Advances in Bioresearch
Adv. Biores., Vol 14 (5) September 2023: 61-67

Print ISSN 0976-4585; Online ISSN 2277-1573
Journal's URL:http://www.soeagra.com/abr.html
CODEN: ABRDC3
DOI: 10.15515/abr.0976-4585.14.5.6167

Advances in Bioresearch

ORIGINAL ARTICLE

Genetic Diversity and Evolutionary Dynamics of Dengue Virus Circulating in Saudi Arabia

Shaia S R Almalki
Department of Laboratory Medicine, Faculty of Applied Medical Sciences Al-Baha University, Al-Baha, Saudi Arabia
Email: shalmalki@bu.edu.sa

ABSTRACT
Dengue fever is a major public health concern in many parts of the world, including Saudi Arabia. The genetic diversity and evolutionary dynamics of dengue virus in the country have not been well characterized, which hinders the development of effective prevention and control strategies. A retrospective study of dengue virus isolates collected and sequenced in Saudi Arabia between 2000 and 2019. The whole genome sequence of DENV-1, 2 and 3 isolates were used to perform phylogenetic and conserved domain analyses to infer the evolutionary relationships and fitness among the viral strains. The DENV isolates could be divided into four distinct genotypes, with genotype IV being the most prevalent. Within each genotype, there are multiple sublineages that had diverged over time. The genetic diversity of DENV strains increased, with the emergence of a specific conserved domain in DENV-1 isolate. Our study provides new insights into the genetic diversity and evolutionary dynamics of dengue virus in Saudi Arabia. The findings suggest that ongoing surveillance and monitoring of dengue virus diversity and evolution in the region are necessary for effective prevention and control of dengue fever. Additionally, the genetic characterization of dengue virus strains can provide valuable information for the development of effective vaccines and antiviral therapies.
Keywords: DENV isolates, Genetic Diversity, Dengue Virus

Received 24.05.2023 Revised 01.06.2023 Accepted 11.06.2023

How to cite this article:
Shaia S R Almalki. Genetic Diversity and Evolutionary Dynamics of Dengue Virus Circulating In Saudi Arabia. Adv. Biores., Vol 14 (5) September 2023: 61-67.

INTRODUCTION

Aedes aegypti and Aedes albopictus mosquitoes are the main vectors of the dengue virus, which causes dengue fever (DF), a tropical acute infectious disease spread by arthropods [1]. High fever, rash, lethargy, joint pain, leucopenia, and lymphadenopathy are the most typical clinical signs of infected people [2]. Additionally, patients with dengue who recover without experiencing any serious consequences are still considered to have the disease, according to the updated WHO categorization system [3]. Severe dengue is defined as the presence of dengue hemorrhagic fever, shock syndrome, an accumulation of serosal fluid severe enough to cause respiratory distress, severe organ dysfunction, or death [4].

DENV is a member of the Flavivirus family, genus Flavivirus. DENV has been divided into four serotypes (DENV-1 to DENV-4) [5]. The most diverse of these serotypes is DENV-1, which has been classified into five different genotypes: genotype I (found in Southeast Asia and East Africa), genotype II (found in Thailand), genotype III (found in Malaysia), genotype IV (found in the South Pacific), and genotype V (found in America and Africa) [6]. These genotypes are further broken down into four lineages (lineages I through lineages IV), with genotype I being one of them [7]. The single-stranded, positive-sense RNA virus DENV has a genome that is roughly 10,000–11,000 nucleotide bases long. The DENV genome also includes 5′ and 3′ untranslated regions (UTRs) and a single open reading frame that codes for a polyprotein. The polyprotein is made up of seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) and three structural proteins (capsid [C], membrane [PrM/M], and envelope [E]) [8-10].

DF is an epidemic illness that tends to spread in tropical and subtropical regions including south America,southern Europe, southeast Asia, the western Pacific, north Africa, Australia, the eastern Mediterranean, and the islands in the Indian Ocean [11]. The morbidity of DF has rapidly increased in recent years as a result of the spread of the DENV [12]. More than 100 nations have recorded DF instances,

and each year, 390 million individuals worldwide contract the disease, including 96 million cases of severe dengue [13, 14]. Three different serotypes of DENV have been identified in Saudi Arabia; however, a robust phylogenetic analysis to compare the genetic diversity and viral fitness of DENV-1, 2 and 3 strains that have been prevalent in Saudi Arabia.

MATERIAL AND METHODS

Multiple Sequence alignment and phylogenetic analysis The complete viral genome sequences of DENV-1 (GenBank accession number KJ649286.1), DENV-2 (GenBank accession number KJ830750.1; MN294937.1) and DENV-3 (GenBank accession number KJ830751.1) along with reference sequences of DENV-1 (GenBank accession number NC_001477.1), DENV-2 (GenBank accession number NC_001474.2), DENV-3 (GenBank accession number NC_001475.2) and DENV-4 (GenBank accession number NC_002640.1) were obtained from NCBI (Table 1 &2).

Table 1: Dengue virus types included in the present study

Name	Accession	Base pairs	percent GC
Dengue virus type I	KJ649286_1	10622	46.63905
Dengue virus type 2	KJ830750_1	10718	46.03471
Dengue virus type 3	KJ830751_1	10635	46.56323
Dengue virus type 2	MN294937_1	10723	45.86403
Dengue virus type 2	NC_001474_1	10723	45.8174
Dengue virus type 3	NC_001475_1	10707	46.70776
Dengue virus type I	NC_001477_1	10735	46.66977
Dengue virus type 4	NC_002640_1	10649	47.1218

Table 2: GenBank Accession of nucleotide and protein sequences of Dengue virus types

GenBank Accession. Nucleotide	GenBank Accession. Protein
KJ649286.1	AIG59667.1[1..3392]
KJ830751.1	AIH13925.1[1..3390]
KJ830750.1	AIH13924.1[1..3391]
MN294937.1	QEV86381.1[1..3391]
NC_001477.1	NP_059433.1[1..3392]
NC_001474.2	NP_056776.2[1..3391]
NC_001475.2	YP_001621843.1[1..3390]
NC_002640.1	NP_073286.1[1..3387]

BLAST software (http://www.ncbi.nlm.nih.gov/BLAST/) was used to compare the whole genome of the DENV-1-3 circulating in Saudi Arabia. ClustalW was used to multiplex align the sequences, and the identity matrix of the nucleotide sequences as well as amino acids substitutions were examined. With 1,000 bootstrap replicates and the Neighbor-Joining approach of the Maximum Composite Likelihood model, gamma-distributed rates among sites, phylogenetic analysis and distance calculations were carried out using the MEGA 7.0 program.

Pairwise comparisons of the nucleotide sequences were conducted using the Genome-BLAST Distance Phylogeny (GBDP) method [15] under settings recommended for prokaryotic viruses [16]. The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME including SPR postprocessing [17] for each of the formulas D0, D4 and D6, respectively. Branch support was inferred from 100 pseudo-bootstrap replicates each. Trees were rooted at the midpoint [18] and visualized with tree [19]

Taxon boundaries at the species, genus and family level were estimated with the OPTSIL program (Göker *et al.,* 2009), the recommended clustering thresholds (Meier-Kolthoff and Göker 2017) and an F value (fraction of links required for cluster fusion) of 0.5 [21].

Conserved Domain Analysis Multiple sequence alignments using a protein sequence in FASTA format or the GI or Accession of a protein sequence that exists in the Entrez Protein database as query was to find conserved domains in the protein. The colored bars that depict the domain footprints (shown in both the concise display and full display of CD-Search results) are active hotlinks that open the corresponding CD summary pages with your query sequence embedded in the multiple sequence alignment of proteins used to create the domain model. CDD was sourced from NCBI Curated Domains, NCBIfams, Pfam, SMART, COG, PRK and TIGRFAMs.

RESULTS

Phylogenomic GBDP (Genome BLAST Distance Phylogeny) trees inferred using the formulas D0 (Figure 1), D4 (Figure 2) and D6 (Figure 3) yielded average support of 51 %, 100 % and 78 %, respectively. The

numbers above branches are GBDP pseudo-bootstrap support values from 100 replications. The branch lengths of the resulting VICTOR (VIral Classification Resource) trees are scaled in terms of the respective distance formula used (Figure 4). The OPTSIL clustering yielded three (D0), eight (D4) and four (D6) species clusters, respectively. At the genus level, one (D0), four (D4) and one (D6) clusters resulted, respectively. The number of clusters determined at the family level were one (D0), four (D4) and one (D6), respectively.

CD-length: 145 **E-value:** 9.72e-75 **Bitscore:** 245

```
                10       20       30       40       50
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   1808 IQDEERDIPERSWNSGYDWITDFPGKTVWFVPSIKSGNDIANCLRKNGKR
1857
Cdd:cd18806      1 IEDVALEIPGRIWFYGKAWITIYGGKTVWFVHSKKKGNEIAACLSGLGKN 50
                60       70       80       90       100
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   1858 VVQLSRKTFDTEYQKTKNNDWDYVVTTDISEMGANFRADRVIDPRRCLKP
1907
Cdd:cd18806     51 VIQLYRKLDDTEYPKIKTIDWDFVVTTDISEMGANFDADRVIDCRTCVKP 100
               110      120      130      140
            ....*....|....*....|....*....|....*....|....*
NP_059433.1 1908 VILKDGPERVILAGPMPVTVASAAQRRGRIGRNQNKEGDQYIYMG 1952
Cdd:cd18806  101 TILFSGDFRVILTGPVPQTAASAAQRRGRTGRNPAQERDIYRFVG 145
```

A specific conserved domain pfam00948, short name Flavi_NS1 accession 279316 was found in Flavivirus non-structural Protein NS1from 777 to 1130 with an E value 0 and bitscore 728.761. The NS1 protein is well conserved amongst the flaviviruses. It contains 12 cysteines, and undergoes glycosylation in a similar manner to other NS proteins.

Mutational analysis has strongly implied a role for NS1 in the early stages of RNA replication. However, this CD was only found in Dengue virus 1 isolate DENV-1-Jeddah, ACCESSION KJ649286, VERSION KJ649286.1

```
                10       20       30       40       50
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   777 SGCVINWKGRELKCGSGIFVTNEVHTWTEQYKFQADSPKRLSAAIGKAWE 826
Cdd:pfam00948   1 QGCAINFGGRELKCGDGIFIFNDSDDWLEKYKFQADDPKKLAAAIGAAFE 50
                60       70       80       90       100
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   827 EGVCGIRSATRLENIMWKQISNELNHILLENDMKFTVVVGDVSGILAQGK 876
Cdd:pfam00948  51 EGKCGINSADRLEHEMWKQIADEINAIFEENDMDFSVVVGDPKGILAQGK 100
               110      120      130      140      150
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   877 KMIRPQPMEH-----KYSWKSWGKAKIIGADVQNTTFIIDGPNTPECPDN 921
Cdd:pfam00948 101 KMIRPHPFEHirdglKYGWKSWGKAKIFGADRKNGSFIIDGKNRKECPDN 150
               160      170      180      190      200
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   922 QRAWNIWEVEDYGFGIFTTNIWLKLRDSYTQVCDHRLMSAAIKDSKAVHA 971
Cdd:pfam00948 151 NRAWNIFEIEDFGFGIFTTNIWLDARDEYTIDCDGRILGAAIKDKKAAHA 200
               210      220      230      240      250
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1   972 DMGYWIES-EKNETWKLARASFIEVKTCIWPKSHTLWSNGVLESEMIIPK 1020
Cdd:pfam00948 201 DMGFWIEShEKNETWKIARAEAIDVKECEWPKSHTIWGNGVEESEMFIPK 250
               260      270      280      290      300
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1  1021 IYGGPISQHNYRPGYFTQTAGPWHLGKLELDFDLCEGTTVVVDEHCGNRG 1070
Cdd:pfam00948 251 IIGGPISQHNHIPGYFTQTAGPWHLGKLELDFDACEGTSVIIDEHCDGRG 300
               310      320      330      340      350
            ....*....|....*....|....*....|....*....|....*....|
NP_059433.1  1071 PSLRTTTVTGKTIHEWCCRSCTLPPLRFKGEDGCWYGMEIRPVKEKEENL 1120
Cdd:pfam00948 301 KSLRSTTDSGKTIHEWCCRSCTLPPLRFHGEDGCWYGMEIRPRKEHEEHL 350
               360
            ....*....|
NP_059433.1  1121 VKSMVSAGSG 1130
Cdd:pfam00948  351 VKSMVSAGEG 360
```


Further phylogenetic analysis revealed that the DENV isolates from Saudi Arabia could be divided into three distinct genotypes (I-III). Within each genotype, multiple sublineages may diverge over time, indicating ongoing viral evolution.

The evolutionary dynamics of DENV-1 in Saudi Arabia are characterized by the emergence of new sublineages and genetic mutations that are associated with changes in viral fitness and disease severity. Further, the study identified a novel Dengue virus 1 isolate DENV-1-Jeddah, ACCESSION KJ649286, VERSION KJ649286.1 harboring conserved domain pfam00948, Superfamily: cl03032 which was not found in any other DENV-2 or DENV-3 isolate.

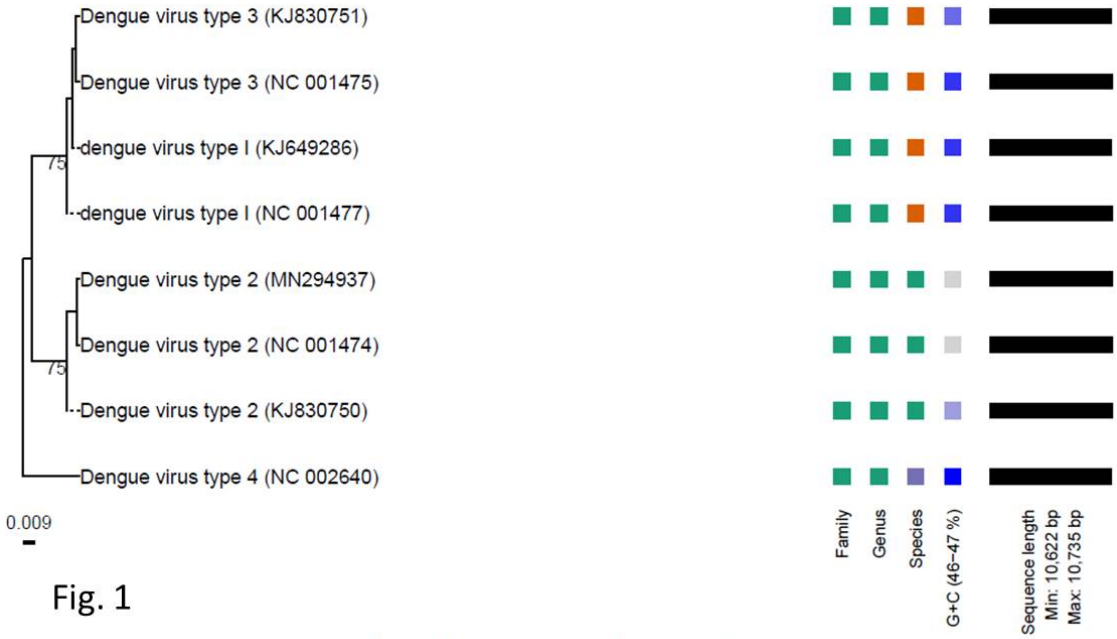

Fig. 1

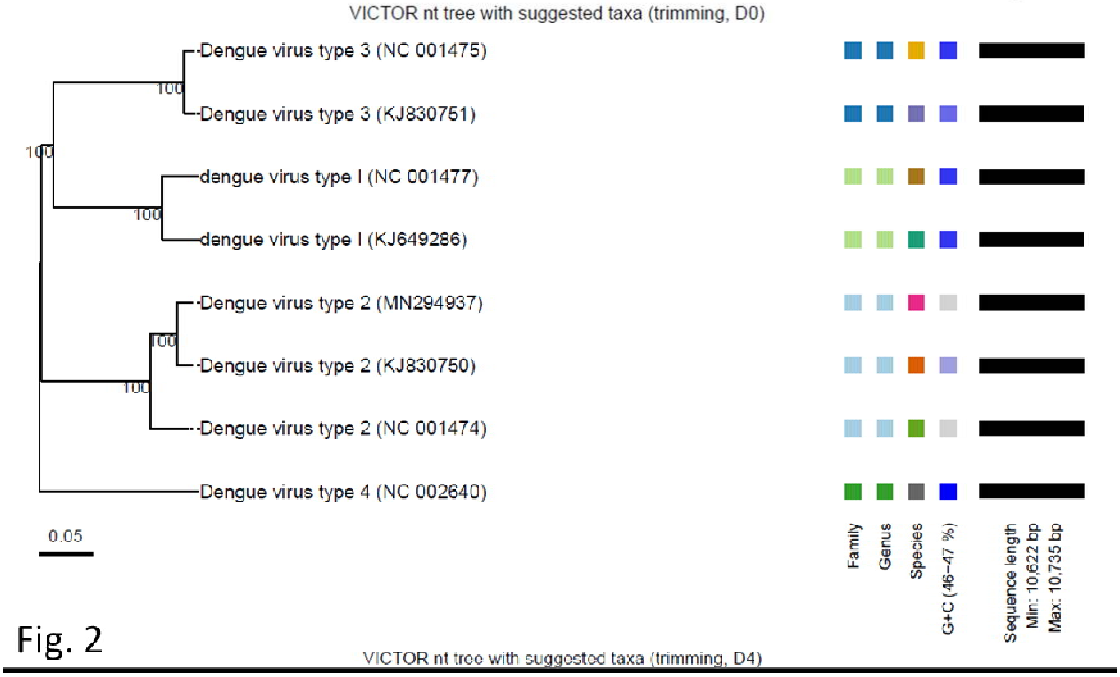

Fig. 2

Fig. 3

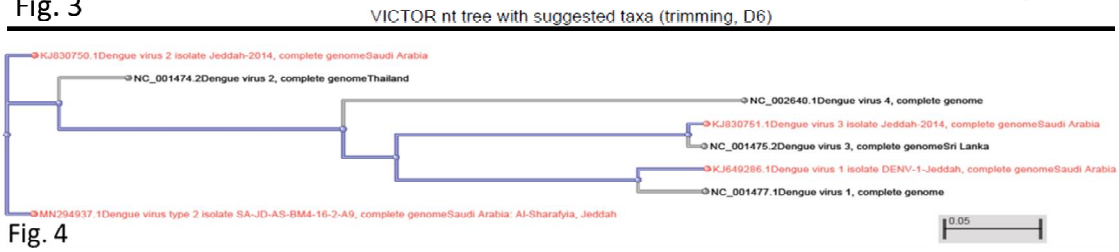

Fig. 4

DISCUSSION

Southeast Asia, the western Pacific, and America are the most afflicted regions globally, with dengue incidence rates increasing in recent decades (Jing and Wang 2019). Multiple sequence alignment (MSA) is a fundamental technique used in phylogenetic analysis to compare and align multiple sequences of DNA, RNA, or protein sequences. It is a crucial step in inferring the evolutionary relationships and constructing phylogenetic trees. The goal of MSA is to identify homologous positions among the sequences, which are believed to have descended from a common ancestor. By aligning these homologous positions, researchers can identify similarities and differences between sequences, which provide insights into their evolutionary relationships.

ClustalW algorithm method and iterative method PSI-BLAST was used to perform MSA in current study. These algorithms used different strategies to optimize the alignment based on maximizing sequence similarity or minimizing the number of gaps. DENV 1, 2 and 3 whole genome sequences from Saudi Arabia were used to generate MSA using maximum likelihood or Bayesian inference. These methods used the aligned sequences to estimate the evolutionary distances leading to the construction of a phylogenetic tree that represents the evolutionary relationships among the sequences. The present results provide a foundation for inferring evolutionary relationships and studying sequence conservation, functional motifs, and evolutionary constraints across related sequences.

Present study conducted Conserved Domain Analysis (CDA) utilizing multiple sequence alignments (MSA) of protein sequences derived from four whole DENV genomes from Saudi Arabia and Four reference genomes of DENV 1 to 4 to identify and analyze conserved domains, a region within a protein sequence that exhibits a high degree of sequence similarity across different organisms or protein families [18]. These domains often correspond to functional units or structural motifs that are important for the protein's activity. A specific conserved domain pfam00948, short name Flavi_NS1 accession 279316 was found in Dengue virus 1 isolate DENV-1-Jeddah, ACCESSION KJ649286, VERSION KJ649286.1. The NS1 protein is well conserved amongst the flaviviruses like Dengue and Yellow fever virus[20,22].

Conserved Domain Analysis (CDA) is essential for understanding the molecular mechanisms of viral replication, pathogenesis, and host-virus interactions. More importantly, CDA can be used to probe evolutionary Relationships. Viruses evolve rapidly, and their genetic diversity can be vast. By comparing the conserved domains across different viral species or strains, CDA allows to infer evolutionary relationships, in understanding viral evolution, transmission patterns, and the emergence of new viral

strains. Equally important application of CDA is vaccine and Antiviral Target Discovery by identifying conserved regions crucial for viral replication or virulence.
Presently found specific conserved domain pfam00948 (accession 279316) in nonstructural protein 1 (NS1) plays a crucial part in viral replication, and controls the host immunological response. By targeting such conserved domains, interventions that have broad-spectrum efficacy against multiple viral strains or species can be developed. CDA can also be utilized for the design of diagnostic assays to detect and identify viral infections. By targeting conserved regions, sensitive and specific tests that can detect a wide range of viral strains, even those with high genetic variability can be developed.

CONCLUSION

The current study provides new insights into the genetic diversity and evolutionary dynamics of dengue virus in Saudi Arabia. The findings suggest that ongoing surveillance and monitoring of dengue virus diversity and evolution in the region are necessary for effective prevention and control of dengue fever. Additionally, Conserved Domain Analysis using MSA plays a crucial role in virology by uncovering important functional and structural aspects of viral proteins. It may lead to an improved understanding of viral evolution, identifying targets for intervention, developing diagnostics, and facilitating comparative genomics studies. Overall, the genetic characterization of dengue virus strains can provide valuable information for the development of effective vaccines and antiviral therapies.

Financial support
The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number MOE- BU- 1- 2020.

CONFLICT OF INTERESTS: No conflict of interest related to this work

REFERENCES

1. Añez G, Volkova E, Jiang Z, Heisey DA, Chancey C, Fares RC, Rios M, Group CS. (2017). Collaborative study to establish World Health Organization international reference reagents for dengue virus Types 1 to 4 RNA for use in nucleic acid testing. Transfusion 57:1977-1987.
2. Bai Z, Liu LC, Jiang L, Luo L, Feng H, Lin P, Jing Q, Xiao X, Zhou H, Su W. (2018). Evolutionary and phylodynamic analyses of Dengue virus serotype I in Guangdong Province, China, between 1985 and 2015. Virus Res. 256:201-208.
3. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O. (2013). The global distribution and burden of dengue. Nature 496:504-507.
4. Brady OJ, Gething PW, Bhatt S, Messina JP, Brownstein JS, Hoen AG, Moyes CL, Farlow AW, Scott TW, Hay SI. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus.
5. Chambers TJ, Hahn CS, Galler R, Rice CM. (1990). Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. 44:649-688.
6. Farris JS. (1972). Estimating phylogenetic trees from distance matrices. The American Naturalist 106:645-668.
7. Göker M, García-Blázquez G, Voglmayr H, Tellería MT, Martín MP. (2009). Molecular taxonomy of phytopathogenic fungi: a case study in Peronospora. PloS one 4:e6319.
8. Guzman MG, Alvarez M, Halstead SB. (2013). Secondary infection as a risk factor for dengue hemorrhagic fever/dengue shock syndrome: an historical perspective and role of antibody-dependent enhancement of infection. Arch. Virol. 158:1445-1459.
9. Harapan H, Michie A, Sasmono RT, Imrie A. (2020). Dengue: a minireview. Viruses 12:829.
10. Holmes EC, Twiddy SS. (2003). The origin, emergence and evolutionary genetics of dengue virus. Infect., Genet. Evol. 3:19-28.
11. Jing Q, Wang M. (2019). Dengue epidemiology. Global Health Journal 3:37-45.
12. Lefort V, Desper R, Gascuel O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. Mol. Biol. Evol. 32:2798-2800.
13. Lustig Y, Wolf D, Halutz O, Schwartz E. (2017). An outbreak of dengue virus (DENV) type 2 Cosmopolitan genotype in Israeli travellers returning from the Seychelles, April 2017. Eurosurveillance 22:30563.
14. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:1-14.
15. Meier-Kolthoff JP, Göker M. (2017). VICTOR: genome-based phylogeny and classification of prokaryotic viruses. Bioinformatics 33:3396-3404.
16. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, Rohde C, Rohde M, Fartmann B, Goodwin LA. (2014). Complete genome sequence of DSM 30083 T, the type strain (U5/41 T) of Escherichia coli, and a proposal for delineating subspecies in microbial taxonomy.
17. Organization WH, Research SPf, Diseases TiT, Diseases WHODoCoNT, Epidemic WHO, Alert P. (2009). *Dengue: guidelines for diagnosis, treatment, prevention and control*. World Health Organization.

18. Parry R, Asgari S. (2019). Discovery of novel crustacean and cephalopod flaviviruses: insights into the evolution and circulation of flaviviruses between marine invertebrate and vertebrate hosts. J. Virol. 93:e00432-19.
19. Srikiatkhachorn A, Gibbons RV, Green S, Libraty DH, Thomas SJ, Endy TP, Vaughn DW, Nisalak A, Ennis FA, Rothman AL. (2010). Dengue hemorrhagic fever: the sensitivity and specificity of the world health organization definition for identification of severe cases of dengue in Thailand, 1994–2005. Clin. Infect. Dis. 50:1135-1143.
20. Tamura T, Torii S, Kajiwara K, Anzai I, Fujioka Y, Noda K, Taguwa S, Morioka Y, Suzuki R, Fauzyah Y. (2022). Secretory glycoprotein NS1 plays a crucial role in the particle formation of flaviviruses. PLoS Path. 18:e1010593.
21. van den Elsen K, Quek JP, Luo D. (2021). Molecular insights into the flavivirus replication complex. Viruses 13:956.
22. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abebo TA, Abera SF. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet 390:1211-1259.
23. Yu G. (2020). Using ggtree to visualize data on tree-like structures. Current protocols in bioinformatics 69:e96.