Comparison of Bioinformatic Approaches for Minimizing Sequence Redundancy in Amplicon Sequencing Data from Oxford Nanopore Technology Platforms Applied to Dengue Virus Research

Table of contents

[Introduction 4](#_Toc162015523)

[Dengue Virus 4](#_Toc162015524)

[Taxonomic classification of Flaviviridae family 4](#_Toc162015525)

[Dengue virus discovery and genus classification 4](#_Toc162015526)

[Dengue virus non-structural and structural proteins 7](#_Toc162015527)

[Dengue Virus Epidemiology 8](#_Toc162015528)

[Transmission routes of infection 8](#_Toc162015529)

[Illness 8](#_Toc162015530)

[Viral Shedding 9](#_Toc162015531)

[Infectivity 9](#_Toc162015532)

[Detection and molecular characterisation 11](#_Toc162015533)

[Oxford Nanopore technology 11](#_Toc162015534)

[Oxford Nanopore technology for Dengue virus 12](#_Toc162015535)

[References 13](#_Toc162015536)

Figures

[Figure 1: Immature flavivirus (left) and Dengue virus (right), created using BioRender 4](#_Toc162015537)

[Figure 2: Genome organization of Orthoflavivirus, created using BioRender 5](#_Toc162015538)

[Figure 3: Genome organization of Dengue Virus, created using BioRender 5](#_Toc162015539)

[Figure 4: Phylogenetic tree of Orthoflavivirus genus, created using iTOL 6](#_Toc162015540)

Tables

[Table 1: Nomenclature for structural and non-structural proteins of the Dengue virus 7](#_Toc162015541)

# Introduction

## Dengue Virus

### Taxonomic classification of Flaviviridae family

The Flaviviridae family comprises viruses with single-stranded, positive-sense RNA genomes, typically within the range of 9.0 to 13 kilobases. These viruses are distinguished by their enveloped structure and icosahedral symmetry. Central to their genomic organization is a singular, extensive open reading frame (ORF) which is translated into a precursor polyprotein. This polyprotein undergoes subsequent cleavage to yield three structural and seven non-structural proteins, each essential for the virus's replication and assembly processes. Within the Flaviviridae family, several members are notable for their transmission via arthropod vectors, such as mosquitoes and ticks, leading to diseases of significant public health concern including Dengue fever, Yellow fever, and West Nile virus infections (International Committee on Taxonomy of Viruses [ICTV], n.d.; Scitable by Nature Education, n.d.).

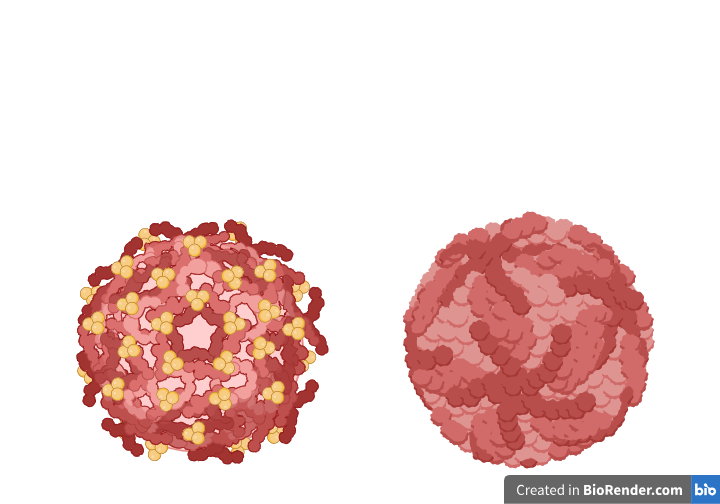


Figure 1: Immature flavivirus (left) and Dengue virus (right), created using BioRender

### Dengue virus discovery and genus classification

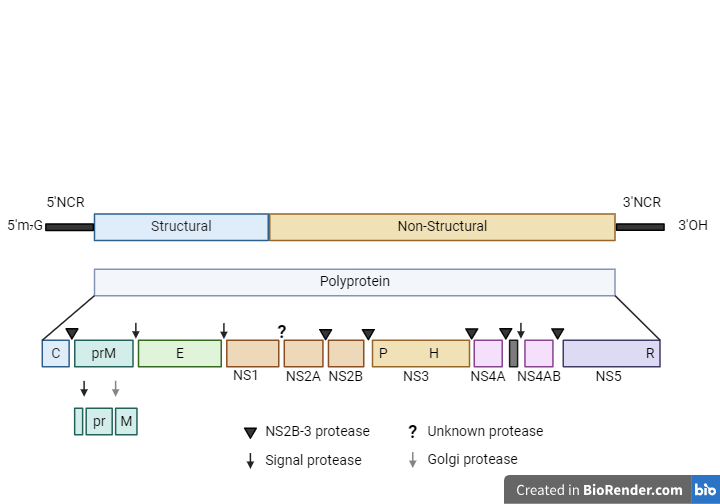
The discovery of the Dengue virus, classified within the Orthoflavivirus genus of the Flaviviridae family, marks a significant milestone in the history of virology. The initial identification occurred during the 1943 epidemic in Nagasaki, with scientists Ren Kimura and Susumu Hotta, alongside Albert B. Sabin and Walter Schlesinger working independently, isolating what would later be designated as the DEN-1 serotype. This enveloped, single-stranded RNA virus, with a genome approximately 11 kilobases in length, exhibits a complex taxonomy rooted in the diversity of its envelope protein E and the NS5 polymerase enzyme. The envelope protein E is instrumental in serotype determination, while NS5 plays a crucial role in the virus's replication process. The phylogenetic distinctions between the four recognized Dengue serotypes are elucidated through the analysis of amino acid variations in the E protein and nucleotide sequences within the NS5 polymerase region. This intricate genomic structure, comprising a single open reading frame, underscores the sophisticated arrangement of structural and non-structural proteins pivotal for the virus's replication and interaction with host cells (International Committee on Taxonomy of Viruses, n.d.; Scitable by Nature Education, n.d.).  
  


Figure 2: Genome organization of Orthoflavivirus, created using BioRender

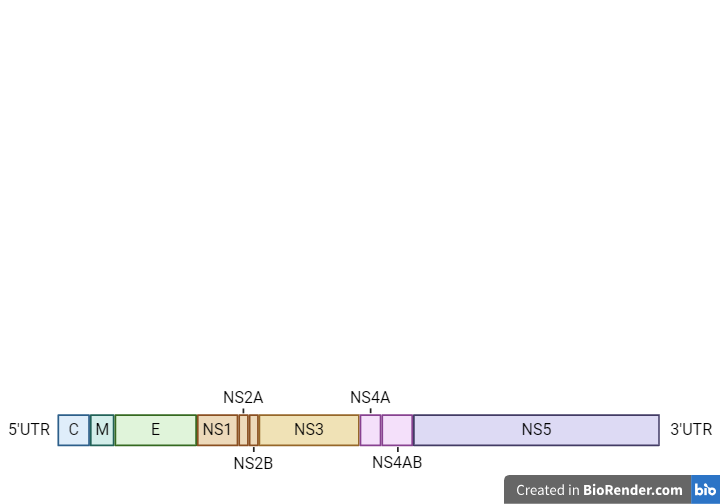


Figure 3: Genome organization of Dengue Virus, created using BioRender



Figure 4: Phylogenetic tree of Orthoflavivirus genus, created using iTOL

### Dengue virus non-structural and structural proteins

Within the Orthoflavivirus genus of the Flaviviridae family, the Dengue virus presents a sophisticated genomic structure that encodes for a polyprotein from its open reading frame (ORF1). This polyprotein undergoes post-translational modifications to yield three structural and seven non-structural proteins, each with distinct roles in the viral life cycle. Notably, the non-structural protein NS3 serves dual functions as a protease and helicase, while NS5 acts as the RNA-dependent RNA polymerase (RdRp), critical for viral replication. The structural envelope protein E is pivotal for mediating viral entry into host cells and determining serotype specificity, with domain III of this protein playing a crucial role in receptor binding. This interaction between the E protein and host cell receptors is a key target for neutralizing antibodies and vaccine development efforts, given its implications for viral pathogenesis and immune evasion (International Committee on Taxonomy of Viruses, n.d.; Scitable by Nature Education, n.d.).

Table 1: Nomenclature for structural and non-structural proteins of the Dengue virus

|  |  |  |
| --- | --- | --- |
| **Protein** | **Nomenclature** | **Function** |
| C | Capsid | Encapsulates the RNA genome, forms the core of the virus particle |
| prM/M | Membrane | Precursor to the membrane protein, involved in virion assembly and maturation |
| E | Envelope | Mediates viral entry through receptor binding; major determinant for serotype specificity |
| NS1 | Non-structural protein 1 | Immune modulation; secreted form is involved in pathogenesis |
| NS2A | Non-structural protein 2A | Viral replication and assembly; modulates host antiviral response |
| NS2B | Non-structural protein 2B | Co-factor for NS3 protease activity |
| NS3 | Non-structural protein 3 | Serine protease; RNA helicase |
| NS4A | Non-structural protein 4A | Induces membrane alterations for replication complex formation |
| NS4B | Non-structural protein 4B | Modulates host antiviral response; contributes to replication complex formation |
| NS5 | Non-structural protein 5 | RNA-dependent RNA polymerase (RdRp); Methyltransferase activity |

## Dengue Virus Epidemiology

### Transmission routes of infection

The transmission of the dengue virus to humans is predominantly through the bites of infected Aedes mosquitoes, specifically Aedes aegypti and Aedes albopictus. When these mosquitoes consume blood from a dengue-infected individual, the virus undergoes an incubation period within the mosquito for about 8-12 days, after which the mosquito becomes capable of transmitting the virus through its bites. The urban adaptability of these mosquitoes, coupled with their breeding in stagnant water found in man-made containers, makes urban and semi-urban areas significant hotspots for dengue transmission (World Health Organization, 2020; Centers for Disease Control and Prevention, 2021).

Besides the vector-based transmission, dengue can also be transmitted via vertical transmission from mother to fetus, blood transfusions, and organ transplants, though these modes are less common (Halstead, 2007). The geographical spread of dengue has been exacerbated by climate change and urbanization, which influence the distribution and breeding patterns of the Aedes mosquitoes (Ryan et al., 2019).

The increase in temperatures and changes in rainfall patterns due to climate change have expanded the distribution and activity of Aedes mosquitoes into new regions, including parts of Europe previously not at risk. The El Nino phenomena and increased humidity also contribute to the favorable conditions for mosquito breeding and survival, complicating the dengue transmission dynamics in Europe and globally. The first instances of local transmission in Europe were noted in France and Croatia in 2010, highlighting the broadening geographical reach of dengue influenced by environmental factors (World Health Organization, 2023).

Prevention strategies focus on reducing mosquito breeding sites and limiting human-mosquito contact through community efforts to eliminate standing water and personal protective measures such as using insect repellent and wearing protective clothing. The development of dengue vaccines offers a potential preventive measure, although their efficacy across all dengue serotypes and age groups poses a challenge (Thomas & Endy, 2011).

### Illness

Illness caused by the dengue virus often begins 4-10 days after a mosquito bite and can last 2-7 days. Common symptoms include high fever, severe headache, pain behind the eyes, muscle and joint pains, nausea, vomiting, and rash. Severe dengue, which occurs in about 1 in 20 cases, can lead to more critical symptoms such as severe abdominal pain, persistent vomiting, and bleeding, necessitating emergency medical care (Centers for Disease Control and Prevention, 2021; World Health Organization, 2023).

There's no specific medication for dengue; treatment primarily involves relieving symptoms. Adequate hydration and acetaminophen (paracetamol) are recommended for fever and pain management. Aspirin and NSAIDs are avoided due to the risk of bleeding. Severe cases may require hospitalization for supportive care (World Health Organization, 2023).

### Viral Shedding

In dengue virus infections, viral shedding refers to the release of the virus into the bloodstream, from where it can be transmitted to mosquitoes upon biting. The primary vectors for dengue virus transmission are Aedes aegypti and Aedes albopictus mosquitoes. After an infected mosquito bites a human, the virus incubates for 3 to 14 days, typically 4 to 7 days, before symptoms manifest. Symptoms can range from asymptomatic to severe, including high fever, headache, muscle and joint pains, nausea, vomiting, and rash. Severe dengue can lead to more critical conditions such as internal bleeding and organ failure (Environment, Health and Safety at Cornell University, 2023; Mayo Clinic, 2022).

There is no specific treatment for dengue; management focuses on symptom relief and supportive care. Prevention strategies emphasize reducing mosquito populations and avoiding mosquito bites, particularly in areas where dengue is endemic. Recent efforts in vaccine development offer hope for more effective dengue prevention in the future (Mayo Clinic, 2022).

### Infectivity

The infectivity of the dengue virus and the complexity of its infectious dose are influenced by various factors, including the virus's transmission mechanisms and the immune response it elicits. The dengue virus, transmitted by Aedes mosquitoes, exhibits diverse transmission cycles such as forest/enzoonotic, rural/endemic, and urban/epidemic/endemic, each contributing differently to the spread of the disease. Following the bite from an infected mosquito, the virus undergoes initial replication in dendritic cells, leading to systemic infection characterized by phases including febrile, critical, and recovery stages. The severity of the disease can range from mild dengue fever to more severe forms like dengue hemorrhagic fever, with the latter involving increased vascular permeability and plasma leakage. Notably, the phenomenon of antibody-dependent enhancement, where pre-existing antibodies to a different serotype can exacerbate the infection, underscores the complex interplay between the virus and the host immune system (Kularatne & Dalugama, 2022; American Academy of Pediatrics, 2022).

Given the absence of a defined infectious dose and the variability in disease manifestation among the four dengue serotypes (DENV-1 to DENV-4), understanding dengue virus infectivity necessitates a comprehensive approach that considers both environmental and individual factors, including prior exposure and immune status. The global incidence of dengue is on the rise, particularly in tropical and subtropical regions, with increasing concerns about its spread to new areas due to environmental changes and urbanization.

## Detection and molecular characterisation

In the realm of virology, the elucidation of the molecular characteristics of dengue virus (DENV) serotypes is pivotal for comprehending its epidemiological landscape and formulating public health strategies. The arsenal of techniques for DENV detection and genotyping is expansive, with conventional and real-time RT-PCR standing out for their proficiency in identifying viral RNA within clinical specimens. Notably, the four-probe Taqman system has garnered acclaim for its adeptness in amplifying viral RNA across diverse serotypes, illustrating the critical role of thermal cycling conditions' optimization in the success of viral detection efforts (Virology Journal, 2021).

Phylogenetic analysis, particularly for DENV-1, leverages complete E gene sequences to construct a comprehensive genotypic framework, thus enhancing our understanding of the virus's evolutionary trajectory and dissemination. This genotypic stratification not only unveils the genetic heterogeneity among DENV-1 genotypes but also serves as a cornerstone for tailored public health initiatives aimed at mitigating dengue outbreaks (Infectious Diseases of Poverty, 2021).

Moreover, the meticulous genotyping of DENV-2 through the amplification of envelope gene fragments exemplifies the rigorous approach to characterizing the virus's genetic constitution. The stringent selection of samples based on cycle threshold (Ct) values, followed by the sequencing of purified PCR products, allows for an in-depth exploration of the virus's genetic diversity, revealing its phylogenetic relationships and epidemiological patterns (Virology Journal, 2021).

### Oxford Nanopore technology

Oxford Nanopore Technologies (ONT) offers a unique approach to DNA and RNA sequencing through its various platforms including MinION, GridION, PromethION, and the compact P2 Solo. At the core of ONT's sequencing technology is the passage of DNA or RNA molecules through a nanopore embedded in a membrane. As each nucleotide passes through the nanopore, it causes a distinctive disruption in the ionic current, which is detected and measured. These electrical signals are then interpreted, or "basecalled," using sophisticated algorithms to determine the sequence of bases in the DNA or RNA strand.

The process of amplicon-based sequencing on ONT platforms involves generating specific fragments of DNA, known as amplicons, through PCR amplification. These amplicons can be designed to target specific regions of interest within a genome, allowing for focused sequencing efforts. This approach is particularly useful for applications such as pathogen detection, genetic variation analysis, and targeted gene studies. In the wet lab, DNA or RNA is extracted from samples, PCR amplification is conducted to generate amplicons, and these amplicons are then prepared for sequencing on ONT devices.

Bioinformatic processing of amplicon ONT data involves handling the raw electrical signal data to derive nucleotide sequences and then further analyzing these sequences for various applications. Challenges in processing ONT data include managing the high volume of data generated, error correction, and the assembly of long reads. Various bioinformatics tools have been developed to address these challenges, offering solutions for basecalling, alignment, variant calling, and data visualization.

Despite the advantages of ONT's long-read sequencing and real-time data analysis capabilities, there are challenges such as higher error rates compared to short-read sequencing technologies and the computational demands of processing large datasets. Researchers continue to develop and refine methodologies to mitigate these challenges, enhancing the utility and accuracy of ONT sequencing data for a wide range of research applications.

### Oxford Nanopore technology for Dengue virus

Oxford Nanopore Technologies (ONT) provides a novel approach to sequencing that leverages the movement of DNA or RNA molecules through a nanopore to generate sequencing data. This method is particularly advantageous for its ability to produce long-read sequences and provide real-time data analysis, which is beneficial for various applications, including the study of the dengue virus.

In the context of dengue virus research, ONT platforms have been utilized to generate near full-length genome sequences, enabling a comprehensive understanding of the virus's genetic makeup. The process involves designing multiplex primer sets to produce overlapping amplicons covering the entire coding region of the dengue virus, facilitating the assembly of sequence data. This approach, detailed in research published on ONT's website, highlights the potential of nanopore sequencing in resource-limited settings for generating complete coding region sequences of the dengue virus, providing valuable insights into the virus's evolution and spread. (Adikari et al., 2020)

Furthermore, a study featured in Virology Journal illustrates the application of a multiplex PCR and Nanopore-based method for sequencing the dengue virus in Indonesia. This method involves extracting RNA, performing reverse transcription and PCR amplification, and then sequencing the resulting amplicons using ONT platforms. The data generated through this process is crucial for understanding the genetic differences of dengue virus serotypes and their implications for disease manifestation and epidemiology. (Stubbs et al., 2020)

Bioinformatic processing of the data from ONT platforms involves base-calling, demultiplexing, alignment to reference sequences, and consensus sequence generation. Despite the benefits of ONT's long-read sequencing capabilities, challenges such as error rates and data volume necessitate the use of specialized bioinformatics tools to ensure accurate sequence assembly and analysis.

# References

International Committee on Taxonomy of Viruses. (n.d.). Flaviviridae. [online] Available at: https://ictv.global/report/chapter/flaviviridae/flaviviridae/orthoflavivirus [Accessed 17 Mar. 2024].

Scitable by Nature Education. (n.d.). Dengue Viruses. [online] Available at: https://www.nature.com/scitable/topicpage/dengue-viruses-22400925/ [Accessed 17 Mar. 2024].

World Health Organization, 2020. Dengue and severe dengue [online]. Available at: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue [Accessed 18 March 2024].

Centers for Disease Control and Prevention, 2021. Dengue [online]. Available at: https://www.cdc.gov/dengue/index.html [Accessed 18 March 2024].

Halstead, S.B., 2007. Dengue Virus–Mosquito Interactions. Annual Review of Entomology, 53, pp.273-291.

Ryan, S.J., Carlson, C.J., Mordecai, E.A., and Johnson, L.R., 2019. Global Expansion of Dengue Virus Serotypes: Revisiting the Role of Aedes aegypti Mosquitoes in Viral Transmission. PLOS Pathogens, 15(10), e1008101.

Thomas, S.J. & Endy, T.P., 2011. Critical Issues in Dengue Vaccine Development. Current Opinion in Infectious Diseases, 24(5), pp.442-450.

World Health Organization, 2023. Dengue- Global situation [online]. Available at: https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON498 [Accessed 19 March 2024].

Centers for Disease Control and Prevention, 2021. Symptoms and Treatment | Dengue | CDC. [online] Available at: https://www.cdc.gov/dengue/symptoms/index.html [Accessed 19 March 2024].

World Health Organization, 2023. Dengue and severe dengue. [online] Available at: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue [Accessed 19 March 2024].

Environment, Health and Safety at Cornell University. (2023). Dengue Virus Biological Agent Reference Sheet (BARS). [online] Available at: https://ehs.cornell.edu/research-safety/biosafety-biosecurity/biological-safety-manuals-and-other-documents/bars-other/dengue-virus [Accessed 19 Mar. 2024].

Mayo Clinic. (2022). Dengue fever - Symptoms and causes. [online] Available at: https://www.mayoclinic.org/ [Accessed 19 Mar. 2024].

Kularatne, S.A. and Dalugama, C. (2022) Dengue infection: Global Importance, immunopathology and management, Clinical medicine (London, England). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8813012/ [Accessed: 19 March 2024].

American Academy of Pediatrics. (2022). Dengue: A Growing Problem With New Interventions. Pediatrics, 149(6), e2021055522. [online] Available at: https://publications.aap.org/pediatrics/article/149/6/e2021055522/187012/Dengue-A-Growing-Problem-With-New-Interventions [Accessed 19 Mar. 2024].

Virology Journal. (2021). 'Genotyping of dengue virus from infected tissue samples embedded in paraffin'. [online] Available at: https://virologyj.biomedcentral.com/articles/10.1186/s12985-023-02072-5 [Accessed 20 Mar. 2024].

Infectious Diseases of Poverty. (2021). 'A unified global genotyping framework of dengue virus serotype-1 for a stratified coordinated surveillance strategy of dengue epidemics'. [online] Available at: https:// https://idpjournal.biomedcentral.com/articles/10.1186/s40249-022-01024-5 [Accessed 20 Mar. 2024].

Oxford Nanopore Technologies. (n.d.). MinION. [online] Available at: https://nanoporetech.com [Accessed 21 Mar. 2024].

Oxford Nanopore Technologies. (n.d.). PromethION 2. [online] Available at: https://nanoporetech.com [Accessed 21 Mar. 2024].

Oxford Nanopore Technologies. (n.d.). PromethION. [online] Available at: https://nanoporetech.com [Accessed 21 Mar. 2024].

Adikari, T.N. et al. (2020) Single Molecule, near full-length genome sequencing of Dengue virus, Nature News. Available at: https://www.nature.com/articles/s41598-020-75374-1 (Accessed: 21 March 2024).

Stubbs, S.C.B. et al. (2020) Assessment of a multiplex PCR and nanopore-based method for dengue virus sequencing in Indonesia - Virology Journal, BioMed Central. Available at: https://virologyj.biomedcentral.com/articles/10.1186/s12985-020-1294-6 (Accessed: 22 March 2024).