Dynamics of Dengue epidemics in Southeast Asia

**Authors:**

**Affiliation:**

**Abstract**

The circulation of four antigenically distinct serotypes of Dengue viruses complicate the development and implementation of vaccination in humans, leaving eradication of mosquito vectors as the main mode of control. Despite intensive mosquito control measures, Singapore has regularly observed dengue with periodic occurrences of epidemic Dengue, indicating a lack of understanding in the epidemiology of Dengue. Through comprehensive genomic analysis of DENV collected globally, we show that S that the diversity of Dengue in Singapore is determined by....

**Introduction**

Dengue is a major public health burden affecting over 3 million people annually in more than 100 endemic countries (1-6), with approximately 40 million people reported annually with dengue infection. The primary vector for dengue is the *Aedes* species of mosquito, which aids in the spread of the virus due to its global abundance, as well as its ability to feed during the day (7-18). DENV is endemic to many regions, including Africa, South and Southeast Asia, the Pacific, and the northern regions of Australia (7, 8, 13, 14, 16, 19-25). Due to factors such as population growth, urbanisation, and the continuous increase in global travel, greater numbers of the population are at risk as a consequence of the upsurge of vectors carrying dengue virus (DENV), and their ability to circulate in new and endemic areas (6, 26-29). There is a need for a safe and effective vaccine in order to combat and prevent the morbidity and mortality resulting from DENV infection, however the development of such a vaccine has been a challenge due to the possibility of antibody-dependent-enhancement, and differences in antigenicity between circulating strains (5, 30-32). The most advanced vaccine is only recommended in highly endemic settings, specifically in population groups with >70% Dengue seroprevalence (33, 34).

Symptoms of dengue infections in humans can range from asymptomatic, or mild illness, to dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS), which can both be fatal if medical attention is not sought immediately (3, 4). Of the 390-million dengue infections estimated to occur annually, only approximately 50-100 million of those cases are reported (6, 35). Many dengue cases are asymptomatic, or cause mild illness, and as such, do not get reported (3, 4). This makes determining the cost of dengue extremely difficult. Even with strong dengue surveillance systems in place, many infections can go unreported or may be misdiagnosed. However, the Break Dengue campaign has estimated a global cost of dengue infection to be around US$8 billion (36). This includes costs for individuals who miss work or school, despite not presenting to doctors with dengue infection.

Dengue virus (DENV) belongs to the genus *Flavivirus*, and of all the arboviruses within this genus – which includes major arthropod-borne human viruses, including Japanese encephalitis, West Nile, Yellow Fever and Zika virus – DENV causes the highest incidence of morbidity and mortality globally (7, 16, 19, 24, 37). Flaviviruses are small (40-65 nm in diameter), enveloped viruses, that have a compact RNA genome (38). The genome contains a single-stranded, positive sense RNA molecule (approximately 10,700 nucleotides in length). The DENV genome contains a single open reading frame (ORF) (23, 38-42) that are flanked by the 5’ and 3’ untranslated regions (UTRs) (23, 38-40, 43), and encodes 10 proteins; three structural proteins, known as the core (C) protein, the membrane (M) protein, and the envelope (E) protein; as well as seven non-structural proteins. The 5’ UTR are relatively short (approximately 100 nts) compared to other viruses within this family (38).

Dengue infections are caused by four closely related but distinct serotypes of dengue virus (DENV-1, DENV-2, DENV-3, and DENV-4) (7, 9-11, 15, 19, 37, 44). Each of the serotypes can be further classified into 1 to 4/5 genotypes (or 'subtypes'), however endemic circulation among humans has only been observed for DENV1 genotypes I-III, DENV2 genotypes I-III, DENV3 genotypes I-II and DENV4 genotypes I-II. In regions where Dengue is endemic, such as Southeast Asia, all four serotypes are commonly observed (4, 26, 45), however genotypes are usually spatially segregated resulting in the circulation of one predominant genotype within the area. In some instances, genotype replacement will occur in endemic and hyperendemic countries (3-6, 46). The mechanisms behind genotype replacement are largely unknown (26). It is thought that selection pressure could play a role, as well as potential genetic bottlenecks of virus populations that have circulated for years and are no longer able to transmit within a population due to immune resistance (26). It is likely that a number of factors underpin genotype replacement, relating to environment, interactions between the host and the virus, as well as the increase of global travel and transport (26).

Singapore is an island city-state in Southeast Asia that has a population of over 5.6 million and a large transient population (45). The government of Singapore have designated strict regulations regarding public health measures, with the aim to prevent disease in the city (45). Control measures are in place including surveillance and warning systems for potential epidemics, spraying of powerful insecticides (known as fogging), and public education programs, which have aided in the control of dengue infections in Singapore. However, dengue infections are still recorded every year, with large epidemics appearing sporadically over time (45). This could be due to the genotypes that are endemic to Singapore switching or replacing clades, resulting in new viral epitopes that are not recognised by immune cells that have previously acquired memory of other dengue viruses (32). As all four dengue serotypes circulate in Singapore, it is considered to be a hyperendemic country in regards to dengue prevalence, as are many other Southeast Asian countries (4, 26, 45).

Intensive dengue surveillance provides an early warning for the impending epidemic, identifying introduction of lineages, and development of endemicity -– which in turn aid in the elimination of mosquitoes. Mosquito control programs have been the main approach to Dengue control in large cities, such as Singapore, and has played a significant role in stopping endemic transmission of Dengue (refs). However, Singapore has experienced an increasing trend of severe epidemics since the 1980s with continuous detection through the years (average infection) despite progresses in control strategies (Jit), indicating a lack of understanding of the epidemiological and evolutionary processes that enable the success of Dengue lineages (ref: Egger et al. Bull World Health Organ, 86 2008 187-196, Jit The Lancet Infectious Diseases).

It is hypothesised that selection pressures could be playing a role in genotype replacement of DENV serotypes circulating in Singapore, resulting in sporadic outbreaks throughout the city in previous years. To understand why dengue remains endemic to Singapore despite the strict control measures in place, we performed phylogenetic studies alongside tests for selection pressure for each of the Dengue serotypes, with specific interest in those genotypes prevalent in Singapore. We aimed to determine the serotypes and genotypes endemic to Singapore, and understand whether any of the sites within these sequences were undergoing positive selective pressure. We also wanted to determine whether genotype replacement had occurred within the serotypes circulating in Singapore.



Results

Genopytic diversity of DENV in Singapore

Since 2005, through multiple surveillance systems in Singapore, including the Early DENgue infection and outcome study (EDEN) (ref) and the Eliminate DENgue project >3000 DENV viruses have been isolated and subtyped envelope protein gene (env) sequences which are routinely generated for DENV genotyping and >xxx whole genomes have been generated (Figure 1). We genotyped and conducted during xxxx and xxxx, and newly generated xxx genomes. We assembled the whole genome seqeunces of all four serotypes (DENV-1=xxxx; 2=xxxx; 3=xxxx; 4=xxxx, including xxxx whole genomes from Singapsore. Separtely, we assembled the envelope protein (*env*) gene sequence dataset contsing >3000 *Env* gene sequences from Singapore and combined these with sequences collected globally DENV-1 (3,924, 1,657 Singapore isolates), DENV-2 (3,967, 1,264 Singapore isolates), DENV-3 (2,139, 319 Singapore isolates), and DENV-4 (1,428, 48 Singapore isolates).

To infer which of the dengue genotypes are responsible for outbreaks in Singapore, envelope (E) sequences taken from human sera from each of the four Dengue virus serotypes were obtained from GenBank and analysed using IQ-TREE to construct individual maximum likelihood phylogenetic trees.

**Phylogeny New Results:**

Maximum likelihood phylogenetic trees were constructed from the envelope sequences of DENV-1 (3,924, 1,657 Singapore isolates), DENV-2 (3,967, 1,264 Singapore isolates), DENV-3 (2,139, 319 Singapore isolates), and DENV-4 (1,428, 48 Singapore isolates). Dengue serotypes, which were downloaded from GenBank, and aligned using MUSCLE algorithm in Geneious v11. The overall topology of each of the phylogenetic trees produced from this analysis shows between 3-6 distinct genotypes within each DENV serotype.

Observation of the collated phylogenetic results show that there was an increase of Singapore isolates during 2013-2014, which were mainly comprised of Dengue viruses from DENV-1 and DENV-2. The number of DENV-3 isolates also increased during 2013, however there were not as many isolates from this serotype compared to DENV-1 and -2. Viruses isolated in Singapore from DENV-4 remained at very low levels. These results were supported by the MCC trees, which depict monophyletic clades of Singapore virus within these years for each serotype.

DENV-1

As shown in Figure 1, there are five recognised, distinct genotypes (I-V) within the DENV-1 serotype, as well as a lineage isolated from Brunei. Of these genotypes, three were found to comprise Singapore isolates.

Genotype III represents another lineage that has been considered a sylvatic genotype, previously shown to predominantly consist of Malaysian sequences (29, 46, 54). There are numerous Singapore sequences (coloured in red, Figure 1), which form monophyletic clades within this lineage. The majority of these Singaporean sequences were isolated in 2013 and 2014. Other sequences in this genotype have been isolated from China, India, and areas of Southeast Asia between 1962-2015.

The overall topology reveals genotypes IV and I split from a major node in the phylogenetic tree (Figure 1). Genotype IV is a small genotype comprised of 313 sequences. There are only four introductions of this genotype to Singapore shown in this phylogenetic tree, two of which occurred in 2011, clustering with Indonesian strains from 2002-2010, one in 2012 and one in 2013, which clustered together alongside sequences from the Philippines and China from 2014 and 2010, respectively. The majority of viruses in genotype IV were isolated from New Caledonia, French Polynesia, and Australia.

The largest genotype within the DENV-1 serotype is genotype I. The viruses in this genotype are isolated from areas in Southeast Asia, including Malaysia, Indonesia, and Thailand. Other Asian strains include those isolated from Japan, China, Viet Nam, and Taiwan. Many introductions to Singapore can be observed (Figure 1), the majority of which occurred in 2013 and 2014. These Singapore sequences are interspersed throughout the lineage, and cluster with DENV-1 isolates from China, Thailand, Myanmar, Malaysia, and Indonesia. Genotype I viruses were also isolated from Russia, Germany, Ireland, and Australia.

DENV-2

As shown in Figure 2, the phylogenetic tree can be delineated into six distinct genotypes. These genotypes were determined to be Sylvatic, American, Asia II, Asia I, Asia/America, and Cosmopolitan (29, 31, 52).

The Asia II genotype consists of 665 sequences, seven of which originated from Singapore. This genotype consists of many clusters of strains from Thailand, Myanmar, Cambodia, Australia, Laos, Viet Nam, and China (multiple individual introductions to China, before small clusters in 2013, 2014). Of the Singaporean sequences found in the Asia II genotype, there were two strains clustered together from 2014, and one strain isolated from 2008. The 2014 sequences clustered with strains from Myanmar isolated in 2013-2015, whilst the 2008 sequence clustered with Thailand strains isolated in 2001 and 2002. There was an individual introduction into Singapore in 2006, shown to be clustering with strains from Myanmar, China, and Thailand. In 2011, another individual introduction into Singapore was observed clustering with viruses sequenced from Thailand from 2010. A second small cluster in Singapore consists of two viruses from 2011 and 2012, and was found to cluster with Thailand isolates from 2011, and Myanmar isolates from 2015.

The final genotype in the Dengue serotype 2 viruses is the Cosmopolitan genotype. There were a total of 1,990 DENV-2 viruses sequenced from this genotype. It has been shown that the lineages within this genotype were derived from Indonesian strains, and have since spread throughout Asia and some countries in South America (29). Singaporean strains of DENV-2 are interspersed throughout this lineage and form four sub-lineages (Figure 2), and can also be seen to cluster within other clades of viruses. Small clusters of DENV-2 viruses sequenced from Singapore can be observed clustering with viruses from India, Bangladesh, China and Malaysia, Indonesia and Malaysia, and the Philippines from 2010 to 2014. The introduction of the DENV-2 strains from Indonesian and Malaysian viruses that circulated between 2005-2010 went on to form a monophyletic sub-lineage, which would appear to be comprised of viruses endemic to Singapore. This sub-lineage of Singaporean DENV-2 isolates shows sequences characterised from 2006 until 2013. A second sub-lineage is comprised of viruses sequenced from countries such as Singapore, China, and Malaysia, all appearing to originate from a single sequence from Guam detected in 2001. This second sub-lineage is comprised of Singaporean sequences from the early 2000’s, right up until 2014, with a single introduction into China in 2015. The results show that viruses from multiple sub-lineages were circulating in Singapore between 2012-2014. Prior to this, only one monophyletic clade likely existed, with viruses circulating in Singapore between 2000-2010.

DENV-3

Three distinct genotypes can be observed, and were determined to be Genotype II, Genotype I, and Genotype III, in descending order of lineages observed in the phylogenetic tree (Figure 3) (52, 55). Although two other DENV3 genotypes (Genotypes IV and V) have been characterised (52, 55), we did not observe them in this phylogenetic analysis. This is likely due to the viruses in this phylogenetic analysis being retrieved from human isolates, not from mosquitoes.

Genotype II is comprised of 546 viruses, and strains from this genotype are known to predominantly circulate in areas such as Thailand, Bangladesh, and Viet Nam (52), which was supported by this phylogenetic analysis, showing viruses belonging to this genotype originating from all of those areas, as well as India and Cambodia. A small number DENV-3 sequences from Singapore were also found to cluster within this genotype, dating from 2008 - 2013, and generally clustering with contemporary strains from Bangladesh, Myanmar, and Thailand isolated between the years of 2006 - 2014. A single Singapore sequence that had been isolated in 1995 was seen to cluster with viruses isolated from Thailand in the early 1980’s. Most of the Singapore isolates sequenced from Genotype II showed only individual introductions, or very small clusters of virus.

Genotype I is comprised of 355 DENV-3 viruses from Indonesia, the Philippines, and Singapore, amongst others. Figure 3 shows multiple clusters of DENV-3 viruses isolated from Singapore, as well as multiple individual isolates representing single introductions. The Singapore sequences in this genotype appear to cluster mainly with contemporary viruses from Indonesia, with some clustering in close proximity to Chinese and Australian strains.

Genotype III is the largest genotype, and is known to be the most successful DENV-3 genotype in terms of prevalence and global spread (55). In this analysis, 1,238 DENV-3 Genotype III viruses were included, having been isolated from Sri Lanka, Singapore, the USA, Nicaragua, Venezuela, Brazil, and other regions of the Americas and Asia. There are three sub-lineages observed. Within this genotype, two large, monophyletic Singaporean clades are shown. Unlike Genotype I, the viruses in Genotype III are not interspersed throughout the genotype, but are mainly restricted to two different locations, each within one of the sub-lineages. These clades also show differing times of isolation. The upper monophyletic clade has isolates that have been circulating Singapore from 2004 until 2014, where the lower monophyletic clade consists of isolates from 2013 and 2014. All of the viruses from Singapore mainly cluster alongside viruses from India and China, isolated during the years from 2008 to 2013.

DENV-4

From observing the topology of the tree in Figure 4, there would appear to be five distinct genotypes. However, there are only four characterised genotypes discussed in literature (52). Observed in descending order shown in the phylogenetic tree (Figure 4), these are Genotype IV, Genotype III, Genotype IIa, Genotype I, and Genotype IIb. Note that I have designated the ‘a’ and ‘b’ subtypes to Genotype II in order to succinctly describe phylogeny within this genotype.

Genotype I mainly consists of viruses sequenced countries within Southeast Asia. Clusters of dengue viruses from Thailand, Cambodia, and Viet Nam form the first sub-lineage of this genotype, with a single introduction into Brazil, as well as an introduction into Singapore in 2011, which is seen to cluster with viruses isolated from Viet Nam in 2011. The next sub-lineage consists of viruses from Thailand during the 1990’s and early 2000’s, as well as a Myanmar cluster of sequences from 2013-2015. A single introduction into Singapore during 2014 appears to cluster with these viruses in Myanmar, alongside two viruses sequenced in Thailand during 2013. The final sub-lineage in this genotype is mainly comprised of dengue virus sequences from Thailand, with clusters of DENV-4 from Cambodia, and small clusters from China and Indonesia. A single Singaporean isolate from 2004 can be seen to cluster within Thailand viruses from 2005 and 2006, and a second Singaporean isolate from 2009 can be found within a Cambodia cluster of viruses from 2001-2013.

The final lineage in this genotype is a second group of genotype II viruses, comprised of 959 DENV-4 sequences isolated from countries including China, Indonesia, the Philippines, New Caledonia, French Polynesia, Singapore, and the Americas. Single introductions into Singapore occurred in 2001, 2005, 2010, 2013, and 2014, generally clustering with viruses sequenced from Indonesia, Thailand, the Philippines, and Malaysia at around the same times. Small clusters of DENV-4 sequences isolated from Singapore also appear in the years of 2010-2011, 2008-2010, 2011-2014, and 2011-2016, all clustering with sequences from Indonesia. A large Singapore cluster can also be observed from 2013

Molecular Clock Analysis

Each of the subsampled genotypes underwent Bayesian molecular clock and skyride analysis to determine the rate of nucleotide substitution, as well as time to most recent common ancestor (TMRCA). Unfortunately, the DENV-1 subsamples dataset did not reach convergence (assessed in Tracer v.1.6). Thus, the results of this analysis could not be considered to be reliable, and further or repeated analysis could not be performed on this dataset due to time constraints.

DENV-2

The Cosmopolitan genotype of DENV-2 was subsampled for molecular clock analysis, as it contained the majority of Singapore isolates. The maximum clade credibility (MCC) tree can be seen in Figure 7. In comparison to the DENV-2 phylogenetic tree in Figure 2, similar topology can be observed in the subsampled MCC tree. The mean rate of nucleotide substitution per site for this dataset is 1.17 x 10-3. The TMRCA for this DENV-2 lineage was estimated to be 1968, 48 years before the most recent sample in 2016, with a 95% highest posterior density (HPD) interval of between 45 – 52 years (1964 – 1971). The oldest Singapore sequences in this MCC tree were isolated in 2007. The majority of Singapore isolates formed monophyletic clades within the tree (Figure 7), with single isolates and small clusters generally clustering alongside Indonesian and other Southeast Asian sequences. The posterior probability values for each of the clades in Figure 7 are high, supporting the clustering pattern of the majority of sequences subsampled from DENV-2. An obvious association between genetic distance and time of sampling can also be observed in the MCC tree, suggesting a very strong temporal structure in this data.

Figure 8 depicts the Bayesian skyride plot of the same DENV-2 dataset. An increase in relative genetic diversity of these virus isolates can be seen to begin in the early 1990’s, before a peak in 1994. Following a minor bottleneck, another peak in genetic diversity arises in 2004 followed by a period of sustained decrease in diversity.

DENV-3

The subsampled population of DENV-3 virus isolates did not quite meet convergence following Bayesian analysis, however, it was close enough that it could still be considered useful for this project.

Despite not meeting convergence (Tree model root height ESS was 162), a maximum clade credibility tree and Bayesian skyride plot were generated from this DENV-3 subsampled dataset due to the results appearing reliable (assessed in Tracer v1.6.0) (Figures 9 and 10). The mean rate of nucleotide substitution for this dataset was determined to be 9.95 x 10-4 per site. The TMRCA for this DENV-3 dataset is approximated to be 1961, 54 years before the most recent sample was isolate. The 95% HPD interval estimates TMRCA to be between 52 and 57 years ago (1958-1963). The MCC tree in Figure 9 shows Singapore viruses are interspersed throughout the tree, with two monophyletic clades forming at around the year 2000, and another in 2011-2013. High posterior probability values (0.71-1.0) at the majority of the nodes indicate full resolution of the branches, and a high degree of confidence at each of the nodes

.

The resultant plot from the Bayesian skyride analysis performed on this dataset reveals the relative genetic diversity of this sample of DENV-3 viruses has slowly increased over time since the early 1960’s (Figure 10). Virus divergence was at its peak in 2002/2003. This coincides with the results shown in the MCC tree for the first monophyletic clade from Singapore forming during the earl 2000’s (Figure 9).

DENV-4

Sequences from Genotype IIb were subsampled from the DENV-4 envelope dataset, due to the predominance of Singapore isolates within this genotype. Singapore isolates within the time-based maximum clade credibility tree (Figure 9) can be observed in three clusters, as well as single isolates that are interspersed within the tree. Two of the clusters are seen to be forming monophyletic lineages. The mean rate of nucleotide substitutions for these sequences was estimated to be 1.61 x 10-3 per site, per year. This estimate of evolutionary rates for DENV-4 is similar to estimates shown by Sasmono *et. al*. (2015), who estimated DENV-4 to be evolving at between 1.5 x 10-4 and -0.8 x 10-3 substitutions per site per year (56). However, other studies have shown evolutionary rates of approximately 7.8 x 10-4 substitutions per site per year (57). The TMRCA for the 333 DENV-4 envelope sequences analysed was estimated to be 1962, 53 years before the most recent sample was taken (95% HPD 1959-1966). High posterior probability values have been applied to the MCC tree (Figure 9), supporting the clustering pattern observed at the majority of the nodes.

To infer the relative rate of genetic diversity within these samples, a Bayesian skyride model was used. Figure 12 shows the viruses have steadily increased in genetic divergence since the late 1970’s, continuing until around 2009. Since 2009, genetic diversity has continued to decrease, however, there is a sign that the most recent samples are beginning to diverge again, with a slight upturn seen towards the end of plot.

Selection Pressure Analysis

Whole Genome

To determine whether sites within the full genomes of Dengue serotypes I-IV were undergoing negative purifying selection or positive diversifying selection, selection pressure analyses were performed using the public webserver, Datamonkey (<http://datamonkey.org>). The total number of sites within each genome undergoing selection pressure is measured by the test statistic dN/dS, where dN is the number of non-synonymous substitutions per non-synonymous site, and dS is the number of synonymous substitutions per synonymous site (58, 59). Four different methods available on the Datamonkey public web server were utilized for this analysis. These were Single Likelihood Ancestor Counting (SLAC), Internal Fixed-Effects Likelihood (IFEL), Mixed Effects Model of Evolution (MEME), and Fast Unconstrained Bayesian AppRoximation (FUBAR).

The results for each of these methods are depicted in Table 2. The majority of Dengue viruses analysed from each of the Dengue serotypes were shown to be undergoing negative or purifying selection, which has also been shown in previous studies (6, 11, 51, 55, 57, 60). This result is determined by the non-synonymous/synonymous nucleotide substitution ratio test statistic, which indicates whether a codon within an alignment has undergone positive, negative, or neutral selection (58, 59). SLAC is the most conservative method of each of the four selection pressure methods used to analyse the envelope sequences from Dengue serotypes 1-4. As a result of this, we considered those sites determined by SLAC to be undergoing positive selection to be of more interest than those determined as undergoing positive selection using the other methods mentioned. Each of the sites selected as undergoing positive selection in SLAC were also shown to be undergoing positive selection in IFEL, MEME, and FUBAR. One site in each of DENV-1 and DENV-4 were found to be undergoing positive selection, while DENV-2 was determined to have 6 sites undergoing positive selection. No sites were found to be undergoing positive selection in DENV-3.

**Table 2.** Total number of sites suggested to be undergoing positive selection within randomly down-sampled full genome sequences of Dengue viruses 1-4. Each sequence was analysed using four immune pressure analysis methods; Single-Likelihood Ancestor Counting (SLAC), Internal Fixed-Effects Likelihood (IFEL), Mixed Effects Method of Evolution (MEME), and Fast Unconstrained Bayesian AppRoximation (FUBAR), available at <http://datamonkey.org>.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method DENV 1 DENV 2 DENV 3 DENV 4** | | | | |
| **SLAC**  **IFEL**  **MEME**  **FUBAR** | 1  4  17  2 | 6  10  40  10 | 0  1  9  0 | 1  3  13  1 |

Envelope Sequences

As mentioned, SLAC is the most conservative selection pressure method available on Datamonkey. Considering this, and the Envelope proteins’ interaction with immune responses during infection, we decided to only use SLAC analysis of the envelope sequences subsampled by either Singapore prominent clades, or by clades that are predominantly made up of other Southeast Asian countries’ sequences. The results from these analyses of DENV-1 to DENV-4 are shown in Table 3. Positive selection was found at 1 site for the Singapore dominant clades in DENV-1 and DENV-3. No positive selection was observed for DENV-2 and DENV-4. The Southeast Asian clades also showed one site undergoing positive selection in each of DENV-1 and DENV-3, whilst DENV-2 and DENV-4 also showed no positive selection for these clades. As has been shown previously (6, 11, 51, 55, 57, 60), the majority of Dengue virus sequences within these datasets appear to be undergoing purifying selection.

**Table 3.** Total number of positively selected sites within envelope sequences of Dengue viruses 1-4 that had been specifically subsampled to clades predominantly consisting of viruses from either Singapore or from other areas of Southeast Asia. Each sequence was analysed using the Single-Likelihood Ancestor Counting (SLAC) method of determining selection pressure available at <http://datamonkey.org>.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **DENV 1 DENV 2 DENV 3 DENV 4** | | | | |
| **Singapore** | 1 | 0 | 1 | 0 |
| **SE Asia** | 1 | 0 | 1 | 0 |

Methods

Sequence Datasets

Whole genome and the complete envelope (Env) sequences and were downloaded for each of the four serotypes of dengue virus using the publically available GenBank database (<https://www.ncbi.nlm.nih.gov/nuccore)> (Data available until June 2017). Only those isolated from human sera and containing known date (year) and geographical location of isolation were included, with total the number of Env and whole genome sequences for each serotype shown in Table 1. The final data sets spanned a temporal range of 57 years (1960-2017), covering a total of 71 countries (data available until June 2017). Multiple sequence alignments were carried out using MUSCLE (47, 48).

**Table 1.** Total number of whole genome and envelope sequences downloaded for each of the four serotypes of dengue virus (DENV) from the GenBank database (https://www.ncbi.nlm.nih.gov/nuccore).

|  |  |  |
| --- | --- | --- |
| **Serotype Number of Sequences** | | |
| **Whole Genome Envelope** | | | |
| DENV 1  DENV 2  DENV 3  DENV 4 | | 1928  1455  998  379 | 5980  4237  2321  1592 |

Phylogenetic Analysis

Maximum likelihood phylogenetic analysis was conducted for each of the datasets using the best-fit nucleotide substitution model (Table 1) in IQ-TREE version 1.5.5 (49). Branch support was estimated using ultrafast bootstrapping (50). Each analysis was repeated thrice to confirm convergence. The phylogenetic trees were visualised using FigTree Version 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Dengue genotypes were delineated through the observation of each virus sample date and origin, and then comparing these with the previous literature (1, 11, 39, 51, 52) and through the metadata published for each viral entry, which is available on the GenBank database.

Molecular Clock and Skyride Analysis

Envelope sequences from DENV 1–4 were subsampled into genotypes that contained clades rich with Dengue sequences and had been isolated from Singapore. Each of these sequences were then thoroughly analysed in Geneious to insure the alignments held no stop codons or misalignments, and maximum likelihood trees were inferred using RAxML v7 under the GTR+G+I (general time-reversible model with gamma-distributed rates of variation among sites and a proportion of invariable sites) nucleotide substitution model. TempEst v1.4 was used to plot the root-to-tip genetic distances between these sequences. Any sequences that did not conform to a linear evolutionary pattern were discarded.. Once it was determined that all of the molecular clock assumptions were met for each of the subsampled populations, they were submitted for molecular clock analysis in BEAST v1.8.4 (53). The uncorrelated relaxed lognormal (UCLD) clock was used alongside the Hasegawa-Kishino-Yano (HKY) nucleotide substitution model with estimated base frequencies and Gamma site heterogeneity to observe evolutionary rates within each of the Dengue clades. Gaussian Markov Random Field (GMRF) priors were also set to produce a smoothed skyride plot from the data. Each run was originally set up to have a chain length of 200,000,000, however this was restricted to 100,000,000 generations with sampling every 10,000 generations due to time constraints. Convergence was assessed using the program Tracer v1.6.0 (<http://tree.bio.ed.ac.uk/software/tracer/>), with an effective sample size (ESS) of at least 200 after the first 10% of chain lengths were discarded at burn-in. The Maximum Clade Credibility (MCC) Trees were generated using TreeAnnotator version 1.7.5 removing the initial 10% of trees as burn-in. The time-ordered maximum clade credibility trees were viewed in FigTree version 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree). Bayesian GMRF skyride plots were performed in Tracer v1.6.0 (<http://tree.bio.ed.ac.uk/software/tracer/>). As a measure of the degree of uncertainty for each of the parameters of the MCC tree, a 95% highest posterior density (HPD) value was applied. A similar measure was applied to the skyride plots as the upper and lower value to the median. All of the Beast output files were assessed for convergence based on an effective sampling size of approximately 200.

Selection Pressure

Positive and negative selection pressures of each codon within the Singapore-rich whole genome sequences were measured as the ratio of synonymous (dS) to non-synonymous (dN) substitutions per site. Selection pressure within each of the samples was determined using selection pressure models available through the Datamonkey web server (<http://datamonkey.org>), including Single-Likelihood Ancestor Counting (SLAC), internal Fixed Effects Likelihood (IFEL), the Mixed Effects Model of Evolution (MEME), and Fast, Unconstrained Bayesian AppRoximation (FUBAR).

To determine selection pressure on E genes from each of the Dengue serotypes, genotypes containing either a majority of Singaporean Dengue strains, or genotypes with strains predominantly from Thailand/Viet Nam/China were subsampled. These genotypes were then down sampled into smaller, representative clades containing <400 sequences so that SLAC analysis could be performed.

**References**

1. Cecilia, D., J. A. Patil, M. B. Kakade, A. Walimbe, K. Alagarasu, B. Anukumar, and A. Abraham. 2017. Emergence of the Asian genotype of DENV-1 in South India. *Virology* 510: 40-45.

2. Bennett, S. N., E. C. Holmes, M. Chirivella, D. M. Rodriguez, M. Beltran, V. Vorndam, D. J. Gubler, and W. O. McMillan. 2003. Selection-driven evolution of emergent dengue virus. *Mol Biol Evol* 20: 1650-1658.

3. Carmo, A. M., R. B. Suzuki, A. D. Cabral, R. T. Costa, G. P. Massari, M. M. Riquena, H. A. Fracasso, A. Eterovic, A. Marcili, and M. A. Speranca. 2017. Co-circulating serotypes in a dengue fever outbreak: Differential hematological profiles and phylogenetic relationships among viruses. *J Clin Virol* 90: 7-13.

4. Choudhary, M. C., E. Gupta, S. Sharma, N. Hasnain, and P. Agarwala. 2017. Genetic signatures coupled with lineage shift characterise endemic evolution of Dengue virus serotype 2 during 2015 outbreak in Delhi, India. *Trop Med Int Health* 22: 871-880.

5. Cruz, C. D., A. Torre, G. Troncos, L. Lambrechts, and M. Leguia. 2016. Targeted full-genome amplification and sequencing of dengue virus types 1-4 from South America. *J Virol Methods* 235: 158-167.

6. Dash, P. K., S. Sharma, M. Soni, A. Agarwal, A. K. Sahni, and M. Parida. 2015. Complete genome sequencing and evolutionary phylogeography analysis of Indian isolates of Dengue virus type 1. *Virus Res* 195: 124-134.

7. Asghar, F., N. Hussain, F. Komal, and Q. Fatima. 2016. Sero-prevalence of Dengue IgM Antibodies in Patients Suspected of having Dengue Fever. *Annual Research & Review in Biology* 11: 1-10.

8. Ashby, J., M. Moreno-Madriñán, C. Yiannoutsos, and A. Stanforth. 2017. Niche Modeling of Dengue Fever Using Remotely Sensed Environmental Factors and Boosted Regression Trees. *Remote Sensing* 9: 328.

9. Chu, H., S. L. George, D. T. Stinchcomb, J. E. Osorio, and C. D. Partidos. 2015. CD8+ T-cell Responses in Flavivirus-Naive Individuals Following Immunization with a Live-Attenuated Tetravalent Dengue Vaccine Candidate. *J Infect Dis* 212: 1618-1628.

10. Holbrook, M. R. 2017. Historical Perspectives on Flavivirus Research. *Viruses* 9.

11. Holmes, E. C. 2006. The evolutionary biology of dengue virus. In *New treatment strategies for dengue and other flaviviral diseases*. Novartis Foundation Symposium Wiley, Chichester. 177-192.

12. Jordan, T. X., and G. Randall. 2017. Dengue Virus Activates the AMP Kinase-mTOR Axis To Stimulate a Proviral Lipophagy. *J Virol* 91.

13. Jupille, H., A. Vega-Rua, F. Rougeon, and A. B. Failloux. 2014. Arboviruses: variations on an ancient theme. *Future Virol* 9: 733-751.

14. Mishra, A., and S. Gakkhar. 2017. A micro-epidemic model for primary dengue infection. *Communications in Nonlinear Science and Numerical Simulation* 47: 426-437.

15. Olagnier, D., D. Amatore, L. Castiello, M. Ferrari, E. Palermo, M. S. Diamond, A. T. Palamara, and J. Hiscott. 2016. Dengue Virus Immunopathogenesis: Lessons Applicable to the Emergence of Zika Virus. *J Mol Biol* 428: 3429-3448.

16. Zhang, B., G. B. Salieb-Beugelaar, M. M. Nigo, M. Weidmann, and P. Hunziker. 2015. Diagnosing dengue virus infection: rapid tests and the role of micro/nanotechnologies. *Nanomedicine* 11: 1745-1761.

17. Behura, S. K., and D. W. Severson. 2013. Overlapping genes of Aedes aegypti: evolutionary implications from comparison with orthologs of Anopheles gambiae and other insects. *BMC Evol Biol* 13: 124.

18. Yakob, L., and T. Walker. 2016. Zika virus outbreak in the Americas: the need for novel mosquito control methods. *The Lancet Global Health* 4: e148-e149.

19. Andrade, C. C., K. I. Young, W. L. Johnson, M. E. Villa, C. A. Buraczyk, W. B. Messer, and K. A. Hanley. 2016. Rise and fall of vector infectivity during sequential strain displacements by mosquito-borne dengue virus. *J Evol Biol* 29: 2205-2218.

20. Guzman, M. G., S. B. Halstead, H. Artsob, P. Buchy, J. Farrar, D. J. Gubler, E. Hunsperger, A. Kroeger, H. S. Margolis, E. Martinez, M. B. Nathan, J. L. Pelegrino, C. Simmons, S. Yoksan, and R. W. Peeling. 2010. Dengue: a continuing global threat. *Nat Rev Microbiol* 8: S7-16.

21. Nealon, J., A. F. Taurel, M. R. Capeding, N. H. Tran, S. R. Hadinegoro, T. Chotpitayasunondh, C. K. Chong, T. A. Wartel, S. Beucher, C. Frago, A. Moureau, M. Simmerman, T. Laot, M. L'Azou, and A. Bouckenooghe. 2016. Symptomatic Dengue Disease in Five Southeast Asian Countries: Epidemiological Evidence from a Dengue Vaccine Trial. *PLoS Negl Trop Dis* 10: e0004918.

22. Thein, T. L., E. L. Ng, M. S. Yeang, Y. S. Leo, and D. C. Lye. 2015. Risk factors for concurrent bacteremia in adult patients with dengue. *J Microbiol Immunol Infect*.

23. Weaver, S. C., and A. D. Barrett. 2004. Transmission cycles, host range, evolution and emergence of arboviral disease. *Nat Rev Microbiol* 2: 789-801.

24. Yong, Y. K., H. Y. Tan, S. H. Jen, E. M. Shankar, S. K. Natkunam, J. Sathar, R. Manikam, and S. D. Sekaran. 2017. Aberrant monocyte responses predict and characterize dengue virus infection in individuals with severe disease. *J Transl Med* 15: 121.

25. Chia, P. Y., H. S. Yew, H. Ho, A. Chow, S. P. Sadarangani, M. Chan, Y. W. Kam, C. Y. Chong, K. C. Thoon, C. F. Yung, J. H. Li, D. C. Lye, P. P. De, L. F. P. Ng, T. W. Yeo, and Y. S. Leo. 2017. Clinical features of patients with Zika and dengue virus co-infection in Singapore. *J Infect* 74: 611-615.

26. Filomatori, C. V., J. M. Carballeda, S. M. Villordo, S. Aguirre, H. M. Pallares, A. M. Maestre, I. Sanchez-Vargas, C. D. Blair, C. Fabri, M. A. Morales, A. Fernandez-Sesma, and A. V. Gamarnik. 2017. Dengue virus genomic variation associated with mosquito adaptation defines the pattern of viral non-coding RNAs and fitness in human cells. *PLoS Pathog* 13: e1006265.

27. Klungthong, C., R. Putnak, M. P. Mammen, T. Li, and C. Zhang. 2008. Molecular genotyping of dengue viruses by phylogenetic analysis of the sequences of individual genes. *J Virol Methods* 154: 175-181.

28. Waman, V. P., S. M. Kasibhatla, M. M. Kale, and U. Kulkarni-Kale. 2016. Population genomics of dengue virus serotype 4: insights into genetic structure and evolution. *Arch Virol* 161: 2133-2148.

29. Waman, V. P., P. Kolekar, M. R. Ramtirthkar, M. M. Kale, and U. Kulkarni-Kale. 2016. Analysis of genotype diversity and evolution of Dengue virus serotype 2 using complete genomes. *PeerJ* 4: e2326.

30. Cunha Mdos, P., V. N. Guimaraes, M. Souza, D. de Paula Cardoso, T. N. de Almeida, T. S. de Oliveira, and F. S. Fiaccadori. 2016. Phylodynamics of DENV-1 reveals the spatiotemporal co-circulation of two distinct lineages in 2013 and multiple introductions of dengue virus in Goias, Brazil. *Infect Genet Evol* 43: 130-134.

31. Diaz, Y., J. Cisneros, H. Guzman, P. Cordoba, J. P. Carrera, B. Moreno, R. Chen, J. C. Mewa, L. Garcia, L. Cerezo, A. T. da Rosa, N. D. Gundacker, B. Armien, S. C. Weaver, N. Vasilakis, S. Lopez-Verges, and R. Tesh. 2018. The reintroduction of DENV-2 in 2011 in Panama and subsequent outbreak characteristic. *Acta Trop* 177: 58-65.

32. Ramakrishnan, L., M. R. Pillai, and R. R. Nair. 2015. Dengue vaccine development: strategies and challenges. *Viral Immunol* 28: 76-84.

33. Aguiar, M., N. Stollenwerk, and S. B. Halstead. 2016. The risks behind Dengvaxia recommendation. *The Lancet Infectious Diseases* 16: 882-883.

34. Flasche, S., M. Jit, I. Rodriguez-Barraquer, L. Coudeville, M. Recker, K. Koelle, G. Milne, T. J. Hladish, T. A. Perkins, D. A. Cummings, I. Dorigatti, D. J. Laydon, G. Espana, J. Kelso, I. Longini, J. Lourenco, C. A. Pearson, R. C. Reiner, Y. T.-R. L. Mier, K. Vannice, and N. Ferguson. 2016. The Long-Term Safety, Public Health Impact, and Cost-Effectiveness of Routine Vaccination with a Recombinant, Live-Attenuated Dengue Vaccine (Dengvaxia): A Model Comparison Study. *PLoS Med* 13: e1002181.

35. Wei, K., and Y. Li. 2017. Global evolutionary history and spatio-temporal dynamics of dengue virus type 2. *Sci Rep* 7: 45505.

36. Booth, A. 2016. Measuring the true cost of dengue. *Break Dengue* 2018: <https://www.breakdengue.org/measuring-the-true-cost-of-dengue/>.

37. Wikan, N., S. Libsittikul, S. Yoksan, P. Auewarakul, and D. R. Smith. 2015. Delayed antibody dependent enhancement of low passage dengue virus 4 isolates. *BMC Res Notes* 8: 399.

38. Fernandez-Sanles, A., P. Rios-Marco, C. Romero-Lopez, and A. Berzal-Herranz. 2017. Functional Information Stored in the Conserved Structural RNA Domains of Flavivirus Genomes. *Front Microbiol* 8: 546.

39. Dettogni, R. S., and I. D. Louro. 2012. Phylogenetic characterization of Dengue virus type 2 in Espirito Santo, Brazil. *Mol Biol Rep* 39: 71-80.

40. Roundy, C. M., S. R. Azar, S. L. Rossi, S. C. Weaver, and N. Vasilakis. 2017. Insect-Specific Viruses: A Historical Overview and Recent Developments. *Adv Virus Res* 98: 119-146.

41. Watson, A. M., and W. B. Klimstra. 2017. T Cell-Mediated Immunity towards Yellow Fever Virus and Useful Animal Models. *Viruses* 9.

42. Aziz, H., A. Zia, A. Anwer, M. Aziz, S. Fatima, and M. Faheem. 2017. Zika virus: Global health challenge, threat and current situation. *J Med Virol* 89: 943-951.

43. Ahmed, H., and E. A. O'Toole. 2014. Recent advances in the genetics and management of harlequin ichthyosis. *Pediatr Dermatol* 31: 539-546.

44. Rodpothong, P., and P. Auewarakul. 2012. Positive selection sites in the surface genes of dengue virus: phylogenetic analysis of the interserotypic branches of the four serotypes. *Virus Genes* 44: 408-414.

45. Lee, K. S., S. Lo, S. S. Tan, R. Chua, L. K. Tan, H. Xu, and L. C. Ng. 2012. Dengue virus surveillance in Singapore reveals high viral diversity through multiple introductions and in situ evolution. *Infect Genet Evol* 12: 77-85.

46. Wang, B., H. Yang, Y. Feng, H. Zhou, J. Dai, Y. Hu, L. Zhang, Y. Wang, Z. Baloch, and X. Xia. 2016. The distinct distribution and phylogenetic characteristics of dengue virus serotypes/genotypes during the 2013 outbreak in Yunnan, China: Phylogenetic characteristics of 2013 dengue outbreak in Yunnan, China. *Infect Genet Evol* 37: 1-7.

47. Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792-1797.

48. Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies, and A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647-1649.

49. Nguyen, L. T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32: 268-274.

50. Minh, B. Q., M. A. Nguyen, and A. von Haeseler. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 30: 1188-1195.

51. Wang, X., D. Ma, X. Huang, L. Li, D. Li, Y. Zhao, L. Qiu, Y. Pan, J. Chen, J. Xi, X. Shan, and Q. Sun. 2017. Complete genome analysis of dengue virus type 3 isolated from the 2013 dengue outbreak in Yunnan, China. *Virus Res* 238: 164-170.

52. Weaver, S. C., and N. Vasilakis. 2009. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect Genet Evol* 9: 523-540.

53. Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29: 1969-1973.

54. Villabona-Arenas, C. J., and P. M. Zanotto. 2013. Worldwide spread of Dengue virus type 1. *PLoS One* 8: e62649.

55. Sittivicharpinyo, T., P. Wonnapinij, and W. Surat. 2017. Phylogenetic analyses of DENV-3 isolated from field-caught mosquitoes in Thailand. *Virus Res* 244: 27-35.

56. Sasmono, R. T., I. Wahid, H. Trimarsanto, B. Yohan, S. Wahyuni, M. Hertanto, I. Yusuf, H. Mubin, I. J. Ganda, R. Latief, P. J. Bifani, P. Y. Shi, and M. J. Schreiber. 2015. Genomic analysis and growth characteristic of dengue viruses from Makassar, Indonesia. *Infect Genet Evol* 32: 165-177.

57. Costa, R. L., C. M. Voloch, and C. G. Schrago. 2012. Comparative evolutionary epidemiology of dengue virus serotypes. *Infect Genet Evol* 12: 309-314.

58. Arafa, A., I. El-Masry, S. Kholosy, M. K. Hassan, G. Dauphin, J. Lubroth, and Y. J. Makonnen. 2016. Phylodynamics of avian influenza clade 2.2.1 H5N1 viruses in Egypt. *Virol J* 13: 49.

59. Singh, J., C. S. Mukhopadhyay, J. S. Arora, and S. Kaur. 2015. Biocomputational characterization and evolutionary analysis of bubaline dicer1 enzyme. *Asian-Australas J Anim Sci* 28: 876-887.

60. Martin, E., M. Chirivella, J. K. G. Co, G. A. Santiago, D. J. Gubler, J. L. Munoz-Jordan, and S. N. Bennett. 2016. Insights into the molecular evolution of Dengue virus type 4 in Puerto Rico over two decades of emergence. *Virus Res* 213: 23-31.