue disease severity determined by the interplay between viral genetics and
633 serotype-specific immunity. *Science translational medicine*. 2011; 3(114):114ra128–114ra128.
- 634 **Olkowski S**, Forshey BM, Morrison AC, Rocha C, Vilcarromero S, Halsey ES, Kochel TJ, Scott TW, Stoddard ST.
635 Reduced risk of disease during postsecondary dengue virus infections. *The Journal of infectious diseases*.
636 2013; 208(6):1026–1033.
- 637 **Pérez F**, Granger BE. IPython: a system for interactive scientific computing. *Computing in Science & Engineering*.
638 2007; 9(3):21–29.
- 639 **Pyke AT**, Moore PR, Taylor CT, Hall-Mendelin S, Cameron JN, Hewitson GR, Pukallus DS, Huang B, Warrilow D, Van
640 Den Hurk AF. Highly divergent dengue virus type 1 genotype sets a new distance record. *Scientific reports*.
641 2016; 6:22356.
- 642 **Reich NG**, Shrestha S, King AA, Rohani P, Lessler J, Kalayanarooj S, Yoon IK, Gibbons RV, Burke DS, Cummings
643 DA. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. *Journal of*
644 *The Royal Society Interface*. 2013; 10(86):20130414.
- 645 **Rico-Hesse R**. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology*. 1990;
646 174(2):479–493.
- 647 **Russell PK**, Nisalak A. Dengue virus identification by the plaque reduction neutralization test. *The Journal of*
648 *Immunology*. 1967; 99(2):291–296.
- 649 **Sabin AB**. Research on dengue during World War II1. *The American journal of tropical medicine and hygiene*.
650 1952; 1(1):30–50.
- 651 **Salje H**, Cummings DA, Rodriguez-Barraquer I, Katzelnick LC, Lessler J, Klungthong C, Thaisomboonsuk B, Nisalak
652 A, Weg A, Ellison D, et al. Reconstruction of antibody dynamics and infection histories to evaluate dengue risk.
653 *Nature*. 2018; 557(7707):719.
- 654 **Salje H**, Lessler J, Berry IM, Melendrez MC, Endy T, Kalayanarooj S, Atchareeya A, Chanama S, Sangkijporn S,
655 Klungthong C, et al. Dengue diversity across spatial and temporal scales: local structure and the effect of
656 host population size. *Science*. 2017; 355(6331):1302–1306.
- 657 **Sangkawibha N**, Rojanasuphot S, Ahandrik S, Viriyapongse S, Jatanasen S, Salitul V, Phanthumachinda B,
658 Halstead SB. Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong, Thailand.
659 *American journal of epidemiology*. 1984; 120(5):653–669.
- 660 **Seabold S**, Perktold J. Statsmodels: Econometric and statistical modeling with python. In: *9th Python in Science*
661 *Conference*; 2010. .
- 662 **Stanaway JD**, Shepard DS, Undurraga EA, Halasa YA, Coffeng LE, Brady OJ, Hay SI, Bedi N, Bensenor IM,
663 Castañeda-Orjuela CA, et al. The global burden of dengue: an analysis from the Global Burden of Disease
664 Study 2013. *The Lancet infectious diseases*. 2016; 16(6):712–723.

- 665 **Tibshirani R.** Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*
666 (Methodological). 1996; 58(1):267–288.
- 667 **Twiddy SS**, Holmes EC, Rambaut A. Inferring the rate and time-scale of dengue virus evolution. *Molecular*
668 *biology and evolution*. 2003; 20(1):122–129.
- 669 **Van Der Walt S**, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation.
670 *Computing in Science & Engineering*. 2011; 13(2):22.
- 671 **Waggoner JJ**, Balmaseda A, Gresh L, Sahoo MK, Montoya M, Wang C, Abeynayake J, Kuan G, Pinsky BA, Harris
672 E. Homotypic dengue virus reinfections in Nicaraguan children. *The Journal of infectious diseases*. 2016;
673 214(7):986–993.
- 674 **Waskom M**, Seaborn visualization library; 2017. <http://seaborn.pydata.org/>.
- 675 **Wearing HJ**, Rohani P. Ecological and immunological determinants of dengue epidemics. *Proceedings of the*
676 *National Academy of Sciences*. 2006; 103(31):11802–11807.
- 677 **Zhang C**, Mammen MP, Chinnawirotisan P, Klungthong C, Rodpradit P, Monkongdee P, Nimmannitya S, Kalaya-
678 narooj S, Holmes EC. Clade replacements in dengue virus serotypes 1 and 3 are associated with changing
679 serotype prevalence. *Journal of virology*. 2005; 79(24):15123–15130.

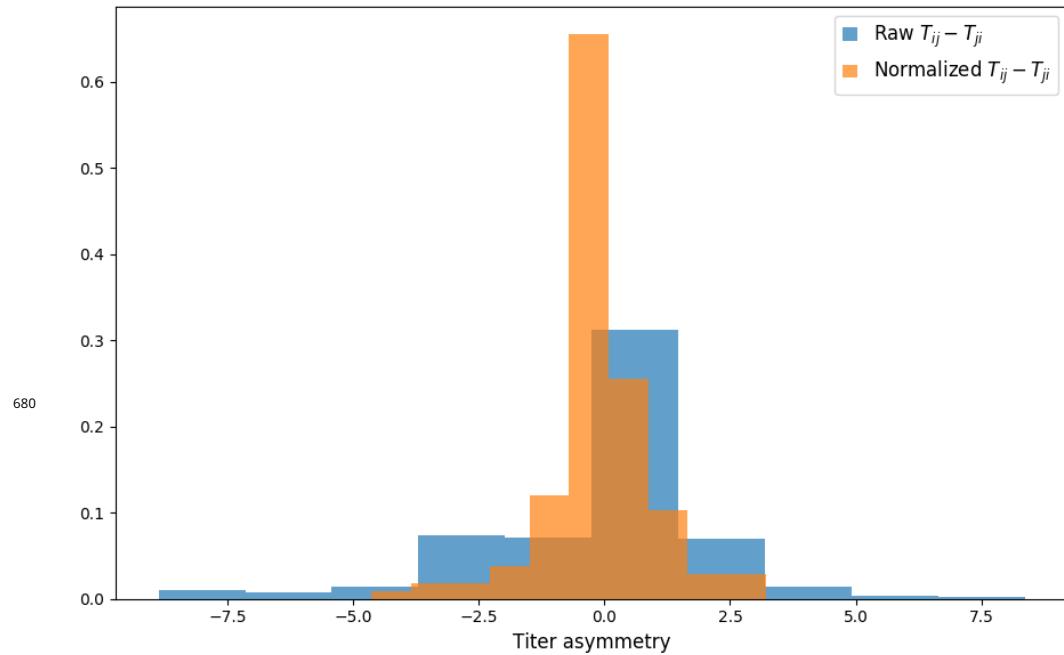


Figure 1–Figure supplement 1. Titer value symmetry. Some viruses have greater avidity overall, and some sera are more potent overall. We normalize for these row and column effects (v_a and p_b , respectively) in the titer model. Once overall virus avidity and serum potency are accounted for, titers are roughly symmetric (i.e., $D_{ij} \approx D_{ji}$).

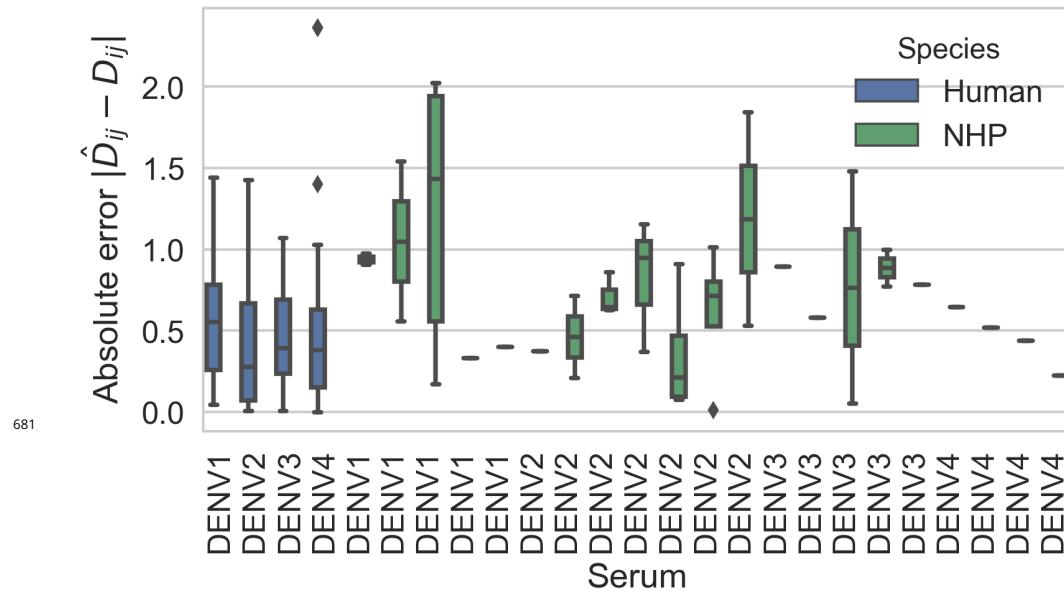


Figure 2–Figure supplement 1. Titer prediction error by serum strain and species. Human sera was raised against four different virus strains (the monovalent vaccine components); non-human primate (NHP) sera was raised against many different virus strains. Here, we excluded NHP sera raised against the monovalent vaccine components, such that each normalized titer measurement is aggregated across individuals, but not across species. We report the out-of-sample titer prediction error for each serum strain (versus all available test viruses), aggregated across 100-fold Monte Carlo cross-validation.

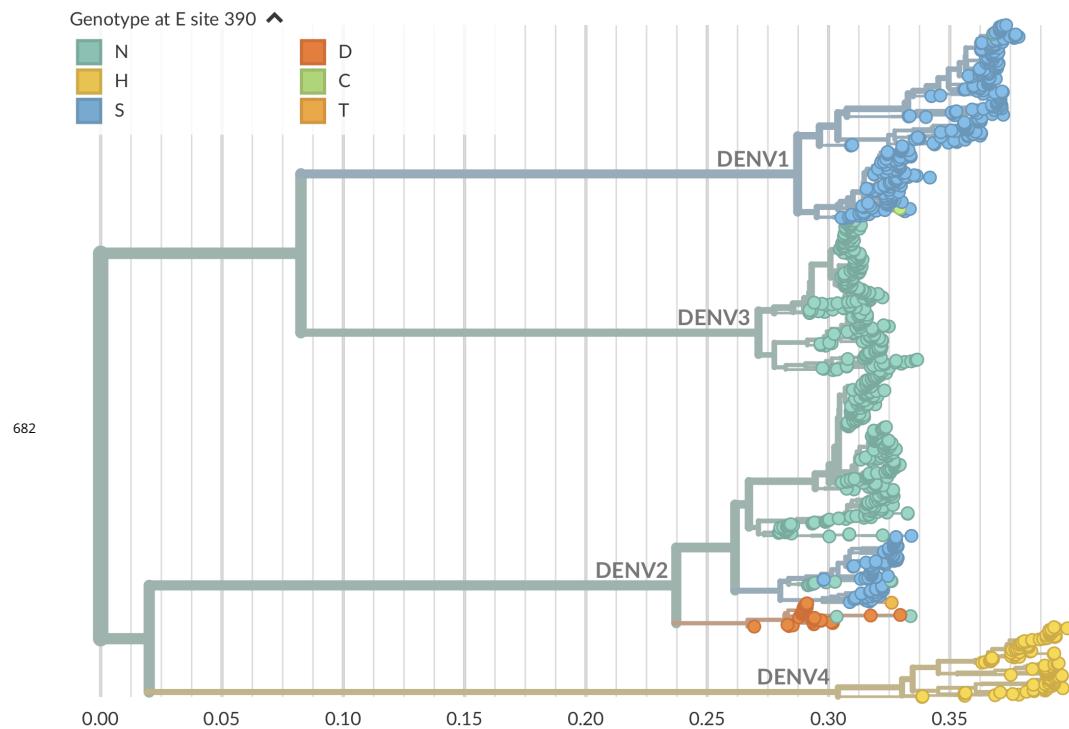


Figure 2–Figure supplement 2. Genotype as site E 390 across dengue phylogeny. Dengue virus genotypes can be seen on Nextstrain ([Hadfield et al., 2018](#)). A live view of this figure is available at nextstrain.org/dengue/.

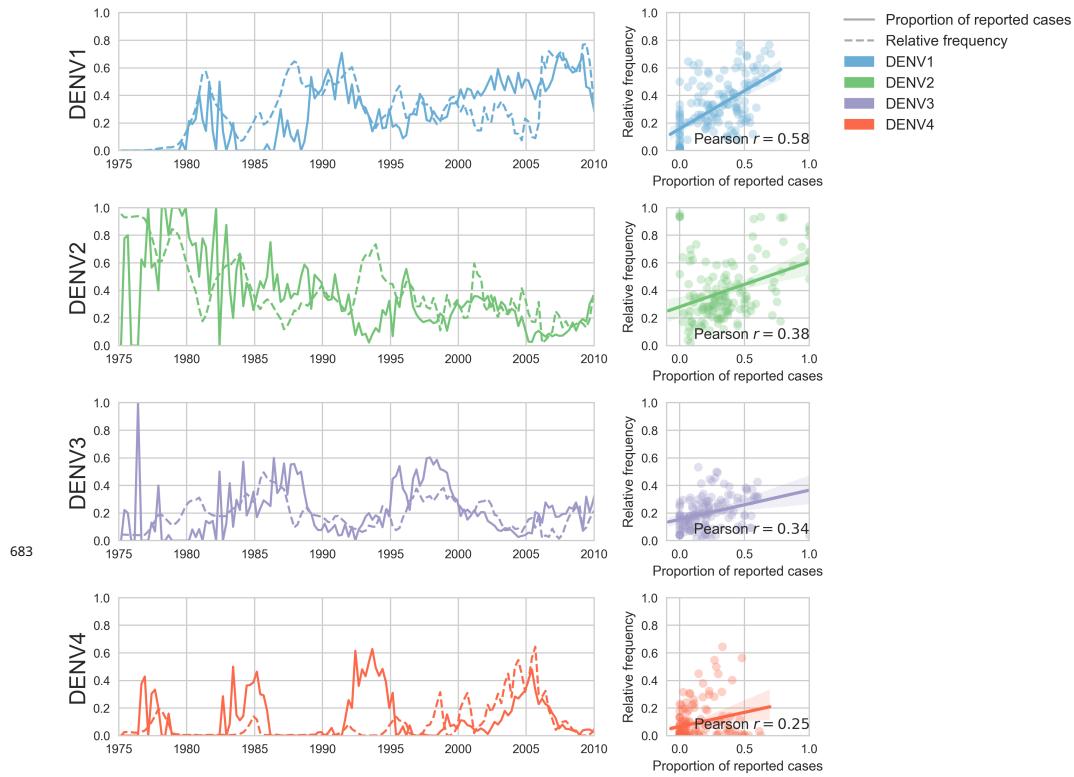


Figure 5-Figure supplement 1. Case counts versus clade frequencies in Thailand. As described in the Methods, we estimate clade frequencies based on observed relative abundance in the 'slice' of the phylogeny at each quarterly timepoint. These frequency estimates are smoothed using a discretized Brownian motion diffusion process. Here, we compare estimated serotype frequencies across Thailand (all available high quality sequences) to case counts from a hospital in Bangkok between 1975–2010 (*Reich et al., 2013*). Biweekly case counts were aggregated into quarterly timepoints, but were not smoothed. While there are some instances where case counts and frequencies diverge (e.g., DENV4 in the early 1990s), the noisy nature of the unsmoothed case counts artificially deflates estimates of concordance.

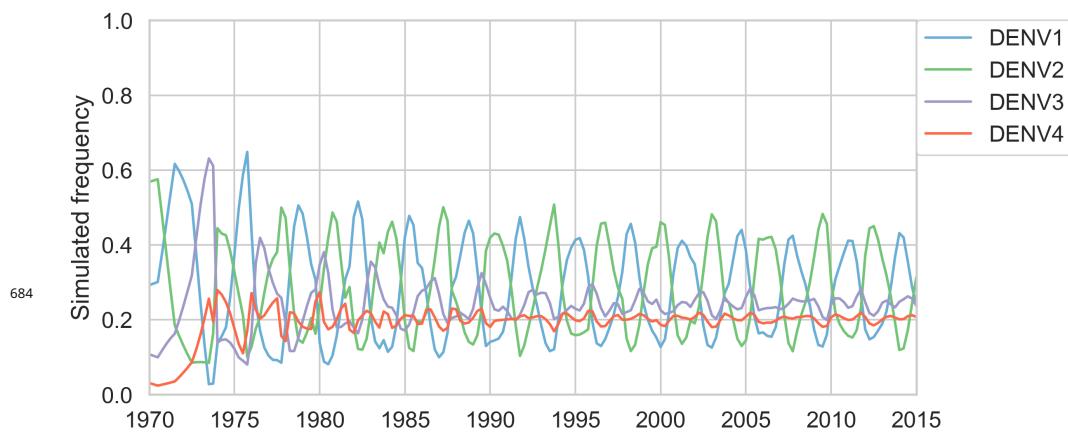


Figure 5-Figure supplement 2. Simulated serotype frequencies. As described in the Methods, we seeded a simulation with two years of empirical frequencies and predicted forward to simulate the remainder of the timecourse. Here, we simulated under the model parameters described in Table 5.

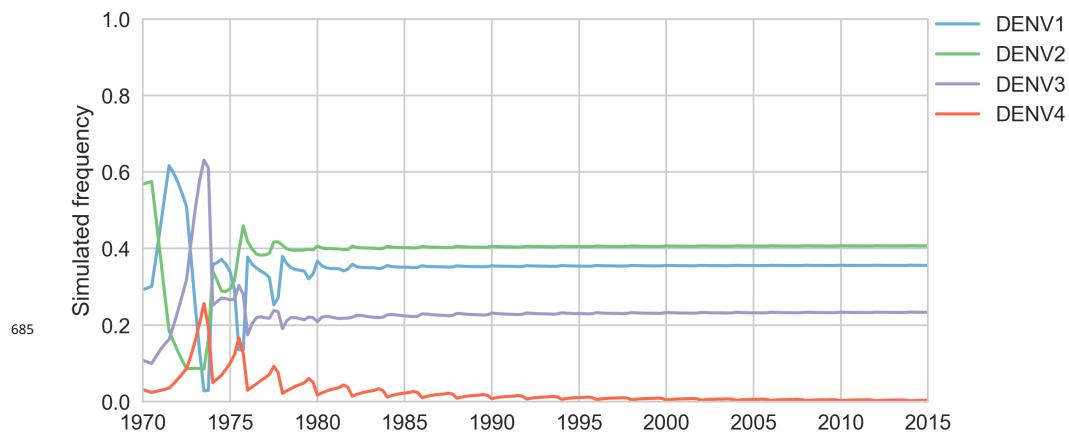


Figure 5-Figure supplement 3. Simulated serotype frequencies (model parameters). As described in the Methods, we seeded a simulation with two years of empirical frequencies and predicted forward to simulate the remainder of the timecourse. Here, we simulated under the model parameters described in Table 4. This results in damped oscillations around the intrinsic fitness value for each serotype, but these intrinsic fitnesses alone are unable to predict observed clade dynamics (Table 2).