APJAI
Asian Pacific Journal of
Allergy and Immunology

ORIGINAL ARTICLE

Genetic diversity of the dengue virus population in dengue fever and dengue hemorrhagic fever patients

Kwanrutai Chin-inmanu,[1] Dumrong Mairiang,[2,3] Phongphak Khongthon,[4] Duangjai Sangsrakru,[5] Sithichoke Tangphatsornruang,[5] Nattaya Tangthawornchaikul,[2,3] Prida Malasit,[2,3,6] Prapat Suriyaphol[1]

Abstract

Background: The error-prone replication of dengue virus (DENV) in host results in the highly diverse viral population. Together with the host factor, intra-host diversity may influence the disease severity. Therefore, it is worth investigating whether there is a correlation between intra-host genetic diversity and disease severity.

Objective: To investigate the genetic diversity in DENV for four serotypes of the dengue population from patients with dengue fever (DF) and dengue hemorrhagic fever (DHF) using next-generation sequencing (NGS) technology.

Methods: Forty RNA samples categorized into eight groups by severity and serotypes were sequenced and analyzed for genetic variation. Analysis on the hot-cold genomic regions, selection pressure and correlation between genotype and disease severity were performed in this study.

Results: Comparison between the NGS data of the DF and DHF specimens showed conservation between their major populations with the consensus sequences for DF and DHF sharing 99% similarity. However, the minor populations in DF and DHF were more diverse. Many genes in DF had an #NS/#S ratio higher than in DHF. Only NS4B of DENV1 DF has #NS/#S ratio higher than one. Hot regions of the DF were detected in NS3 of DENV1, DENV2 and Envelope of DENV3, whereas the hot regions of the DHF samples were detected in the small region in 3'UTR of DENV2 and DENV3.

Conclusions: Various explorations of the variations of DF and DHF were performed in this study. However, we have not yet found any specific characteristics of intra-host diversity associated with disease severity.

Key words: dengue fever, dengue hemorrhagic fever, minor population, genetic diversity, quasispecies, single nucleotide variation

From:
[1] Division of Bioinformatics and Data Management for Research, Research Group and Research Network Division, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[2] Siriraj Center of Research Excellence in Dengue and Emerging Pathogens, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[3] Molecular Biology of Dengue and Flaviviruses Research Team, Medical Molecular Biotechnology Research Group, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand
[4] Research Group and Research Network Division, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[5] National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand
[6] Division of Dengue Hemorrhagic Fever Research, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

Corresponding author:
Prapat Suriyaphol
Division of Bioinformatics and Data Management for Research, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University,
Room 205, SiMR Building, 2 Wang Lang road, Siriraj, Bangkoknoi, Bangkok 10700, Thaland
E-mail: prapat.sur@mahidol.ac.th

Abbreviations:

DF	Dengue fever
DHF	Dengue hemorrhagic fever
DENV	Dengue virus
DENV1	Dengue virus serotype 1
DENV2	Dengue virus serotype 2
DENV3	Dengue virus serotype 3
DENV4	Dengue virus serotype 4
SNV	Single nucleotide variation
3'UTR	3' Untranslated region
C	Capsid
prM	Pre-membrane

Abbreviations (Continued):

E	Envelope
NS1	Non-structural 1
NS2A	Non-structural 2A
NS2B	Non-structural 2B
NS3	Non-structural 3
NS4A	Non-structural 4A
NS4B	Non-structural 4B
NS5	Non-structural 5
5'UTR	5'Untranslated region

Introduction

Dengue virus is an RNA virus that requires RNA-dependent RNA polymerase (RdRp) for replication. Since the enzyme has an inefficient proofreading activity during copying, the mutation rate of viral RdRp during one round of RNA replication is approximately 10^{-3} to 10^{-5} substitutions per nucleotide.[1] Therefore, each round of viral replication can introduce more variations in the viral genome, thus increasing genetic diversity in the virus population. Consequently, a highly diverse population of related viral genotypes will exist within a single host after multiple rounds of replications. This phenomenon is called "viral quasispecies".[2,3]

The factors involved in the disease severity of dengue virus infection are not completely understood, but are believed to arise from both the host and pathogen.[4] Host factors, such as cytokine storm[5] and antibody-dependent enhancement,[6,7] were usually used to explain pathogenesis in DHF. There is evidence suggesting that the health status and variations in the host affect the severity of the disease.[8-10] Virus factors are also important for disease severity. For example, the American genotype of DENV2 is less virulent than its Southeast Asian counterpart.[11,12]

Another possible virus factor for disease severity is the intra-host diversity of the DENV population. The variations in intra-host virus population can also be derived from selective pressure from the environment, such as the host immune system.[13,14] Taken together, intra-host diversity may influence the virulence and emerging capability of disease. The correlation between intra-host genetic diversity and disease severity is thus worth investigating.

In this study, quasispecies of dengue virus were studied in Thai dengue patients between 2006 and 2009. Four serotypes of dengue virus were extracted from the plasma of DF and DHF patients. The whole dengue viral genome was sequenced by 454 pyrosequencing technology. The population diversity of the dengue virus from the DF and DHF patients was compared and contrasted to investigate the relationship between the genome and the virulence.

Methods

Sample collection and classification

Dengue RNA specimens were obtained from plasma samples collected at two clinical cohort study sites for pediatric dengue infection located at Khon Kaen and Songkhla hospitals in Thailand. In the cohort study, a serotype of dengue virus in every specimen was identified by nested PCR.[15] In addition, the disease severity of every patient was classified as either dengue fever (DF) or dengue hemorrhagic fever (DHF) according to the World Health Organization (WHO) 1997 guideline.[16] To select appropriate specimens from the cohort study, five DF specimens and five DHF specimens were chosen from each serotype. In each serotype group, all the specimens were collected during the same outbreak season (**table 1**). In total, the specimens from 40 patients were used in this study.

Table 1. Details of sequenced dengue samples

	DENV1	DENV2	DENV3	DENV4
Type of infection	Secondary	Secondary	Secondary	Secondary
Pool sequencing				
DF Pool (n)	5	5	5	5
DHF Pool (n)	5	5	5	5
DHF grade II (n)	4	5	2	4
DHF grade III (n)	1	0	3	1
Day of defervescence	-2 and -3	-1	-1	-1
Outbreak season	2006	2007	Aug 2008 to Jul 2009	2006
Site	KK	SK	SK	KK

Abbreviations: KK, Khon Kaen Hospital; SK, Songkhla Hospital

Template preparation

The plasma samples were derived from blood samples collected from clinical cohort mentioned above. The samples were EDTA-treated, aliquoted to microcentrifuge tubes and stored in the -70°C deep freezer until use. Dengue RNA was extracted from the plasma samples by the QIAamp Viral RNA Mini Kit (QIAGEN Inc.), and the concentration and purity of the extracted RNA were determined with a NanoDrop 1000 Spectrophotometer. The 40 RNA samples were grouped by disease severity and virus serotype into eight subgroups. In each subgroup, the RNA samples were then pooled in equal volume and used for preparing dengue DNA template. Two overlapping fragments of dengue virus genome were amplified to cover the whole genome.[17] Finally, eight pooled samples were obtained, consisting of dengue serotype 1 from DF (DENV1DF) and DHF (DENV1DHF) patients, dengue serotype 2 from DF (DENV2DF) and DHF (DENV2DHF) patients, dengue serotype 3 from DF (DENV3DF) and DHF (DENV3DHF) patients, dengue serotype 4 from DF (DENV4DF) and DHF (DENV4DHF) patients, for sequencing.

Massively parallel sequencing

The dengue genome in each sample pool was sequenced by 454 pyrosequencing technology (GS FLX sequencing system, Roche Applied Science). Library preparation, clonal amplification, and sequencing were performed according to the manufacturer's protocols for whole genome shotgun sequencing.

Data QC and variant calling

Sequence reads were processed to trim low-quality bases and remove primers using PRINSEQ (version 0.20.4)[18] and Trimmomatic (version 0.38),[19] respectively (**figure 1A**). Then cleaned reads longer than 40 bases were re-corrected homopolymer errors produced by the sequencing platform using Pollux (version 1.0.2)[20] (**figure 1A**). The cleaned data were aligned to the reference sequences of dengue prototype strains using BWA-MEM algorithm (version 0.7.17),[21] then constructed the consensus sequence using BCFTools version 2.27.1[22] (**figure 1B**). The Genome Analysis Toolkit (version 4.1.1.0)[23] was then used for recalibration of the base quality scores. Finally, SNVs calling was performed using LoFreq (version 2.1.3.1)[24] (**figure 1C**). SNVs with an allele frequency ≥ 1% were selected for the hot and cold region detection and #NS/#S analysis.

Consensus sequence analysis

The consensus sequence represents the main dengue virus population. The consensus sequences were aligned against the non-redundant nucleotide database by web-based BLASTn to find the most similar DENV group reported in Genbank. Then, the consensus sequences from different severities, i.e., DF and DHF, of each serotype were aligned against each other using the BL2SEQ program. The second analysis aimed to compare and contrast between the major virus populations in DF and DHF patients at the same cohort site within the same outbreak season.

Detection of the hot and cold regions on the DENV genome

Hot regions are parts of the genome containing more SNVs than other parts of the genome. Conversely, cold regions are parts of the genome that are relatively more conserved than the other regions of the genome. To identify the hot and cold regions, we followed the method reported by Wilm et al.[24] Briefly, the dengue genome was sliced into small windows for the SNV counting. The windows were then slided with five nucleotides overlapping until covering the whole genome. The window size for counting the SNVs for hot region identification was 20 bases. For cold region identification, the dengue genome was scanned with a minimum size of window of 40 bases. A binomial test followed by Bonferroni correction at a p-value < 0.05 was used to examine the excess or depleted SNVs in the window compared to the whole genome in the hot and cold region detection, respectively.

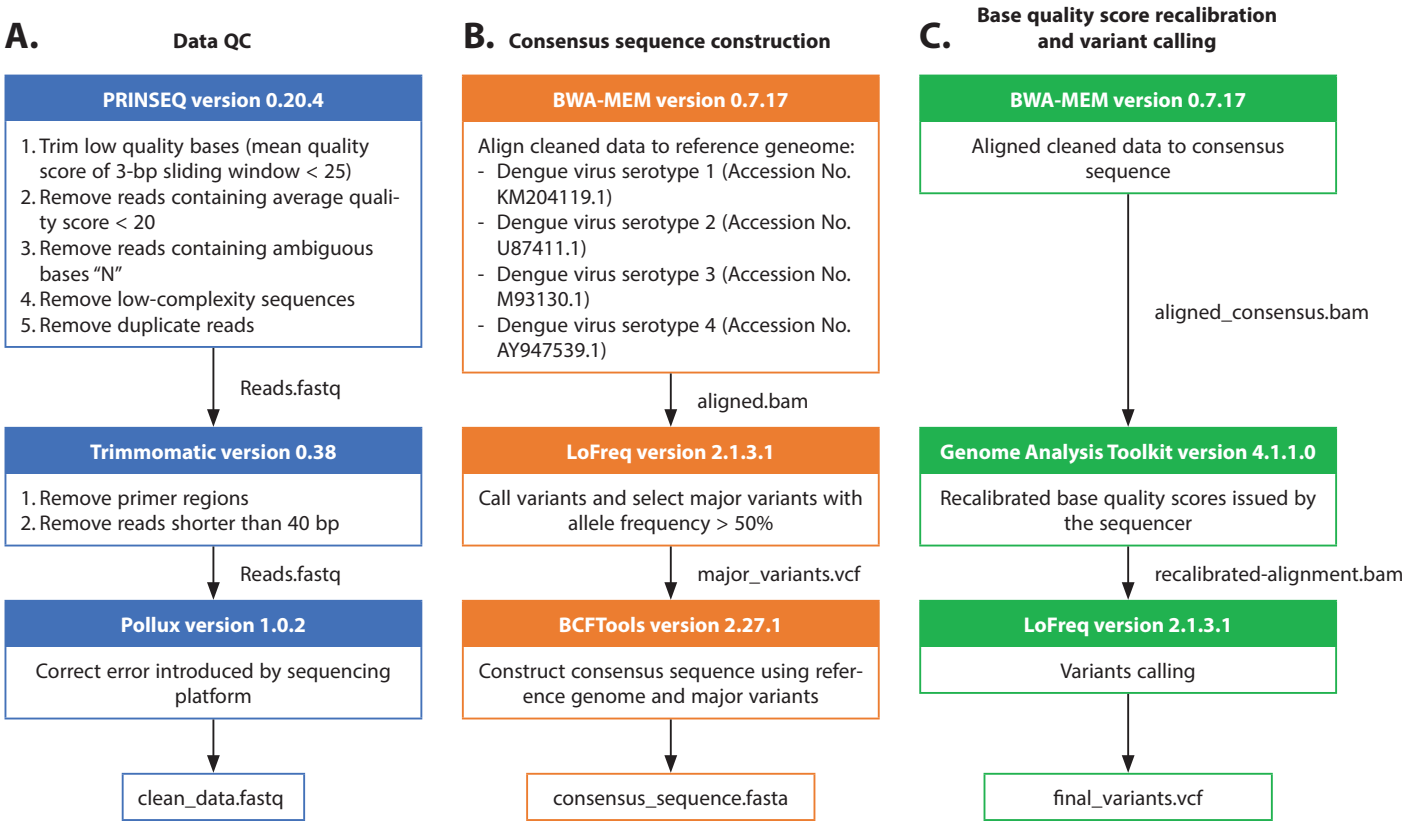


Figure 1. The data QC and variant calling workflow.
(A) Raw sequencing data were cleaned from low quality bases and primer regions using PRINSEQ and Trimmomatic, and homopolymer errors were corrected using and Pollux programs. (B) The cleaned sequencing data were used for assembling the consensus sequence. (C) Cleaned data were aligned to the consensus sequence, and then the base quality scores were recalibrated using BWA-MEM and GATK. Finally, the SNVs were called using LoFreq program.

#NS/#S analysis

For investigating the positive selection pressure on the dengue genome resulting in diverse minor populations, the ratio of the number of non-synonymous to the number of synonymous mutations (#NS/#S) was calculated for each gene.[13] The #NS/#S ratio of each gene was compared between DF and DHF to investigate the selection pressure of the different severities.

Results

The forty samples of each serotype were chosen from the same outbreak season. However, the samples of serotype 3 were selected from two consecutive outbreak seasons from August 2008 to July 2009 because the number of DENV3 patients enrolled were very limited during the time of this study. The samples from DF and DHF (grade II and III) patients were used. The DF and DHF samples of serotypes 2, 3, and 4 were selected from 1 day prior to defervescence. The samples of serotype 1 were selected from 2 to 3 days prior to defervescence. Details of the selected samples are shown in **table 1**.

Consensus sequence comparison

Consensus sequences from the DF and DHF data of each serotype were matched best with the same DENV isolate in the nt database. The best match of the consensus sequences from the DF and DHF data of dengue serotypes 1, 2, 3, and 4 were DENV1 strain ThD1_0102_01, DENV2 strain ThD2_0078_01, DENV3 isolate CNR_15418, and DENV4 strain ThD4_0485_01, respectively. When comparing the consensus sequences between DF and DHF within the same serotype, the nucleotide sequences were almost identical (99%), implying that the major DENV populations in the DF and DHF samples were highly similar. In addition, all of the positions of the consensus sequences differing between the DF and DHF samples still had shared a set of SNVs. In other words, minor SNVs in the DF samples may be major SNVs in the DHF samples and vice versa. Details of the comparison are shown in supplement data.

The consensus nucleotide sequences were translated to amino acid sequences by the ExPASy translate tool.[25] The comparison between the amino acid sequences of DF and those of DHF showed different amino acid positions on the genome. The comparison showed thirteen, nine, five, and six different positions in DENV1, DENV2, DENV3, and DENV4, respectively (**table 2**). There were ten positions in DENV1, DENV2, and DENV3 where the DF and DHF samples had different amino acids along with different chemical groups (R groups) as shown in **table 2**. However, the effect of these SNVs to the fitness of the dengue virus is still unknown and needs to be explored further.

Table 2. Chemical groups of the different amino acids detected from the consensus amino acid sequence comparison between DF and DHF

Sample	Allele		R-Group		Gene	Position on gene
	DF	DHF	DF	DHF		
DENV1	R	G	Positively charged	Nonpolar, aliphatic	Capsid	76*
	S	N	Polar, uncharged	Polar, uncharged	Capsid	90
	I	L	Hydrophobic	Hydrophobic	Capsid	106
	A	T	Hydrophobic	Polar, uncharged	NS1	98*
	Y	H	Hydrophobic	Positively charged	NS1	111*
	A	V	Hydrophobic	Hydrophobic	NS1	213
	K	R	Positively charged	Positively charged	NS2A	29
	P	S	Nonpolar, aliphatic	Polar, uncharged	NS2A	142*
	I	V	Nonpolar, aliphatic	Nonpolar, aliphatic	NS2A	155
	T	I	Polar, uncharged	Nonpolar, aliphatic	NS3	7*
	M	V	Nonpolar, aliphatic	Nonpolar, aliphatic	NS3	323
	K	R	Positively charged	Positively charged	NS3	562
	G	S	Nonpolar, aliphatic	Polar, uncharged	NS5	746

Table 2. (Continued)

Sample	Allele		R-Group		Gene	Position on gene
	DF	DHF	DF	DHF		
DENV2	V	I	Nonpolar, aliphatic	Nonpolar, aliphatic	Capsid	72
	K	R	Positively charged	Positively charged	PrM	57
	I	M	Nonpolar, aliphatic	Nonpolar, aliphatic	NS2A	48
	M	V	Nonpolar, aliphatic	Nonpolar, aliphatic	NS4B	152
	A	V	Nonpolar, aliphatic	Nonpolar, aliphatic	NS4B	181
	M	L	Nonpolar, aliphatic	Nonpolar, aliphatic	NS5	20
	K	Q	Positively charged	Polar, uncharged	NS5	469*
	T	I	Polar, uncharged	Nonpolar, aliphatic	NS5	554*
	Y	H	Aromatics	Positively charged	NS5	556*
DENV3	T	I	Polar, uncharged	Nonpolar, aliphatic	Envelope	82*
	V	I	Nonpolar, aliphatic	Nonpolar, aliphatic	NS2A	61
	S	P	Polar, uncharged	Nonpolar, aliphatic	NS3	107*
	S	T	Polar, uncharged	Polar, uncharged	NS5	231
	V	A	Nonpolar, aliphatic	Nonpolar, aliphatic	NS5	267
DENV4	E	D	Negatively charged	Negatively charged	NS1	92
	S	N	Polar, uncharged	Polar, uncharged	NS2B	93
	K	R	Positively charged	Positively charged	NS3	538
	K	R	Positively charged	Positively charged	NS5	74
	I	V	Nonpolar, aliphatic	Nonpolar, aliphatic	NS5	373
	S	C	Polar, uncharged	Polar, uncharged	NS5	651

*Different chemical groups (R groups)

Diversity comparisons between DF and DHF

The number of SNVs with allele frequency ≥ 1% of each serotype consisted of DENV1 (DF 77 SNVs, DHF 247 SNVs), DENV2 (DF 142 SNVs, DHF 134 SNVs), DENV3 (DF 87 SNVs, DHF 25 SNVs), and DENV4 (DF 85 SNVs, DHF 132 SNVs). The number of synonymous SNVs was higher than the number of nonsynonymous SNVs in the samples for all four serotypes. The SNVs that changed amino acid to stop codon were detected in DENV1 DHF and DENV4 DF. The SNVs consisted of L171* NS1 and S205* NS1 of DENV1 DHF and DENV4 DF, respectively.

Selection pressure investigation of DF and DHF

The #NS/#S ratio of each gene was calculated from the number of nonsynonymous and synonymous SNVs (**table 3**). There was no common pattern for #NS/#S among all four serotypes. An #NS/#S ratio equal to one was detected in many genes, such as capsid, envelope, NS2B, and NS4B of DENV1 DF, DENV2 DF, and DENV4 DF (**figure 2**). In DENV3, both DF and DHF had low #NS/#S ratios. Especially for DHF, only the #NS/#S of NS5 could be calculated, while the other genes did not have nonsynonymous SNVs. All four serotypes had the same trend of a higher #NS/#S level in the DF than

Table 3. Number of nonsynonymous or synonymous SNVs detected in each gene

Sample	NS | S	C	prM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5
DENV1 DF	NS	3	0	7	2	2	0	9	0	2	3
	S	3	0	7	0	3	0	21	0	1	6
DENV1 DHF	NS	3	1	5	8	4	0	11	2	2	12
	S	7	2	22	11	13	5	50	7	12	54

Table 3. (Continued)

Sample	NS | S	C	prM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5
DENV2 DF	NS	3	2	1	1	1	1	11	0	2	7
	S	3	5	14	6	6	1	24	6	6	31
DENV2 DHF	NS	3	2	1	1	1	0	3	0	1	7
	S	4	5	16	3	8	1	16	6	10	33
DENV3 DF	NS	0	0	4	0	1	0	1	0	0	2
	S	2	6	24	4	4	1	9	4	7	13
DENV3 DHF	NS	0	0	0	0	0	0	0	0	0	1
	S	1	2	7	1	2	1	2	0	2	4
DENV4 DF	NS	1	0	1	2	0	1	2	1	1	2
	S	1	3	16	12	7	5	7	5	1	16
DENV4 DHF	NS	1	1	2	2	1	0	2	1	1	4
	S	0	7	21	14	7	5	13	6	8	32

Abbreviations: NS, Nonsynonymous; S, Synonymous

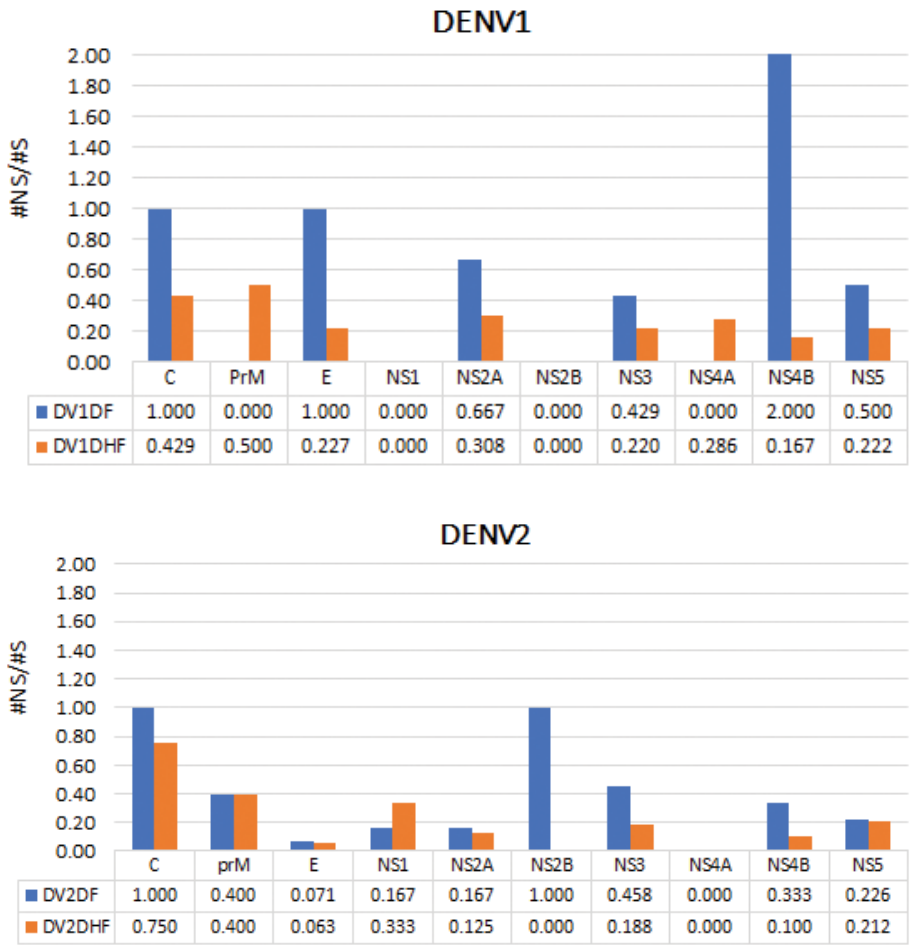


Figure 2. #NS#S ratios of all the genes on the dengue genome. The x-axis is the gene name and the y-axis is the #NS/#S ratio.

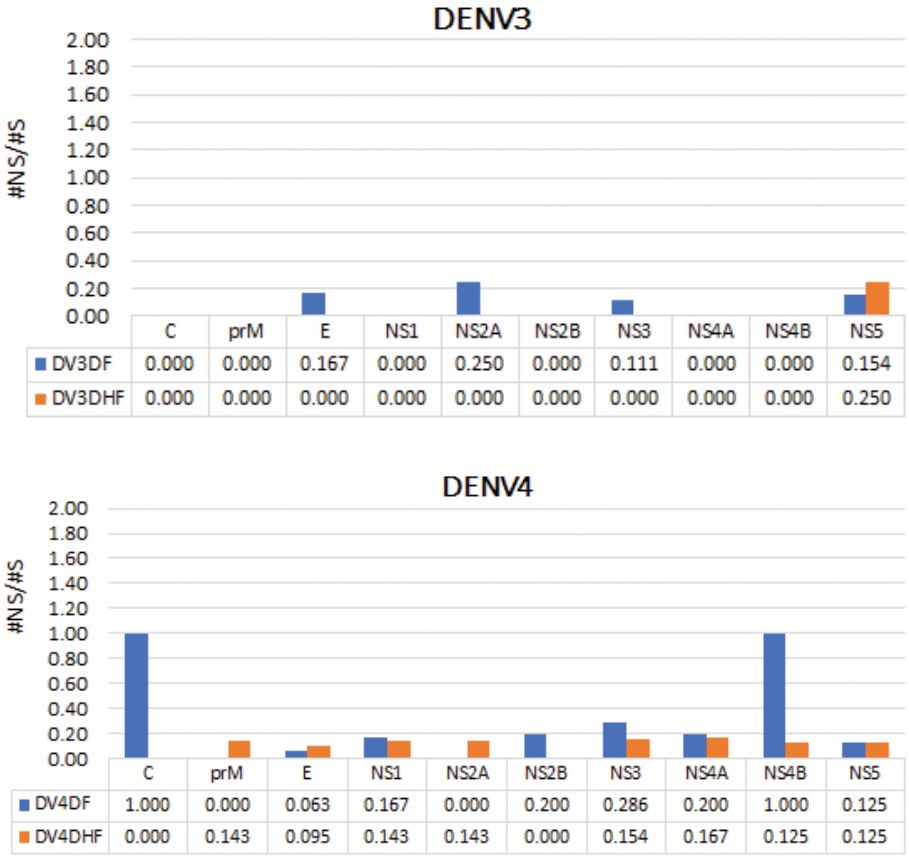


Figure 2. (Continued)

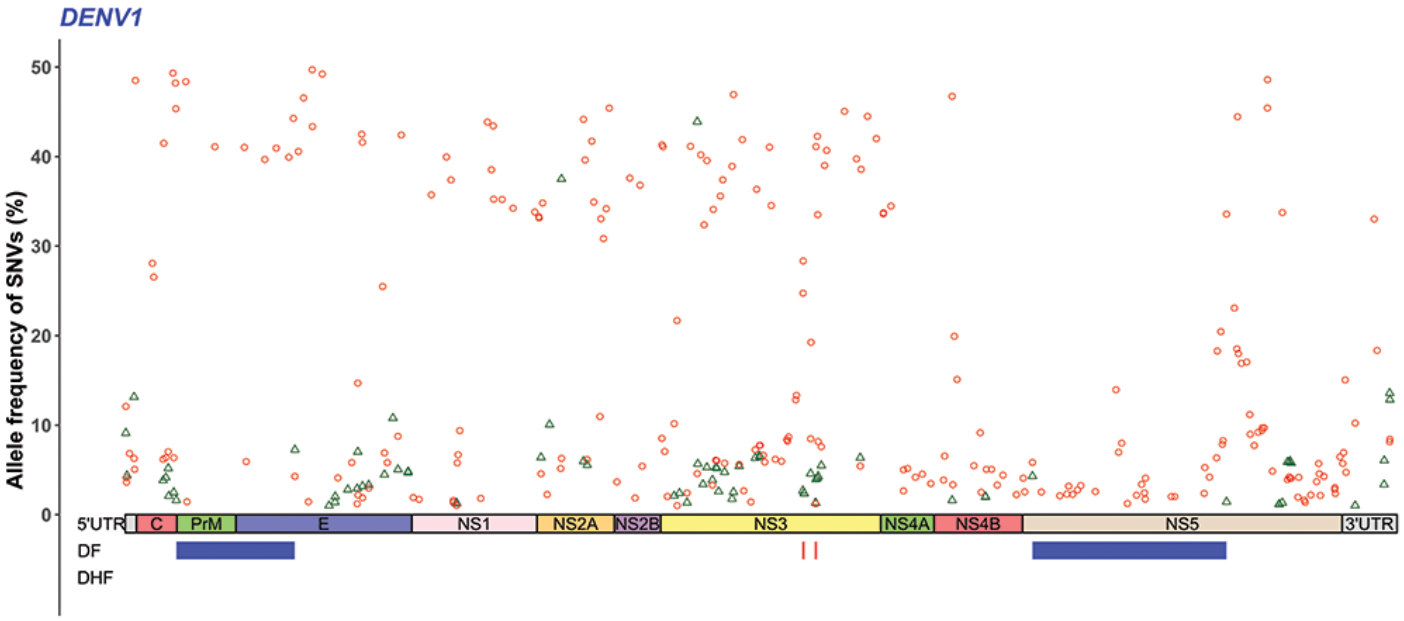


Figure 3. Hot and cold regions on the dengue genome.
Each gene was plotted as a rectangle with a width corresponding to the length of the gene. Rectangles below the genes are the hot and cold regions of DF and DHF, respectively. The red color is the hot region and the blue color is the cold region. The SNVs distribution of each disease severity are also plotted above the genes. The green triangles represent the SNVs of DF and the red circles represent the SNVs of DHF. The y-axis indicates the percent allele frequency of the SNVs.

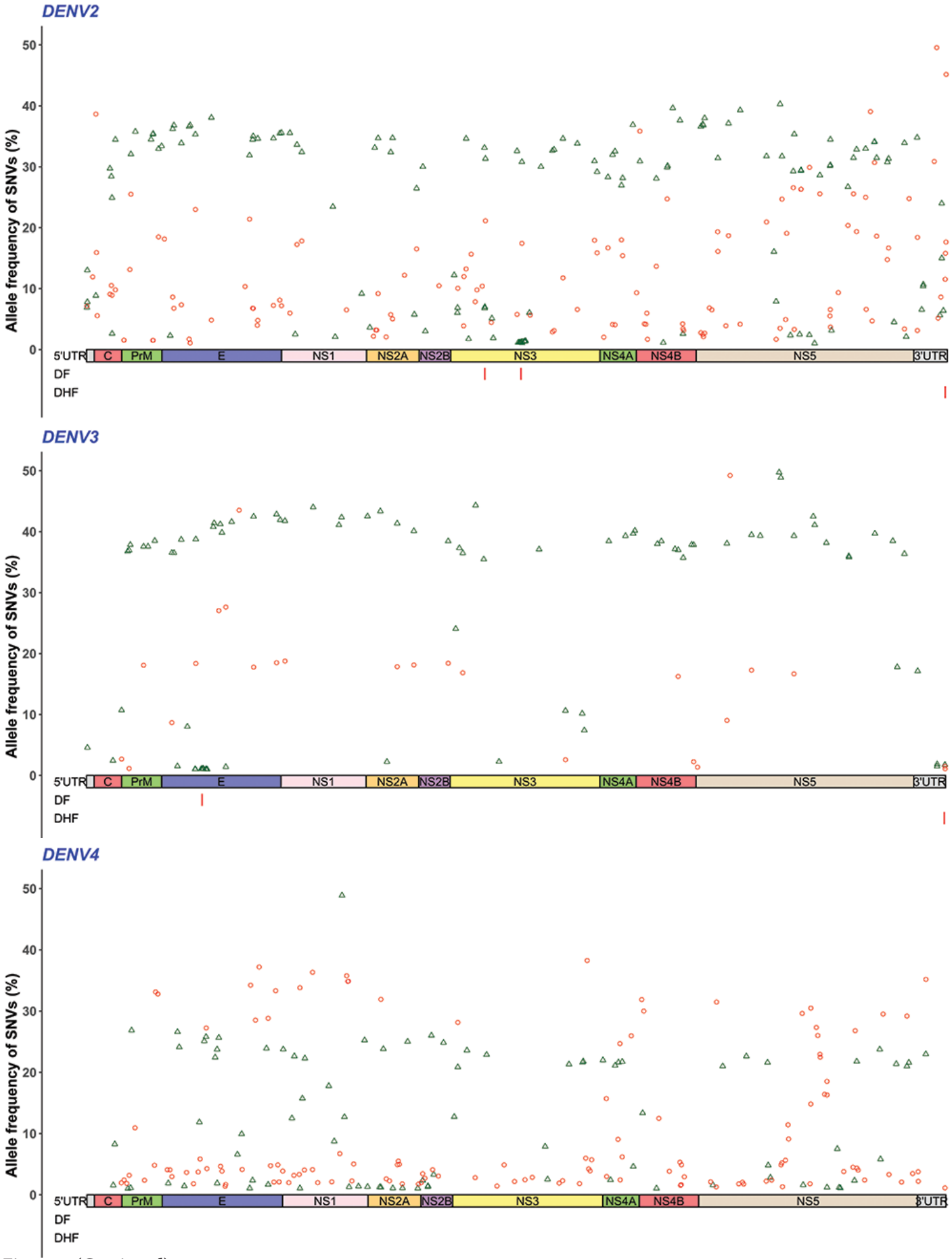


Figure 3. (Continued)

in the DHF sample. However, only NS4B of DENV1 DF has #NS/#S ratio higher than one. The #NS/#S values are shown in **figure 2**.

Hot and cold region analysis

Hot and cold regions were investigated throughout the whole dengue genomes (**figure 3**). For the DHF samples, hot regions could be identified in 3'UTR of DENV2 and DENV3, whereas cold regions could not be found in any of the serotypes. The cold region can be detected only in the prM – envelope and NS5 of DENV1 DF. The hot regions in DF samples were detected in NS3 of DENV1 and DENV2, and envelope of DENV3. Hot and cold regions could not be detected in the DENV4 genome.

Discussion

The quality of NGS data

NGS has a higher error rate than Sanger sequencing. Thus, using data without cleaning may incorrectly identify SNVs from error reads. Many processes for data QC were performed before the variant calling and other analyses. Due to the error generated by the 454 pyrosequencing machine, especially for the homopolymer region, an error correction step was performed to reduce system errors using the Pollux program.[20]

SNV detection

In this study, the number of detected SNVs did not show any trend between the number of SNVs and the severity. However, since the levels of viral load between the DF and DHF samples were not equalized prior to sequencing, the viral load levels could be the factor that may affect the number of SNVs. A higher viral load and more prolonged virus titer were detected in the DHF patients than in the DF patients.[26-28] A higher viral load was suspected to be correlated with a higher risk of bias toward recovering and enriching the minor SNVs.[29] Therefore, a similar amount of viral titer in both samples may need to be considered in order to compare variations of dengue in DF and DHF. Unfortunately, this study did not control for the viral titer between the DF and DHF samples. Hence, the consensus sequences of DF and DHF were mainly used to compare the nucleotide and amino acid variations to avoid the effect from viral titer that may occur. Moreover, sequencing each sample individually might provide more details of viral diversity.

Six SNVs positions detected in this study, consisting of DENV1 DF and DHF D437N NS3; DENV3 DF S124P, H132Y, V169I/A, I380I Envelope; DENV3 DHF I380I Envelope; and DENV4 DHF T329T Envelope, were reported by Durães-Carvalho et al. in 2019. Interestingly, the major and minor alleles of the variants reported by Durães-Carvalho et al. were the opposite from our detected SNVs, especially for the SNVs from the DF samples. The reported variations, consisting of DENV1 D437E NS3; DENV3 P124S/L/F/K, Y132H, T169V/A/H, I380T/V Envelope; and DENV4 A329T Envelope, were reported to be affected by Darwinian selection and may be associated to the virulence and pathogenesis of DENV.[30] However, there have been no reports on DENV virulence according to these variants. Therefore, further laboratory experiments are need to examine this.[30]

There are mutations that change amino acid to stop codon detected at the amino acid positions L171* NS1 and S205* NS1 of DENV1 DHF and DENV4 DF, respectively. The stop codon in the envelope gene of DENV1, which indicates a defective virus due to the quasispecies of dengue virus, has already been reported.[31] However, there are no reports of stop codons detected in the NS1 gene.

Selection pressure

NS4B gene of DENV1 DF has #NS/#S ratio higher than one. This gene may evolve under positive selection, whereby they are affected by selection pressure of the host immune. This gene has been previously reported to have a #NS/#S ratio > 1 in samples from dengue patients.[13] In this study, ratios of #NS/#S were small in all DHF samples. In addition, #NS/#S ratios could not be calculated in some genes due to the lack of synonymous change. Overall, few SNVs were found in this study (**table 3**). Pooling several samples into a sequencing reaction might result in a loss of sample specific SNVs due to a reduced frequency respected to a pool sample and a limited sequencing throughput. Thus, sequencing each sample individually might increase SNVs detection and improve #NS/#S ratio results.

Consensus comparison

The comparison between the consensus nucleotide sequences of DF and DHF showed a high similarity with 99% similarity. These data showed that the major populations of DF and DHF were mostly conserved. All the different positions between the consensus for DF and DHF contained the shared variations. This may imply that even though the majority are different, DF and DHF still contained the same group of dengue population, but in different ratios.

However, small differences in the consensus nucleotide sequences between the DF and DHF samples resulted in some non-synonymous changes. These positions may be linked to the severity. However, since there is no report supporting this yet in the literature, further experiments need to be performed to examine their function. However, the minor DENV populations in DF and DHF patients were apparently different.

Rodriguez-Roche et al. found that a minor DENV population found in patients at the beginning of epidemic season had become a predominant population found in patients by the end of the season.[32] The minor population could influence the drug resistance in HIV.[33] Therefore, the minor population plays a critical role in the evolution of quasispecies as the virus adapts to changes in the selective pressure. However, in this study, it is still unclear whether the difference in minor DENV populations of DF and DHF resulted in the difference in disease severity.

Hot and cold regions on the dengue genome

Hot and cold regions were identified throughout the whole dengue genome. Sim et al. reported hot spots of a mosquito-derived dengue serotype 2 population on NS2B, NS3, and

3'UTR.[13] In this study, we also detected hot spots on NS3 of DENV2 DF and 3'UTR of DENV2 DHF, but in different positions. Only the DF sample of dengue serotype 1 contained a long range of cold spots, which represented the conserved region. This cold region overlapped with the cold spot in DENV2 reported by Sim et al.[13] Hot and cold regions could not be detected in DENV4 samples. The SNVs of DENV4 were dispersed through the whole genome.

Dengue diversity studies using NGS with clinical samples

A few studies had applied NGS with dengue clinical samples to investigate virus diversity and its association with disease severity. Parameswaran et al. compared variations of dengue virus extracted from PBMC and from plasma of 77 DENV3 infected patients. The study found variants consisting of G100S (prM), M101L (prM), and H513L (envelope) which appeared to be immune escape variants.[14] In our study, we also detected the hotspot region at envelope of DENV3DF sample and found a nonsynonymous variant, V169I/A in this hotspot. Rodriguez-Roche et al. studied the role of dengue virus diversity in increasing the disease severity during dengue epidemics of DENV1 followed by DENV3 (DENV1/DENV3). The viral population change occurred during the course of epidemic which could affect viral fitness. The higher genetic variability was observed in the secondary infection cases. Therefore, the evolution of viral populations due to selective pressure of heterotypic antibodies could be related to the increase of severity.[32] We did not find the significant intra-host variations reported by Rodriguez-Roche et al. in our study.

The dengue incidence during 2006 – 2009 in north eastern and southern of Thailand

The annual incidence of dengue infection in north eastern and southern of Thailand reported by Bureau of Epidemiology, Department of Disease Control (DDC), Ministry of Public Health, Thailand during 2006 to 2009 is showed in **supplement figure 1**. DHF had the highest number of cases follow by DF and DSS in both north eastern and southern region in 2006 to 2009 (**supplement figure 1A, 1B**). There was no countrywide epidemiological survey for dengue serotype in Thailand during 2006-2009. However, the epidemiological cohort study of children in primary schools in central Thailand (Ratchaburi province) during 2006-2009 found that DENV1 was predominant for all four years (2006 – 2009) starting from 50%, 34%, 51%, and 38% for 2006, 2007, 2008 and 2009, respectively.[34] The number of dengue cases enrolled to our cohort between 2006-2009 is showed in the **supplement figure 2**. Our cohort study specifically recruited pediatric patients with duration of fever shorter than four days at the time of enrollment so the characteristics of cohort participants, including serology and serotype, were expected to be dissimilar from those of patients in dengue epidemiological reports during the same period (**supplement figure 2A**). In addition, cohort participants were closely monitored by experienced clinicians. Thus, the rate of DSS was expected to be low because of adequate and prompt fluid replacement. The number of cases grouping in each serotype is showed in **supplement figure 2B**. DENV1 was the most common serotype during 2006 – 2009 in both study sites. DENV1 and DENV2 samples were selected from KK in 2006 and SK in 2007 which were the predominant serotype of that time. DENV3 and DENV4 samples were selected from SK in 2008-2009 and KK in 2006 which were the second dominant serotype.

Conclusion

This study explored various aspects of the intra-host diversity of DF and DHF patients. However, we have not yet found specific characteristics of intra-host diversity associated with disease severity.

Acknowledgments

The authors are thankful to Prof.Wasun Chantratita and his colleague at the Virology and Molecular Biology unit, Ramathibodi Hospital, Thailand, for advice on next-generation sequencing. The authors would like to thank Dr.Sirijitt Vasanawathana and team at Pediatric Department, Khon Kaen Hospital and Dr.Wannee Limpitikul and team at Pediatric Department, Songkhla Hospital for dengue sample collection

Conflict of Interests

The authors declare that they have no conflict of interests.

Ethics approval

Specimens used in this study were from a cohort study that was approved by the ethical committee of Faculty of Medicine Siriraj Hospital, Mahidol University (SiEC number: 349/2550) and the Ministry of Public Health, Thailand (Reference Number: 92/2550).

Consent for Publication

Not applicable.

Funding

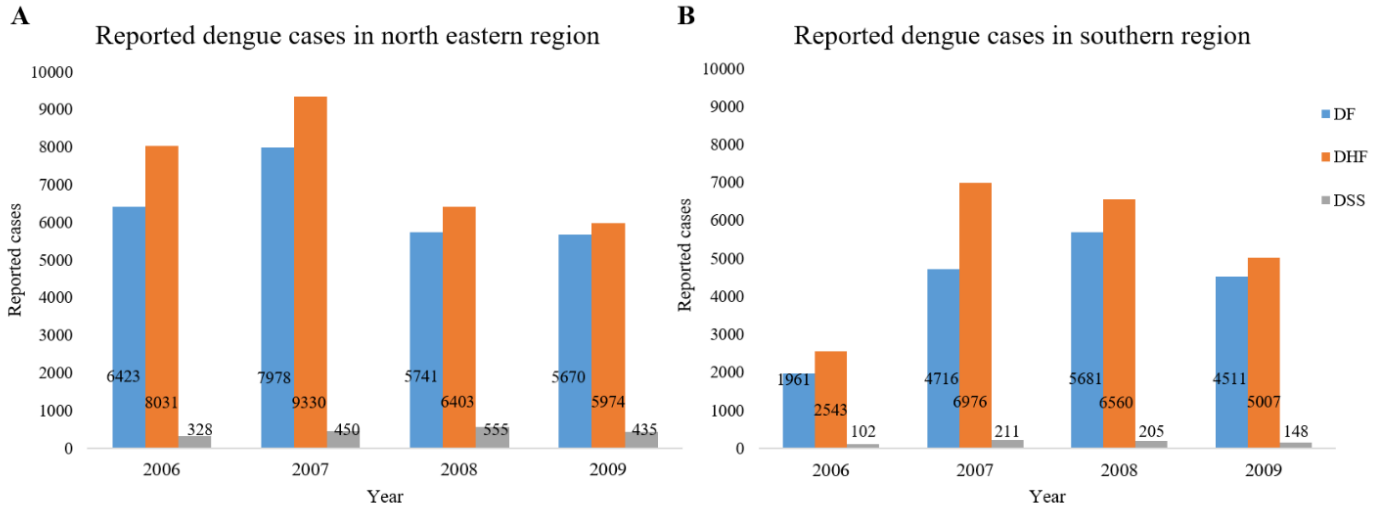
This work was supported by the National Science and Technology Development Agency (NSTDA) grant number P-09-00837, the Office of the Higher Education Commission and Mahidol University under the National Research Universities Initiative, and TRF-office for R&D. PM is supported by NSTDA Chair Professor Grant. PS is supported by a Siriraj Chalermprakiat Grant.

Authors' Contributions

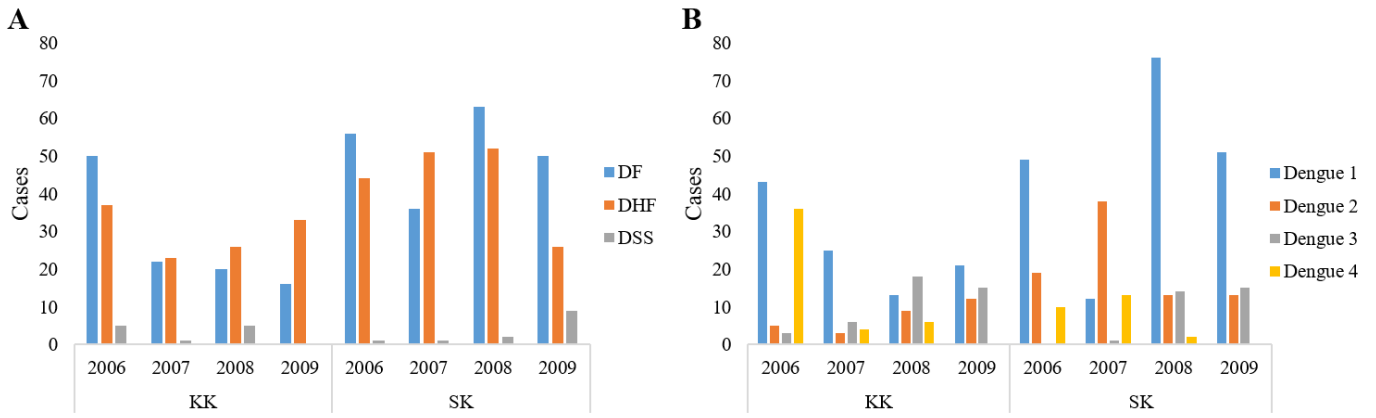
- KC performed experiments, analyzed the data, and wrote the manuscript.
- DM analyzed the data and revised the manuscript.
- PK analyzed the data and wrote the manuscript.
- ST and DS performed the NGS.
- NT collected and prepared specimens for this study.
- PM conceptualized the project and provided specimen resources.
- PS conceptualized the project, acquired the funding, oversaw the investigation, and revised the manuscript.
- All the authors read and approved the final manuscript.

References

1. Drake JW. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci U S A. 1993;90:4171-5.
2. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. Microbiol Mol Biol Rev. 2012;76:159-216.
3. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog. 2010;6:e1001005.
4. Martina BE, Koraka P, Osterhaus AD. Dengue virus pathogenesis: an integrated view. Clin Microbiol Rev. 2009;22:564-81.
5. Rothman AL. Immunity to dengue virus: a tale of original antigenic sin and tropical cytokine storms. Nat Rev Immunol. 2011;11:532-43.
6. Halstead SB, O'Rourke EJ. Antibody-enhanced dengue virus infection in primate leukocytes. Nature. 1977;265:739-41.
7. Halstead SB, O'Rourke EJ. Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralizing antibody. J Exp Med. 1977;146:201-17.
8. Arayasongsak U, Naka I, Ohashi J, Patarapotikul J, Nuchnoi P, Kalambaheti T, et al. Interferon lambda 1 is associated with dengue severity in Thailand. Int J Infect Dis. 2020;93:121-5.
9. Wang WH, Lin CY, Chang K, Urbina AN, Assavalapsakul W, Thitithanyanont A, et al. A clinical and epidemiological survey of the largest dengue outbreak in Southern Taiwan in 2015. Int J Infect Dis. 2019;88:88-99.
10. Kurnia B, Suryawan IWB. The Association between Obesity and Severity of Dengue Hemorrhagic Fever in Children at Wangaya General Hospital. Open Access Maced J Med Sci. 2019;7:2444-6.
11. Rico-Hesse R, Harrison LM, Salas RA, Tovar D, Nisalak A, Ramos C, et al. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. Virology. 1997;230:244-51.
12. Cologna R, Armstrong PM, Rico-Hesse R. Selection for virulent dengue viruses occurs in humans and mosquitoes. J Virol. 2005;79:853-9.
13. Sim S, Aw PP, Wilm A, Teoh G, Hue KD, Nguyen NM, et al. Tracking Dengue Virus Intra-host Genetic Diversity during Human-to-Mosquito Transmission. PLoS Negl Trop Dis. 2015;9:e0004052.
14. Parameswaran P, Wang C, Trivedi SB, Eswarappa M, Montoya M, Balmaseda A, et al. Intrahost Selection Pressures Drive Rapid Dengue Virus Microevolution in Acute Human Infections. Cell Host Microbe. 2017;22:400-10 e5.
15. Yenchitsomanus PT, Sricharoen P, Jaruthasana I, Pattanakitsakul SN, Nitayaphan S, Mongkolsapaya J, et al. Rapid detection and identification of dengue viruses by polymerase chain reaction (PCR). Southeast Asian J Trop Med Public Health. 1996;27:228-36.
16. World Health Organization. Dengue haemorrhagic fever : diagnosis, treatment, prevention and control. 2nd ed. Geneva: World Health Organization; 1997.
17. Chin-inmanu K, Suttitheptumrong A, Sangsrakru D, Tangphatsornruang S, Tragoonrung S, Malasit P, et al. Feasibility of using 454 pyrosequencing for studying quasispecies of the whole dengue viral genome. BMC Genomics. 2012;13 Suppl 7:S7.
18. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27:863-4.
19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-20.
20. Marinier E, Brown DG, McConkey BJ. Pollux: platform independent error correction of single and mixed genomes. BMC Bioinformatics. 2015;16:10.
21. Li H, Durbin R. Fast and accurate short read alignment with Burrows -Wheeler transform. Bioinformatics. 2009;25:1754-60.
22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078-9.
23. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297-303.
24. Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012;40:11189-201.
25. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 2012;40:W597-603.
26. Vaughn DW, Green S, Kalayanarooj S, Innis BL, Nimmannitya S, Suntayakorn S, et al. Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J Infect Dis. 2000;181:2-9.
27. Wang WK, Chao DY, Kao CL, Wu HC, Liu YC, Li CM, et al. High levels of plasma dengue viral load during defervescence in patients with dengue hemorrhagic fever: implications for pathogenesis. Virology. 2003;305: 330-8.
28. Wang WK, Chen HL, Yang CF, Hsieh SC, Juan CC, Chang SM, et al. Slower rates of clearance of viral load and virus-containing immune complexes in patients with dengue hemorrhagic fever. Clin Infect Dis. 2006;43:1023-30.
29. Mani I, Gilbert P, Sankale JL, Eisen G, Mboup S, Kanki PJ. Intrapatient diversity and its correlation with viral setpoint in human immunodeficiency virus type 1 CRF02_A/G-IbNG infection. J Virol. 2002;76:10745-55.
30. Duraes-Carvalho R, Ludwig-Begall LF, Salemi M, Lins RD, Marques ETA. Influence of directional positive Darwinian selection-driven evolution on arboviruses Dengue and Zika virulence and pathogenesis. Mol Phylogenet Evol. 2019;140:106607.
31. Aaskov J, Buzacott K, Thu HM, Lowry K, Holmes EC. Long-term transmission of defective RNA viruses in humans and Aedes mosquitoes. Science. 2006;311:236-8.
32. Rodriguez-Roche R, Blanc H, Borderia AV, Diaz G, Henningsson R, Gonzalez D, et al. Increasing Clinical Severity during a Dengue Virus Type 3 Cuban Epidemic: Deep Sequencing of Evolving Viral Populations. J Virol. 2016;90:4320-33.
33. Briones C, de Vicente A, Molina-Paris C, Domingo E. Minority memory genomes can influence the evolution of HIV-1 quasispecies in vivo. Gene. 2006;384:129-38.
34. Sabchareon A, Sirivichayakul C, Limkittikul K, Chanthavanich P, Suvannadabba S, Jiwariyavej V, et al. Dengue infection in children in Ratchaburi, Thailand: a cohort study. I. Epidemiology of symptomatic acute dengue infection in children, 2006-2009. PLoS Negl Trop Dis. 2012;6:e1732.



Supplement figure 1. The annual incidence of dengue infection in north eastern (A) and southern (B) regions of Thailand as reported by Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand during 2006 to 2009



Supplement figure 2. The number of dengue cases enrolled in the dengue cohort during 2006-2009 grouped by severity (A) and serotype (B)