Main Manuscript for

Climate, demography, immunology, and virology combine to drive two decades of dengue virus dynamics in Cambodia

Cara E. Brook[1*], Carly Rozins[2], Jennifer A. Bohl[3], Vida Ahyong[4], Sophana Chea[5], Liz Fahsbender[6], Rekol Huy[7], Sreyngim Lay[5], Rithea Leang[7], Yimei Li[1], Chanthap Lon[5], Somnang Man[5,7], Mengheng Oum [5], Graham R. Northrup[8], Fabiano Oliveira[3], Andrea R. Pacheco[5], Daniel M. Parker[9,10], Katherine Young[11], Michael Boots[12], Cristina M. Tato[4], Joseph L. DeRisi[4], Christina Yek[3], Jessica E. Manning[3,5]

Affiliations:
[1] Department of Ecology and Evolution, University of Chicago, Chicago, Illinois, USA
[2] Department of Science and Technology Studies, York University, Toronto, Canada
[3] Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland, USA
[4] Chan Zuckerberg Biohub, San Francisco, California, USA
[5] International Center of Excellence in Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Phnom Penh, Cambodia
[6] Chan Zuckerberg Initiative, San Francisco, California, USA
[7] National Center for Parasitology, Entomology, and Malaria Control, Phnom Penh, Cambodia
[8] Center for Computational Biology, University of California, Berkeley, California, USA
[9] Department of Population Health and Disease Prevention, University of California, Irvine, California, USA
[10] Department of Epidemiology and Biostatistics, University of California, Irvine, California, USA
[11] Department of Biological Sciences, University of Texas, El Paso, Texas, USA
[12] Department of Integrative Biology, University of California, Berkeley, California, USA

*Corresponding author: Cara E. Brook
Email: cbrook@uchicago.edu

Author Contributions: CEB and JEM conceived the study. JEM, JAB, HR, RL, and SC carried out the febrile surveillance study and collected samples for mNGS in 2019. VA, CMT, JLD, JAB, SC, SL, CY, OM, SM, and JEM carried out RT-qPCR to identify dengue-positive samples, as well as library preparation and mNGS for raw sequencing. Other ancillary demographic data collection and analyses were performed by RL, KS, CY, FO, and ARP. DP, CL, and JEM set up geospatial system for patients. CEB and JAB generated consensus dengue genomes for upload to NCBI. KY and CEB downloaded climate data. CEB carried out TSIR, FOI, and simulation modeling, as well as all phylogenetic analyses, with help from YL, GN, CR, and MB. CR and CEB extended the catalytic model to include waning multitypic immunity. CEB and JEM wrote the first draft of the manuscript. All authors approved and contributed to subsequent drafts of the manuscript and agree with the results presented.

Competing Interest Statement: No competing interests.

Classification: Biological Sciences / Genetics
Keywords: dengue, genomic epidemiology, force of infection, wavelet decomposition, arbovirus

This PDF file includes:

Main Text
Figures 1 to 5

ABSTRACT

The incidence of dengue virus disease has increased globally across the past half-century, with highest number of cases ever reported in 2019. We analyzed climatological, epidemiological, and phylogenomic data to investigate drivers of two decades of dengue virus disease in Cambodia, an understudied endemic setting. Using epidemiological models fit to a 19-year dataset, we first demonstrate that climate-driven transmission alone is insufficient to explain three epidemics across the time series. We then use wavelet decomposition to highlight enhanced annual and multiannual synchronicity in dengue cycles between provinces in epidemic years, suggesting a role for climate in homogenizing dynamics across space and time. Assuming reported cases correspond to symptomatic secondary infections, we next use an age-structured catalytic model to capture a trend of declining force of dengue infection through time, which drives observed patterns of increasing mean age of reported cases in Cambodia. Reported cases in >70 year-old individuals in epidemic years are best explained when additionally allowing for waning multitypic immunity and repeat symptomatic infections in older cohorts. We support this work with time-resolved phylogenetic trees incorporating 122 dengue virus (DENV) genomes sequenced in the 2019-2020 epidemic, which document introduction of DENV-2 Cosmopolitan genotype into Cambodia, yielding localized transmission and decreased genomic diversity compared to endemic DENV-1. Finally, we simulate an age-structured, mechanistic model of dengue dynamics to demonstrate how introduction of a genetically distinct lineage into a population with waning multitypic immunity could drive repeat infections within a serotype in older-age individuals and recover patterns from reported data.

CLINICAL TRIAL NUMBERS: NCT04034264 and NCT03534245.

SIGNIFICANCE STATEMENT

The year 2019 witnessed the highest number of dengue cases ever reported, including in Cambodia, a Southeast Asian country with endemic transmission. We analyzed two decades of national dengue surveillance data for Cambodia to demonstrate how increasing temperature and precipitation enhance spatiotemporal synchronicity in epidemic years. We document how twenty years of demographic transition has increased the age of reported dengue infection and correspondingly depressed the force of infection through time. We show that introduction of a divergent DENV-2 genotype in Cambodia in 2019 likely drove repeated symptomatic infections in older-age individuals that contributed to a high burden epidemic. As climates warm, we are likely to see more synchronized dynamics globally and a shifting burden of symptomatic disease into older cohorts.

MAIN TEXT

INTRODUCTION

Dengue virus (DENV) transmission has increased dramatically over the past two decades, culminating in the year 2019 with the highest number of global cases ever reported (>5.2 million) to the World Health Organization (WHO) (1). Since nearly three-quarters of DENV infections are estimated to be clinically inapparent (thereby unreported), these counts represent a vast underestimate of the true scale of dengue burden on public health systems (2). DENV is a flavivirus primarily transmitted by the *Aedes aegypti* mosquito, an ubiquitous arthropod vector in tropical and subtropical regions (2). DENV is comprised of four antigenically distinct serotypes—each of which is further subdivided among four to seven distinct genotypes; infection with one serotype is thought to result in lifelong immunity to that same serotype (homotypic immunity) but only temporary (up to two-year) protection against different serotypes (heterotypic immunity) (3). Heterotypic secondary infections are often more clinically severe due to an interaction with pre-existing flavivirus-specific antibodies known as antibody-dependent enhancement (ADE) (4, 5). In regions with multiple circulating endemic flaviviruses, this interplay makes the dynamics of dengue difficult to interpret at local, subnational, or national levels. Dynamical inference is further challenged by the multifaceted disease ecology characteristic of any arbovirus—for which infections are also impacted by changes in human land use, behavior, or movement, as well as the distribution and abundance of arthropod vectors. Nonetheless, national health systems in DENV-endemic regions rely on mathematical models to inform resource allocation on a year-to-year basis—ranging from vector control to hospital capacity for supportive care, given no currently available treatments for dengue. Improved understandings of the drivers of dengue transmission are therefore crucial for public health efforts, especially in resource-scarce settings.

Southeast Asia (SEA) represents ~70% of global reported dengue cases (1). Surprisingly, while dengue is largely classified as an urban disease, only 46% of the SEA population is considered urban. For most SEA countries, roughly 35-55% of the total population is concentrated in urban areas co-localizing with the highest dengue burdens; Singapore reports the largest urbanized population percentage in SEA (100%), while Cambodia—with 16 million people—reports the lowest (24%) (6). Like many SEA countries, Cambodia nonetheless exhibits significant peri-urban sprawl surrounding major metropolitan centers like the capital city, Phnom Penh (7, 8). Dengue was first detected in Cambodia in 1963 (9), though passive surveillance—via reporting of clinically diagnosed cases from public health centers and hospitals to the national level—was not adopted until 1980, following political instability and civil war in the late 1970s (10). In 2001, the National Dengue Control Program (NDCP) was inaugurated in Cambodia, and in 2002, the NDCP formally adopted the WHO clinical case definitions of dengue and its complications as criteria for the surveillance program (10, 11). The NDCP surveillance system is largely limited to clinicosyndromic diagnoses, meaning that reported cases often correspond to more severe heterotypic secondary DENV infections (12), which are reported collectively without the ability to systematically distinguish between serotypes (11). Nonetheless, the NDCP also instituted some active surveillance efforts in Cambodia in 2001, which included limited virological testing to identify distinct serotypes at sentinel sites in four provinces, which were expanded to 15 provinces by 2021 (10, 11).

Over the past two decades, all four dengue serotypes have been detected in Cambodia by virological surveillance (11), though cases have been largely dominated by one or two

serotypes in a given year—with DENV-1 and DENV-3 most common in the early 2000s and DENV-1 and DENV-2 predominant in the last decade (11). Cambodia witnessed three major dengue epidemics across this period—in 2007, 2012, and 2019 (11). The first epidemic, in 2007, occurred coincidentally with a genotype replacement event in DENV-1 (13, 14), though cases were dominated by DENV-3, marking the last year of this serotype's dominance in the region (10, 14). The second epidemic, in 2012, has been largely attributed to DENV-1 (11, 15–17), and the third, in 2019, appears to have been driven by co-circulation of both DENV-1 and DENV-2 (11). Consistent with the global dengue phenomenon of 2019 (1), Cambodia suffered its worst dengue epidemic on record in this year, with approximately 40,000 total cases reported across all 25 provinces—a likely drastic underestimation of the true disease burden (18). Indeed, the past twenty years of surveillance in Cambodia have witnessed, on average, steadily increasing dengue incidence, coupled with a steady increase in the mean age of reported dengue infection (11). The latter has been anecdotally attributed to an aging population following demographic transition, as similar phenomena have been reported previously in neighboring Thailand (19, 20), as well as in Nicaragua (21).

Explosive periodic outbreaks are a hallmark of dengue virus disease, though the multiple drivers of these phenomena have long been debated (20, 22–28). Seasonal climate cycles are a strong predictor of annual cycling for many arboviruses, including dengue (29, 30), and climate has been implicated as a possible driver of multiannual dengue periodicity, as well. In Thailand, multiannual dengue cycles demonstrate coherence with El Niño phenomena (26), and epidemic years exhibit more synchronized dynamics across latitudes (26), as well as higher correlation with local temperature than do inter-epidemic periods (25). The interaction of demography and heterotypic immunity is also thought to play an important role in driving multiannual cycles for dengue (20, 22, 31, 32), which, in Thailand, show elongated periodicity as a result of declining birth rates and slower build-up of the susceptible population over the past half-century (20). Virology also plays a key role in many dengue epidemics, which have, historically, been linked to turnover in the dominant regional serotype (32–34) or to replacement of a dominant viral genotype with a serologically homotypic but phylogenetically divergent viral lineage (14, 35). Indeed, recent work links the magnitude of periodic dengue epidemics to antigenic evolution; large epidemics tend to result from the takeover of those DENV lineages most antigenically distinct from previously circulating strains of the same serotype or most antigenically similar to lineages of a different serotype (35).

The past two decades of changing dengue dynamics in Cambodia have taken place alongside extensive, rapid, heterogenous development across much of SEA, with significant peri-urban sprawl expanding beyond central mega-city limits and high-density population centers springing up around industry hubs (often factories) in otherwise rural areas (7, 8). Longitudinal phylodynamic studies of georeferenced DENV sequences from Thailand have demonstrated the importance of microscale transmission—particularly at the household level—in generating DENV diversity in urban Bangkok, predicting that DENV transmission will intensify as peri-urban settings in SEA become better connected (36). Nonetheless, investigations into the dynamical drivers of dengue transmission in rural and peri-urban settings have been limited to date. As climatic changes drive increases in the population at-risk for dengue infection in SEA and elsewhere (29, 30, 37), better-designed active surveillance programs and improved targeting and evaluation of vector control interventions are greatly needed (38).

Here, we explore the dynamics of dengue virus transmission across the past two decades in peri-urban Cambodia, investigating the potential mechanisms that underly periodic epidemics, particularly the epidemic of 2019. We queried a 19-year dataset (2002-2020) of serotype-agnostic dengue case counts from the NDCP, aggregated at the province level, to, first, interrogate the role that climate played in driving epidemic spikes in 2007, 2012, and 2019, and second, more generally, explore the impact of climate on annual and multiannual cycling across the time series. We, third, fit catalytic models to the age-structured incidence of reported dengue disease to estimate the annual force of infection (FOI), at the province level, across this time series (20, 39, 40), then considered the national surveillance data in the context of our own active febrile surveillance study carried out in peri-urban Kampong Speu province from 2019-2020 (41). Whole genome sequencing of DENV from serum samples collected in part with our active surveillance program identified the first record of the DENV-2 Cosmopolitan lineage ever documented in Cambodia (42). Phylodynamic analysis of the spread of the DENV-2 Cosmopolitan lineage in this region suggests that this new genotype drove pathogenic infections in older age individuals in 2019, contributing to the largest documented dengue outbreak on record for Cambodia. Finally, we constructed an age-structured discrete-time, dynamical model to simulate the interplay of climate, demography, immunology, and virology which combine to structure two decades of dengue dynamics in Cambodia.

RESULTS

Though warmer than average, epidemic years were not major climate anomalies.

Because of the widely acknowledged role of climate as a driver of arboviral disease globally (29, 30), coupled with recent work out of Thailand highlighting coherence between El Niño climate anomalies and dengue epidemics (26), we first explored the influence of changing temperature and precipitation on dengue caseload in Cambodia. To this end, we first aggregated high-resolution temperature and precipitation data at the Cambodian province level, across two-week intervals from 2002-2019 and sought to address the extent to which epidemic years represented climatic anomalies over this time period. We identified a characteristic annual variation in both temperature and precipitation, which was largely recapitulated across provinces and years. Temperature peaked in the first half of each year, between April and July, consistently across all provinces and preceding the peak in dengue caseload (Fig. S1). Precipitation peaked in the latter half of each year, between August and October (Fig. S2). Generalized additive modeling (GAMs) (43) demonstrated that, after controlling for intra-annual variation, temperature has increased significantly across the past two decades in all provinces; no significant interannual changes were detected for precipitation (Fig. S3-S4; Table S1). Additional GAMs and climate data normalized into z-scores indicated that the epidemic years of 2012 and 2019 were hotter than average for the time series, while 2007 was cooler, consistent with the observed interannual increase in temperature (Fig. S5). Years 2015-2016, which spanned a major El Nino event in SEA (but did not correspond to a dengue epidemic) (44), were also hotter than average. By contrast, 2007 had higher-than-average precipitation, while 2019 had lower-than-average precipitation, and 2012 did not deviate from the mean; in the years preceding all three major dengue epidemics, precipitation was not significantly different from average (Fig. S6). We concluded that, while both temperature and dengue incidence increased across our time series, epidemic years were not characterized by any remarkably aberrant climatic profile.

Climate-informed transmission rates failed to recover epidemic dynamics in a TSIR model.

To more mechanistically interrogate the role of climate as a driver of dengue epidemics in Cambodia, we developed a simple, climate-informed time series Susceptible-Infected-Recovered (TSIR) model (45–48), which we fit to dengue case counts, aggregated over two week intervals at the province level, for the three inter-epidemic periods (2002-2006, 2008-2011, and 2013-2018). Though typically used to describe the transmission dynamics of perfectly immunizing childhood infections, such as measles (46, 47), TSIR has been applied to dengue dynamics previously and offers an effective means by which to isolate the impact of climate on transmission, despite oversimplifying the multiple serotype dynamics of dengue virus disease (49–53).

We used biweekly transmission rates recovered from TSIR fits to the inter-epidemic periods, as well as climate-informed transmission rates for epidemic years projected at the province level from lagged temperature and precipitation data, to predict epidemic-year cases (Fig. 1; *Methods; SI Appendix*). We found that TSIR successfully recaptured the timing of annual dengue epidemics across the 22 provinces considered, with the recovered transmission rate peaking between May and August, slightly preceding reported cases. The magnitude and timing of transmission varied by province and between the three inter-epidemic periods, showing no consistent pattern of directional change in magnitude with time (Table S2). TSIR-estimated transmission was significantly positively associated with higher temperature and precipitation (lagged, respectively, such that climate variables preceded transmission by a median 3.5 months for temperature and 1 month for precipitation) in the corresponding province across all inter-epidemic periods (Table S3-S4). More rapid transmission gains were observed for corresponding increases in temperature vs. precipitation (Fig. S7-S9; Table S4). Nonetheless, climate-informed transmission rates for epidemic years were not substantially different from rates fitted to inter-epidemic periods, reflecting the absence of major climate anomalies across the time series (Fig. S10).

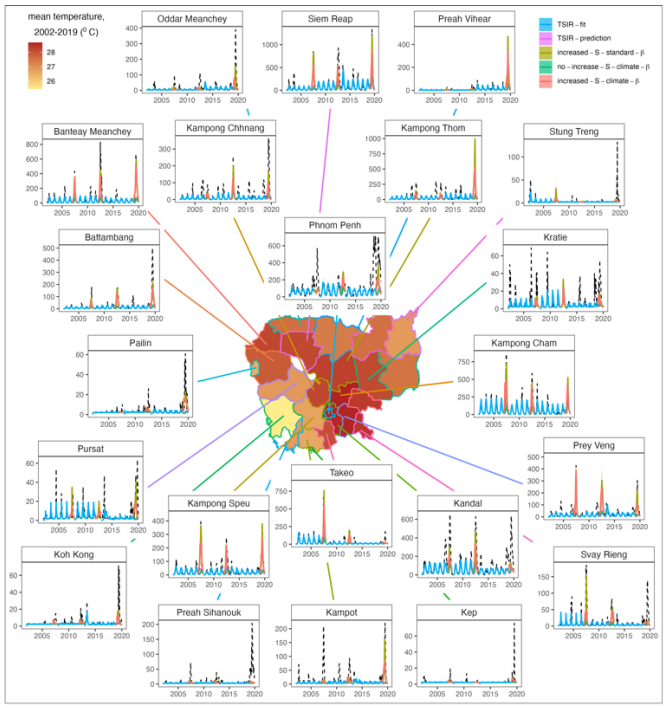


Figure 1. Climate-informed TSIR insights into epidemic dynamics of DENV in Cambodia. Inset panels show

province level clinicosyndromic reported DENV cases (*dashed black lines*) with fitted TSIR output to the three inter-epidemic periods (2002-2006, 2008-2011, 2013-2018; *blue lines*) and TSIR projections for epidemic years (2007, 2012, 2019) under trained biweekly transmission rate (β) estimates (*pink lines*), incorporating a factorial increase in the susceptible population (*gold lines*), using a climate-projected β estimated from lagged temperature and precipitation data by province (*green lines*), or, finally, using both climate-projected β and a factorial increase in the susceptible population (*red lines*). Parameter estimates are available for viewing in *SI Appendix,* Table S2, S5. The center map shows province level administrative boundaries for Cambodia, shaded by the mean biweekly temperature from the 2020 climate data.

Despite recovering the intra-annual timing of transmission, as expected, TSIR underpredicted the magnitude of epidemic year caseloads for all provinces (Fig. 1). Climate-informed transmission rates improved TSIR projections but still failed to recover epidemic peaks (Fig. 1). Inclusion of an amplification term in the susceptible population at the beginning of each epidemic year facilitated TSIR's recovery of the epidemic year caseloads (*Methods*) (53). In the absence of climate input into the transmission rate, the proportional increase in the susceptible population needed to recover the epidemic peak was a median 1.8x for 2007, 2.10x for 2012, and 2.17x for 2019 (Table S5). The proportional increase in the susceptible population needed to recover the epidemic peak was not significantly lower when epidemic year cases were projected using climate-informed transmission rates for 2007 and 2012 (paired student's t-test: [2007] $t=0.33$, $p>0.1$; [2012] $t=-0.52$;$p>0.1$) and only marginally lower for 2019 ($t=1.31$, $p=0.1$), suggesting (at most) a minimal role for temperature and precipitation in driving epidemic dynamics (Table S5). For many provinces, the combination of climate-informed transmission rate with susceptible augmentation still fell short of effectively capturing epidemic caseloads, highlighting the need for alternative explanations for these high transmission years.

Wavelet analysis showed enhanced synchrony in dengue dynamics across provinces and climate time series in epidemic years.

Consistent with previous reports out of neighboring Thailand (26), wavelet decomposition of the biweekly dengue time series, at the province level, demonstrated statistically significant peaks in the amplitude and average wavelet power, of, respectively, both annual (Fig. 2A, S11A) and multiannual (Fig. 2B, S11B) cycles in dengue incidence during the three epidemic years (54). As also witnessed in Thailand (26), we observed significantly elevated synchronicity between provinces in the timing and amplitude of yearly dengue incidence, as measured both by pairwise Pearson's correlation coefficient (Fig. 2C) and cross-wavelet power (Fig. S11), in epidemic vs. non-epidemic years. Synchronicity in the annual incidence data was negatively associated with geographic distance between provinces, though patterns were less clear than have been previously described in neighboring regions (26) (Fig. S12). High synchronicity between annual case data for paired provinces was significantly positively associated with high temperature, precipitation, and population size of a focal province (Fig. S12; Table S6).

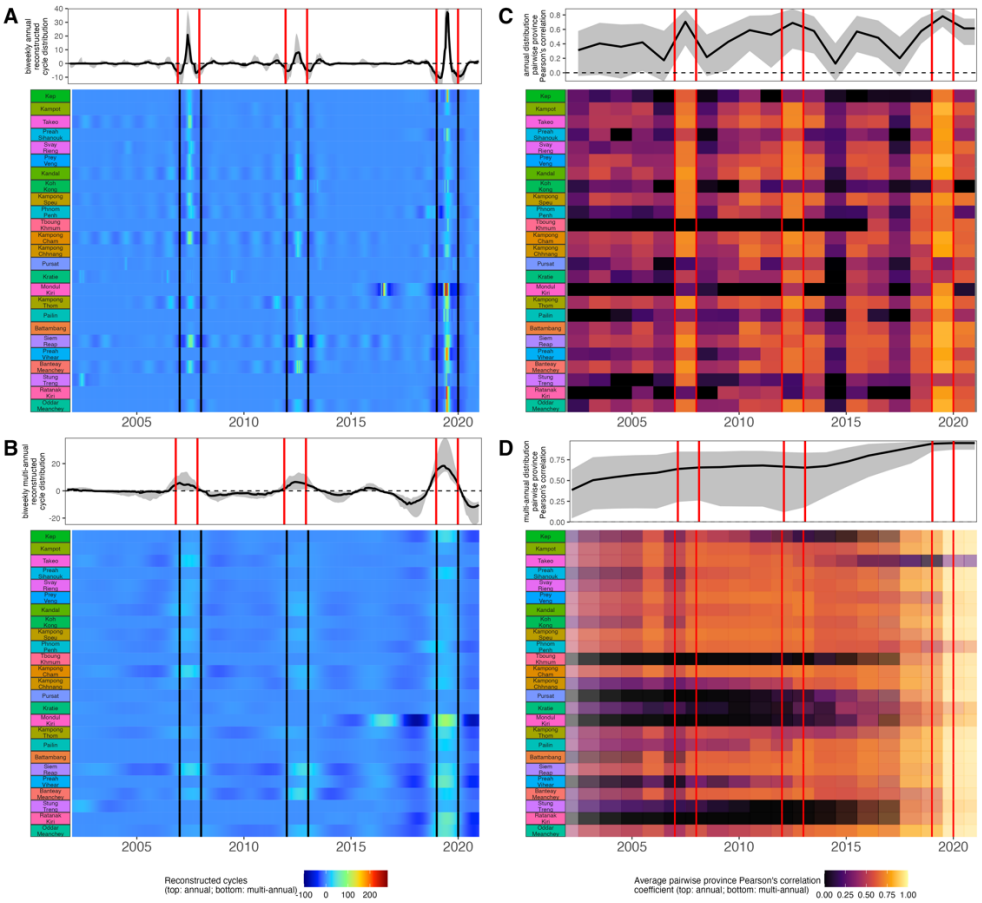


Figure 2. Wavelet reconstructions show heightened synchrony in epidemic years for Cambodian dengue. Reconstructed **A** annual and **B** multiannual dengue cycles by province, by year from NDCP incidence per 100,000 population. **C** Mean pairwise Pearson's correlation coefficient (ρ) for annual dengue incidence between focal province and all other provinces through time. **D** Mean ρ comparing province-to-province reconstructed multiannual dengue cycles across a 5-year sliding window, with overlapping window frames plotted (partially translucent) atop one another. In all panels, epidemic years are highlighted by vertical red or black bars. Top panels indicate the distribution of corresponding values (median = solid line; max-to-min range = gray shading) observed across all provinces within each timestep. X-axis labels are marked on January 1 of the corresponding year.

For multiannual cycles, we observed a trend of steadily increasing synchronicity (Fig. 2D) and cross-wavelet power (Fig. S11D) between provinces through time, though no peaks occurred in epidemic years. Again consistent with recent work from Thailand (25), cross-wavelet power between raw dengue incidence and the corresponding, province level mean temperature and total precipitation (grouped biweekly) peaked in epidemic years (Fig. S13 AB). At the multiannual scale, cross-wavelet power between province level reconstructed dengue cycles and the time series of temperature and precipitation largely increased across the time series (Fig. S13CD), though values were highest slightly preceding the 2019 epidemic and overlapping the 2015-2016 El Niño event (44). The same pattern was observed when comparing monthly reconstructed multiannual dengue cycles and the Oceanic Niño Incidence (ONI), the monthly measure of El Niño activity (Fig. S13E). For temperature, precipitation, and ONI, cross-wavelet power with multiannual dengue cycles appeared to peak earlier in more southern provinces and move gradually northward over several years (Fig. S13DE). All told, wavelet analyses suggested a role

for climate in synchronizing annual dengue epidemics across provinces, but only minimal support for a hypothesis of multiannual climate cycles driving periodic epidemics across our time series. In contrast to reports from two decades prior in Thailand (20), the periodicity of multiannual dengue cycles across provinces also did not increase in duration throughout our time series, despite decreasing birth rates over the same period (Fig. S14). This perhaps reflects the relative slowdown in both birth and death rate declines in Cambodia over the past two decades, when compared with the dramatic declines witnessed twenty years prior (Fig. S14).

Mean age of reported dengue infection increased across the study period, corresponding to a declining force of infection.

Previous work has reported a trend of increasing mean age of reported dengue infection across the past two decades in Cambodia at the national level (11); we confirmed this to be consistent at the province level, as well (Fig. 3A,B; Fig. S15; Table S7). At the national level, the mean age of reported dengue infection increased from 6.79 years (95% CI: 5.85-7.72) in 2002 to 10.34 years (95% CI: 9.41-11.27) in 2020 ($p<0.001$). Increases in age of reported cases were even more severe in more remote provinces: in the distant northeastern province of Mondul Kiri, for example, the mean reported age of infection increased from 3.72 years (95% CI: 1.97-5.47) in 2002 to 21.18 years (95% CI: 20.84-21.53) in 2020 (Fig. 3A; Fig. S15; Table S7).

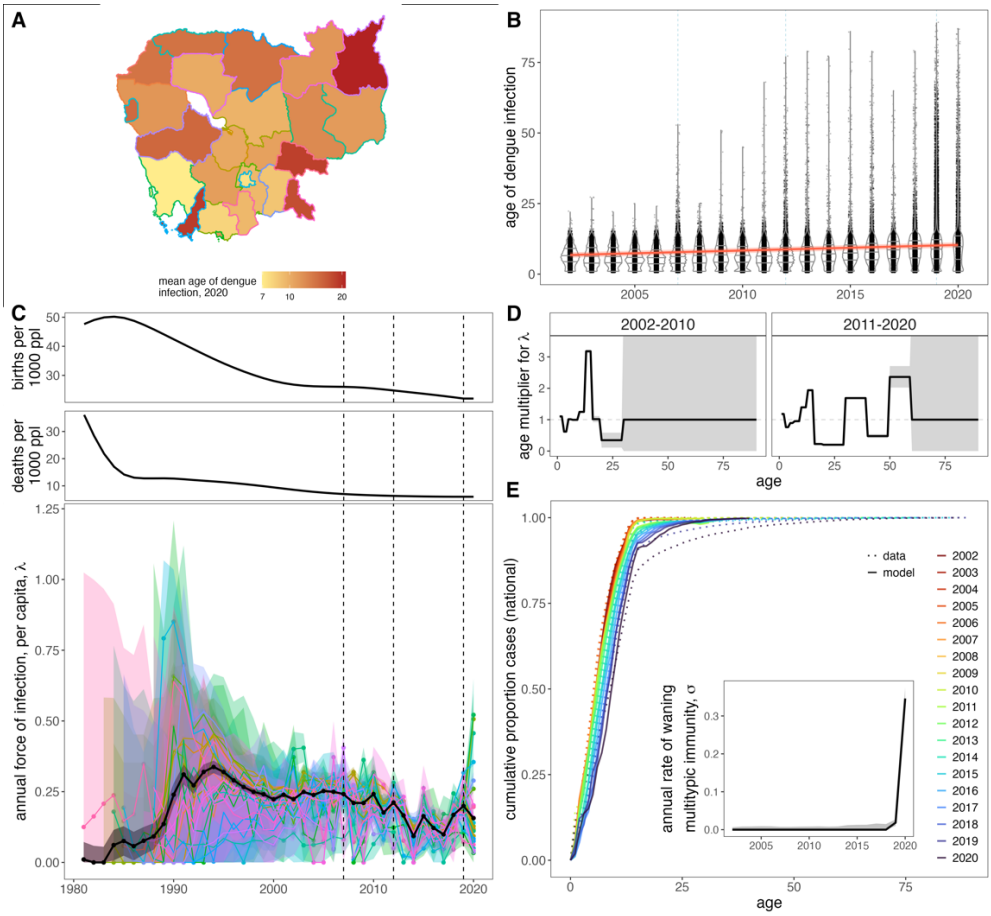


Figure 3. Demographic transition underpins declining force of infection and increasing age of reported dengue incidence in Cambodia. A Mean age of reported dengue infection, by province in the last year of the NDCP time series

(2020). **B** Age distribution of reported dengue cases by year, with violin plots highlighting changes in the interquartile range of cases. The interannual trend in the mean age of dengue infection is plotted as a solid red line, with 95% confidence intervals by standard error shown as a narrow, translucent band behind it (Table S7). Epidemic years (2007, 2012, 2019) are highlighted by a light blue, dashed line in the background. **C** National (black) and province level (colored) estimates for the annual force of infection from 1981 (the birth year of the oldest individual in the first year of the NDCP time series) to 2020. 95% confidence intervals from the hessian matrix are shown as translucent shading. FOI estimates are compared against national birth and death rates for Cambodia across the time series, with epidemic years highlighted by vertical dashed lines. **D** Age modifiers to the FOI fit as shared across all provinces for 2002-2010 and 2011-2020 subsets of the data, with 95% confidence intervals by profile likelihood shown as translucent shading (note that the likelihood profile was flat in age categories 30+ for years 2002-2010 and 60+ for years 2011-2020 so no FOI modification was applied here). **E** Cumulative increase in the proportion of cases reported by age at the national level, colored by year. Data are shown as dotted lines and model output as solid lines. Model includes national FOI estimates from **C**, age modification terms from **D**, and time-varying waning multitypic immunity as shown in the inset.

To interrogate the mechanistic drivers of this pattern, we estimated the annual force of infection (FOI, or λ), the rate at which susceptibles become infected, by fitting a catalytic model with multiple serotype exposures to the province level data for every year in the dataset (2002-2020) and the 22 years preceding the onset of the time series, dating back to the birth year of the oldest individual in the first year of the data (Fig. 3C) (20, 39, 40). As in previous models for dengue (20, 39), we assumed that reported cases represented secondary, symptomatic infections more likely to report to public hospitals and clinics. Resulting patterns in FOI by province largely mirrored those recovered at the national level (Fig. S16), demonstrating a consistently high FOI in the early 1990s, roughly corresponding to the years in which individuals born in Cambodia's 1980s birth pulse (which followed severe population reductions resulting from civil war in the late 1970s) would likely be experiencing secondary infections. The FOI subsequently demonstrated a gradual decline across the time series, interspersed with minor local peaks during epidemic years.

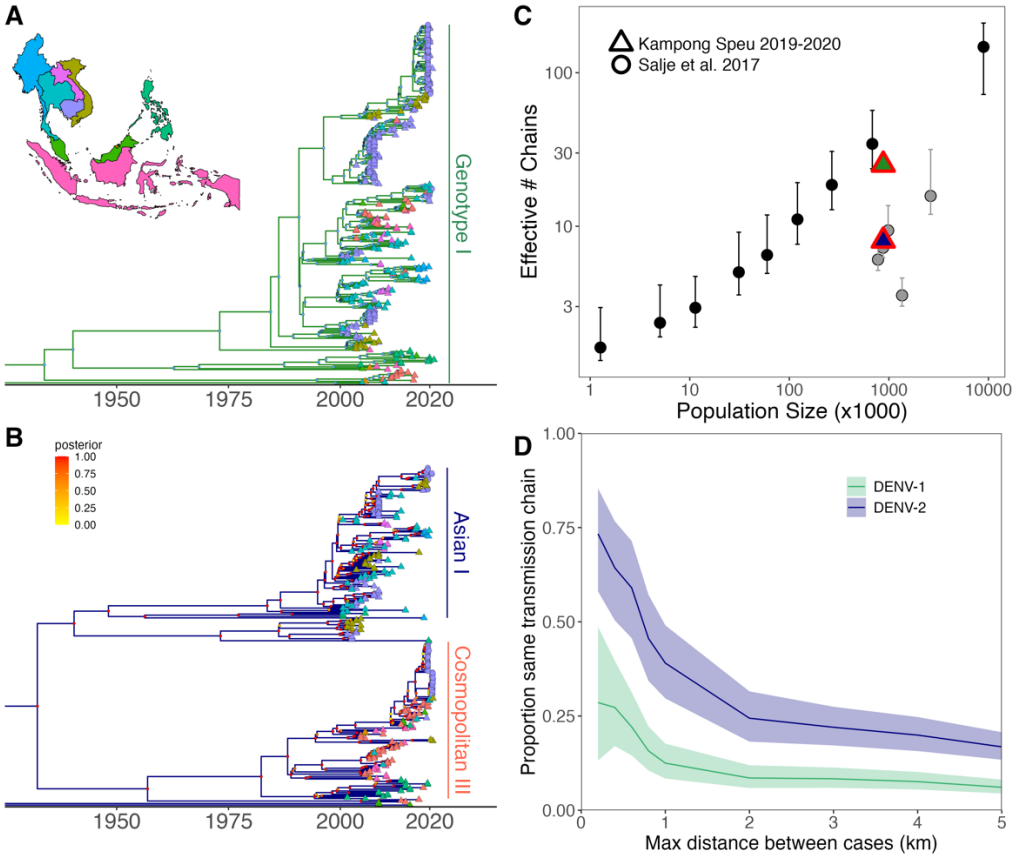
Age-structured FOI modifications and waning multitypic immunity in 2019-2020 improved the model's ability to recapture observed data.

Inspired by previous studies in other systems (19–21), we next fit 26 (thirteen per decade) multiplicative age modification parameters, shared across all provinces and a subset of years, to allow for modulation of annual FOI across individuals in different age categories (Fig. 3D). Incorporation of age modifiers improved model fits to the data across all provinces (Table S8). Consistent with recent work in Thailand (19), we identified a high hazard of infection in adolescents (13-15 year-olds) across the time series. In the second decade of the time series, we additionally noted an elevated age-specific hazard of infection in 30-39 year-olds and 50-59 year-olds, corresponding to reported symptomatic infections in older age individuals.

This support for elevated infection rates in older age individuals led us to revisit our assumption that reported cases corresponded to clinically apparent secondary infections only. Visualization of the age distribution of reported cases by year at both the national (Fig. 3B) and province levels (Fig. S15) indicated that, in addition to increases in the mean reported age of dengue infection, the past two decades of dengue incidence in Cambodia have also witnessed expansion in the age *range* of reported cases—such that approximately 1% of reported cases (712/68597) in 2019 occurred in individuals over the age of 45, with 61 infections reported in individuals >70 years in age. We also observed an isolated spike in the age distribution of reported cases in some provinces in the earlier epidemic years of 2007 and 2012 (Fig. S15). To allow for symptomatic infections in older age-individuals, we modified the two-serotype catalytic modeling framework presented in previous work (20, 21, 39) to permit individuals to wane from a

384 state of multitypic immunity back to that of monotypic immunity subject to renewed hazard of
385 symptomatic re-infection (*SI Appendix*). Using this new modeling framework, we estimated a
386 shared parameter across all provinces, which corresponded to the rate of waning multitypic
387 immunity (σ). No signature of waning multitypic immunity was observed across the time series as
388 a whole; however, when we allowed σ to vary by year, we estimated a significant signature of
389 waning multitypic immunity in 2019 and 2020 data only (Fig. 3E, *inset*). In general, incorporating
390 σ improved model fits to the data when added both to the FOI-only model and the FOI model
391 modified by age class (Table S8). Our best fit FOI model including both age modifications and
392 time-varying multitypic waning immunity effectively recaptured the observed cumulative
393 proportion of cases per age class per year at national (Fig. 3E) and province (Fig. S17) levels. In
394 general, cases accumulated more slowly across age classes with advancing years, tracking
395 declining FOI.
396
397
398 **DENV genotyping in Kampong Speu province from 2019-2020 suggests a possible clade**
399 **replacement event in DENV-2.**
400 The high burden of pathogenic cases in older individuals in 2019 next led us to ask
401 whether the virology of the DENV strains associated with the epidemic could play a role in driving
402 observed patterns in the data. To address this question, we leveraged serum samples amassed
403 during an active febrile surveillance study that we carried out in Kampong Speu province,
404 adjacent to the Cambodian capital city of Phnom Penh, between July 2018 and December 2020
405 (41, 42). In this study, we screened 760 samples collected from 697 unique individuals who
406 reported to participating sites with fever during the study period, identifying 123 PCR-positive
407 DENV infections. All patients reported within a self-identified window of five days since fever
408 onset. From this data subset of positive DENV cases, we sequenced and published 122 whole
409 DENV genomes (57 DENV-1, 61 DENV-2, and 4 DENV-4; Table S9), representing one-third of
410 full genome DENV-1 sequences and over half of DENV-2 sequences currently available for
411 Cambodia—thus emphasizing the considerable challenges faced in undertaking genomic
412 epidemiology in resource-scarce settings (42). Maximum likelihood phylogenetic analysis of the
413 resulting genomes demonstrated that most DENV-2 sequences recovered in 2019 and 2020
414 belonged to the DENV-2 Cosmopolitan III lineage, the first record of this genotype reported in
415 Cambodia, though its introduction has been reported recently in several other neighboring
416 countries in SEA (55–58). All DENV-1 sequences clustered in the Genotype 1 lineage, consistent
417 with previously reported genotype records in Cambodia (Fig. S18). In our study, sequences
418 collected from 2019 were largely split between DENV-1 (N=52) and DENV-2 (N=35) serotypes
419 (59.8% vs. 40.2% respectively), while sequences collected in 2020 were dominated by DENV-2
420 (N=26/31, 83.9 %). In 2019 and 2020, respectively, 26/35 (74.3%) and 24/26 (92.3%) of DENV-2
421 sequences belonged to the Cosmopolitan III lineage, with the remaining sequences clustering in
422 the Asian-1 DENV-2 lineage previously reported in Cambodia.
423 We constructed serotype-specific Bayesian timetrees (59, 60) from DENV sequences to
424 assess the divergence time of 2019-2020 lineages from sequences previously reported from
425 Cambodia and from neighboring SEA countries across our 2002-2020 study period (Fig. 4; Table
426 S9). The majority of DENV-1 lineages (detailed in Fig. S18) diverged relatively recently from 2015
427 and 2016 Cambodian sequences last reported within this serotype (Fig. 4A), with a time to Most
428 Recent Common Ancestor (tMRCA) of approximately 7.9 years (MRCA at February 2013, 95%

429 HPD: July 2012 – September 2013). BLAST analysis indicated that this cluster of DENV-1
430 Cambodia sequences demonstrated the highest identity to sequences recovered from China in
431 2019, while another, rarer subset of DENV-1 sequences more closely resembled those recovered
432 from Thailand in 2019 and 2020 (61). By contrast, DENV-2 sequences in the Cosmopolitan III
433 lineage demonstrated high divergence from previously reported sequences for Cambodia (which
434 have been recorded only in the Asian-1 lineage), with a tMRCA of approximately 88.5 years
435 (MRCA at 1932, 95% HPD: 1927-1937). BLAST analysis indicated that DENV-2 Cosmopolitan
436 sequences recovered from Cambodia in 2019 and 2020 showed the highest similarity to
437 sequences previously derived from recent DENV-2 Cosmopolitan outbreaks in Singapore,
438 Malaysia, Sri Lanka, and Thailand (53, 57, 61). Locally, within our regional data subset for SEA,
439 sequences tightly clustered in both geographic space and time were highly phylogenetically
440 related (Fig. S19), consistent with previous studies emphasizing the importance of microscale
441 transmission dynamics for DENV (36).
442



443
444 **Figure 4. Bayesian time trees highlight geospatial structuring in evolutionary relationships for Cambodian**
445 **dengue. A** Map of Southeast Asia with countries colored corresponding to sequences derived from each country, as
446 shown in tip points on phylogenetic timetrees constructed using BEAST 2 for DENV-1 and **B** DENV-2. X-axis highlights
447 divergence times between corresponding sequences. Reference sequences from GenBank are represented as triangle
448 tips and sequences contributed by active febrile surveillance in this study as circles. Cambodia and corresponding
449 sequences are shaded purple. Clade bars indicate the genotype of corresponding sequences within each serotype:
450 genotype-1 for DENV-1 and Asian-1 and Cosmopolitan III for DENV-2. A detailed inset of geographic localities for 2019-
451 2020 Cambodia sequences can be viewed in Fig. S19. **C** Number of effective transmission chains for circulating DENV
452 estimated across populations of varying densities. Black (urban) and gray (rural) circles with corresponding 95%
453 confidence intervals depict estimates for Thailand from Salje et al. 2017 (36), while triangles depict estimates from our
454 Kampong Speu active febrile surveillance study for DENV-1 (*green*) and DENV-2 (*blue*). **D** Proportion of sequence pairs

for Kampong Speu DENV-1 (*green*) and DENV-2 (*blue*) genomes which are derived from the same transmission chain (*i.e.* share a MRCA <6 months since the earlier sequence in the pair) across progressively longer Euclidean distances separating the localities from which the sequences were recovered.

Prior genomic analysis of geolocated DENV sequences in Thailand has demonstrated the importance of microscale, household-level transmission in driving DENV dynamics, particularly in the highly urbanized setting of Bangkok (36). In Bangkok, human population size was shown to be positively correlated with the number of distinct DENV transmission chains (a measure of phylogenetic diversity) circulating in the region (Fig. 4C). In more rural settings in Thailand, this relationship was less pronounced, but the authors predicted that increasing urbanization of these regions would drive corresponding increases in DENV diversity in the future. We compared patterns in DENV genomic diversity captured in geolocated sequences from 2019 and 2020 in peri-urban Cambodia with those previously investigated in Thailand (Fig. 4C), building transmission chains from sequence pairs which shared a MRCA within the past six months. The number of transmission chains recovered for endemic DENV-1 in our study region closely matched that predicted by population size in urban Bangkok. For DENV-2, which showed less overall diversity than DENV-1, the lower number of transmission chains per population size better approximated that recovered from rural regions in Thailand, in keeping with a hypothesis of recent invasion (Fig. 4C). We further demonstrated a tighter coupling between phylogenetic relatedness and the physical, geographic distance between paired sequences for DENV-2 vs. DENV-1 sequences, consistent with epidemic invasion behavior and the tight clustering of cases in space and time (Fig. 4D).

Mechanistic simulations of clade replacement in an age-structured model captured patterns qualitatively similar to those seen in the data.

To consolidate findings from TSIR and FOI analysis with insights gleaned from genomic epidemiology in 2019-2020, we constructed a mechanistic, age-structured discrete time deterministic epidemic model in biweekly timesteps (*SI Appendix*) (62–64) to simulate two- and three-serotype dengue infections in a population demographically structured to mimic that of Cambodia over the past half-century. We initiated simulations incorporating seasonal variation in transmission as estimated by TSIR (Fig. 1) and annual variation in FOI corresponding to the time series of national estimates (Fig. 3), then modified dynamics to test hypotheses of potential drivers underlying the spike in cases and corresponding expansion in age structure witnessed in the three epidemics identified in the NDCP data (Fig. 5). As with our FOI analysis, we modeled secondary infections as reported cases, except when testing hypotheses of tertiary case detection or waning multitypic immunity and reinfection.

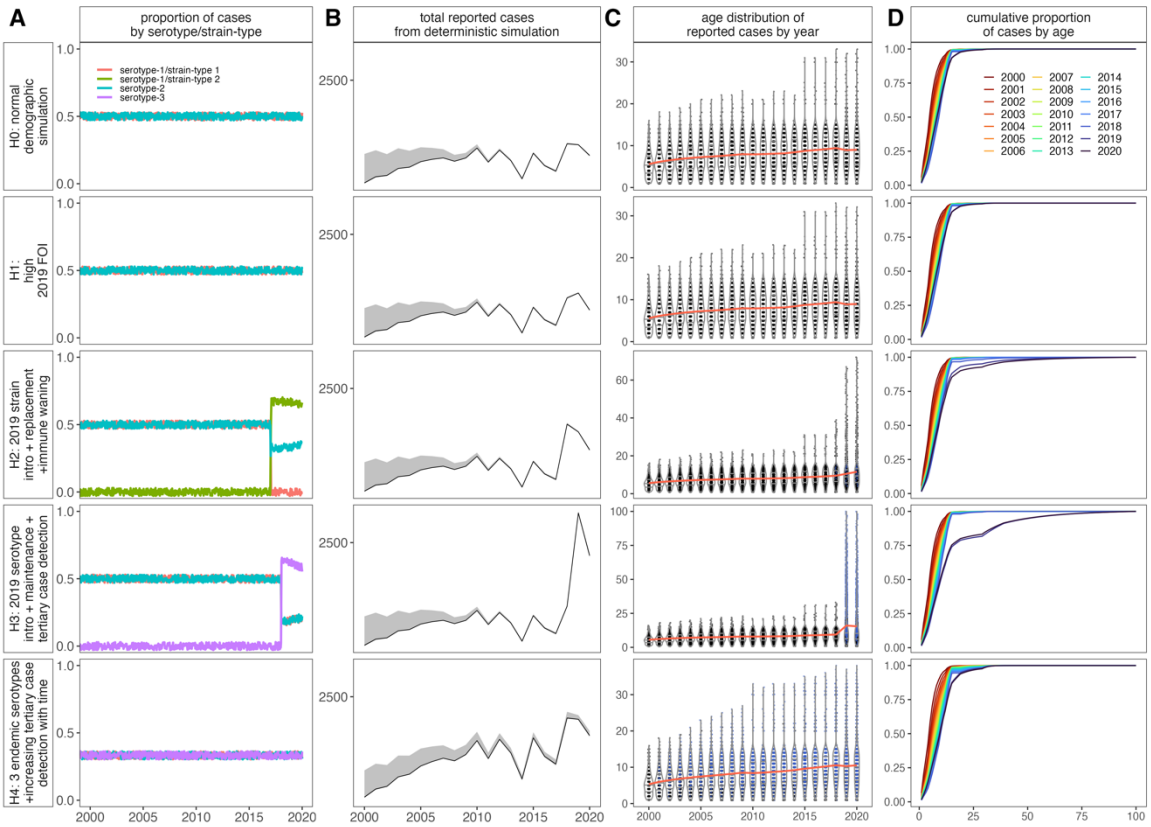


Figure 5. Simulations of genotype or serotype introduction events recapitulate observed variation in the age structure of reported dengue cases from Cambodian data. A Structure of deterministic simulation outlining hypotheses of possible mechanisms underlying patterns in observed data for Cambodia: (H0) normal demographic simulation, (H1) 2-circulating serotypes with a climate-driven FOI spike in epidemic year 2019, (H2) novel genotype invasion and clade replacement with waning immunity in 2019, (H3) novel serotype invasion and 10% tertiary case detection in 2019, (H4) 3 endemic circulating serotypes and progressively increasing rates of tertiary case detection (from 0 to 100%) across time series. **B** Total reported cases (solid line = mean FOI; translucent shading = 95% confidence interval for FOI), **C** age distribution of reported cases by year (black = secondary; blue = tertiary), and **D** cumulative proportion of reported cases by age, by year from each corresponding simulation.

For simplicity, we focused our analysis on drivers of the 2019 epidemic, simulating scenarios by which 2-circulating serotype dynamics were interrupted by a climate-driven spike in FOI in the epidemic year (H1), a genotype invasion and clade-replacement event with waning immunity within the serotype (H2), or a third serotype invasion (H3), then compared these to a hypothesis of 3-serotype circulation and increasing detectability of tertiary infections through time (H4) (Fig. 5A). We found that all hypotheses could recover high overall caseloads in epidemic years (Fig. 5B), though climate-driven increase in FOI (H1) did little to alter the mean age or age distribution of reported infection beyond baseline (Fig. 5C). By contrast, H2, H3, and H4 all elevated the mean age of infection, expanded the age distribution of cases, and produced a more gradual proportional accumulation of cases by age (Fig. 5D), in keeping with patterns in the observed data. Impacts on the age distribution of reported cases were less extreme for our H4 simulations but could be easily modified by varying the rate of increase in the proportion of tertiary cases detected with time, which we did not estimate from the data. To highlight the link between the underlying simulated transmission dynamics and our prior inference from the

observed data in Fig. 3, we additionally fit our two-serotype catalytic FOI model with waning immunity to the simulated data generated for each hypothesis. Using this approach, we successfully recovered the input FOI in all simulations (Fig. S20), and the model structure estimated a signature of waning immunity for H2 and H3 simulations, which produced simulated data that closely resembled the true data (e.g. elevated mean age of reported infection and expansion of age distribution of reported cases in epidemic year).

Though we focused simulations on hypothesized drivers of the 2019 epidemic, we also repeated analyses with H1, H2, and H3 alterations introduced in 2007 to highlight their downstream impacts on the rest of the time series (Fig. S21). By definition, H4 did not define a change in a specific year. High epidemic FOI (H1) simulations for 2007 recovered the same spike in cases but had minimal impact on the age of reported infection, while both hypotheses of genotype invasion and clade replacement with waning immunity (H2) and third serotype invasion (H3) recovered a pattern consistent with the observed data: that of increasing mean and distribution of reported infection age concentrated around the epidemic year, which returned to baseline following the perturbation. Age increases were more extreme for H3, despite being suppressed by the increased overall FOI resulting from the introduction of a third serotype with equivalent FOI into the system. To date, no evidence of serotype invasion or turnover has been presented for the 2019 Cambodian epidemic, though serotype exchange is thought to have played a role in the 2007 epidemic (10, 14). Genotype invasion and clade replacement events within a serotype have been described in Cambodia in 2007 (13, 14) and 2019 (this paper).

DISCUSSION

Drawing from two decades of national surveillance data and a more recent genomic cohort study, we queried the mechanisms that underly dengue virus transmission and drive periodic dengue epidemics in Cambodia. All told, our study highlights the complex interplay of climate, demography, immunology, and virology that dictates the dynamics of dengue virus disease in endemic settings.

Our investigations of climate effects on dengue transmission in Cambodia mirror those previously reported in Thailand (25, 26), Sri Lanka (53), and China (49), emphasizing the role of temperature, and—to a lesser extent—precipitation, in synchronizing annual dynamics in epidemic years. Consistent with prior work that highlights a role for El Niño in driving multiannual dengue cycles in other systems (26, 27, 65, 66), we observed synchrony between reconstructed multiannual dengue cycles in Cambodia and the ONI; however, this synchrony appeared to peak during the robust 2015-2016 El Niño, which did not correspond to one of the three major dengue epidemic years in our dataset. The muted impact of El Niño on dengue dynamics in Cambodia could reflect an increased homogenization of caseloads at higher temperatures in more recent years (25)—which is likely to mask more subtle effects on transmission—or may also be a consequence of less pronounced latitudinal variation (and corresponding climate differences) in Cambodia, as compared to other countries where these phenomena have been studied (e.g. Thailand). Nonetheless, our analysis suggests that warmer temperatures do play a role in driving epidemic dynamics for Cambodian dengue, and likely contributed to high caseloads in 2019. As temperatures increase globally, dengue transmission is likely to accelerate, and variation in both the timing of annual dengue transmission and multiannual peaks in caseload may become more homogenous.

In addition to climate, our data and corresponding analysis emphasizes the importance of human demography and prior immunity in structuring dengue transmission globally. Consistent with studies conducted elsewhere (19–21), we observed a pattern of increasing mean age of reported dengue infection that we effectively recaptured by modeling a declining force of infection through time—which we attributed to declining birth and death rates consistent with Cambodia's demographic transition. Certainly, other mechanisms of demographic change (e.g. immigration and emigration) could also impact infection dynamics. Cambodia experienced a net negative migration rate across the time series investigated (67); as the majority of Cambodian migrants are young adults who seek employment in neighboring Thailand (68), it is possible that their emigration could lower the force of dengue infection even further. However, the Cambodian migrant population comprises <0.5% of the total national population (67), suggesting that birth and death rates are still likely the dominant drivers of observed patterns in the data. Nonetheless, migration rates can be 10-fold higher in border provinces and could have correspondingly elevated impacts in these regions (68).

The declining force of infection for dengue through time is somewhat counterintuitive considering recent explosive epidemics in Cambodia and across the globe. Nonetheless, we provide evidence for a high frequency of reported cases in older age individuals that underlies the high case burden in epidemic years. Our study is unique from previous investigations in recognition, not only of escalation of the mean age of reported dengue infection, but also of the expanding age range of reported disease. Indeed, symptomatic individuals > 70 years in age during the 2019 epidemic are difficult to reconcile under assumptions by which reported cases correspond to clinically apparent secondary infections only. In our modeling framework, we take a novel approach to account for these infections by allowing for waning multitypic immunity and repeated symptomatic infections in older age individuals.

Recent analyses out of the Thailand system highlight a similarly surprising uptick in case reports among older individuals in more recent years (19), which the authors attribute to increased detectability in tertiary and quaternary infections with time. The authors offer a series of hypothetical explanations for increasing case detectability across their time series: that tertiary and quaternary dengue infections might be more pathogenic (and therefore more detectable) due to an abundance of comorbidities in older individuals, that immunopathology may be exacerbated in older patients who experienced longer durations between repeat infections; or that waning multitypic immunity could allow for repeat infections in the oldest age cohorts (our hypothesis). Our analyses in the Cambodia system are not mutually exclusive with any of these explanations; however, we hypothesize that, were comorbidities or immunopathogenesis in older individuals driving patterns in the observed data, we would expect to see an hour-glass shape in symptomatic cases, with reduced reporting in middle-aged individuals who have progressed beyond secondary exposures but are at lower risk for both comorbidities and immunopathology. Instead, we see a gradual tapering in the age-frequency of infection, which expands in range with time (Fig. S15), consistent with the hypothesis of waning multitypic immunity. This latter hypothesis, when considered in light of a genotype replacement event, can also explain isolated expansion in the age range of reported cases witnessed in conjunction with epidemic year 2007 and 2012 in our time series, while the first two hypotheses cannot.

Our deterministic model simulations demonstrate how short-term (1–3 year) spikes in the age distribution of reported cases, independent of the mean annual trend in the time series, can be achieved both via introduction of a novel genotype with waning immunity or via introduction of

606 a novel tertiary serotype into an endemic, two-serotype system (equally plausible would be
607 introduction of a quaternary serotype into an endemic three-serotype system). Though limited in
608 scope, PCR-based sentinel surveillance data provide no support for serotype turnover associated
609 with the 2019 dengue epidemic in Cambodia (11). By contrast, genome sequencing data from our
610 own febrile cohort study in Kampong Speu province offer support for a genotype replacement
611 event. Recent analyses of the Thailand dengue system link transitions in serotype dominance
612 and clade replacement of genotype sublineages within a single serotype to epidemic magnitude
613 (35). This work suggests that large epidemics typically result from the invasion, expansion, and
614 evolution of viral lineages that diverge from previously circulating lineages of the same serotype
615 but are more closely related to co-circulating viruses of endemic heterotypic serotypes, reflecting
616 the viral fitness advantage afforded by ADE (31). In Thailand, the most severe epidemics were
617 linked to clade replacement of a resident viral genotype by an evolutionarily fitter DENV lineage
618 within the same serotype, resulting in a 'selective sweep', which subsequently reduced overall
619 viral diversity and, consequently, diminished between-serotype antigenic differences (35). In the
620 2019 epidemic in Cambodia, the Cosmopolitan genotype introduction, followed by serotypic and
621 genotypic homogeneity in 2020, are consistent with the dynamics of clade replacement within the
622 DENV-2 serotype. Indeed, clade replacement dynamics have been previously witnessed in
623 conjunction with epidemic outbreaks in Cambodia (13, 14) and elsewhere (53, 57, 69–71).
624 As one major limitation of this study, we did not obtain virus isolates and undertake
625 antigenic cartography (72), which would be needed to resolve whether the invading Cosmopolitan
626 III DENV-2 genotype in Cambodia in 2019 was, indeed, antigenically distinct from previously
627 circulating Asian-1 lineages and, potentially, more antigenically similar to endemic heterotypic
628 viruses. Indeed, because much of our inference is derived from serotype-agnostic national
629 surveillance data, we are unable to rigorously evaluate the role of a possible genotype
630 replacement event in particularly driving a high burden of disease in the oldest-age (>70 years)
631 individuals in the dataset. Unfortunately, our febrile cohort was derived from an ongoing childhood
632 cohort study (41), largely limiting the resulting sequence data to younger cases. Genotyped
633 infections in the oldest-age individuals would do much to illuminate these hypotheses—and could
634 provide even more insights if paired with serotyping prior to the 2019 epidemic.
635 Several published studies offer regional explanations for the global dengue phenomenon
636 of 2019. Recent analyses from Brazil, for example, argue that a low FOI in 2017 and 2018
637 resulting from new public health interventions and behavioral modifications implemented in the
638 wake of the Zika virus epidemic, drove a resurgence in cases in 2019 and an expansion of
639 specific lineages of DENV-1 and DENV-2 that had been circulating cryptically for much of the
640 past decade (73). Independent work in the same region indicates that the DENV-2 lineage
641 responsible for the Brazilian outbreak additionally caused a clade replacement event within the
642 serotype (70). Reports from other 2019 dengue epidemics in Bhutan and Bangladesh point to
643 increased monsoon activity coincident with mass movement of people for school and religious
644 holidays as key drivers of the surge in South Asia (74, 75)–while others still highlight a role for
645 reintroduction of previously extinct serotypes, combined with high rainfall and insecticide
646 resistance in these same regions (76). Notably, these latter studies lack data on both host
647 susceptibility and viral phylogenetics needed to query the hypotheses presented here. Dengue
648 surged worldwide in 2019, though the factors driving this surge appear to be somewhat
649 heterogenous across ecosystems. Nonetheless, our study lends support for the role of climate in
650 synchronizing epidemic dynamics across landscapes; it is possible that optimal climate conditions

may facilitate expansion of low prevalence lineages which possess an intrinsic fitness advantage (e.g. higher replication rate, shorter incubation period (77)) over resident genotypes, thus driving epidemics. Previous modeling studies have suggested that invading dengue lineages may circulate at low prevalence in a host population for many years prior to detection, expansion, and displacement of resident clades in the same serotype (78).

Here, we propose the novel introduction of the DENV-2 Cosmopolitan genotype, coupled with a subtle climate-driven increase in FOI, and overlaid on the background of an aging population with correspondingly aging multitypic immunity, as one possible explanation for Cambodia's largest recorded dengue epidemic to date. Our study emphasizes the extraordinary dearth of publicly available DENV sequence data for Southeast Asia; indeed, DENV is sequenced so infrequently in Cambodia that it is impossible to know whether 2019 truly marked the first year of DENV-2 Cosmopolitan introduction to the country, or simply the year of intensified expansion and consequential epidemic dynamics. More broadly, our work illustrates the importance of the combined forces of climate, demography, immunology, and virology in driving increasingly severe dengue epidemics. As the global burden of dengue continues to expand, ongoing serological and genomic surveillance is needed to improve epidemic forecasting in Southeast Asia and around the world.

METHODS

Ethics

This study was approved by the National Ethics Committee on Human Research and the National Institutes of Health (NIH) Institutional Research Board. Written informed consent was obtained from the individual participant or the parent or guardian of the child participants enrolled in this study. This study was registered at clinicaltrials.gov as NCT04034264 and NCT03534245.

Enrollment and sample collection

Nursing staff at the Kampong Speu (KPS) Referral Hospital identified, consented, enrolled, and collected demographic data from study participants. Participants, aged 6 months to 65 years of age, presented to the outpatient department with a documented fever of 38°C or greater in the previous 24 hours. Participants with clinical symptoms and signs consistent with dengue were first screened for infection using SD Bioline DengueDuo rapid tests for NS1 antigen, pan-dengue IgM and IgG. Sera was collected and processed as described elsewhere for RNA extraction, and confirmatory qRT-PCR testing for DENV-1 – 4 was performed for rapid-test positive participants (41).

Cambodia national dengue data

We obtained a 2002-2020 time series of age-structured dengue cases reported at the national level from clinicosyndromic surveillance efforts administered by the National Dengue Control Program (NDCP) of the Cambodian Ministry of Health. Since inception of the national surveillance system in 2002, each Cambodian province reports cases on a monthly basis to the national authorities who compile these reports for public health use and reporting to the World Health Organization. The date of each case and the corresponding age and gender of each

patient are reported, in addition to the province in which the case was diagnosed (25 total). We binned cases to every-two-week (biweekly) intervals for application to TSIR modeling and summarized by age class within each year for application to FOI models. We note that one province (Tboung Khmum) was created only recently, such that the corresponding time series spanned from 2016-2020. As a result, this province was eliminated from both TSIR and FOI analyses.

Spatially-resolved climate data

Climate variables, including daily mean temperature (^{0}C) and total daily precipitation (mm) from 2002 to 2019 were downloaded from the ERA5 Daily Aggregates dataset for each Cambodian province (N=25) using Google Earth Engine (79). Comparable data were not available for 2020. The spatial resolution for the ERA5 data set is approximately 31 km; however, we aggregated climate data to the size of each province using the Cambodia Administrative Boundaries Level 1 shapefile (80) during the extraction process, then computed the mean temperature (^{0}C) and total precipitation (mm) over biweekly intervals, from January 1, 2002 to December 31, 2019, while accounting for leap years.

Climate analyses

We first investigated the climate data to identify intra- and interannual trends and identify any years within the dataset that might be considered climatic anomalies. For both temperature and precipitation, we constructed two generalized additive models (GAMs) in the Gaussian family. The first GAMs incorporated a response variable of, respectively, biweekly mean temperature or total precipitation, at the province level, across the time series, with a fixed predictor of the interaction of year (as a numeric) with province, a cyclic smoothing spline by biweek of year, and a random effect of province. As a result, each GAM fit 25 distinct slopes and 25 distinct y-intercepts to the corresponding temperature or precipitation data to describe the interannual climate trend for each province, while controlling for intra-annual variation (Fig. S3-S4; Table S1). We next constructed two GAMs with the same response variables but incorporating predictor variables of year (a factor) as a random effect and a cyclic smoothing spline by biweek of year. These GAMs enabled us to identify specific years that significantly deviated from the mean climate trends (Fig. S5-6; Table S1). Finally, we reduced temperature and precipitation time series by province into z-scores by subtracting, respectively, the mean biweekly temperature and precipitation of the entire time series from each biweekly datapoint per province, then dividing by the standard deviation. We visualized z-scores to highlight years in which many provinces exhibited z-scores outside the 95% confidence intervals, representing a climate anomaly (Fig. S5-6).

TSIR modeling

We next fit time series Susceptible-Infected-Recovered models (45–47) to the province level time series for the three inter-epidemic periods (2002-2006, 2008-2011, and 2013-2018), in order to highlight the extent to which epidemic year case loads deviated from projections based on susceptible reconstruction by birth rates alone. The TSIR model leverages an input time series of case counts, births, and total population to estimate the susceptible population and disease transmission rate at intervals corresponding to the generation time of the pathogen in question (here, roughly two weeks (81)). Under TSIR assumptions, the transmission rate is held constant

740 from year to year, though allowed to vary intra-annually to reflect seasonal dynamics. Using the R
741 package tsiR (45), we fit the TSIR model individually to an input time series of biweekly dengue
742 case counts by province. Because births were reported only annually at the national level for
743 Cambodia (67), we divided these among biweeks of each year and scaled them spatially by the
744 relative population size of each province. For each inter-epidemic period, we reconstructed the
745 susceptible population from the regression of cumulative cases on cumulative births, using either
746 a Gaussian or a linear regression, depending on which approach provided the better fit to each
747 data subset, as summarized in Table S2. The new province, Tboung Khmum, was excluded from
748 TSIR analysis because case reporting only began in this region in 2016. Provinces Ratanak Kiri
749 and Mondul Kiri were additionally excluded, due to poor performance of regression models for
750 susceptible reconstruction within each region (Table S2; R^2 for 2007/2012/2019 data subsets for
751 Ratanak Kiri: 0.24/0.54/0.36; and for Mondul Kiri: 0.15/0.20/0.51).
752 While fitting TSIR, we allowed the transmission rate (β) to vary across 26 biweekly
753 intervals in a single year and the homogeneity parameter (α) to vary with each model fit. Using
754 the model trained on each inter-epidemic period, we first projected cases in the subsequent
755 epidemic years (2007, 2012, 2019) while reconstructing the susceptible population from
756 corresponding epidemic year births alone. This first exercise demonstrated the ineffectiveness
757 with which simple TSIR assumptions were able to recover case counts reported from epidemic
758 years; therefore, we next sought to quantify the increase in the assumed susceptible count
759 needed to capture reported cases. Previously, Wagner et al. 2020 (22) demonstrated analytically
760 that, under short time horizons, the susceptible population (S') in a two-strain DENV system, in
761 which a new secondary serotype invades a population previously dominated by a single endemic
762 serotype, can be approximated by the equation:

763 $$S' \approx S + \frac{R_1 \rho_2 N}{I_1} \approx S^* \left[1 + \frac{\rho_2 \gamma \beta}{\mu(\gamma + \mu)}\right]$$

764 in which S corresponds to individuals naïve to all prior infection, while R_1 and I_1 correspond to
765 individuals recovered from or infected with the endemic strain only, and N and ρ_2 respectively
766 represent the total population size and the rate of exogenous importation of a new serotype. S^*
767 then corresponds to the steady-state value of the Susceptible population under assumptions of a
768 single endemic serotype only, while μ, γ, and β encompass the rates of birth, death, and
769 transmission of exogenous infections of the new serotype. Building on these assumptions, we
770 first calculated the factorial increase in the reconstructed susceptible population needed to
771 recover case counts for the three epidemic years.
772
773 *Panel regression and climate-informed TSIR*
774 To evaluate whether consideration of epidemic-year climate factors could improve TSIR
775 predictions of epidemic cases, we next followed previous work to explore the extent to which
776 biweekly transmission was predicted by climate (53). Because dengue is a vector-borne disease,
777 we hypothesized that transmission should be lagged from climate effects on the vector population
778 and, correspondingly, first conducted cross correlation analysis to determine the optimal lag
779 between the time series of biweekly mean temperature and total precipitation on transmission at
780 the province level. In keeping with prior work for dengue (53), we considered only lag times of up
781 to one year in which the climate predictor led the empirical transmission rate; however, we
782 deviated from previous studies by allowing the optimal lag to vary by province and inter-epidemic

period. We found that the optimal time lag between peak climate predictor and peak dengue transmission ranged from 0.5 to 5 months (1 – 11 biweeks) for temperature (median = 3.5 months) and from 0.5 to 9 months (1 – 18 biweeks) for precipitation (median = 1 month), consistent with prior analyses from Sri Lanka (53) (Table S3).

Once optimal lags for temperature and precipitation were obtained, inspired by previous work (48, 49, 53, 82), we next constructed a suite of regression models, incorporating a response variable of the log of biweekly province level transmission with the corresponding predictors of optimally lagged biweekly mean temperature and total precipitation for the province in question. We constructed separate regressions for each inter-epidemic period (2002-2006, 2008-2011, and 2013-2018) and tested the sensitivity of transmission to climate predictors using a standard linear regression, a linear regression for precipitation paired with a Brière function for temperature (53, 83), and a GAM (Fig. S8-10; Table S4). We then used lagged temperature and precipitation corresponding to each epidemic year in each province to project climate-informed transmission rates for 2007, 2012, and 2019. Results were qualitatively similar across all regression structures; as such, we report only GAM-projected transmission rates here (*SI Appendix*). Using climate-informed transmission rates, we again simulated TSIR to recover epidemic year predictions of caseload at the province level, which we evaluated against projections by TSIR fitted to the inter-epidemic periods and incorporating a proportional increase in the susceptible population. Because climate-derived transmission rates still failed to recover epidemic year peaks, we allowed for yet another susceptible amplification term to determine the relative increase in the susceptible population still needed to recover epidemic dynamics even after accounting for climate (Table S5).

Wavelet analyses

We next used wavelet decomposition in the R package 'WaveletComp' (54) to explore annual and multiannual periodicity in dengue epidemics and corresponding climate variables. We converted biweekly case totals by province from the 2002-2020 NDCP dataset into incidence rates per 100,000 population, then, as in previous work (25, 26), used a Morlet wavelet with nondimensional frequency ($\omega = 6$) to extract, detrend, and reconstruct annual cycles in dengue incidence with a maximum period of two years (52 biweeks) and multiannual cycles with periods ranging from two to 20 years (Fig. 2AB). We additionally calculated the average wavelet power in each biweekly timestep for both annual and multiannual cycles per province (Fig. S11AB). Next, we investigated synchronicity in dengue incidence across space by computing the Pearson's correlation coefficient (ρ) between province pairs, using the annual raw incidence and the reconstructed annual and multiannual cycles. All results for annual cycles were qualitatively similar to those for the annual raw incidence and are, therefore, not reported here. For annual incidence, we computed ρ for each province pair combination in yearly timesteps, then calculated the annual average ρ for each focal province compared against all other provinces (Fig. 2C). For multiannual cycles, we computed ρ for all pairwise province combinations across a sliding 5-year window. As with annual incidence, we then calculated the average ρ for each focal province per year, compared against all other provinces. Because 5-year cycles overlapped, ρ was averaged over multiple overlapping comparisons for each pairwise combination for all but the first and last year in the time series (Fig. 2D). For another measure of synchronicity, we additionally computed the cross-wavelet power spectrum for all province pairs, using both annual incidence rates in yearly timesteps and multiannual reconstructed cycles in 5-year intervals. As with the Pearson's

correlation coefficient, we averaged the output of these analyses (here, significant values for cross-wavelet power) for each focal province, as compared with all other provinces, at the appropriate timestep (Fig. S11CD). Because patterns in synchronicity were most pronounced at the annual incidence scale, in order to identify statistical correlates of high synchronicity years and localities, we finally constructed a GAM with a response variable of ρ, as reported in Fig. 2C, a fixed predictor of the interaction of focal province and geographic distance to the province under comparison; smoothing predictors of biweekly mean temperature, total precipitation, and mean population size for the focal province; and a random effect of year (Fig. S12; Table S6).

Because this last analysis indicated some predictive role for climate in driving synchrony in dengue incidence, we next investigated coherence between case data and climate variables, as previously compiled for TSIR from 2002-2019 (Fig. S13). We first computed mean cross-wavelet power as a measure of synchrony between the biweekly time series of raw dengue incidence and mean temperature (Fig. S13A) and total precipitation (Fig. S13B) per province. We next calculated mean cross-wavelet power between reconstructed multiannual dengue cycles over a 5-year interval and the same two climate variables, again per province (Fig. S13CD). Finally, we reconstructed *monthly* multiannual dengue cycles for each province to compute average cross-wavelet power with the Oceanic Niño Index (ONI), a monthly time series that quantifies the intensity of the El Niño Southern Oscillation, which has been previously associated with dengue dynamics in SEA (Fig. S13E) (26, 27, 65, 66).

Lastly, we extracted the mean period from reconstructed multiannual dengue cycles by province and at the national level, to compare with publicly available demographic data (birth rates, death rates, population size (67)) that informed downstream analyses (Fig. S14).

Quantifying mean age of reported infection

Our next analyses shifted the focus from climate to demographic drivers of dengue incidence. To this end, we first quantified the interannual trend in the mean age of reported DENV infection, by province, across our 2002-2020 time series. We fit a GAM with a response variable of age to a fixed predictor of the interaction of year and province, while also including a random effect of province, to allow both slope and y-intercept to vary by the locality over which the date were compiled. We visualized these interannual trends across the age distribution of reported cases by province (Fig. S15) and also summarized at the national level (Fig. 3AB; Table S7).

Force of Infection estimation

We next used the age-stratified, province level surveillance data to estimate the annual FOI for DENV in Cambodia across the 19-year time series from 2002-2020. Methods for estimating FOI from age-stratified serological data for single-strain pathogens are well established (40, 84–87), and prior work has adapted these methods to account for the role of pre-existing heterotypic immunity in DENV infection (39) and modified them for application to age-structured incidence data, in lieu of serology (20). We applied the model developed by Ferguson et al. 1999 (39) and Cummings et al. 2009 (20) to age-structured incidence recovered from the NDCP data, at the province level, assuming reported cases to represent secondary infections and all individuals in the dataset to eventually experience exposure to multiple DENV serotypes across their lifetimes. We allowed for a unique FOI across each year in the time series but first assumed a constant FOI across all age cohorts within a single year. As in prior work, this method additionally supported the estimation of FOIs that predated our data time series, albeit with

873 diminished confidence as compared to our data-associated years. Specifically, we estimated one
874 FOI per year per province, in addition to a summary FOI at the national scale, for all years
875 predating the start of the 2002 time series and corresponding to the year of birth for the oldest
876 individual in the first year of the corresponding subset of the data (Fig. S16; Table S8). For the
877 national data (and for Takeo province), this corresponded to a 40-year time series, dating back to
878 1981, the birth year of the oldest individual (22 years) in the first year of the dataset (*SI Appendix;*
879 Table S8). No FOIs were estimated for the Tboung Khmum province for which the NDCP only
880 began reporting data in 2016. We estimated mean FOI per serotype under assumptions of two
881 circulating dengue serotypes in the region. Because FOI was estimated as a serotype average,
882 these estimates would be proportionally depressed if instead accounting for three or four
883 circulating dengue serotypes in the system. To highlight links to demography, we plotted FOI
884 estimates collectively in comparison to national reported birth and death rates in Fig. 3C (67).
885 After fixing FOI by province, we next followed prior work (20) to estimate FOI modifiers by
886 age cohort, shared across all provinces and all years of the data. Because of the wide range in
887 the age distribution of cases across the time series under consideration, we fit thirteen age-
888 specific modifiers to the fixed FOI values, which we allowed to vary between the first and second
889 decades of the dataset, such that 26 FOI-modification terms were estimated in total (Table S8).
890 Finally, after observing the extremely high age distribution of reported cases in the later years of
891 the time series, we modified our previous model to allow for a rate of waning immunity from
892 multitypic back to monotypic immunity (σ), such that older individuals could experience renewed
893 pathogenic infections likely to be reported in the data (*SI Appendix*). As with age-modification
894 terms, we estimated waning multitypic immunity shared across all provinces and all years in the
895 dataset after fixing FOI. Because little signal of waning multitypic immunity was evident across
896 the entire time series, we next allowed σ to vary by year across the dataset. We compared fits of
897 the FOI-only, FOI with age modification, FOI with waning multitypic immunity, and FOI with both
898 age modification and waning multitypic immunity models to the data (Table S8), then simulated
899 the resulting accumulation of cases with age from the best fit model at national (Fig. 3E) and
900 province levels (Fig. S17).
901
902 *Viral sequencing*
903 Metagenomic Next-Generation Sequencing (mNGS) was applied to serum samples
904 collected from all patients reporting with symptoms in our febrile cohort study. Briefly, pathogen
905 mNGS libraries were prepared from isolated pathogen RNA and converted to cDNA Illumina
906 libraries using the NEBNext Ultra II DNA Library Prep Kit (E7645) according to manufacturer's
907 instructions. Library size and concentration were determined using the 4150 Tapestation system,
908 Agilent, and Qubit 4 Flurometer, Invitrogen (for quantitation only). External RNA Controls 103
909 Consortium collection, ERCC, ThermoFisher, were used as indicators of potential library
910 preparation errors and for input RNA mass calculation. Samples were sequenced on a NovaSeq
911 (Illumina) instrument and an iSeq100 (Illumina) instrument using 150 nucleotide paired-end
912 sequencing. A water ("no template") control was included in each library preparation.
913 Raw fastq files were uploaded to the CZID portal, a cloud-based, open-source
914 bioinformatics platform, to identify microbes from metagenomic data (https://czid.org) (88).
915 Potential pathogens were distinguished from commensal flora and contaminating microbial
916 sequences from the environment by establishing a z-score metric based on a background
917 distribution derived from 16 non-templated "water-only" control libraries. Data were normalized to

reads mapped per million input reads for each microbe at both species and genus levels. Taxa with z-score less than 1, base pair alignment less than 50 base pairs, NT log(1/e) less than 10 and reads per million (rpM) less than 10 were removed from analysis. Microbial sequences from the samples are available for access in the National Center for Biotechnology Information (NCBI) Sequence Read Archive.

Construction of consensus genomes

We attempted to construct full-genome DENV sequences from any samples which were confirmed to be DENV-positive by RT-qPCR and which generated at least one reliable contig mapping to any serotype of DENV in the CZID pipeline. To generate full genome consensus sequences, we ran the ARTIC network's Nextflow consensus genome pipeline (89), mapping each sequence to the closest GenBank accession number hit in the original mNGS run of CZID, using a cutoff of 5 reads per nucleotide site to make a consensus call (sites with <5 reads were called as "N"). Sequences were additionally run through the CZID integrated consensus genome pipeline, again mapping to the closest hit identified in GenBank from the original mNGS assembly. Resulting consensus sequences from both assembly pipelines were then aligned with reference sequences and visually examined in Geneious Prime. Raw reads from mNGS were then mapped to each full genome contig in turn and examined manually to determine the correct call for each base pair.

Using these methods, we generated full or near-full genome sequences for 57 DENV-1, 61 DENV-2, and 4 DENV-4 samples, representing 122 of the 123 DENV-positive patients identified in our dataset. All contributed genomes were >10,000 bps in length and had a maximum of 90 Ns (corresponding to <1% of the DENV genome). Resulting sequences were uploaded to NCBI as individual FASTA files (Table S9).

Phylodynamic and phylogenetic analysis

We supplemented our own Cambodia sequences with all other Cambodian sequences for DENV-1 and DENV-2 available in GenBank at the time of analysis, selecting all full or partial genome nucleotide sequences >10,000 bp in length up to a collection date of December 31, 2020 (DENV-1: tax id 11053, 192 sequences, including 57 contributed by this study; DENV-2: tax id 11060, 116 sequences, including 61 contributed by this study). We further supplemented these Cambodia sequences with genomes collected from other major Southeast Asian countries (nine), which were Laos, Myanmar, Malaysia, Thailand, Vietnam, Brunei, Indonesia, the Philippines, and Singapore. All countries were represented in both the DENV-1 and DENV-2 datasets. To avoid overrepresenting certain countries outside of Cambodia, we limited sequence selection to a maximum of three randomly selected genomes collected per year from each available year per country, beginning in 2002, the year in which we began our national time series.

After selection, sequences were aligned separately by serotype in the program MAFFT (90), and the best fit nucleotide substation model for each set of data was evaluated in the program ModelTest-NG (91). For both DENV-1 and DENV-2, a GTR+I+G4 nucleotide was determined to offer the best fit to the data. Using this best fit nucleotide substitution model, we next built a Bayesian phylogenetic tree for each set of DENV genomes in the program BEAST 2 (60). We incorporated the date of sample collection for each sequence (or the midpoint of the year of collection if the date was not reported), and, after Salje et al. 2017 (36), we specified a strict molecular clock at a rate of 7.9×10^{-4} s/s/y (92) and a Coalescent Bayesian skyline prior in

our models. We ran Markov Chain Monte Carlo chains in BEAST 2 for 150 million iterations, logging results every 10,000 iterations. After chains completed, we removed the initial 10% of iterations as burn-in and evaluated parameter convergence in Tracer v1.6 (parameters were deemed to have converged at ESS values ≥ 200). We summarized resulting phylogenetic trees in TreeAnnotator and visualized summary trees in the R package ggtree (93) (Fig. 4AB; Fig. S19).

Finally, we computed a Maximum Likelihood phylogenetic tree to illustrate the phylogenetic placement of our new Cambodia sequences in relation to all previously described genotypes of DENV-1 and DENV-2. For this analysis, we included all Cambodian sequences of DENV-1 and DENV-2 available in NCBI, as well as a broadly representative subset of sequences within all known genotypes of DENV-1 (genotypes I, II, and III) and DENV-2 (genotypes Asian-I, Asian-II, Asian-American, American, Cosmopolitan I, Cosmopolitan II, and Cosmopolitan III). As with Bayesian timetrees, sequences were first aligned in MAFFT (90), and the best fit nucleotide substitution model was computed in ModelTest-NG (91). As before, we found that a GTR+I+G4 best represented both sequence subsets; using this model, corresponding phylogenetic trees were then constructed in the program RAxML (94). Following best practices outlined in the RAxML-NG manual, 20 ML inferences were made on each original alignment and bootstrap replicate trees were inferred using Felsenstein's method (95), with the MRE-based bootstopping test applied after every 50 replicates (96). Resulting phylogenies were then visualized in ggtree (93) (Fig. S18).

Estimating transmission chains from sequence data

Lastly, we followed methods outlined in Salje et al. 2017 (36) to calculate the proportion of sequences within each DENV serotype that could be attributed to the same transmission chain on our Bayesian timetrees, defined as having a most recent common ancestor within the past six months in the same season. The reciprocal of the proportion of sequences sharing a transmission chain corresponds to the effective number of transmission chains circulating in a given population. Thus, to compare our estimates of transmission chain density against those previously reported for Thailand (36), we computed the total effective number of transmission chains observed in our 2019-2020 sequence dataset, separately for DENV-1 and DENV-2, at the World Bank reported population density for Kampong Speu province (67) (Fig. 4C). We also compared the proportion of DENV-1 vs. DENV-2 sequences determined to share a transmission chain to the Euclidean distance separating the precise GPS coordinates of the collection points of each sequence pair (Fig. 4D).

Mechanistic modeling of age-structured dengue dynamics

Finally, we constructed a mechanistic, age-structured discrete time deterministic epidemic model in biweekly timesteps (*SI Appendix*) (62–64) to simulate two- and three-serotype dengue infections in a population demographically structured to mimic that of Cambodia over the past half-century. Using publicly-available national data for Cambodian birth rates (1960-2020) (67), population age structure (1950-2020) (97), and age-specific mortality rates (1950-2020) (98), we simulated the dynamics of two or three circulating DENV serotypes out to endemic equilibrium, incorporating seasonal intra-annual variation in transmission as estimated from TSIR. We then introduced the annual national FOI from 1999-2020, as previously estimated by fitting catalytic models to the NDCP data (Fig. 3) and tracked the corresponding increase in the mean

age of simulated dengue cases. We modified dynamics to test hypotheses of potential drivers underlying the spike in cases and corresponding expansion in age structure witnessed in the three epidemics identified in the NDCP data (Fig. 5). As highlighted in the 'Results' section, we focused our analysis on drivers of the 2019 epidemic, simulating scenarios by which 2-circulating serotype dynamics were interrupted by a climate-driven spike in FOI in the epidemic year (H1), a genotype invasion and clade-replacement event with waning immunity within the serotype (H2), or a third serotype invasion (H3), then compared these to a hypothesis of 3-serotype circulation and increasing detectability of tertiary infections through time (H4) (Fig. 5A). To highlight the link between the underlying dynamics and our prior inference from the observed data, as well as check that our simulated dynamics recapitulated those witnessed in the actual data, we fit our original Ferguson-Cummings two-serotype catalytic FOI model to the simulated data generated from our mechanistic model for each hypothesis to recapture the input FOI (Fig. S20). Holding FOI constant, we then estimated a time-varying signature of waning monotypic immunity for all simulated time series (Fig. S20) and compared results. Finally, we repeated analyses with H1, H2, and H3 alterations introduced in 2007 to highlight their downstream impacts on the rest of the time series (Fig. S21).

DATA AVAILABILITY

All genome sequence data from this study have been submitted to the NCBI Sequence Read Archive under Bioproject ID PRJNA681566. Consensus DENV sequences are also available in GenBank, under accession numbers reported in Table S9. All bioinformatics code for the initial mNGS assembly is available at https://github.com/chanzuckerberg/idseq-workflows and all wet lab bench protocols are updated at https://docs.google.com/document/d/1RtNQc1D4or_ys7OxCCBjh4SDIdy7JaI4IE7if8EkHgE/edit. Further information on how to use ½ reaction volumes and FastSelect® are available upon request. Code for generation of consensus genomes is available from the Artic Network at https://artic.readthedocs.io/en/latest/. Raw data and detailed instructions of all specific analyses undertaken to produce the figures and results presented here are available in our open access github repository at: https://github.com/brooklabteam/cambodia-dengue-national. All other reasonable data requests can be made to authors directly.

ACKNOWLEDGEMENTS

This research is supported by the Division of Intramural Research at the National Institute of Allergy and Infectious Diseases at the National Institutes of Health and the Bill and Melinda Gates Foundation [grant numbers OPP1211806, OPP1211841]. We thank patients and families of Kampong Speu District Referral Hospital who participated in this study. We thank members of the National Dengue Control Program not listed in the author byline and the Provincial Health Department of Kampong Speu province in Cambodia. We thank all the other employees at the Chan Zuckerberg Biohub and Chan Zuckerberg Initiative not listed in the author byline. We thank Brian Moyer and the NIAID Office of Cyberinfrastructure and Computational Biology (OCICB) for their assistance in improving the cyberinfrastructure of our Cambodian field sites. This work was completed with resources provided by the University of Chicago's Research Computing Center.

1053 **REFERENCES**

1054 1. World Health Organization, Dengue and severe dengue. *WHO Fact Sheets* (2022).
1055 2. S. Bhatt, *et al.*, The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
1056 3. A. B. Sabin, Research on dengue during World War II. *American Journal of Tropical*
1057 *Medicine and Hygiene* **1**, 30–50 (1952).
1058 4. L. C. Katzelnick, *et al.*, Antibody-dependent enhancement of severe dengue disease in
1059 humans. *Science (New York, N.Y.)* **932**, 929–932 (2017).
1060 5. L. C. Katzelnick, *et al.*, Zika virus infection enhances future risk of severe dengue disease.
1061 *Science* **369**, 1123–1128 (2020).
1062 6. statista, Cambodia: Urbanization from 2010 to 2020.
1063 *https://www.statista.com/statistics/455789/urbanization-in-cambodia/.* (2022).
1064 7. T. N. Peng, "Southeast Asia's demographic situation, regional variations, and national
1065 challenges" in D. Singh, M. Cook, Eds. (ISEAS Publishing, 2017), pp. 55–82.
1066 8. G. W. Jones, "The Population of Southeast Asia" (2013).
1067 9. Hahn, H., Chastel, C., Dengue in Cambodia in 1963. Nineteen laboratory-proved cases.
1068 *American Journal of Tropical Medecine and Hygiene* **19**, 106–109 (1970).
1069 10. R. Huy, *et al.*, National dengue surveillance in Cambodia 1980–2008: epidemiological and
1070 virological trends and the impact of vector control. *Bull. World Health Organ.* **88**, 650–657
1071 (2010).
1072 11. C. Yek, *et al.*, National dengue surveillance, Cambodia 2002–2020. *Bull World Health Org*
1073 **101**, 605–616 (2023).
1074 12. D. S. Burke, A. Nisalak, D. E. Johnson, A prospective study of dengue infections in
1075 Bangkok. *The American Journal of Tropical Medicine and Hygiene* **38**, 172–180 (1988).
1076 13. V. Duong, *et al.*, Genetic diversity and lineage dynamic of dengue virus serotype 1 (DENV-
1077 1) in Cambodia. *Infection, Genetics and Evolution* **15**, 59–68 (2013).
1078 14. O. O'Connor, *et al.*, Potential role of vector-mediated natural selection in dengue virus
1079 genotype/lineage replacements in two epidemiologically contrasted settings. *Emerging*
1080 *Microbes and Infections* **10**, 1346–1357 (2021).
1081 15. A. A. Lover, *et al.*, Spatial epidemiology and climatic predictors of paediatric dengue
1082 infections captured via sentinel site surveillance, Phnom Penh Cambodia 2011-2012. *BMC*
1083 *Public Health* **14** (2014).
1084 16. Y. Arima, M. Chiew, T. Matsui, Epidemiological update on the dengue situation in the
1085 Western Pacific Region, 2012. *Western Pacific surveillance and response journal : WPSAR*
1086 **6**, 82–89 (2015).
1087 17. S. Ly, *et al.*, Asymptomatic dengue virus infections, Cambodia, 2012–2013. *Emerging*
1088 *Infectious Diseases* **25**, 1354–1362 (2019).
1089 18. S. Vong, *et al.*, Under-recognition and reporting of dengue in Cambodia: A capture-
1090 recapture analysis of the National Dengue Surveillance System. *Epidemiology and Infection*
1091 **140**, 491–499 (2012).
1092 19. A. T. Huang, *et al.*, Assessing the role of multiple mechanisms increasing the age of dengue
1093 cases in Thailand. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2115790119 (2022).
1094 20. D. A. T. Cummings, *et al.*, The impact of the demographic transition on dengue in Thailand:
1095 Insights from a statistical analysis and mathematical modeling. *PLoS Medicine* **6** (2009).

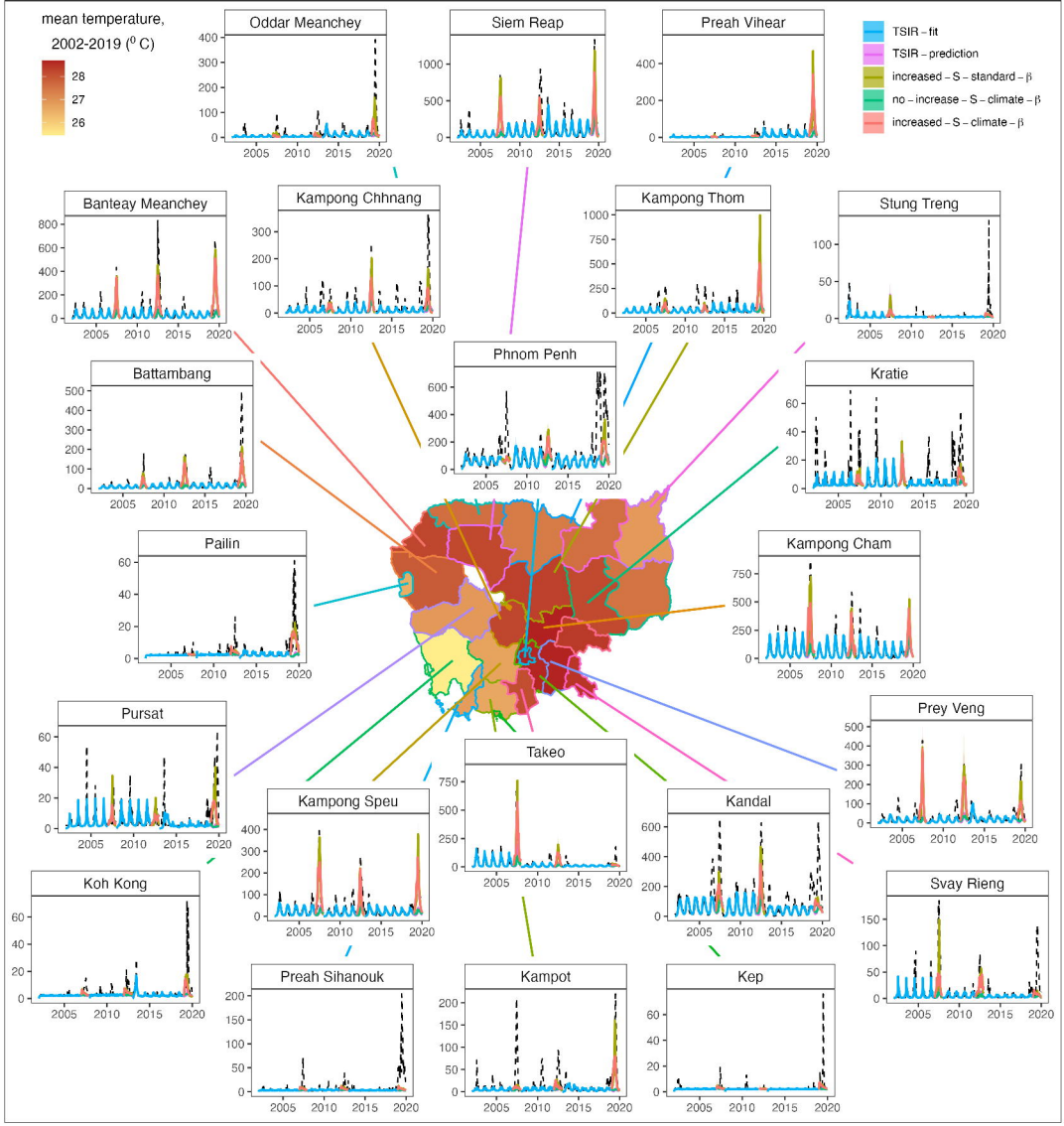
1096 21. L. C. Katzelnick, *et al.*, Dynamics and determinants of the force of infection of dengue virus
1097 from 1994 to 2015 in Managua, Nicaragua. *Proceedings of the National Academy of*
1098 *Sciences of the United States of America* **115**, 10762–10767 (2018).
1099 22. N. Ferguson, R. Anderson, S. Gupta, The effect of antibody-dependent enhancement on
1100 the transmission dynamics and persistence of multiple-strain pathogens. *Proc. Natl. Acad.*
1101 *Sci. U.S.A.* **96**, 790–794 (1999).
1102 23. S. I. Hay, *et al.*, Etiology of interepidemic periods of mosquito-borne disease. *Proc. Natl.*
1103 *Acad. Sci. U.S.A.* **97**, 9335–9339 (2000).
1104 24. M. A. Johansson, D. A. T. Cummings, G. E. Glass, Multiyear climate variability and
1105 dengue—El Niño Southern Oscillation, weather, and dengue incidence in Puerto Rico,
1106 Mexico, and Thailand: A longitudinal data analysis. *PLoS Med* **6**, e1000168 (2009).
1107 25. B. García-Carreras, *et al.*, Periodic synchronisation of dengue epidemics in Thailand over
1108 the last 5 decades driven by temperature and immunity. *PLoS Biol* **20**, e3001160 (2022).
1109 26. W. G. van Panhuis, *et al.*, Region-wide synchrony and traveling waves of dengue across
1110 eight countries in Southeast Asia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13069–13074 (2015).
1111 27. B. Cazelles, M. Chavez, A. J. McMichael, S. Hales, Nonstationary Influence of El Niño on
1112 the Synchronous Dengue Epidemics in Thailand. *PLoS Med* **2**, e106 (2005).
1113 28. D. A. T. Cummings, *et al.*, Travelling waves in the occurrence of dengue haemorrhagic
1114 fever in Thailand. *Nature* **427**, 344–347 (2004).
1115 29. J. H. Huber, M. L. Childs, J. M. Caldwell, E. A. Mordecai, Seasonal temperature variation
1116 influences climate suitability for dengue, chikungunya, and Zika transmission. *PLoS Negl*
1117 *Trop Dis* **12**, e0006451 (2018).
1118 30. E. A. Gould, S. Higgs, Impact of climate change and other factors on emerging arbovirus
1119 diseases. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **103**, 109–
1120 121 (2009).
1121 31. D. A. T. Cummings, I. B. Schwartz, L. Billings, L. B. Shaw, D. S. Burke, Dynamic effects of
1122 antibody-dependent enhancement on the fitness of viruses. *Proceedings of the National*
1123 *Academy of Sciences* **102**, 15259–15264 (2005).
1124 32. B. Adams, *et al.*, Cross-protective immunity can account for the alternating epidemic pattern
1125 of dengue virus serotypes circulating in Bangkok. *Proceedings of the National Academy of*
1126 *Sciences of the United States of America* **103**, 14234–9 (2006).
1127 33. K. L. McElroy, *et al.*, Endurance, refuge, and reemergence of dengue virus type 2, Puerto
1128 Rico, 1986–2007. *Emerg. Infect. Dis.* **17**, 64–71 (2011).
1129 34. C. V. F. Carrington, J. E. Foster, O. G. Pybus, S. N. Bennett, E. C. Holmes, Invasion and
1130 maintenance of Dengue virus type 2 and type 4 in the Americas. *Journal of Virology* **79**,
1131 14680–14687 (2005).
1132 35. L. C. Katzelnick, *et al.*, Antigenic evolution of dengue viruses over 20 years. *Science* (2021).
1133 36. H. Salje, *et al.*, Dengue diversity across spatial and temporal scales: Local structure and the
1134 effect of host population size. *Science (New York, N.Y.)* **355**, 1302–1306 (2017).
1135 37. J. P. Messina, *et al.*, The current and future global distribution and population at risk of
1136 dengue. *Nature Microbiology* **4**, 1508–1515 (2019).
1137 38. L. Cattarino, I. Rodriguez-Barraquer, N. Imai, D. A. T. Cummings, N. M. Ferguson, Mapping
1138 global variation in dengue transmission intensity. *Science Translational Medicine* **12**,
1139 eaax4144 (2020).

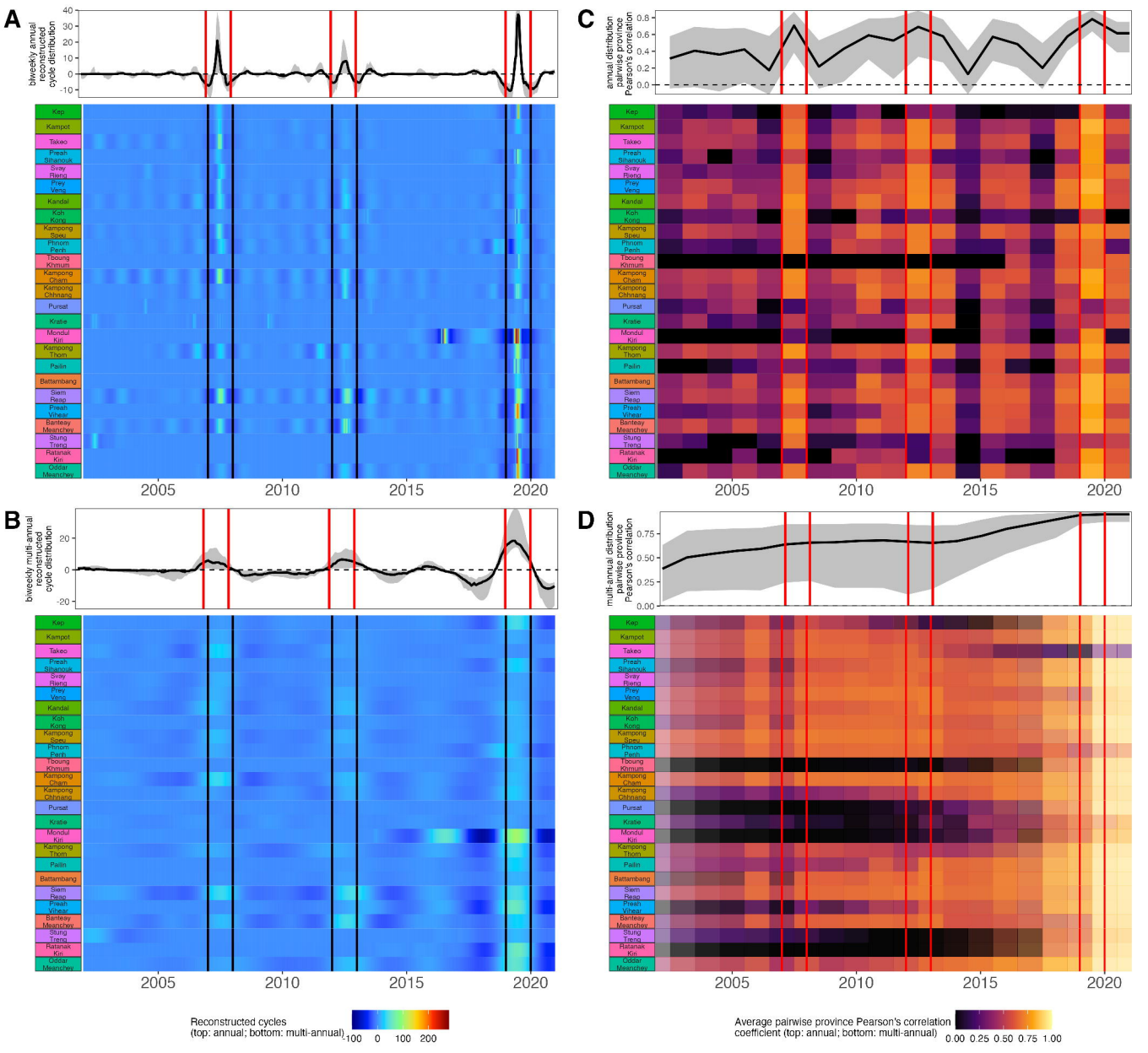
39. N. M. Ferguson, C. A. Donnelly, R. M. Anderson, Transmission dynamics and epidemiology of dengue: insights from age-stratied sero-prevalence surveys. *Proceedings of the Royal Society B* **354**, 757–768 (1999).
40. Hugo. Muench, *Catalytic models in epidemiology.* (Harvard University Press, 1959).
41. J. E. Manning, *et al.*, Development of inapparent dengue associated with increased antibody levels to *Aedes aegypti* salivary proteins: A longitudinal dengue cohort in Cambodia. *The Journal of Infectious Diseases* (2021) https:/doi.org/10.1093/infdis/jiab541.
42. J. A. Bohl, *et al.*, Discovering disease-causing pathogens in resource-scarce Southeast Asia using a global metagenomic pathogen monitoring system. *Proceedings of the National Academy of Sciences* **119**, e2115285119 (2022).
43. S. N. Wood, mgcv: GAMs and Generalized Ridge Regression for R. *R News* **1/2**, 20–24 (2001).
44. NOAA: National Weather Service*Climate Predication Scenter: Cold & Warm Episodes by Season* (2023).
45. A. D. Becker, B. T. Grenfell, TSIR: An R package for time-series susceptible-infected-recovered models of epidemics. *PLoS ONE* **12**, 1–10 (2017).
46. O. N. Bjornstad, B. F. Finkenstadt, B. T. Grenfell, Dynamics of measles epidemics: Estimating scaling of transmission rates using a times series SIR model. *Ecological Monographs* **72**, 169–184 (2002).
47. B. T. Grenfell, O. N. Bjornstad, B. F. Finkenstadt, Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the TSIR Model. *Ecological Monographs* **72**, 185–202 (2002).
48. R. E. Baker, A. S. Mahmud, C. J. E. Metcalf, Dynamic response of airborne infections to climate change: predictions for varicella. *Climatic Change* **148**, 547–560 (2018).
49. R. J. Oidtman, *et al.*, Inter-annual variation in seasonal dengue epidemics driven by multiple interacting factors in Guangzhou, China. *Nature Communications* **10** (2019).
50. N. A. Rehman, *et al.*, Quantifying the localized relationship between vector containment activities and dengue incidence in a real-world setting: A spatial and time series modelling analysis based on geo-located data from Pakistan. *PLoS Neglected Tropical Diseases* **14**, 1–22 (2020).
51. M. U. G. Kraemer, *et al.*, Big city, small world: density, contact rates, and transmission of dengue across Pakistan. *J. R. Soc. Interface.* **12**, 20150468 (2015).
52. M. U. G. Kraemer, *et al.*, Inferences about spatiotemporal variation in dengue virus transmission are sensitive to assumptions about human mobility: a case study using geolocated tweets from Lahore, Pakistan. *EPJ Data Science* **7** (2018).
53. C. E. Wagner, *et al.*, Climatological, virological and sociological drivers of current and projected dengue fever outbreak dynamics in Sri Lanka. *Journal of the Royal Society Interface* **17** (2020).
54. Roesch, Angi, Schmidbauer, Harald, WaveletComp: Computational Wavelet Analysis. R package version 1.1. (2018).
55. J. Phadungsombat, *et al.*, Emergence of genotype cosmopolitan of dengue virus type 2 and genotype III of dengue virus type 3 in Thailand. *PLoS ONE* **13**, 1–26 (2018).
56. V. T. Tran, *et al.*, Reemergence of Cosmopolitan genotype dengue virus serotype 2, southern Vietnam. *Emerg. Infect. Dis.* **29** (2023).

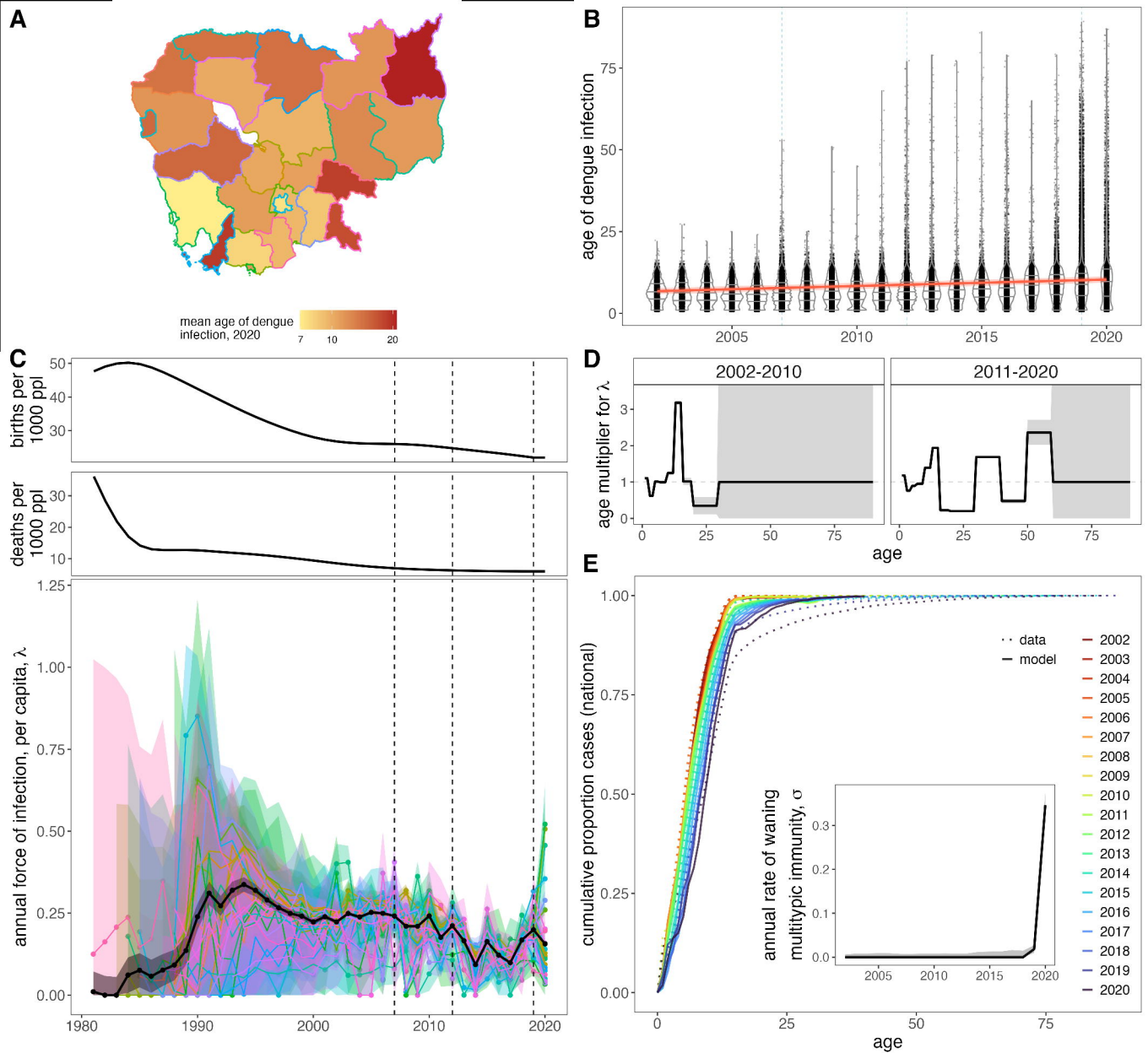
1184 57. A. Wijewickrama, *et al.*, Emergence of a Dengue virus serotype 2 causing the largest ever
1185 dengue epidemic in Sri Lanka. *bioRxiv* (2018) https:/doi.org/10.1101/329318.
1186 58. H. A. Tissera, *et al.*, Severe dengue epidemic, Sri Lanka, 2017. *Emerg. Infect. Dis.* **26**, 682–
1187 691 (2020).
1188 59. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti
1189 and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969–1973 (2012).
1190 60. R. Bouckaert, *et al.*, BEAST 2: A software platform for Bayesian evolutionary analysis.
1191 *PLoS Computational Biology* **10** (2014).
1192 61. K. Poltep, *et al.*, Genetic diversity of dengue virus in clinical specimens from Bangkok,
1193 Thailand, during 2018–2020: Co-circulation of all four serotypes with multiple genotypes
1194 and/or clades. *Tropical Medicine and Infectious Disease* **6** (2021).
1195 62. P. Klepac, H. Caswell, The stage-structured epidemic: Linking disease and demography
1196 with a multi-state matrix approach model. *Theoretical Ecology* **4**, 301–319 (2011).
1197 63. C. J. E. Metcalf, *et al.*, Structured models of infectious disease: Inference with discrete data.
1198 *Theoretical Population Biology* **82**, 275–282 (2012).
1199 64. P. Klepac, *et al.*, Stage-structured transmission of phocine distemper virus in the Dutch
1200 2002 outbreak. *Proceedings. Biological sciences / The Royal Society* **276**, 2469–2476
1201 (2009).
1202 65. M. Tipayamongkholgul, C.-T. Fang, S. Klinchan, C.-M. Liu, C.-C. King, Effects of the El
1203 Niño-Southern Oscillation on dengue epidemics in Thailand, 1996-2005. *BMC Public Health*
1204 **9**, 422 (2009).
1205 66. A. Gagnon, A. Bush, K. Smoyer-Tomic, Dengue epidemics and the El Niño Southern
1206 Oscillation. *Clim. Res.* **19**, 35–43 (2001).
1207 67. World Bank, Cambodia. *World Development Indicators* (2021) (April 12, 2022).
1208 68. I. A. Sampson, G. M. Miles, E. Piano, “Undocumented, unregistered and invisible”: an
1209 exploratory study of the reasons for and effects of migrating to Thailand of Cambodian
1210 young people. *IJSSP* **41**, 862–874 (2021).
1211 69. L. Lambrechts, *et al.*, Dengue-1 virus clade replacement in Thailand associated with
1212 enhanced mosquito transmission. *Journal of Virology* **86**, 1853–1861 (2012).
1213 70. J. G. de Jesus, *et al.*, Genomic detection of a virus lineage replacement event of dengue
1214 virus serotype 2 in Brazil, 2019. *Memorias do Instituto Oswaldo Cruz* **115** (2020).
1215 71. K. Suzuki, *et al.*, Genotype replacement of dengue virus type 3 and clade replacement of
1216 dengue virus type 2 genotype Cosmopolitan in Dhaka, Bangladesh in 2017. *Infection,*
1217 *Genetics and Evolution* **75** (2019).
1218 72. L. C. Katzelnick, *et al.*, Dengue viruses cluster antigenically but not as discrete serotypes.
1219 *Science* **349**, 1338–1343 (2015).
1220 73. A. F. Brito, *et al.*, Lying in wait: the resurgence of dengue virus after the Zika epidemic in
1221 Brazil. *Nature Communications* **12**, 1–13 (2021).
1222 74. T. Tsheten, *et al.*, Epidemiological analysis of the 2019 Dengue epidemic in Bhutan.
1223 *International Journal of Environmental Research and Public Health* **18**, 1–13 (2021).
1224 75. M. S. Hossain, *et al.*, Dengue in a crowded megacity: Lessons learnt from 2019 outbreak in
1225 Dhaka, Bangladesh. *PLoS Neglected Tropical Diseases* **14**, 1–5 (2020).
1226 76. A. Ahsan, N. Haider, R. Kock, C. Benfield, Possible drivers of the 2019 Dengue outbreak in
1227 Bangladesh: the need for a robust community-level surveillance system. *Journal of Medical*
1228 *Entomology* **58**, 37–39 (2021).

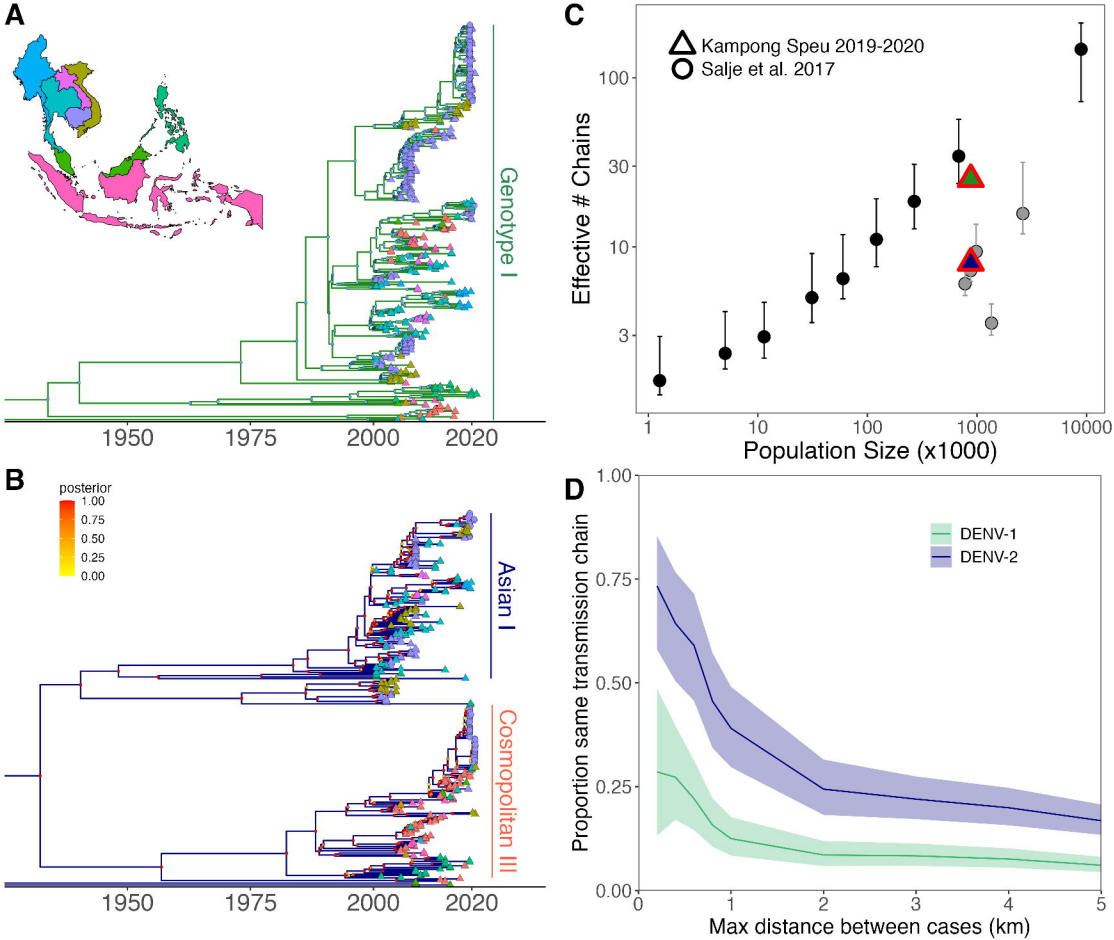
1229 77. R. Rico-Hesse, *et al.*, Origins of Dengue type 2 viruses associated with increased
1230 pathogenicity in the Americas. *Virology* **230**, 244–251 (1997).
1231 78. J. Lourenço, M. Recker, Viral and epidemiological determinants of the invasion dynamics of
1232 Novel Dengue Genotypes. *PLoS Neglected Tropical Diseases* **4** (2010).
1233 79. Copernicus Climate Change Service Climate Data Store (CDS): Copernicus Climate
1234 Change Service (C3S), ERA5: Fifth generation of ECMWF atmospheric reanalyses of the
1235 global climate. (2017) (March 1, 2023).
1236 80. AmeriGEOSS Community Platform DataHub., Cambodia Admin Boundaries Level 1 (2023).
1237 81. H. Salje, *et al.*, Reconstructing unseen transmission events to infer dengue dynamics from
1238 viral sequences. *Nature Communications*, 1–10 (2021).
1239 82. R. E. Baker, *et al.*, Epidemic dynamics of respiratory syncytial virus in current and future
1240 climates. *Nat Commun* **10**, 5512 (2019).
1241 83. E. A. Mordecai, *et al.*, Detecting the impact of temperature on transmission of Zika, dengue,
1242 and chikungunya using mechanistic models. *PLoS Negl Trop Dis* **11**, e0005568 (2017).
1243 84. B. T. Grenfell, R. M. Anderson, The estimation of age-related rates of infection from case
1244 notifications and serological data. *The Journal of Hygiene* **95**, 419–36 (1985).
1245 85. L. W. Pomeroy, *et al.*, Serotype-specific transmission and waning immunity of endemic foot-
1246 and-mouth disease virus in Cameroon. *PLoS ONE* **10**, 1–16 (2015).
1247 86. G. H. Long, *et al.*, Identifying the age cohort responsible for transmission in a natural
1248 outbreak of Bordetella bronchiseptica. *PLoS Pathogens* **6** (2010).
1249 87. D. M. Heisey, D. O. Joly, F. Messier, The fitting of general force-of-infection models to
1250 wildlife disease prevalence data. *Ecology* **87**, 2356–2365 (2006).
1251 88. K. L. Kalantar, *et al.*, IDseq-An open source cloud-based pipeline and analysis service for
1252 metagenomic pathogen detection and monitoring. *GigaScience* **9**, 1–14 (2021).
1253 89. ARTIC-network, The ARTIC field bioinformatics pipeline (2021) (July 10, 2021).
1254 90. K. Katoh, J. Rozewicki, K. D. Yamada, MAFFT online service: Multiple sequence alignment,
1255 interactive sequence choice and visualization. *Briefings in Bioinformatics* **20**, 1160–1166
1256 (2018).
1257 91. Di. Darriba, *et al.*, ModelTest-NG: A new and scalable tool for the selection of DNA and
1258 protein evolutionary models. *Molecular Biology and Evolution* **37**, 291–294 (2020).
1259 92. A. A. Sall, *et al.*, Yellow fever virus exhibits slower evolutionary dynamics than dengue
1260 virus. *Journal of Virology* **84**, 765–772 (2010).
1261 93. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Y. Lam, Ggtree: an R Package for visualization
1262 and annotation of phylogenetic trees with their covariates and other associated data.
1263 *Methods in Ecology and Evolution* **8**, 28–36 (2017).
1264 94. A. M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: A fast, scalable
1265 and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**,
1266 4453–4455 (2019).
1267 95. J. Felsenstein, Confidence limits on phylogenies: An approach using the bootstrap.
1268 *Evolution* **39**, 783–791 (1985).
1269 96. N. D. Pattengale, M. Alipour, O. R. P. Bininda-Emonds, B. M. E. Moret, A. Stamatakis, How
1270 many bootstrap replicates are necessary? *Journal of Computational Biology* **17**, 337–354
1271 (2010).
1272 97. United Nations, United Nations Department of Economic and Social Affairs: Population by
1273 Age Groups–Both Sexes. (2022).

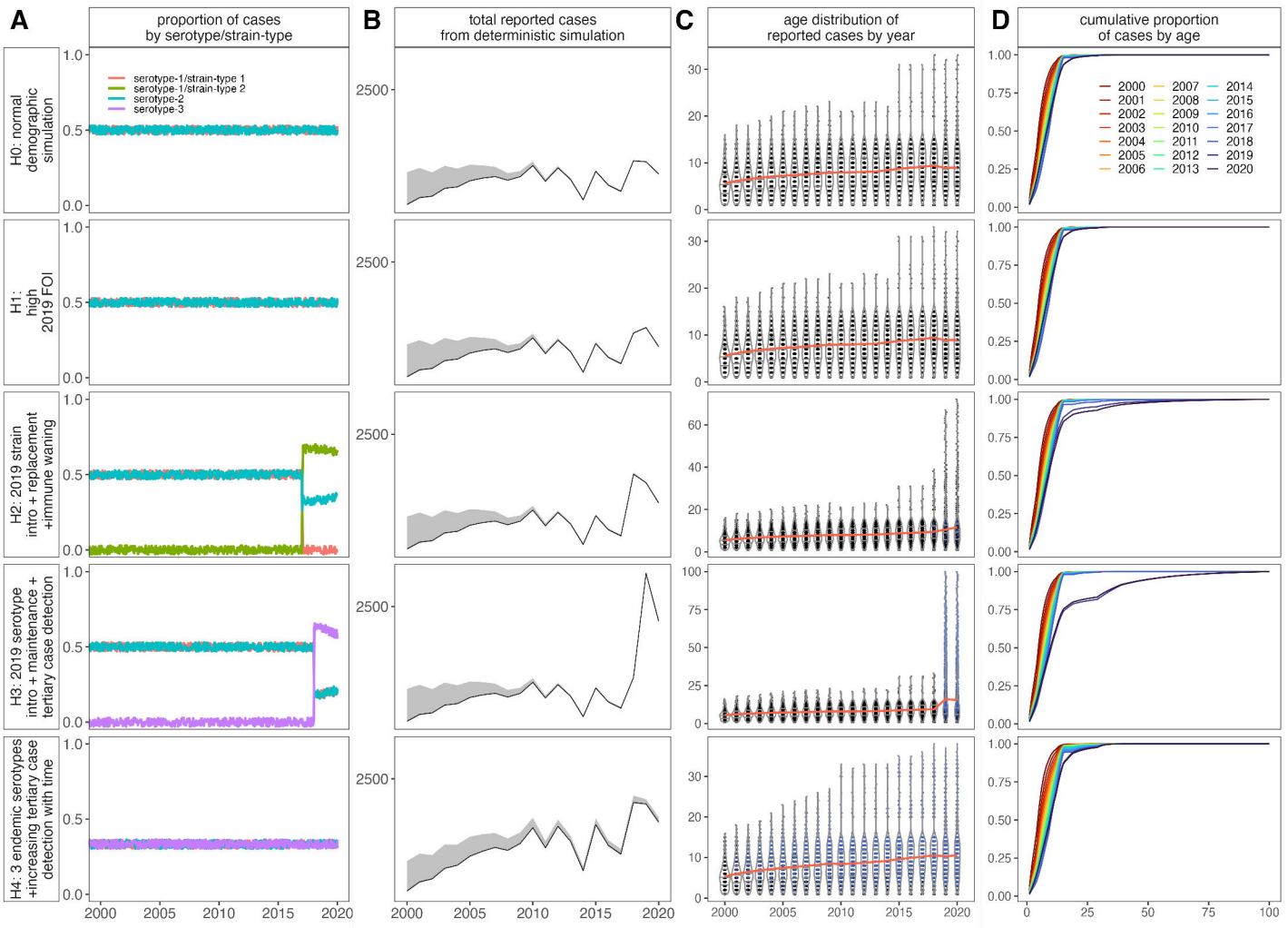
98. United Nations, United Nations Department of Economic and Social Affairs: Deaths by Single Age - Both Sexes. (2022).


mean temperature, 2002-2019 (⁰ C)
28
27
26
TSIR – fit
TSIR – prediction
increased – S – standard – β
no – increase – S – climate – β
increased – S – climate – β
Oddar Meanchey
Siem Reap
Preah Vihear
Banteay Meanchey
Kampong Chhnang
Kampong Thom
Stung Treng
Battambang
Phnom Penh
Kratie
Pailin
Kampong Cham
Pursat
Prey Veng
Takeo
Kampong Speu
Kandal
Koh Kong
Svay Rieng
Preah Sihanouk
Kampot
Kep


A
biweekly annual reconstructed cycle distribution
40
30
20
10
0
-10
C
annual distribution pairwise province Pearson's correlation
0.8
0.6
0.4
0.2
0.0
B
biweekly multi-annual reconstructed cycle distribution
20
0
-20
D
multi-annual distribution pairwise province Pearson's correlation
0.75
0.50
0.25
0.00
Kep
Kampot
Takeo
Preah Sihanouk
Svay Rieng
Prey Veng
Kandal
Koh Kong
Kampong Speu
Phnom Penh
Tboung Khmum
Kampong Cham
Kampong Chhnang
Pursat
Kratie
Mondul Kiri
Kampong Thom
Pailin
Battambang
Siem Reap
Preah Vihear
Banteay Meanchey
Stung Treng
Ratanak Kiri
Oddar Meanchey
2005
2010
2015
2020
Reconstructed cycles
(top: annual; bottom: multi-annual)
-100
0
100
200
Average pairwise province Pearson's correlation
coefficient (top: annual; bottom: multi-annual)
0.00
0.25
0.50
0.75
1.00


A
mean age of dengue infection, 2020
7
10
20
B
age of dengue infection
0
25
50
75
2005
2010
2015
2020
C
births per 1000 ppl
deaths per 1000 ppl
annual force of infection, per capita, λ
1980
1990
2000
2010
2020
D
2002-2010
2011-2020
age multiplier for λ
age
E
cumulative proportion cases (national)
data
model
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
annual rate of waning multitypic immunity, σ
age


A
Genotype I
1950
1975
2000
2020
B
posterior
1.00
0.75
0.50
0.25
0.00
Asian I
Cosmopolitan III
1950
1975
2000
2020
C
Kampong Speu 2019-2020
Salje et al. 2017
Effective # Chains
100
30
10
3
1
10
100
1000
10000
Population Size (x1000)
D
DENV-1
DENV-2
Proportion same transmission chain
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
Max distance between cases (km)


A
proportion of cases by serotype/strain-type
B
total reported cases from deterministic simulation
C
age distribution of reported cases by year
D
cumulative proportion of cases by age
serotype-1/strain-type 1
serotype-1/strain-type 2
serotype-2
serotype-3
H0: normal demographic simulation
H1: high 2019 FOI
H2: 2019 strain intro + replacement +immune waning
H3: 2019 serotype intro + maintenance + tertiary case detection
H4: 3 endemic serotypes +increasing tertiary case detection with time
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020