Unifying the Epidemiological, Ecological and Evolutionary Dynamics of Dengue

A Dissertation Submitted to the Department of Zoology and the

Committee on Graduate Studies of the University of Oxford

in Partial Fulfilment of the Requirements for the

Degree of Doctor of Philosophy, D. Phil.

José Lourenço

Wolfson College

Trinity 2013

Abstract

In under 6 decades dengue has emerged from South East Asia to become the most widespread arbovirus affecting human populations. Recent dramatic increases in epidemic dengue fever have mainly been attributed to factors such as vector expansion and ongoing ecological, climate and socio-demographic changes. The failure to control the virus in endemic regions and prevent global spread of its mosquito vectors and genetic variants, underlines the urgency to reassess previous research methods, hypotheses and empirical observations. This thesis comprises a set of studies that integrate currently neglected and emerging epidemiological, ecological and evolutionary factors into unified mathematical frameworks, in order to better understand the contemporary population biology of the dengue virus. The observed epidemiological dynamics of dengue are believed to be driven by selective forces emerging from within-host cross-immune reactions during sequential, heterologous infections. However, this hypothesis is mainly supported by modelling approaches that presume all hosts to contribute equally and significantly to the selective effects of cross-immunity both in time and space. In the research presented in this thesis it is shown that the previously proposed effects of cross-immunological reactions are weakened in agent-based modelling approaches, which relax the common deterministic and homogeneous mixing assumptions in host-host and host-pathogen interactions. Crucially, it is shown that within these more detailed models, previously reported universal signatures of dengue's epidemiology and population genetics can be reproduced by demographic

and natural stochastic processes alone. While this contrasts with the proposed role of cross-immunity, it presents demographic stochasticity as a parsimonious mechanism that integrates, for the first time, multi-scale features of dengue's population biology. The implications of this research are applicable to many other pathogens, involving challenging new ways of determining the underlying causes of the complex phylodynamics of antigenically diverse pathogens.

for us.

Acknowledgements

I am particularly thankful to Mario Recker for his reassuring presence and for allowing me to explore my ideas and motivations in this work. Past colleagues, such as Ricardo Águas and Daniel Damineli, also deserve mention for the valuable comments and sometimes insensitive criticism. A warm hug to Manuela Cordeiro for quickly reassuring me of the unnecessary self-inflicted stresses I frequently cause myself when challenged with bureaucracy and paperwork. This thesis would not have been possible without the trust put in me by Jorge Carneiro.

Contents

List of Figures

CHAPTER 1

Introduction to the Population Biology of Dengue

The dengue virus (DENV) is now the most widespread mosquito-born virus affecting human populations. During the last few decades it has increasingly become a major public health problem with a significant economic and social impact in South America and South East Asia [Halasa et al., 2012; San Martín et al., 2010]. Insufficient resources for mosquito control in key regions, lack of political will, increasing urbanization, climate change and misguided public health policies are some of the reasons why prevention and control have been generally disappointing [Chareonsook et al., 1999; Gubler, 2002, 2011; Guzman et al., 2010; Halstead, 2008; San Martín et al., 2010].

Whereas hyper-endemicity was classically a unique epidemiological property of South East Asia, globalization trends of the last few decades have changed this situation [Maramorosch et al., 2008; Rico-Hesse, 2010]. Its two main vectors, the *Aedes aegypti* and *albopictus* mosquito-species, have witnessed rapid expansion worldwide allowing for DENV endemicity in more than 100 countries [Halstead, 2007]. The World Health Organization (WHO) now estimates that approximately 3.5 billion people, or ≈55% of the world's population, live in countries affected or on the brink of being affected by dengue (figure 1.1 on page 15) [Gubler, 2002, 2011; Halstead, 2008].

Apart from the fast geographical expansion of recent years, dengue has also increased

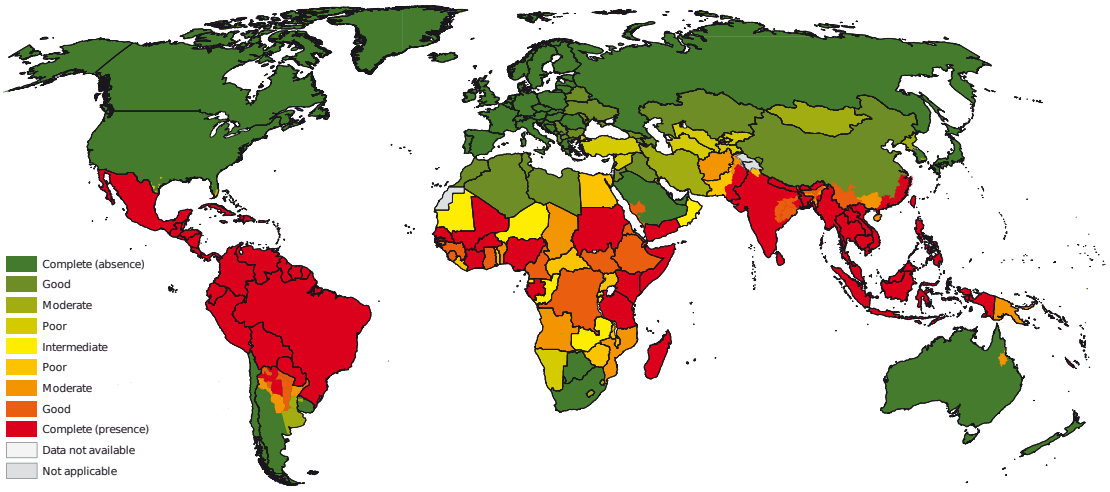


Figure 1.1: **Consensus on Dengue Virus Geographical Presence** - The map combines information from different sources to show the degree of consensus as to whether dengue transmission occurs in each region. Categorization is made over a spectrum: complete consensus on absence in dark green, indeterminate consensus in yellow, and complete consensus on presence in dark red. Figure reproduced from Brady's *et al.* evidence-based literature review [Brady et al., 2012]. The searches were performed for the period 1920-2009 and last updated on the 8th of February 2012. Geo-spatial information was extracted using site descriptions within 2,838 published articles. Data sources were included only if cases were confirmed as resulting from indigenous (not imported) transmission events.

in epidemic activity in long standing endemic regions of South America and South East Asia (figure 1.2 on page 16). Here, most endemic countries have absent or poor dengue surveillance systems such that unexpected epidemics often occur, creating overwhelming situations in health systems [Gubler, 2002]. A common response is the use of ultra low volume (ULV) insecticide sprays to control the vector [Beatty et al., 2010; Horstick et al., 2010]. However, due to the associated high costs, these initiatives are generally of short duration and insufficient to prevent future outbreaks. Furthermore, from the perspective of the local populations, the visible nature of such approaches render a false sense of security, which in time reduces awareness and control against vector breeding sites, in and around habitations [Gubler, 2002; Horstick

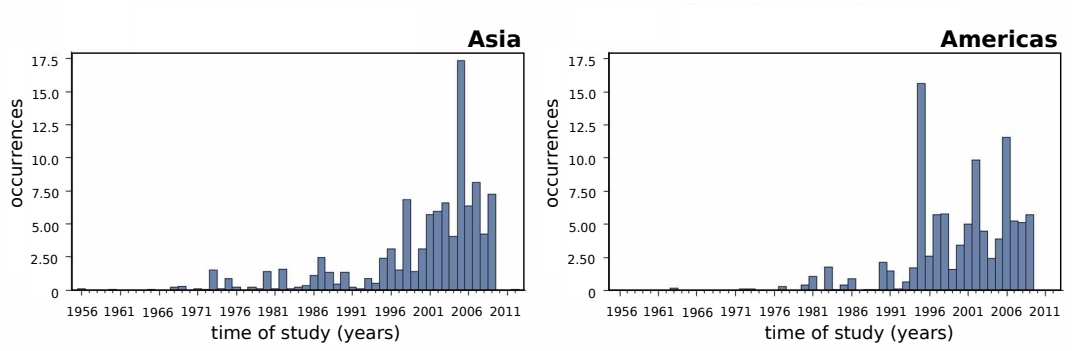


Figure 1.2: **Dengue Outbreak Occurrences in Asia and Americas** - Number of occurrences, i.e. independent dengue outbreaks, in regions of Asia, North and South America. Asia excludes regions of Australasia. A steep increase can be seen in the last 2 decades. Figure reproduced from Brady's *et al.* evidence-based literature review [Brady et al., 2012]. The searches were performed for the period 1920-2009 and last updated on the 8th of February 2012. Geo-spatial information was extracted using site descriptions within 2,838 published articles. Data sources were included only if cases were confirmed as resulting from indigenous (not imported) transmission events.

et al., 2010; Spiegel et al., 2010].

The epidemiological picture of dengue in other areas of the world is variable. In Africa, dengue has always been overshadowed by other disease related problems and characterized by poor surveillance with no official reporting to the World Health Organization. Recently, outbreaks of dengue-like illness were reported in dry countries of Eastern Africa and South West Asia, such as Pakistan, Saudi Arabia, Yemen, Sudan and Madagascar [Guzman et al., 2010]. In 2009, two major epidemics occurred in the isolated and dengue-free populations of Cape Verde [Franco et al., 2010] and Mauritius islands [Issack, 2010]. Given these reports and the presence of the *aegypti* and *albopictus* vector-species, the current epidemiological situation in Africa is believed to be worse than previously acknowledged [Gubler, 2002; Guzman et al., 2010; Kyle and Harris, 2008; Maramorosch et al., 2008]. In North America, dengue has

been endemic in Mexico and has recently emerged in Florida and Texas in the United States, following multiple reports on the settlement of *Aedes aegypti* [Guzman et al., 2010]. In Europe, sporadic epidemics in Italy, France and Greece have been described in recent years, some attributed to *Aedes albopictus* [Cavrini et al., 2009; Gautret et al., 2010; Gjenero-Margan et al., 2011]. The increasing number of reports on autochthonous and imported infections to non-endemic countries highlight the need to study the ecological barriers that determine DENV's invasion success.

Of extreme importance is the worldwide emergence of previously unreported severe forms of disease, cases of which first appeared in Thailand and the Philippines in the 1950s, turning into a worldwide phenomenon within two decades [Gubler, 2002]. Some researchers believe this to be partially explained by increased epidemic activity, which could have led to selection of viral variants with greater epidemic potential and concurrently more virulent phenotypes [Gubler, 2002; Rico-Hesse, 2010]. In contrast, others believe this to be the result of complex immuno-reactions dependent on the host's history of infection and increasing propensity of sequential challenge, promoted by heightened persistence and co-circulation of viral variants [Halstead and O'Rourke, 1977; Rothman and Ennis, 1999].

1.1 The Dengue Virus

Classification based on genomic organization, homology and antigenic cross-reactivity places the dengue virus (DENV) into the *Flavivirus* genus of the viral family *Flaviviridae* [Calisher et al., 1989]. More than 70 flaviviruses have been so far identified [Kuno et al., 1998], many of which are important human pathogens, including the hepatitis C virus (HCV), the Japanese encephalitis virus (JEV) and yellow fever virus (YFV). Amongst the ones that require blood-feeding mosquitoes for transmission, DENV is responsible for the highest human morbidity and mortality [Maramorosch et al., 2008]. The pathogen's population comprises four antigenically distinct viral groups (serotypes DENV1-4, figure 1.3 on page 19) that present similar population dynamics and are the cause of clinically indistinguishable illnesses in humans. Serotypes were initially defined by their limited cross-reactivity in various serological tests [Calisher et al., 1989]. More recently, nucleic acid sequencing techniques have confirmed sequence similarity and also identified genetic subgroups or lineages, commonly denominated as dengue genotypes (figure 1.4 on page 20) [Weaver and Vasilakis, 2009]. The overall amino acid sequence similarity among serotypes ranges between $\approx 60 - 80\%$ [Irie et al., 1989].

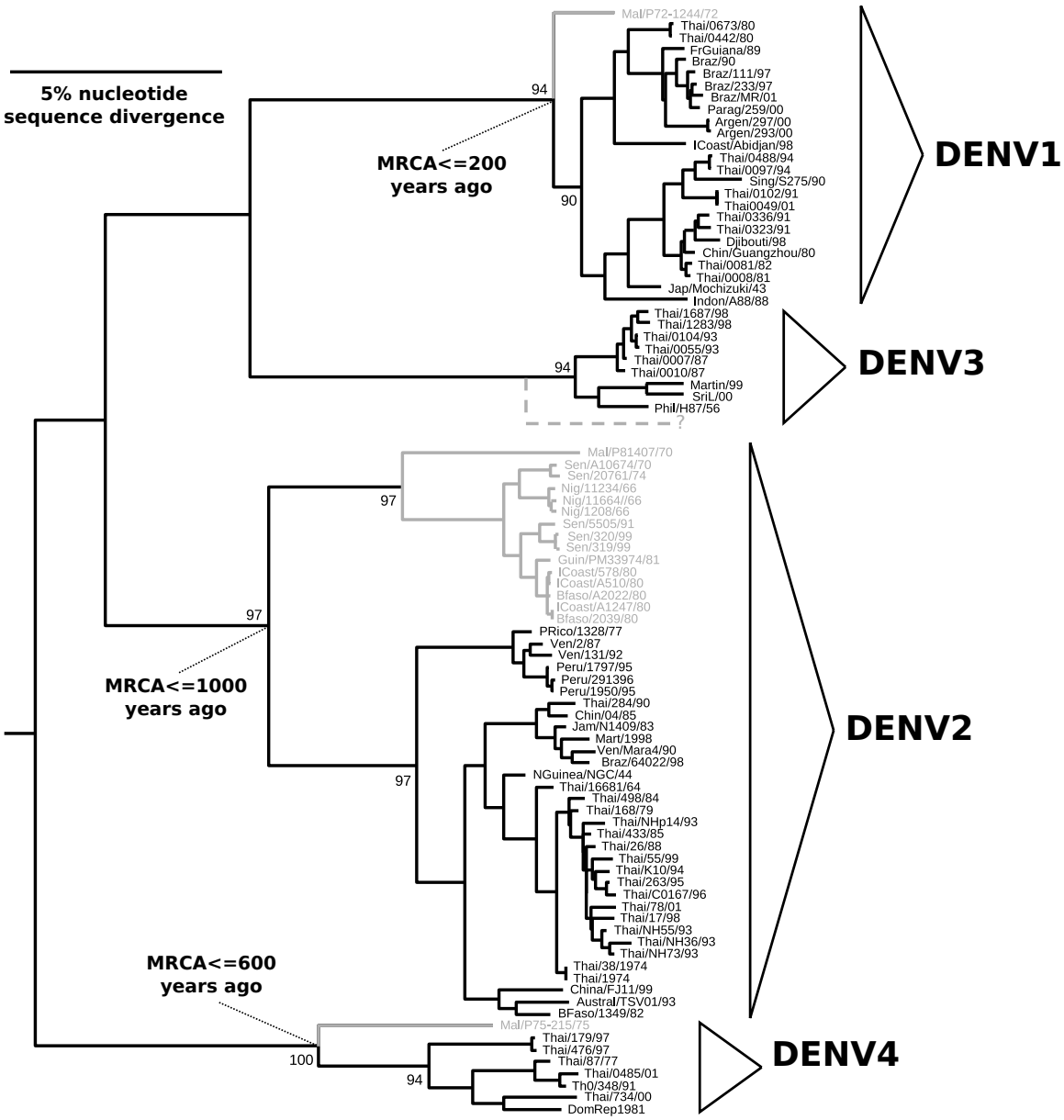


Figure 1.3: **Phylogeny of the Dengue Virus (serotypes)** - Phylogeny of all DENV strains isolated and available in the GenBank library (until 2008/2009). Serotypes are identified on the far right. Strains are either in grey or black, belonging to the sylvatic and epidemic cycle, respectively. Sylvatic strains of DENV3, coded with "?", have not been isolated but believed to circulate in Malaysia [Maramorosch et al., 2008]. Reproduced from figure 2 of Weaver *et al.* with permission from the authors [Weaver and Vasilakis, 2009]. **Original legend:** *Phylogenetic tree of DENV strains from all 4 serotypes derived from complete open reading frames available in the GenBank library. The phylogeny was inferred using Bayesian analysis (1 million reiterations) and all horizontal branches are scaled according to the number of substitutions per site. Bayesian probability values are shown for key nodes. Virus strains are coded by abbreviated country of collection/strain name/year of collection.*

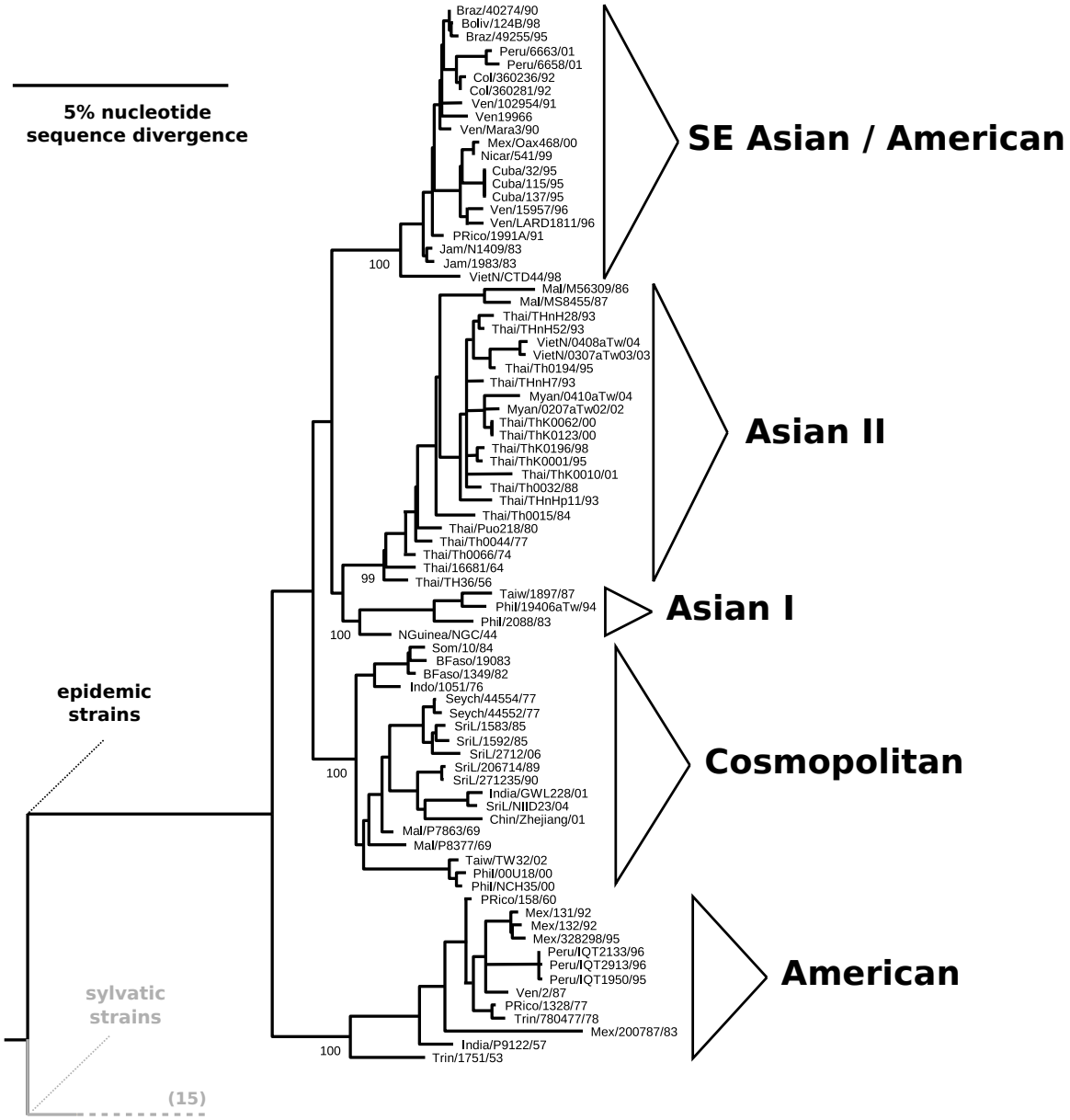


Figure 1.4: **Phylogeny of the Dengue Virus Serotype 2 (genotypes)** - Phylogeny of all DENV2 strains isolated and available in the GenBank library (until 2008/2009). Genotype classification/name is on the far right. Strains that are identified in black belong to the epidemic cycle. 15 sylvatic strains are hidden for visualization purposes (dashed branch, grey). Genotypes were originally defined by Rico-Hesse *et al.* as clusters of viruses with sequence divergence lower than 6% within a short genome region in the E/NS1 genome junction (figure 1.5 on page 21) [Rico-Hesse, 1990]. Reproduced from figure 5 of Weaver *et al.* with permission from the authors [Weaver and Vasilakis, 2009]. **Original legend:** *Phylogenetic relationships of DENV-2 strains from the GenBank library. The phylogeny was inferred based on complete E gene nucleotide sequences and Bayesian analysis (1 million reiterations) using MrBayes 3.1. Bayesian probability values are shown for key nodes. Virus strains are coded by abbreviated country of collection/strain name/year of collection.*

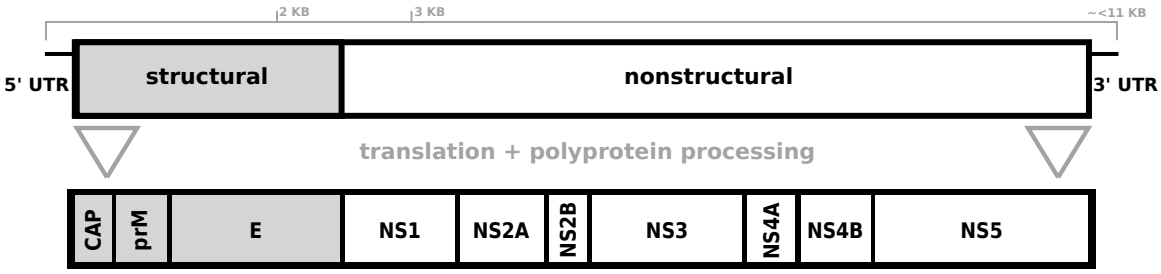


Figure 1.5: **DENV Genome and Viral Proteins** - Translation of the open reading frame results in a single polyprotein that is cleaved by proteases from the host and virus to produce both structural (shaded grey) and nonstructural (white) proteins. The E-protein is responsible for binding and fusion and is involved in key determinants of host range, tropism and virulence. A short sequence of the genome region in the E/NS1 junction is used to define DENV genotypes as originally proposed by Rico-Hesse *et al.* [Rico-Hesse, 1990]. Diagram based on figure 1 from [Weaver and Vasilakis, 2009] and figure 1 from [Fernandez-Garcia et al., 2009].

A single positive-sense RNA molecule of ≈11 kilobases in length, encoding one open reading frame, composes DENV's genome (figure 1.5 on page 21). The structural proteins - core, premembrane and envelope - constitute the viral particle and are the main target of antibody-mediated responses. The remaining seven non-structural proteins are involved in viral replication, assembly and modulation of host cell responses [Fernandez-Garcia et al., 2009]. Experimental recombinant variants have not yet been reproduced, and the role of recombination in the population biology of DENV remains controversial [Maramorosch et al., 2008; Weaver and Vasilakis, 2009].

Although RNA-based viruses generally present rapid evolution occurring at a similar time-scale as their host's ecological dynamics [Grenfell et al., 2004], studies have demonstrated that strong purifying selection dominates DENV's recent evolution [Bennett et al., 2003; Holmes, 2003; Holmes and Twiddy, 2003; Weaver and Vasilakis, 2009; Zhang et al., 2005]. A prevailing hypothesis is that vector-borne

flaviviruses have evolved to replicate efficiently in both the vertebrate and arthropod hosts, expressing a compromise genome in which most amino acid changes are expected to be deleterious and selectively removed from the population [Holmes, 2003; Holmes and Twiddy, 2003; Weaver and Vasilakis, 2009]. The lack of recombination further limits the fixation and emergence of single, positively selected mutations due to clonal interference [Grenfell et al., 2004; Miralles et al., 1999]. Since mutations that take place in one host may be selected against in the other, one of the major implications of purifying selection is low, nonsynonymous diversity [Woelk and Holmes, 2002]. Therefore, as described by Holmes *et al.*, more variation can be found within hosts than amongst them [Holmes, 2003]. The consequences of deleterious pressures are illustrated in figure 1.6 (on page 23), which compares the genetic diversity along the genomes of DENV and hepatitis C virus (HCV), a directly transmitted flavivirus. Two main differences are observed. First, in sharp contrast with HCV, diversity is uniformly distributed along DENV's genome, suggesting weak or no asymmetries in selective pressures, even between structural and functional genes. Second, HCV's antigenic types present slightly different landscapes, indicative of dissimilar evolutionary histories or selective pressures amongst the variants.

At the same time, reconstruction of DENV's evolutionary history has been difficult due to the weak resolution of the *Flaviviridae* phylogeny [Kuno et al., 1998; Weaver and Vasilakis, 2009]. Observations were made in the late 1970s of dengue-like viruses and respective animal reservoirs in Asia and West Africa [Aiken and Leigh, 1978].

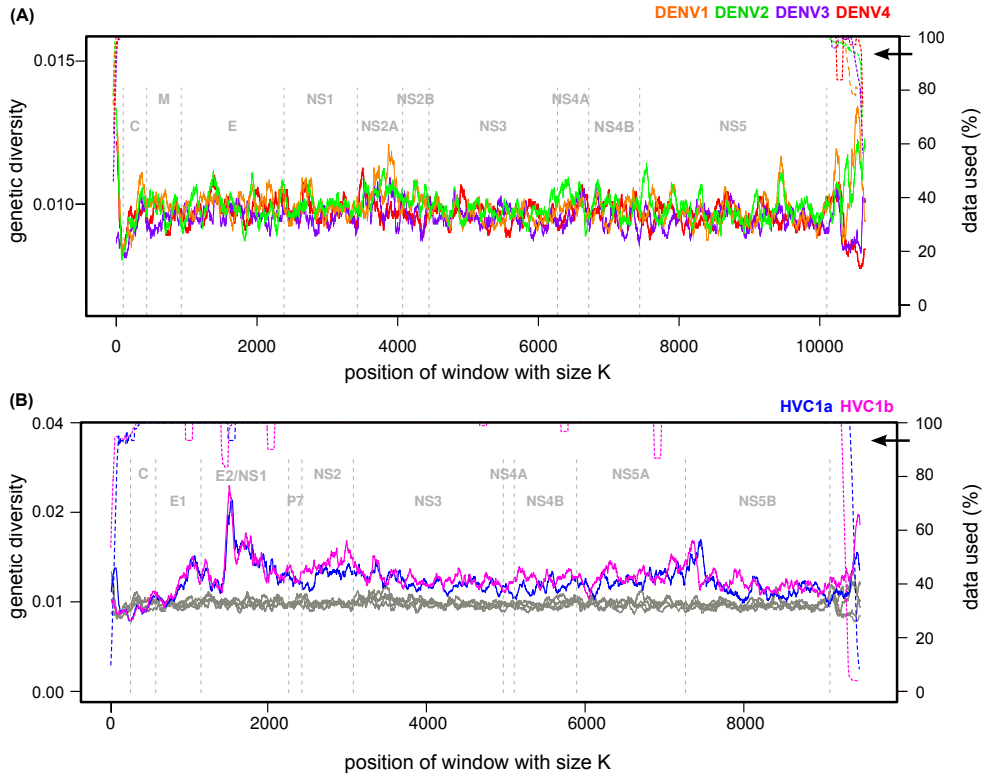


Figure 1.6: **Genetic Diversity of HCV and DENV Antigenic Types** - The genetic diversity of DENV1-4 and hepatitis C virus (flavivirus) genotypes 1a (HCV1a) and 1b (HCV1b) is presented. The different genome segments are identified between vertical, dashed lines. Diversity, D, is estimated with a sliding window approach (in this example, windows are K=90 nucleotides). There are $(N-K)+1$ windows in a genome of size N. D is the average, given $Nseqs$ sequences, of the mean number of different nucleotides (nDN) found in all positions i of a given window w (see equation bellow). For each variant, 100 different full genome sequences were randomly selected from data available in Genbank. The percentage of data used (Y-axis, right) recognizes gaps in the data. **(A)** DENV data contained sequences from Africa (AF), South America (SA), South East Asia (SEA), East Asia (EA), the Caribbean (CA) and North America (NA), for a period spanning 1956-2011. **DENV1** - SA, NA, AF, CA, EA, SEA; 1987-2009. **DENV2** - SA, SEA, NA, CA; 1972-2009. **DENV3** - SA, SEA, CA; 1983-2010. **DENV4** - SA, SEA, EA; 1956-2011. **(B)** HCV data contained sequences from the United States of America (USA), China (CH) and Europe (EU) for a period spanning 1989-2008. **HCV1a** - USA, EU; 1989-2007. **HCV1b** - USA, CH, EU; 1989-2008. For comparison, DENV1-4 diversity is presented in grey.

$$D_w = (\sum_{i=1}^{K} nDN)/K/Nseqs$$

Since then, a wide number of sylvatic strains (i.e. of animal origin) have been isolated and reported to belong to each of the human-bearing serotypes, proposing a zoonotic origin (figure 1.3 on page 19) [Holmes and Twiddy, 2003]. Introduction of dengue (serotypes) into the human population was extrapolated to have taken place at differ-

ent time points, with DENV2 being the oldest with endemic potential. The age of (DENV2's) most recent common ancestor was further estimated to be approximately 1000 years, roughly corresponding to some of the first reports of dengue-like illnesses (Section 1.2 on page 24) [Twiddy, 2003].

It is likely that the genetic distance between currently circulating serotypes is a by-product of allopatric evolution due to ecological partitioning of ancestral, sylvatic strains [Holmes and Twiddy, 2003; Weaver and Vasilakis, 2009]. This hypothesis is sustained not only by DENV's current population structure, but also by studies that report sylvatic and endemic (i.e. human bearing) strains to exhibit similar evolutionary rates [Twiddy, 2002]. Furthermore, experimental surrogates for human and vector infection have shown the presence of small reproductive differences [Vasilakis et al., 2007], which suggests that their infection potential is similar and the low number of reported sylvatic cases may be due to ecological barriers.

1.2 Endemic Transmission Cycle

The oldest record of dengue-like illness dates from the 3^{rd} century of the common era in the Chinese literature [Holmes and Twiddy, 2003]. More recently, historical records propose that between the 19^{th} and the second decade of the 20^{th} centuries, at least five pandemics of dengue-like illness have occurred amongst Africa, India, Oceania and the Americas [Maramorosch et al., 2008; Weaver and Vasilakis, 2009]. Epidemics

have been described around areas with large reservoirs of water, along coastal lines, inland upstream rivers and seaports. Even though the clinical descriptions fit with contemporary diagnosis of dengue, other mosquito-borne viral diseases with similar symptomatic manifestations, such as chikungunya or yellow fever, might have been the cause of some of the described occurrences. Awareness of dengue as a public health issue is recent and a growing phenomenon initially promoted by events that took place during and immediately after World War II (WWII) [Maramorosch et al., 2008]. Movement of supplies and military vehicles over long distances resulted in the importation of the mosquito vectors into new geographical regions. The last years of the war coincided with reports of several epidemic events affecting military personnel in regions such as the Caribbean, East Africa, Japan, New Guinea, Hawaii and Australia [Maramorosch et al., 2008; Weaver and Vasilakis, 2009]. The end of WWII in the Pacific setting led to general patterns of uncontrolled urbanization. Millions of individuals moving into urban areas with inadequate water distribution systems resulted in domestic water storage practices that still persist. Such demographic changes allowed for the establishment of high population densities of the *aegypti* mosquito. At the same time, the absence of reports in the Americas was due to the initiation of eradication measures against *aedes* mosquitoes by the Pan American Health Organization (PAHO). Successful eradication of *aegypti* was achieved in most countries, but relaxation and complete suspension of the program in the late 1960s led to a gradual reinfestation of almost the entire South American continent [Maramorosch et al., 2008]. Not surprisingly, this resurgence led to the reintroduction of DENV. Its epi-

demiology in the Americas until the beginning of the 1980s was characterized by the absence of severe epidemics and the circulation of single serotypes at any point in time and space [Maramorosch et al., 2008; Weaver and Vasilakis, 2009]. This trend changed with the arrival of viral variants from South East Asia. By the end of the 20th century, multiple DENV serotypes were present in most endemic regions with the annual incidence of dengue and severe haemorrhagic episodes increasing dramatically every year [San Martín et al., 2010].

As described earlier, transmission of DENV to humans is mainly attributed to the mosquito vector *Aedes aegypti*. Other species, which generally appear to express reduced susceptibility, a longer extrinsic incubation period (EIP) and lower transmission capabilities, have also been identified as viable vectors [Anderson and Rico-Hesse, 2006; Halstead, 2008; Paupy et al., 2009]. *Aedes albopictus*, considered DENV's secondary vector, has gained notoriety in the last decade after observations of fast worldwide spread and sporadic sustained transmission in Europe and the Americas [Gautret et al., 2010; Gjenero-Margan et al., 2011; Guzman et al., 2010]. A second transmission route occurs in healthcare-related environments through unsafe management of blood products [de Wazières et al., 1998], although the frequency of such occurrences and epidemiological importance are largely unknown [Wilder-Smith et al., 2009]. *Aegypti's* breeding and feeding behaviour allows it to surpass in epidemiological importance all other viable vector-mosquito species. Individual mosquitoes are highly domesticated and lay eggs in any type of artificial water container found within

and around human habitations, which work as safe and large-number producing spots. Adults are optional hematophages, mainly feeding on pollen, fruit juice and other biological fluids. Females require blood for egg production and are anthropophilic with a characteristic 'nervous' feeding behaviour that generally leads to multiple bites per single blood meal, possibly in different individuals [Scott et al., 2000]. Reports of limited susceptibility within and between *aegypti* populations suggests that its recent success as an epidemic vector might be due to behavioural dynamics and current ecological context rather than viral adaptation to the species [Anderson and Rico-Hesse, 2006; Maramorosch et al., 2008].

Dengue incidence patterns within endemic regions generally present epidemic cycles highly correlated with seasonal variations in vector population size (figure 1.7 B on page 28), in addition to oscillatory prevalence levels of its serotypes (figure 1.7 A on page 28) [Nisalak et al., 2003]. Incidence patterns are driven by a complex interplay of local factors, from the abundance of susceptible humans and their age, to climate, viral traits, density and distribution of vector individuals [Johansson et al., 2009; Morrison et al., 2010; Nagao et al., 2008; Tsuzuki et al., 2009; Williams et al., 2010]. However, none of these seem to directly influence the qualitative properties of serotype cyclical behaviour, which seem to be a universal feature of dengue epidemiology. Given the presence of strong, long lasting, seasonal epidemic troughs, a subject of great discussion is how viral persistence is achieved. Phylogenetic analysis show that the absence of reported cases can be followed by resurgence of pre-

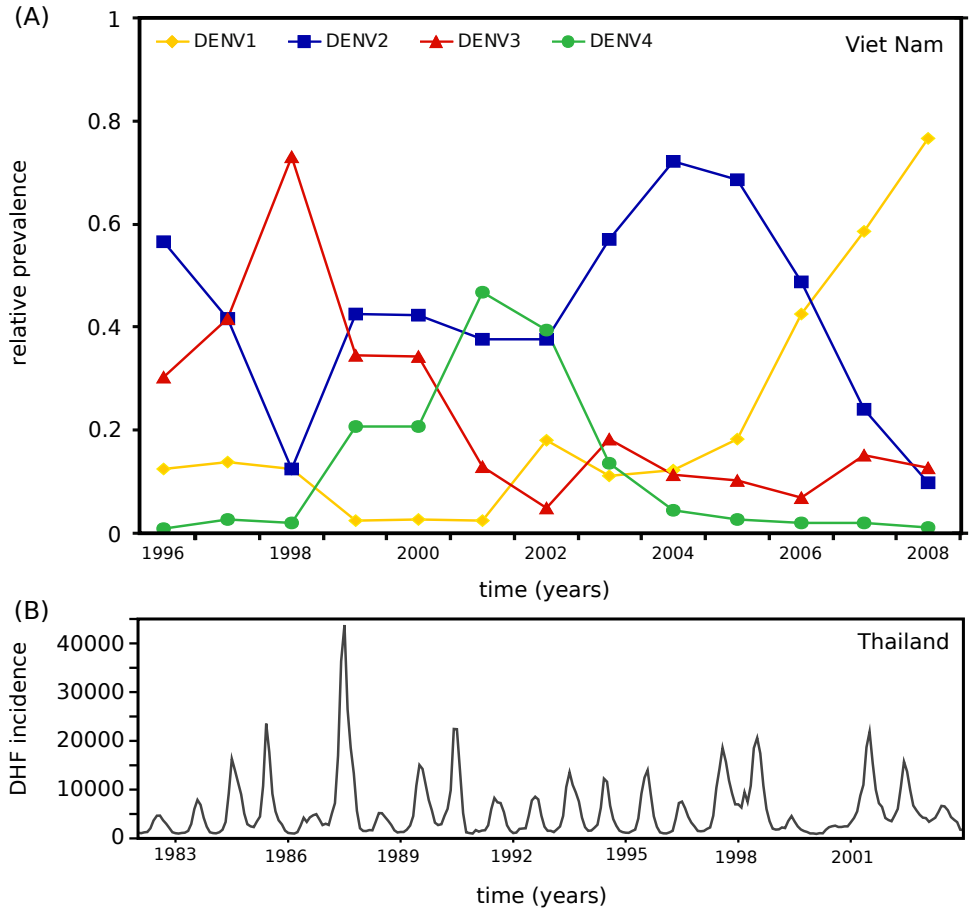


Figure 1.7: **Dengue Epidemiological Patterns** - **(A)** Total number of hospitalised cases between 1994-2008 show the sequential replacement in dominance of dengue's four co-circulating serotypes (DENV1-4). Data from [Lourenço and Recker, 2010]. **(B)** Monthly DHF incidence reveals a strong seasonal component, which is not reflected in serotype dynamics. Description of the data can be found in Section 1.5 on page 58.

viously circulating genetic variants, proposing persistence in the vector population, reintroduction from nearby locations or continuous silent transmission [Bennett et al., 2010, 2003; Carrington et al., 2005]. The latter could be due to the high number of asymptomatic DENV infections (Section 1.4 on page 30) that shadow the true levels of ongoing transmission during off-season periods. Therefore, the frequent absence of clinical cases might not necessarily reflect viral extinction. A possible contribut-

ing factor for persistence could also be vertical transmission (from female adults to offspring), which has been proposed for multiple mosquito species, including *aegypti* and *albopictus* [Bosio et al., 1992], and may enhance DENV's persistence by providing a human-independent reservoir for the pathogen. There are, however, contrasting field and laboratory measures of its efficiency [see Table 1 in Adams and Boots, 2010, for review], possibly due to experimental artificial selection or asymmetries between viral and mosquito strains. Nonetheless, recurrent loss of diversity has been reported in the viral population [Bennett et al., 2010; Raghwani et al., 2011], which demonstrates that population bottlenecks are real. Although these are generally attributed to ecological features, such as fluctuations in vector densities, viral factors also appear to be relevant. For instance, the extrinsic incubation period is temperature dependent due to constraints in viral replication, becoming longer during cold seasons and therefore restricting the virus' transmission potential [Watts et al., 1987].

1.3 Sylvatic Transmission Cycle

The transmission cycle of sylvatic viruses takes place in forest environments of South East Asia and West Africa between nonhuman primates and arboreal *Aedes* mosquitoes (Section 1.1 on page 18, and figure 1.1 on page 22). Rural populations living in close proximity to forested or plantation areas are periodically exposed to such variants [Holmes and Twiddy, 2003; Weaver and Vasilakis, 2009]. Studies have shown that the canopy-dwelling *Aedes furcifer* is not only highly susceptible to sylvatic DENV

but also regularly disperses from the forests to human villages, suggesting a role in sporadic primary transmission [Maramorosch et al., 2008]. Secondary transmission is rarely reported, which may be a result of limited human-to-vector transmission potential or underdetection due to low pathogenicity. An alternative explanation could be the rarity of peridomestic vectors (*aegypti* and *albopictus*) in such small human populations. Both the ability of peridomestic mosquitoes to transmit sylvatic DENV and the lack of evidence that genetic adaptation is required for sylvatic strains to replicate efficiently in humans, propose an ecological barrier for the rarity of zoonotic transmission [Vasilakis et al., 2007].

1.4 Human Infection, Immunity and Disease

Human infection with any of the 4 serotypes is generally mild with the majority of cases believed to be asymptomatic [Kyle and Harris, 2008]. It results in a short-lived febrile illness commonly denominated as *dengue fever* (DF), generally characterized by arthralgia (joint pain), myalgia (muscle pain), low-grade fever and retro-orbital headaches [Balmaseda et al., 2006; Burke et al., 1988; Wichmann et al., 2011]. In some cases it may progress to more severe and life-threatening forms - *dengue haemorrhagic fever* (DHF) or *dengue shock syndrome* (DSS) - with manifestations of circulatory failure, vascular permeability and haemorrhagic symptoms [Gubler, 1998; Halstead, 2007].

Several risk factors for developing DHF/DSS have been described, including: host genetic background [Gubler, 1998; Guzman and Kouri, 2008], secondary heterologous challenge [Dejnirattisai et al., 2010; Halstead et al., 2010; Halstead and O'Rourke, 1977; Rothman and Ennis, 1999], viral genotype [Armstrong and Rico-Hesse, 2001; Cologna et al., 2005], order of infecting serotype [Gibbons et al., 2007; Guzman and Kouri, 2008; Nisalak et al., 2003], time between infections [Guzman and Kouri, 2008; Kyle and Harris, 2008] and age of infection [Anders et al., 2011; Nisalak et al., 2003]. From these, the most widely cited factor is secondary heterologous infection, with strong evidence that the risk of developing severe disease can be several times higher when compared to primary infection [Burke et al., 1988; Guzman and Kouri, 2008; Kyle and Harris, 2008; Sangkawibha et al., 1984; Thein et al., 1997].

A typical first immune reaction is characterized by an initially slow antibody response during which Immunoglobulin M (IgM) antibodies (the generic response to the presence of any antigen), are the first to arise (figure 1.8 on page 32). DENV specific Immunoglobulin G (IgG) antibodies are generally detected at low titre by the end of the first week, thereafter increasing until week two. Antibody concentration then slowly decreases, stabilizing at levels that confer life-long *homologous immunity* (i.e. sero-specific protection to reinfection) [Halstead, 2007]. A few studies have demonstrated the existence of sero-specific immune responses lasting for up to 40 years in individuals from Cuba [Guzman et al., 2007; Sierra et al., 2002], Japan and Greece [Kyle and Harris, 2008].

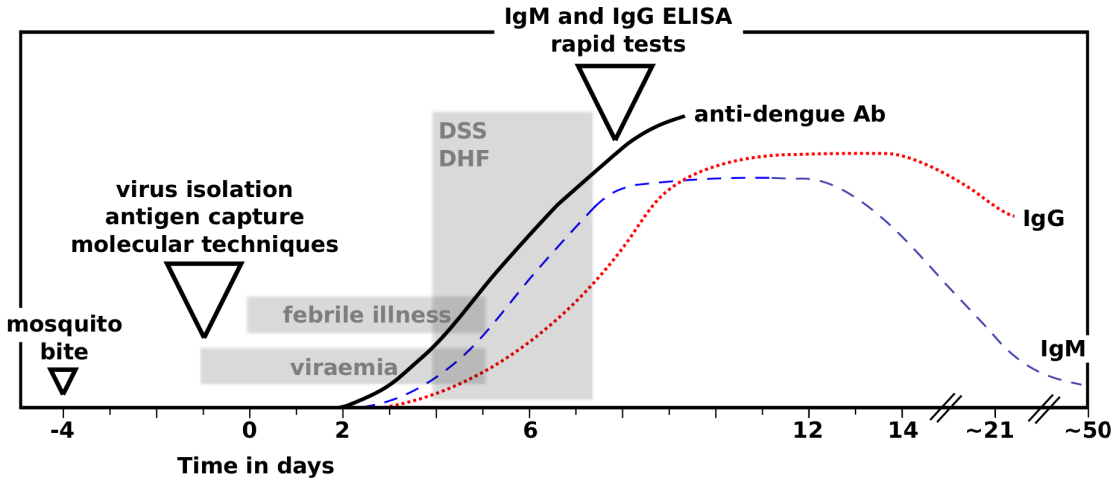


Figure 1.8: **Immune Response to DENV Primary Infection** - Typical progression of primary challenge with DENV in human hosts. IgG and IgM stand for Immunoglobulin G (responsive to any exposure to antigen) and Immunoglobulin G (DENV specific), respectively. ”Febrile illness” can manifest either as DF or mild/asymptomatic disease prone to be missed by passive surveillance systems. Progression to severe disease (DHF/DSS) is rare and generally associated with post-viraemia period (Section 1.4 on page 30). Figure reproduced and changed with permission from Scott Halstead [Halstead, 2007].

In addition to *homologous immunity*, Sabin *et al.* suggested the existence of a short period of cross-protective immunity to all serotypes upon recovery to infection (transcending immunity). Evidence was collected after World War II in human volunteers, which described complete heterologous immunity lasting up to 2 months and partial immunity up to 9 months after primary infection [Sabin, 1952]. In more recent studies, for example, the neutralization of DENV2 viruses by sera containing monospecific DENV1 antibodies has also been demonstrated *in vitro* [Kochel et al., 2005, 2002]. In this instance, however, immunity did not appear to prevent infections but rather down-regulates their clinical outcome.

So far there are contrasting reports and no consensus on the qualitative behaviour of acquired immunity after secondary infection. This is mainly due to the rarity of third and fourth infections in clinical data and difficulties in assessing the history of infection of single individuals given short periods of viraema and high antibody cross-reactivity (see bellow) [Gibbons et al., 2007; Nisalak et al., 2003]. It is therefore commonly assumed that recovery from secondary infection offers protection to further challenges. Whether this completely prevents infection or merely provides clinical immunity, is still an open question.

Although reports on concurrent human infections with more than one DENV serotype are present in the literature, not much is known on its impact in disease progression and outcome [Bharaj et al., 2008; Loroño Pino et al., 1999].

Flaviviruses are known to induce the phenomenon of *original antigenic sin* (OAS) in which B-cell clones sequentially challenged by viruses of the same family will initially respond by synthesizing antibodies with higher affinity to a previous infecting virus [Francis, 1960; Halstead et al., 1983]. Such subneutralizing antibodies, when attached to viral particles, form an immune complex that can facilitate the infection of monocytes, macrophages and dendritic cells through their *fragment crystalizable* (Fc) receptors [Dejnirattisai et al., 2010; Halstead and O'Rourke, 1977; Jessie et al., 2004; Rothman and Ennis, 1999; Whitehorn and Simmons, 2011]. Cell entry via such routes further suppresses innate immune responses and activates interleukin-10 production

at an early phase of infection [Green et al., 1999a,b; Ubol and Halstead, 2010]. This creates a cascade of regulatory changes that facilitates expression and synthesis of pro-inflammatory cytokines, and also biases *cellular-based* (TH1) to *humoral-based* (TH2) responses, inducing a positive feedback that favours an increase in the rate of intracellular infection and consequently in viraemia levels [Fernandez-Garcia et al., 2009; Whitehorn and Simmons, 2011].

In dengue infection, this cascade is recognized and commonly referred to as *antibody-dependent enhancement* (ADE) and can occur during heterologous DENV challenges [Balsitis et al., 2010; Halstead et al., 2010; Wrammert et al., 2012] or in the presence of transplancental antibodies of maternal origin [Chau et al., 2009]. As a result, secondary challenges are generally dominated by cross-reactive immunological responses [Bukowski et al., 1989; Dejnirattisai et al., 2010], as well as higher viraemia levels [Libraty et al., 2002a; Vaughn et al., 2000] and more severe disease manifestations [Gubler, 1998; Rothman and Ennis, 1999]. Some researchers have proposed that the delayed specific response due to OAS may increase the chances of establishing infection upon an infectious mosquito bite, while the ADE-mediated increase in viraemia may affect the human-to-vector transmission probability per bite. Both these hypotheses - *increased susceptibility and transmissibility in heterologous infections* - have been widely explored in modelling approaches to DENV's population dynamics (Section 1.5 on page 37). Importantly, viral titres have been shown to peak earlier and clearance is faster in secondary challenges, a possible consequence of the OAS

phenomenon [Tricou et al., 2011].

At the same time, the relatively high sequence similarity between DENV1-4 also contributes significantly to the cross-reactivity of memory T-cells [Bukowski et al., 1989]. This has been explored by Rothman and Ennis to propose a model of immunopathogenesis of vascular permeability, a key signature of severe disease manifestations [Rothman and Ennis, 1999]. The model posits that upon monocytic cell infection, either through normal or ADE-mediated means, these target cells directly stimulate T-cells that proliferate and produce pro-inflammatory type 1 cytokines, which directly affect vascular endothelial cells to produce plasma leakage. The correlation of the latter with secondary challenge could then be partially explained by the quicker kinetics in activation of cross-reactive memory T-cells extant from first infection. The three main expectations of this model have been explored in several studies, which demonstrated that secondary challenges developing into DHF/DSS often present (i) higher T-cell activation [Green et al., 1999a; Kurane et al., 2011], (ii) pro-inflammatory cytokine storms [Green et al., 1999a; Kurane et al., 2011], and, (iii) preferential expansion of cross-reactive T-cells [Loke et al., 2001; Mathew et al., 1998]. Nonetheless, the phenomenon of increased risk for severe disease in infants cannot be easily reconciled with this model, as in these cases, the individual's immune reactions to infection are not ruled by cross-reactive TH1 responses but are rather antibody-based of maternal origin [Chau et al., 2009].

Discriminating between the protective and potentially disease enhancing role of TH1 and TH2 responses is difficult due to inadequacies of experimental animal models. For example, research in immunocompetent mice has shown that genetic differences in antibody Fc regions deny the ADE phenomenon [Balsitis et al., 2010] and that DENV replication is poor [Gubler, 1998]. The use of genetically modified host lineages is a partial solution [Cassetti et al., 2010; Shresta et al., 2006] but introduces doubts about the generalization of the disease-related immune responses to the human counterpart. On the other hand, research conducted on nonhuman primates has demonstrated that these animals do not develop disease that resembles DF/DHF. Also, these experimental models are expensive and suffer from a limited characterization of DENV-specific immune responses, restricting their usefulness for understanding the human counterpart [Cassetti et al., 2010].

Finally, there are major difficulties in assessing the history of infection of single human individuals. For example, in both primary and secondary challenge the period of viraemia coincides with the development of mild disease but not necessarily with manifestation of severe symptoms, as the latter tends to be associated with periods of exacerbated immune response which typically appear after the virus has been cleared. Given that mild manifestations of dengue are easily mistaken for other uncomplicated febrile illnesses and that individuals generally seek medical assistance only upon severe symptoms, case detection based on passive surveillance systems is difficult as the majority of infections are not detected or being so too late for virus isolation. Sero-

logical assays are thus the most common method for diagnosis given their inexpensive nature and feasibility for longer periods after exposure [Guzman et al., 2010; Halstead, 2007]. However, persistence of IgG, which cross-react between DENV serotypes, severely hampers the evaluation of the host's history of infection via Hemagglutination assays [Halstead, 2007]. An emerging alternative are the *non-structural protein 1* (NS1) enzyme-linked immunosorbent assays (ELISA). This soluble protein is highly persistent in the blood after the onset of symptoms [Tricou et al., 2011] and has also been implicated with the *complement system* and pathways related to vascular leakage [Avirutnan et al., 2006; Lin et al., 2012]. Yet, its role in development of severe disease is still highly debatable. For instance, NS1 levels in plasma can correlate with viral titres and may be higher in patients with DHF [Libraty et al., 2002b], but its clearance is faster in secondary infections [Tricou et al., 2011]. The protein's kinetics also appear to be dependent of the infecting serotype [Duyen et al., 2011; Tricou et al., 2011] and there are reports of low concentrations during DENV2 infections, a serotype generally acknowledged to be statistically linked to severe disease [Rico-Hesse, 2010; Vaughn et al., 2000].

1.5 Theoretical Approaches

Mathematical models of disease transmission are useful tools to analyse epidemiological and experimental data. They are also crucial simulators of complex scenarios that allow to circumvent real world constraints such as lack of resources or hazard effects

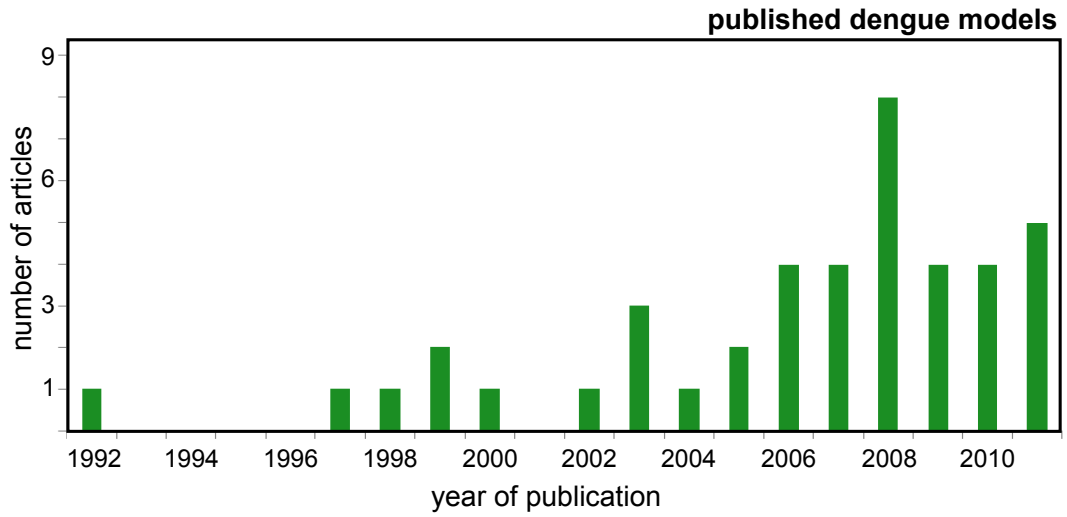


Figure 1.9: **Number of Published Dengue Models (1992-2011)** - The number of published models on dengue's population biology has increased in the last decade. This figure is a reproduction of material published in a review on dengue modelling by Andraud *et al.* [Andraud et al., 2012].

of environmental, human and animal experimentation. Driven by a raising trend in awareness to dengue, the number of mathematical models on its population biology has steadily increased in recent years (figure 1.9 on page 38) [for reviews, see Andraud et al., 2012; Johansson et al., 2011]. The most commonly used approach is that of dynamical systems defined with deterministic ordinary-differential equations (ODE). This section contextualises the results of a small subset of such studies that are of relevant nature to the research presented in this thesis.

Immunity and Population Dynamics

As with many other antigenically diverse pathogens such as the influenza A virus [Ferguson et al., 2003] or the malaria parasite [Gupta and Day, 1994; Recker et al., 2004], dengue's population dynamics present a seemingly competitive behaviour between its

antigenic types (Section 1.2 on page 27). The majority of studies have therefore focused on cross-immunological mechanisms to explain dengue's multi-annual oscillations in serotype prevalence. In fact, a recent review on dengue modelling by Andraud *et al.* has showed that 24 out of the 42 relevant studies have focused on the population consequences of the ADE phenomenon [Andraud et al., 2012].

Even though observations of varying prevalence levels of each serotype in time had been present in the literature since the 1950s, the first study to address a possible mechanism for serotype oscillatory behaviour was based on a mathematical model of transmission. By modelling an endemic population with two circulating serotypes, Ferguson *et al.* showed that varying degrees of ADE in secondary infections could generate oscillatory patterns qualitatively similar to the empirical ones [Ferguson et al., 1999a]. Analysis of the model's dynamics further suggested that ADE could act to favour the coexistence of serotypes. An important assumption was that higher viraemea levels in secondary heterologous infections, due to the ADE phenomenon, can lead to an increase in the probability of human-to-vector transmission per mosquito bite - the enhanced transmissibility hypothesis. Thus, strains circulating at low frequency can be rescued by inducing moderate levels of enhancement in secondary infections. With evidence from a short time series of dengue incidence from Mexico and an hypothesis for a biological mechanism, the authors proposed serotype oscillatory behaviour as a universal feature of dengue's epidemiology. Soon after, Nisalak and colleagues presented the first long time series on serotype dynamics confirming the existence of

such data patterns [Nisalak et al., 2003].

Further expanding Ferguson's *et al.* mathematical framework to allow for the circulation of *N* serotypes, Cummings *et al.* revised the ADE model [Cummings et al., 2005]. The authors showed that the selective advantage of immune enhancement is higher when multiple serotypes co-circulate by supporting persistence during periods of low-frequency circulation, thus reinforcing the original observations. Evolutionary constraints on ADE were also proposed, since greater levels of enhancement can lead to large amplitude oscillations and stochastic extinction. This in turn suggested that ADE may provide a mechanism that favours the co-circulation of antigenically distinct viruses but might also limit the total number and nature / phenotype of such agents in single host populations.

Other studies have analysed variations of the ADE model by assuming different outcomes from immune enhancement. For example, Kawaguichi *et al.* showed that ADE can explain the genetic dissimilarity between DENV serotypes when enhancement leads to a higher mortality rate due to manifestations of severe disease [Kawaguchi et al., 2003]. Varying levels of cross-immunity were used as a proxy for serotype antigenic similarity to explore a spectrum of antigenic distance, with the extremes of high and low protective immunity representing similar or dissimilar viruses, respectively. It was shown that serotypes circulating at low-frequency must override the pressures of both cross-immunity and enhanced mortality in order to emerge. As a consequence,

there was a strong selective drive for greater immunological distance (i.e. low cross-immunity). The results demonstrated that the effects of ADE-induced mortality, in terms of serotype coexistence, were opposite to what had been previously suggested for transmission enhancement. The authors thus demonstrated that high mortality due to ADE may lead to an evolutionarily stable viral population comprising dissimilar variants.

Adams and Boots unified (partially) protective cross-immunity with previously hypothesized ADE assumptions of increased transmissibility, susceptibility and mortality into a single framework [Adams and Boots, 2006]. With this approach, the authors demonstrated that the proposed effects of enhanced mortality (as assumed by Kawaguchi *et al.*) were crucially dependent on other factors, such as co-infection and a very long infectious period. Therefore, the selective pressure towards higher immunological distance derived from ADE-mediated mortality was found to be too small, under more strict dengue-like assumptions. It was also established that higher mortality was the only type of enhancement that could not reproduce the empirical cyclical epidemic behaviour. Importantly, susceptibility and transmissibility enhancement favoured stable coexistence of different serotypes even at high levels of cross-immunity, corroborating the studies by Ferguson *et al.* and Cummings *et al.* [Cummings et al., 2005; Ferguson et al., 1999a].

Further expanding on previous work, a comprehensive comparison of a multitude

of ecological and immunological determinants for dengue's epidemiological patterns was done by Wearing and Rohani [Wearing and Rohani, 2006]. Based on Sabin *et al.* observations, their framework included a short period of serotype transcending immunity after recovery, rather than a life-long, partially cross-protective immunity [Sabin, 1952]. Their results showed that for a fixed amplitude in seasonality, short periods of heterologous immunity were sufficient to reproduce dynamics consistent with data. Inclusion of moderate amounts of enhancement due to ADE were still compatible but not necessary to explain the desired patterns. Using a stochastic formulation and varying host population size, they further demonstrated that moderate levels of enhancement in transmissibility together with a short period of transcending immunity robustly reproduce viral dynamics and persistence patterns observed in Thai province data.

While most theoretical studies concentrate on patterns of competitive exclusion in serotype cyclical replacement, Adams *et al.* have focused on particular phase observations amongst serotypes [Adams et al., 2006]. By correlating epidemiological and phylogenetic signatures between DENV4 and other serotypes, such as oscillatory prevalence and clade replacement events, the authors proposed significant asymmetric, cross-immunological reactions. The effect of such immune interactions was analysed using a two-serotype mathematical model with seasonal forcing and environmental stochasticity. Numerical simulations demonstrated that moderate levels of cross-immunity could explain the observed out-of-phase patterns between DENV1

and DENV4 in prevalence data. In contrast, the in-phase oscillations observed for DENV1-3 required either weak or very strong cross-immunity. As sequential infections are frequently observed for DENV1-3, the latter hypothesis was not further pursued. Comparison between protective and enhancing effects of immunity also revealed a determinant role on the epidemic period, with moderate levels of cross-protection better reproducing the empirically observed periods of 8 to 10 years in serotype prevalence. Finally, the authors suggested the use of large-scale individual-based models, as the ones used in the context of influenza A [Ferguson et al., 2003; Koelle et al., 2006a], to further address the links between the drivers of the population dynamics and phylogenetics of dengue variants.

A commonly criticized feature of frameworks based solely on transmissibility increases due to ADE is their dependence on high enhancement levels to desynchronise serotype prevalence. By assuming that increases in susceptibility and transmissibility might be a concurrent phenomenon in individual hosts, Recker *et al.* showed, under a model system, that the required levels of such enhancements could be much lower and closer to biologically plausible values than the ones assumed in previous studies [Recker et al., 2009]. The resulting dynamics were therefore less prone to high amplitude oscillations, either in overall incidence or sero-prevalence, lowering the risk of viral extinction, an essential consequence of ADE enhancement, as discussed by other authors [Cummings et al., 2005; Ferguson et al., 1999a].

Common to the above studies, is the assumption that complete immunity is acquired after recovery from secondary infection. In Andraud's *et al.* recent review on influential models of dengue, only 2 out of the 42 considered studies had used more than 2 sequential infections in their frameworks [Andraud et al., 2012]. This trend has emerged mainly due to the rarity of third and fourth infections making model parametrization difficult, but also from its simplifying power in mathematical systems. Interestingly, one of the first theoretical approaches was also one of the few considering the possibility of more than two sequential infections [Fischer and Halstead, 1970]. The main hypothesis of the study was whether fitting a model to age-stratified data could address the relative weight of secondary and tertiary infections in the development of severe disease. The results demonstrated that the best fit could only be achieved when assuming a strong correlation to secondary heterologous infection.

Recently the twice-infected protected assumption was challenged by two theoretical studies [Chikaki and Ishikawa, 2009; Wikramaratna et al., 2010]. Wikramaratna *et al.* focused on the qualitative differences of relaxing the assumption of complete immunity after secondary infection. Their results, based on a 4 serotype modelling framework, demonstrated that relaxation of this assumption more easily derived into the epidemiological patterns of dengue. Inclusion of third and fourth cases significantly increased the force of infection even though the reproductive number (see bellow), remained the same. This in turn reconciled realistic sero-conversion age profiles in the four-infected protected simulations with low values for dengue's transmission po-

tential. There are a number of reasons why high values for the basic reproductive ratio of dengue are unlikely. For instance, while other vector-borne pathogens, such as the malaria parasite, have evolved immune evasion mechanisms to prolong their infectious period and overcome transmission bottlenecks, dengue has a relatively short period of infection of just a few days (Section 1.4 on page 30). Furthermore *aegypti's* life-span and long extrinsic incubation period severely constrain the vectors' success in transmitting to a significant number of hosts. Overall, this study makes a strong case that high seroconversion levels to dengue viruses, characteristic of many endemic regions, may not reflect high reproductive ratios.

Endemicity in the Light of Host Demography, Ecology and Control

As described in the previous section, most theoretical work on dengue's epidemiology has focused on cross-immunological and competitive factors between the virus' variants. However, the worldwide increases in clinical dengue are more generally a consequence of large-scale trends, such as vector expansion, demography and climate change [Halstead, 2007, 2008; Kyle and Harris, 2008; Maramorosch et al., 2008]. Local epidemiologies thus emerge from a mixture of regional and global determinants. Notably, only a few studies have focused on such factors to better understand the drivers of dengue's contemporary population dynamics.

An interesting and recent trend in epidemiological data is the apparent increase in mean age of infection in endemic countries of South East Asia [Cummings et al., 2009;

Nisalak et al., 2003]. Based on such observations in Thai data, Nagao and Koelle have proposed that decreases in local transmission, for instance due to control, may act to increase the number of clinical cases [Nagao and Koelle, 2008]. The authors based their hypothesis on the crucial assumption that dengue is in *endemic stability* . This is an epidemiological state characterized by low levels of disease under the premise that the probability of developing symptoms increases with age and that transmission intensities are high such that individuals get infected at young ages [Aguas et al., 2006; Coleman et al., 2001]. Pursuing this conjecture, the authors started by identifying a concurrent raise in DF/DHF and mean age of infection in Thai data, and developed a mathematical model to reproduce the increase in age of infection and clinical cases using non-stationary decreases in transmission potential. The model was based on the assumption that individuals undergoing a period of transcending immunity could sero-convert if challenged with another serotype. This was based on experimental work in monkeys, reporting that sero-conversion is possible in cross-protected individuals and that detectable viraemia is not necessary to sero-convert [Kochel et al., 2005; Kraiselburd et al., 1981]. However, it is so far not clear if this is possible in humans as well. More so, the correlation of symptomatic episodes with increasing age is still unclear due to contrasting observations and the strong correlation with secondary challenge, prone to happen at older ages [Burke et al., 1988; Gibbons et al., 2007; Lin et al., 2010; Thai et al., 2011]. Nonetheless, this study highlighted how moderate reductions in transmission can have counterproductive effects in the context of dengue.

The results of Nagao and Koelle are reminiscent of dengue's epidemiological situation in Singapore. In this island city-state, the pathogen's prevalence was successfully reduced between the 1970s and mid 1990s by implementing rigorous vector-control measures. Although entomological indices have been maintained at low levels since (e.g. $\approx$ 1–2% of house index), major epidemics re-emerged in the late 1990s [Ooi et al., 2006]. Recent modelling work has suggested that successful control resulted in an estimated 10% decrease in Singaporean herd-immunity per decade [Oki and Yamamoto, 2012]. Resurgence may therefore be an expected consequence of a reduced epidemic threshold for dengue viruses.

In a more recent publication, Cummings *et al.* focused on human demographics to explain the increases in age of infection [Cummings et al., 2009]. Under an age-stratified transmission model, it was shown that demographic shifts towards lower birth and death rates could easily accommodate the observed age changes in clinical incidence. This study thus presented a parsimonious mechanism that is independent of reductions in transmission (as proposed by [Nagao and Koelle, 2008]). Importantly, if symptomatic dengue truly correlates with older ages, as populations across the world are undergoing a general ageing trend, these results would predict a continuous rise in symptomatic dengue.

Small islands frequently present intermittent outbreaks due to an increased risk for pathogen extinction. These ecologically isolated populations are therefore useful

models to address determinants of persistence. Likewise, as susceptibility is generally high and accumulates between major epidemics, transmission is less influenced by (unobserved and unmeasurable) herd-effects, simplifying the estimation of key factors such as transmission potential or dispersal rates. In a study by Favier *et al.*, typical characteristics of dengue outbreaks were analysed for the Easter Island, where a small and immunologically naive population is concentrated in a single village [Favier et al., 2005]. It was demonstrated that homogeneous mixing models could not reproduce the epidemic shape and duration in a naive host population. In contrast, using the methodology of individual-based modelling, the authors proposed that contact heterogeneity, either at the individual or population levels, is necessary. These results highlighted the importance of considering natural variation in dengue modelling, specially in the light of the spatio-temporal, heterogeneous nature of mosquito distributions, even within small spatial scales.

In contrast to small populations, hyper-endemic regions present recurrent epidemics with patterns emerging at wider spatial ranges. Using spatial incidence data from Thailand, Cummings *et al.* demonstrated the existence of a spatio-temporal travelling wave of DHF emanating from Bangkok to outer regions with a periodicity of approximately 3 years [Cummings et al., 2004]. According to the authors, such fixed periodicities might reflect important host-pathogen features, such as serotype replacement or human densities, although the influence of climatic fluctuations could not be ruled out. Interestingly, other studies have demonstrated inverse wave-like behaviour

[Thai et al., 2010], from rural to urban centres in Viet Nam.

The contrasting observations on the origin of epidemic waves propose that human density is not the sole driver of such phenomena and urban centres are not necessarily their epicentre. Using an approach based on ODE meta-population modelling of both human and mosquito individuals, Adams and Kapan were able to address the importance of host segregation and movement [Adams and Kapan, 2009]. Within the meta-population setting, human hosts were allowed to disperse amongst subpopulations. Using numerical simulations, it was found that incidence levels were practically independent of the time that human hosts spent in distant subpopulations, based on the assumed frequency-dependent biting rate of the vector (i.e. mosquito individuals bite with a fixed frequency, independently of human availability, both in numbers and time). Together with reports of the patchy nature of vector populations [Higa, 2011; Nagao et al., 2008], this led to the proposal of hubs or reservoirs for transmission. These localities, frequently visited by humans and infested with the vector, would contribute massively to the persistence of the pathogen, independently of the duration of the visits. Importantly, the study also highlighted how human mobility together with heterogeneous mosquito populations might explain the difficulties in disease control despite increased investment in vector control.

A final important feature of the mosquito's ecology is the possibility of viral vertical transmission from female individuals to their offspring (VT). Due to contrasting

reports from field studies and consequent difficulties in model parametrization, VT has generally not been considered in theoretical studies. In the most relevant publication to date, Adams and Boots reviewed a variety of field and experimental VT measures. Using a mathematical model, the authors found that the most commonly reported range for adult-to-egg infection per oviposition ($1-4\%$) does not allow VT to be a key determinant for long term viral persistence [Adams and Boots, 2010]. These results emerged from the vector's rapid life-cycle in which vertically acquired infections are multiplicatively diluted with every generation. Consequently, DENV might be rapidly lost in isolated populations in the absence of amplification in the human hosts. Alternative mechanisms such as silent transmission or reintroduction from external reservoirs were suggested to be more relevant determinants of viral persistence. In an alternative discrete and stochastic scenario, persistence of DENV might occur in a disjoint set of subpopulations and asynchrony in such a system could prolong the positive advantage of VT. This, however, remains do be demonstrated.

The Basic Reproductive Number

A key concern for the area of theoretical epidemiology is whether a single infectious host may start a population-wide epidemic. In theory, an outbreak is likely to happen if the average number of infections (R_n) produced by an infectious individual is bigger than one; since, otherwise, transmission fades away. If the population in consideration is completely susceptible, R_n is referred to as the basic reproductive number or R_0. In contrast, in the presence of herd-immunity, it is referred to R_t, the effective

reproductive number. Importantly, in the case of mosquito-born pathogens, the reproductive number covers a complete vector-host transmission cycle and will therefore take into account several mosquito and human factors. These include incubation periods, transmission probabilities from one host-species to the other, infectious periods, vector feeding behaviour, life-expectancies, etc.

Differences in R_n estimations for a given pathogen may be attributable to a variety of multi-scale factors, from real geographic and temporal variation in transmission, to differences in methodology or assumptions in modelling frameworks, not withstanding biological asymmetries in host and pathogen genetic backgrounds. Another relevant issue is the frequent discrepancy between the real and the detected number of infections, since surveillance systems, specially those based in passive strategies, do not capture the full extend of outbreaks.

For pathogens, such as the dengue virus, that frequently cause asymptomatic or mild disease, underestimation of incidence rates severely hampers the evaluation of the transmission potential and R_n from incidence data. Moreover, field and experimental quantification of biological parameters, specially those related to the vector, is generally affected by high levels of uncertainty. This has resulted in a wide range of R_0 estimations for dengue, from values close to 1 to bigger than 20, depending on location and the methods used (table 1.1 on page 52). One way of overcoming these problems is by analysing seroconversion data. The latter allows for an approximate

Table 1.1: R_0 **estimates from modelling studies.**

Study	N	Cross-protection	Enhancement	Median R_0	Method
[1]	1	No	No	1.3 (0, 2.4)	final epidemic size
[2]	1	No	No	1.9	hypothetical
[3]	1	No	No	2.0 (1.6, 2.5)	epidemic growth curve
[4]	4	Susceptibility	Susceptibility	4.8 (1.4, 8.5)	age stratified seroprevalence
[5]	1	No	No	6.3 (3.6,12.9)	epidemic growth curve
[6]	1	No	No	4.8 (2.7, 11.6)	epidemic growth curve
[7]	1	No	No	3.0 (2.0, 103)	epidemic growth curve
[8]	1	No	No	2.0 (0.5, 3.3)	epidemic growth curve
[9]	1	No	No	3.69 (2.26, 11.39)	epidemic growth curve
[10]	4	Susceptibility	Both*	(4, 12)	age distribution of cases
[11]	1	No	No	5.17 (1.56, 22.89)	fitting SIR model
[12]	4	Susceptibility	Susceptibility	(5.2, 6.7)	age distribution of cases

N - number of serotypes; * - Susceptibility and Infectiousness.
Table reproduced from Johansson *et al.* [Johansson et al., 2011].
References: [1] Koopman et al. [1991] ; [2] Newton and Reiter [1992]; [3] Marques et al. [1994]; [4] Ferguson et al. [1999a]; [5] Massad et al. [2001]; [6] Massad et al. [2003]; [7] Favier et al. [2006]; [8] Chowell et al. [2007]; [9] Coelho et al. [2008]; [10] Nagao and Koelle [2008]; [11] Degallier et al. [2009]; [12] Cummings et al. [2009]

examination of real case numbers and of the current immunological landscape. By using age-stratified indexation of such seroconversion rates one can estimate the time-dependent and serotype-specific forces of infection and infer possible R_t values. This has been done by Ferguson *et al.* in one of the most solid quantifications of the reproductive numbers so far, which lead to estimations between 4 and 6 [Ferguson et al., 1999b].

”The purpose of models is not to fit data but to sharpen the questions.”

Samual Karlin

Thesis Outline

The ideas explored in this thesis derive from a growing need to better understand the contemporary epidemiology and evolutionary dynamics of the dengue virus. The majority of theoretical work on the epidemic behaviour of dengue has concentrated on viral competition effects, arising from protective and/or enhancing cross-immunological reactions within the human host. Mathematical models have generally been successful in reproducing DENV's population dynamics, which, alike other antigenically diverse pathogens such as the influenza A virus or rotavirus, are characterized by seasonal population fluctuations in the form of epidemic outbreaks in addition to oscillatory prevalence levels of the constituent antigenic variants (serotypes). However, these approaches have also derived into a multitude of slightly different biological assumptions that reproduce the observed patterns equally well. So far, discerning between alternative or potentially complementary mechanisms has been difficult due to our lack of understanding on the processes of immunity to dengue viruses. The failure to ascertain a definite mechanism for dengue's population-wide dynamics, through approaches based on incidence time series alone, highlights the need to shift theoretical research towards multi-scale data integration. The aim of this thesis is to integrate epidemiological, ecological and evolutionary factors into unified mathematical frameworks.

While previous models have only focused on the dynamics of dengue's antigenic

types (serotypes), Chapter 2 introduces a new approach to address the determinants of the population dynamics of newly emerging and potentially advantageous viral lineages (genotypes) within serotypes. Results from numerical simulations show that the success of new genetic variants is mainly determined by stochastic processes, strong clonal interference (i.e. viral competition) and the immunity landscape of the human population at the time of emergence. These findings show that apart from the well documented purifying pressures acting on dengue viruses, other factors severely restrict the emergence of genetic variants and dictate the observed dynamical behaviour of genotypes. Importantly, the results suggest that the effects of demographic stochasticity could extend to the population dynamics of dengue's antigenic types (serotypes), which has not been considered in previous modelling approaches. To address this, an agent-based framework, allowing variation to be considered at multiple biological scales, is developed and analysed throughout Chapters 3 to 5 in the light of the eco-epidemiological context of dengue.

In Chapter 3, a discrete, agent-based single-strain mathematical model allowing for host-population structure is introduced. Before expanding its epidemiological framework to a dengue-like system in Chapter 4, a simple analysis is performed on model behaviour. It is demonstrated that demographic stochasticity is a parsimonious driver of persistent, oscillatory epidemic behaviour for single strain pathogens. In addition, when considering a spatial dimension, stochasticity is shown to play a major role in the emergence of spatially heterogeneous immunity landscapes. These are crucial but of-

ten overlooked factors of multi-strain systems, specially in those whose empirical observations are believed to be driven by viral competition due to cross-immunological mechanisms, as their strength is dependent on the host's histories of infection.

In Chapter 4 the agent-based model is further expanded to include four co-circulating antigenic types. Contrary to expectations from ODE-based multi-serotype models, in which the dynamics inevitably converge towards a stable equilibrium in the absence of explicit competition, this model exhibits persistent oscillatory behaviour in serotype prevalence. These observations can be explained by stochastic differences in each serotype's effective transmission, which are amplified in time and generate highly asymmetric immunity profiles within the population, thus maintaining each serotype in an independent transient state driven by susceptibility. In the specific context of dengue, it is demonstrated that the inclusion of a mosquito vector and host-population structure under dengue-like eco-epidemiological parameter values is compatible with the emergent oscillatory behaviour. Apart from qualitatively and quantitatively reproducing general epidemiological characteristics of the dynamical behaviour of dengue, the inclusion of a spatial dimension advances on model validation by reproducing well documented dengue spatial features, including wave-like epidemics, heterogeneous sero-conversion landscapes and spatial epidemic clustering. Importantly, when compared to observations from ODE systems, the effects of a short period of transcending immunity and infection enhancement due to ADE were weakened within the agent-based framework, due to random effects and population structure.

Aiming at developing a thorough description of the expected phylodynamic behaviour of dengue serotypes under a model system, the framework of Chapter 4 is expanded in Chapter 5 to include neutral molecular evolution of viral sequences. This multi-scale approach is used to analyse the ubiquitous phylodynamic patterns emerging from the interaction of key demographic, epidemiological and evolutionary processes, by comparing *in silico* summary measures of population genetics and phylogenetics to *ex novo* empirical estimations from serotype genetic data. Adding to the results of earlier chapters, it is demonstrated that factors mostly of stochastic and non-selective nature are sufficient and parsimonious determinants of frequently observed multi-scale features of dengue's population biology. These include signatures of purifying selection, temporal patterns of genetic diversity and typical phylogenetic shape. Importantly, while the results of Chapter 4 suggest that cross-immunological mechanisms are not necessary to reproduce the epidemiological behaviour of dengue in the presence of demographic stochasticity, their consideration when modelling viral evolution lead to significant deviations from empirical expectations on phylogenetic shape.

Together, the results present in Chapter 2, 4 and 5 do not endeavour in questioning the pathological or clinical significance of cross-immunological interactions between dengue viruses. However, this material strongly suggests that such mechanisms are unlikely to be the main drivers of key universal signatures in dengue's population biology, and reiterates host-related demographic and ecological factors instead.

Reproduced Data and Figures in This Thesis

- The consensus map on DENV presence (figure 1.1 on page 15) and outbreak occurrences (figure 1.2 on page 16) were reproduced from material published in the Public Library of Science (PLoS) to which the Creative Commons Attribution License (CCAL) applies.

- The general phylogenetics of DENV serotypes (figure 1.3 on page 19) and DENV2's genotypes (figure 1.4 on page 20) were reproduced from Weaver and Vasilakis with permission from the authors [Weaver and Vasilakis, 2009].

- The immune response to DENV during primary infection is illustrated in figure 1.8 (on page 32). It is a reproduction of published material by Scott Halstead, with permission from the author [Halstead, 2007].

- Data on yearly patterns of dengue disease hospitalization and relative serotype abundance in Southern Viet Nam were kindly provided by the dengue surveillance programme at the Institute Pasteur, Ho Chi Minh City - figures 1.7 A on page 28, and 2.1 on page 63.

- Data from Thailand, was reproduced from Aguiar *et al.* with permission from the authors (figures 1.7 B on page 28) [Aguiar et al., 2011; dos Santos, 2012].

- Spatial serotype data was provided by C. Simmons from the Oxford University Clinical Research Unit, Ho Chi Minh City (HCMC), Viet Nam, representing 290 dengue cases in children presenting to the outpatients department of three

large hospitals in HCMC during the season 2010/2011 (figures 4.4 D on page 127). Spatial information was based on residency of the patient.

- The illustrations on counts of DENV2 genotypes, during a replacement event in Southern Viet Nam, were reproduced with authorization of the authors (figure 2.1 on page 63) [Hang et al., 2010].

- Data for serotype spatial distributions in Thailand, during the 2005/2006 season, was obtained from Fried *et al.* and Veeraseatakul *et al.* and was further reproduced in figure 4.4 C (on page 127) [Fried et al., 2010; Veeraseatakul et al., 2007].

- A range of estimates on R_0 from relevant theoretical studies is presented in table 1.1 on page 52. Values and references were taken from a review by Johansson *et al.* [Johansson et al., 2011].

- The number of published mathematical models on dengue's population dynamics, illustrated in figure 1.9 (on page 38), was taken from Andraud's *et al.* review [Andraud et al., 2012] published under the Creative Commons Attribution License (CCAL).

CHAPTER 2

The Determinants of the Invasion Dynamics of Novel Dengue Genotypes

Abstract

In a recent report by Hang *et al.*, a major replacement event was described for South East Asia where the Asian-1 genotype of serotype 2 replaced the resident Asian/American type. When comparing viraemia in human patients, the authors found significantly higher levels amongst the ones infected with the Asian-1 genotype. It was proposed that this altered phenotype could justify the replacement event by enhancing the per-bite probability of transmission from humans to mosquitoes per bite. In this Chapter, a model system is developed to investigate the general determinants of the invasion dynamics of a novel, advantageous DENV genotype in an endemic setting. The results demonstrate that replacement events can be explained by competition between genotypes of relatively small fitness differences which, although sufficient for displacement, do not interfere with the overall serotype dynamics. Furthermore, the invasion success and total time required for fixation are strongly influenced by viral competition and the epidemiological landscape at time of introduction. As a consequence, novel genotypes can face a high risk of stochastic extinction despite their fitness advantage. The rarity of markers for positive selection in DENV's molecular evolution has often been explained by strong purifying selection whereby the constraints imposed by a two-host life cycle are expected to result in a high rate of deleterious mutations. The

results presented here demonstrate that even highly beneficial mutants are under severe threat of extinction, suggesting that stochastic effects and genetic drift beyond seasonal bottlenecks are equally important in shaping dengue's viral evolution.

2.1 Introduction

Despite a general bias in the literature towards studies based on single-gene approaches, spatio-temporal patterns of genotype replacement in endemic regions have been widely recovered from data [Bennett et al., 2003; Holmes and Twiddy, 2003; Myat Thu et al., 2005; Weaver and Vasilakis, 2009; Zhang et al., 2005]. With the extrinsic pressures on DENV, such as seasonal or human-forced reductions in vector population size or abundance of susceptible hosts, it has been proposed that genetic drift plays a major role in the observed phylodynamics [Myat Thu et al., 2005; Wittke, 2002].

Following earlier reports on inter-serotypic difference in virulence (e.g. [Gubler et al., 1978]), one of the first convincing evidence for a genetic determinant in disease outcome came from epidemiological studies suggesting that the DENV2 Asian genotype was associated with higher frequencies in DHF compared to the American type [Rico-Hesse et al., 1997]. *In vitro* studies have since shown that the Asian genotype's replication rate in human monocyte-derived macrophages and dendritic cells is higher [Cologna et al., 2005]. There is also evidence that the vector is more susceptible to this genotype [Armstrong and Rico-Hesse, 2001] and may also present a shorter EIP [An-

derson and Rico-Hesse, 2006]. These reports provided a rational explanation for the replacement patterns observed in the Americas, where displacement of the American genotype (by the Asian) has taken place in several countries in recent years [Bennett et al., 2006; Cologna et al., 2005; Rico-Hesse et al., 1997].

A similar replacement event has occurred in South East Asia, with Asian-1 having displaced the resident Asian/American viruses within Viet Nam, Cambodia and Thailand [Hang et al., 2010]. In the report by Hang *et al.*, some intriguing aspects about the invasion dynamics of Asian-1 were demonstrated. A phylogenetic analysis suggested that the genotype was introduced into the population years before it had been detected, and once discovered, it reached fixation within a relatively short period of time (figure 2.1 B on page 63). The rate at which the replacement took place suggested a significant reproductive advantage. While this could have predicted signatures in total incidence dynamics, there was no discernible change in the period before or after the fixation event (figure 2.1 A on page 63). In search for fitness asymmetries between the genotypes, the authors quantified both the infectivity level to mosquitoes and degree of viraemia in human patients. While the first provided weak support, viraemia was significantly higher for the invading type, which sustained an hypothesis of higher probability of transmission from humans to mosquitoes.

Even if fitness advantages in a specific viral trait played a decisive role in this event, the emergence of advantageous genotypes is also likely to be influenced by the epidemi-

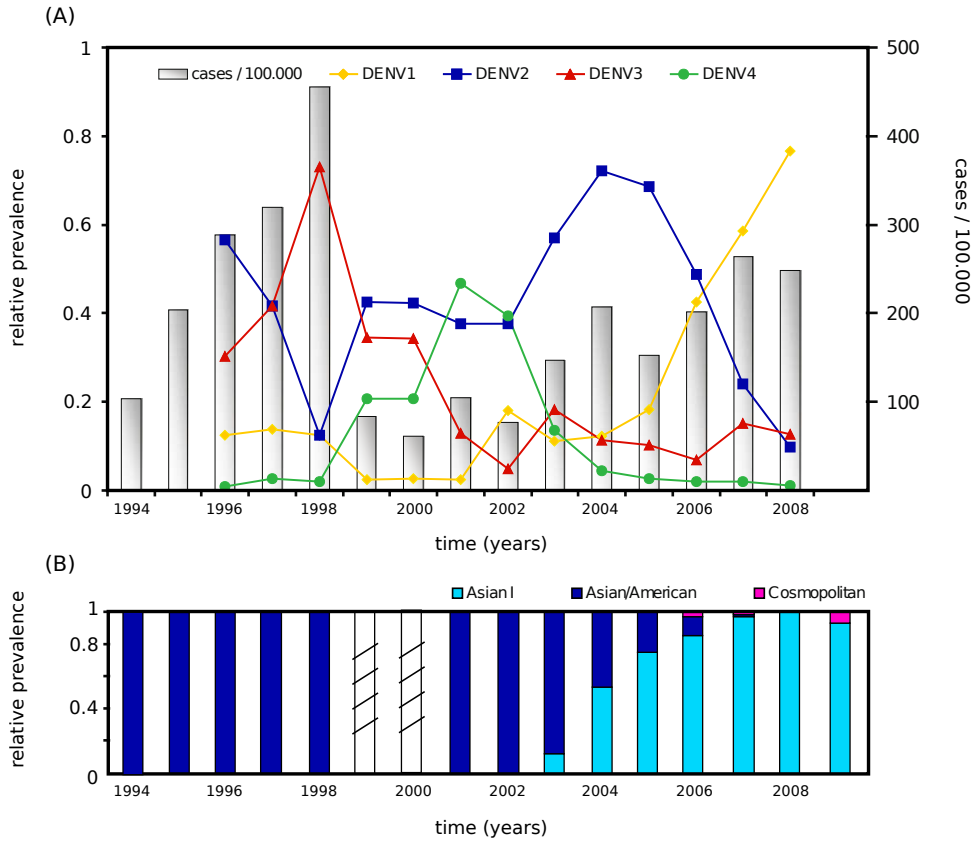


Figure 2.1: **Epidemiology and Genotype Replacement in Viet Nam - (A)** Epidemiological patterns from total number of hospitalised cases in Southern Viet Nam. **(B)** Between 2002 and 2008, the Asian-1 genotype of serotype 2 (cyan) replaced the resident Asian/American type (blue) (data from 1999-2000 is missing). No major epidemiological signatures were observed during the replacement. See Section 1.5 on page 58 for details on data.

ological landscape in which they arise. Here, using a model system, the combined impact of immunity, viral fitness and transmission on the invasion and replacement patterns of novel, advantageous DENV genotypes, is investigated.

2.2 Materials and Methods

2.2.1 Description of the Model

The model is an extension of the mathematical framework analysed by Recker *et al.*, here expanded to include a period of transcending immunity and a vector component with seasonal forcing in mosquito biting [Recker et al., 2009]. In summary, the effect of maternal antibodies is disregarded and human individuals are assumed to be born susceptible to all viruses. After recovery from primary infection the host retains life-long immunity to the infecting serotype and short-term transcending protection against all other viruses. As temporary immunity wanes, individuals become susceptible to secondary, heterologous infection. Due to the rarity of reported third and fourth infections, it is assumed that recovery from secondary infection fully protects against further challenges [Gibbons et al., 2007; Halstead, 2008].

To investigate the invasion patterns of an emerging variant, two genotypes of serotype 2 are considered, an invading (DENV2') and a resident (DENV2). It is assumed that there are no antigenic differences between the competing genotypes, in accordance with a lack of evidence for recurrent infection with variants of the same serotype.

The system is given by the following set of differential equations describing the rate of change in susceptible, infected with, temporarily immune or recovered from, dengue

viruses i (i =DENV1, DENV2, DENV2', DENV3 or DENV4):

$$\frac{dS^h}{dt} = \mu^h N^h - \left(\sum_i \lambda_i^v + \mu^h\right) S^h \quad (2.1)$$

$$\frac{dI_i^h}{dt} = \lambda_i^v S^h - (\sigma_i^h + \mu^h) I_i^h \quad (2.2)$$

$$\frac{dX_i^h}{dt} = \sigma_i^h I_i^h - (\alpha^h + \mu^h) X_i^h \quad (2.3)$$

$$\frac{dR_i^h}{dt} = \alpha^h X_i^h - \left(\sum_{j \neq i} \gamma^h \lambda_j^v + \mu^h\right) R_i^h \quad (2.4)$$

$$\frac{dI_{ji}^h}{dt} = \gamma^h \lambda_i^v R_j^h - (\sigma_i^h + \mu^h) I_{ji}^h \quad (2.5)$$

$$\frac{dR^h}{dt} = \sum_{j \neq i} \sigma_i^h I_{ji}^h - \mu^h R^h \quad (2.6)$$

The force of infection of serotype i affecting the human population, λ_i^v, is given as

$$\lambda_i^v = \eta \beta_i^{v \to h} \frac{I_i^v}{N^h}, \quad (2.7)$$

with η representing the mosquito biting rate and $\beta_i^{v \to h}$ the vector-to-human transmission probability. $1/\sigma_i^h$ and $1/\alpha^h$ are the respective durations of infection and cross-immunity. Given dengue's short period of infection, the possibility of co-infections by two or more serotypes is not considered. The human population size ($N^h = S^h + \sum_i (I_i^h + X_i^h + R_i^h + \sum_j I_{ij}^h) + R^h$) is assumed constant and infection has a negligible effect on the average death rate, μ^h.

To account for seasonal variation, a periodically forced biting rate is set to:

$$\eta = \eta_0 \left(1 + \epsilon \sin{(\pi t)}^k\right), \tag{2.8}$$

where k is a positive integer influencing the length of the transmission season, such that values >> 1 result in shorter and more pronounced seasons.

The dynamics of the mosquito population are given as follows:

$$\frac{dS^v}{dt} = \mu^v N^v - \left(\sum_i \lambda_i^h + \mu^v\right) S^v \tag{2.9}$$

$$\frac{dI_i^v}{dt} = \lambda_i^h S^v - \mu^v I_i^v \tag{2.10}$$

with the force of infection from humans to mosquitoes given as:

$$\lambda_i^h = \eta \frac{\beta_i^{h \to v}}{N^h} \left(I_i^h + \sum_i \phi I_{ji}^h\right) \tag{2.11}$$

In accordance with Recker *et al.*, antibody-dependent enhancement is assumed to act by increasing both susceptibility to, and transmissibility of secondary, heterologous infection by factors γ and ϕ, respectively.

Four different fitness traits are considered and varied independently: (i) transmissibility from human to mosquito, e.g. through increased viral load, $\beta_{2'}^{h \to v}$, (ii) longer

life-expectancy of mosquitoes infected with DENV2' to emulate a shorter EIP, $\mu^v_{2'}$, (iii) longer infectious period in humans, $1/\sigma_{2'}$, and (iv) an increased level of enhancement of secondary infections, ϕ'_2. These are formulated around a degree of fitness advantage, ρ_i, given by:

$$\beta^{h\to v}_{2'} = \beta^{h\to v}_2(1+\rho_\beta) \tag{2.12}$$

$$\mu^v_{2'} = \mu^v_{2'}/(1+\rho_\mu) \tag{2.13}$$

$$\sigma_{2'} = \sigma_2/(1+\rho_\sigma) \tag{2.14}$$

$$\phi_{2'} = \phi_2(1+\rho_\phi) \tag{2.15}$$

In line with the suggestion by Hang *et al.*, most of the analysis is concentrated on the fitness advantage due to increased viral load and thus transmissibility from the infected human individual to the mosquito vector, $\beta^{h\to v}_{2'}$ [Hang et al., 2010].

Parameter values are described in Table 2.1 on page 87.

2.2.2 Stochastic Simulations

To address certain aspects of the invasion process of a more probabilistic nature, such as invasion success, the above model was further implemented in a stochastic framework using a Gillespie algorithm [Gillespie, 1977]. Stochastic simulations are initialized with equilibrium population status derived from the deterministic framework with parameter values the same as given in Table 2.1 on page 87.

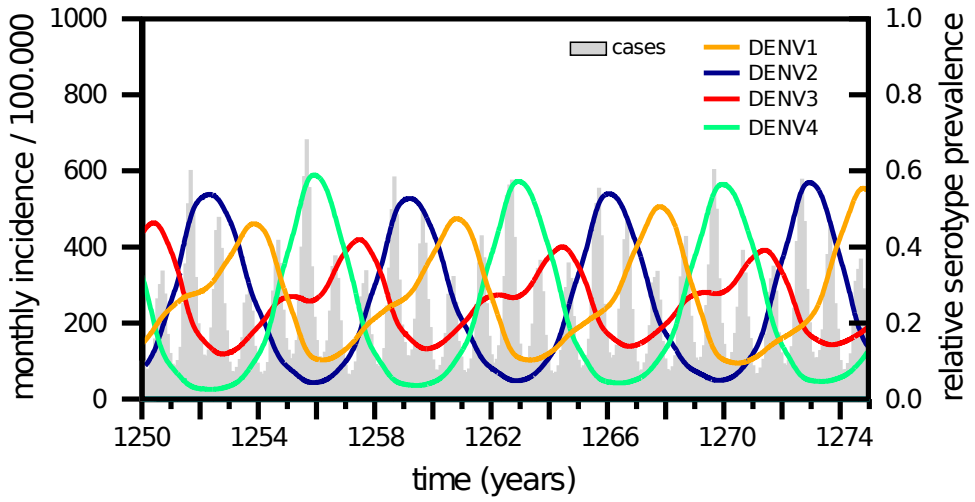


Figure 2.2: **General Model Behaviour** - Under parameter values given in Table 2.1 (on page 87) the model reproduces the typical epidemiological dynamics of dengue, presenting cyclical behaviour in serotype prevalence (coloured lines) and semi-regular epidemic outbreaks (total incidence per month, grey).

2.3 Results

2.3.1 General Model Behaviour

The general dynamics generated by the model under parameter values given in Table 2.1 (on page 87), prior to the introduction of a novel DENV2 genotype, are characterised by semi-regular epidemic outbreaks and asynchronous cyclical behaviour in serotype prevalence (figure 2.2 on page 68). In accordance with previous studies [Ferguson et al., 1999a; Recker et al., 2009; Wearing and Rohani, 2006], a wide range of incidence and serotype dynamics with different periodicities can be reproduced under changes to key parameters values, especially those relating to the degree of enhancement of secondary infection or the period of temporary cross-immunity (not shown

here, see [Lourenço and Recker, 2010]). For the remainder of this chapter, however, parameter values are kept constant to allow better comparisons between invasion patterns and their viral and epidemiological determinants.

2.3.2 Genotype Invasion and Replacement Dynamics

Figure 2.3 (on page 70) shows the result of an invasion scenario where the invading genotype (DENV2') has a small fitness advantage over the resident type (DENV2). In this case, higher fitness is realised through enhanced transmissibility (4.5%, $\rho_\beta = 0.045$) from infected human individuals to the mosquito vectors, i.e. $\beta_{2'}^{h\to v} > \beta_2^{h\to v}$.

In agreement with empirical observations, important features of the invasion dynamics are highlighted in figure 2.3 A (on page 70). Despite the eventual fast rate at which the advantageous genotype replaces the resident type, there is a significant lag between the point of introduction and the time when DENV2' would reach a detectable level of prevalence within the population - here referred to as *detection threshold*. Furthermore, in spite of an expected temporary rise in dengue incidence, compared to the situation without invasion, the overall dynamics in both disease incidence and serotype prevalence remain largely invariant (figure 2.3 B on page 70). This suggests that both the time lag between introduction and detection, as well as the rapid exclusion of the resident genotype, as reported by Hang *et al.*, can be reproduced by a relatively small fitness advantage of the invading type [Hang et al., 2010].

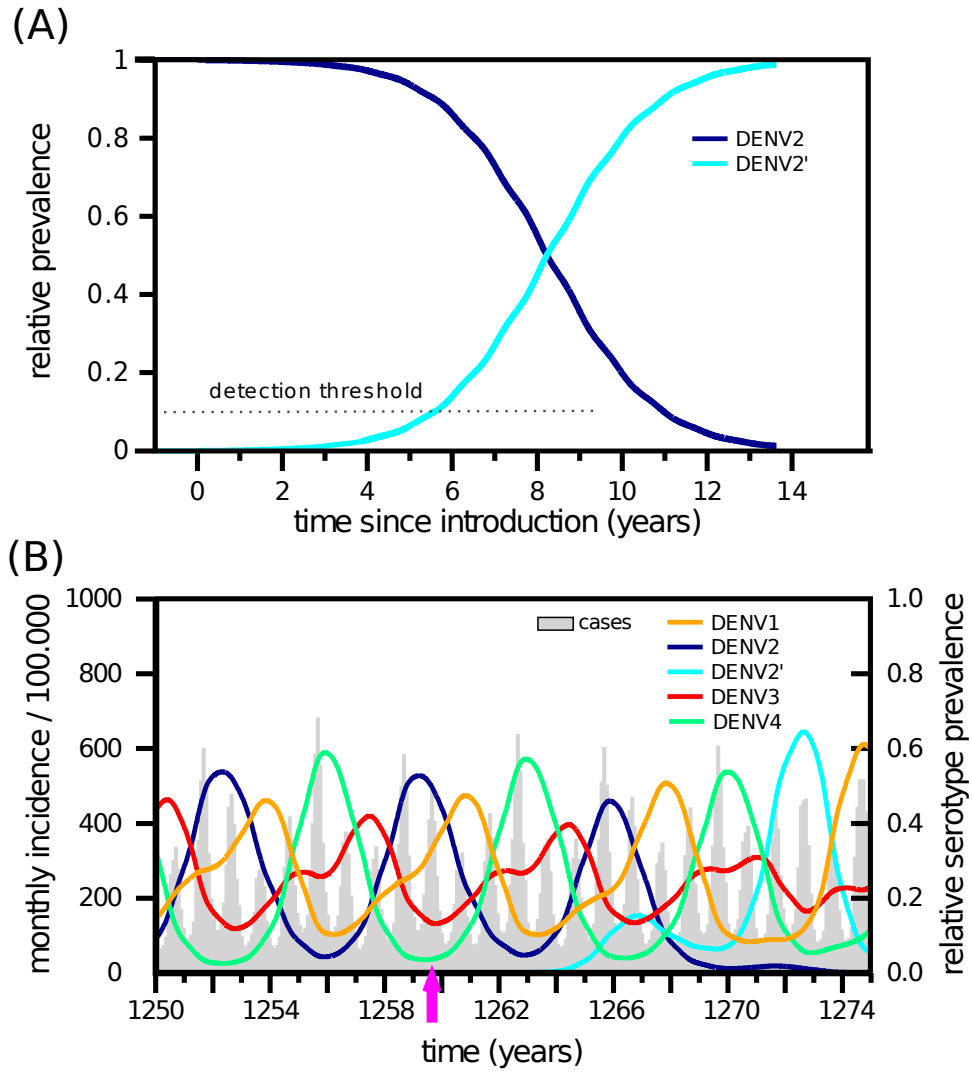


Figure 2.3: **Dynamics of an Invading Genotype - (A)** Frequency of DENV2' relative to DENV2 highlights two phases of the invasion process: a period of very low frequency and a subsequent shift in dominance and competitive exclusion. **(B)** The cyclical serotype behaviour remains invariant to the introduction of DENV2' (cyan line), which enters the population at time $t = 1259.5$ (pink arrow) and drives the resident type, DENV2 (blue line), to extinction after ≈ 13 years. In accordance with empirical observations (figures 2.1 A,B on page 63), no major changes in the epidemiology are observed. The fitness advantage is due to increased human-to-vector transmission rate ($\rho_\beta = 0.045$). Other parameters as in Table 2.1 on page 87.

The same qualitative behaviour can be found when considering other viral traits, that might confer a fitness advantage. That is, shortening the EIP (ρ_μ), increasing the duration of infection (ρ_σ) or the level of enhancement of secondary infection (ρ_ϕ),

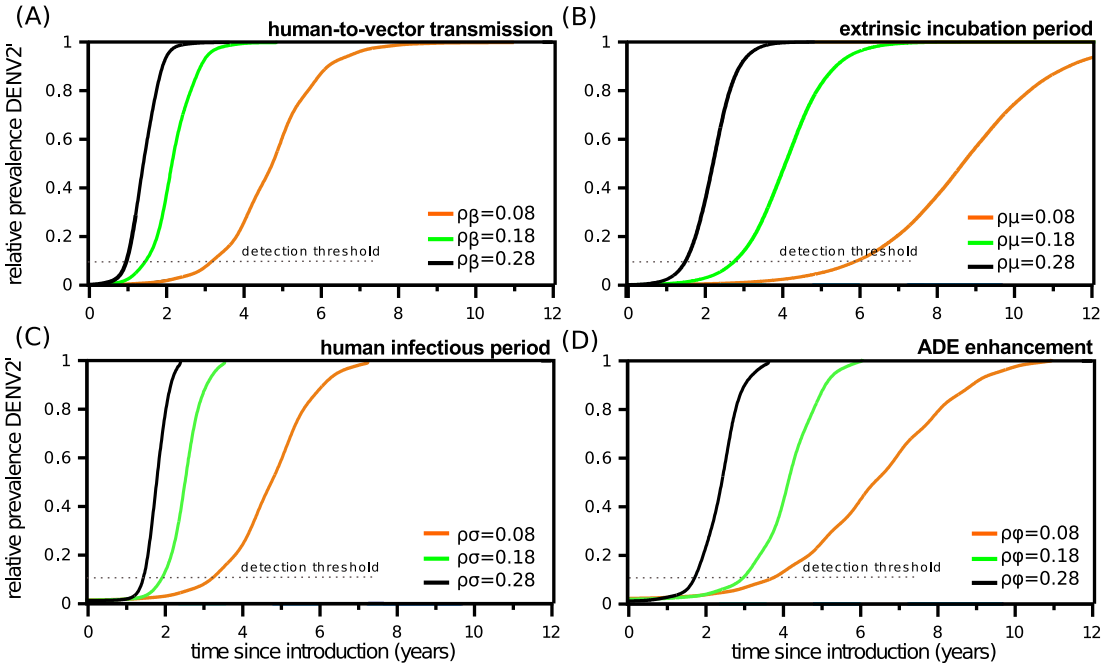


Figure 2.4: **Dynamics of an Invading Genotype Given Differences in Viral Traits** - Frequency of DENV2' relative to DENV2 during replacement. In general, the invasion dynamics are similar when considering different viral traits expressing higher fitness, in this case **(A)** human-to-vector transmissibility ρ_β, **(B)** EIP ρ_μ, **(C)** duration of infection ρ_σ, and **(D)** level of enhancement of secondary infection ρ_ϕ. The system was more responsive to the viral traits ρ_β and ρ_σ.

have similar effects as increasing the transmission probability from infected humans to mosquitoes (ρ_β) (figure 2.4 on page 71). Notably, the human-to-vector transmission (ρ_β) and human infections period (ρ_σ) viral traits are the most responsive, leading to fast replacement under similar degrees of fitness advantage (figure 2.4 A,C on page 71). It is also observed that the synergistic effect of combining different viral traits can be additive (figure 2.5 on page 72). Hence, the same invasion dynamics can be obtained with combinations of very small enhancements in a multitude of traits.

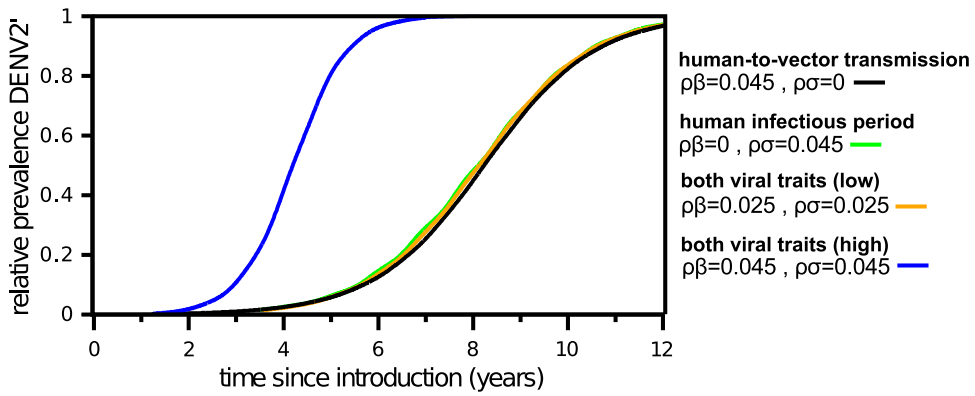


Figure 2.5: **The Synergistic Effect of Viral Fitness Assuming Changes in the Human Infectious Period and Human-to-Vector Transmission** - The graph demonstrates the increased rate in competitive exclusion of the resident genotype DENV2 for varying levels of viral fitness of DENV2' expressed as a longer human infectious period (ρ_σ) and increased human-to-vector transmission (ρ_β). Equal fitness differences, either expressed as longer ρ_σ or increased ρ_β lead to similar emergence and fixation times. The effect of ρ_σ and ρ_β on the invasion dynamics is additive. Other parameter values as in table 2.1 on page 87.

2.3.3 The Effect of Viral Fitness and Time of Introduction

As shown in figure 2.3 (on page 70), a small increase in transmissibility from human to mosquito seems sufficient for a novel genotype to displace a resident type within a short period of time. The actual rate of competitive exclusion and overall time from introduction to fixation, however, is likely to depend on multiple factors such as fitness level, transmission settings, viral competition and the immune profile of the human population. Increasing viral fitness, for instance, accelerates the rate at which the invading genotype drives the resident type to extinction, resulting in a shorter period between introduction and fixation (figure 2.4 A on page 71). Varying the fitness advantage from 8% to 28% reduces the time to fixation from $\approx$ 8 years down to $\approx$ 2 years, for example. However, this increase in viral fitness has a major effect on dengue incidence patterns and the dynamics of the other serotypes (figure

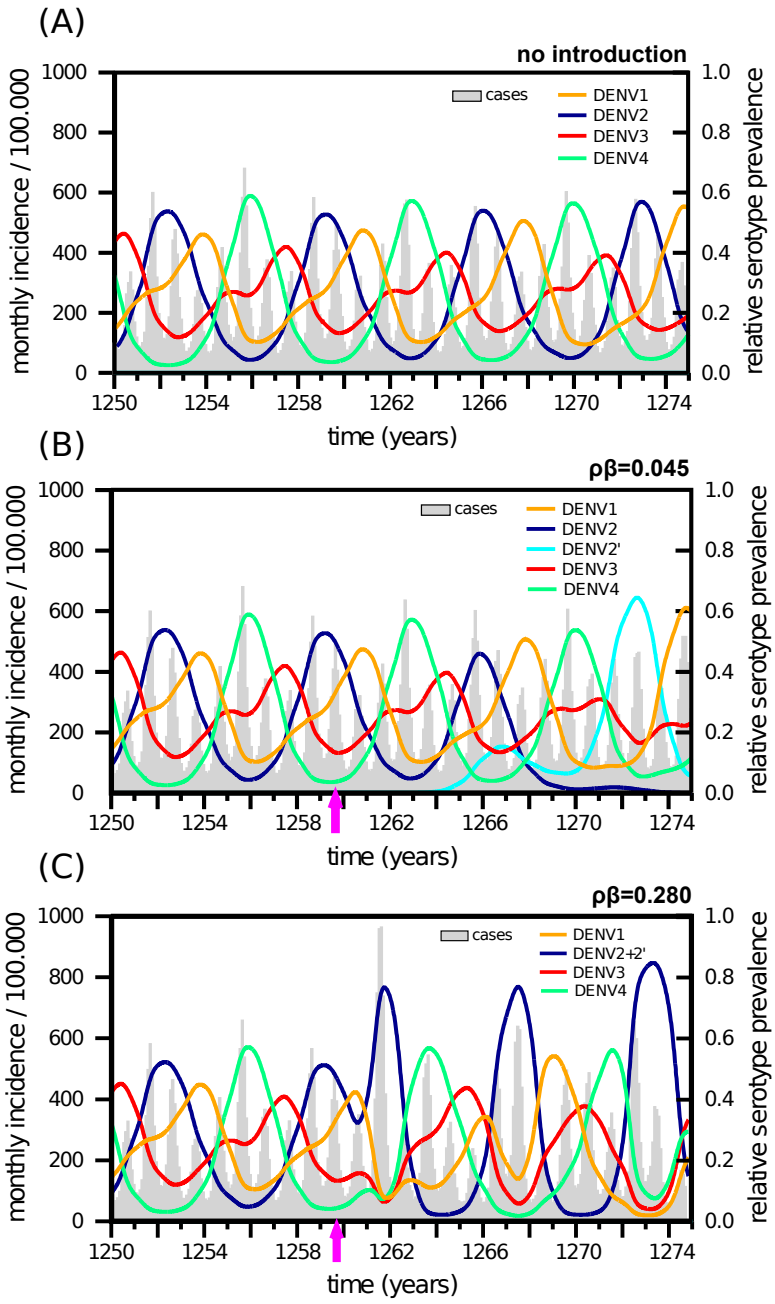


Figure 2.6: **The Effect of Viral Fitness on Fixation Time and Epidemiology** - **(A)** Reproduction of figure 2.2 (on page 68) for direct comparison. **(B)** Reproduction of figure 2.3 (on page 70) for direct comparison. It presents an introduction at the same time point with $\rho_\beta = 0.045$. No epidemiological signatures are observed. **(C)** Higher fitness advantages, here $\rho_\beta = 0.28$, can have a significant effect on both incidence and serotype dynamics - invasion develops into a clear outbreak followed by a severe trough in the serotype's prevalence (for visualization purposes, the frequency of both DENV2 and DENV2' are aggregated into the relative prevalence of serotype 2, dark blue line). Other parameter values as in Table 2.1 on page 87.

2.6 A on page 73). In this case it leads to a major epidemic outbreak shortly after the introduction, followed by a long period of low prevalence of serotype 2, which could endanger its continuous persistence.

At the same time, the *time point of introduction* (TPI) also influences the time taken for a novel genotype to reach fixation. Figure 2.7 A (on page 75) presents the increase in frequency of DENV2', relative to the resident DENV2, for two different TPIs. While the overall time to fixation is dependent on the TPI, the actual rate of replacement remains largely constant. In other words, the time taken from DENV2' passing a detection threshold to nearly achieving fixation (replacement time) is independent of the TPI and therefore independent of the overall epidemiological status of the host population (figure 2.7 B on page 75). On the other hand, the time lag between introduction and the point when it has spread sufficiently for detection, or *emergence time*, is strongly influenced by the epidemiological status of the population.

2.3.4 The Effect of Competition on Emergence Time and Invasion Success

To further investigate the determinants of fixation time, a number of invasion events are simulated at various TPIs over several years. The resulting time to fixation is recorded for each event with respect to several population profiles: (i) the number of naive individuals and (ii) serotype 2 susceptible individuals, the (iii) overall prevalence

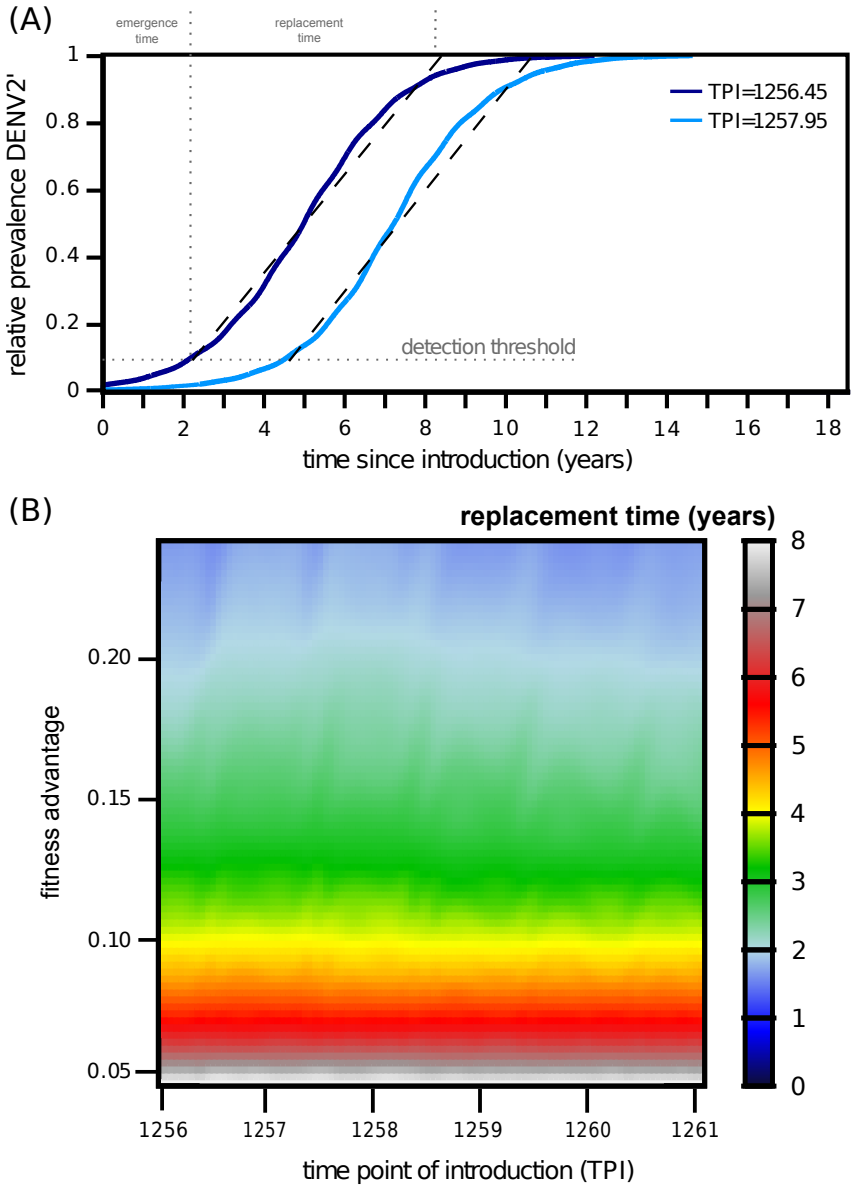


Figure 2.7: **The Effect of Time of Introduction on Fixation Time - (A)** The graph shows the frequency of DENV2', relative to DENV2, for two different time points of introduction (TPI). Despite a discernible difference in the total time for DENV2' to reach fixation, the actual replacement time (i.e. the time taken from passing a detection threshold to nearly reaching fixation; highlighted as parallel dashed lines) remains approximately the same. That is, the differences in fixation times are mostly due to differences in the initial expansion period of the invading genotype before it reaches widespread detection level (here arbitrarily set at 10% relative prevalence). **(B)** The relative fitness advantage of the invading genotype has a significant effect on replacement time as seen by the changing colours in vertical lines. However, it remains mostly invariant to the time TPI (horizontal lines). All parameters as in Table 2.1 (on page 87) and $\rho_\beta = 0.045$ for (A).

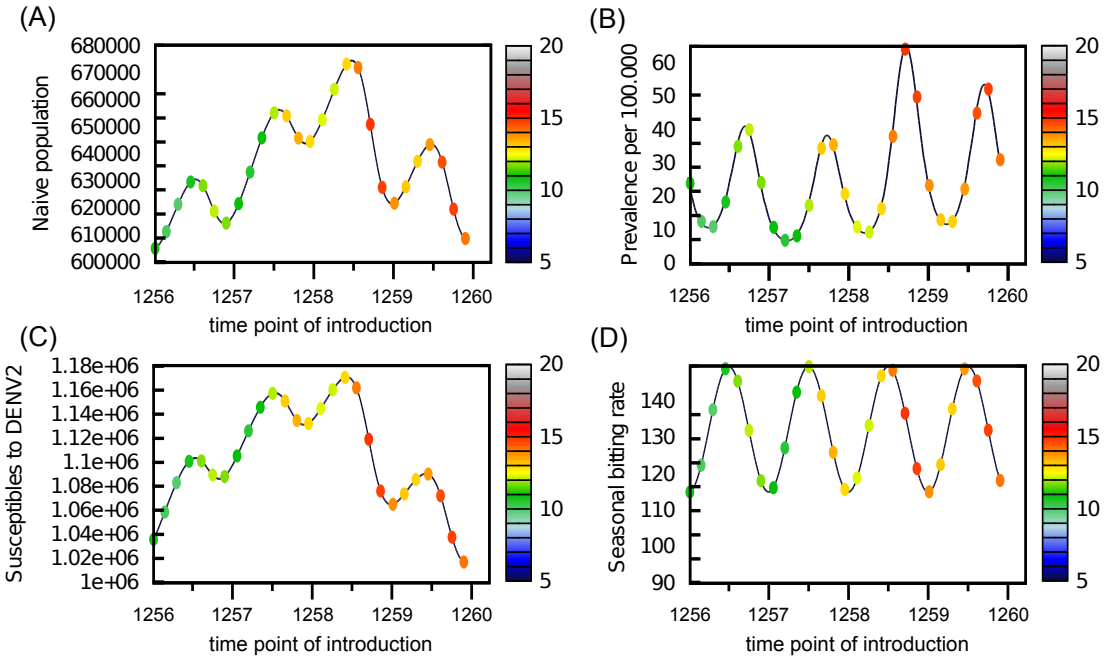


Figure 2.8: **The Effect of Population Status in Fixation Time** - Fixation time in years (coloured solid circles) for different introduction times (TPI). The plots present the evolution of variables that define the host population status at each TPI. **(A)** Number of susceptible individuals. **(B)** Dengue prevalence. **(C)** Serotype 2 susceptibles. **(D)** Seasonality (biting rate, y-axis in years). No correlation is found between natural oscillations in the variables and fixation time. The latter increases within the time range considered. All parameters as in Table 2.1 (on page 87) and $\rho_\beta = 0.045$.

and (iv) seasonality (i.e. mosquito biting frequency) at the time of introduction (figure 2.8 on page 76). During this period of introduction events, a wide range of fixation times are observed, which do not appear related to the seasonal oscillations of the population profiles. Instead, fixation time appears linked to periods of increasing and decreasing susceptibility (figure 2.8 A, C on page 76), which suggests a possible role of viral competition for susceptible hosts.

Stochastic Simulations

The results so far suggest that novel advantageous genotypes can face long periods at very low prevalence before emerging and can also suffer competition for susceptible hosts from the resident viruses. However, these observations are based on a deterministic system. Within a more realistic setting, they signify an enhanced risk of stochastic extinction for the invading virus despite a possible advantageous phenotype. To better address the probability of success of the invading genotype, a stochastic formulation of the model is used in the following sections (see Materials and Methods, page 67).

A number of invasion events are simulated over a period of several years and the *success rate of invasion* (here defined as the successful introduction into a population followed by competitive exclusion of the resident type) is recorded. As demonstrated in figure 2.9 A (on page 79), the invasion success shows an oscillatory behaviour which is negatively correlated to total dengue prevalence at the time of introduction. This suggests that the invasion of a newly advantageous genotype can be hampered by viral competition during epidemics and favoured during off-season periods. Moreover, the maximum success rates are found to be dependent on, and negatively correlated with, serotype 2 prevalence (figure 2.9 B on page 79). Since the replacement time (i.e the time taken from passing a detection threshold to nearly reaching fixation) appears to be independent of the time of introduction (figure 2.7 B on page 75), the relationship between serotype 2 prevalence and the time to emergence is further

explored. Figure 2.9 C (on page 79) clearly illustrates that an invading genotype of DENV2, entering the population during periods of high prevalence of that serotype, faces significantly longer emergence times than if introduced during periods of low prevalence. In general, these observations do not change when considering fitness advantages in other viral traits (figures 2.10, 2.11, 2.12, on pages 80, 81, 82), apart from ADE enhancement which translates in the lowest success rates of invasion, possibly due to its restricted effect to secondary infection. Together these results indicate that the fate of the novel genotype is strongly determined by both inter- and intra-serotype competition at the time of introduction.

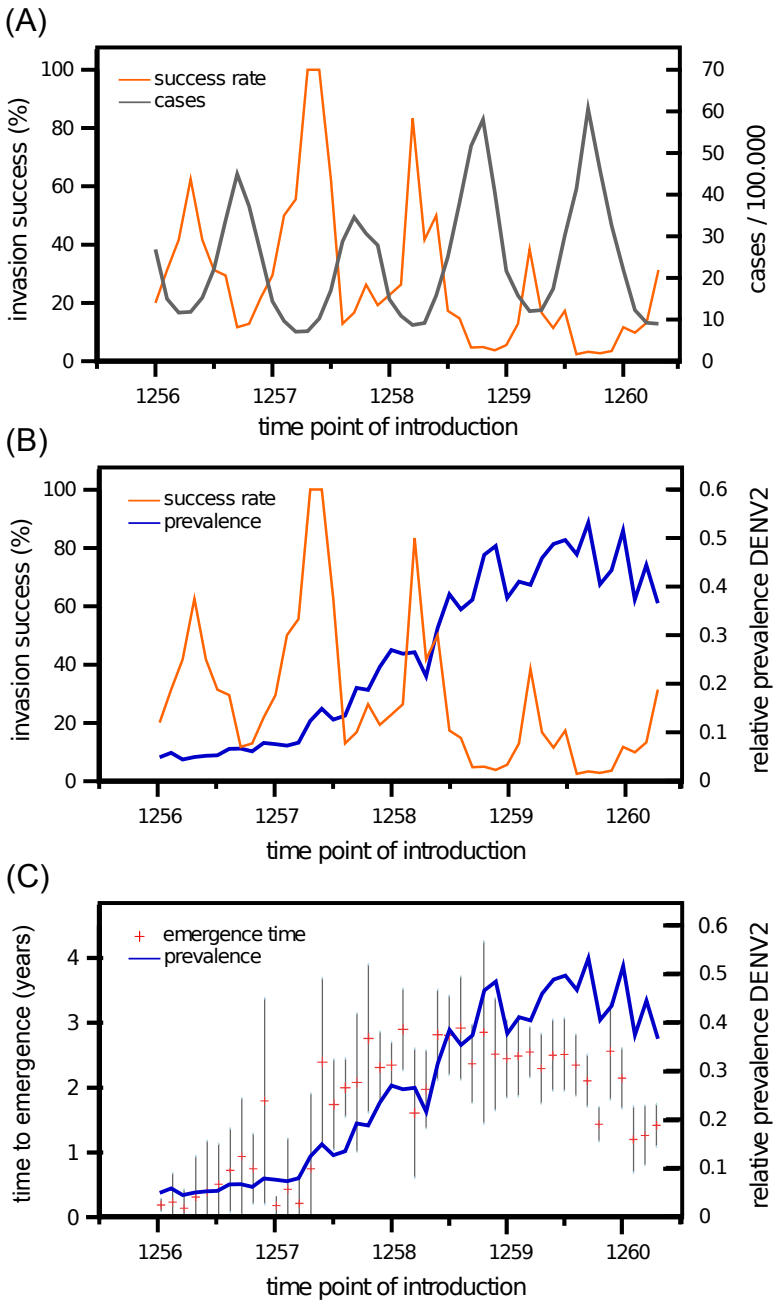


Figure 2.9: **The Effect of Viral Competition on Invasion Success and Emergence Time For Fitness in Transmission - (A)** The invasion success (orange line) of the invading genotype oscillates out of phase with total dengue incidence (grey line) at the time of introduction (TPI). **(B)** The highest rates of successful invasions can be observed for low relative prevalence of serotype 2 (blue line) at the TPI. **(C)** The red crosses show how the average emergence times, i.e. the period between introduction and reaching a 10% detection threshold, of successful invasion events increases with the relative prevalence of serotype 2 at the TPI (blue line). Standard deviations, based on 10 simulated successful invasion events, are shown as grey bars. Parameters as in Table 2.1 (on page 87) and $\rho_\beta = 0.045$.

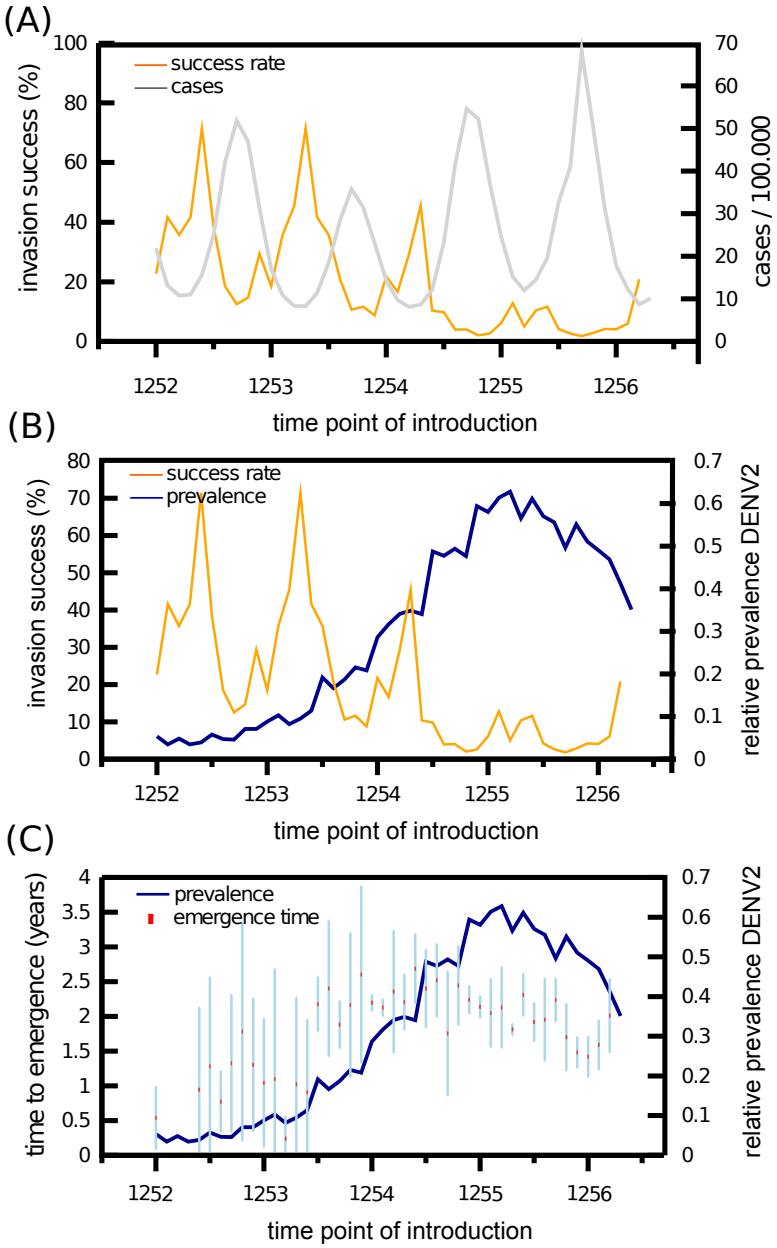


Figure 2.10: **The Effect of Viral Competition on Invasion Success and Emergence Time For Fitness in EIP** - **(A)** The invasion success (orange line) of the invading genotype oscillates out of phase with total dengue incidence (grey line) at the time of introduction (TPI). **(B)** The highest rates of successful invasions can be observed for low relative prevalence of serotype 2 (blue line) at the TPI. **(C)** The red points show how the average emergence times, i.e. the period between introduction and reaching a 10% detection threshold, of successful invasion events increases with the relative prevalence of serotype 2 at the TPI (blue line). Standard deviations, based on 10 simulated successful invasion events, are shown as light blue bars. Parameters as in Table 2.1 (on page 87) and $\rho_\mu = 0.045$.

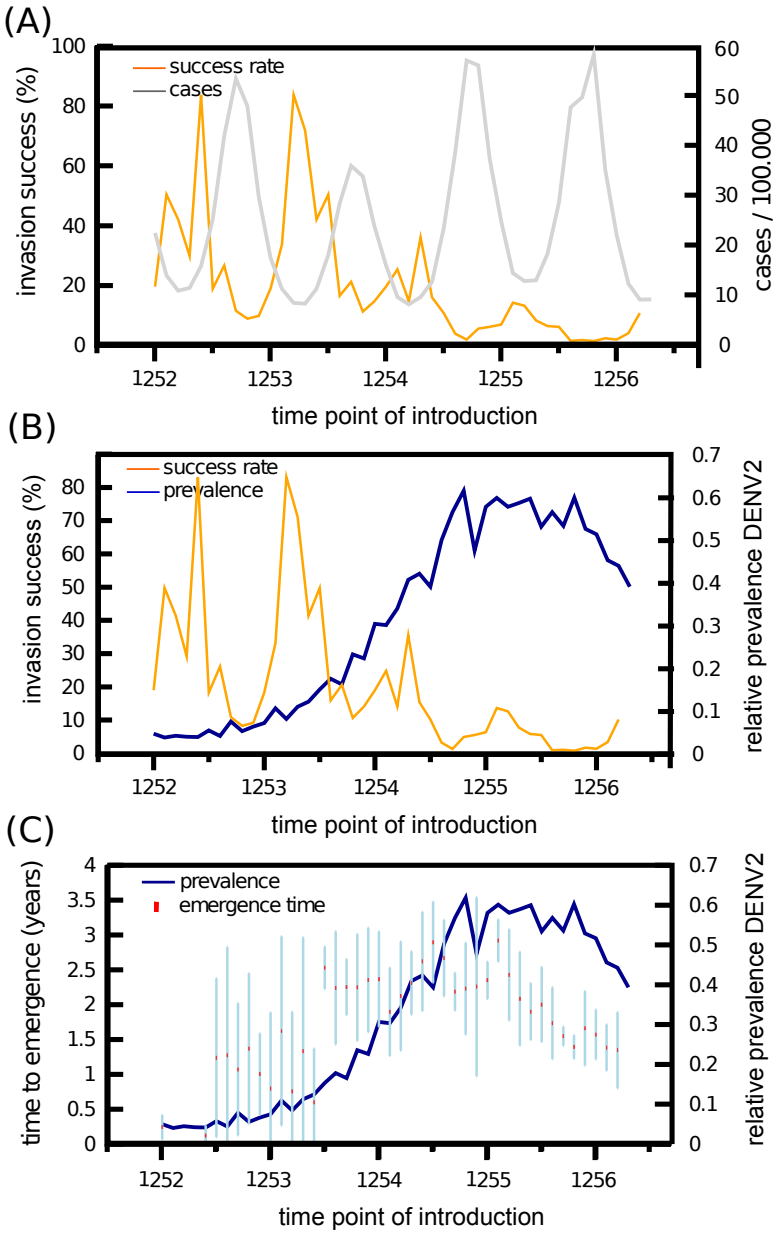


Figure 2.11: **The Effect of Viral Competition on Invasion Success and Emergence Time For Fitness in Human Infectious Period - (A)** The invasion success (orange line) of the invading genotype oscillates out of phase with total dengue incidence (grey line) at the time of introduction (TPI). **(B)** The highest rates of successful invasions can be observed for low relative prevalence of serotype 2 (blue line) at the TPI. **(C)** The red points show how the average emergence times, i.e. the period between introduction and reaching a 10% detection threshold, of successful invasion events increases with the relative prevalence of serotype 2 at the TPI (blue line). Standard deviations, based on 10 simulated successful invasion events, are shown as light blue bars. Parameters as in Table 2.1 (on page 87) and $\rho_\sigma = 0.045$.

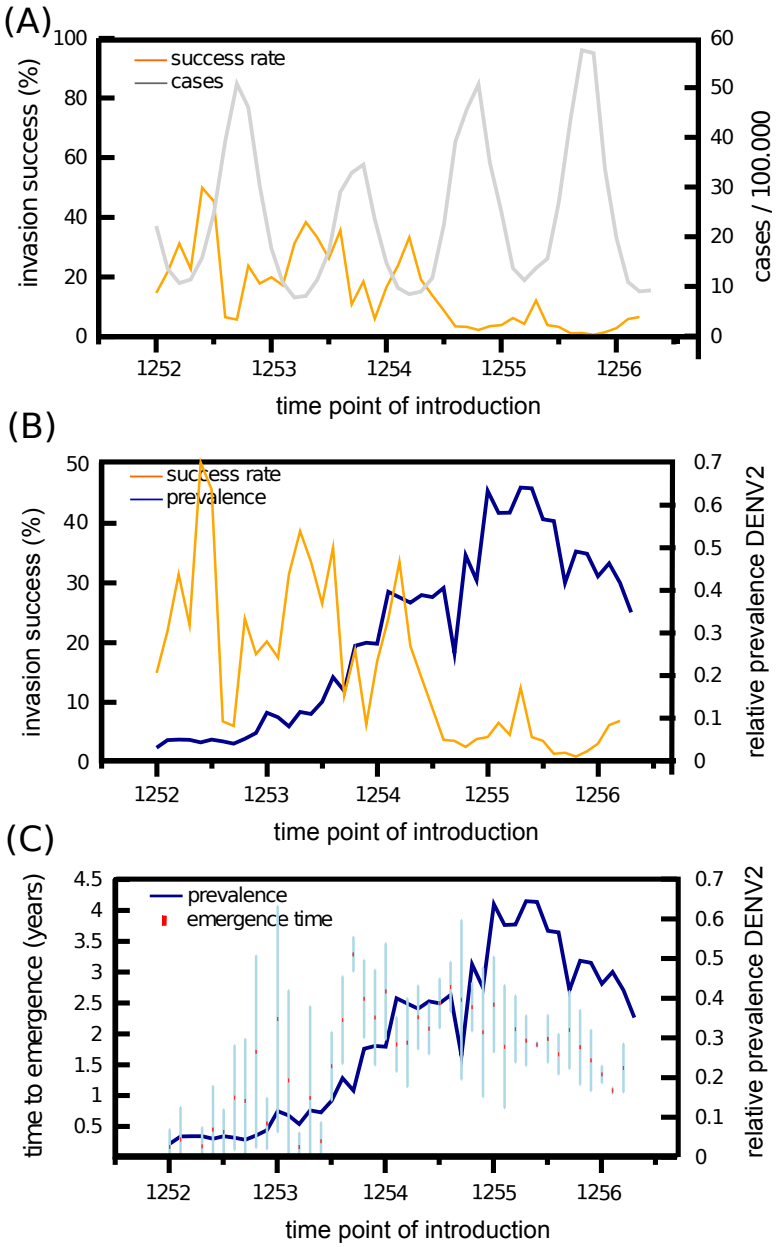


Figure 2.12: **The Effect of Viral Competition on Invasion Success and Emergence Time For Fitness in ADE** - **(A)** The invasion success (orange line) of the invading genotype oscillates out of phase with total dengue incidence (grey line) at the time of introduction (TPI). **(B)** The highest rates of successful invasions can be observed for low relative prevalence of serotype 2 (blue line) at the TPI. **(C)** The red points show how the average emergence times, i.e. the period between introduction and reaching a 10% detection threshold, of successful invasion events increases with the relative prevalence of serotype 2 at the TPI (blue line). Standard deviations, based on 10 simulated successful invasion events, are shown as light blue bars. Parameters as in Table 2.1 (on page 87) and $\rho_\phi = 0.045$.

2.4 Discussion

The invasion pattern of a novel dengue genotype into an endemic population was analysed. In the modelling framework it was considered that the invading genotype possesses a fitness advantage over the resident type. Most of the results were based on the assumption of enhanced transmissibility from infected human individuals to the mosquito vectors. This followed the findings by Hang *et al.*, who reported increased plasma viraemia levels in patients infected by Asian-1 DENV2 viruses during a genotype replacement event [Hang et al., 2010]. In contrast to other studies, Hang and colleagues did not find increased infectivity of Asian-1 viruses to *aegypti* mosquitoes *per se*. However, it is easy to envisage how higher viral titres could enhance the per bite probability of human-to-vector transmission [Armstrong and Rico-Hesse, 2001, 2003].

By focusing on a small increase in transmissibility during primary and secondary infections, the results presented here demonstrated that the total time for genotype replacement is composed of two phases. First, the invading type circulates at very low prevalence levels for several years, after which a rapid shift in dominance and competitive exclusion of the resident type takes place. The same pattern of invasion was described by Hang *et al.*, as the Asian-1 genotype was first detected in 2003 despite phylogenetic analyses dating the introductory event sometime during the late 1990s [Hang et al., 2010]. If this is a universal signature of the population dynam-

ics of dengue viruses remains an open question, mostly because replacement events are generally identified in phylogenetic analysis and the prevalence frequencies of the intervening types are rarely measured in time.

Of particular interest is the time lag between introduction and emergence, when detection of the new genotype might be difficult by surveillance systems based on low viral sampling numbers and/or infrequent genotyping. Furthermore, as the epidemiological patterns would remain largely invariant, passive surveillance systems based simply on case numbers could also easily fail to detect this intra-serotype replacement event.

Apart from increased transmission from infected humans to the mosquito vectors, other viral traits were also considered, including longer infectious periods, increases in ADE enhancement and shorter EIP. While the actual viral trait did not alter the overall replacement dynamics, it was found that different traits can have asymmetric effects on time and success of fixation. Importantly, they can also have addictive effects and thus the combination of very small individual enhancements in multiple viral traits may be sufficient to explain selective replacements. This is particularly the case for events, such as the one of the Asian-1 genotype in South East Asia, that do not present major shifts in the sero-epidemiological patterns of dengue.

In addition to viral fitness, the time point at which a novel genotype enters a population was found to be crucially important in determining its invasion dynamics and

ultimate success. Whereas the relative fitness advantage affected the overall time between introduction and fixation, the epidemiological profile more strongly determined the period of low level prevalence before the advantageous genotype emerged. Various epidemiological factors were tested for their influence on the emergence time but only the relative prevalence of serotype 2 was found to have a strong effect. In fact, while transmission intensities strongly affected invasion success, the level of prevalence of serotype 2 within the population determined both the invasion success rate and, independently, the period before the invading genotype would reach a sufficient level to be widely detectable. These results thus demonstrated that serotype interactions, clonal interference and the current epidemiological landscape can have a big influence on intra-serotype dynamics and therefore the observed viral evolution.

There is considerable interest in determining the evolutionary processes that underlie the observed structures and genetic variation of dengue virus populations (both inter- and intra-serotypic). Overall, low estimates of positive selection and the fact that dengue has a two-host life-cycle are commonly used to place purifying selection as the strongest selective force acting on DENV's molecular evolution [Holmes, 2003; Klungthong et al., 2004; Zhang et al., 2005]. It is also clear that dengue viruses exhibit strong spatio-temporal variations, and genetic drift has been proposed to play a major part in dengue evolution where the replacement of viral lineages or clades could be explained through stochastic processes alone. For example, repeated bottlenecks due to large seasonal fluctuations in mosquito densities imply that the emergence of novel

and possibly advantageous genotypes could be a recurrent phenomenon followed by a strong probability for extinction in the subsequent circulating seasons, explaining the weak signature for positive selection in the data (compared to purifying selection). This in turn would also suggest that the success of a genotype does not always reflect its viral fitness [Bennett et al., 2010; Holmes and Twiddy, 2003].

Indeed, the results presented here have clearly shown that novel genotypes, specially those that arise during large epidemic outbreaks, can face high risks of extinction despite possessing a fitness advantage. Even successful genotypes, i.e. those that eventually reach fixation, potentially undergo prolonged periods of low frequency which can span for several transmission seasons independently of the epidemics therein. Therefore, low measures of adaptive evolution for dengue viruses can be partially explained by eco-epidemiological factors.

DENV's two-host life-cycle implies a significant evolutionary constraint whereby the majority of newly arising variants are likely to be deleterious and selectively removed from the population. The results presented here indicate that in addition to purifying selection, the epidemiological landscape, clonal interference and stochastic effects might be equally important determinants of dengue's viral evolution.

2.5 Tables

Table 2.1: **Parameters used in the deterministic and stochastic simulations.**

	definition	**value**
μ	host lifespan	70 years
α	temporary heterologous immunity	5 months
σ	infectious period	3.65 days
γ	susceptibility enhancement	1.33
ϕ	transmissibility enhancement	1.66
ρ_β	increase in human-to-mosquito transmission	$0 \leq \rho_\beta \leq 1$
ρ_σ	increase in infectious period	$0 \leq \rho_\sigma \leq 1$
ρ_ϕ	increase in enhancement of secondary infections	$0 \leq \rho_\phi \leq 1$
N^h	host population size	9 million
N^v	vector population size	22.5 million
μ^v	vector lifespan	16 days
ϵ	amplitude in seasonality	0.3
η_0	biting rate	115 per year (0.31 d^{-1})
$\beta^{h\to v}$	transmission probability human→mosquito	0.9
$\beta^{v\to h}$	transmission probability mosquito→human	0.8
k	speed in seasonality change	2
	detection threshold (relative frequency)	10%
	deterministic fixation threshold (relative frequency)	99%
	deterministic number of introduced DENV2' (cases)	1 infected mosquito
	stochastic fixation threshold (cases)	0
	stochastic number of introduced DENV2' (cases)	2 per infectious class

For references on the expected biological ranges see Table 4.2 on page 141.

CHAPTER 3

Developing an Individual-Based, Spatially Explicit Framework

Abstract

The results of Chapter 2 suggest that non-selective pressures are a significant determinant of the population structure and dynamics of dengue genotypes. It is thus possible that stochasticity plays a previously unforeseen role in dengue's overall population and epidemiological dynamics. As a first step to study the impact of demographic stochasticity in the epidemiological dynamics of dengue, this Chapter describes the development and output of an individual-based, spatially explicit and single-strain host-pathogen model. Contrasting to deterministic expectations of convergence towards a stable equilibrium, the individual-based model readily presents persistent oscillatory behaviour in disease incidence, solely driven by demographic stochasticity. When host population structure is introduced, the stochastic nature of local transmission leads to the emergence of spatially heterogeneous immunity landscapes. These results highlight the role of natural variation and stochasticity in commonly observed spatio-temporal features of single-strain host-pathogen systems, with important implications for multi-strain systems.

3.1 Introduction

The most usual modelling methodology in host-pathogen systems is the use of ordinary-differential equations. Common to the continuous and deterministic class of such models is their propensity to exhibit damped oscillations around an approaching equilibrium, in which the rate of new infections equals the loss from the infectious pool due to (host) recovery and death [Anderson and May, 1992]. In reality, however, many infectious diseases will not remain in this state of equilibrium but instead exhibit persistent oscillations, ranging from seasonal increases in incidence rates, to multi-annual outbreaks. Empirical evidence points to a variety of natural mechanisms by which host–pathogen interactions and their ecological context can be significantly altered in time and space, possibly inducing the observed oscillatory patterns. The most common are changes in host social behaviour, climatic fluctuations, host-pathogen adaptation, strain competition and pulses in host natality, mortality or migration [Altizer et al., 2006]. Accordingly, *periodic forcing* is often incorporated into deterministic susceptible-infectious-recovered (SIR) models, which moves the underlying system away from its natural equilibrium into a regime characterised by periodic oscillations, akin to those observed in nature [Altizer et al., 2006; Bauch and Earn, 2003].

While ad-hoc forcing is generally sufficient, its inclusion in mathematical frameworks has the drawback of extra parametrization. Moreover, it is not always clear which, amongst possible biological drivers, such as host-seasonal behaviour or climate fluc-

tuations, justify the introduced periodicity chosen to reflect the empirical one [Altizer et al., 2006]. Alternatively, *demographic stochasticity* or natural variability, inherently present in individual host-pathogen interactions, has been shown to be a parsimonious driver for the emergence of persistent oscillatory behaviour in dynamical systems [Alonso et al., 2007; Bartlett, 1957; Verdasca et al., 2005].

From the point of view of continuous deterministic frameworks, stochasticity is a source of perturbations that allow the system to remain in a transient regime instead of approaching an expected equilibrium. Seemingly global behaviours can thus emerge as bottom up phenomena based on a wide spectrum of local erratic behaviours [Breckling et al., 2006; DeAngelis and Mooij, 2005; Hanski, 1998]. Importantly, stochastic formulations have often generated richer dynamical properties than their deterministic counterparts, and have contributed to a more complete understanding of the modelled systems [Bolker and Grenfell, 1993; Buckee et al., 2011; Chesson, 1981; Durret and Levin, 1994; Favier et al., 2005; Verdasca et al., 2005].

In recent years, mathematical modellers have become increasingly concerned with the role of stochasticity and heterogeneity, slowly shifting towards frameworks that incorporate such factors at multiple biological levels [Huston et al., 1988; Levin et al., 1997]. Mathematical frameworks that readily incorporate a spatial dimension, stochasticity and variation at multiple scales, such as agent-based models (ABM), are becoming of particular interest [DeAngelis and Mooij, 2005]. In theoretical ecology, for

example, some studies have addressed the effects of demographic stochasticity on the population dynamics of predator-prey systems, or in the qualitative behaviours of idealised *meta-populations* of single and multiple species [Levins, 1969; McKane and Newman, 2005; Simonis, 2012]. In theoretical epidemiology, on the other hand, studies have mostly focused on single-strain host-pathogen systems [Alonso et al., 2007; Keeling, 2000; Verdasca et al., 2005] and the role of natural variation in the population dynamics of multi-strain pathogens has not yet been addressed.

In this chapter, an agent-based model with spatially explicit host-population structure is developed and the effect of stochastic amplification on the epidemic behaviour and spatial distribution of a single-strain pathogen are analysed. The underlying framework will serve as the starting point of Chapter 4 (on page 113), in which the effects of demographic stochasticity are studied in the context of a multi-strain dengue-like pathogen.

3.2 Materials and Methods

3.2.1 Epidemiological Framework

The model presented is an agent-based, discrete and stochastic *single-strain host-pathogen system*. The population is divided into different classes, each representing an epidemiological state, akin to those defined in a classical SIR framework and extended to include an *exposed* set of carriers not yet infectious [Anderson and May, 1992; Ker-

mack and McKendric, 1927; Otto and Day, 2007]. At each time step, individuals may transit between states, expressing only one at the time. The population-level deterministic system provides the backbone for host state transitions, defined as follows:

$$\frac{dS}{dt} = \mu N - (\lambda + \mu + \theta)S \quad (3.1)$$

$$\frac{dE}{dt} = \lambda S + \theta S - (\gamma + \mu)E \quad (3.2)$$

$$\frac{dI}{dt} = \gamma E - (\delta + \mu)I \quad (3.3)$$

$$\frac{dR}{dt} = \delta I - \mu R \ , \quad (3.4)$$

where $\lambda, 1/\gamma, 1/\mu, 1/\delta$ are the force of infection (see bellow), the latency period, the expected life-span and infectious period of the host, respectively. The host population size, N, is assumed to be constant. In discrete systems, such as this one, pathogen extinction is possible. When the disease-free states are not the focus of study, high values for transmission parameters or externally acquired infections (a proxy of host migration or mobility), are often considered to assure persistence [Anderson and May, 1992]. Here, a small number of externally acquired infections (θ) is used.

State changes are addressed as probabilistic events per individual, with probabilities equal to the corresponding rates in the deterministic formulation. For example, the probability of an infectious individual recovering is $P_{I \to R} = \delta$ per time step. On average, each infectious host will spend $1/\delta$ time steps in state I (although the values of the model's parameters are fixed, demographic stochasticity is assured from the

probabilistic and discrete nature of state transitions).

A *standard incidence* is considered, where the average *contact rate* between individuals is a constant C, and does not scale with population size (as opposed to *mass-action incidence*). The force of infection, λ, is defined as $\lambda = \epsilon C \times I/N$, with ϵ being the probability of infection given an infectious contact. With demographic stochasticity present, the number of new infections (incidence) can deviate from the deterministic potential ($\lambda S + \theta S$) and can be quantified in time by the *effective incidence ratio*:

$$R_\lambda(t) = \frac{incidence(t)}{\lambda S(t) + \theta S(t)} \quad . \tag{3.5}$$

3.2.2 Population Structure and Connectivity

In ecology, a meta-population is a conceptual population of single or multiple species whose distributions can be formalized as a system composed of a set of patches, connected by dispersion rates and subjected to local extinction [Grenfell and Harwood, 1997; Hanski, 1998; Levins, 1969]. Setting a pathogen as the species of interest, the direct analogy in epidemiology is to consider hosts as pathogen-harbouring patches, subject to infection, recovery and death [Grenfell and Harwood, 1997].

An alternative approach, such as the one taken here, defines patches as communities of hosts and dispersion as transmission events between distant patches. This allows to explore both the microscopic (individual) and macroscopic (eco-spatial) effects of

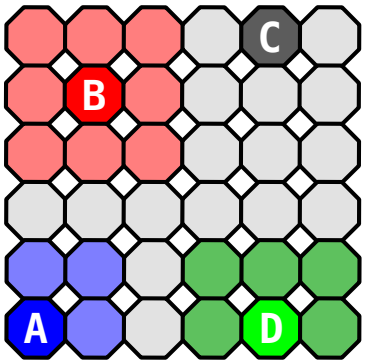


Figure 3.1: **Approach to Meta-Population Structure** - Schematic of a meta-population in a 6x6 lattice. Each community C_j has K_j neighbours. Localities are exemplified by the coloured communities surrounding C_A, C_B and C_D. Each community includes a population of homogeneously mixing hosts. Contacts of individuals residing in C_j are mainly local, with equal probability within the locality ($\{C_j, K_j\}$). Long distance contacts due to mobility, for example between a host in C_D and C_C can happen with probability ω.

demographic stochasticity, by segregating hosts into well-defined social units - for example families, villages, towns, cities or countries - therefore modelling ecologically relevant host-population units for pathogen persistence and evolution.

In this framework, spatial structure is added by subdividing the host population into a spatially organized set of communities (patches), forming a squared and non-wrapping lattice with x rows/columns and community size L. In this setting, each community C_j has K_j neighbours, with the set $\{C_j, K_j\}$ defining a *locality* (see figure 3.1 on page 94).

To include host movement, a connectivity parameter ω is considered. Here, ω defines the probability of an individual's contact happening with another individual from a distant community, and is therefore the expected proportion of long distance contacts

of a host/community per time step. It contrasts with other meta-population formulations that generally define connectivity as a constant (movement) rate between all subpopulations [Earn et al., 2000a; Keeling, 2000]. Hosts in community C_j have contacts within their locality with probability $(1 - \omega)$, i.e. amongst individuals within C_j and its neighbouring communities K_j; and outside with probability ω. Mixing within the locality is assumed to be homogeneous

The parameter ω can also be interpreted as the *outward connectivity* level of the host population, i.e. the number of infective contacts a single infectious individual is expected to have outside its locality during the infectious period:

$$\Omega_{out} = \frac{C\epsilon\omega}{\delta} \in [0, \approx R_0] \ , \tag{3.6}$$

where

$$R_0 = \frac{C\epsilon\gamma}{(\mu + \gamma)(\mu + \delta)} \tag{3.7}$$

is the basic reproductive number.

For the analysis of spatial clustering across the meta-population, such as susceptibility or herd-immunity pockets, the Ripley's H-function is used, which has been successfully applied to multiple biological scales in host-pathogen systems [Kiskowski et al., 2009; Liebman et al., 2012]. With the original function (the K-function), aggregation within a certain area is identified if the average number of points (occurrences) within

a distance r of another existing point (occurrence) is statistically greater than that expected for a random distribution. The H-function is a normalization of the K-function such that the expected value is 0:

$$K(r) = \frac{1}{n}\sum_{i=1}^{n}\frac{N_{pi}(r)}{\lambda}, \quad L(r) = \sqrt{K(r) - \pi}, \quad H(r) = L(r) - r. \tag{3.8}$$

This allows for the point distribution to be tested against the null hypothesis that the points are distributed randomly and independently. That is, a positive value of $H(r)$ indicates clustering over the spatial scale r, whereas a negative value indicates dispersion.

3.2.3 Host Mortality

Initially, age-independent mortality is assumed, in which death events have a fixed probability of μ per time step, according to the mean life expectancy of $1/\mu$. An age-dependent method is later considered, in which the age-dependent part of the Gompertz-Makeham (GM) law of mortality is used to derive age-associated death probabilities [Gompertz, 1825]:

$$\mu(x) = (ae^{bx})e^{-\frac{a}{b}(e^{bx}-1)} \quad , \tag{3.9}$$

where x is the host age, a and b the Gompertz function variables. Figure 3.2 on page 97 presents an example of this function, showing the death probabilities per age and

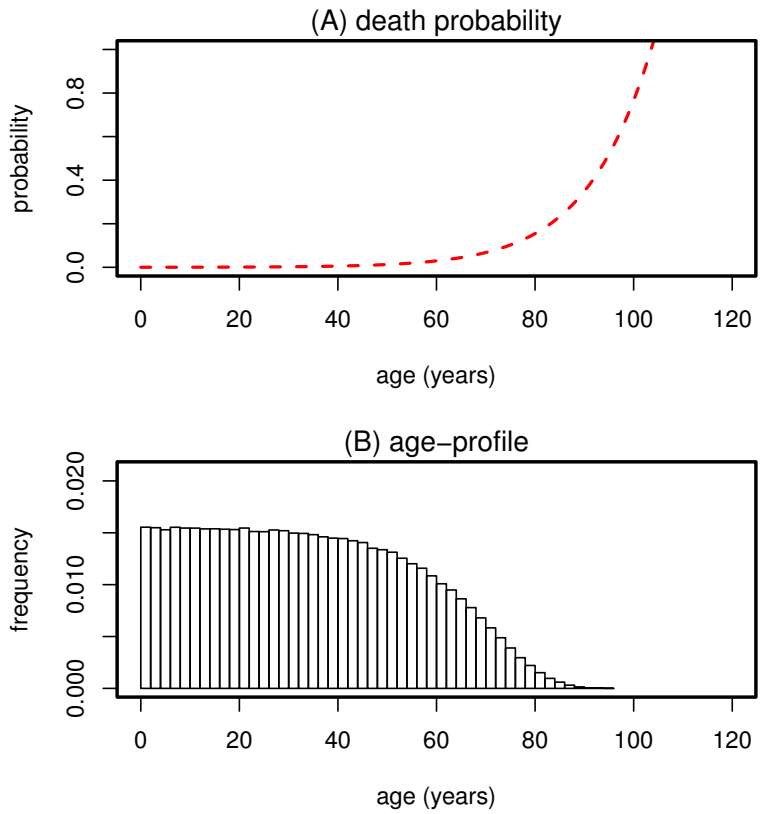


Figure 3.2: **Approach to Age-Dependent Mortality** - Example on the use of Gompertz-Makeham's law of mortality to model age-dependent mortality rates in a population of N $\approx$ 1 million individuals. **(A)** Death probability per age. Death events are assessed per time step (day), with probability $p = \mu(x)/364$, where x is the individual's age. Parameters $a = 0.0002$ and $b = 10$. **(B)** Resulting age-profile with mean $\approx$ 33 and life-span $\approx$ 67.5 years.

resulting age-profile of the population.

Since total population size is constant, each death is followed by a birth event independently of the method used.

3.3 Results

3.3.1 General Behaviour

The progression of agent-based SEIR epidemics is illustrated in figure 3.3 (on page 99) in the absence of population structure, i.e. in an homogeneously mixing scenario. As expected, the dynamics of hosts with specific epidemiological states approximate the deterministic counterpart. Nonetheless, some numerical simulations do not develop into outbreaks due to pathogen extinction, while others present variations in peak height or epidemic length, due to amplified stochastic effects happening at the individual level. That is, during outbreaks, demographic stochasticity causes the transmission success of the pathogen to deviate from the deterministic potential (figure 3.4 A on page 100), especially when the number of infected individuals is small. The same effect is observed in the length of the latency periods that individuals experience (figure 3.4 B). While its average is close to the deterministic expectation of $1/\gamma$ days, significant variations are observed simply due its probabilistic nature in the model.

This variability, arising at multiple scales within a single epidemic, has important effects in the long term dynamics of the single-strain system [Alonso et al., 2007; Verdasca et al., 2005]. An example is illustrated in figure 3.5 on page 101, where the discrete homogeneous model is shown to follow the expected mean behaviour of the continuous model but does not converge to the deterministic equilibrium.

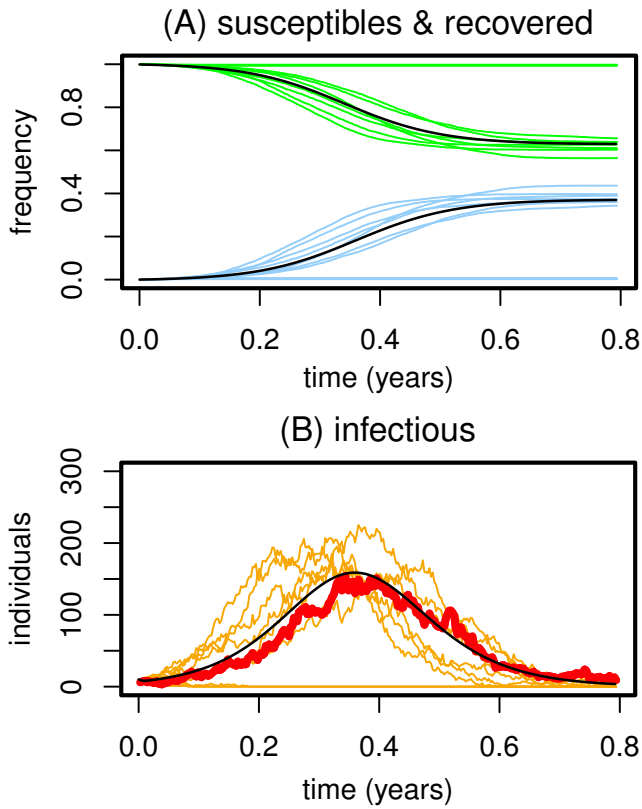


Figure 3.3: **General Dynamic Behaviour of Single Epidemics When Assuming Homogeneous Mixing - (A,B)** Comparison of the output of deterministic (black lines) and agent-based SEIR simulations. Ten solutions are presented for the ABM, 2 non-persistent. The dynamics of the susceptible (green), recovered (blue) and infectious (orange) classes are presented. Parameters: $\epsilon = 0.05$, $\delta = 5d^{-1}$, $\gamma = 2d^{-1}$, $\mu = 60y^{-1}$, $C = 5$, $\theta = 0$, $N = 10000$, initial infectious $I = 10$ ($R_0 \approx 1.25$).

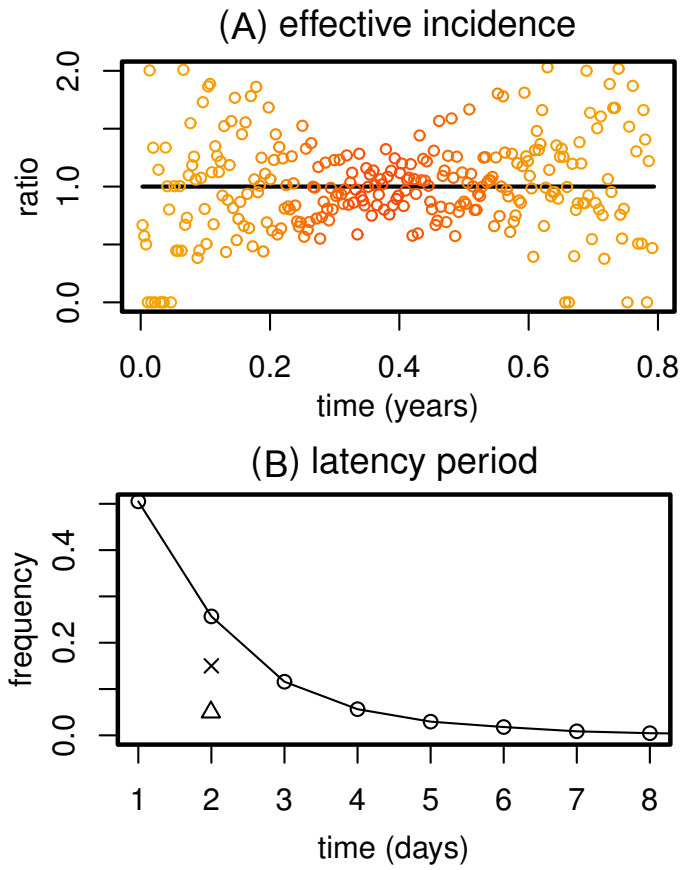


Figure 3.4: **Short Term Effects of Demographic Stochasticity Within Single Epidemics When Assuming Homogeneous Mixing - (A,B)** Demographic stochasticity acting on the effective incidence and on the time spent in the exposed class for one agent-based simulation (thick line in figure 3.3 **B** on page 99). **(A)** Black line is the expected effective incidence ratio, circles are the deviations during the epidemic. **(B)** Distribution of latency periods experienced by hosts. Projections on the horizontal axis of the cross and diamond are the expected and resulting mean period, respectively. Parameters: $\epsilon = 0.05$, $\delta = 5d^{-1}$, $\gamma = 2d^{-1}$, $\mu = 60y^{-1}$, $C = 5$, $\theta = 0$, $N = 10000$, initial infectious $I = 10$ ($R_0 \approx 1.25$).

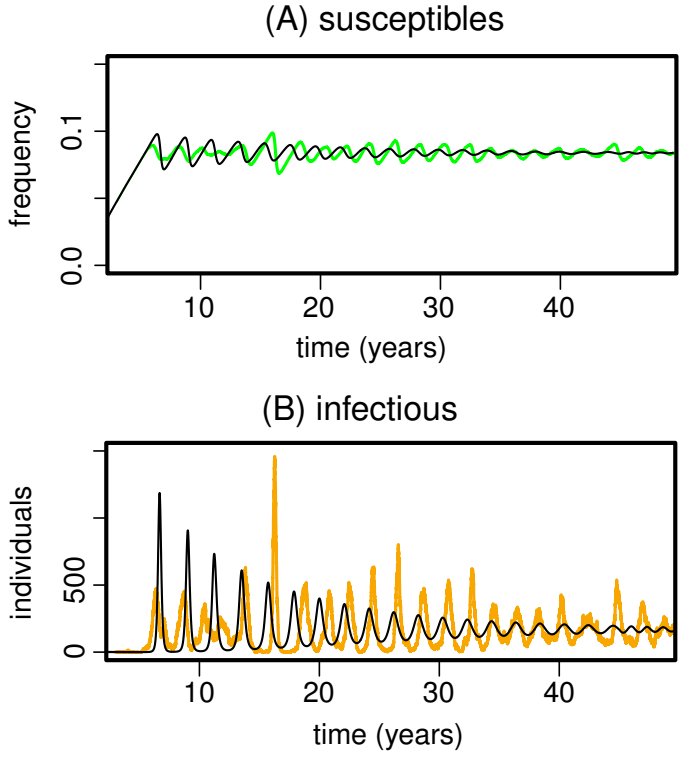


Figure 3.5: **Long Term Effects of Demographic Stochasticity When Assuming Homogeneous Mixing** - Comparison of the output of one deterministic (black, solid) and agent-based SEIR simulations. The dynamics of the **(A)** susceptible (green) and **(B)** infectious (orange) classes are presented. The ABM follows the expected mean behaviour but presents persistent oscillatory behaviour. Parameters: $\epsilon = 0.3$, $\delta = 4d^{-1}$, $\gamma = 2d^{-1}$, $\mu = 60y^{-1}$, $C = 10$, $\theta = 1$ per month, $N \approx 1$ million ($R_0 \approx 12$).

3.3.2 Expanding the Framework: Age-Dependent Mortality

In a continuous age-independent framework, such as the one illustrated in figures 3.3 to 3.5, the parameter μ defines the rate at which population turn-over takes place, thus defining the mean life expectancy of the individuals as $1/\mu$ years. Within a discrete and agent-based system, however, this does not equate to realistic age-profiles and may introduce undesired biases. Figure 3.6 (on page 103) compares the use of age-dependent and independent mortality rates and their age profiles. Although, the resulting mean life expectancies at equilibrium are similar for both methods (figure 3.6 A), the age distributions are significantly different (figure 3.6 B). With age-independent mortality the age span is widely unrealistic and the mean age therefore high. Biases in model output are therefore expected depending on transmission levels. That is, for instance, if transmission is high and exposure expected at early ages, immune individuals might be alive for twice as long as defined by μ. To avoid such effects, the age-dependent method is preferred for the discrete approach and used hereafter.

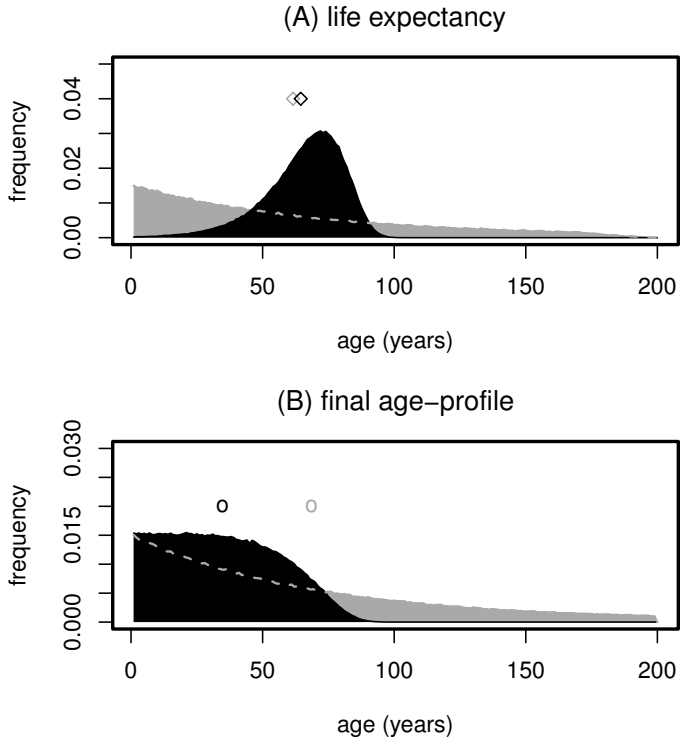


Figure 3.6: **Effects of Age-Dependent and Age-Independent Mortality Rates** - Comparison between age-dependent (black) and age-independent (grey) mortality using the ABM for $N \approx 1$ million individuals. Age-independent parameter $\mu = 67.5y^{-1}$. **(A)** Frequency of death events per age at equilibrium, diamonds are the mean life span. Deviations from expected mean is due to demographic stochasticity. Parameters as in figure 3.2 on page 97. **(B)** Age-profiles at equilibrium. Circles are mean age. The age-independent method produces an unrealistic age profile that may introduce undesired biases (see 3.3.2 on page 102).

3.3.3 Expanding the Framework: Population Structure

Both population structure and patch connectivity modify host-mixing and can therefore alter global dynamics. To start addressing their effects, mixing is assumed to be independent of spatial structure by setting the probability of long distance contacts, or connectivity, to be maximum ($\omega = 1$). In this case, the system's dynamics are expected to be invariant to the level of (explicit) structure added to the population, and behave similarly to an homogeneously mixing system. This is illustrated in figures 3.7 A,B (page 105), which show the dynamics of the susceptible and infectious portion of the population when considering three different levels of structuring. As expected, the mean epidemic size and period is invariant to structure changes and equal to the homogeneous model (single community) of figure 3.5 on page 101.

When structure is present and connectivity is maximum ($\omega = 1$), the probability of a transmission event taking place in a specific community is $1/N_C$, with N_C being the number of communities. Hence, even though the pathogen keeps circulating independently of the spatial dimension, not all communities or hosts experience equal forces of infection in time and space. This is shown in figure 3.8, which presents the susceptibility levels within single communities for meta-populations with low and medium structure when $\omega = 1$. Although both meta-populations present the same average level of susceptibility, as demonstrated by their dynamics in figure 3.7 (on page 105), the more structured population shows higher spatial heterogeneity. This is a result of the

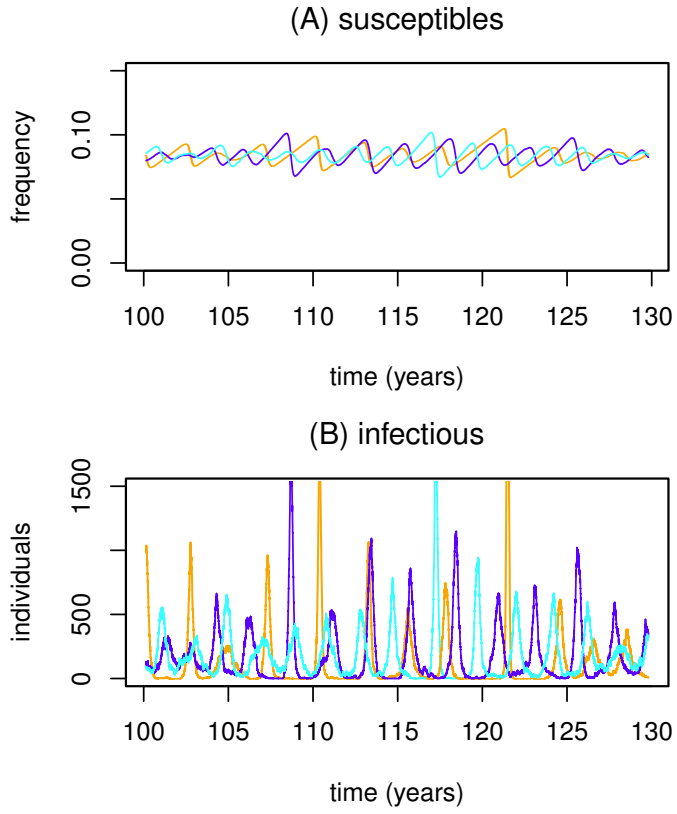


Figure 3.7: **General Dynamics When Assuming Homogeneous Mixing With Explicit Population Structure** - Examples of similar dynamics of **(A)** susceptible and **(B)** infectious hosts in a homogeneous mixing system ($\omega = 1$). The level of explicit structure in the host population is varied: no structure, low structure (community size 62500, 4x4) and medium structure (community size 5000, 20x20). Quantitative measures of the epidemic behaviour - mean and epidemic period - are independent of structure (no structure: 169.7, 2.4; low structure: 169.2, 2.5; medium structure: 168.9, 2.3; average of 25 simulations). Parameters as in figure 3.5 on page 101 with age-dependent mortality as in figure 3.2 on page 97.

reduced probability of cases happening in each community with increasing N_C, allowing for local variations to accumulate. Hence, it is easy to envisage how relaxing the homogeneous mixing assumption ($\omega = 1$) and allowing transmission to become more localised, as is often the case for real host-pathogen systems, can result in significant spatio-temporal dynamical changes.

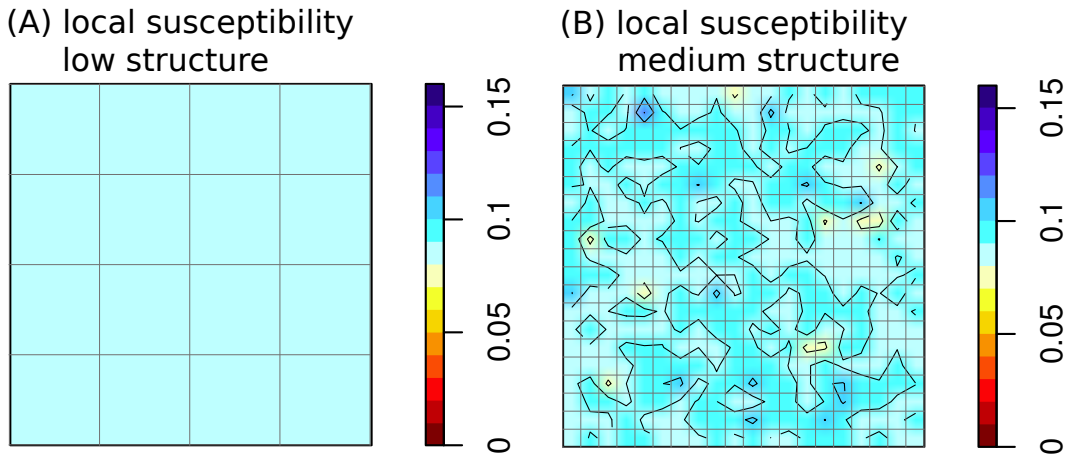


Figure 3.8: **Emergence of Susceptibility Landscapes When Assuming Homogeneous Mixing With Explicit Population Structure** - Susceptibility landscapes while assuming homogeneous mixing ($\omega = 1$). The probabilistic nature of transmission events happening at each community ($p = 1/N_C, N_C$=number of communities) dictates that spatial variation can emerge even though the pathogen keeps circulating independently of the spatial dimension. The level of explicit structure in the host population is varied: **(A)** low structure (community size 62500, 4x4) and **(B)** medium structure (community size 5000, 20x20). The epidemic dynamics are in figure 3.7 on page 105. Parameters as in figure 3.5 on page 101 with age-dependent mortality as in figure 3.2 on page 97.

3.3.4 Expanding the Framework: Host Mobility

The effects of relaxing the assumption of homogeneous mixing on local variation are illustrated in figure 3.9 (on page 107). It is shown that variation in local susceptibility across the meta-population increases with decreasing connectivity, a result of reduced community synchronization (i.e. timing and epidemic size). Notably, the relationship between connectivity and local variation appears to be separated into three regimes: (1) when contacts outside the locality dominate and local variation does not build up (white area, $\Omega_{out} > R_0/2$); (2) when contacts inside the locality dominate with at least one contact outside being expected, leading to slight increases in variation for decreasing connectivity (light-grey area, $\Omega_{out} \in [1, R_0/2]$); and, (3) when less than

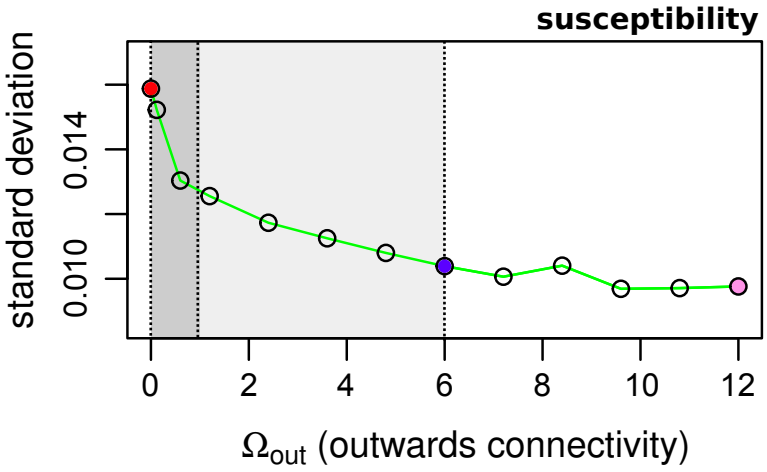


Figure 3.9: **Effects of Relaxing Homogeneous Mixing With Explicit Population Structure** - Decreasing Ω_{out} enhances the effects of local demographic stochasticity on susceptibility as the system moves away from the homogeneous mixing behaviour (mean of 25 simulations). Shaded areas: regimes in which varying connectivity affects local susceptibility with different rates. Coloured points are the scenarios in which $\Omega_{out} = 0$ (no mobility, red), $\Omega_{out} = 6 = R_0/2$ (balanced local and distant contacts, blue), $\Omega_{out} = 12 = R_0$ (homogeneous mixing, pink). Parameters as in figure 3.7 on page 105 but now for high structuring (community size 800, 50x50), $N \approx 2$ million.

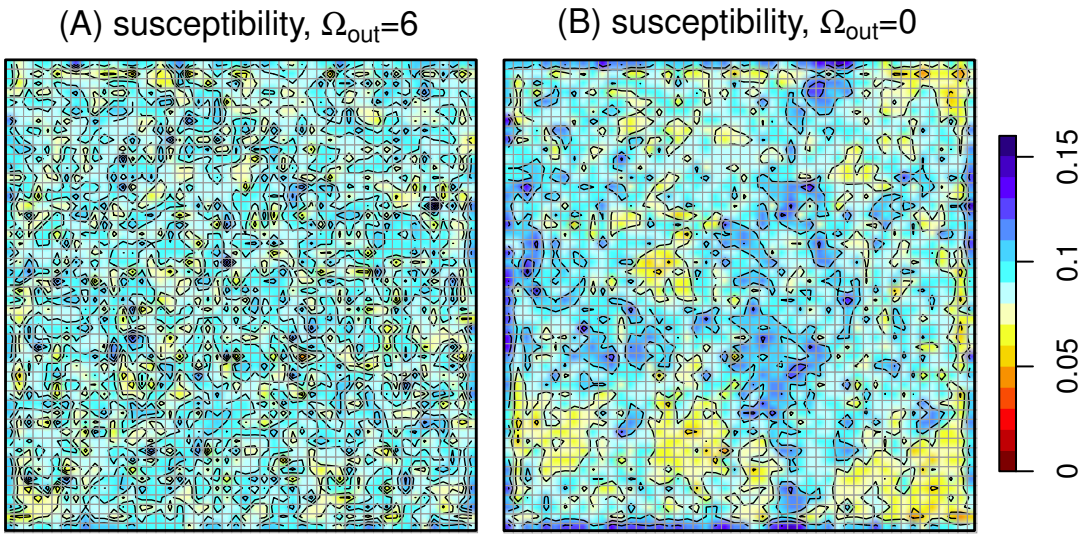


Figure 3.10: **Spatial Patterns of Susceptibility Landscapes Depending on Connectivity - (A,B)** Susceptibility landscape for 1 simulation with varying connectivity (blue and red points in figure 3.9 on page 107). Mean susceptibility ≈ 0.08, light blue, as before (figure 3.7 A on page 105). Parameters as in figure 3.7 on page 105 for high structuring (community size 800, 50x50), $N \approx 2$ million.

one contact outside the locality is expected, resulting in significant increases in local variation for slight decreases in connectivity (dark-grey area, $\Omega_{out} < 1$).

The spatial patterns of the emergent local variation are also different depending on the level of connectivity. Figure 3.10 (on page 107) presents an example that compares the susceptibility landscapes that result from meta-populations that differ by whether long distant contacts are possible. Qualitatively, there are significant differences between the two scenarios with the meta-population with higher connectivity presenting high variation between adjacent communities. These differences can be quantified using he H-function (see Materials and Methods on page 95), which measures the spatial scale, i.e. radius in terms of number of communities, of areas with similar susceptibility levels (figure 3.11 on page 109). As expected from the fine-grained distribution of the susceptibility landscape of the population with frequent distant contacts, variation between close ranged communities is so common that clustering is not detected. In contrast, the patchy distribution from the population in which transmission is restricted within localities reflects that communities in close range have similar susceptibility levels and thus clustering is detected.

Overall, these results suggest that the level of heterogeneity in the susceptibility landscape of the meta-population is significantly influenced by the connectivity between individuals, and a wide range of landscapes can be reproduced independently of other biological parameters, including the ones that define the pathogen under focus.

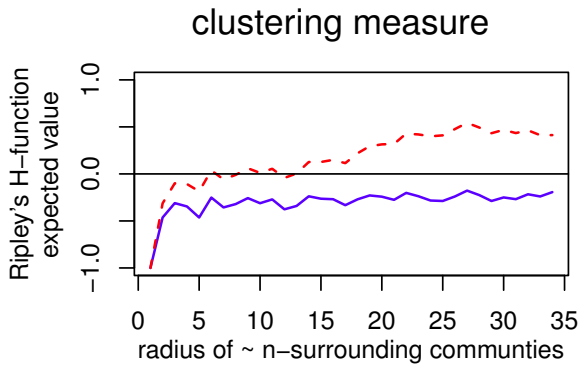


Figure 3.11: **Clustering Within Susceptibility Landscapes** - Analysis of susceptibility patterns from the meta-populations of figure 3.10 on page 107. Clustering of susceptibility is consistently detected for radius above 12 communities when $\Omega_{out} = 0$ (H-function > 0, red) but the fine-grained distribution for Ω_{out} only allows to detect dispersion (blue).

3.4 Discussion

Following the observations of Chapter 2 that natural stochasticities may play a significant role in the population structure and dynamics of dengue viruses, the question of whether demographic stochasticity plays a more general role in the overall epidemiological behaviour of dengue was raised. To address this, however, a stochastic, spatially explicit and multi-strain framework is required. As a first step towards the implementation of such a framework, this Chapter presented the development and analysis of an agent-based, single-strain model with explicit host-population structure. The results here presented have confirmed and expanded previous reports on the effects of demographic stochasticity in the epidemiological dynamics of single-strain pathogens.

Specifically, it was found that demographic stochasticity is a parsimonious mechanism that can lead to persistent oscillatory behaviour in disease incidence. It allows the dynamical system to remain under the influence of the existing attractors or repellors, while generating sufficient perturbations to be kept in a transient, non-converging regime. In the case of this framework, demographic stochasticity was sourced from the discreteness of individuals and the probabilistic nature of the host-host and host-pathogen interactions.

Adding to the results of previous studies that did not account for a spatial dimension, it was demonstrated how amplification of stochastic effects at the individual level can drive the emergence of heterogeneous susceptibility landscapes. This is a common property of real host-pathogen systems that often present spatial asymmetries in seroconversion rates [Anderson and May, 1992; Veeraseatakul et al., 2007]. Heterogeneous spatial distributions further reflect the existence of spatially segregated susceptibility pockets from which various epidemiological phenomena can emerge, such as wave-like epidemics [Cummings et al., 2004; Grenfell et al., 2001; Thai et al., 2010] and allopatric persistence and emergence of pathogen strains [Grenfell et al., 2004; McElroy, 2011; Verdasca et al., 2005; Weaver and Vasilakis, 2009].

It was also shown that a variety of susceptibility landscapes, from practically homogeneous to highly heterogeneous, can be generated solely by varying demographic factors such as population structure and host mobility, in the absence of biological

changes. As spatial variation in susceptibility further mirrors asymmetries in the immunity profile of hosts across the meta-population, this framework becomes a relevant tool for studying the role of eco-demographic factors in spatio-temporal patterns of sero-conversion, which are often found to be significantly different between urban and rural or developed and developing regions, for instance. In addition, due to the age-dependent mortality here considered, herd-immunity is age associated, which means that the framework readily allows to consider age-effects on transmission probabilities or contact patterns, which are crucial for control and vaccination strategies.

Given these observations for a single-strain pathogen, it is possible that within a multi-strain host-pathogen system demographic stochasticity could also result in significant spatio-temporal differences in the immunity landscapes to each strain. Hence, an expansion of this model into a multi-strain framework is particularly useful for the study of antigenically diverse pathogens whose population dynamics are often believed to be driven by within-host cross-immunological reactions that are dependent on the host's history of infection. The dengue virus is one example, whose theoretical research has mainly focused on mathematical models of transmission addressing the temporal and dynamical effects of protective and/or enhancing cross-immunological interactions between its four co-circulating serotypes (Section 1.5 on page 37). Crucially, these models have not considered demographic variation and have thus not been able to account for the reported differences in serotype spatial distribution and consequent heterogeneous sero-conversion rates [Fried et al., 2010; Raghwani et al., 2011;

Veeraseatakul et al., 2007]. Hence, the impact of heterogeneous immunity landscapes on the proposed effects of cross-immunological reactions is so far not known in the context of dengue. To address this, the next Chapter presents an expansion of the single-strain framework to include the co-circulation of multiple antigenic strains.

CHAPTER 4

Modelling the Dengue Virus Epidemiology

Abstract

To address the effects of demographic stochasticity in multi-strain systems in general and dengue in particular, the agent-based, single-strain framework of Chapter 3 is modified to allow for the co-circulation of multiple strains of the same pathogen. The expanded framework is shown to exhibit all of the expected epidemiological dynamics of continuous multi-strain models, even in the absence of immune competition. After a second framework expansion to include host-population structure, a mosquito-vector and dengue-like biological parameters, it is also shown that the previously proposed populational effects of cross-immunological reactions between dengue serotypes are weakened when considering a spatially explicit host-population structure. Importantly, the addition of the spatial dimension further advances on model validation by reproducing well documented spatio-temporal features of dengue data, such as wave-like epidemics and spatially heterogeneous sero-conversion distributions. The obtained results integrate, for the first time, key temporal and spatial characteristics of dengue's epidemiology and propose that cross-immunological reactions are not the main drivers for the complex dynamic behaviour of dengue serotypes and argue for a critical role of demographic and ecological factors instead.

4.1 Introdution

The epidemiological dynamics of many antigenically diverse pathogens are often characterized by sequential dominance of strains in time. While a pathogen's epidemic periodicity can be partially explained by temporary changes such as climatic fluctuations or host social behaviour [Altizer et al., 2006], the timing by which one dominant antigenic strain is replaced by another is often out of sync with such pressures. Influenza A viruses, for instance, present yearly epidemics driven by a complex interplay of favourable climate and host-behaviour [Altizer et al., 2006], while at the same time exhibiting multi-annual oscillations of its antigenic subtypes [Ferguson et al., 2003]. Similarly, cases of dengue fever (mild and severe) follow a strong seasonal signature, while dengue serotypes generally present oscillations in prevalence that peak every 8-9 years (Section 1.2 on page 27) [Nisalak et al., 2003].

In order to explain strain replacement dynamics, theoretical studies have often concentrated on factors such as cross-protection or cross-enhancement of infection between strains. In time, these within-host mechanisms can offer temporary advantages to subsets of strains through competition for susceptible hosts and can thus cause their amplification and subsequent decline. This selective mechanism has been proposed to underlie the population biology of a variety of human pathogens, including the influenza A virus [Ferguson et al., 2003], *Plasmodium falciparum* [Gupta and Day, 1994], *Vibrio cholerae* [Koelle et al., 2006b], dengue virus (Section 1.5 on page

37), respiratory syncytial virus [White et al., 2005] and rotavirus [Pitzer et al., 2011]. In the case of dengue, different immunological interactions in the form of antibody-dependent enhancement or temporary and/or partial cross-immunity have been proposed independently as the driving forces behind its complex epidemiology (Section 1.5 on page 37) [Adams et al., 2006; Aguiar et al., 2008; Cummings et al., 2005; Ferguson et al., 1999a; Nagao and Koelle, 2008; Recker et al., 2009; Wearing and Rohani, 2006]. Although the ODE-based models in these studies have qualitatively captured the virus' epidemiological dynamics, they have not considered natural variability in the host population or in disease transmission across time and space, and thus have not accounted for observed differences in incidence and serotype distributions within endemic regions [Fried et al., 2010; Raghwani et al., 2011; Veeraseatakul et al., 2007].

The consideration of spatio-temporal variability is of particular importance for the understanding and control of vector-borne pathogens such as dengue, where the underlying drivers of the observed epidemiologies may be confounded by substantial spatial heterogeneities and time fluctuations in vector densities through space and time. In addition, from the results of Chapter 2 (on page 60) as well as phylogenetic studies [Bennett et al., 2010; Myat Thu et al., 2005], it is possible that demographic stochasticity is a determinant factor of the observed population structure of the pathogen. Following the results of Chapter 3, there is also a general expectation that the natural emergence of oscillatory behaviour by stochastic means in single-strain pathogen systems could be replicable for a system with multiple co-circulating strains. That

is, each strain may be kept in a transient, non-converging regime, leading to independent oscillations in time. This, however, has not yet been explored in modelling frameworks and the role of demographic stochasticity in the observed population dynamics of multi-strain pathogens is unknown. Moreover, the results of Chapter 3 have also implicated demographic stochasticity in the emergence of heterogeneous immunity landscapes, which, if presenting significant asymmetries between the circulating strains, must influence the proposed effects of within-host cross-immunological reactions that are dependent on the host's history of infection. However, this has also not been addressed in previous modelling studies.

Here, using an extension of the agent-based framework of Chapter 3, stochasticities underlying host-pathogen ecologies are shown to give rise to persistent oscillations in multi-strain pathogen systems, even in the absence of strong competition between antigenic types. In the context of dengue, natural asymmetries in the local transmission success of serotypes, both in time and space, are shown to generate heterogeneous immunity landscapes that weaken the selective effects of cross-immunological reactions previously described for continuous, deterministic and homogeneously mixing frameworks. Complimentary to immune interaction, host demographic factors and vector ecologies thus emerge as important drivers of dengue's spatio-temporal dynamics.

4.2 Materials and Methods

In order to model the population dynamics of dengue, the agent-based model developed and described in Chapter 3 (on page 91) is expanded. It initially considers a directly transmitted pathogen consisting on four antigenic types, later being modified to include a mosquito-vector and host-population structure to explicitly address dengue epidemiology. A full parameter list with values, expected biological ranges and references can be found in tables 4.1 and 4.2 (on page 141) and the models are described bellow.

4.2.1 Individual-based Models

The direct-transmission model is based on the SIR epidemiological framework of Chapter 3 with no population structure. For simplicity, biological asymmetries between the co-circulating strains as well as co-infection are assumed to be negligible. Recovery from infection results in strain-specific complete immunity. There is no incubation period, the transmission probability is $\beta = 0.1175$, the daily contact rate per host $c = 10$, the infectious period $1/\gamma = 3$ days, the average host life span $1/\nu = 50$ years, a daily external infection rate is $\mu = 1E - 5$, and host population size $N = 100000$.

For the dengue model, mosquitoes are added to the system and dengue-like eco-epidemiological parameter values used (see Table 4.2 on page 141). Only the sus-

ceptible, exposed and infectious states of the adult life-stage of the mosquitoes are considered. To account for seasonal variation in vector densities, an annually driven mosquito birth rate is used, with the maximum number of vectors per human in the range $M = [0.7 - 1.2]$. Mosquitoes have per-day biting rate of $a_v = 0.6$, and the probabilities in human-to-vector and vector-to-human transmission are considered to be equal ($\epsilon_{h \leftrightarrow v} = 0.5$). The intrinsic and extrinsic incubation periods are assumed to be $\delta_h = 2$ and $\delta_v = 6$ days, respectively. The resulting basic reproductive number is $R_0 \approx 4$ in the absence of seasonality and $R_0 \approx 6$ with seasonality included [Ferguson et al., 1999b]. Following infection, individuals retain full immunity against the infecting serotype but remain susceptible to all others. For simplicity, co-infection with two or more serotypes is not considered. Immunological interactions are considered in the forms of (i) temporary serotype-transcending immunity, where individuals are fully protected against all other serotypes for a period $1/\alpha_h$ (months); and (ii) antibody-dependent enhancement through an increase in susceptibility and infectivity of heterologous, secondary infections, ϕ_h.

With the introduction of the mosquito-vector, the basic reproductive ratio needs to accommodate relevant human-vector interactions for the transmission of the dengue virus. The resulting expression is a variation of the Ross-MacDonald model's expression for the human malaria parasite:

$$R_0 = \frac{M\, a_v^2\, \epsilon_{v\to h}\epsilon_{h\to v}\, \delta_v\delta_h}{\nu_v(\gamma_h + \nu_h)(\delta_h + \nu_h)(\delta_v + \nu_v)} \tag{4.1}$$

where ν_v and ν_h are the life-expectancies resulting from the age-dependent mortality rate. This same approach as been used by Wearing and Rohani [Wearing and Rohani, 2006] in a deterministic model of dengue exploring ecological and immunological determinants of dengue epidemics.

4.2.2 Population Structure and Connectivity

As described in Chapter 3, host-population structure is added by subdividing hosts into a spatially organized set of communities of size L, organized into a squared and non-wrapping lattice with x rows/columns. Each community C_j has k_j neighbours, with the set $\{C_j, k_j\}$ defining a *locality*. For each host in community C_j, contacts take place randomly amongst individuals of the locality, that is, within C_j and its neighbouring communities k_j. Contacts outside the locality take place with probability ω. A detailed description can be found in the Methods section of Chapter 3 on page 93.

4.2.3 Host and Vector Mortality

Age-dependent mortality rates are used, as described in the methods of Chapter 3 in Section 3.2.3 on page 96.

4.2.4 Spatial Epidemic Synchrony

To measure the spatial epidemic synchrony of single serotypes across the population, the cross-correlation coefficient, or Pearson's correlation coefficient, is used. The coefficient is a measure of the correlation between two variables, X and Y, and is defined as the covariance divided by the product of their standard deviations:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\delta_X \delta_Y} \quad . \tag{4.2}$$

Its values are between -1 and 1, indicating perfect negative correlation or perfect positive correlation between the two variables, respectively. If zero, the variables are unrelated. Here, the incidence time series of communities are used as input variables with the expectation that values towards 1 signify increasing epidemic synchrony between communities.

4.2.5 Dengue Serotype Data

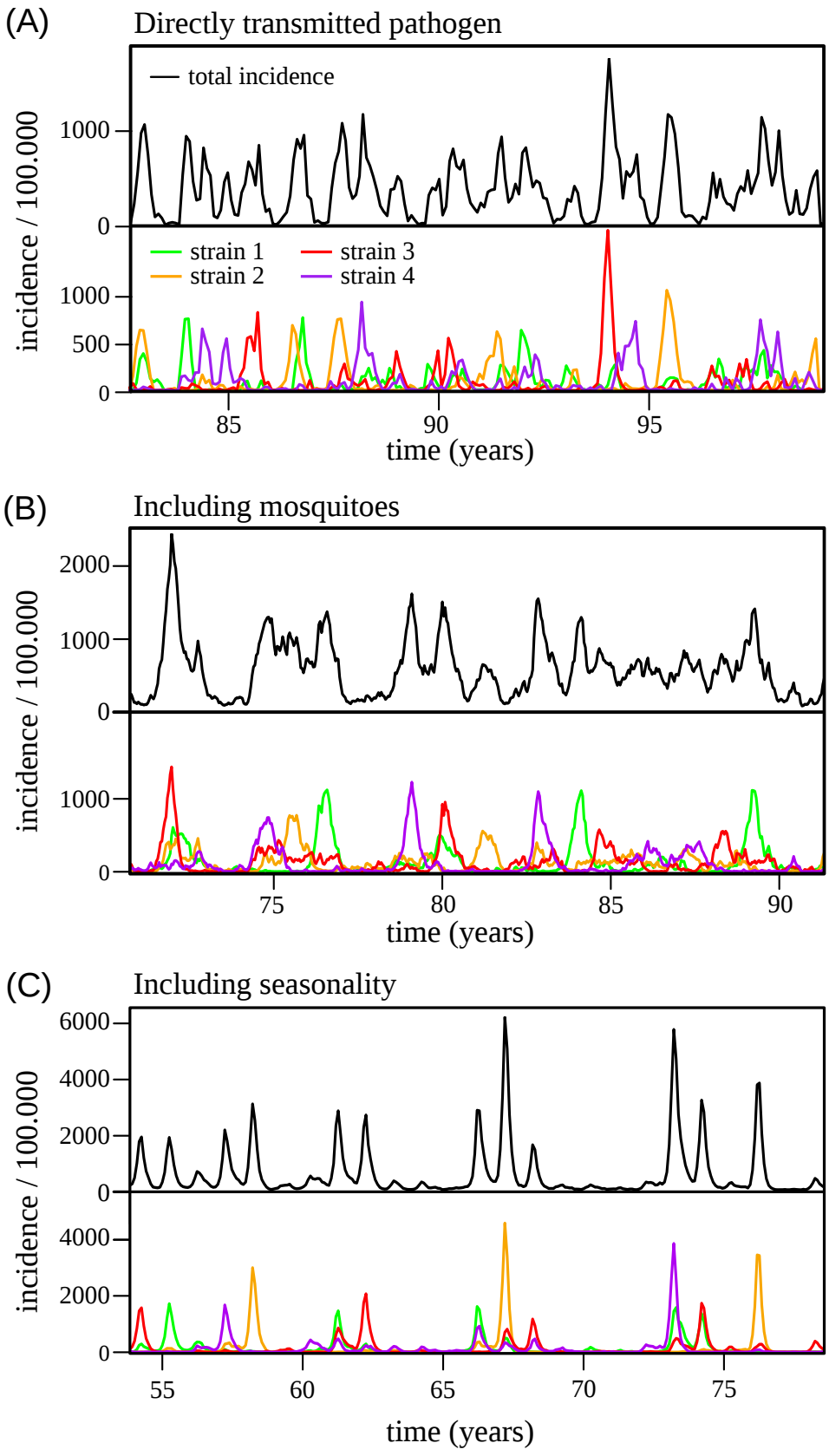
Spatial serotype data was provided by C. Simmons from the Oxford University Clinical Research Unit (OUCRU), Ho Chi Minh City (HCMC), Viet Nam, representing 290 dengue cases in children presenting to the outpatients department of three large hospitals in HCMC during the season 2010-2011; spatial information was based on residency of the patient. Data for Thailand during the 2005/2006 season was obtained from [Fried et al., 2010; Veeraseatakul et al., 2007].

4.3 Results

4.3.1 Non-Spatial Multi-Strain Dynamics

First, the dynamic behaviour of a directly transmitted pathogen, comprising four co-circulating strains under the assumption of homogeneous mixing, is analysed. Unless stated otherwise, immunological interactions between strains are restricted to the prevention of super-infection, i.e. cross-immunity is not considered to have an effect beyond the duration of infection.

Contrasting the predictions of deterministic multi-strain models, in which the dynamics inevitably converge towards a stable equilibrium in the absence of explicit competition, the system exhibits oscillations in the total number of infections and relative strain prevalence (figure 4.1 A on page 122). In agreement with the results of Chapter 3 as well as previously studied single-strain systems [Alonso et al., 2007; Verdasca et al., 2005], these dynamics are driven by the amplification of stochastic effects at the individual level, which keep each strain in a transient regime instead of approaching the expected deterministic equilibrium. At the same time, strain desynchronization is dictated by short-term differences in each strain's transmission success that accumulate in time, generating significant asymmetries in the immunity profile of the host population. Consequently, the epidemic periodicity of the individual strains is found to be in the same range of the one obtained for total incidence, i.e. for the pathogen


(A)
Directly transmitted pathogen
total incidence
strain 1
strain 2
strain 3
strain 4
incidence / 100.000
1000
0
1000
500
0
85
90
95
time (years)
(B)
Including mosquitoes
incidence / 100.000
2000
1000
0
1000
0
75
80
85
90
time (years)
(C)
Including seasonality
incidence / 100.000
6000
4000
2000
0
4000
2000
0
55
60
65
70
75
time (years)

(figure 4.2 on page 124).

To investigate the effect of stochastic amplifications on dengue's epidemiological dynamics and inter-epidemic periods, a vector population is further incorporated into the model, using dengue-relevant eco-epidemiological parameter values (Table 4.2 on page 141). The resulting dynamics show a significant increase in average disease prevalence (figure 4.1 B on page 122), which is driven by a slight increase in R_0 ($R_0 \approx 3.5$ vs. $R_0 \approx 4$ in the vector model) and, more importantly, a reduction in the risk of stochastic extinction due to the inclusion of viral incubation periods.

Further including seasonality, through annual variations in mosquito densities, results in dengue-like epidemiological dynamics with a distinct seasonal signature, strong multi-annual periodicities in incidence and fluctuating distribution in serotype prevalence (figure 4.1 C on page 122). Again, due to variation in R_0 (now ≈ 6), this is accompanied by a considerable increase in peak incidence levels and more pronounced epidemic troughs.

Figure 4.1 *(preceding page)*: **Non-Spatial Multi-Strain Dynamics - (A)** The simulated time series of an agent-based, multi-strain model of a directly transmitted pathogen show irregular oscillations in total (black lines, top graphs) and strain-specific (coloured lines, bottom graphs) incidences even in the absence of immunological interactions or asymmetries between pathogen strains ($R_0 \approx 3.5$). **(B)** Changing the system to describe a vector-transmitted pathogen and including the associated intrinsic and extrinsic incubation periods results in an overall increase in mean incidence and decrease in the risk of stochastic extinction ($R_0 \approx 4$). **(C)** Further including seasonal variations in mosquito densities results in multi-annual epidemic outbreaks followed by severe transmission bottlenecks ($R_0 \approx 6$).

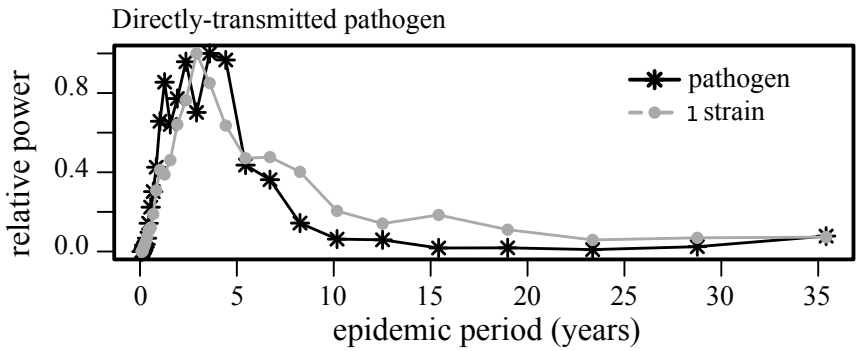


Figure 4.2: **Non-Spatial, Directly Transmitted Multi-Strain Epidemic Periodicities** - The global wavelet spectrum (GWS) is obtained by averaging the local wavelet power spectrum (LWPS) across time and is analogous to a traditional Fourier spectrum. For a directly transmitted homogeneous mixing framework, each strain's periodicity is dominated by a single frequency that overlaps a wider range of periods presented by the pathogen.

4.3.2 Spatial Dengue Model

Disease Dynamics

The occurrence of large epidemic outbreaks demonstrates the synchronising effects of vector ecology on the persistence and epidemiological patterns of vector-born pathogens within an homogeneous mixing setting. To examine the effect of host spatial segregation, the framework is restructured into a meta-population formulation by subdividing the human and mosquito hosts into sets of spatially arranged communities (see Section 4.2 on page 117). Within this set-up, individuals are assumed to get infected predominantly by mosquitoes of their own and surrounding communities, with long distance contacts, due to host movement, taking place with a low and distance independent probability, ω. These assumptions are in accordance with previous studies showing that human movement, in contrast to vector dispersal, is a stronger determi-

nant of dengue's epidemiology [Adams and Kapan, 2009], backed by reports of rapid geographical spread of specific viral lineages [Rabaa et al., 2010] despite the limited flight range of *aedes* mosquitoes [Harrington et al., 2005; Maciel-De-Freitas et al., 2007].

With the addition of this spatial setting, more defined seasonal dynamics are observed, as well as a lower variability in the epidemic behaviour (figure 4.3 B on page 126). In general, the incidence dynamics are qualitatively analogous to empirical time series (figure 4.3 A). The periodicity in serotype prevalence also increases and settles onto a 8-9 year cycle (figure 4.3 C), in line with that observed in dengue endemic countries [Nisalak et al., 2003] and dengue's phylodynamics in Thailand Adams et al. [2006].

In agreement with earlier work and the results of Chapter 3, spatial segregation between hosts and demographic stochasticity are found to enhance global persistence but at the same time facilitate local extinction [Earn et al., 2000a; Keeling, 2005; Verdasca et al., 2005]. This creates spatially heterogeneous immunity landscapes, upon which individual serotypes can be sequentially selected and amplified. Consequently, the population-wide serotype prevalence is highly heterogeneous and frequently exhibits locally propagating waves (figures 4.4 A, B on page 127), which can offer a parsimonious explanation for the geographical differences observed in real serotype distributions (figures 4.4 C, D).

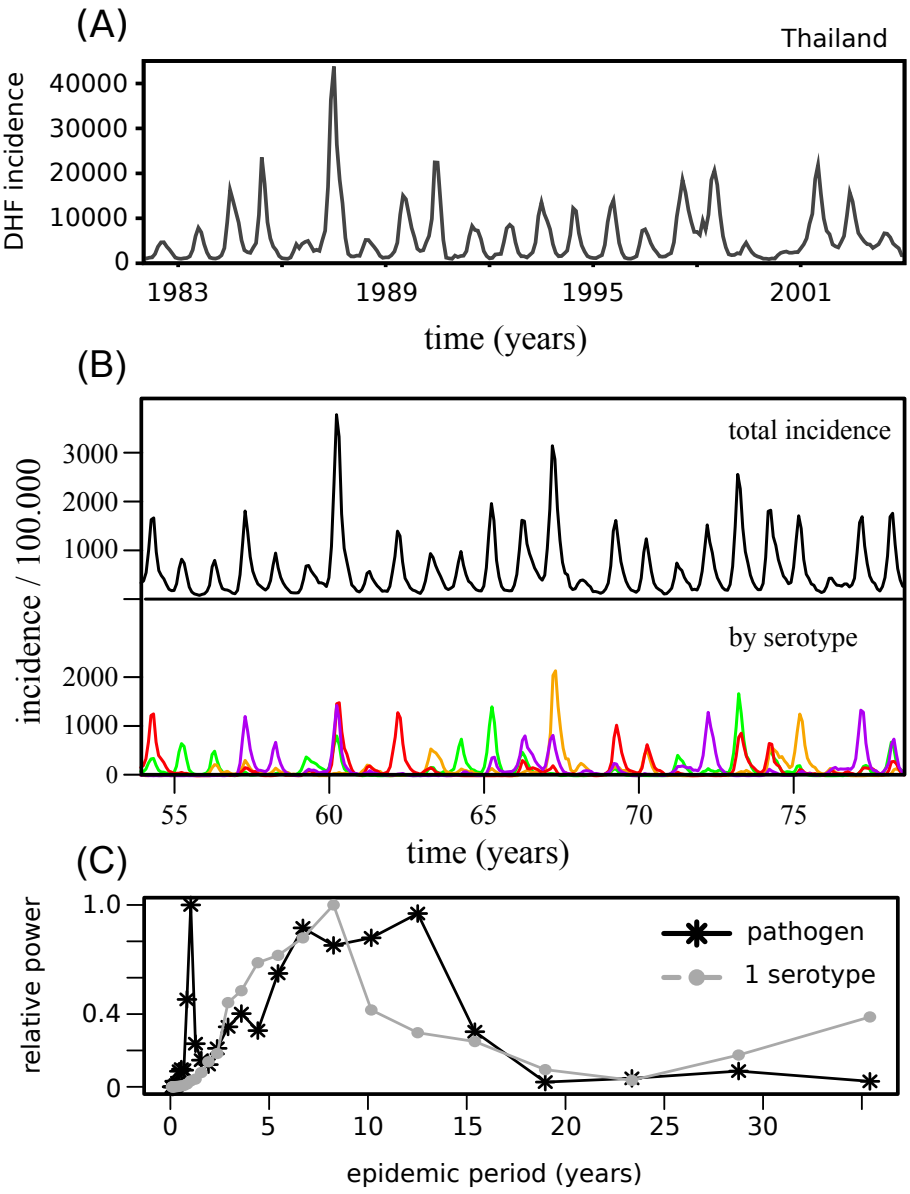


Figure 4.3: **Dengue Epidemiological Dynamics With Population Structure** - **(A)** Monthly DHF incidence reveals a strong seasonal component, which is not reflected in serotype dynamics. Description of the data can be found in Section 1.5 on page 58. **(B)** Structuring the host population into a (20 by 20) lattice of smaller sub-communities results in lower epidemic variability in the simulated, epidemiological time series and higher out-of-season viral persistence. The incidence dynamics are qualitatively analogous to empirical time series. **(C)** The global wavelet spectrum (GWS) is obtained by averaging the local wavelet power spectrum (LWPS) across time and is analogous to a traditional Fourier spectrum. The strongest period in total incidence, i.e. for the pathogen, is determined by the annual variation in vector densities, whereas the serotype cycling settles close to a 8-9 year periodicity, similar to what is observed in dengue-endemic regions [Nisalak et al., 2003].

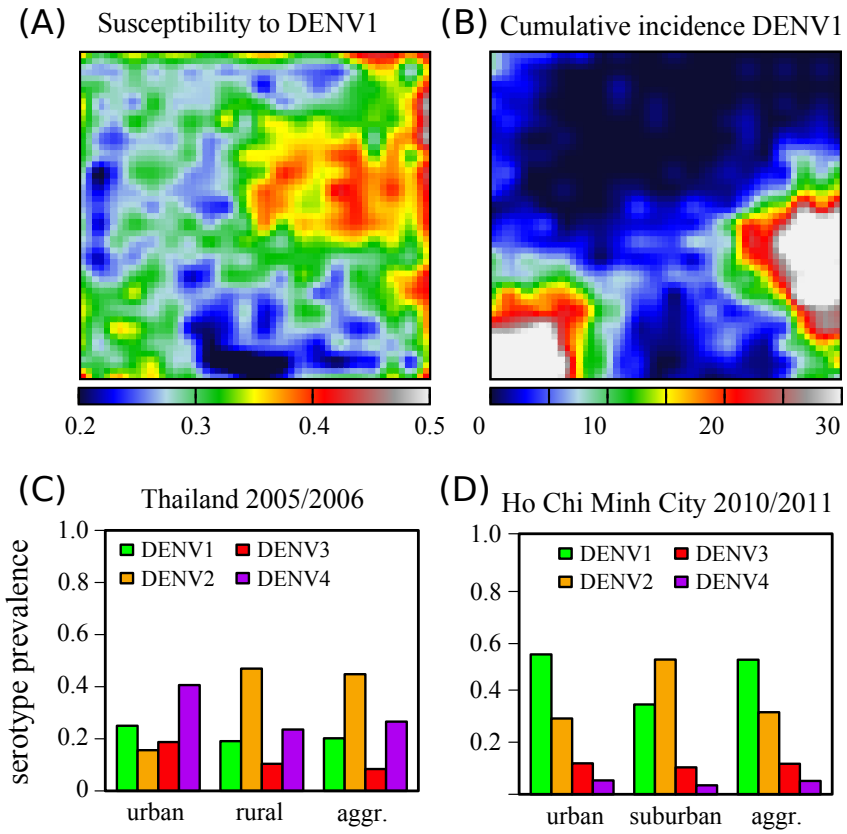


Figure 4.4: **Dengue Spatial Dynamics With Population Structure** - **(A)** Local viral extinction generates a highly heterogeneous immunity landscape, shown as a snapshot of the population-wide susceptibility level to DENV1 (year=80 of the time series in figure 4.3 on page 126). **(B)** The spatial prevalence of individual serotypes is equally heterogeneous and driven by serotype-specific susceptibility, as demonstrated by the cumulative incidence of DENV1 for the following 3 seasons. **(C,D)** Significant differences in serotype prevalence within multiple geographical scales during a single season in data from both Thailand and Ho Chi Minh City. These would be hidden by just considering aggregated data.

Spatial Synchrony

Figure 4.4 B (on page 127) further suggests the presence of localised spatial coherence in the epidemiological dynamics of close-ranged communities. To address this, the spatial synchrony in the incidence of one serotype is measured across the diagonal of the meta-population using the Pearson's r correlation coefficient (See Methods on page

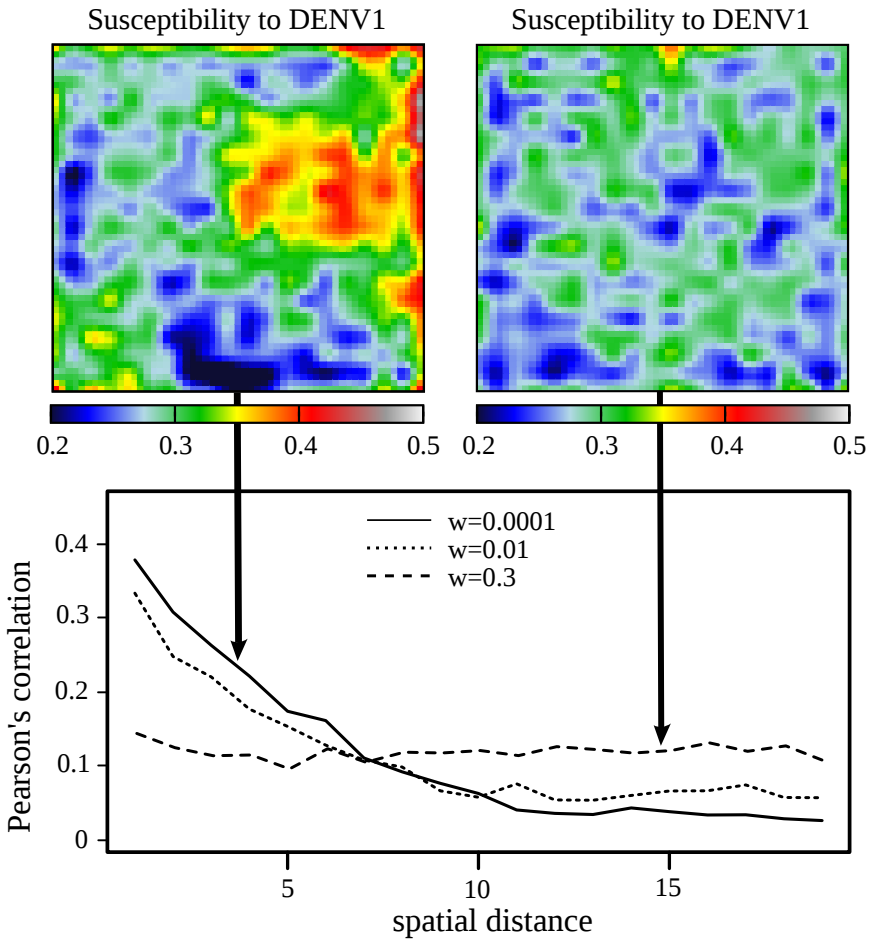


Figure 4.5: **The Effect of Mobility in Serotype Spatial Coherence and Immunity Landscape** - Spatial synchrony in serotype prevalence (DENV1) decreases with increasing distance between communities (Pearson's r between communities along the diagonal as exemplified in figure 4.6 on page 130). Low host mobility, $\omega = 0.0001$, forces transmission to be local and results in a highly heterogeneous susceptibility landscape (top left, reproduction of figure 4.4 A on page 127). In contrast, high host mobility, $\omega = 0.3$, has an homogenizing effect, which also results in a less variable, fine-grained immunity landscape of the population (top right).

120). As highlighted by the results of Chapter 3, the rate at which hosts acquire infections in geographically distant communities has a significant effect on the immunity landscape of the population. This is illustrated in figure 4.5 (on page 128) for varying values of mobility, which shows a transition to less variable and fine-grained distributions of susceptibility levels to DENV1 with an increasing rate of long-distance

transmission events. As a result of a reduction in locally acquired infection with increasing mobility ($\omega = 0.3$), epidemic synchrony between neighbouring communities is disrupted, causing an overall low but homogeneous spatial coherence across the population. In contrast, if disease transmission is predominantly local ($\omega = 0.0001$), a significant reduction in spatial synchrony is observed with geographic distance between communities (an example of desynchronized local outbreaks is presented in figure 4.6 on page 130). This corroborates recent studies showing that dengue cases cluster at the level of households or neighbourhoods [Salje et al., 2012], whereas epidemics across larger regions might follow a power-law distribution [Massad et al., 2008] and/or wave-like behaviour [Cummings et al., 2004; Thai et al., 2010], which implies that outbreaks are predominantly driven by a limited set of spatial clusters in time.

Contrasting with these results, migration (or dispersal), has previously been shown to increase synchronization between local populations [Earn et al., 2000a; Hastings, 2001; Jansen, 1999]. Although this might seem counterintuitive at first, it can be explained by the nature of long-distance transmission considered in this model. That is, compared to a continuous migratory term often considered in previous studies [Earn et al., 2000a; Keeling, 2000], the acquisition of an infection from a host in a distant community is here considered to be of probabilistic nature. Hence, increasing mobility does not necessarily mean that introductions from distant communities will result in sustained local transmission. As a consequence, high mobility results in less vari-

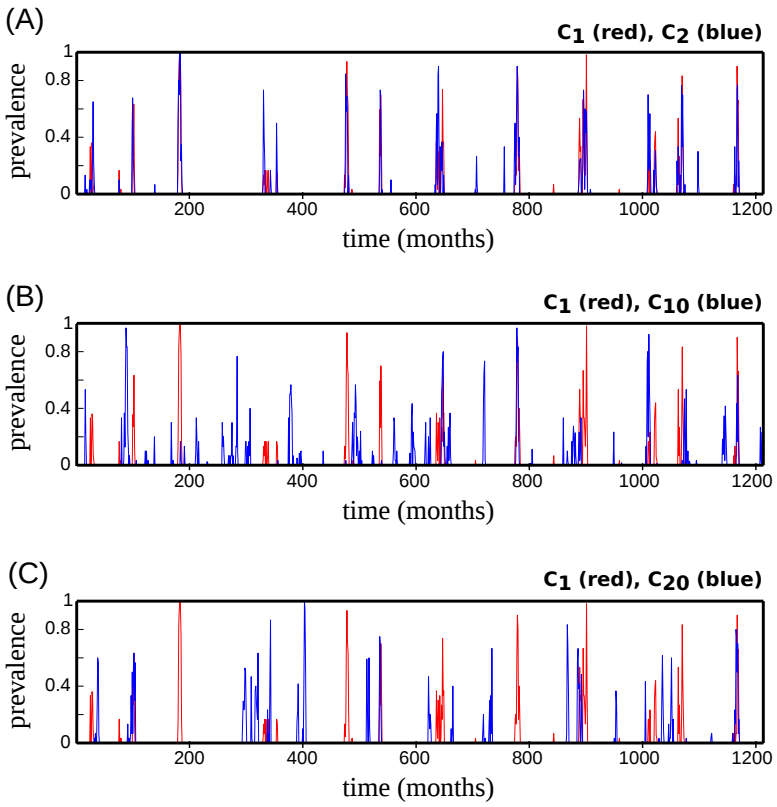


Figure 4.6: **Example of Serotype Spatial Synchrony Across the Meta-Population** - Time series of DENV1 in different communities over a 100 years period. Synchrony along the diagonal **(A-C)** decreases with distance to community 1 (C_1, bottom left corner of meta-population) (this is an example for $\omega = 0.0001$ as described for the results of figure 4.5 on page 128). C_2, C_{10} and C_{20} stand for communities 2, 10 and 20 across the diagonal.

able immunity landscapes that also present patterns indicative of the occurrence of frequent short outbreaks (top right, figure 4.5 on page 128). Importantly, the desynchronizing effects of considering demographic stochasticity in dispersal rates have also been found in other, non agent-based modelling frameworks [Simonis, 2012].

The Effects of Host Demographics

The influence of host demographic factors is investigated in more detail by independently varying host-population structure and mobility. Increasing spatial structuring

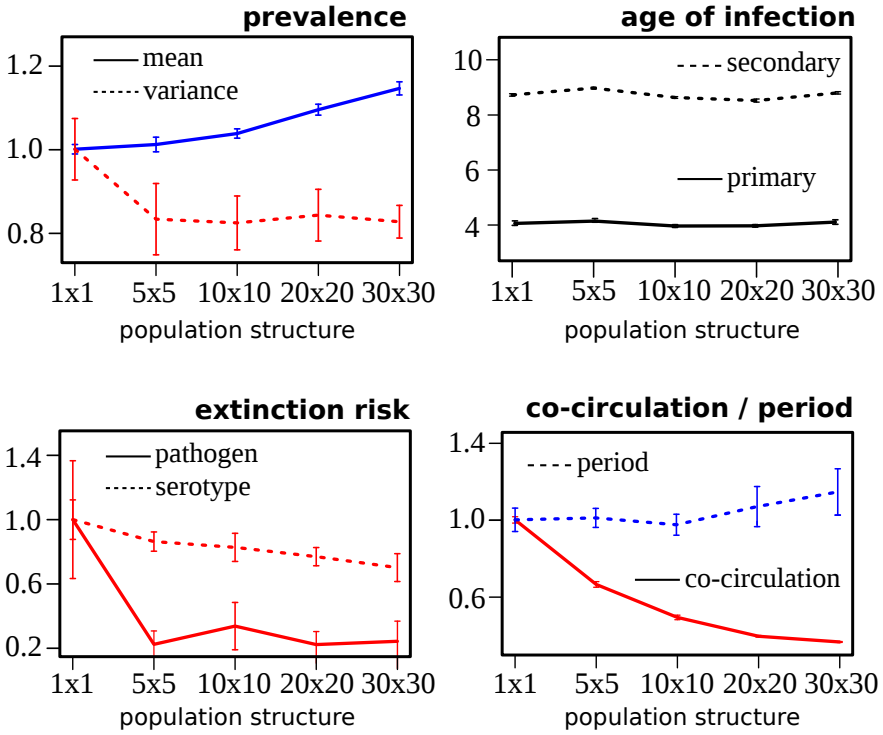


Figure 4.7: **The Effects of Host Demographics - Population Structure** - Increasing structure results in a significant reduction in epidemic variability **(top left)**, extinction risk **(bottom left)**, longer serotype periodicities **(bottom right)** and serotype co-circulation **(bottom right)**. Increases in persistence also cause higher mean incidence rates **(top left)**. Age of primary or subsequent infections is not affected by structuring as the force of infection remains constant **(top right)**. For ease of comparison, except for age of infection, all results are normalised to the values derived from a homogeneous mixing model, with ratios above 1 representing increases (blue) and below 1 decreases (red). Extinction risk is defined as the amount of time the pathogen or individual serotype remains bellow a critical threshold of 10 infected hosts (human or mosquito). Co-circulation is defined as the average percent time where > 1 serotypes are present in a given community. Points are the mean and deviation for 25 stochastic simulations.

and thereby decreasing the size of each community, forcing transmission to become localized, results in a decrease in serotype co-circulation and variability in total annual outbreak size (figure 4.7 on page 131). The latter is a consequence of a reduced propensity for large-scale, population-encompassing outbreaks due to restricted access to the susceptible pool (the opposite mechanism caused long periods of stochastic fade-out in the homogeneous framework as illustrated in figure 4.1 C on page 122).

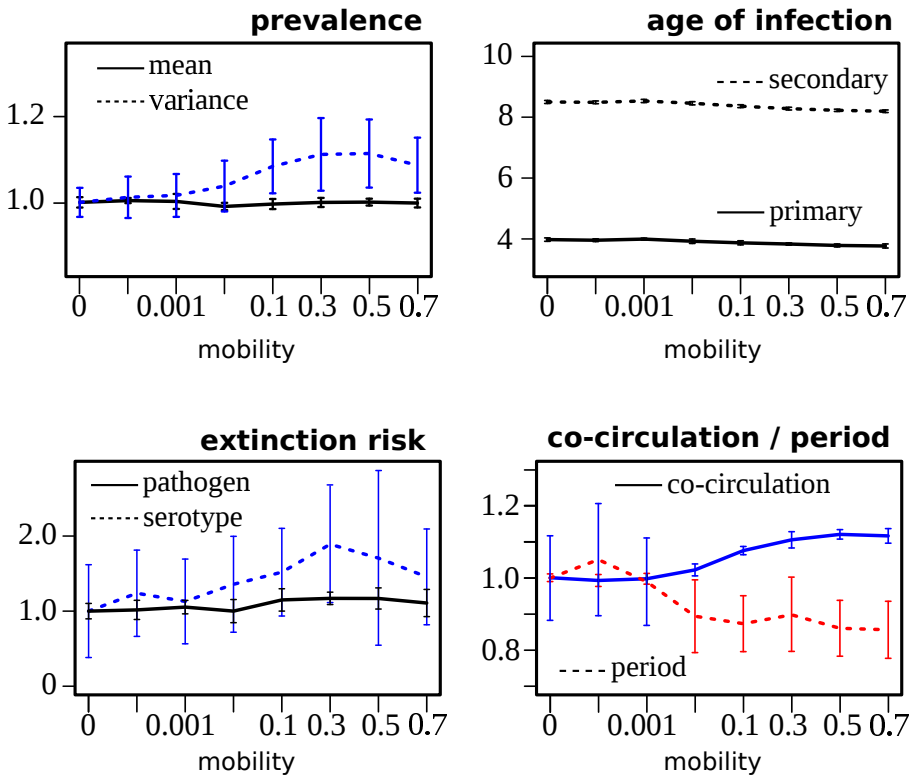


Figure 4.8: **The Effects of Host Demographics - Host Mobility** - For details on the y-axis of the plots see legend of figure 4.7 on page 131. Mobility, ω, counteracts the effects of population structure (here a 20x20 lattice). The increase in epidemic variability does not coincide with higher mean incidence rates **(top left)**, but is due to more frequent epidemic sparks (see immunity landscape, right, of figure 4.5 on page 128 for an example with $\omega = 0.3$). As mobility increases, serotype co-circulation also increases and their epidemic periodicity is shortened **(bottom-right)**. The average age of infection is not affected as the force of infection remains constant **(top right)**.

Although infection prevalence increases as a result of a reduced risk of stochastic extinction, the force of infection remains unchanged, as seen by the average age of primary and secondary infection, which rests constant as structure is increased.

In contrast to population structuring, increasing the probability of transmission between distant communities, as a proxy for daily host mobility, has a homogenizing effect (figure 4.8 on page 132). That is, it increases viral co-circulation and the fre-

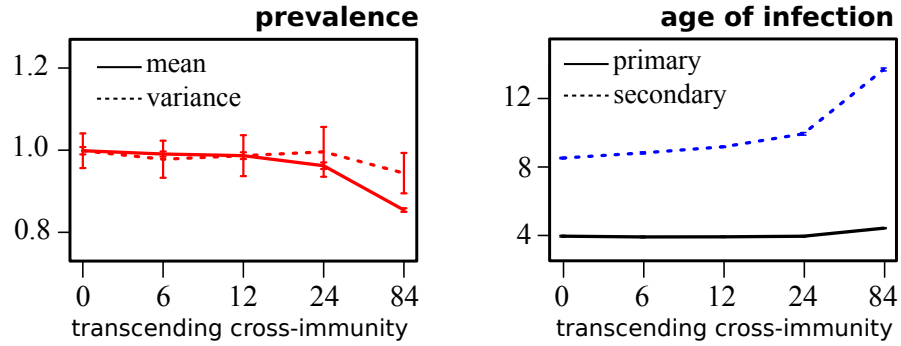


Figure 4.9: **The Effects of Immunological Interactions - Transcending Immunity** - For ease of comparison, the mean and variance of prevalence are normalised to the case of no immunological cross-reactions ($x = 0$), with ratios above 1 representing increases (blue) and below 1 decreases (red). The effects of temporary cross-immunity, $1/\alpha_h$, on mean prevalence, epidemic variability and average age of infection, only become apparent for periods beyond 12-24 months. Longer periods hamper variant transmission, decreasing mean incidence rates and significantly increasing the age of heterologous infection.

quency of brief localized outbreaks. As explained above, this results in a less variable but finely grained immunity landscapes (figure 4.5 top right, on page 128). The consequent increase in epidemic variability does not equate to a similar change in mean prevalence levels, however, highlighting again the importance of localized herd-immunity and local stochastic extinction [John and Samuel, 2000]. That is, the heterogeneous distribution of immunity within the spatially structured population (figure 4.4 A on page 127) severely restricts each serotype's potential pool of susceptibles and thus counteracts the occurrence of population-wide outbreaks that are otherwise expected from the synchronising effect of higher mobility [Earn et al., 2000a; Keeling, 2000].

The Effect of Immunological Interactions

In the previous sections, demographic stochasticity is shown to be sufficient for the emergence of dengue-like epidemiological dynamics (figure 4.3 on page 126), but its influence on the previously proposed effects of cross-immunological interactions is not addressed. Here, transcending immunity and enhancement due to ADE are considered and their impact on incidence dynamics and age of infection is analysed. Transcending immunity is assumed in the form of a short period of protection to all viruses after recovery, as reported by Sabin et al [Sabin, 1952], while enhancement due to ADE is assumed to take effect during secondary, heterologous infections and to result in increased transmissibility and susceptibility of the host.

Contrary to previous modelling predictions, temporary cross-immunity is not found to have a strong effect on the simulated incidence and age of infection. Only once the period is considered to be longer than 12 months does it have an effect, by decreasing incidence levels and epidemic variation, as well as reducing the probability of secondary infection (figure 4.9 on page 133). From the effects of host demographics (population structure and mobility), described in the previous section (figures 4.7 and 4.8 on pages 131, 132), it is clear that the time required for individuals to acquire a secondary, heterologous infection (here referred to as *heterologous exposure period*, or HEP) is in the order of $> 4 years$, even with the relatively high basic reproductive number of ≈ 6 assumed throughout the simulations. Crucially, this is much

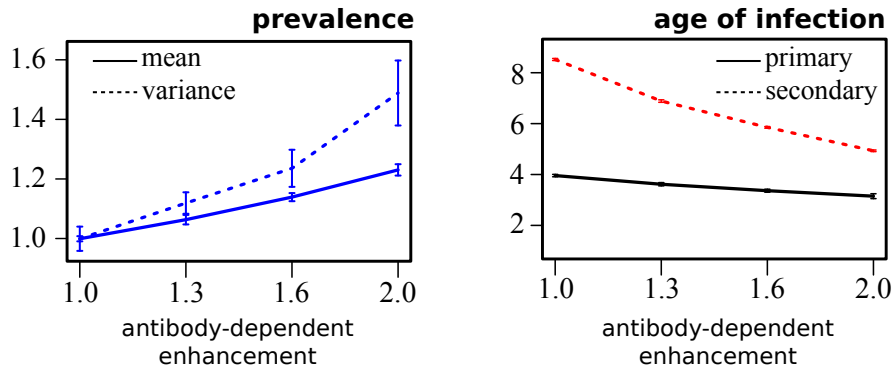


Figure 4.10: **The Effects of Immunological Interactions - Antibody-Dependent Enhancement** - For ease of comparison, the mean and variance of prevalence are normalised to the case of no immunological cross-reactions ($x = 0$), with ratios above 1 representing increases (blue) and below 1 decreases (red). Antibody-dependent enhancement, ϕ_h, simultaneously increases susceptibility to and transmissibility of secondary, heterologous infections. It causes an overall increase in the force of infection and more variable epidemic behaviour. Due to higher susceptibility and co-circulation, it also leads to a drop in age of primary, and particularly secondary infection.

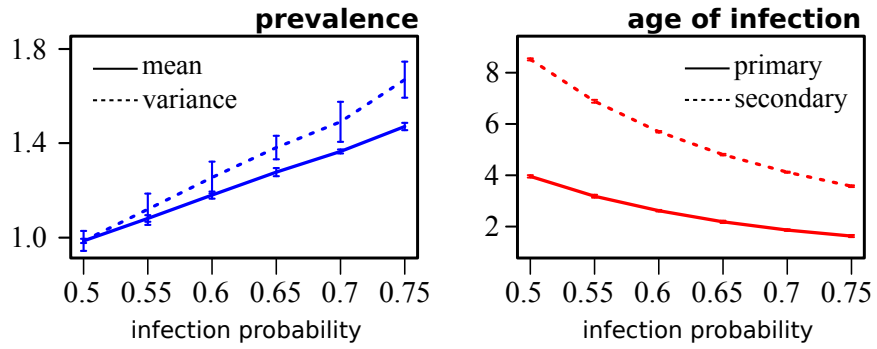


Figure 4.11: **The Effects of Infection Probability** - For ease of comparison, except for age of infection, all results are normalised to the case of no immunological cross-reactions, with ratios above 1 representing increases (blue) and below 1 decreases (red). Increasing overall transmission by increasing the per-bite probability of infection, $\epsilon_{h \leftrightarrow v}$, leads to higher mean and variance in prevalence but, in contrast to ADE, affects the ages of primary and subsequent infections equally.

higher than the reported 2-9 months of serotype-transcending immunity following a primary infection [Sabin, 1952]. Hence, within this discrete, probabilistic and spatial framework, it emerges that individual hosts are infrequently challenged by multiple serotypes during the proposed period.

Immune enhancement, on the other hand, here considered to affect both susceptibility to, and infectivity of, secondary infections through the process of ADE, has a more noticeable and anticipatory effect in increasing both disease prevalence and epidemic variability (figure 4.10 on page 135), which is in agreement with the results from ODE-based approaches (Section 1.5 on page 37). Increasing levels of ADE results only in small decreases on the average age at which individuals experience their first infection, whereas the HEP shows a dramatic reduction. The decreases are due to the combined effect of elevated serotype co-circulation and increase in the per-bite probability of infection during secondary, heterologous infections. When compared to the effects of increasing the probability of transmission per mosquito bite (figure 4.10 on page 135), the impact of ADE on the age of first infection is noticeably weaker, which can be explained by its dependence on secondary infections only.

4.4 Discussion

Stochastic effects have previously been suggested to be an important driver for population oscillations in epidemiological systems of single-strain pathogens (previous Chapter and see also [Alonso et al., 2007; Bauch and Earn, 2003; Earn et al., 2000b]). This Chapter has advanced upon those findings by studying the effect of demographic stochasticity on the dynamical behaviour of a multi-strain pathogen within a spatially explicit framework with particular focus on dengue.

Adding to the results of Chapter 3, stochasticity was found to be a parsimonious driver of oscillatory behaviour in strain prevalence, which is commonly found in epidemiological time series of antigenically diverse pathogens. This self-emergent phenomenon was driven by the temporal amplification of stochastic effects generated by the underlying probabilistic nature of the interactions between the hosts and pathogen variants. In the context of dengue, protective and infection-enhancing effects of cross-reacting antibodies and T-cell responses to dengue viruses have been well documented both *in vivo* and *in vitro* [Dejnirattisai et al., 2010; Halstead and O'Rourke, 1977; Sabin, 1952]. Therefore, immunological interactions have long been hypothesized to drive dengue's complex epidemiological dynamics. However, given the problematic experimental models (see Chapter 1 on page 36) and the readiness of continuous modelling tools to replicate the dynamics under different biological assumptions (see Chapter 1 on page 37), their contributing effect is still unclear. Notably, dengue viruses are under strong purifying selection with no clear trend for continuous antigenic change [Weaver and Vasilakis, 2009], while ecological bottlenecks and clonal interference severely hamper the emergence of advantageous phenotypes (Chapter 2 on page 60). Both these factors strongly support that the cyclical replacement of dengue's four serotypes may not be driven by the same inter-strain, immuno-based selective forces that shape the phylodynamics of other antigenically rapidly evolving pathogens, such as influenza A, for example. In accordance with the results presented here, it instead argues for a critical role of demographic and ecological factors.

Important for this discussion are the findings of this Chapter with regards to the impact of demographic stochasticity and spatial structure on the previously proposed effects of protective and enhancing immunity to dengue viruses. For instance, ADE was found to induce a signature in the epidemiological age-profiles that is characterised by a longer period for first infection than the time required for heterologous exposure. This highlighted that, contrary to the expectations from homogeneously mixing approaches, the populational effects of ADE are not simply to increase the overall force of infection but are rather manifested asymmetrically between primary and secondary infections, reflecting a weakened effect in primary cases within structured host populations. At the same time, both the mean ages of primary and secondary infections in the model system appear to suggest a low probability for sequential challenge during the empirically proposed 2-9 months period of serotype-transcending immunity [Sabin, 1952]. In fact, in line with this framework's results, others have estimated the heterologous exposure period to be around 3-4 years [Gibbons et al., 2007].

While these findings do not question the pathological or clinical significance of immune interactions *per se*, they suggest that the strength of within-host serotype interactions, and therefore the consequences of acquired immunity, are unlikely to be the sole drivers of the complex epidemiological dynamics of dengue. Moreover, the discrepancy of these results to previous ODE-based modelling predictions, which focused primarily on reproducing the qualitative dynamics of incidence time series, fur-

ther highlight the need to consider multi-scale data integration in modelling systems. A starting point, pursued in this Chapter, is to consider well described spatial signatures, such as highly heterogeneous sero-conversion rates and serotype prevalence levels. Another promising direction is to simultaneously model the epidemiology and population genetics of the pathogen (see Chapter 5 on page 142). Similar approaches have significantly contributed to the understanding of the population dynamics of the influenza A virus, for example, by allowing to differentiate which, of the alternative immunological mechanisms that reproduce its population dynamics, also reproduce key phylogenetic signatures [Ferguson et al., 2003; Koelle et al., 2006a; Minayev and Ferguson, 2009].

In summary, these results have highlighted the importance of considering host spatial segregation and stochasticities in disease transmission to understand the epidemiology of dengue and other related pathogens. While previous studies have demonstrated how immune interactions may influence the population dynamics of multi-strain pathogens, the inclusion of host and vector ecologies adds to this understanding and provides complimentary hypotheses about the underlying causes for the oscillatory nature in incidence and serotype distributions that commonly characterize the complex epidemiologies of multi-strain host-pathogen systems.

4.5 Tables

Table 4.1: **Parameters used in the directly transmitted multi-strain framework.**

	definition	**value**
β	transmission rate	0.1175
c	contacts per host	10 per day
$1/\gamma$	infectious period	3 days
$1/\nu$	average life-span	50 years
μ	external infection rate	0.00001 per day
N	human host population size	100000

Values based on Alonso et al. [2007].

Table 4.2: **Parameters used in the vector-born multi-strain framework.**

	definition	**value**	**ref.**
$1/\delta_h$	intrinsic incubation period	2 days (2-7)	[1]
$1/\gamma_h$	human infectious period	4 days (4-12)	[2]
$1/\alpha_h$	period of temporary cross-immunity	2-12 months (2-12)	[3]
$\epsilon_{h\to v}$	human-to-vector transmission	0.5 per bite (0.33-1)	[4]
ϕ_h	ADE enhancement	1-2	- -
M	vectors per human host	0.7-1.2 (0.3-20)	[5]
a_v	mosquito biting rate	0.6 per day (0.33-1)	[6]
$1/\delta_v$	extrinsic incubation period	6 days (6-12)	[7]
$\epsilon_{v\to h}$	vector-to-human transmission	0.5 per bite (0.33-1)	[8]
θ	external infection rate	0.000005 per day	- -
N_h	human host population size	100000	- -
ω	daily mobility	0.0001 (0-1)	- -
a	Gompertz parameter (human)	0.0002	- -
b	Gompertz parameter (human)	10	- -
ν_h	human average life-span	67.5 years	- -
a	Gompertz parameter (vector)	0.003	- -
b	Gompertz parameter (vector)	10	- -
ν_v	vector average life-span	23 days (8-42)	[9]

References: [1] Halstead [2007]; [2] Gubler et al. [1981]; Vaughn et al. [2000]; [3] Sabin [1952]; [4] Armstrong and Rico-Hesse [2003]; Sabin [1952]; [5] Focks et al. [2000]; [6] Trpis and Hausermann [1986]; [7] Watts et al. [1987]; [8] Armstrong and Rico-Hesse [2003]; Sabin [1952]; [9] Trpis and Hausermann [1986]

CHAPTER 5

Modelling the Dengue Virus Phylodynamics

Abstract

As demonstrated in the previous chapters, demographic stochasticity may be a determinant of the population dynamics of dengue viruses. To further explore its role at the scale of viral genetics, the discrete multi-strain framework of Chapter 4 is here expanded to implicitly model the transmission and neutral molecular evolution of the virus. By comparing model output with *ex novo* empirical estimations from serotype genetic data, it is shown that non-selective factors are sufficient and parsimonious drivers of DENV's typical phylogenetic structure and qualitative dynamics of genetic diversity. Importantly, cross-immunological reactions are found to cause significant signatures on simulated phylogenies, which are incompatible with the empirical expectations. By integrating, for the first time, key multi-scale factors of DENV's population biology, these results reinforce the hypothesis that demographic and ecological factors are the most important drivers of the complex dynamic behaviour of dengue viruses.

5.1 Introduction

RNA-based pathogens, such as the influenza A and dengue viruses, present rapid evolution occurring at a similar time-scale as their host's ecological dynamics. Their genetic diversity is thus shaped in direct response to changes in host demographic fac-

tors, such as population size, segregation or movement, which provides measurable connections between population dynamics, viral evolution and transmission.

In recent years, a multitude of studies have utilized sequence data to investigate such connections, unifying key evolutionary and epidemiological processes into a single framework known as phylodynamics [Grenfell et al., 2004]. Understanding a pathogen's phylodynamic behaviour can provide relevant information on the origin of epidemic, resistant and virulent strains, as well as elucidate how persistence is achieved in the face of herd-immunity or demographic and climate change. These approaches have therefore the potential to predict transmission patterns and provide new opportunities for control.

The foundation of the phylodynamic framework is the analysis of correlations between genetic and epidemiological patterns in time and space. Pathogens under strong positive selection have had a great deal of attention in phylodynamic approaches, given their propensity to exhibit consistent and universal patterns, both at the genetic and epidemiological levels. Two prime examples are the influenza A virus [Bedford et al., 2011; Grenfell et al., 2004] and the immuno-deficiency virus 1 [Edwards et al., 2006; Gray et al., 2011], for which selection, at different ecological scales, promotes the continuous generation of antigenically novel strains that temporarily escape the host's immunity. As a consequence, phylogenetic trees present highly imbalanced (ladder-like) shapes, characterized by temporal clustering of similar variants with fast and

recurrent lineage turnover. In contrast, pathogens under weak positive selection, such as DENV or the measles virus (MEV), can present inconsistent but generally more balanced phylogenetic structures, with varying degrees of variant co-existence in time and space [Bedford et al., 2011; Grenfell et al., 2004].

As discussed in previous chapters, a multitude of alternative and potentially complementary mechanisms, based on cross-immunological reactions, can reproduce dengue's complex epidemiological dynamics within model systems. Importantly, as is the case for influenza, these imply that within-host processes can generate significant selective pressures at the level of the population, which could affect the molecular evolution of the viruses. However, long term genetic or phylogenetic signatures, linked to the proposed mechanisms, are yet to be demonstrated for dengue.

In line with other vector-borne flaviviruses, such as the yellow fever virus (YFV) and the japanese encephalitis virus (JEV), strong purifying selection dominates DENV's molecular evolution [Bennett et al., 2003; Holmes, 2003; Holmes and Twiddy, 2003; Weaver and Vasilakis, 2009; Zhang et al., 2005]. While this could predict generally constant intra-serotype population dynamics, genotype spatio-temporal patterns such as emergence and competitive exclusion, have nonetheless been recovered from data [Bennett et al., 2003; Holmes and Twiddy, 2003; Myat Thu et al., 2005; Weaver and Vasilakis, 2009; Zhang et al., 2005]. Interestingly, major phylodynamic studies on the global behaviour of DENV3 and DENV4 have reported negative spatial correlations

between genotypes [Araújo et al., 2009; Villabona-Arenas and Zanotto, 2011]. This reported rarity of genotype co-circulation, of which DENV2 appears to be an exception [Hang et al., 2010; Rico-Hesse et al., 1997], has been suggested to be the result of intra-serotypic cross-protective immunity that hampers the co-circulation of viral variants of the same serotype [Villabona-Arenas and Zanotto, 2011; Zanotto et al., 1996] . In contrast, within genotypes, the diversity of viral subgroups appears to be abundant with lineage co-circulation and intermittent turnover being a common phenomena [Bennett et al., 2010, 2003; Foster et al., 2003; Messer et al., 2003; Myat Thu et al., 2005; Sittisombut et al., 1997; Weaver and Vasilakis, 2009; Wittke, 2002].

While some reports have suggested the action of weak positive selection [Bennett et al., 2003; Twiddy et al., 2002], the majority seem to attribute turnover to stochastic effects. For example, a recent report that explored the phylodynamics of DENV4 from *in situ* evolution in Puerto Rico [Bennett et al., 2010], found recurrent losses in sequence diversity, highlighting population bottlenecks and arguing for an active and frequent role of genetic drift, meaning that the success of viral variants will not always reflect their fitness. By adding a spatial dimension into similar analysis for Viet Nam sequence data, others have demonstrated that the genetic diversity of different lineages can be maintained at unequal levels, even in the absence of differential fitness, pointing in the direction of demographic heterogeneities as a determinant factor [Rabaa et al., 2010; Raghwani et al., 2011]. This is in agreement with recent studies suggesting that throughout time, dengue outbreaks tend to be spatially limited [Cummings

et al., 2004; Massad et al., 2008; Salje et al., 2012; Thai et al., 2010].

These and other studies have contributed to clarify DENV's epidemiology from phylogenetic observations (and vice-versa), although a unifying perspective is yet to be developed. This Chapter presents a thorough description of the expected phylodynamic behaviour of DENV serotypes. For this, a meta-population of human and mosquito hosts is modelled as in Chapter 4, with explicit neutral molecular evolution of viral sequences. This discrete and multi-scale approach is used to analyse the ubiquitous phylodynamic patterns emerging from the interaction of key demographic, epidemiological and evolutionary processes that shape DENV's population biology.

5.2 Materials and Methods

5.2.1 Sequence Data

Sets of sequence data were obtained by querying Genbank, per viral species and geographical region, to represent *in situ* evolution as much as possible. Sequences were aligned with the Bio3d package for R [Grant et al., 2006]. DENV data sets were analysed per serotype. Summary information on the species and data can be found in tables 5.3 and 5.4 (on pages 186, 186).

5.2.2 Empirical Phylogenetics

Maximum likelihood phylogenetic reconstruction was performed in *R* with the *phangorn* package, using a Tamura 3-parameter model (TN93) of substitution [Tamura et al., 2011].

5.2.3 Epidemiological Model

The modelling framework is the same as the one analysed in Chapter 4 (on page 113) but with significantly larger host-population sizes of 3.5 million humans and, on average, $\approx$ 4 million mosquitoes. A burn-in period of 25 years is discarded for model output analysis. The basic reproductive ratio is kept at $\approx$ 6 in line with Chapter 4.

5.2.4 Evolutionary Model

Evolution is assumed to be neutral at all biological levels. Recombination is not considered given the reported rarity of such events for DENV (see Chapter 1 on page 14). For simplicity, a single consensus sequence is modelled per infected individual host.

The transmission chain from the epidemiological model is used to define the discrete dynamics of the viral population. In here, hosts are seen as viral demes undergoing a birth (transmission to host or mutation) and extinction (recovery or death of host) process. At evolutionary time $t = 0$, all viral demes are initialized with the same genetic sequence. With time, molecular evolution follows an infinite sites model, with

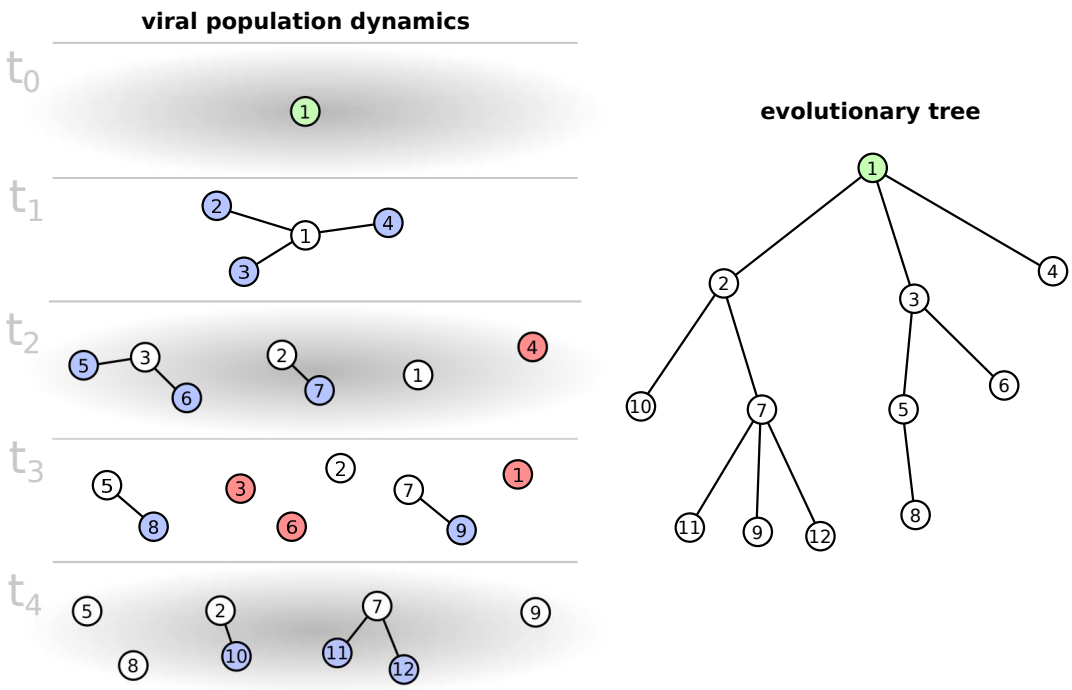


Figure 5.1: **Example of Simulated Viral Evolutionary Dynamics and Tree** - Short example on the simulated evolutionary process within the evolution model and the resulting evolutionary tree. **(left)** Mutational process for $t \in [0, 4]$ at the level of the population. At each time step, viruses can generate offspring (mutate) with a fixed probability creating new variants (blue virions) and may also get extinct (red virions). The probability of generating offspring is linearly dependent on the number of copies in the population, which depends on the number of hosts it infects. **(right)** The resulting multichotomous evolutionary tree, in which the branches have no scale.

changes occurring according to a fixed molecular clock [Gillespie, 2004; Page and Holmes, 1998]. As all viruses are initially the same, an evolutionary burn-in period of 10 years is considered before using any simulated sequences.

Since evolution is modelled forwards in time [Arenas, 2012], the final output is the complete, multichotomous evolution tree, in which: (i) the root is the initial sequence; (ii) all internal nodes represent ancestors of mutational events; (iii) tips represent either extinct or extant viral sequences (see an example in figure 5.1 on page 148). Individual

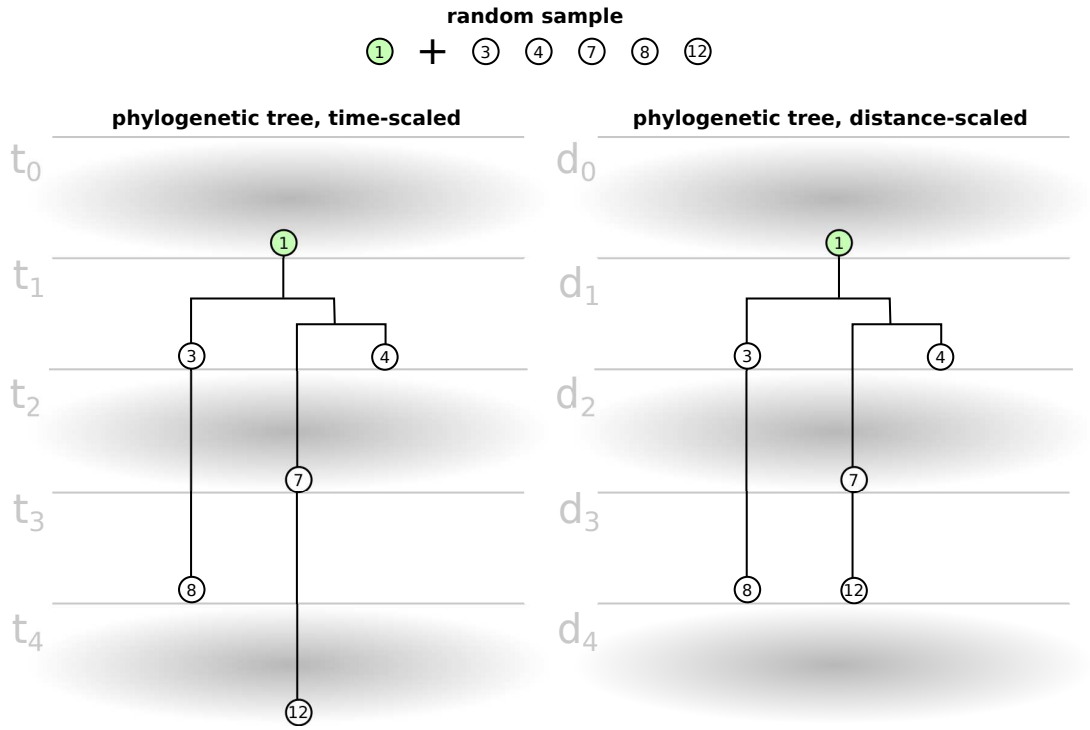


Figure 5.2: **Example of Simulated Phylogenetic Tree** - To obtain an example of a simulated phylogeny, the evolutionary tree in figure 5.1 (on page 148) is pruned by sampling 5 out of 12 possible viruses. The multichotomies are solved by adding randomly small branch lengths in series of dichotomies, keeping the total branch lenghts between viruses (nodes). Branch lenghts can be considered in time units **(left)** or genetic distance, i.e. number of mutations from the parent node **(right)**. Each node's depth can be different depending on the unit used (node 12) but the relationship between the nodes, and thus the tree's topological shape, is preserved.

nodes are tagged with meta-information, such as generation time and spatial location.

Inwards migratory events are assumed to not result in gene flow. That is, externally introduced incidence cases are adopted from within the meta-population, chosen randomly from the circulating variants. This allows to model and analyse a virtually isolated viral population.

5.2.5 Simulated Phylogenetics

The complete multichotomous evolution tree from the evolutionary model is pruned to obtain a representative phylogenetic tree (see an example in figure 5.2 on page 149). Multichotomies are solved by conversion into a series of dichotomies with potential branches of length zero, preserving the order in which nodes have been generated and their position in the appropriate scale (mutational distance or time). For the results, tree size is chosen to be within $\{10, 25, 50, 100, 150, 200\}$, to mirror the existing range in empirical data sets collected from Genbank.

5.2.6 Simulated Population Genetics

Within the evolutionary model, N_{seq} sequences are randomly sampled from the human population per time step, which are used to estimate: (i) sequence pairwise differences, Π, (ii) segregating sites, S, and (iii) the Tajima D statistic (using the Watterson estimator for segregating sites) [Gillespie, 2004; Page and Holmes, 1998; Tajima, 1989].

5.2.7 Empirical and Simulated Phylogenetic Shape

Tree shape is estimated using the Colless Index, a measure of topological imbalance independent of branch lengths and hence of the underlying molecular clock assumption [Colless, 1982]. The index considers the sizes B and S of the 2 groups of terminal nodes that descend from each of the $n-1$ internal nodes, with $B >= S$. Its quantifica-

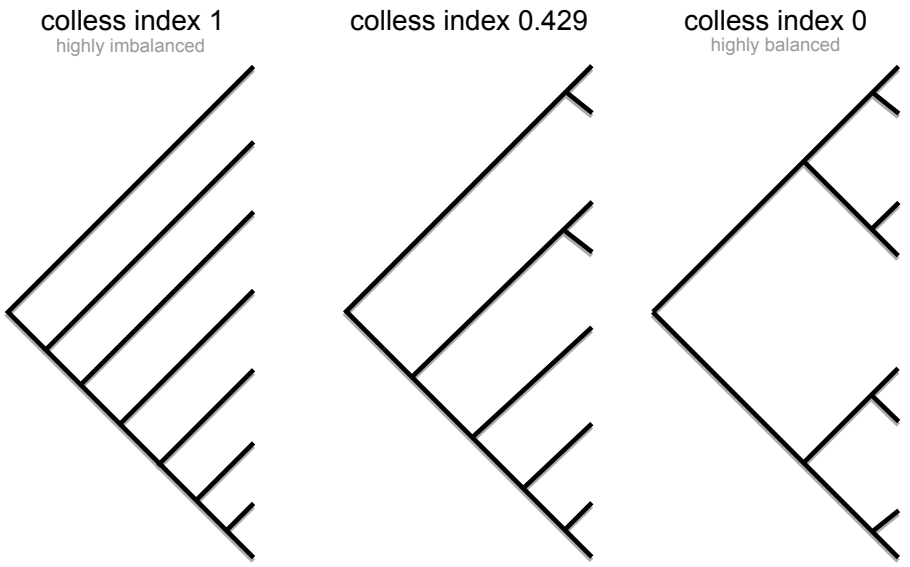


Figure 5.3: **Examples of Colless Index Measures** - Examples of phylogenetic trees and their Colless Index (see main text on page 151). The topologies are increasingly balanced from left to right (decreasing Colless Index).

tion is then based on the difference between B and S, summed over all $n-1$ internal nodes, for a total of n terminal nodes:

$$C = \frac{2}{(n(n-3)+2)} \sum_{i=1}^{n-1} (B_i - S_i) \ . \tag{5.1}$$

For a given species, the index converges to an expected value given a sufficiently large number of samples, and varies between 0 and 1 for maximum balance and imbalance, respectively [Rogers, 1994, 1996]. An example can be found in figure 5.3 on page 151.

To fit the index's convergence and estimate its expected value for a sample size 200 sequences, a weighted, two parameter (a, b) nonlinear least-squares estimation ($index \sim a * N^b$) is performed (weights being the sample size). The same method is applied for both the empirical and simulated trees.

5.2.8 Surrogates of Meta-population Dynamics

Four summary measures, here referred to as *dynamical surrogates*, were created to summarize the dynamic behaviour of the meta-population when varying model parameters. An example on how these measures are used to understand changes in the topological shape of the simulated phylogenetic trees is presented in figure 5.4 on page 153. The description of each dynamical surrogate can be found bellow.

Risk of Extinction

This surrogate proxies the risk of local stochastic extinction. It is a measure of the percentage of time a serotype spends circulating bellow a threshold, defined as 10 human individuals per 100.000. Hence, high percentage indicates high risk for local extinction and low percentage otherwise.

Variant Spatial Range

To address the average spatial dispersion of variants of the same serotype, each community is conceptually assigned to a district, from 16 possible (4x4) of equal size. The number of unique districts in which each variant is persistent is measured per

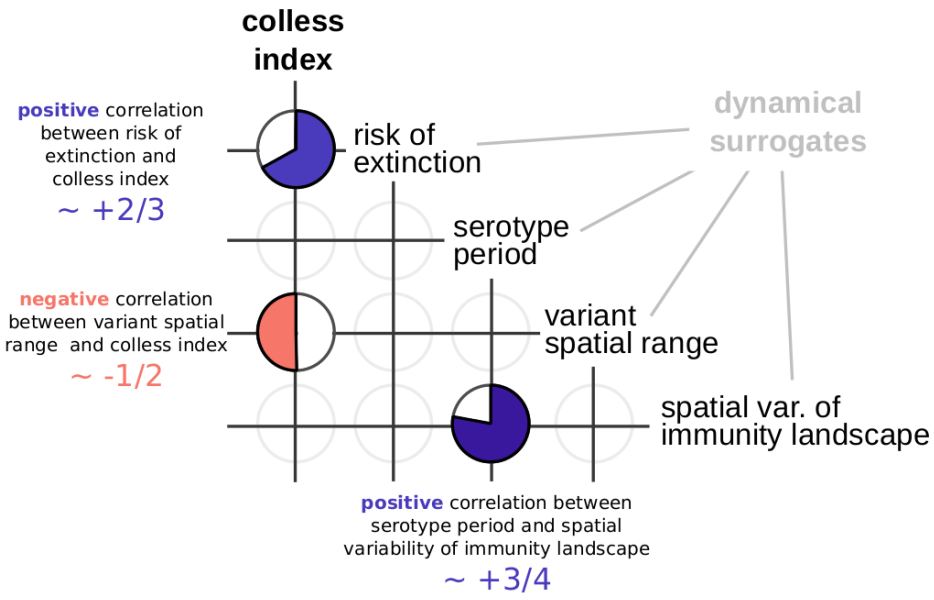


Figure 5.4: **Exemplifying the Use of Dynamical Surrogates** - Dynamical surrogates (risk of extinction, serotype period, variant spatial range and spatial variation of the immunity landscape) are calculated and averaged for sets of simulations when varying model parameters (see main text on page 152 for their description). The correlation of these measures is then addressed two-by-two, as well as with the measure of topological shape (Colless Index) of the resulting phylogenetic trees. This figure illustrates an example where only tree relationships are addressed (Colless Index with risk of extinction and variant spatial range, as well as serotype periodicity with spatial variation of the immunity landscape). The pie charts show the strenght and type of correlation, from 0 to 1, with blue for positive and red for negative correlations.

time step. The average for all variants is used to quantify this surrogate, which therefore proxies the degree of spatial dispersion.

Spatial Variability of Immunity Landscape

This surrogate proxies the spatial heterogeneity in the immunity landscape of a single serotype. It is calculated through the standard deviation of susceptibility across communities and thus mirrors to what degree transmission is localised or widespread. That is, if spatial variation is high, viral circulation is spatially restricted in time. In

contrast, if variation is low, spatial restrictions for transmission are expected to be low.

Serotype Period

This surrogate is simply the dominant frequency in serotype dynamics. It is relevant only within the context of other dynamical surrogates which summarize spatial information. This is because the frequency with which a serotype oscillates may have ambiguous meanings. For instance, longer periods can reflect two scenarios: slow consumption of the susceptible pool causing flatter oscillations in serotype prevalence, or, fast consumption of the susceptible pool which causes sharp prevalence oscillations with long periods of waiting for population turn-over. However, the speed at which susceptibles are infected leaves signatures in the variation of the immunity landscape of the population, and together these surrogates are informative for the behaviour of serotypes in time and space.

5.3 Results

Validation of the simulated epidemiology, population genetics and phylogenetic shape is performed by comparing the model output to empirical expectations. While the first is based on literature reports and the results of Chapter 4, the expectations of phylogenetic shape and population genetics are estimated *ex novo* from Genbank genetic sequences. Initially, fourteen different viral species are considered, of which nine are included in the results and general discussion, while the rest are excluded on the basis

of insufficient information (table 5.3 on page 186).

5.3.1 Empirical Phylodynamics

Phylogenetic Shape Expectation

To address the expectation on phylogenetic shape, the Colless Index is used as a measure of topological imbalance on trees obtained from a maximum likelihood (ML) reconstruction (see Materials and Methods) [Colless, 1982]. A useful expectation is the index's value for trees generated under an equal-rates Markov model (ERM), which assumes equal branching (extinction and speciation) rates for all genetic variants. Within a single species, measures that deviate from the ERM's expectation indicate the existence of asymmetric biological pressures amongst the extant variants [Rogers, 1994, 1996]. Between species, divergence in measures of Colless Index are a reflection of multi-scale biological asymmetries, ranging from genetic constraints, to ecological and selective pressures [Grenfell et al., 2004; Kirkpatrick and Slatkin, 1993]. In figure 5.5 (on page 156), Colless Index measures for phylogenetic trees of varying sizes (open circles) for four viruses - rotavirus A (ROTVA), influenza A (H1N1), DENV1 and DENV2 - are exemplified. The estimated measures are fitted to derive an expected value of the index for trees of size 200 (see Materials and Methods for details on page 151). The figure further summarizes the resulting indexes for 9 viruses, including DENV serotypes. More fittings are illustrated in figure 5.6 (on page 158) and the abbreviations of the viruses' names can be found in table 5.3 on

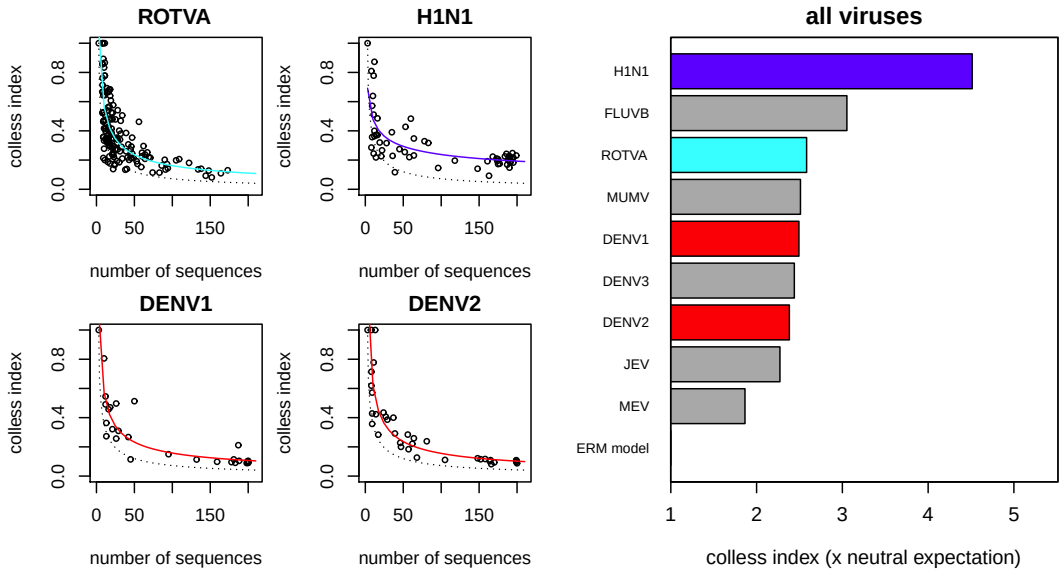


Figure 5.5: **Colless Index Expectation for Several Viruses** - Estimation of the Colless Index for empirical data sets of varying size (open circles). **(left)** Example of fittings for four viruses. **(right)** Summary of the estimated Colless Index for 9 viruses. The index is normalized to the expected value under the equal-rates Markov model (ERM, dotted lines in **left**). Multiple levels are estimated, with influenza A H1N1 viruses presenting the highest phylogenetic imbalance. For fittings of all viruses see figure 5.6 on page 158. Abbreviation of virus' names can be found in table 5.3 on page 186.

page 186.

As expected from reports of strong positive selection and highly imbalanced phylogenies [Ferguson et al., 2003], tree imbalance is found to be higher for the *Orthomyxovirus* family (influenza viruses). Within this group, H1N1 from the *Influenzavirus A* genus presents $\approx$ 4.5 times more imbalanced topologies than the neutral expectation. It is also notably higher than FluVB from the *Influenzavirus B* genus, which is in agreement with weaker signatures of selection amongst those viruses [Ferguson et al., 2003; Nobusawa and Sato, 2006]. On the other hand, the childhood pathogens MUMV and MEV, as well as the arboviruses DENV and JEV, present more balanced

topologies. The fact that species with significantly different life-cycles show similar topologies suggests shared genetic constraints, such as purifying selection or clonal interference [Grenfell et al., 2004; Miralles et al., 1999]. However, it does not easily explain the high deviation from the ERM's expectation observed for all viruses, which could be due to shared host demographic or ecological factors, for instance.

Population Genetics Expectation

To address the expectation on general signatures of the population genetics of DENV, the Tajima D statistic is calculated for sets of full genome sequences per region and per year (see Materials and Methods) [Tajima, 1989]. The statistic, $D = \Pi - S$, considers measures of genetic diversity based on the number of pairwise differences (Π) and number of segregating sites (S) in the sample. While the first takes into account the frequencies of variants, the second accounts simply for the number of polymorphic sites. Under a neutral model of evolution and for idealized populations of constant size, the expectation of this statistic is zero. Measurements of negative values in real populations are indicative of an excess of low frequency polymorphisms, while positive values are attributed to an excess of polymorphisms in intermediate frequencies. The Tajima D is therefore a valuable measure for the population dynamics of a pathogen's genetic variants.

The Tajima D's estimation for DENV1-4 genetic data is presented in figure 5.7 (on page 159). Two types of distribution are found. While DENV1, 3 and 4 demon-

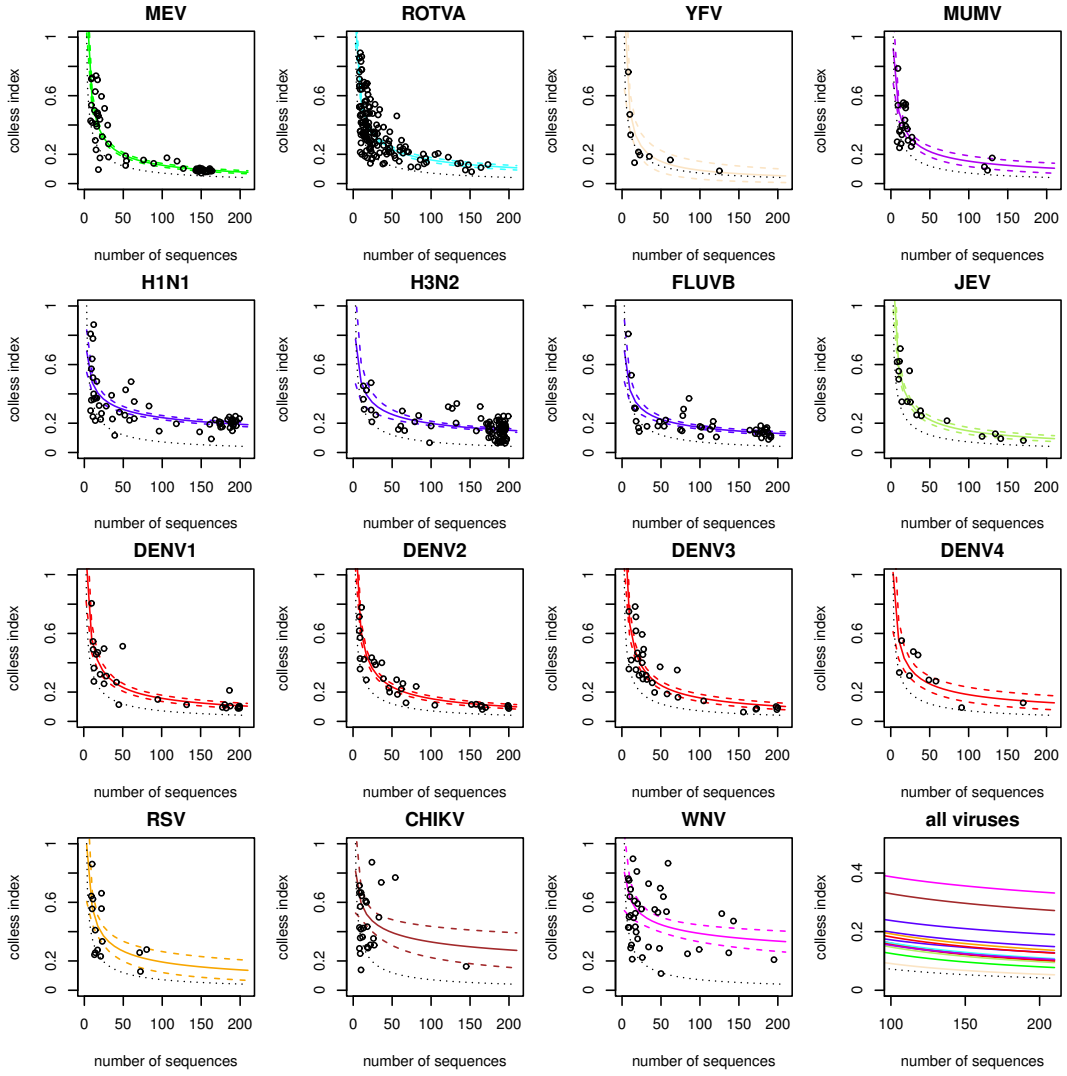


Figure 5.6: **Colless Index Expectation for All Viruses** - Fittings of the expected Colless Index for all data obtained from Genbank (see Materials and Methods). The data of 5 viruses presented unexpected or insufficient material for discussion and were therefore excluded: (i) CHIKV, (ii) RSV and (iii) YFV presented a lack of data sets with sufficient size; (iv) WNV did not present a clear convergence with increasing number of sequences, which may have resulted from the presence of numerous animal derived samples; (v) the estimation of DENV4 was above all other dengue serotypes, which appeared to be caused by a lack of sets with sufficient size. Abbreviation of virus' names can be found in table 5.3 on page 186. The expected value under the equal-rates Markov model (ERM) is represented by the dotted line.

strate a strong bias towards negative values, DENV2 presents a distinctive population genetics signature with a zero centred distribution. The fact that all distributions overlap in frequent negative values is indicative of shared evolutionary pressures amongst

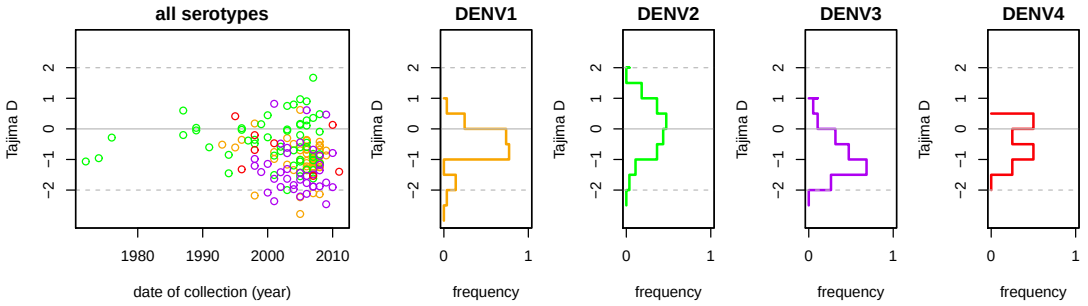


Figure 5.7: **Tajima D Expectation for Dengue Serotypes** - The neutral expectation of the Tajima D statistic is for zero centred distributions (solid grey lines). DENV1, 3 and 4 present negative biased distributions. Negative values are indicative of an excess of low frequency polymorphisms in the population, which may be caused by purifying selection, population growth after bottleneck and/or inwards gene flow. Positive values can be attributed to balancing selection, ongoing bottlenecks or positive selection, population segregation and/or inwards gene flow. DENV2 is often reported to be under weak positive selection and the only serotype presenting significant genotype co-circulation (see text, page 145), which may explain the balanced distribution. Data includes full genome sequences, per region, per year (see table 5.4 on page 186).

serotypes. Since viral persistence must take place in similar ecological contexts, it is possible that pathogen extrinsic factors (i.e. non-genetic) may play a determinant role. Alternatively, negative Tajima D values could be a consequence of purifying selection pressure acting on DENV's genome, which can hamper the emergence of variants and result in a viral population rich in low frequency polymorphisms.

As mentioned earlier in this Chapter, phylodynamic studies have highlighted the unique behaviour of DENV2 genotypes, which demonstrate frequent co-circulation [Hang et al., 2010; Rico-Hesse et al., 1997]. Moreover, in contrast to the other serotypes, genetic variation within DENV2 has long been implicated in episodes of weak positive selection [Cologna et al., 2005; Gubler et al., 1978; Rico-Hesse et al., 1997]. Both these facts will move the Tajima D to the positive spectrum, and may therefore explain

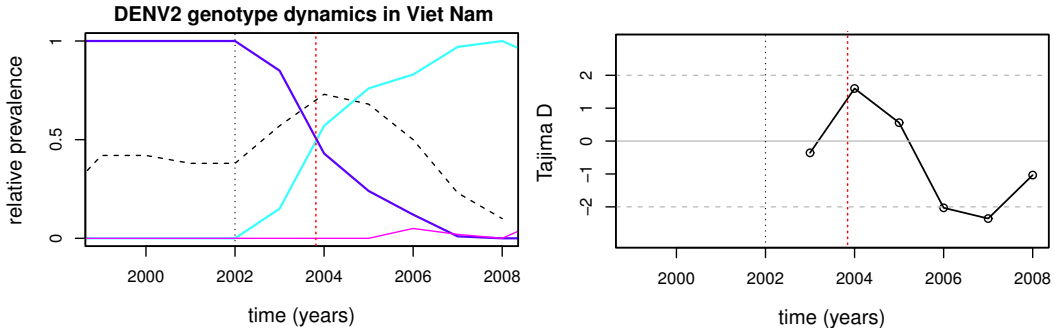


Figure 5.8: **Tajima D Dynamics During DENV2 Genotype Replacement Event in Southern Viet Nam - (left)** Population dynamics of serotype 2 in southern Viet Nam between 1999-2008 as described in chapter 2 on page 60. The invading Asian-1 genotype (cyan) is detected in 2002 (black dotted vertical line) and replaces the resident, Asian/American type (blue). The relative prevalence of DENV2 is represented by the black dashed line. The cosmopolitan genotype (magenta) appears in the population in the mid 2000s. **(right)** Tajima D during the replacement event, calculated with data obtained from Genbank using the sequences published by Rabaa *et al.* between 2001-2008 [Rabaa et al., 2010]. The statistic is positive during the period in which DENV2's population is composed of intermediate frequencies of each genotype, peaking when mixing is maximum (red dotted vertical line), returning to negative thereafter.

DENV2's centred distribution. To illustrate this, the Tajima D is calculated for the genetic sequences collected during the Viet Nam genotype replacement event described in Chapter 2 (figure 5.8 on page 160). Positive values are found during the period of competitive exclusion, peaking when genotype frequencies are intermediate.

5.3.2 Modelled Phylodynamics

The agent-based epidemiological model is set within the context of a structured host-population of human and mosquito individuals within a highly endemic scenario. In order to model the evolution of viral sequences, host numbers are significantly increased to achieve large and effective pathogen population sizes. Analysis of the modelled phylodynamics is performed in 3 major steps:

(i) Simulated Topological Shape - to measure the shape of simulated phylogenies, the epidemiological and evolution models are run with parameters set to values within a biologically relevant interval (see table 5.2 on page 185). A parameter sweep approach is used to find default values of community size, host mobility and mutation rate. This is performed by approximating the Colless Index of simulated trees to the empirical estimation for DENV1, which is used as reference. The index measure resulting from the default parameter set is then used as a second reference to quantify the effect of independent model parameters on tree shape.

(ii) Simulated Population Genetics - based on the range found for the Colless Index in step (i), three parameter sets with contrasting topological observations are chosen to analyse their population genetics in detail.

(iii) Simulated and Empirical Phylodynamics - the obtained measures of topological shape (i) and population genetics (ii) are integrated and placed within the context of DENV's empirical expectations.

Simulated Topological Shape

Estimation of the index for simulated trees is performed with the same approach as with the empirical data - by fitting convergence with increasing tree size (see Materials and Methods and figure 5.5 on page 156)). The first example is illustrated in figure 5.9 (on page 163) where changes in the mutation rate, host mobility and

population structure are considered. For each parameter variation, 10 trees of sizes $s \in \{10, 25, 50, 100, 150, 200\}$ are generated from stochastic simulations of the epidemiological and evolutionary models. The Colless Index values obtained for each tree size (grey points) are used for fitting, which is represent by the lines in the figures: the best approximation to DENV1's empirical estimation is in green (also called the default parameter set), the empirical estimation for DENV1 in orange (as in figure 5.5 on page 156) and the remaining parameter sets in black (dashed). The index fittings for other parameter variations are shown in figure 5.10 on page 164. The parameter ranges used in all figures can be found in table 5.1 in page 163.

The Colless Index that best approximates DENV1's estimation (in green, solid) is further used has reference to normalize the obtained change in tree shape when varying all model parameters. The effect of each parameter variation is presented in % change of the Colless Index in figure 5.11 (on page 165), where the ordered changes highlight both increasing and decreasing effects on topological shape. The example phylogenies are from the extreme sides of the obtained Colless Index range and demonstrate how variable the topologies can be.

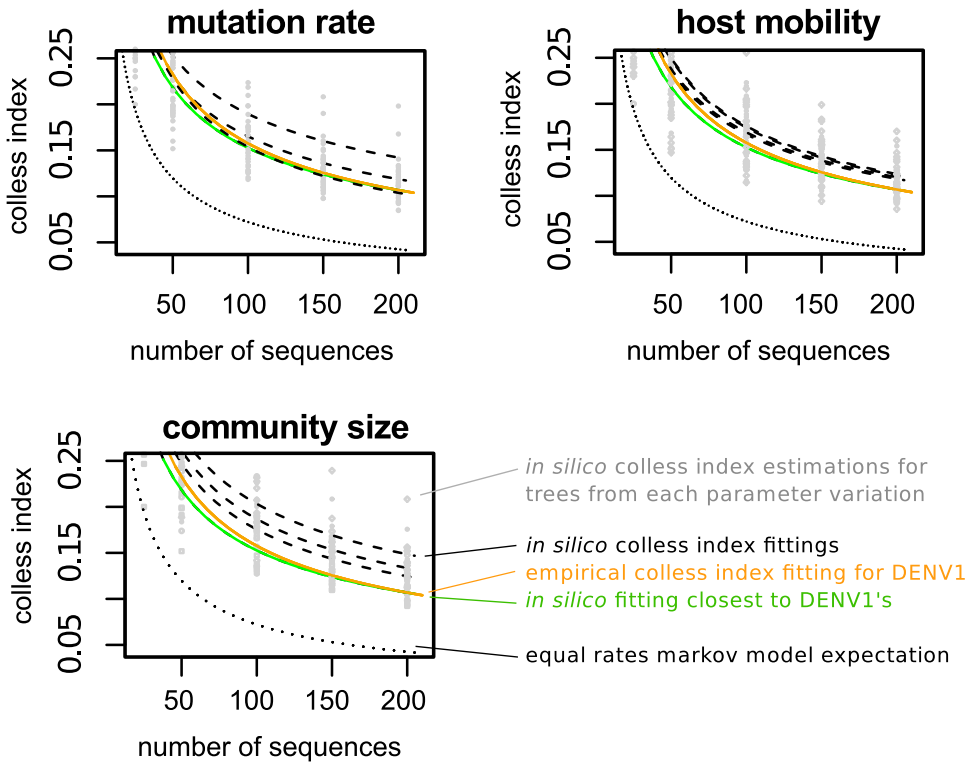


Figure 5.9: **Colless Index Fitting for Simulated Phylogenies - mutation rate, community size and host mobility** - Parameters were set to a default value within biologically relevant ranges (table 5.2 on page 185). A parameter sweep approach was used to find default values of community size, host mobility and mutation rate. For each parameter set, the model run for 10 independent stochastic simulations and Colless Index measures (grey points) were made on trees of varying sizes $(10, 25, 50, 100, 150, 200)$. Fitting of the simulated and empirical trees was therefore performed in the same way, as described in figure 5.5 on page 156. DENV1's empirical estimation is in orange (solid) and the ERM expectation in black (dotted). The green line represents the fitting that best approximates DENV1's estimation and was used to define the default parameter set.

Table 5.1: **Parameters ranges used to generate trees for Colless Index estimation.**

parameter	**range used**	**parameter**	**range used**
community size	108 - 2187	mobility ω	0.0001 - 0.2
seasonality M	0.7,1.4 - 1.4,1.4	mutation rate μ	0.001 - 0.02
transmissibility ϵ	0.4 - 0.8	latency period δ_h	1 - 12
transcending period α_h	0 - 9	ADE enhancement ϕ_h	0 - 1.5
partial immunity	0 - 0.75	n. infections	1 - 4

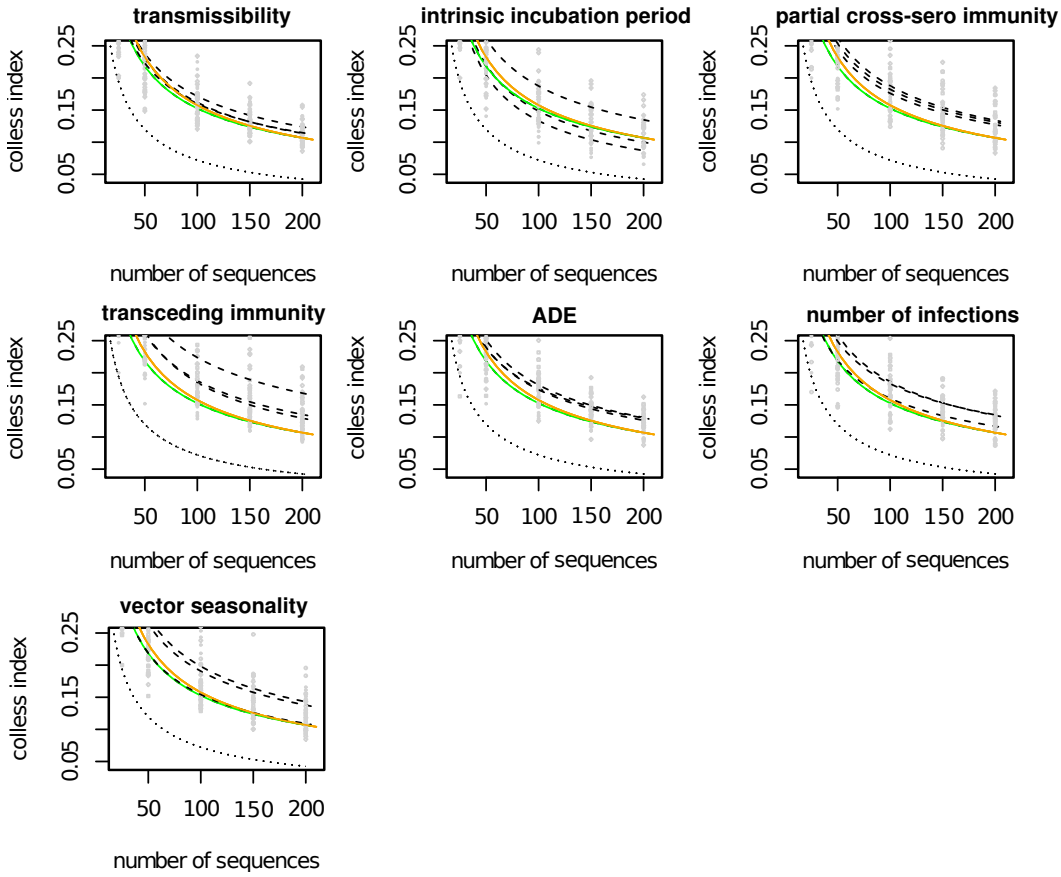


Figure 5.10: **Colless Index Fitting for Simulated Phylogenies - eco-epidemiological parameters** - Parameters were set to a default value within biologically relevant ranges (table 5.2 on page 185). Here, changes in various model parameters are considered and their effect on the converge value of the Colless Index measured. For each parameter set, the model run for 10 independent stochastic simulations and Colless Index measures (grey points) were made on trees of varying sizes $(10, 25, 50, 100, 150, 200)$. Fitting of the simulated and empirical trees was therefore performed in the same way, as described in figure 5.5 on page 156. DENV1's empirical estimation is in orange (solid) and the ERM expectation in black (dotted). The green line represents the fitting that best approximates DENV1's estimation.

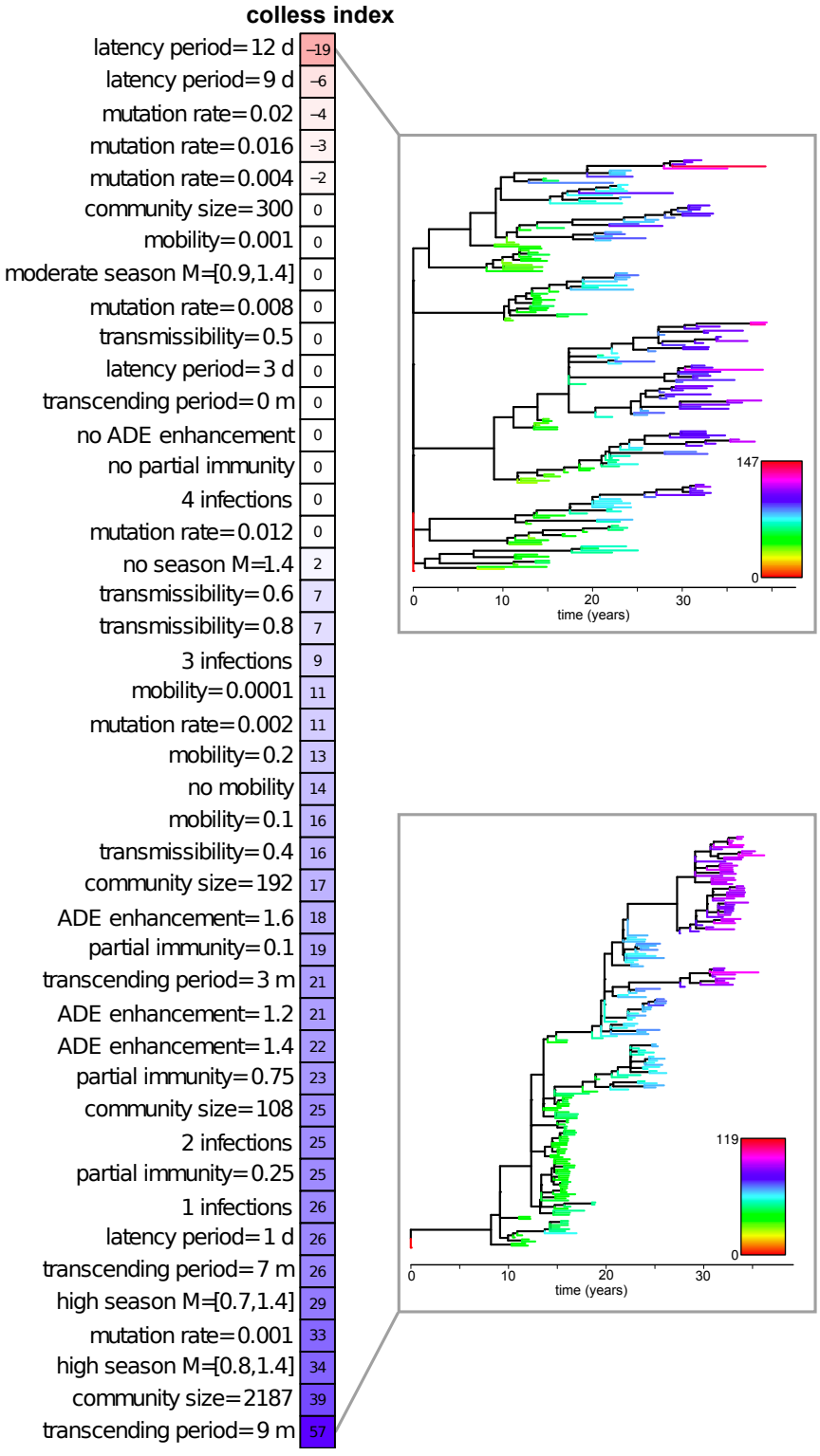


Figure 5.11: **Colless Index Ranges for Simulated Phylogenies Under Parameter Changes - (left)** Best approximation to DENV1's imbalance of figures 5.9 and 5.10 (pages 163, 164) is used to normalize all the resulting index estimations. The effect of each parameter variation is presented in % change. **(right)** Example of trees with low (top) and high (bottom) values of Colless Index.

When varying model parameters, the dynamical features of the meta-population change. To understand which changes may justify the resulting topological shapes, the correlations of the Colless Index and four surrogates (i.e. summary measures) of meta-population dynamics are investigated (see Materials and Methods on page 152 for surrogate definition). As shown in 5.12 (on page 167), the Colless Index (left column) is found to be positively correlated with extinction risk (first row) and negatively correlated with variant spatial range (left, third row). This means that, in general, phylogenetic trees with imbalance above the expected for DENV serotypes are a result of meta-population dynamics presenting both increased risk of extinction and narrow spatial ranges for the co-circulating variants. However, as shown by the unordered list of parameters along the Colless Index scale of figure 5.11 (on page 165), each parameter presents a specific relationship with the resulting topological shape. To addressed this in more detail, parameters are further categorized into three classes - immunological, ecological and viral - and their separate effect on the Colless Index and surrogates is summarized and described bellow.

Ecological Parameters and Topological Shape

For ecological parameters, i.e. community size, host mobility and seasonality, the Colless Index is found to be mostly correlated with risk of extinction (figure 5.13 on page 167). In the presence of considerable risk, local extinction is not experienced equally amongst viral variants, not only due to stochasticity but also local herd-effects. This in turn translates into asymmetric branching probabilities for the co-circulating

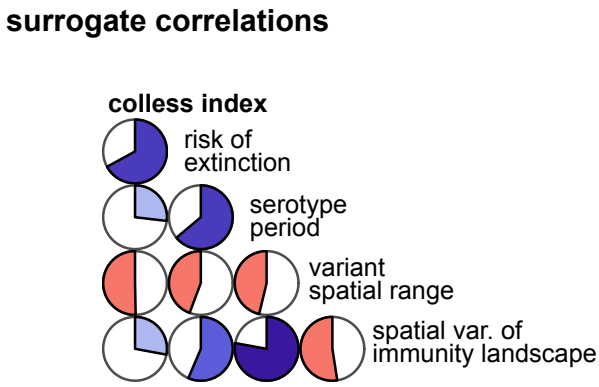


Figure 5.12: **Effects of All Parameters in Topological Shape and Surrogates of Meta-population Dynamics** - Four surrogates are used to summarize meta-population dynamics (see Materials and Methods). The correlation of dynamical surrogates with the range in Colless Index of figure 5.11 (on page 165) are presented. In general, tree imbalance is associated with increased risk of extinction and narrower spatial ranges. However, each parameter has a specific relationship with the Colless Index. This is further analysed in figures figures 5.13 on page 167, 5.14 on page 169 and 5.15 on page 171.

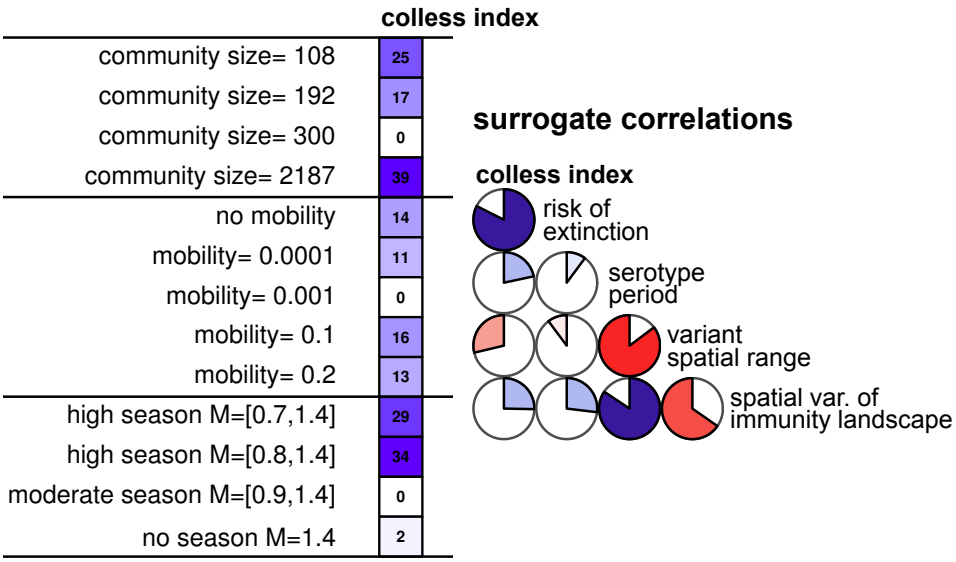


Figure 5.13: **Effects of Ecological Parameters in Topological Shape and Surrogates of Meta-Population Dynamics** - Four surrogates are used to summarize meta-population dynamics (see Materials and Methods). **(left)** Changes to topological shape when varying parameters - increases in blue, decreases in red. Values are the normalized changes to simulations with the default parameter set (lines with zeros, table 5.2 on page 185, see main text). **(right)** The surrogate of risk of extinction is the strongest determinant of topological shape when varying this class of parameters.

variants and thus explains the positive correlation with topological imbalance. Interestingly, the relationship of each parameter with the index (and therefore the risk

of extinction) is reminiscent of previous observations in the dynamical behaviour of meta-populations. For instance, both community size and mobility are reported to have a non-monotonic relationship with extinction, such that persistence is generally maximized for intermediate values of such factors [Grenfell and Harwood, 1997; Keeling, 2000]. The same is observed here, with intermediate population structures and host mobility deriving the lowest index measures, that is, more balanced phylogenetic trees.

Viral Parameters and Topological Shape

The parameters considered to be intrinsic to the virus are the mutation rate, the probability of transmission between hosts and the latency period. Under a slow molecular clock, it is expected that the majority of viruses will not experience genetic change within their life-time. As a consequence, phylogenetic shape is highly influenced by the sporadic occurrence of mutations resulting in imbalanced topologies as seen by the negative and monotonic relationship between low mutation rates and the Colless Index in figure 5.14 (on page 169). Interestingly, the index appears insensitive to intermediate rates within the interval $]0.004, 0.016[$, corresponding to a substitution rate of $\approx\ ]1.4, 5.5[\times 10^{-4}$ per site/year. This range not only derives imbalance levels that are closest to the empirical expectation but also overlaps with DENV's real substitution rates $[4.55, 11.58] \times 10^{-4}$ [Twiddy, 2003], which suggests that the *ad-hoc* values here used for the mutation rate are within a reasonable domain.

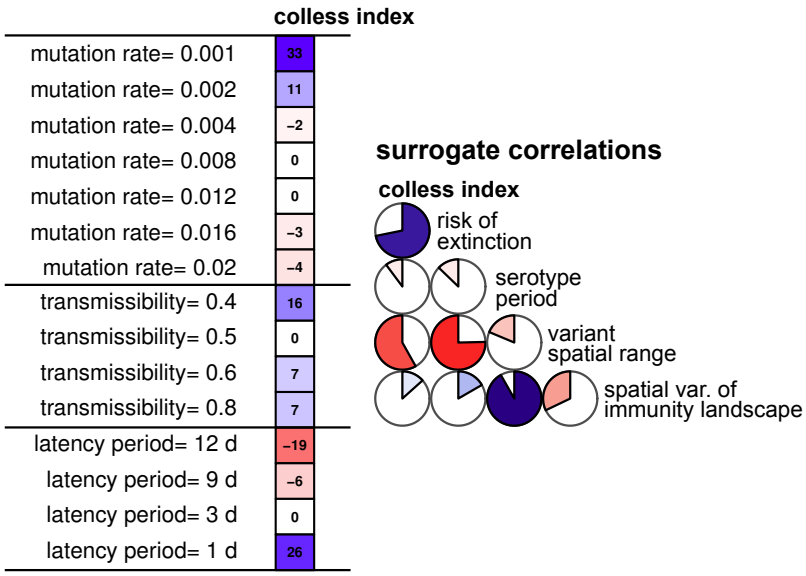


Figure 5.14: **Effects of Viral Parameters in Topological Shape and Surrogates of Meta-Population Dynamics** - Four surrogates are used to summarize meta-population dynamics (see Materials and Methods). **(left)** Changes to topological shape when varying parameters - increases in blue, decreases in red. Values are the normalized changes to simulations with the default parameter set (lines with zeros, table 5.2 on page 185, see main text). **(right)** The surrogates of risk of extinction and variant spatial range are the strongest determinants of topological shape when varying this class of parameters.

As the mutation rate does not have an effect on the meta-population dynamical behaviour, changes in the latency period and transmissibility are entirely responsible for the strong correlation of topological shape with risk of extinction and variant spatial range shown in figure 5.14. Intuitively, increases in these parameters result in lower extinction risk and wider spatial ranges (as seen by the negative correlation between the two), and thus favour balanced branching probabilities for the circulating viruses, deriving equilibrated topologies. Hence, similarly to the effects of the ecological parameters, the Colless Index is strongly and positively correlated with the risk of extinction and negatively correlated with variant spatial range.

Immunological Parameters and Topological Shape

From the three parameter classes, the ones related to immunity result in the clearest monotonic and positive relationship with topological shape as well as the highest imbalance measures (figure 5.15, left, on page 171). However, the Colless Index appears only correlated with risk of extinction and spatial range. The significant influence of variant spatial range on topological shape can be explained by the dependency of the cross-immunological reactions on local histories of infection. That is, as variant spatial range changes, so does the strength of immunological parameters in generating asymmetries in variant success and therefore branching probabilities. For example, when the spatial range is low, variant segregation is prone to happen and viruses will experience different local cross-immunological effects and therefore branching probabilities, resulting in highly asymmetric trees. In contrast, in a setting presenting wider spatial range, local histories will be less heterogeneous and variants will experience weakened herd-effects and less asymmetries in selective pressures, resulting in more balanced topologies. The slightly weaker relationship with the risk of extinction can be justified by the contrasting dynamical effects of the parameters in this class. For instance, all parameters have similar outcomes on tree imbalance but ADE is expected to contrast with the other parameters in its effects on viral persistence.

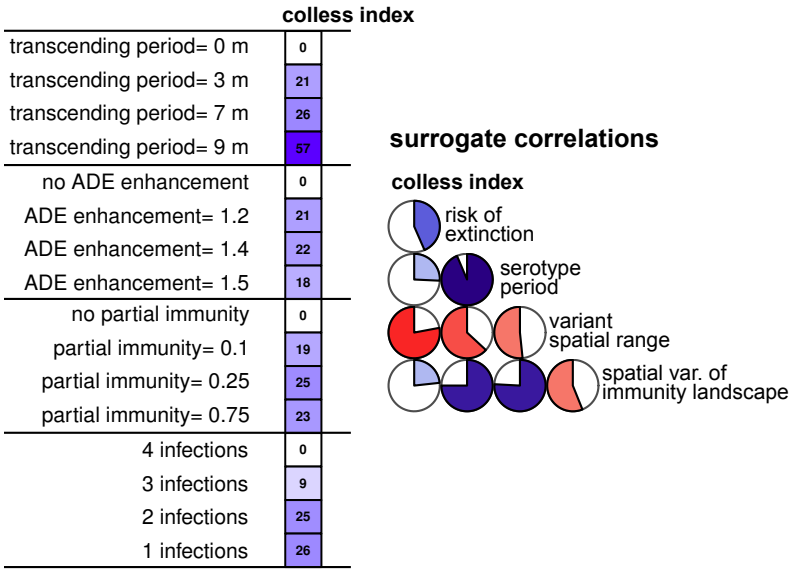


Figure 5.15: **Effects of Immunological Parameters in Topological Shape and Surrogates of Meta-Population Dynamics** - Four surrogates are used to summarize meta-population dynamics (see Materials and Methods). **(left)** Changes to topological shape when varying parameters - increases in blue, decreases in red. Values are the normalized changes to simulations with the default parameter set (lines with zeros, table 5.2 on page 185, see main text). **(right)** The surrogates of risk of extinction and variant spatial range are the strongest determinants of topological shape when varying this class of parameters.

Simulated Population Genetics

As described above, the risk of extinction and variant spatial range are the two most influencial dynamical surrogates on topological shape. The relationship of each with the Colless Index is summarized in figure 5.16 (on page 172) for all combinations of parameters tested. Three parameter sets (identified in figure 5.16 by the red, orange and blue points) are used as reference to analyse in detail the simulated population genetics: (i, orange) the parameter set that derives the lowest imbalance, which considers 12 days for the intrinsic incubation period; (ii, blue) the default parameter set that results in the same imbalance as the empirical expectation of DENV1; and (iii,

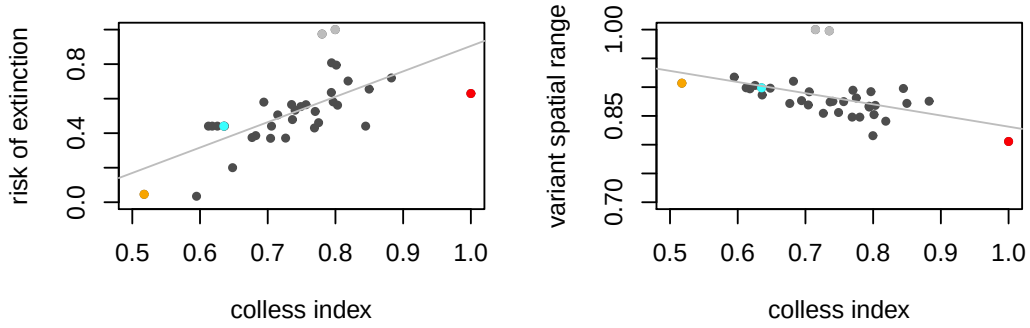


Figure 5.16: **Determinant Dynamical Surrogates of Topological Shape** - Normalized values of the Colless Index and of the two most influential dynamical surrogates - **(left)** risk of extinction and **(right)** variant spatial range. Coloured points present parameter sets that reproduced the lowest (orange), the closest to DENV1 (cyan) and highest (red) imbalances. Points in grey (outliers) highlight parameter sets with less than expected imbalance for high surrogate measures. **(left)** Risk of extinction. Outliers are parameter sets considering 1 infection and partial immunity of 0.75. **(right)** Variant spatial range. Outliers are parameter sets considering host mobility 0.1 and 0.2.

red) the parameter set that derives the highest imbalance, which considers 9 months of temporary, transcending cross-immunity.

When considering a 12 day intrinsic incubation period (orange point in figure 5.16 on page 172), the population genetics of single serotypes show general signatures that are expected from a very low risk of extinction. For example, the serotype dynamics are characterised by a high prevalence base-line, low variability in epidemic behaviour and enhanced co-circulation (figure 5.17 A on page 173). As a consequence, the two measures of genetic diversity present only weak signatures of population bottlenecks. That is, the natural fluctuations in transmission potential do not equate to similar variation in diversity, as the number of segregating sites and pairwise differences are generally increasing in time (figure 5.17 B). There is also a dominance of negative Tajima D, which reflects an excess of low frequency polymorphisms (figure 5.17 C). This is a

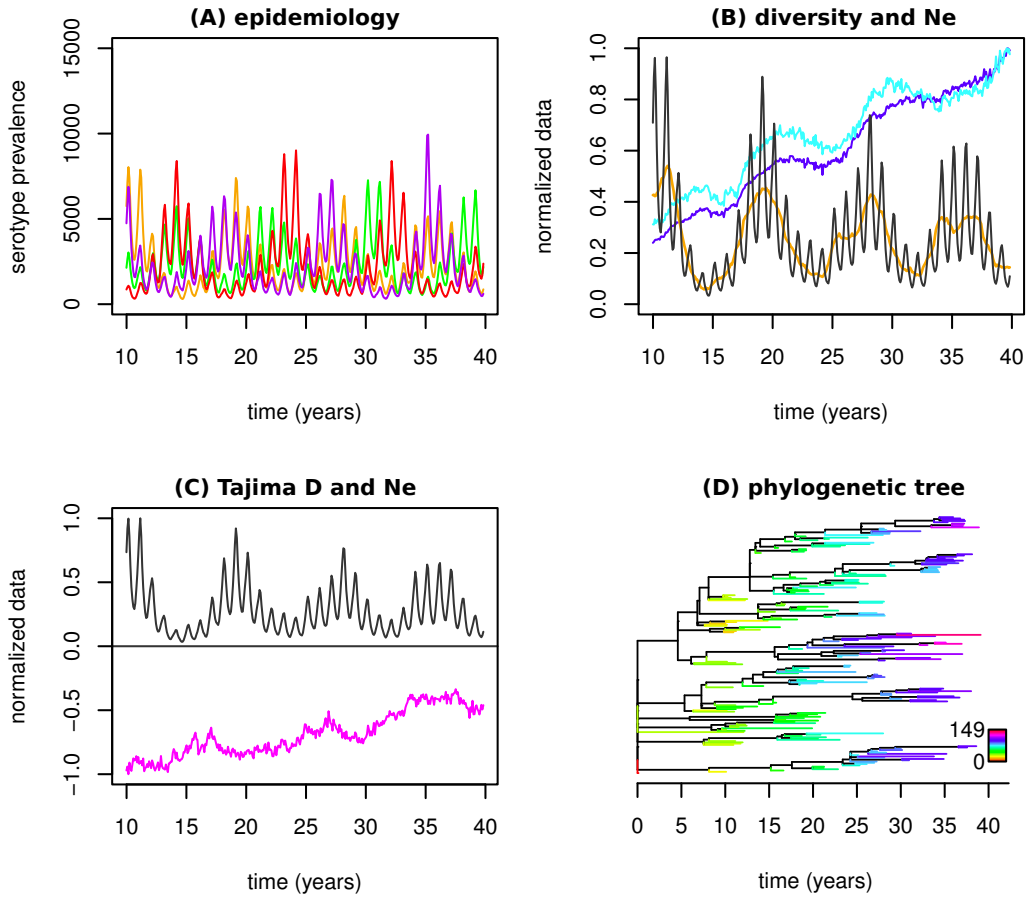


Figure 5.17: **Simulated Population Genetics and Phylogenies When Using 12 Days of Intrinsic Incubation Period (Lowest Colless Index - (A)** Serotype dynamics per month. **(B)** Effective population size *Ne* (human and vector carriers, black); serotype relative prevalence (orange); pairwise genetic diversity Π (blue); number of segregating sites S (cyan). Normalization with maximum value per series. **(C)** Effective population size *Ne* (human and vector carriers, black); Tajima D (magenta). Normalization with maximum value per series. **(D)** Phylogenetic tree with colours indicating the number of mutations accumulated (genetic distance).

result of the low risk of extinction that maintains each variant's population size stable, such that most variants persist long enough to diversify. As illustrated in the phylogenetic tree of figure 5.17 D, the intra-serotype population is composed of several variants in time, resulting in a highly balanced phylogenetic structure, similar to those observed for childhood pathogens such as MEV, but contrasting with the expectation

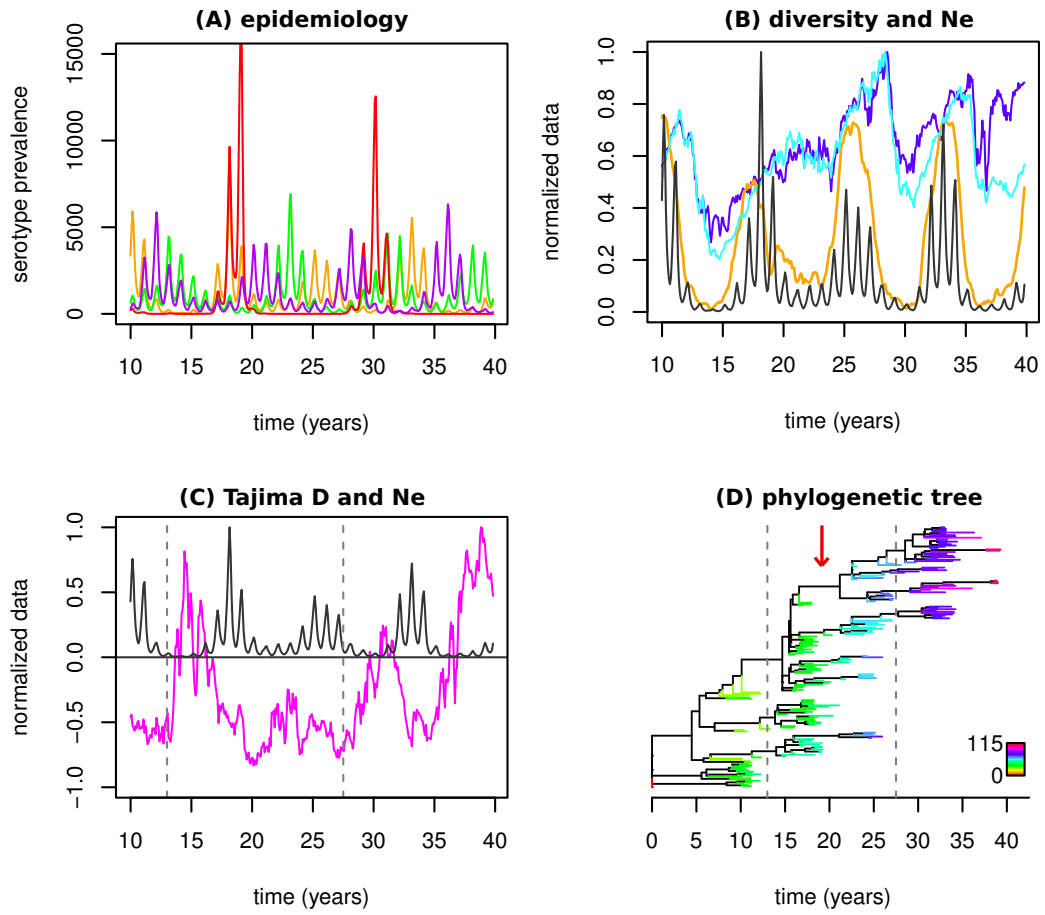


Figure 5.18: **Simulated Population Genetics and Phylogenies When Using the Default Parameter Set (Closest Colless Index to DENV1)** - Vertical dashed lines highlight population bottlenecks with phylogenetic signatures. **(A)** Serotype dynamics per month. **(B)** Effective population size Ne (human and vector carriers, black); serotype relative prevalence (orange); pairwise genetic diversity Π (blue); number of segregating sites S (cyan). Π and S had a lagged response to Ne of $\approx$ 15 months. Normalization with maximum value per series. **(C)** Effective population size Ne (human and vector carriers, black); Tajima D (magenta). Normalization with maximum value per series.**(D)** Phylogenetic tree with colours indicating the number of mutations accumulated (genetic distance). Red arrow marks lineage recycling (see main text).

for DENV [Grenfell et al., 2004].

Under the default parameter set (blue point in figure 5.16 on page 172), the simulated population genetics and phylogenetics present contrasting patterns to what is

observed when considering 12 days of intrinsic incubation period, which is mainly due to the presence of a moderate risk of local extinction. For instance, serotypes often present very low prevalence levels (figure 5.18 A on page 174) and the number of segregating sites and pairwise differences are no longer an increasing function of time, but show instead a mixed behaviour between periods of linear increase and significant reduction, indicative of population bottlenecks (figure 5.18 B). The latter are mostly associated with the natural troughs in serotype prevalence dynamics and not with seasonal variation in transmission potential. A closer inspection to the time series reveals that the response of the measures of genetic diversity lag $\approx$ 15 months behind population size variation. Overall, the Tajima D has a tendency towards negative values, indicating a population rich in low frequency polymorphisms, but with recurrent incursions into the positive spectrum (figure 5.18 C). The periods of positive Tajima D, which occur for small population sizes, are a result of substantial reductions in the number of segregating sites in comparison to changes in variant frequencies (figure 5.18 B). Within the neutral framework, this can be explained by the nature of population crashes, which induce extinction pressures that should affect equally all co-circulating variants, therefore maintaining the relative frequencies of dominant lineages but eliminating low frequency polymorphisms to a higher degree. The resulting phylogeny also presents signatures coincident with crashes in effective population size (figure 5.18 D). These are characterized by severe pruning of the co-circulating variants, prior to periods of positive Tajima D. The topological shape is nonetheless semi-balanced with varying degrees of variant co-existence in time. No-

tably, the phenomenon of lineage recycling, previously reported in dengue endemic regions [Bennett et al., 2010, 2003; McElroy, 2011] and characterized by the dominance of previously circulating variants that have undergone periods of several years with low or absent reporting, is observed sporadically (red arrow, figure 5.18 D). This phenomenon is possible under the default parameter set that induces moderate risks of local extinction and thus allows variants to circulate at different frequencies, possibly in different locations where transmission pressures vary and may allow allopatric diversification and temporary refuge.

Finally, when considering 9 months of transcending immunity (red point in figure 5.16 on page 172), population bottlenecks driven by enhanced viral competition and risk of extinction result in stronger genetic and phylogenetic signatures (figure 5.19 on page 177). As described above for the default parameter set (figure 5.18 on page 174), variant frequencies are once again less sensitive to population bottlenecks than the number of segregating sites. In this case, however, this is also observed during population expansions in which the number of pairwise differences increase slower than the number of segregating sites (figure 5.19 B). This means that the frequencies of the circulating variants have the tendency to remain stable for long periods of time independently of population size, which may be a reflection of the selective pressures now present due to transcending immunity. As a consequence, and contrasting the observations for the default parameter set in which immune competition was not present, the time lag between diversity by pairwise comparison and population size

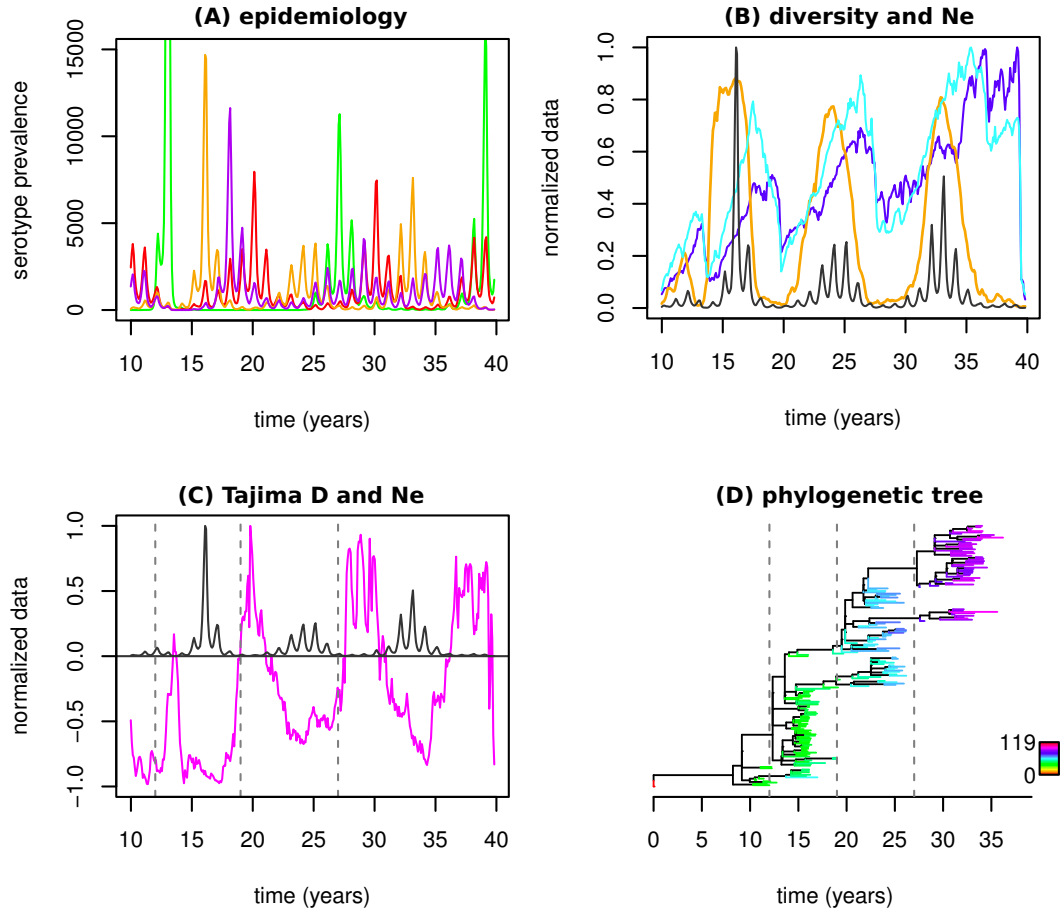


Figure 5.19: **Simulated Population Genetics and Phylogenies When Using 9 Months of Transcending Immunity (Highest Colless Index)** - Vertical dashed lines highlight population bottlenecks with phylogenetic signatures. **(A)** Serotype dynamics per month. **(B)** Effective population size Ne (human and vector carriers, black); serotype relative prevalence (orange); pairwise genetic diversity Π (blue); number of segregating sites S (cyan). Π and S had a lagged response to Ne of $\approx$ 22 and $\approx$ 16 months. Normalization with maximum value per series. **(C)** Effective population size Ne (human and vector carriers, black); Tajima D (magenta). Normalization with maximum value per series. **(D)** Phylogenetic tree with colours indicating the number of mutations accumulated (genetic distance).

is now found to have increased from $\approx$ 15 to $\approx$ 22 months. In contrast, the delay in response of the number of segregating sites remains the same ($\approx$ 16 months). The resulting phylogenetics present a highly imbalanced topology (figure 5.19 D), which is characterized by temporal clustering of similar variants during population expansion and severe pruning during bottlenecks. Both of these signatures can be explained by

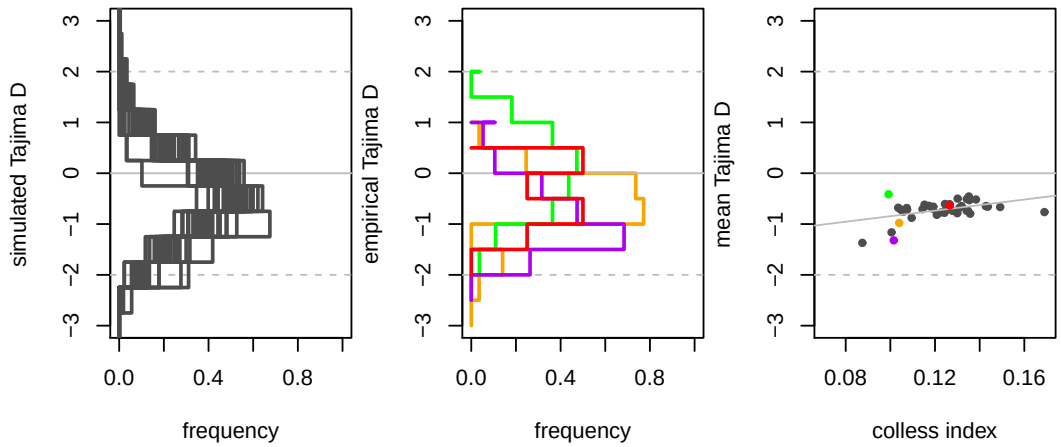


Figure 5.20: **Summarizing Empirical and Simulated Dengue Phylodynamics - (left)** Tajima D distributions (joint of 10 simulations) for each of the parameter sets modelled (table 5.2 on page 185 and figure 5.9 on page 163). **(center)** Estimations of the empirical Tajima D distribution for each DENV serotype (figure 5.7 on page 159). **(right)** Relationship of the Tajima D and Colless Index for both the empirical and simulated data. A positive relationship is found. In addition, model output covers most of the variation found in the empirical data.

the presence of viral competition due to transcending immunity. On one hand, pruning reflects more severe population bottlenecks since variants now face both the seasonal and inter-serotype troughs plus transcending herd-effects. On the other hand, population expansions are now populated by variants that successfully evade locations with significant (temporary transcending) herd-immunity and are thus characterized by growth and diversification of genetically similar viruses, explaining the phylogenetic clustering over time.

Simulated and Empirical Phylodynamics

As described above, the Tajima D distributions of the simulated population genetics are generally biased towards negative values (figure 5.20, left, page 178), in line with the empirical expectation of DENV1, 3 and 4 (figure 5.20, centre, and figure 5.7 on

page 159). Furthermore, the mean simulated Tajima D is found to be positively correlated with the Colless Index (figure 5.20, right), which can be explained by the general effects of the parameter variations here considered. That is, starting from the default parameter set that best approximated DENV1 empirical expectations, the majority of parameter variations derives into higher risks of extinction and more restricted spatial ranges. This, as demonstrated in the previous sections, results in asymmetries in variant success and more imbalanced topologies. At the same time, these asymmetries induce differences in variant relative prevalence levels in time and space, which is a driver for higher Tajima D values. However, it should be noted that some of the considered parameter sets may deviate from the dengue-relevant biological ranges. This relationship between the Tajima D and Colless Index is therefore a general expectation of host-pathogen systems and not necessarily specific to the dengue setting here modelled.

5.4 Discussion

Understanding the population structure and dynamics of a pathogen within a phylodynamic perspective can provide relevant information on its epidemic and evolutionary behaviour. So far, the links between DENV's population genetics, phylogeny and complex epidemiological dynamics have been hard to address given the relative small scale of the reported studies, which are often restricted by the numbers of genetic samples, data from a single serotype or short periods of observation including only a

few seasons. In this chapter, the general phylogenetic shape and qualitative population genetics of dengue viruses were estimated *ex novo* from sequence data and their expectations compared to simulations from an epidemiological and neutral evolutionary model.

The resulting empirical estimations for phylogenetic shape showed that DENV's typical levels of topological imbalance are at least two times lower than those presented by populations of influenza A viruses. In fact, DENV's phylogenies exhibited similar imbalance to other genetically conserved pathogens, such as MUMV and JEV. When simulating the epidemiology and evolutionary process of DENV in a stochastic and spatial setting within the meta-population model developed in Chapters 3 and 4, a wide range of topological shapes could be obtained from *in silico* phylogenies. It was found that the variation in topological shape was predominantly correlated with two concurrent dynamical features of the meta-population - the risk of local extinction and variant spatial range (the average degree of spatial dispersion amongst viral variants). Past studies have often reported the presence and role of spatial heterogeneities in dengue's epidemiological behaviour [Johansson et al., 2009; Morrison et al., 2010; Nagao et al., 2008; Tsuzuki et al., 2009; Williams et al., 2010], as well as recurrent population bottlenecks with high risk for viral extinction [Bennett et al., 2010; Williams et al., 2010] along with demonstrations that outbreaks may be driven by a limited set of spatial clusters [Cummings et al., 2004; Massad et al., 2008; Salje et al., 2012; Thai et al., 2010]. It is therefore plausible that these spatio-temporal features,

common to dengue endemic regions, are natural determinants of dengue's observed phylogenetic structures.

Importantly, the consideration of cross-protective or enhancing immunity between dengue serotypes was not necessary to simulate phylogenies with DENV-like topological shape. In fact, while the results of Chapter 4 suggest that immuno-based mechanisms are not necessary but still compatible with the epidemiological behaviour of dengue, their consideration when modelling viral evolution in this Chapter has lead to significant deviations from empirical expectations. By comparing model results to the empirical imbalance found for phylogenies of viruses from other viral families, it was found that even intermediate levels of ADE or short periods of transcending immunity, well within reasonable biological ranges for dengue, would generate phylogenetic trees with imbalance levels close to, or higher than, influenza B viruses. It cannot be ruled out, however, that the strong effects of immunological interactions are an artefact of the underlying framework. For instance, real populations may have substantially different structural organizations or might not be as well connected locally, while mosquito populations are generally patchy and volatile. Thus, it is possible that the simplifying spatial assumptions of the framework enhance co-circulation and the strength of cross-immunological interactions in the simulated phylogenies. However, their effects on the epidemiological dynamics, as described in Chapter 4, suggests that this is not the case. It should also be noted that, while these results are innovative for the understanding of the roles of both protective and enhancing immunity in the phylodynamic behaviour of dengue, they are specially so for ADE. That is, protec-

tive immunity has been previously reported in modelling studies to be the main driver of the high imbalance present in influenza A phylogenies [Bedford et al., 2011; Ferguson et al., 2003; Koelle et al., 2006a; Minayev and Ferguson, 2009]. Importantly, such studies rely on concurrent observations that influenza viruses undergo continuous antigenic change, while there is no evidence of similar genetic trends in DENV's molecular evolution that could justify the temporal clustering of similar variants or the fast and recurrent turnover reflected in the *in silico* trees in the presence of cross-immunity.

In this modelling framework, various parameter combinations were tested and their impact on dynamical properties of the meta-population were measured and compared to the resulting topological shapes. As mentioned earlier in this section, it was found that from all the observed dynamic changes taking place while varying parameters, topological shape was mainly and negatively correlated with changes in variant spatial range and positively correlated with changes in the risk of extinction. These observations followed the expectation that the branching probabilities of the co-circulating variants will be increasingly asymmetric when the degree of local extinction is considerable and the variant's spatial range is narrow.

The *ex novo* estimations of general signatures of the population genetics of dengue serotypes was done through measurements of the Tajima D statistic on empirical genetic data. A general tendency for negative Tajima D measurements was found for

serotypes 1, 3 and 4, which suggests that in real populations a vast number of low frequency polymorphisms exist that suffer significant extinction pressures and rarely emerge, which can be partially explained by the strong purifying selection imposed by DENV's two-host life cycle (see Chapter 1 on page 21). Interestingly, this evolutionary model readily reproduced Tajima D in similar ranges to those estimated from the genetic data. With evolution and transmission assumed to be neutral at all biological levels and following the results of Chapter 2, it is thus possible that demographic stochasticity, in the context of dengue, is sufficient to maintain a population rich in low frequency polymorphisms.

At the same time, a close analysis of the simulated population genetics demonstrated a natural tendency for local stable variant frequencies, which was reflected in a general time lag between the effective population size and measures of genetic diversity. Similar lagged responses have been demonstrated for endemic countries between estimated effective population size and sample counts from epidemics [Bennett et al., 2010; Raghwani et al., 2011]. Interestingly, some authors have reported that troughs in estimated population sizes are less pronounced than in case counts, arguing for a limitation in the inference approach to resolve amplitudes of highly variable (epidemic) dynamics [Bennett et al., 2010]. However, the same phenomenon is observed in the measures of *in silico* population genetics in this Chapter, and others described elsewhere [Raghwani et al., 2011], suggesting that dengue serotypes may be able to maintain genetic diversity between seasons, possibly through silent transmission, as

the majority of infections are mild or asymptomatic and not detected by surveillance systems (see Chapter 1 on page 30).

The dengue virus is a surprisingly conserved pathogen for its highly persistent nature and complex population dynamics that are similar to other fast evolving and antigenically diverse pathogens. The material presented here proposes that previously hypothesised determinants of dengue's epidemic behaviour may not be compatible with general signatures of its population genetics and phylogenetic structures. It does not, however, argue against their existence or possible role in clinical outcome. Instead, it demonstrates that their selective effect on viral variants must be negligible at the level of the population in order to conform with empirical observations. Adding to results of Chapters 2-4, it is therefore demonstrated that factors extrinsic to the virus, mostly of stochastic and non-selective nature, are sufficient and parsimonious determinants of the observed multi-scale features of its population biology.

5.5 Tables

Table 5.2: **Parameters used in the framework.**

	definition	**value**	**ref.**
$1/\delta_h$	intrinsic incubation period	3 days (2-7)	[1]
$1/\gamma_h$	human infectious period	4 days (4-12)	[2]
$1/\alpha_h$	period of temporary cross-immunity	0-9 months (2-12)	[3]
$\epsilon_{h\to v}$	human-to-vector transmission	0.5 per bite (0.33-1)	[4]
ϕ_h	ADE enhancement	1-1.5	- -
M	vectors per human host	0.9-1.4 (0.3-20)	[5]
a_v	mosquito biting rate	0.55 per day (0.33-1)	[6]
$1/\delta_v$	extrinsic incubation period	6 days (6-12)	[7]
$\epsilon_{v\to h}$	vector-to-human transmission	0.5 per bite (0.33-1)	[8]
θ	external introductions	2.2 per sero/year	- -
N_h	human host population size	3.5 million	- -
ω	daily mobility	0.001 (0-1)	- -
μ	infinite sites mutation rate	0.008	- -
N_{seq}	(evo model) human samples per day	20	- -
a	Gompertz parameter (human)	0.0002	- -
b	Gompertz parameter (human)	10	- -
ν_h	human average life-span	67.5 years	- -
a	Gompertz parameter (vector)	0.003	- -
b	Gompertz parameter (vector)	10	- -
ν_v	vector average life-span	23 days (8-42)	[9]

References: [1] Halstead [2007]; [2] Gubler et al. [1981]; Vaughn et al. [2000]; [3] Sabin [1952]; [4] Armstrong and Rico-Hesse [2003]; Sabin [1952]; [5] Focks et al. [2000]; [6] Trpis and Hausermann [1986]; [7] Watts et al. [1987]; [8] Armstrong and Rico-Hesse [2003]; Sabin [1952]; [9] Trpis and Hausermann [1986]

Table 5.3: **Abbreviations of virus' names.**

abbreviation	name	family
CHIKV	Chikungunya virus	*Togaviridae*
WNV	West nile virus	*Flaviviridae*
YFV	Yellow fever virus	*Flaviviridae*
JEV	Japanese encephalitis virus	*Flaviviridae*
RSV	Respiratory syncytial virus	*Paramyxoviridae*
ROTVA	Rotavirus A	*Reoviridae*
MUMV	Mumps virus	*Paramyxoviridae*
MEV	Measles virus	*Paramyxoviridae*
FLUB	Influenza virus B	*Orthomyxoviridae*
H3N2,H1N1	Influenza virus A	*Orthomyxoviridae*

Table 5.4: **Summary information of sequence data.**

virus	number of sequences	time period	number of regions
YFV	520	1954 - 2010	36
CHIKV	2072	1953 - 2011	100
DENV1	3575	1944 - 2012	200
DENV2	4101	1944 - 2012	242
DENV3	3179	1966 - 2012	168
DENV4	1133	1956 - 2012	92
DENV1*	1389	1944 - 2011	63
DENV2*	906	1944 - 2010	83
DENV3*	672	1973 - 2012	72
DENV4*	109	1956 - 2011	20
H1N1	6272	1933 - 2008	208
H3N2	14319	1968 - 2008	315
FLUB	6212	1940 - 2012	189
JEV	1920	1935 - 2011	86
MEV	7559	1954 - 2012	275
MUMV	1165	1958 - 2012	80
ROTVA	7616	1977 - 2012	92
RSV	1720	1999 - 2010	17
WNV	18396	1931 - 2012	342

* full genome data sets.

CHAPTER 6

Summary and Conclusion

The main aim of this thesis was to better understand the contemporary population biology of the dengue virus through the expansion of previously proposed mathematical models of transmission. Given the recent increase in the number of studies including both epidemiological and genetic data, focus was placed on possible multi-scale determinants of the ubiquitous patterns found in such data, by integrating epidemiological, ecological and evolutionary factors into unified mathematical frameworks. The results presented in Chapters 2, 4 and 5 have advanced on the current literature by proposing that the population dynamics of dengue viruses, contrary to other antigenically diverse RNA viruses, are mainly driven by the natural variation present in the transmission cycle and by underlying host-related demographic and ecological factors. While this contrasts with previous hypotheses from modelling approaches, there are key points in dengue's population biology and evolutionary history, often overlooked in the theoretical literature, that clearly suggest this to be the case.

Due to the lack of proof-reading mechanisms in the viral polymerases of RNA viruses, these pathogens present rapid evolution that occurs at a similar time-scale as their host's population dynamics. Genetic diversity is thus shaped in direct response to changes in host demographic factors, such as population size, segregation or movement, thereby providing measurable connections between population dynamics, evolution and transmission. A promising venue to study the behaviour and epidemic

potential of RNA-based viruses is therefore to perform phylodynamic studies that integrate data from multiple biological scales (i.e. epidemiological, ecological, demographic and genetic). While this approach has been particularly successful to understand the population biology of other pathogens such as the influenza and west nile viruses, the scarcity of studies presenting relevant genetic data has hindered its usefulness in the context of dengue.

At the same time, the predominance of epidemiological studies in dengue's literature has lead to the general focus of mathematical approaches on the possible determinants of its seasonal signatures and seemingly competitive dynamics of its four antigenic types (serotypes). Diverging from this trend, the stochastic model presented in Chapter 2 is the first to address the dynamic behaviour of dengue's viral lineages (genotypes). In general, empirical studies suggest the within serotype diversity to be mostly neutral and genotype co-circulation not common. The major contributing factor restricting co-circulation is believed to be cross-protective immunity between variants of the same serotype, while the overall conserved nature of the viral population is generally attributed to purifying selection. The later is a common feature of vector-transmitted pathogens, which have evolved to replicate and transmit amongst vertebrate and arthropod hosts, thus expressing a compromised genome in which most non-synonymous mutations are expected to have deleterious effects. By analysing the dynamical properties of the invasion of novel dengue genotypes into endemic populations, Chapter 2 has demonstrated that even viruses with advantageous mutations can

experience periods of several seasons (years) of low frequency circulation (waiting time) before reaching sufficiently high numbers to outcompete the already established (resident) viruses.

The fitness advantage expressed by the invading virus was found to be negatively correlated with the length of the waiting time, such that highly advantageous mutants could readily fixate. There are, however, strong reasons to believe that new dengue variants do not harbour mutations that significantly increase their relative fitness. For instance, the pressures of purifying selection negate the emergence of mutants with significant fitness advantages as these would imply concurrent substantial genetic/phenotypic change. Furthermore, as demonstrated in the modelling system of Chapter 2, moderate to high fitness advantages produce fixation times and significant signatures in the overall epidemiological dynamics that do not conform with published data on genotype replacement events. Hence, one of the main expectations from the results of Chapter 2 is that positively selected dengue mutants express low fitness advantages that, although sufficient for competitive exclusion, will induce relatively long waiting periods for emergence and will not interfere with the overall serotype dynamics. The latter implies that significant changes in the within-serotype population structure are silent at the epidemiological level, thus explaining why such events are rarely reported in the literature as their detection would require frequent and numerous genotyping in time and space.

In reality, the waiting times described above also signify an elevated risk of stochastic extinction for the emerging dengue variants. Accordingly, when demographic stochasticity was included in the analysis, the success (fixation) rate of the newly emerging viruses was found to be negatively correlated with the waiting time, and surprisingly, with the prevalence levels of the resident viruses at the time point of emergence (introduction into the population). The later highlighted the presence of strong viral competition, or clonal interference, driven by the similar levels of fitness amongst the co-circulating viruses.

In combination, these results suggested that the reported low measures of genetic diversity and rare evidence of positive selection are not necessarily a unique result of purifying selection. The generation and temporary circulation of advantageous mutants may therefore be a more common phenomenon than previously acknowledged, but their detection and emergence is highly restricted by clonal interference and a high risk of stochastic extinction. In fact, the expected low fitness advantages of the emergent genotypes implies that differences between viruses would be difficult to measure in experimental settings, which could help explain why the majority of studies searching for positively selected mutations in dengue viruses rarely find significant evidence. The case study of the recent replacement event in South East Asia, included in Chapter 2, is one example, as the reported speed of displacement suggested the Asian-1 genotype to be under positive selection although no positively selected mutations could be found.

The effects of demographic stochasticity in the population dynamics of dengue genotypes had never been demonstrated in a modelling system. Chapter 2 thus raised the question of whether the natural variation in host-host or host-pathogen interactions could also significantly affect or shape the observed epidemiological behaviour of dengue's serotypes. Based on both *in vivo* and *in vitro* observations, temporary cross-immunity and ADE enhancement have been proposed to be the main determinants of their dynamics. Nonetheless, two problems remain in the current literature: (1) a multitude of different biological assumptions based on these cross-immunological mechanisms have been shown to reproduce the observed epidemiological patterns equally well within modelling systems and thus a consensus on the exact driver(s) has not been reached; and (2) because modelling studies have disregarded the less well reported molecular scale, the link between dengue's epidemiology and molecular evolution is not well understood and it is therefore not known whether genetic data supports the hypothesis of the determinant role of cross-immunological reactions in dengue's epidemiological dynamics.

If temporary cross-immunity and ADE enhancement are thriving mechanisms of the epidemiological behaviour of dengue's serotypes, evolutionary theory predicts the presence of significant selective pressures for antigenic change. However, the current antigenic distance of human-bearing serotypes is a consequence of allopatric evolution prior to zoonosis, contrasting with influenza A viruses for example, whose antigenic

distance in time is a product of continuous genetic change selected by pressures for immune escape. Without antigenic novelty, there is little evidence that the recurrent oscillations in serotype prevalence are driven by immune escape (which in turn is sustained by no observations of human reinfection with the same serotype). Hence, our current understanding of dengue's molecular evolution suggests these oscillations to be mainly a natural consequence of susceptibility turn-over to each antigenic type, independently of whether immunological cross-reactions determine certain features of the observed dynamics. This was precisely the focus of the research in Chapter 4, in which a dengue-like modelling system is shown to sustain oscillatory behaviour of four serotypes in the absence of cross-immunological reactions.

However, it was in Chapter 3 that the fundamental mechanism for the sustained oscillatory behaviour was exemplified. There, in the context of a single-strain pathogen under an agent-based modelling approach, demographic stochasticity was shown to be sufficient to maintain the pathogen in a regime characterized by recurrent, semi-periodic epidemics. Mechanistically, the effects of stochasticity were to generate sufficient perturbations as to keep the pathogen in a transient and non-converging regime. Importantly, in a natural system, the effects of demographic stochasticity should be independent of the number of co-circulating variants. That is, stochasticity is expected to be independent between viral variants as it emerges from the natural variation in host-host and host-pathogen interactions, which involve different individuals in time and space for each variant. Hence, its effects must also be independent, and, in an

ideal scenario in which variant interactions are minimal or non-existing, each virus should remain in its own transient regime. In Chapter 4, the expectations of the effects of demographic stochasticity in a multi-strain system were confirmed. There, the agent-based model of Chapter 3 was expanded to include the circulation of four antigenic strains with minimal interactions (i.e. only restricting co-infection) and the resulting epidemiological dynamics were found to be characterized by persistent and independent oscillatory behaviour of the co-circulating viruses.

With the goal of addressing the question raised in Chapter 2 of whether demographic stochasticity was an important determinant of the observed epidemiological behaviour of dengue serotypes, the model was further expanded to include a mosquito population. The oscillatory behaviour was found to be compatible with the dengue-like framework, proposing demographic stochasticity as a possible determinant of the observed epidemiological behaviour of dengue serotypes. Importantly, stochasticity presented itself as a parsimonious explanation since its presence in the natural transmission cycle of the pathogen is certain, while the immuno-based hypothesis requires that the within-host phenomena generate significant dynamical pressures at the level of the population in time and space, which is dependent on a multitude of factors. For instance, the proposed effect of ADE enhancement to change the transmission potential of specific serotypes is dependent on viral and host genotype, as well as the local density of hosts undergoing heterologous infections, while the relevance of temporary cross-immunity in blocking a significant number of infections is limited by the

chances of hosts being challenged during its period. Crucially, factors such as these are tied to intrinsic assumptions in the modelling approaches of the studies from which the immuno-based hypothesis was originally formulated: the deterministic nature of transmission and the homogeneous mixing in the population.

To this point, demographic factors such as host-population structure and mobility were also considered in Chapter 4 and their qualitative effects analysed. Overall, the dynamic behaviour of the model was greatly influenced by demographic changes, presenting the most quantitatively similar patterns to dengue epidemiological data for host populations with intermediate- to high-structuring and with low host mobility. Importantly, this followed the conclusions of an increasing number of publications demonstrating that dengue transmission is highly localised at the level of houses and neighbourhoods. Analysis of the dynamic behaviour of serotypes in space further validated the modelling approach by reproducing, for the first time in a dengue-like model system, previously observed spatial features of dengue epidemiology such as wave-like epidemics of the four serotypes and heterogeneous immunity (or sero-conversion) landscapes.

Following the expectations from previous ODE-based modelling approaches, the epidemiological behaviour of dengue was also found to be influenced by immunological factors in the agent-based framework of Chapter 4. However, the previously proposed effects of transcending cross-immunity were found to be weakened in the presence of

stochasticity and population structure. The temporary nature of transcending immunity dictates that its protective effects are dependent on the probability of hosts being challenged by multiple serotypes within its duration. However, the average time for sequential challenge (heterologous exposure period) emerging from the model system was in the range of several years, such that the effects of transcending immunity on the epidemiological dynamics of dengue only became relevant when its duration was considered to be above 2 years, well outside the period of 2-9 months proposed by experimental reports. For this, the relaxation of the homogeneous mixing assumption played an essential role, as it allowed for the possibility of local stochastic extinction, significantly reducing serotype co-circulation and therefore increasing the heterologous exposure period. Importantly, this natural consequence of structured host-populations is not considered in deterministic and homogeneous mixing models of transmission, which suggested that the previously proposed effects of temporary, transcending cross-immunity are exacerbated in such modelling approaches. Crucial for the validation of the self-emergent heterologous exposure period of the model system was the fact that other quantitatively relevant features commonly used to validate dengue modelling approaches were reproduced for the same parameter values, which included the ages of primary and secondary infection as well as epidemic and serotype periodicities. Although there have been no studies focused on the estimation of the time required for heterologous challenge in dengue endemic regions, some have reported indirect quantifications. Importantly, these are in the range of years, close to the periods obtained in Chapter 4.

In contrast to transcending immunity, ADE in heterologous infections was found to comply to previous expectations from ODE models, which was due to the everlasting predisposition of hosts to undergo immune enhancement after the first infection. In fact, these results were evocative of the Cuban experience from the 1977's DENV1 and 1997's DENV2 epidemics in which it was demonstrated that antibodies from the first outbreak influenced the clinical outcome of secondary infections in the second outbreak, even after 20 years. What was not clear, however, was if ADE also contributed to overall elevated transmission in the second epidemic [Vaughn, 2000].

As mentioned earlier, the link between dengue's epidemiology and molecular evolution is not well understood, mainly due to insufficient data on the virus' molecular evolutionary history. In recent years, the arrival of numerous next-generation sequencing techniques has started to change this scenario; however, it is still not known which of the numerous hypothesis concerning the drivers of dengue's epidemiological dynamics are supported by genetic data. To address this, Chapter 5 presented an *in silico*, unifying perspective on dengue's phylodynamic behaviour by integrating neutral evolution of viral sequences into the mathematical framework of Chapter 4. For this, summary statistics of phylogenetic shape and population genetics were compared between model output and *ex novo* empirical estimations from publicly available serotype genetic data.

The analysis of the topological shape of dengue's phylogenies revealed these to be highly balanced, indicative of weak positive selection and close to the levels obtained for other vector-born or genetically conserved pathogens such the Japanese encephalitis, the mumps and the measles viruses. Furthermore, dengue's population genetics showed a general tendency towards negative Tajima D values, which is in accordance with purifying selection maintaining a population rich in low frequency polymorphisms that rarely emerge.

One of the most relevant conclusions from the phylodynamic model system was that cross-immunity between dengue serotypes was not necessary to generate data patterns that are quantitatively similar to the *ex novo* empirical estimations. As the modelling framework assumed evolution and transmission to be neutral (i.e. all viral variants were equally fit), selective pressures were not present. Therefore, other factors intrinsic to the modelled system were fully responsible for the balanced phylogenies and dominance of low frequency polymorphisms. Notably, the analysis of meta-population dynamical features revealed that the most determinant factor for the topological shape of *in silico* phylogenies was the risk for local stochastic extinction of the circulating viruses.

These phylodynamic results were in perfect agreement with the overall conclusion of Chapters 2 and 4 on the role of demographic stochasticity for the ubiquitous patterns found in dengue's epidemiological data and genotype dynamics. However, while

the results of Chapter 4 suggested that cross-immunity is compatible but not necessary to reproduce the epidemiological behaviour of dengue's serotypes in the presence of demographic stochasticity, in the phylodynamic approach this was found not to be the case. When modelling neutral viral evolution, it was demonstrated that even intermediate levels of cross-protective or cross-enhancing immunity, well within reasonable biological ranges for dengue, resulted in phylogenetic signatures that significantly deviated from the *ex novo* empirical estimations. In fact, the phylogenetic trees reproduced by the model system in the presence of varying levels of cross-immunity were found to have ladder-like topologies, with imbalance levels between the ones presented by empirical phylogenies of influenza B and influenza A viruses.

While not challenging the existence or role of transcending immunity and ADE enhancement in the within-host viral dynamics and clinical outcome of dengue infections, these results strongly suggested that in order to conform to empirical observations from genetic data, the selective effect of cross-immunity must be negligible at the level of the population. In reality, this meant that the effects of cross-enhancing and cross-protective immunity on the transmission success of dengue viruses is weaker than commonly acknowledged in mathematical models of transmission. In the case of ADE enhancement, it is possible that the assumption of higher levels of viraemia in hosts undergoing a secondary heterologous infection does not necessarily correlate with the success of transmission to mosquito individuals. In fact, this relationship is yet to be empirically demonstrated, while the shorter viraemia period of sec-

ondary heterologous infections further suggests a smaller time window for transmission that may counter-balance the relative increase in the probability of transmission per mosquito bite. At the same time, the exact nature of transcending immunity amongst dengue viruses is not fully understood. Experiments on human volunteers have suggested that heterotypic (cross-reactive) antibodies remain elevated for short periods after primary infection. Their presence appears to result in complete cross-protective immunity to secondary heterologous infection for only 1-2 months and partial immunity for up to 9 months that results in mild disease manifestations. Hence, the full extend of the protective effects of transcending immunity after 2 months is not known, and there is not sufficient evidence to reject the hypothesis that protection is only partially clinical and transmission still possible. Unfortunately, due to restrictions on human experimentation and unsuited animal models, this phenomenon has not been the focus of extensive *in vivo* scientific research.

The phylodynamic approach further reproduced important temporal features of the population genetics of dengue viruses. In general, genetic diversity was found to be insensitive to the seasonal variations in viral population size due to changing vector-population densities. At the same time, there were strong signatures indicative of viral population bottlenecks during the periodic multi-year troughs in serotype prevalence. Interestingly, a significant time lag was found between genetic diversity and effective population size, similarly to previous reports of endemic regions. Together, these results suggested that dengue serotypes may be able to maintain genetic diver-

sity between seasons even when facing significant risk for local stochastic extinction. In reality, as introduced in Chapter 1, several factors could partially explain these observations, from the generally mild and asymptomatic nature of dengue infections that allow clinical counts to be virtually absent during off-season periods, to vertical transmission in the mosquito-vector. However, as these factors are not taken into account implicitly in the modelling system of Chapter 5, the *in silico* results propose that meta-population dynamics may play a determinant role in the maintenance of dengue's viral diversity. This is also in accordance with previous reports on allopatric viral persistence, with variants circulating at different frequencies in different locations, and where transmission pressures vary due to local demographic of ecological heterogeneities [McElroy, 2011; Raghwani et al., 2011].

The impact of meta-population features described in this thesis, such as host population structures and host mobility, further corroborate the hypothesis that the ongoing emergence and world-wide success of dengue is mainly due to current demographic and ecological trends rather than viral adaptation. To understand its epidemiology in the long-term, it is thus crucial to establish how dengue's meta-population disease dynamics correlate with more realistic host population structures. Importantly, the discrete agent-based formulation of this framework allows to readily consider different structural organizations, such as small-world or scale-free topologies, as well as spatial heterogeneities, such as natural variations in population size and density or availability in vector breeding sites. Moreover, these are essential factors that need to

be accounted for the planning of community-specific vector control strategies as well as drug and vaccine intervention policies. The future use of the framework developed in Chapters 3, 4 and 5 is therefore promising, specially in the light of the undergoing trials of vector control through *wolbachia* infected mosquitoes [Frentiu et al., 2010; Hancock et al., 2011; Hedges et al., 2008] and sterile-mosquito methods [Alphey et al., 2010], as well as the forthcoming antiviral drugs [Lescar et al., 2008; Pastorino et al., 2010] and multivalent human vaccines [Johansson et al., 2011; Schmitz et al., 2011].

Chapters 2, 4 and 5 have explored the pros and cons of old hypothesis posed by ODE-based approaches in more complex host-pathogen systems and, importantly, have suggested new parsimonious explanations for existing features of dengue data that advance our understanding of its population biology. Besides the positive contribution to the current literature, there are three general implications from this material that merit the future attention of empirical and theoretical research:

(I) The strong effects of demographic factors on dengue's population dynamics obtained from all the modelling approaches of this thesis corroborate the idea that a natural ecological barrier is the main determinant for the rarity of sylvatic outbreaks amongst humans. If so, with the ongoing demographic and ecological changes across the globe, the sylvatic transmission cycles remain as a source for future reintroduction. This is partially relevant in the light of vaccine introduction, as it is not known if the vaccine-induced immunity is protective and/or enhancing against sylvatic viruses.

(II) The phylodynamic model system presented in Chapter 5 demonstrated, that the strong signatures in total incidence due to seasonal effects do not necessarily result in significant reductions in serotype genetic diversity. This, together with other reports, proposes that even within moderate transmission settings and highly structured host-populations allowing for local extinction, the degree of persistence is still adequate to maintain the pathogen population stable. To better understand this phenomenon, improved estimates on the frequency of silent (non-clinical) transmission are required, as well as a definite consensus on the role of vertical transmission in the vector population. This is particularly important in the face of vaccine deployment as the vaccination threshold needs to be known to induce sufficient herd-immunity to effectively reduce transmission.

(III) Given the newly identified factors (stochastic / drift in Chapter 2 and demographic in Chapter 5) that contribute to the purifying signatures present in the population genetics of dengue viruses, the emergence of advantageous variants may be a more common phenomenon than previously acknowledged (but their detection difficult). This is likely to have implications for the emergence of vaccine escape or *wolbachia* resistance once the natural selective pressures of endemic populations change upon mass interventions that may take place in the near future.

The overall results described in this thesis confirm the usefulness of multi-scale bi-

ological modelling in the context of infectious disease epidemiology. Its results redefine some of the presumptions present in dengue's literature and strongly reiterates human- and mosquito-related demographic and ecological trends as the most relevant determinants of the virus' population biology. The implications of this research material go beyond the studied pathogen, involving challenging new ways of looking at the underlying causes of the complex phylodynamics of antigenically diverse pathogens.

Glossary

allopatric or geographic evolution/speciation/transmission - Occurs when populations of the same species become geographically isolated to an extent that local evolution/speciation/transmission occurs independently. 24, 110, 176, 191, 200

antibody-dependent enhancement or ADE - Subneutralizing antibodies attached to clones of the infecting virus form an immune complex that can facilitate the infection of monocytic cells through their Fc receptors. Innate immunity responses are suppressed when cell infection happens with such immune complexes and can lead to an increase in the rate of intracellular infection and viraemia levels. 34

cellular-based or TH1-based response - Maximizes the killing efficacy of the macrophages and the proliferation of cytotoxic CD8+ T cells. Also promotes the production of *opsonizing* antibodies. 34

complement system - The complement system is part of the innate immune system. It consists of a number of small proteins found in the blood that when stimulated by certain triggers, proteases in the system cleave specific proteins to release cytokines and initiate an amplifying cascade of immunological reactions that include the adaptive system. . 37

contact rate - The number of individuals contacted by an infective host per unit of time. 93

demographic stochasticity - Inherent noise, or natural variability, in the reproductive rates of individuals within a population. 90

detection threshold - Level of viral variant prevalence within the population necessary to detect its present circulation (dependent on the surveillance system in place). See emergence time. 69

eco-spatial - The association of ecological properties or determinants, such as resource availability or climate, to a spatial dimension. Linked to the concept of spatially structured populations in ecology. 93

effective incidence ratio - Ratio between the potential and effective number of new infections in a population. The potential is determined by the deterministic backbone, i.e. the force of infection λ. The effective depends on population status and will vary in a stochastic framework. 93

emergence time - Time lag between the point of introduction/invasion of a viral variant and the time when it reaches a detectable level of prevalence within the population. See detection threshold. 74, 78

endemic - The characteristic of being endogenously persistent in a given area. Attributed to diseases or pathogens. May also refer to pathogen variants that circulate amongst human hosts, in contrast with *sylvatic* variants. 24

endemic stability - This epidemiological state is defined by patterns of lower levels of disease at higher transmission settings; possible if immunity is long lasting

and the probability of developing symptoms increases with age. Within such scenario, decreases in transmission either by natural phenomena or human intervention can increase the age at which individuals are challenged and therefore the number of symptomatic cases. 46

exposed - Or *latent*, is the epidemiological state of a host, defined by being a carrier of a pathogen but not (yet) able to transmit. It pre-seeds the infectious state. It is sometimes referred as the intrinsic incubation period for the human host, or, extrinsic incubation period for the vector host. 91

extrinsic incubation period - The time required for the virus to develop to infectivity within mosquito individuals [Anderson and Rico-Hesse, 2006]. 26

fragment crystalizable receptor - Protein at the surface of cells from the immune system that participate in the protective functions. 33

genotypes - Viral subgroups within each serotype. Defined by Rico-Hesse *et al.* as clusters of viruses with sequence divergence lower than 6% within a short genome region in the E/NS1 junction (figure 1.5 on page 21) [Rico-Hesse, 1990]. 18

H-function - The H-function is a normalization of Ripley's K-function, such that the expected resulting value is 0. It allows for a point distribution to be tested against the null hypothesis that the points are distributed randomly and independently. If positive, it indicates clustering over a spatial scale r, if negative it

indicates dispersion.. 95, 96, 108, 109

heterologous exposure period or HEP - The time required for individuals to acquire a secondary, heterologous infection. 134, 138, 195

homologous immunity - Life-long sero-specific immunity to reinfection. 31

humoral-based or TH2-based response - Stimulates B-cells into proliferation, to induce B-cell antibody class switching, and to increase neutralizing antibody production. 34

locality - Area, within the meta-population, related to a certain community C_j, defined by the set $\{C_j, K_j\}$, i.e. the community and its immediate neighbours. 94, 119

mass-action incidence - Type of incidence used in mathematical frameworks to define incidence as dependent of total population size. It implicitly assumes that individuals have a contact rate that scales linearly with total population size. 93

meta-population - A population composed of a set of local communities (patches) linked by dispersal rates (colonization rates) and subject to the possibility of local extinction. 91

non-structural protein 1 - is a DENV complement-fixing protein. It presents itself in both membrane-associated and soluble forms. Its presence is detected in

blood of patients around the time of viraemia increase and remains detectable post-viraemia and symptomatic periods. See *complement system* entry. 37

ordinary-differential equations or ODE - One equation is differential if it relates to an unknown function and one or more of its derivatives. An ordinary differential equation has only one independent variable. 38

original antigenic sin - Common phenomenon amongst infections with flaviviruses. Under this biological mechanism, B-cell clones of human hosts sequentially challenged by viruses of this family will respond by synthesizing antibodies with higher affinity for a previous infecting virus than for the current one [Halstead et al., 1983] . 33

outward connectivity - The number of possibly infective contacts a single infectious individual is expected to have outside its locality during the infectious period. 95

periodic forcing - Periodic changes added to mathematical models of disease transmission to reflect changes generally observed in data [Anderson and May, 1992] (for example: explicit modeling of yearly increases in the number of vectors leading to increases in incidence). 89

serotypes - Antigenically distinct viral groups that are epidemiologically very similar and the cause of clinically indistinguishable illnesses in humans (figure 1.3 on page 19). There are 4 DENV serotypes. . 18

silent transmission - Transmission that is not detected by surveillance systems. 28, 50, 183

single-strain host-pathogen system - Name given to an epidemiological framework which considers a pathogen whose population is assumed to be comprised of a single antigenic variant. When modelled, reinfection is not due to genetic change enabling the pathogen to escape the immune system, but rather from host-intrinsic loss of immunity with time. 91

standard incidence - Type of incidence used in mathematical frameworks to define incidence independent of total population size. It implicitly assumes that individuals have a constant contact rate independent of total population size. 93

success rate of invasion - Defined as the successful introduction/invasion of a viral variant into a population followed by competitive exclusion of the resident types (fixation). 77

sylvatic - Attributed to diseases or pathogens exclusively (or predominantly) circulating amongst non-human hosts. 19, 20, 23, 24, 29, 201

time point of introduction or TPI - Point in time in which a novel viral variant enters the host population. 74

transcending immunity - Name given to immunity that protects against any antigenic variant of the same pathogen. For the context of dengue, see *heterologous immunity*.. 32, 46

vertical transmission - Transmission from adult female mosquitoes to eggs. 29, 49, 200, 202

Acronyms

R_0 Basic Reproductive Ratio. 43

R_t Effective Reproductive Number. 43

ADE antibody-dependent enhancement. 27–29, 32, 33

DENV dengue virus. 16

DF dengue fever. 24

DHF dengue haemorrhagic fever. 24

DSS dengue shock syndrome. 24

EIP extrinsic incubation period. 21

Fc fragment, crystalizable. 27

GM Gompertz-Makeham law of mortality. 96

HEP heterologous exposure period. 125

NS1 non-structural protein 1. 30

OAS original antigenic sin. 26

ODE Ordinary differential equation(s). 31

TH1 cellular-based immune response. 27

TH2 humoral-based immune response. 27

TPI time point of introduction. 72

VT Vertical Transmission. 42

WHO World Health Organization. 12

WWII World War II. 19

Bibliography

Adams, B. and Boots, M. (2006). Modelling the relationship between antibody-dependent enhancement and immunological distance with application to dengue. *Journal of theoretical biology*, 242(2):337–46.

Adams, B. and Boots, M. (2010). How important is vertical transmission in mosquitoes for the persistence of dengue? Insights from a mathematical model. *Epidemics*, 2(1):1–10.

Adams, B., Holmes, E. C., Zhang, C., Mammen, M. P., Nimmannitya, S., Kalayanarooj, S., and Boots, M. (2006). Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. *Proceedings of the National Academy of Sciences of the United States of America*, 103(38):14234–9.

Adams, B. and Kapan, D. D. (2009). Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics. *PloS ONE*, 4(8):e6763.

Aguas, R., Gonçalves, G., and Gomes, M. G. M. (2006). Pertussis: increasing disease as a consequence of reducing transmission. *The Lancet infectious diseases*, 6(2):112–7.

Aguiar, M., Ballesteros, S., Kooi, B. W., and Stollenwerk, N. (2011). The role of seasonality and import in a minimalistic multi-strain dengue model capturing differences between primary and secondary infections: complex dynamics and its implications for data analysis. *Journal of theoretical biology*, 289:181–96.

Aguiar, M., Kooi, B. W., and Stollenwerk, N. (2008). Epidemiology of Dengue Fever: A Model with Temporary Cross-Immunity and Possible Secondary Infection Shows Bifurcations and Chaotic Behaviour in Wide Parameter Regions. *Mathematical Modelling of Natural Phenomena*, 3(4):48–70.

Aiken, S. and Leigh, C. (1978). Dengue haemorrhagic fever in South-east Asia. *Transactions of the Institute of British Geographers*, 3(4):476–497.

Alonso, D., McKane, A. J., and Pascual, M. (2007). Stochastic amplification in epidemics. *Journal of The Royal Society Interface*, 4(14):575–82.

Alphey, L., Benedict, M., Bellini, R., Clark, G. G., Dame, D. A., Service, M. W., and Dobson, S. L. (2010). Sterile-insect methods for control of mosquito-borne diseases: an analysis. *Vector borne and zoonotic diseases (Larchmont, N.Y.)*, 10(3):295–311.

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., and Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology letters*, 9(4):467–84.

Anders, K. L., Nguyet, N. M., Chau, N. V. V., Hung, N. T., Thuy, T. T., Lien, L. B., Farrar, J., Wills, B., Hien, T. T., and Simmons, C. P. (2011). Epidemiological factors associated with dengue shock syndrome and mortality in hospitalized dengue patients

in Ho Chi Minh City, Vietnam. *The American journal of tropical medicine and hygiene*, 84(1):127–34.

Anderson, J. R. and Rico-Hesse, R. (2006). Aedes aegypti vectorial capacity is determined by the infecting genotype of dengue virus. *The American journal of tropical medicine and hygiene*, 75(5):886–92.

Anderson, R. M. and May, R. M. (1992). *Infectious diseases of humans: dynamics and control*. Oxford University Press.

Andraud, M., Hens, N., Marais, C., and Beutels, P. (2012). Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches. *PloS ONE*, 7(11):e49085.

Araújo, J. M. G., Nogueira, R. M. R., Schatzmayr, H. G., Zanotto, P. M. D. a., and Bello, G. (2009). Phylogeography and evolutionary history of dengue virus type 3. *Infection, Genetics and Evolution*, 9(4):716–25.

Arenas, M. (2012). Simulation of Molecular Data under Diverse Evolutionary Scenarios. *PLoS Computational Biology*, 8(5):e1002495.

Armstrong, P. and Rico-Hesse, R. (2001). Differential susceptibility of Aedes aegypti to infection by the American and Southeast Asian genotypes of dengue type 2 virus. *Vector Borne and Zoonotic Diseases*, 1(2):159–168.

Armstrong, P. M. and Rico-Hesse, R. (2003). Efficiency of dengue serotype 2 virus strains to infect and disseminate in Aedes aegypti. *The American journal of tropical medicine and hygiene*, 68(5):539–44.

Avirutnan, P., Punyadee, N., Noisakran, S., Komoltri, C., Thiemmeca, S., Auethavornanan, K., Jairungsri, A., Kanlaya, R., Tangthawornchaikul, N., Puttikhunt, C., Pattanakitsakul, S.-N., Yenchitsomanus, P.-T., Mongkolsapaya, J., Kasinrerk, W., Sittisombut, N., Husmann, M., Blettner, M., Vasanawathana, S., Bhakdi, S., and Malasit, P. (2006). Vascular leakage in severe dengue virus infections: a potential role for the nonstructural viral protein NS1 and complement. *The Journal of infectious diseases*, 193(8):1078–88.

Balmaseda, A., Hammond, S. N., Tellez, Y., Imhoff, L., Rodriguez, Y., Saborío, S. I., Mercado, J. C., Perez, L., Videa, E., Almanza, E., Kuan, G., Reyes, M., Saenz, L., Amador, J. J., and Harris, E. (2006). High seroprevalence of antibodies against dengue virus in a prospective study of schoolchildren in Managua, Nicaragua. *Tropical medicine & international health*, 11(6):935–42.

Balsitis, S. J., Williams, K. L., Lachica, R., Flores, D., Kyle, J. L., Mehlhop, E., Johnson, S., Diamond, M. S., Beatty, P. R., and Harris, E. (2010). Lethal antibody enhancement of dengue disease in mice is prevented by Fc modification. *PLoS pathogens*, 6(2):e1000790.

Bartlett, M. S. (1957). Measles periodicity and community size. *Journal of the Royal Statistical Society A*, 120(1):48–70.

Bauch, C. T. and Earn, D. J. D. (2003). Transients and attractors in epidemics. *Proceedings of the Royal Society B*, 270(1524):1573–8.

Beatty, M. E., Stone, A., Fitzsimons, D. W., Hanna, J. N., Lam, S. K., Vong, S., Guzman, M. G., Mendez-Galvan, J. F., Halstead, S. B., Letson, G. W., Kuritsky, J., Mahoney, R., and Margolis, H. S. (2010). Best Practices in Dengue Surveillance: A Report from the Asia-Pacific and Americas Dengue Prevention Boards. *PLoS Neglected Tropical Diseases*, 4(11):e890.

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

Bennett, S. N., Drummond, a. J., Kapan, D. D., Suchard, M. a., Muñoz Jordán, J. L., Pybus, O. G., Holmes, E. C., and Gubler, D. J. (2010). Epidemic dynamics revealed in dengue evolution. *Molecular biology and evolution*, 27(4):811–8.

Bennett, S. N., Holmes, E. C., Chirivella, M., Rodriguez, D. M., Beltran, M., Vorndam, V., Gubler, D. J., and McMillan, W. O. (2003). Selection-driven evolution of emergent dengue virus. *Molecular biology and evolution*, 20(10):1650–8.

Bennett, S. N., Holmes, E. C., Chirivella, M., Rodriguez, D. M., Beltran, M., Vorndam, V., Gubler, D. J., and McMillan, W. O. (2006). Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope accompanies clade reintroduction. *The Journal of general virology*, 87(Pt 4):885–93.

Bharaj, P., Chahar, H. S., Pandey, A., Diddi, K., Dar, L., Guleria, R., Kabra, S. K., and Broor, S. (2008). Concurrent infections by all four dengue virus serotypes during an outbreak of dengue in 2006 in Delhi, India. *Virology*, 5:1.

Bolker, B. M. and Grenfell, B. T. (1993). Chaos and biological complexity in measles dynamics. *Proceedings of the Royal Society B*, 251(1330):75–81.

Bosio, C. F., Thomas, R. E., Grimstad, P. R., and Rai, K. S. (1992). Variation in the efficiency of vertical transmission of dengue-1 virus by strains of Aedes albopictus (Diptera: Culicidae). *Journal of medical entomology*, 29(6):985–9.

Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., Moyes, C. L., Farlow, A. W., Scott, T. W., and Hay, S. I. (2012). Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLoS Neglected Tropical Diseases*, 6(8):e1760.

Breckling, B., Middelhoff, U., and Reuter, H. (2006). Individual-based models as tools for ecological theory and application: Understanding the emergence of organisational properties in ecological systems. *Ecological Modelling*, 194(1-3):102–113.

Buckee, C. O., Recker, M., Watkins, E. R., and Gupta, S. (2011). Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37):15504–9.

Bukowski, J. F., Kurane, I., Lai, C. J., Bray, M., Falgout, B., and Ennis, F. a. (1989). Dengue virus-specific cross-reactive CD8+ human cytotoxic T lymphocytes. *Journal of virology*, 63(12):5086–91.

Burke, D. S., Nisalak, A., Johnson, D. E., and Scott, R. M. (1988). A prospective study of dengue infections in Bangkok. *The American journal of tropical medicine and hygiene*, 38(1):172–80.

Calisher, C. H., Karabatsos, N., Dalrymple, J. M., Shope, R. E., Porterfield, J. S., Westaway, E. G., and Brandt, W. E. (1989). Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. *The Journal of general virology*, 70 (Pt 1):37–43.

Carrington, C., Foster, J., Pybus, O., Bennett, S., and Holmes, E. (2005). Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. *Journal of virology*, 79(23):14680.

Cassetti, M. C., Durbin, A., Harris, E., Rico-Hesse, R., Roehrig, J., Rothman, A., Whitehead, S., Natarajan, R., and Laughlin, C. (2010). Report of an NIAID workshop on dengue animal models. *Vaccine*, 28(26):4229–34.

Cavrini, F., Gaibani, P., Pierro, A. M., Rossini, G., Landini, M. P., and Sambri, V. (2009). Chikungunya: an emerging and spreading arthropod-borne viral disease. *The Journal of Infection in Developing Countries*, 3(10).

Chareonsook, O., Foy, H. M., Teeraratkul, a., and Silarug, N. (1999). Changing epidemiology of dengue hemorrhagic fever in Thailand. *Epidemiology and infection*, 122(1):161–6.

Chau, T. N. B., Hieu, N. T., Anders, K. L., Wolbers, M., Lien, L. B., Hieu, L. T. M., Hien, T. T., Hung, N. T., Farrar, J., Whitehead, S., and Simmons, C. P. (2009). Dengue virus infections and maternal antibody decay in a prospective birth cohort study of Vietnamese infants. *The Journal of infectious diseases*, 200(12):1893–900.

Chesson, P. (1981). Models for spatially distributed populations: the effect of within-patch variability. *Theoretical Population Biology*, 325(198 1):288–325.

Chikaki, E. and Ishikawa, H. (2009). A dengue transmission model in Thailand considering sequential infections with all four serotypes. *The Journal of Infection in Developing Countries*, 3(9).

Chowell, G., Diaz-Dueñas, P., Miller, J. C., Alcazar-Velazco, A., Hyman, J. M., Fenimore, P. W., and Castillo-Chavez, C. (2007). Estimation of the reproduction number of dengue fever from spatial epidemic data. *Mathematical biosciences*, 208(2):571–89.

Coelho, G. E., Burattini, M. N., Teixeira, M. D. G., Coutinho, F. A. B., and Massad, E. (2008). Dynamics of the 2006/2007 dengue outbreak in Brazil. *Memórias do Instituto Oswaldo Cruz*, 103(6):535–9.

Coleman, P. G., Perry, B. D., and Woolhouse, M. E. (2001). Endemic stability–a veterinary idea applied to human public health. *Lancet*, 357(9264):1284–6.

Colless, D. H. (1982). Review of "Phylogenetics: the theory and practice of phylogenetic systematics". *Systematic Zoology*, 31:100–104.

Cologna, R., Armstrong, P., and Rico-Hesse, R. (2005). Selection for virulent dengue viruses occurs in humans and mosquitoes. *Journal of virology*, 79(2):853.

Cummings, D. a. T., Iamsirithaworn, S., Lessler, J. T., McDermott, A., Prasanthong, R., Nisalak, A., Jarman, R. G., Burke, D. S., and Gibbons, R. V. (2009). The impact of the demographic transition on dengue in Thailand: insights from a statistical analysis and mathematical modeling. *PLoS Medicine*, 6(9):e1000139.

Cummings, D. A. T., Irizarry, R. A., Huang, N. E., Endy, T. P., Nisalak, A., Ungchusak, K., and Burke, D. S. (2004). Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature*, 427(6972):344–347.

Cummings, D. A. T., Schwartz, I. B., Billings, L., Shaw, L. B., and Burke, D. S. (2005). Dynamic effects of antibody-dependent enhancement on the fitness of viruses. *Proceedings of the National Academy of Sciences of the United States of America*, 102(42):15259–64.

de Wazières, B., Gil, H., Vuitton, D. A., and Dupond, J. L. (1998). Nosocomial transmission of dengue from a needlestick injury. *Lancet*, 351(9101):498.

DeAngelis, D. L. and Mooij, W. M. (2005). Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology, Evolution, and Systematics*, 36(1):147–168.

Degallier, N., Favier, C., Boulanger, J.-P., and Menkes, C. (2009). Imported and autochthonous cases in the dynamics of dengue epidemics in Brazil. *Revista de saúde pública*, 43(1):1–7.

Dejnirattisai, W., Jumnainsong, A., Onsirisakul, N., Fitton, P., Vasanawathana, S., Limpitikul, W., Puttikhunt, C., Edwards, C., Duangchinda, T., Supasa, S., Chawansuntati, K., Malasit, P., Mongkolsapaya, J., and Screaton, G. (2010). Cross-reacting antibodies enhance dengue virus infection in humans. *Science*, 328(5979):745–8.

dos Santos, M. (2012). *Rich Dynamics in Multi-strain Models: Non-linear Dynamics and Deterministic Chaos in Dengue Fever Epidemiology*. Centro de Matematica e Aplicações Fundamentais.

Durret, R. and Levin, S. (1994). The importance of being discrete (and spatial). *Theoretical population biology*, 46:363–394.

Duyen, H. T. L., Ngoc, T. V., Ha, D. T., Hang, V. T. T., Kieu, N. T. T., Young, P. R., Farrar, J. J., Simmons, C. P., Wolbers, M., and Wills, B. a. (2011). Kinetics of plasma viremia and soluble nonstructural protein 1 concentrations in dengue: differential effects according to serotype and immune status. *The Journal of infectious diseases*, 203(9):1292–300.

Earn, D. J. D., Levin, S. A., and Rohani, P. (2000a). Coherence and Conservation. *Science*, 290(5495):1360–1364.

Earn, D. J. D., Rohani, P., Bolker, B. M., and Grenfell, B. T. (2000b). A simple model for complex dynamical transitions in epidemics. *Science*, 287(5453):667–670.

Edwards, C. T. T., Holmes, E. C., Pybus, O. G., Wilson, D. J., Viscidi, R. P., Abrams, E. J., Phillips, R. E., and Drummond, a. J. (2006). Evolution of the human immunodeficiency virus envelope gene is dominated by purifying selection. *Genetics*, 174(3):1441–53.

Favier, C., Degallier, N., Rosa-Freitas, M. G., Boulanger, J. P., Costa Lima, J. R., Luitgards-Moura, J. F., Menkès, C. E., Mondet, B., Oliveira, C., Weimann, E. T. S., and Tsouris, P. (2006). Early determination of the reproductive number for vector-borne diseases: the case of dengue in Brazil. *Tropical medicine & international health : TM & IH*, 11(3):332–40.

Favier, C., Schmit, D., Müller-Graf, C. D. M., Cazelles, B., Degallier, N., Mondet, B., and Dubois, M. A. (2005). Influence of spatial heterogeneity on an emerging infectious disease: the case of dengue epidemics. *Proceedings of the Royal Society B*, 272(1568):1171–7.

Ferguson, N. M., Anderson, R. M., and Gupta, S. (1999a). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, 96(2):790–4.

Ferguson, N. M., Donnelly, C. A., and Anderson, R. M. (1999b). Transmission dynamics and epidemiology of dengue: insights from age-stratified sero-prevalence surveys. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 354(1384):757–68.

Ferguson, N. M., Galvani, A. P., and Bush, R. M. (2003). Ecological and immunological determinants of influenza evolution. *Nature*, 422(6930):428–433.

Fernandez-Garcia, M.-D., Mazzon, M., Jacobs, M., and Amara, A. (2009). Pathogenesis of flavivirus infections: using and abusing the host cell. *Cell host & microbe*, 5(4):318–28.

Fischer, D. B. and Halstead, S. B. (1970). Observations related to pathogenesis of dengue hemorrhagic fever. Examination of age specific sequential infection rates using a mathematical model. *The Yale journal of biology and medicine*, 42(5):329–49.

Focks, D. a., Brenner, R. J., Hayes, J., and Daniels, E. (2000). Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. *The American journal of tropical medicine and hygiene*, 62(1):11–8.

Foster, J. E., Bennett, S. N., Vaughan, H., Vorndam, V., McMillan, W. O., and Carrington, C. V. F. (2003). Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean. *Virology*, 306(1):126–34.

Francis, T. (1960). On the doctrine of original antigenic sin. *Proceedings of the American Philosophical Society*, 104(6):572–578.

Franco, L., Di Caro, A., Carletti, F., Vapalahti, O., Renaudat, C., Zeller, H., and Tenorio, A. (2010). Recent expansion of dengue virus serotype 3 in West Africa. *Euro surveillance*, 15(7):2005–2008.

Frentiu, F., Robinson, J., Young, P., McGraw, E., O'Neill, S., and Coffey, L. (2010). Wolbachia-Mediated Resistance to Dengue Virus Infection and Death at the Cellular Level. *PloS one*, 5(10):480–496.

Fried, J. R., Gibbons, R. V., Kalayanarooj, S., Thomas, S. J., Srikiatkhachorn, A., Yoon, I., Jarman, R. G., Green, S., Rothman, A. L., and Cummings, D. A. T. (2010). Serotype-specific differences in the risk of dengue hemorrhagic fever: an analysis of data collected in Bangkok, Thailand from 1994 to 2006. *PLoS Neglected Tropical Diseases*, 4(3):e617.

Gautret, P., Botelho-Nevers, E., Charrel, R. N., and Parola, P. (2010). Dengue virus infections in travellers returning from Benin to France, July-August 2010. *Euro surveillance*, 15(36):2–3.

Gibbons, R. V., Kalanarooj, S., Jarman, R. G., Nisalak, A., Vaughn, D. W., Endy, T. P., Mammen, M. P., and Srikiatkhachorn, A. (2007). Analysis of repeat hospital admissions for dengue to estimate the frequency of third or fourth dengue infections resulting in admissions and dengue hemorrhagic fever, and serotype sequences. *American Journal of Tropical Medicine and Hygiene*, 77(5):910–3.

Gillespie, D. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361.

Gillespie, J. H. (2004). *Population Genetics - A concise guide*. The John Hopkins University Press, 2nd edition.

Gjenero-Margan, I., Aleraj, B., Krajcar, D., Lesnikar, V., Klobučar, A., Pem-Novosel, I., Kurečić-Filipović, S., Komparak, S., Martić, R., Duričić, S., Betica-Radić, L., Okmadžić, J., Vilibić-Čavlek, T., Babić-Erceg, A., Turković, B., Avsić-Županc, T., Radić, I., Ljubić, M., Sarac, K., Benić, N., and Mlinarić-Galinović, G. (2011). Autochthonous dengue fever in Croatia, August-September 2010. *Euro surveillance*, 16(9).

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the royal society of london*, 115(1825):513–583.

Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006). Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21):2695–6.

Gray, R. R., Pybus, O. G., and Salemi, M. (2011). Measuring the temporal structure in serially sampled phylogenies. *Methods in Ecology and Evolution*, 2(5):437–445.

Green, S., Vaughn, D. W., Kalayanarooj, S., Nimmannitya, S., Suntayakorn, S., Nisalak, A., Lew, R., Innis, B. L., Kurane, I., Rothman, a. L., and Ennis, F. a. (1999a). Early immune activation in acute dengue illness is related to development of plasma leakage and disease severity. *The Journal of infectious diseases*, 179(4):755–62.

Green, S., Vaughn, D. W., Kalayanarooj, S., Nimmannitya, S., Suntayakorn, S., Nisalak, a., Rothman, a. L., and Ennis, F. a. (1999b). Elevated plasma interleukin-10 levels in acute dengue correlate with disease severity. *Journal of medical virology*, 59(3):329–34.

Grenfell, B. T., Bjø rnstad, O. N., and Kappey, J. (2001). Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414(6865):716–23.

Grenfell, B. T. and Harwood, J. (1997). (Meta) population dynamics of infectious diseases. *Trends in ecology & evolution*, 12(10):395–399.

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. a., and Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, 303(5656):327–32.

Gubler, D. J. (1998). Dengue and dengue hemorrhagic fever. *Clinical microbiology reviews*, 11(3):480–96.

Gubler, D. J. (2002). Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends in microbiology*, 10(2):100–3.

Gubler, D. J. (2011). Emerging vector-borne flavivirus diseases: are vaccines the solution? *Expert review of vaccines*, 10(5):563–5.

Gubler, D. J., Reed, D., Rosen, L., and Hitchcock, J. R. (1978). Epidemiologic, clinical, and virologic observations on dengue in the Kingdom of Tonga. *The American journal of tropical medicine and hygiene*, 27(3):581–9.

Gubler, D. J., Suharyono, W., and Tan, R. (1981). Viraemia in patients with naturally acquired dengue infection. *Bulletin of the World Health Organization*, 59(2):623–630.

Gupta, S. and Day, K. P. (1994). A strain theory of malaria transmission. *Parasitology today (Personal ed.)*, 10(12):476–81.

Guzman, M. G., Alvarez, M., Rodriguez-Roche, R., Bernardo, L., Montes, T., Vazquez, S., Morier, L., Alvarez, A., Gould, E. a., Kouri, G., and Halstead, S. B. (2007). Neutralizing antibodies after infection with dengue 1 virus. *Emerging infectious diseases*, 13(2):282–6.

Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., Hunsperger, E., Kroeger, A., Margolis, H. S., Martínez, E., Nathan, M. B., Pelegrino, J. L., Simmons, C., Yoksan, S., and Peeling, R. W. (2010). Dengue: a continuing global threat. *Nature Reviews Microbiology*, 8(12):S7–S16.

Guzman, M. G. and Kouri, G. (2008). Dengue haemorrhagic fever integral hypothesis: confirming observations, 1987-2007. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102(6):522–3.

Halasa, Y. a., Shepard, D. S., and Zeng, W. (2012). Economic cost of dengue in Puerto Rico. *The American journal of tropical medicine and hygiene*, 86(5):745–52.

Halstead, S. B. (2007). Dengue. *Lancet*, 370(9599):1644–52.

Halstead, S. B. (2008). Dengue virus-mosquito interactions. *Annual Review of Entomology*, 53:273–91.

Halstead, S. B., Mahalingam, S., Marovich, M. a., Ubol, S., and Mosser, D. M. (2010). Intrinsic antibody-dependent enhancement of microbial infection in macrophages: disease regulation by immune complexes. *The Lancet Infectious Diseases*, 10(10):712–722.

Halstead, S. B. and O'Rourke, E. J. (1977). Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralizing antibody. *Journal of Experimental Medicine*, 146(1):201–17.

Halstead, S. B., Rojanasuphot, S., and Sangkawibha, N. (1983). Original antigenic sin in dengue. *The American journal of tropical medicine and hygiene*, 32(1):154–6.

Hancock, P. a., Sinkins, S. P., and Godfray, H. C. J. (2011). Strategies for introducing Wolbachia to reduce transmission of mosquito-borne diseases. *PLoS neglected tropical diseases*, 5(4):e1024.

Hang, V. T. T., Holmes, E. C., Duong, V., Nguyen, T. Q., Tran, T. H., Quail, M., Churcher, C., Parkhill, J., Cardosa, J., Farrar, J., Wills, B., Lennon, N. J., Birren, B. W., Buchy, P., Henn, M. R., and Simmons, C. P. (2010). Emergence of the Asian 1 genotype of dengue virus serotype 2 in viet nam: in vivo fitness advantage and lineage replacement in South-East Asia. *PLoS neglected tropical diseases*, 4(7):e757.

Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396:41–49.

Harrington, L. C., Scott, T. W., Lerdthusnee, K., Coleman, R. C., Costero, A., Clark, G. G., Jones, J. J., Kitthawee, S., Kittayapong, P., Sithiprasasna, R., and Edman, J. D. (2005). Dispersal of the dengue vector Aedes aegypti within and between rural communities. *The American journal of tropical medicine and hygiene*, 72(2):209–20.

Hastings, A. (2001). Transient dynamics and persistence of ecological systems. *Ecology Letters*, 4(3):215–220.

Hedges, L. M., Brownlie, J. C., O'Neill, S. L., and Johnson, K. N. (2008). Wolbachia and virus protection in insects. *Science (New York, N.Y.)*, 322(5902):702.

Higa, Y. (2011). Dengue Vectors and their Spatial Distribution. *Tropical medicine and health*, 39(4 Suppl):17–27.

Holmes, E. (2003). Patterns of intra-and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. *Journal of virology*, 77(20):11296.

Holmes, E. C. and Twiddy, S. S. (2003). The origin, emergence and evolutionary genetics of dengue virus. *Infection, Genetics and Evolution*, 3(1):19–28.

Horstick, O., Runge-Ranzinger, S., Nathan, M. B., and Kroeger, A. (2010). Dengue vector-control services: how do they work? A systematic literature review and country case studies. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104(6):379–86.

Huston, M., DeAngelis, D., and Post, W. (1988). New computer models unify ecological theory. *BioScience*, 38(10):682–691.

Irie, K., Mohan, P. M., Sasaguri, Y., Putnak, R., and Padmanabhan, R. (1989). Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C strain). *Gene*, 75(2):197–211.

Issack, M. I. (2010). Reemergence of Dengue in Mauritius. *Emerging Infectious Diseases*, 16(4):3–5.

Jansen, V. A. (1999). Phase locking: another cause of synchronicity in predator–prey systems. *Trends in Ecology & Evolution*, 14(7):278–279.

Jessie, K., Fong, M. Y., Devi, S., Lam, S. K., and Wong, K. T. (2004). Localization of dengue virus in naturally infected human tissues, by immunohistochemistry and in situ hybridization. *The Journal of infectious diseases*, 189(8):1411–8.

Johansson, M. a., Dominici, F., and Glass, G. E. (2009). Local and global effects of climate on dengue transmission in Puerto Rico. *PLoS neglected tropical diseases*, 3(2):e382.

Johansson, M. a., Hombach, J., and Cummings, D. a. T. (2011). Models of the impact of dengue vaccines: A review of current research and potential approaches. *Vaccine*, pages 1–9.

John, T. J. and Samuel, R. (2000). Herd immunity and herd effect: new insights and definitions. *European Journal of Epidemiology*, 16(7):601–6.

Kawaguchi, I., Sasaki, A., and Boots, M. (2003). Why are dengue virus serotypes so distantly related? Enhancement and limiting serotype similarity between dengue virus strains. *Proceedings. Biological sciences / The Royal Society*, 270(1530):2241–7.

Keeling, M. J. (2000). Metapopulation moments: coupling, stochasticity and persistence. *Journal of Animal Ecology*, 69(5):725–736.

Keeling, M. J. (2005). The implications of network structure for epidemic dynamics. *Theoretical population biology*, 67(1):1–8.

Kermack, A. and McKendric, W. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*, 115(772):700–721.

Kirkpatrick, M. and Slatkin, M. (1993). Searching for evolutionary patterns in the shape of a phylogenetic tree. *Evolution*, 47(4):1171–1181.

Kiskowski, M. a., Hancock, J. F., and Kenworthy, A. K. (2009). On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophysical journal*, 97(4):1095–103.

Klungthong, C., Zhang, C., Mammen, M. P., Ubol, S., and Holmes, E. C. (2004). The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand. *Virology*, 329(1):168–79.

Kochel, T. J., Watts, D. M., Gozalo, A. S., Ewing, D. F., Porter, K. R., and Russell, K. L. (2005). Cross-serotype neutralization of dengue virus in Aotus nancymae monkeys. *The Journal of infectious diseases*, 191(6):1000–4.

Kochel, T. J., Watts, D. M., Halstead, S. B., Hayes, C. G., Espinoza, A., Felices, V., Caceda, R., Bautista, C. T., Montoya, Y., Douglas, S., and Russell, K. L. (2002). Effect of dengue-1 antibodies on American dengue-2 viral infection and dengue haemorrhagic fever. *Lancet*, 360(9329):310–2.

Koelle, K., Cobey, S., Grenfell, B., and Pascual, M. (2006a). Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science*, 314(5807):1898–903.

Koelle, K., Pascual, M., and Yunus, M. (2006b). Serotype cycles in cholera dynamics. *Proceedings of the Royal Society B*, 273(1603):2879–86.

Koopman, J. S., Prevots, D. R., Vaca Marin, M. a., Gomez Dantes, H., Zarate Aquino, M. L., Longini, I. M., and Sepulveda Amor, J. (1991). Determinants and predictors of dengue infection in Mexico. *American journal of epidemiology*, 133(11):1168–78.

Kraiselburd, E. N., Lavergne, J. a., Woodall, J. P., Kessler, M. J., Meier, G., Chiriboga, J., Moore, C. G., Sather, G. E., Pomales, A., Maldonado, E., and Rivera, R. (1981). Lack of greater seroconversion of rhesus monkeys after subcutaneous inoculation of dengue type 2 live-virus vaccine combined with infection-enhancing antibodies. *Infection and immunity*, 33(2):389–94.

Kuno, G., Chang, G. J., Tsuchiya, K. R., Karabatsos, N., and Cropp, C. B. (1998). Phylogeny of the genus Flavivirus. *Journal of virology*, 72(1):73–83.

Kurane, I., Matsutani, T., Suzuki, R., Takasaki, T., Kalayanarooj, S., Green, S., Rothman, A. L., and Ennis, F. a. (2011). T-cell responses to dengue virus in humans. *Tropical medicine and health*, 39(4 Suppl):45–51.

Kyle, J. L. and Harris, E. (2008). Global spread and persistence of dengue. *Annual review of microbiology*, 62:71–92.

Lescar, J., Luo, D., Xu, T., Sampath, A., Lim, S. P., Canard, B., and Vasudevan, S. G. (2008). Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target. *Antiviral research*, 80(2):94–101.

Levin, S. A., Grenfell, B. T., Hastings, A., and Perelson, A. S. (1997). Mathematical and computational challenges in population biology and ecosystems science. *Science*, 275(5298):334–343.

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, pages 237–240.

Libraty, D. H., Endy, T. P., Houng, H.-S. H., Green, S., Kalayanarooj, S., Suntayakorn, S., Chansiriwongs, W., Vaughn, D. W., Nisalak, A., Ennis, F. a., and Rothman, A. L. (2002a). Differing influences of virus burden and immune activation on disease

severity in secondary dengue-3 virus infections. *The Journal of infectious diseases*, 185(9):1213–21.

Libraty, D. H., Young, P. R., Pickering, D., Endy, T. P., Kalayanarooj, S., Green, S., Vaughn, D. W., Nisalak, A., Ennis, F. a., and Rothman, A. L. (2002b). High circulating levels of the dengue virus nonstructural protein NS1 early in dengue illness correlate with the development of dengue hemorrhagic fever. *The Journal of infectious diseases*, 186(8):1165–8.

Liebman, K. A., Stoddard, S. T., Morrison, A. C., Rocha, C., Minnick, S., Sihuincha, M., Russell, K. L., Olson, J. G., Blair, P. J., Watts, D. M., Kochel, T., and Scott, T. W. (2012). Spatial dimensions of dengue virus transmission across interepidemic and epidemic periods in Iquitos, Peru (1999–2003). *PLoS Neglected Tropical Diseases*, 6(2):e1472.

Lin, C.-C., Huang, Y.-H., Shu, P.-Y., Wu, H.-S., Lin, Y.-S., Yeh, T.-M., Liu, H.-S., Liu, C.-C., and Lei, H.-Y. (2010). Characteristic of dengue disease in Taiwan: 2002-2007. *The American journal of tropical medicine and hygiene*, 82(4):731–9.

Lin, S.-W., Chuang, Y.-C., Lin, Y.-S., Lei, H.-Y., Liu, H.-S., and Yeh, T.-M. (2012). Dengue virus nonstructural protein NS1 binds to prothrombin/thrombin and inhibits prothrombin activation. *The Journal of infection*, 64(3):325–34.

Loke, H., Bethell, D. B., Phuong, C. X., Dung, M., Schneider, J., White, N. J., Day, N. P., Farrar, J., and Hill, a. V. (2001). Strong HLA class I–restricted T cell responses in dengue hemorrhagic fever: a double-edged sword? *The Journal of infectious diseases*, 184(11):1369–73.

Loroño Pino, M. a., Cropp, C. B., Farfán, J. a., Vorndam, a. V., Rodríguez-Angulo, E. M., Rosado-Paredes, E. P., Flores-Flores, L. F., Beaty, B. J., and Gubler, D. J. (1999). Common occurrence of concurrent infections by multiple dengue virus serotypes. *The American journal of tropical medicine and hygiene*, 61(5):725–30.

Lourenço, J. and Recker, M. (2010). Viral and epidemiological determinants of the invasion dynamics of novel dengue genotypes. *PLoS Neglected Tropical Diseases*, 4(11):e894.

Maciel-De-Freitas, R., Codeço, C. T., and Lourenço-De-Oliveira, R. (2007). Body size-associated survival and dispersal rates of Aedes aegypti in Rio de Janeiro. *Medical and veterinary entomology*, 21(3):284–92.

Maramorosch, K., Shatkin, A. J., and Murphy, F. A. (2008). *Advances in Virus Research Vol. 72*. Elsevier Science.

Marques, C. A., Forattini, O. P., and Massad, E. (1994). The basic reproduction number for dengue fever in São Paulo state, Brazil: 1990-1991 epidemic. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 88(1):58–9.

Massad, E., Burattini, M. N., Coutinho, F. A. B., and Lopez, L. F. (2003). Dengue and the risk of urban yellow fever reintroduction in São Paulo State, Brazil. *Revista de saúde pública*, 37(4):477–84.

Massad, E., Coutinho, F. A., Burattini, M. N., and Lopez, L. F. (2001). The risk of yellow fever in a dengue-infested area. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 95(4):370–4.

Massad, E., Ma, S., Chen, M., Struchiner, C., Stollenwerk, N., and Aguiar, M. (2008). Scale-free network of a dengue epidemic. *Applied Mathematics and Computation*, 195(2):376–381.

Mathew, a., Kurane, I., Green, S., Stephens, H. a., Vaughn, D. W., Kalayanarooj, S., Suntayakorn, S., Chandanayingyong, D., Ennis, F. a., and Rothman, a. L. (1998). Predominance of HLA-restricted cytotoxic T-lymphocyte responses to serotype-cross-reactive epitopes on nonstructural proteins following natural secondary dengue virus infection. *Journal of virology*, 72(5):3999–4004.

McElroy, K. L. (2011). Endurance, Refuge, and Reemergence of Dengue Virus Type 2, Puerto Rico, 1986–2007. *Emerging Infectious Diseases*, 17(1):1–4.

McKane, A. J. and Newman, T. J. (2005). Predator-Prey Cycles from Resonant Amplification of Demographic Stochasticity. *Physical Review Letters*, 94(21):1–4.

Messer, W. B., Gubler, D. J., Harris, E., Sivananthan, K., and de Silva, A. M. (2003). Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerging infectious diseases*, 9(7):800–9.

Minayev, P. and Ferguson, N. M. (2009). Incorporating demographic stochasticity into multi-strain epidemic models: application to influenza A. *Journal of The Royal Society Interface*, 6(40):989–96.

Miralles, R., Gerrish, P. J., Moya, A., and Elena, S. F. (1999). Clonal interference and the evolution of RNA viruses. *Science*, 285(5434):1745–7.

Morrison, A. C., Minnick, S. L., Rocha, C., Forshey, B. M., Stoddard, S. T., Getis, A., Focks, D. a., Russell, K. L., Olson, J. G., Blair, P. J., Watts, D. M., Sihuincha, M., Scott, T. W., and Kochel, T. J. (2010). Epidemiology of dengue virus in iquitos, peru 1999 to 2005: interepidemic and epidemic patterns of transmission. *PLoS neglected tropical diseases*, 4(5):e670.

Myat Thu, H., Lowry, K., Jiang, L., Hlaing, T., Holmes, E. C., and Aaskov, J. (2005). Lineage extinction and replacement in dengue type 1 virus populations are due to stochastic events rather than to natural selection. *Virology*, 336(2):163–72.

Nagao, Y. and Koelle, K. (2008). Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever. *Proceedings of the National Academy of Sciences of the United States of America*, 105(6):2238–43.

Nagao, Y., Svasti, P., Tawatsin, a., and Thavara, U. (2008). Geographical structure of dengue transmission and its determinants in Thailand. *Epidemiology and infection*, 136(6):843–51.

Newton, E. a. and Reiter, P. (1992). A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics. *The American journal of tropical medicine and hygiene*, 47(6):709–20.

Nisalak, A., Endy, T. P., Nimmannitya, S., Kalayanarooj, S., Thisayakorn, U., Scott, R. M., Burke, D. S., Hoke, C. H., Innis, B. L., and Vaughn, D. W. (2003). Serotype-specific dengue virus circulation and dengue disease in Bangkok, Thailand from 1973 to 1999. *American Journal of Tropical Medicine and Hygiene*, 68(2):191–202.

Nobusawa, E. and Sato, K. (2006). Comparison of the mutation rates of human influenza A and B viruses. *Journal of virology*, 80(7):3675–3678.

Oki, M. and Yamamoto, T. (2012). Climate Change, Population Immunity, and Hyperendemicity in the Transmission Threshold of Dengue. *PLoS ONE*, 7(10):e48258.

Ooi, E. E., Goh, K. T., and Gubler, D. J. (2006). Dengue prevention and 35 years of vector control in Singapore. *Emerging infectious diseases*, 12(6):887–93.

Otto, S. P. and Day, T. (2007). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Number 6. Princeton University Press.

Page, D. R. and Holmes, C. E. (1998). *Molecular Evolution - A Phylogenetic Approach*. Blackwell Publishing, i edition.

Pastorino, B., Nougairède, A., Wurtz, N., Gould, E., and Lamballerie, X. D. (2010). Role of host cell factors in flavivirus infection : Implications for pathogenesis and development of antiviral drugs. *Antiviral Research*, 87(3):281–294.

Paupy, C., Delatte, H., Bagny, L., Corbel, V., and Fontenille, D. (2009). Aedes albopictus, an arbovirus vector: from the darkness to the light. *Microbes and infection*, 11(14-15):1177–85.

Pitzer, V. E., Patel, M. M., Lopman, B. A., Viboud, C., Parashar, U. D., and Grenfell, B. T. (2011). Modeling rotavirus strain dynamics in developed countries to understand the potential impact of vaccination on genotype distributions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48):19353–8.

Rabaa, M. a., Ty Hang, V. T., Wills, B., Farrar, J., Simmons, C. P., and Holmes, E. C. (2010). Phylogeography of Recently Emerged DENV-2 in Southern Viet Nam. *PLoS neglected tropical diseases*, 4(7):e766.

Raghwani, J., Rambaut, A., Holmes, E. C., Hang, V. T., Hien, T. T., Farrar, J., Wills, B., Lennon, N. J., Birren, B. W., Henn, M. R., and Simmons, C. P. (2011). Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission. *PLoS pathogens*, 7(6):e1002064.

Recker, M., Blyuss, K. B., Simmons, C. P., Hien, T. T., Wills, B., Farrar, J., and Gupta, S. (2009). Immunological serotype interactions and their effect on the epidemiological pattern of dengue. *Proceedings of the Royal Society B*, 276(1667):2541–8.

Recker, M., Nee, S., Bull, P. C., Kinyanjui, S., Marsh, K., Newbold, C., and Gupta, S. (2004). Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature*, 429(6991):555–8.

Rico-Hesse, R. (1990). Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology*, 174(2):479–93.

Rico-Hesse, R. (2010). Dengue virus virulence and transmission determinants. *Current topics in microbiology and immunology*, 338:45–55.

Rico-Hesse, R., Harrison, L. M., Salas, R. a., Tovar, D., Nisalak, a., Ramos, C., Boshell, J., de Mesa, M. T., Nogueira, R. M., and da Rosa, a. T. (1997). Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology*, 230(2):244–51.

Rogers, J. (1994). Central moments and probability distribution of Colless's coefficient of tree imbalance. *Evolution*, 48(6):2026–2036.

Rogers, J. S. (1996). Central Moments and Probability Distributions of Three Measures of Phylogenetic Tree Imbalance. *Systematic Biology*, 45(1):99.

Rothman, A. and Ennis, F. (1999). Immunopathogenesis of Dengue hemorrhagic fever. *Virology*, 257(1):1.

Sabin, A. B. (1952). Research on dengue during World War II. *The American journal of tropical medicine and hygiene*, 1(1):30–50.

Salje, H., Lessler, J., Endy, T. P., Curriero, F. C., Gibbons, R. V., Nisalak, A., Nimmannitya, S., Kalayanarooj, S., Jarman, R. G., Thomas, S. J., Burke, D. S., and Cummings, D. a. T. (2012). Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24):9535–8.

San Martín, J. L., Brathwaite, O., Zambrano, B., Solórzano, J. O., Bouckenooghe, A., Dayan, G. H., and Guzmán, M. G. (2010). The epidemiology of dengue in the americas over the last three decades: a worrisome reality. *The American journal of tropical medicine and hygiene*, 82(1):128–35.

Sangkawibha, N., Rojanasuphot, S., Ahandrik, S., Viriyapongse, S., Jatanasen, S., Salitul, V., Phanthumachinda, B., and Halstead, S. B. (1984). Risk factors in dengue shock syndrome: a prospective epidemiologic study in Rayong, Thailand. I. The 1980 outbreak. *American journal of epidemiology*, 120(5):653–69.

Schmitz, J., Roehrig, J., Barrett, A., and Hombach, J. (2011). Next generation dengue vaccines: a review of candidates in preclinical development. *Vaccine*, 29(42):7276–84.

Scott, T. W., Morrison, a. C., Lorenz, L. H., Clark, G. G., Strickman, D., Kittayapong, P., Zhou, H., and Edman, J. D. (2000). Longitudinal studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: population dynamics. *Journal of medical entomology*, 37(1):77–88.

Shresta, S., Sharar, K. L., Prigozhin, D. M., Beatty, P. R., and Harris, E. (2006). Murine model for dengue virus-induced lethal disease with increased vascular permeability. *Journal of virology*, 80(20):10208–17.

Sierra, B., García, G., Pérez, A. B., Morier, L., Rodríguez, R., Alvarez, M., and Guzmán, M. G. (2002). Long-term memory cellular immune response to dengue virus after a natural primary infection. *International Journal of Infectious Diseases*, 6(2):125–128.

Simonis, J. (2012). Demographic stochasticity reduces the synchronizing effect of dispersal in predator-prey metapopulations. *Ecology*, 93(7):1517–1524.

Sittisombut, N., Sistayanarain, a., Cardosa, M. J., Salminen, M., Damrongdachakul, S., Kalayanarooj, S., Rojanasuphot, S., Supawadee, J., and Maneekarn, N. (1997). Possible occurrence of a genetic bottleneck in dengue serotype 2 viruses between the 1980 and 1987 epidemic seasons in Bangkok, Thailand. *The American journal of tropical medicine and hygiene*, 57(1):100–8.

Spiegel, J. M., Dharamsi, S., Wasan, K. M., Yassi, A., Singer, B., Hotez, P. J., Hanson, C., and Bundy, D. a. P. (2010). Which New Approaches to Tackling Neglected Tropical Diseases Show Promise? *PLoS Medicine*, 7(5):e1000255.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 595(3):585–595.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28(10):2731–9.

Thai, K. T. D., Cazelles, B., Nguyen, N. V., Vo, L. T., Boni, M. F., Farrar, J., Simmons, C. P., van Doorn, H. R., and de Vries, P. J. (2010). Dengue dynamics in Binh Thuan

province, southern Vietnam: periodicity, synchronicity and climate variability. *PLoS Neglected Tropical Diseases*, 4(7):e747.

Thai, K. T. D., Nishiura, H., Hoang, P. L., Tran, N. T. T., Phan, G. T., Le, H. Q., Tran, B. Q., Nguyen, N. V., and de Vries, P. J. (2011). Age-Specificity of Clinical Dengue during Primary and Secondary Infections. *PLoS neglected tropical diseases*, 5(6):e1180.

Thein, S., Aung, M. M., Shwe, T. N., Aye, M., Zaw, A., Aye, K., Aye, K. M., and Aaskov, J. (1997). Risk factors in dengue shock syndrome. *The American journal of tropical medicine and hygiene*, 56(5):566–72.

Tricou, V., Minh, N. N., Farrar, J., Tran, H. T., and Simmons, C. P. (2011). Kinetics of viremia and NS1 antigenemia are shaped by immune status and virus serotype in adults with dengue. *PLoS neglected tropical diseases*, 5(9):e1309.

Trpis, M. and Hausermann, W. (1986). Dispersal and other population parameters of Aedes aegypti in an African village and their possible significance in epidemiology of vector-borne diseases. *American Journal of Tropical Medicine and Hygiene*, 35(6):1263.

Tsuzuki, A., Vu, T. D., Higa, Y., Nguyen, T. Y., and Takagi, M. (2009). High potential risk of dengue transmission during the hot-dry season in Nha Trang City, Vietnam. *Acta tropica*, 111(3):325–9.

Twiddy, S. (2002). Phylogenetic Relationships and Differential Selection Pressures among Genotypes of Dengue-2 Virus. *Virology*, 298(1):63–72.

Twiddy, S. S. (2003). Inferring the Rate and Time-Scale of Dengue Virus Evolution. *Molecular Biology and Evolution*, 20(1):122–129.

Twiddy, S. S., Woelk, C. H., and Holmes, E. C. (2002). Phylogenetic evidence for adaptive evolution of dengue viruses in nature. *The Journal of general virology*, 83(Pt 7):1679–89.

Ubol, S. and Halstead, S. B. (2010). How innate immune mechanisms contribute to antibody-enhanced viral infections. *Clinical and vaccine immunology*, 17(12):1829–35.

Vasilakis, N., Holmes, E. C., Fokam, E. B., Faye, O., Diallo, M., Sall, A. a., and Weaver, S. C. (2007). Evolutionary processes among sylvatic dengue type 2 viruses. *Journal of virology*, 81(17):9591–5.

Vaughn, D. W. (2000). Invited commentary: Dengue lessons from Cuba. *American journal of epidemiology*, 152(9):800–3.

Vaughn, D. W., Green, S., Kalayanarooj, S., Innis, B. L., Nimmannitya, S., Suntayakorn, S., Endy, T. P., Raengsakulrach, B., Rothman, A. L., Ennis, F. A., and Nisalak, A. (2000). Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. *The Journal of infectious diseases*, 181(1):2–9.

Veeraseatakul, P., Wongchompoo, B., Thichak, S., Yananto, Y., Waneesorn, J., and Chutipongvivate, S. (2007). Circulation of dengue serotypes in five provinces of northern Thailand during 2002-2006. *Dengue Bulletin*, 31.

Verdasca, J., Telo da Gama, M. M., Nunes, A., Bernardino, N. R., Pacheco, J. M., and Gomes, M. C. (2005). Recurrent epidemics in small world networks. *Journal of theoretical biology*, 233(4):553–61.

Villabona-Arenas, C. J. and Zanotto, P. M. D. A. (2011). Evolutionary history of dengue virus type 4: insights into genotype phylodynamics. *Infection, Genetics and Evolution*, 11(5):878–85.

Watts, D. M., Burke, D. S., Harrison, B. A., Whitmire, R. E., and Nisalak, A. (1987). Effect of temperature on the vector efficiency of Aedes aegypti for dengue 2 virus. *The American journal of tropical medicine and hygiene*, 36(1):143.

Wearing, H. J. and Rohani, P. (2006). Ecological and immunological determinants of dengue epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, 103(31):11802–7.

Weaver, S. C. and Vasilakis, N. (2009). Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infection, Genetics and Evolution*, 9(4):523–40.

White, L. J., Waris, M., Cane, P. A., Nokes, D. J., and Medley, G. F. (2005). The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. *Epidemiology and infection*, 133(2):279–89.

Whitehorn, J. and Simmons, C. P. (2011). The pathogenesis of dengue. *Vaccine*.

Wichmann, O., Yoon, I.-K., Vong, S., Limkittikul, K., Gibbons, R. V., Mammen, M. P., Ly, S., Buchy, P., Sirivichayakul, C., Buathong, R., Huy, R., Letson, G. W., and Sabchareon, A. (2011). Dengue in Thailand and Cambodia: An Assessment of the Degree of Underrecognized Disease Burden Based on Reported Cases. *PLoS Neglected Tropical Diseases*, 5(3):e996.

Wikramaratna, P., Simmons, C., Gupta, S., Recker, M., and Schneider, B. (2010). The Effects of Tertiary and Quaternary Infections on the Epidemiology of Dengue. *PloS ONE*, 5(8):172–180.

Wilder-Smith, A., Chen, L. H., Massad, E., and Wilson, M. E. (2009). Threat of dengue to blood safety in dengue-endemic countries. *Emerging infectious diseases*, 15(1):8–11.

Williams, C. R., Bader, C. a., Kearney, M. R., Ritchie, S. a., and Russell, R. C. (2010). The extinction of dengue through natural vulnerability of its vectors. *PLoS neglected tropical diseases*, 4(12):e922.

Wittke, V. (2002). Extinction and Rapid Emergence of Strains of Dengue 3 Virus during an Interepidemic Period. *Virology*, 301(1):148–156.

Woelk, C. and Holmes, E. (2002). Reduced positive selection in vector-borne RNA viruses. *Molecular biology and evolution*, 19(12):2333.

Wrammert, J., Onlamoon, N., Akondy, R. S., Perng, G. C., Polsrila, K., Chandele, A., Kwissa, M., Pulendran, B., Wilson, P., Wittawatmongkol, O., Yoksan, S., Angkasekwinai, N., Pattanapanyasat, K., Chokephaibulkit, K., and Ahmed, R. (2012). Rapid and massive virus specific plasmablast responses during acute dengue virus infection in humans. *Journal of virology*.

Zanotto, P. M., Gould, E. a., Gao, G. F., Harvey, P. H., and Holmes, E. C. (1996). Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America*, 93(2):548–53.

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