Development of an in-silico pipeline for assessing specificity of dengue virus genotyping via conventional PCR assay

Nelson Krisanto – 23203786

Table of contents

[Introduction 6](#_Toc172928055)

[Dengue virus 6](#_Toc172928056)

[Taxonomic classification of Flaviviridae family 6](#_Toc172928057)

[Dengue virus discovery and genus classification 7](#_Toc172928058)

[Dengue virus non-structural and structural proteins 9](#_Toc172928059)

[Dengue virus Epidemiology 11](#_Toc172928060)

[Transmission routes of infection 11](#_Toc172928061)

[Prevention 13](#_Toc172928062)

[Illness 15](#_Toc172928063)

[Viral Shedding 16](#_Toc172928064)

[Infectivity 17](#_Toc172928065)

[Detection and molecular characterisation 18](#_Toc172928066)

[Materials and Methods 19](#_Toc172928067)

[Execution Environment 19](#_Toc172928068)

[CPU Information: 20](#_Toc172928069)

[Network Configuration: 20](#_Toc172928070)

[Developing an in-house Dengue virus genotyping pipeline 23](#_Toc172928071)

[Fetching Dengue virus Sequences 23](#_Toc172928072)

[Clustering Dengue virus Sequences 23](#_Toc172928073)

[Fetching Orthoflavivirus Sequences 23](#_Toc172928074)

[Assigning Taxonomy Using Genome Detective 24](#_Toc172928075)

[Generating VADR Amplicons 24](#_Toc172928076)

[Extracting VADR Amplicons 25](#_Toc172928077)

[Creating Taxonomy Database 25](#_Toc172928078)

[Converting and Creating BLAST Database 26](#_Toc172928079)

[Assigning Taxonomy Using BLAST 27](#_Toc172928080)

[Fetching and Indexing Dengue virus Reference Genome – Process 2 28](#_Toc172928081)

[Fetching Reference Genomes 28](#_Toc172928082)

[Indexing Reference Genomes 28](#_Toc172928083)

[Dengue virus-Specific Primers and Amplicons – Process 3 29](#_Toc172928084)

[Literature Mining for Dengue virus-Specific Primers 29](#_Toc172928085)

[Converting Cleaned Primer Data to FASTA Format 29](#_Toc172928086)

[Checking Primer Positions 29](#_Toc172928087)

[Determining Primer Melting Temperatures 30](#_Toc172928088)

[Generating Primer Pairs 30](#_Toc172928089)

[In-silico PCR 31](#_Toc172928090)

[Results/Analysis Generation – Process 4 32](#_Toc172928091)

[Conventional PCR 32](#_Toc172928092)

[Phylogenetic Tree Creation 39](#_Toc172928093)

[Results 43](#_Toc172928094)

[Distribution of Dengue Virus Primers by Year and Serotype 43](#_Toc172928095)

[Region Distribution of Conventional PCR Primers 46](#_Toc172928096)

[Dengue Virus Nomenclature 50](#_Toc172928097)

[Development of an in-house dengue virus genotyping pipeline 51](#_Toc172928098)

[Clustering Dengue Virus Sequence 51](#_Toc172928099)

[Evaluation of In-House Genotyping Pipeline Performance 53](#_Toc172928100)

[In-silico comparison of PCR assays for dengue virus genotyping 54](#_Toc172928101)

[Serotype Confusion Matrix 54](#_Toc172928102)

[Genotype Confusion Matrix 56](#_Toc172928103)

[Missed and Mismatched Classification 57](#_Toc172928104)

[F1 Score for Each Primer Pair for Each Serotype and Genotype 59](#_Toc172928105)

[Discussion 65](#_Toc172928106)

[References 68](#_Toc172928107)

[Appendix 77](#_Toc172928108)

[Primer Pair for Dengue Virus Overall Metrics 77](#_Toc172928109)

[Missed Classification of Conventional PCR Primer Pairs 78](#_Toc172928110)

[Mismatched Classification of Conventional PCR Primer Pairs 81](#_Toc172928111)

Figures

[Figure 1: Immature flavivirus (left) and Dengue virus (right), created using BioRender 6](#_Toc172928112)

[Figure 2: Genome organization of Orthoflavivirus, created using BioRender 7](#_Toc172928113)

[Figure 3: Phylogenetic tree of Orthoflavivirus genus, created using iTOL 8](#_Toc172928114)

[Figure 4: Genome organization of Dengue virus, created using BioRender 9](#_Toc172928115)

[Figure 5: Process diagram 22](#_Toc172928116)

[Figure 6: Literature Mining – Dengue Virus PCR Primers by Year and Serotype 43](#_Toc172928117)

[Figure 7: OLS Regression for (a) Number of Primers Over the Years, (b) Number of Publications Over the Years, and (c) Weighted Number of Primers Over the Years 44](#_Toc172928118)

[Figure 8: (a) Number of Primers and Publications Over the Years, and (b) Weighted Number of Primers by Number of Publications Over the Years 45](#_Toc172928119)

[Figure 9: Region Distribution of Conventional PCR Primers for (a) Individual Primers, and (b) Primer Combination. 46](#_Toc172928120)

[Figure 10: PCA Analysis of Conventional PCR Primers for (a) Individual Primer, and (b) Primer Combination 47](#_Toc172928121)

[Figure 11: Dengue Virus Sequence Genotyping between Unclustered vs Clustered 51](#_Toc172928122)

[Figure 12: Whole vs Clustered Dengue Virus Sequence Dataset – Genotype Distribution 52](#_Toc172928123)

[Figure 13: Serotype Confusion Matrix 55](#_Toc172928124)

[Figure 14: Genotype Confusion Matrix 56](#_Toc172928125)

[Figure 15: Missed Classification Heatmap 57](#_Toc172928126)

[Figure 16: Mismatch Classification Heatmap 58](#_Toc172928127)

Tables

[Table 1: Nomenclature for structural and non-structural proteins of the Dengue virus 9](#_Toc172928128)

[Table 2: Table of Dengue Outbreaks in Non-endemic Regions 11](#_Toc172928129)

[Table 3: Summary of Dengue virus Shedding 16](#_Toc172928130)

[Table 4: DENV infectivity by genotype 17](#_Toc172928131)

[Table 5: Overview of Dengue virus Detection and Genotyping Methods 18](#_Toc172928132)

[Table 6: Orthoflaviviruses grouped by vector and host 39](#_Toc172928133)

[Table 7: Contribution of Variables for Individual Primers 48](#_Toc172928134)

[Table 8: NCBI Taxonomy, New Lineage, and Genome Detective Nomenclature 50](#_Toc172928135)

[Table 9: Top Cutoff (75-70-65) 53](#_Toc172928136)

[Table 10: Top 25 Primer Pair for Dengue Virus Overall Metrics 54](#_Toc172928137)

[Table 11: F1 Score for Each Primer Pair for Each Serotype 59](#_Toc172928138)

[Table 12: F1 Score for Each Primer Pair for Each Genotype 61](#_Toc172928139)

[Table 13: Primer Pair for Dengue Virus Overall Metrics 77](#_Toc172928140)

[Table 14: Missed Classification of Conventional PCR Primer Pairs 78](#_Toc172928141)

[Table 15: Missed Classification of Conventional PCR Primer Pairs 81](#_Toc172928142)

# 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 (kb). 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. The significant public health impact of these arthropod-borne viruses is underscored by their mechanisms of transmission and the diseases they induce, primarily transmitted through bites from infected mosquitoes and ticks. This transmission cycle is crucial for their maintenance in nature, facilitating the spread of viruses among vertebrate hosts, including humans. Moreover, more than half of the identified flaviviruses are associated with human diseases, including severe conditions such as hemorrhagic fevers and encephalitis, highlighting the serious health risks these pathogens pose. (Organization, 1997; Gubler, 1998; Mukhopadhyay, Kuhn and Rossmann, 2005; Lindenbach *et al.*, 2013)

As seen in Figure 1, the structure of flaviviruses can be visualized, highlighting both immature flavivirus (left) and Dengue virus (right) forms.

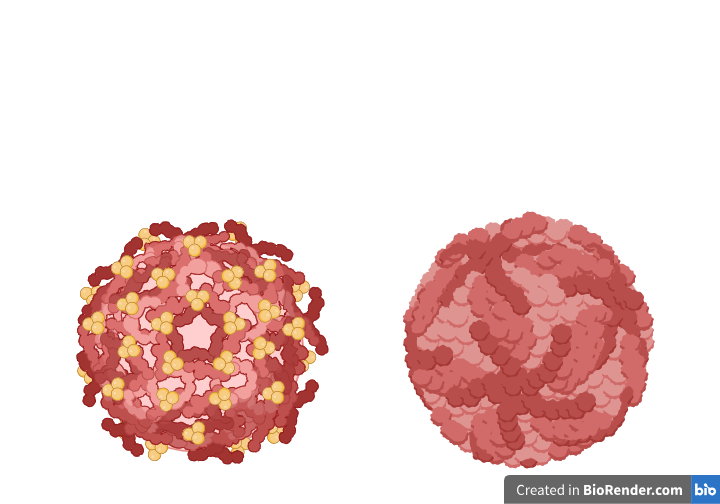


Figure 1: Immature flavivirus (left) and Dengue virus (right), created using BioRender

### Dengue virus discovery and genus classification

The Dengue virus, part of the Orthoflavivirus genus in the Flaviviridae family, was first identified during a 1943 epidemic in Nagasaki by Ren Kimura, Susumu Hotta, Albert B. Sabin, and Walter Schlesinger (Hotta, 1952; Mukhopadhyay, Kuhn and Rossmann, 2005). They discovered what is now known as the DENV-1 serotype. This virus is an enveloped, single-stranded RNA virus with a genome approximately 11 kilobases long. Its classification relies heavily on the envelope protein E, crucial for serotype determination, and the NS5 polymerase enzyme, essential for viral replication. Variations in the amino acid sequences of the E protein and nucleotide sequences in the NS5 region distinguish the four known Dengue serotypes. This delineation reveals a complex genomic structure characterized by a single open reading frame, fundamental for the virus's replication and interaction with host cells. (Hotta, 1952; Mukhopadhyay, Kuhn and Rossmann, 2005)

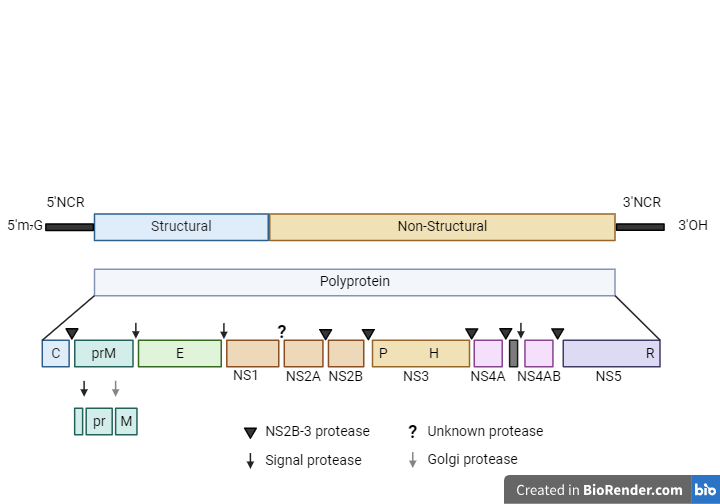
As seen in Figure 2, the genome organization of Orthoflavivirus is illustrated, showing the structural and non-structural protein coding regions within the polyprotein.  
  


Figure 2: Genome organization of Orthoflavivirus, created using BioRender

The phylogenetic tree depicted in Figure 3 illustrates the evolutionary relationships within the Orthoflavivirus genus. This tree was created using iTOL, showcasing the genetic diversity and lineage differentiation among various strains of the virus.



Figure 3: Phylogenetic tree of Orthoflavivirus genus, created using iTOL

### Dengue virus non-structural and structural proteins

The Dengue virus, belonging to the Orthoflavivirus genus in the Flaviviridae family, has a complex genomic structure encoding a polyprotein from its open reading frame (ORF1), as illustrated in Figure 4. This polyprotein is modified post-translationally to produce three structural and seven non-structural proteins, each essential for the virus's lifecycle. Among these, the non-structural protein NS3 functions as both a protease and helicase, while NS5 serves as the RNA-dependent RNA polymerase, critical for viral replication. The structural envelope protein E is crucial for virus entry into host cells and determines serotype specificity, particularly through domain III, which is key in receptor binding. This binding is a central focus for neutralizing antibodies and vaccine development due to its role in viral pathogenesis and immune evasion. Modifications to the E protein to minimize antibody cross-reactivity are vital in developing vaccines. Mapping the epitopes recognized by neutralizing antibodies has been instrumental in designing vaccines that effectively target multiple serotypes, highlighting the significance of the E protein in Dengue's pathology and its role in immune response strategies. (Lin *et al.*, 2012)

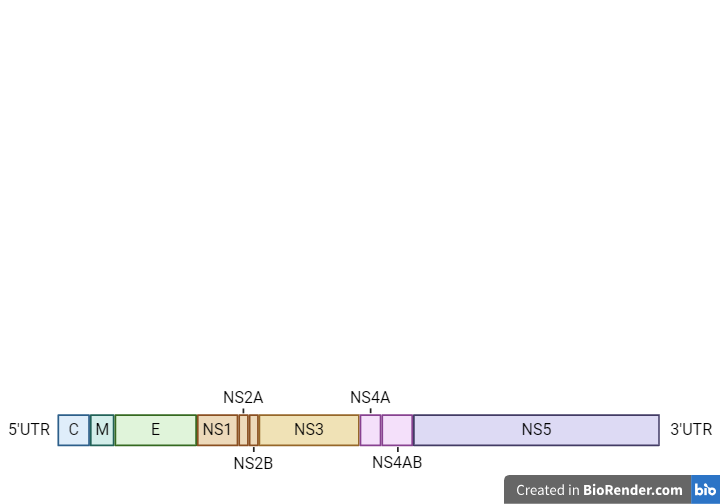


Figure 4: Genome organization of Dengue virus, created using BioRender

Table 1 provides the nomenclature for the structural and non-structural proteins of the Dengue virus, outlining their respective functions.

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

Dengue virus is primarily transmitted to humans through bites from infected *Aedes* mosquitoes, notably *Aedes aegypti* and *Aedes albopictus* (Gubler, 1998). After feeding on blood from an infected person, these mosquitoes undergo an incubation period of about eight to 12 days before they can transmit the virus (Rohani *et al.*, 2009). Their adaptability to urban environments and propensity to breed in stagnant water in man-made containers make urban and semi-urban areas prime spots for Dengue virus transmission. Urban expansion often leads to more stagnant water due to poor waste management, enhancing mosquito breeding sites (Mayer, Tesh and Vasilakis, 2017). Climate change further complicates this by altering temperatures and rainfall patterns, which can accelerate mosquito life cycles and increase their biting frequency, thus heightening the risk of Dengue virus outbreaks (Messina *et al.*, 2019). Additionally, socio-economic factors like migration increase human mobility, raising the potential for virus spread, especially in densely populated areas (Mondini and Chiaravalloti-Neto, 2008; Piovezan-Borges *et al.*, 2022; Langlois, 2024).

Though less common, Dengue virus can also be transmitted from mother to fetus, through blood transfusions, and via organ transplants (Ellis *et al.*, 2016). The expansion of Dengue virus’s geographical reach is driven by climate change and urbanization, affecting the distribution and breeding of *Aedes* mosquitoes and complicating disease control (Caminade, McIntyre and Jones, 2019). Addressing this challenge requires an integrated approach that considers the biological and socio-environmental factors influencing Dengue virus transmission (Malavige *et al.*, 2023; Langlois, 2024).

Climate change, by increasing global temperatures and altering rainfall patterns, has allowed *Aedes* mosquitoes to expand into previously unaffected regions, including parts of Europe, as seen in Table 2 (Caminade *et al.*, 2014). The El Niño phenomenon, along with higher humidity levels, further supports conditions conducive to mosquito growth and survival (Anyamba *et al.*, 2012). This shift was highlighted by the first recorded local Dengue virus transmissions in France and Croatia in 2010, indicating a significant geographic spread of the virus due to environmental changes. Developing a comprehensive table that correlates outbreaks in non-endemic areas with climate variables could shed light on these patterns. Such data is essential for improving surveillance systems that combine environmental, virological, and entomological information, thereby enhancing the ability to predict, detect early, and respond to Dengue virus outbreaks, as evidenced by studies on integrated disease surveillance. (Badurdeen *et al.*, 2013; Marques-Toledo *et al.*, 2019; *Dengue- Global situation*, 2023; *Dengue worldwide overview*, 2024; *Chikungunya, dengue et zika - Données de la surveillance renforcée en France hexagonale 2024*, 2024; EpiCentro, 2024; SPF, 2024)

Table 2: Table of Dengue Outbreaks in Non-endemic Regions

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **Region** | **Climate Factors** | **Outbreak Details** | **Source** |
| 2010 | France, Croatia | Warm summer, high humidity | First local transmissions recorded | (*Dengue- Global situation*, 2023) |
| 2014 | Portugal | Mild winter, wet spring | Multiple cases in Madeira |
| 2016 | Italy | Unusually hot summer | Localized outbreak in northern regions |
| 2019 | Germany | Extended warm periods | First autochthonous case detected |
| 2023 | Italy, France, Spain | Favorable warm conditions | Continued local transmissions in Southern Europe | (EpiCentro, 2024; SPF, 2024; *Chikungunya, dengue et zika - Données de la surveillance renforcée en France hexagonale 2024*, 2024) |

### Prevention

#### Vector Control Techniques

Efforts to manage dengue fever involve innovative approaches such as releasing genetically modified mosquitoes, which have substantially reduced mosquito populations in trial regions, showcasing a sustainable option (Alphey *et al.*, 2016). Another method is the Sterile Insect Technique (SIT), where sterilized male mosquitoes are released to cut down the population through non-productive mating (Hendrichs, Dyck and Robinson, 2005). Additionally, larvivorous fish in stagnant waters help by consuming mosquito larvae, further curtailing breeding sites. These strategies are recognized in research for their effectiveness in managing mosquito populations that transmit Dengue virus. (Eisen et al., 2009; Rather et al., 2017)

#### Chemical Control Methods

Traditional chemical controls remain fundamental in vector management, primarily using insecticides. However, increasing resistance among mosquitoes has driven the exploration of alternatives like plant-based insecticides. These botanical solutions are environmentally benign and aid in slowing resistance development. Insect Growth Regulators (IGRs) are another chemical approach, aimed at disrupting the mosquitoes' developmental stages, preventing them from becoming adults capable of spreading the virus. (Rather *et al.*, 2017)

#### Integrated Pest Management (IPM)

IPM merges physical, biological, and chemical tactics to control mosquito populations efficiently (Ghosh and Dash, 2007). It includes "attract-and-kill" strategies where pheromones draw mosquitoes into traps that subsequently eliminate them (Ghosh and Dash, 2007). IPM strives to minimize harmful chemical use while maintaining effective vector control, positioning it as a crucial approach for sustainable Dengue virus prevention. (Rather *et al.*, 2017)

#### Vaccine Development

Creating vaccines against Dengue virus remains a critical preventive measure, despite the challenges posed by the virus's four distinct serotypes and the need for efficacy across all age groups. Effective vaccines must provide protection against all Dengue virus serotypes to ensure comprehensive immunity. The deployment of these vaccines, combined with robust vector control strategies, significantly enhances the effectiveness of Dengue virus prevention programs (Achee *et al.*, 2015).

Currently, there are two prominent vaccines that have shown efficacy against Dengue virus. The first is Dengvaxia (CYD-TDV), which was the initial Dengue vaccine to receive licensing. It has demonstrated higher efficacy against DENV-3 and DENV-4, while showing moderate efficacy against DENV-1 and DENV-2. Dengvaxia is recommended for individuals aged 9 to 45 years who live in endemic areas and have had a prior Dengue virus infection (Sridhar *et al.*, 2018).

Another significant development in Dengue vaccination is the Takeda Dengue Vaccine (TAK-003), which recently received prequalification from the World Health Organization on May 10, 2024. TAK-003 has shown to offer protection across all four serotypes of Dengue virus. Clinical trials have demonstrated its safety and efficacy in both seropositive and seronegative individuals, making it a promising option for widespread use in endemic regions (Biswal *et al.*, 2019).

These vaccines represent significant progress in the fight against Dengue virus. The integration of vaccine deployment with effective vector control strategies, such as the use of genetically modified mosquitoes and the Sterile Insect Technique (SIT), can create a synergistic effect, greatly reducing the incidence of Dengue fever and aiding in the control of outbreaks (Achee *et al.*, 2015).

### Illness

#### Illness and Symptoms

Dengue virus infection typically manifests four to ten days after a mosquito bite, with symptoms lasting two to seven days. Common symptoms include high fever, severe headaches, eye pain, muscle and joint pains, nausea, vomiting, and skin rashes (Roy and Bhattacharjee, 2021). Approximately 5% of cases develop into severe dengue, characterized by intense abdominal pain, continuous vomiting, and bleeding, which require urgent medical attention. (Brady *et al.*, 2012; Bhatt *et al.*, 2013; Waggoner *et al.*, 2016; CDC, 2021)

#### Treatment and Medication

Dengue does not have a specific treatment; management focuses on alleviating symptoms. Key recommendations include staying well-hydrated and using acetaminophen for fever and pain relief. Aspirin and Non-steroidal anti-inflammatory drugs (NSAID)s are avoided to prevent bleeding risks. Severe dengue often necessitates hospitalization for supportive care. Promisingly, the experimental antiviral JNJ-1802 has shown potential in preliminary trials to reduce viremia and lessen disease severity. (Bhatt *et al.*, 2013; Goethals *et al.*, 2023; McCormack *et al.*, 2023; Kesteleyn *et al.*, 2024)

#### Antiviral Research and Microbiome Impact

Ongoing research into how antivirals can induce mutations presents a constant challenge in treating Dengue virus, emphasizing the need for continuous monitoring to address resistance (Beaucourt and Vignuzzi, 2014). Additionally, the effectiveness of Dengue virus vaccines may be significantly affected by the individual’s microbiome. Recent studies indicate that gut microbiota might influence the immune response to vaccines, potentially impacting their effectiveness (Zimmermann and Curtis, 2018). This suggests the need for personalized strategies in vaccine development and disease management to accommodate these complex interactions. (Hong, 2023)

### Viral Shedding

In Dengue virus infections, the transmission process is crucially dependent on the virus being shed into the bloodstream, which then can be picked up by *Aedes aegypti* and *Aedes albopictus* mosquitoes, the primary vectors (Gubler, 1998; Simmons *et al.*, 2012). After a human is bitten by an infected mosquito, the virus incubates for three to 14 days, usually manifesting symptoms within four to seven days. These symptoms range from none to severe and include high fever, headaches, muscle and joint pains, nausea, vomiting, and rash. In severe cases, Dengue virus can escalate to internal bleeding and organ failure. Research has shown that while the virus primarily sheds into the bloodstream, it is shed at much lower levels in other bodily fluids like saliva and urine, which are not generally involved in transmission (Halstead, Shotwell and Casals, 1973). This indicates that focusing on mosquito control and monitoring individuals with symptoms could effectively manage the spread of Dengue virus. Further studies confirm that non-vector transmission through bodily fluids other than blood is rare, underlining the importance of controlling mosquito populations and protecting against mosquito bites to interrupt the disease's transmission cycle (Godói *et al.*, 2017; Wilder-Smith *et al.*, 2017; Vogels *et al.*, 2023).

Table 3 provides a summary of Dengue virus shedding across various studies. The DengueSeq Study examined blood samples, revealing viral loads ranging from 10 to 100 copies/mL with moderate infectious virus levels. This study focused on genomic surveillance and provided insights into viral loads across different serotypes (Vaughn *et al.*, 2000; Guzman *et al.*, 2010). The CYD-TDV Vaccine Study, which looked at saliva and urine, found very low viral loads and no detectable infectious virus, suggesting a minimal environmental risk of virus uptake by mosquitoes due to limited replication in these vectors (Hadinegoro *et al.*, 2015). General virology research indicates that fecal shedding is generally low and not typically associated with transmission, as Dengue virus primarily transmits through vectors.

Table 3: Summary of Dengue virus Shedding

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study Reference** | **Fluids Examined** | **Viral Load (copies/mL)** | **Infectious Virus (copies/mL)** | **Remarks** |
| DengueSeq Study | Blood | 10^1 - 10^2 | Moderate | This study focused on genomic surveillance and provided insights into viral loads across different serotypes. |
| CYD-TDV Vaccine Study | Saliva, Urine | Very Low | Absent | The study noted that despite low-level viremia, the environmental risk of virus uptake by mosquitoes is low due to the limited replication in these vectors. |
| General Virology Research | Fecal | Not specified | Not typically associated with transmission | While not specific to Dengue virus, fecal shedding is generally low for viruses like dengue that primarily transmit through vectors. |

### Infectivity

The infectivity of the Dengue virus (DENV) significantly varies among its genotypes, which is crucial for understanding the dynamics of disease transmission and the severity of outbreaks. This variability is evident when comparing three distinct genotypes of DENV-2 genotype Cosmopolitan, DENV-2 genotype Asian I, and DENV-2 genotype Asian II. Each genotype exhibits unique characteristics in replication efficiency and disease-causing potential, impacting public health (Horstick, Tozan and Wilder-Smith, 2015).

The Cosmopolitan genotype, widespread across various regions, shows high replication capabilities in human hosts, indicating a strong infectivity profile that contributes to its global prevalence (Lee *et al.*, 2010). On the other hand, the Asian-I genotype, once the dominant strain in areas like Thailand, displays variable infectivity, which can greatly affect its epidemic potential and influence outbreak development and spread (Horstick, Tozan and Wilder-Smith, 2015).

Additionally, the DENV-2 genotype Asian II, mainly found in Malaysia and parts of the Indian subcontinent, is known for its adaptability to changing environmental conditions (Domingues *et al.*, 2008; Shahid *et al.*, 2020). This adaptability may enhance its ability to spread more efficiently than other genotypes, possibly leading to more frequent and widespread outbreaks. (Sarkar *et al.*, 2023; Samune *et al.*, 2024)

These variations highlight the need for genomic surveillance and targeted public health strategies that account for the specific infectivity and replication characteristics of different DENV genotypes. Implementing such strategies can aid in developing more effective preventive measures and treatments, specifically tailored to the dynamics of the DENV strains prevalent in various regions. (Sarkar *et al.*, 2023; Samune *et al.*, 2024)

Table 4 provides a summary of DENV infectivity by genotype, detailing the unique infectivity characteristics of each genotype.

Table 4: DENV infectivity by genotype

|  |  |  |
| --- | --- | --- |
| **Genotype** | **Description** | **Infectivity Notes** |
| DENV-2 genotype Cosmopolitan | Common in diverse geographic locations. | High replication and medium infectivity in human hosts |
| DENV-2 genotype Asian I | Previously dominant in Asia, particularly Thailand. | Varied infectivity with significant epidemic potential |
| DENV-2 genotype Asian II | Predominantly found in Malaysia and the Indian subcontinent areas. | High adaptability to changing environmental conditions |

## Detection and molecular characterisation

In virology, understanding the molecular characteristics of Dengue virus (DENV) serotypes is essential for grasping its epidemiological dynamics and devising public health strategies. The range of detection and genotyping techniques for DENV is broad, with conventional RT-PCR and real-time RT-PCR (RT-qPCR) being particularly effective for identifying viral RNA in clinical samples. The four-probe Taqman system used in RT-qPCR is notable for its ability to efficiently amplify viral RNA across various serotypes, highlighting the need to optimize thermal cycling conditions for successful detection. (Gurukumar *et al.*, 2009; Alm *et al.*, 2015)

Furthermore, detailed genotyping of DENV-2 is typically performed using techniques such as conventional PCR or nested PCR to amplify specific regions of the viral genome, particularly the envelope gene. This approach is followed by Sanger sequencing of the purified PCR products, which allows for an extensive examination of genetic diversity. This meticulous process elucidates phylogenetic relationships and epidemiological trends, offering vital data for vaccine development and outbreak management. The careful selection of samples based on specific criteria ensures the reliability and accuracy of the genotyping results. (Ko *et al.*, 2020; Rivera *et al.*, 2023)

Table 5 provides an overview of various DENV detection and genotyping methods, detailing their advantages and disadvantages.

Table 5: Overview of Dengue virus Detection and Genotyping Methods

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **Description** | **Advantages** | **Disadvantages** | **References** |
| Conventional RT-PCR | Standard polymerase chain reaction method for amplifying RNA sequences. | Widely used, cost-effective, good for initial screening. | Lower sensitivity compared to real-time methods, post-amplification processing required. | (Mullis and Faloona, 1987; Notomi *et al.*, 2000) |
| RT-qPCR | Real-time quantitative PCR that quantifies RNA in samples. | High sensitivity and specificity, quantitative results, rapid turnaround. | More expensive, requires calibration, and specialized equipment. | (Mackay, Arden and Nitsche, 2002) |
| RT-LAMP | Loop-mediated isothermal amplification; does not require thermal cycling. | Fast, does not require sophisticated equipment, can be used in field settings. | Lower specificity, potential for carry-over contamination. | (Notomi *et al.*, 2000; Parida *et al.*, 2005) |
| ELISA | Enzyme-linked immunosorbent assay to detect dengue-specific antibodies or antigens. | Simple, suitable for large-scale screening, cost-effective. | Lower sensitivity and specificity compared to molecular methods, indirect virus detection. | (Kuno, 1995) |
| Sanger Sequencing | Sequencing of PCR products to determine the genetic makeup of the virus. | High accuracy, allows for phylogenetic studies. | Time-consuming, expensive, requires sophisticated setup. | (Sanger, Nicklen and Coulson, 1977; Ewing and Green, 1998) |
| Short Read Amplicon Sequencing | Uses next-generation sequencing (NGS) to sequence specific regions of the genome. | Provides detailed insights into genetic variability, high throughput. | Requires complex data analysis, more expensive than Sanger. | (Margulies *et al.*, 2005; Shendure and Ji, 2008) |
| Short Read Metagenomics | NGS of all genetic material in a sample. | Comprehensive analysis, identifies co-infections. | Complex sample preparation, extensive computational resources needed. | (Turnbaugh *et al.*, 2007; Thomas, Gilbert and Meyer, 2012) |
| Long-Read Tiling Amplicon | Long-read sequencing technologies to cover larger genomic regions. | Allows for characterization of complex genomic regions, fewer assembly issues. | Higher error rates than short-read methods, expensive. | (Eid *et al.*, 2009; Koren *et al.*, 2012) |
| Long Read | Long-read sequencing to obtain much larger sequences in a single read. | Simplifies mapping and assembly of virus genomes, useful for identifying structural variants. | High cost, higher error rate than short-read technologies. | (Goodwin, McPherson and McCombie, 2016; Jain *et al.*, 2018) |

# Materials and Methods

Data extraction and analysis were conducted using custom scripts within an Anaconda-controlled environment, ensuring reproducibility and accuracy. Sequences were retrieved from the NCBI database using Entrez Direct and indexed with tools facilitated by BWA for efficient alignment. Clustering was performed using MMseqs2, and VADR amplicons were generated and extracted using the VADR library. A custom taxonomy database was created (Cuypers *et al.*, 2018; Schoch *et al.*, 2020; Hill *et al.*, 2024), mapping taxonomy data to the latest conventions. Primer pairs were designed and verified using a series of tools for literature mining, data conversion, alignment verification, melting temperature calculation, and primer pair generation.

## Execution Environment

All computational tasks were executed on the Sonic High-Performance Computing (HPC) cluster at University College Dublin (UCD). The Sonic HPC cluster is equipped with:

* CPU Nodes: 20 standard compute nodes, each with 2 Intel Xeon Gold 6152 processors (22 cores each) and 384GB of RAM.
* GPU Nodes: 3 GPU servers, each with 2 Nvidia Tesla V100 GPUs and 256GB of RAM.
* Storage: A parallel storage system (BeeGFS) across 4 storage servers, providing a total of 180GB shared storage, optimized for computational tasks.
* Network: InfiniBand interconnect for high-speed data transfer between nodes.

Access to the cluster is managed via a research VPN for off-campus access.

### CPU Information:

* Architecture: x86\_64
* CPU op-mode(s): 32-bit, 64-bit
* Byte Order: Little Endian
* CPU(s): 24
* On-line CPU(s) list: 0-23
* Thread(s) per core: 1
* Core(s) per socket: 12
* Socket(s): 2
* NUMA node(s): 2
* Vendor ID: GenuineIntel
* CPU family: 6
* Model: 85
* Model name: Intel(R) Xeon(R) Gold 5118 CPU @ 2.30GHz
* Stepping: 4
* CPU MHz: 1505.865
* CPU max MHz: 3200.0000
* CPU min MHz: 1000.0000
* BogoMIPS: 4600.00
* Virtualization: VT-x
* L1d cache: 32K
* L1i cache: 32K
* L2 cache: 1024K
* L3 cache: 16896K
* NUMA node0 CPU(s): 0,2,4,6,8,10,12,14,16,18,20,22
* NUMA node1 CPU(s): 1,3,5,7,9,11,13,15,17,19,21,23

### Network Configuration:

1: lo: <LOOPBACK,UP,LOWER\_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000

link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00

inet 127.0.0.1/8 scope host lo

valid\_lft forever preferred\_lft forever

inet6 ::1/128 scope host

valid\_lft forever preferred\_lft forever

2: em1: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc mq state UP group default qlen 1000

link/ether b0:26:28:4e:67:b8 brd ff:ff:ff:ff:ff:ff

inet 192.168.8.6/24 brd 192.168.8.255 scope global dynamic em1

valid\_lft 355921sec preferred\_lft 355921sec

inet6 fe80::b226:28ff:fe4e:67b8/64 scope link

valid\_lft forever preferred\_lft forever

3: em3: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN group default qlen 1000

link/ether b0:26:28:4e:67:b6 brd ff:ff:ff:ff:ff:ff

4: em2: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc mq state UP group default qlen 1000

link/ether b0:26:28:4e:67:b9 brd ff:ff:ff:ff:ff:ff

inet 137.43.92.104/23 brd 137.43.93.255 scope global dynamic em2

valid\_lft 27020sec preferred\_lft 27020sec

inet6 fe80::b226:28ff:fe4e:67b9/64 scope link

valid\_lft forever preferred\_lft forever

5: em4: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN group default qlen 1000

link/ether b0:26:28:4e:67:b7 brd ff:ff:ff:ff:ff:ff

6: ib0: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 2044 qdisc pfifo\_fast state UP group default qlen 256

link/infiniband 80:00:02:08:fe:80:00:00:00:00:00:00:98:03:9b:03:00:e2:63:a1 brd 00:ff:ff:ff:ff:12:40:1b:ff:ff:00:00:00:00:00:00:ff:ff:ff:ff

inet 192.168.12.6/24 brd 192.168.12.255 scope global ib0

valid\_lft forever preferred\_lft forever

inet6 fe80::9a03:9b03:e2:63a1/64 scope link

valid\_lft forever preferred\_lft forever

7: vrrp.86@em1: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000

link/ether 00:00:5e:00:01:56 brd ff:ff:ff:ff:ff:ff

inet 192.168.8.5/24 scope global vrrp.86

valid\_lft forever preferred\_lft forever

8: vrrp.87@em2: <BROADCAST,MULTICAST,UP,LOWER\_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000

link/ether 00:00:5e:00:01:57 brd ff:ff:ff:ff:ff:ff

inet 137.43.92.172/23 scope global vrrp.87

valid\_lft forever preferred\_lft forever

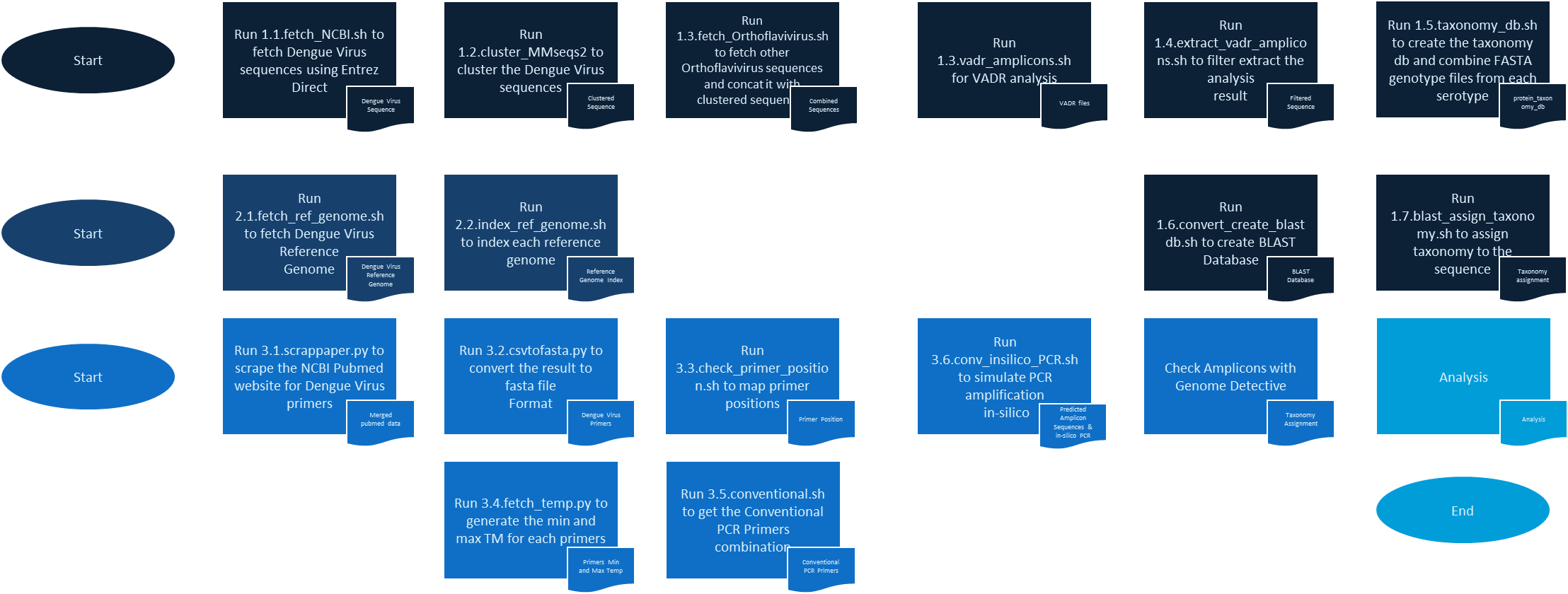


Figure 5: Process diagram

## Developing an in-house Dengue virus genotyping pipeline

### Fetching Dengue virus Sequences

Dengue virus sequences were retrieved from the National Center for Biotechnology Information (NCBI) database using the script 1.1.fetch\_NCBI.sh. This script was executed on a SLURM scheduler within an Anaconda environment that includes the Entrez Direct suite (version 13.2). The sequences were filtered by a predefined size range of 300 to 11000 bases and downloaded in FASTA format. Subsequent processing refined FASTA headers and extracted key sequence information, preparing data for further analysis.

* Script: 1.1.fetch\_NCBI.sh
* Packages and Libraries: Entrez Direct (version 13.2)
* Input File: NCBI nucleotide database query results
* Output File: 1.1.dengue\_virus\_sequences.fasta

### Clustering Dengue virus Sequences

The script 1.2.cluster\_MMseqs2.sh was used to cluster Dengue virus sequences using MMseqs2 (version 13-45111). The script created a database from the input FASTA file, performed clustering with a 97% identity threshold, and generated an output FASTA file containing representative sequences of each cluster.

* Script: 1.2.cluster\_MMseqs2.sh
* Packages and Libraries: MMseqs2 (version 13-45111)
* Input File: 1.1.dengue\_virus\_sequences.fasta
* Output Files:
  + MMseqs2 database: db/input\_db
  + Clustering results: clusters
  + Clustered sequences FASTA file: clustered\_sequences.fasta

### Fetching Orthoflavivirus Sequences

Orthoflavivirus sequences were retrieved from the National Center for Biotechnology Information (NCBI) database using the script 1.3.fetch\_combine\_orthoflavivirus.sh. This script was executed on a SLURM scheduler within an Anaconda environment that includes the Entrez Direct suite. The sequences were fetched based on a predefined list of accession numbers and downloaded in FASTA format. Subsequent processing included renaming FASTA headers to retain only accession IDs and extracting key sequence information for further analysis.

* Script: 1.3.fetch\_combine\_orthoflavivirus.sh
* Packages and Libraries: Entrez Direct, Anaconda (version 3.5.2)
* Input File: List of Orthoflavivirus accession numbers
* Output Files:
  + Orthoflavivirus sequences in FASTA format: orthoflavivirus.fasta
  + Orthoflavivirus sequence information in TSV format: orthoflavivirus.tsv

Combining FASTA Files The script combined the fetched Orthoflavivirus sequences with previously clustered dengue sequences to create a single FASTA file. This step ensured that all relevant sequences were included in the final dataset for subsequent analyses.

* Script: 1.3.fetch\_combine\_orthoflavivirus.sh
* Input Files:
  + Orthoflavivirus sequences in FASTA format: orthoflavivirus.fasta
  + Clustered sequences in FASTA format: clustered\_sequences.fasta
* Output File: Combined sequences in FASTA format: combined\_sequences.fasta

### Assigning Taxonomy Using Genome Detective

To set a ground truth, we used the Genome Detective Dengue virus Typing Tool, which is designed to utilize Blastx and phylogenetic methods to identify the Dengue virus serotypes, genotypes, and major lineages of nucleotide sequences. (Fonseca *et al.*, 2019; Vilsker *et al.*, 2019)

* Input Files:
  + Clustered Dengue Virus sequences: FASTA File
  + In-silico PCR Amplicons: Filtered in-silico PCR amplicon FASTA File
* Output Files:
  + Clustered Dengue Virus sequences: FASTA File, and Genotyping CSV Result
  + In-silico PCR Amplicons: FASTA File, Genotyping CSV Result

### Generating VADR Amplicons

The script 1.4.vadr\_amplicons.sh generated amplicons for VADR analysis. This script set up the environment for Perl scripts required for the analysis and utilized the v-annotate.pl script using VADR models specific to Flaviviridae from the VADR library (Schäffer *et al.*, 2020).

* Script: 1.4.vadr\_amplicons.sh
* Packages and Libraries: Anaconda (version 3.5.2), VADR (version 1.5.1), Perl
* Input File: filtered\_conv\_insilico\_PCR\_amplicons.fasta
* Output Files:
  + VADR output directory: vadr\_output
  + VADR annotated files

### Extracting VADR Amplicons

Following the generation of amplicons, the script 1.5.extract\_vadr\_amplicons.sh was employed to extract VADR amplicons. This script merged all .sqa and .sqc files and used a Python script (extract\_vadr.py) to filter and retain only the sequences that passed quality checks.

* Script: 1.5.extract\_vadr\_amplicons.sh
* Packages and Libraries: Python (version 3.9.15), Perl
* Tools: cat (for merging files), custom Python script extract\_vadr.py
* Input Files:
  + Merged .sqa files: vadr\_combined.sqa
  + Merged .sqc files: vadr\_combined.sqc
  + Original FASTA file: filtered\_conv\_insilico\_PCR\_amplicons.fasta
* Output File: passed\_sequences.fasta

### Creating Taxonomy Database

The 1.6.taxonomy\_db.sh script was utilized to create the protein\_db\_taxonomy.tsv file and to combine FASTA files of each genotype from their respective serotypes. The script performed several key tasks:

1. Downloading Protein Sequences: The script identified and downloaded protein sequences for various Dengue virus genotypes from the NCBI database. The genotypes included were carefully selected to match the new lineage nomenclature, avoiding duplicates.
2. Renaming Sequence Headers: Each downloaded sequence was renamed according to the format "DENV-X\_genotype\_Y", where "X" represents the serotype number and "Y" represents the genotype.
3. Combining FASTA Files: The renamed sequences were then combined into serotype-specific FASTA files, resulting in files such as combined\_denv1.fasta, combined\_denv2.fasta, etc.

The taxonomy data were sourced from NCBI and formatted as follows: Amarillovirales;Flaviviridae;Orthoflavivirus;Orthoflavivirus\_denguei;dengue\_virus;dengue\_virus\_type\_X;Genotype\_name. This taxonomy was mapped to the latest naming convention described in the paper "A new lineage nomenclature to aid genomic surveillance of Dengue virus." Additionally, the nomenclature used in Genome Detective was mapped to the new lineage nomenclature, as seen in Table 6: Nomenclature of Dengue virus Genotype. (Cuypers *et al.*, 2018; Schoch *et al.*, 2020; Hill *et al.*, 2024)

During the process, multiple genotypes with the same nomenclature (e.g., "DENV-2 genotype Asian I", "DENV-2 genotype Asian II") were found. To ensure clarity and avoid redundancy, representative genotypes were selected for the protein taxonomy database. The chosen representative genotypes are as follows:

* Dengue virus 1 Brazil/97-11/1997 for "DENV-1 genotype V"
* Dengue virus 1 Jamaica/CV1636/1977 for "DENV-1 genotype II"
* Dengue virus 1 Nauru/West Pac/1974 for "DENV-1 genotype I"
* Dengue virus 1 Singapore/S275/1990 for "DENV-1 genotype IV"
* Dengue virus 1 Thailand/AHF 82-80/1980 for "DENV-1 genotype III"
* Dengue virus 2 16681-PDK53 for "DENV-2 genotype Asian I"
* Dengue virus 2 China/D2-04 for "DENV-2 genotype Asian II"
* Dengue virus 2 Jamaica/1409/1983 for "DENV-2 genotype American"
* Dengue virus 2 Peru/IQT2913/1996 for "DENV-2 genotype Cosmopolitan"
* Dengue virus 3 China/80-2/1980 for "DENV-3 genotype III"
* Dengue virus 3 Martinique/1243/1999 for "DENV-3 genotype IV"
* Dengue virus 3 Philippines/H87/1956 for "DENV-3 genotype I"
* Dengue virus 3 Singapore/8120/1995 for "DENV-3 genotype II"
* Dengue virus 3 Sri Lanka/1266/2000 for "DENV-3 genotype V"
* Dengue virus 4 Dominica/814669/1981 for "DENV-4 genotype II"
* Dengue virus 4 Philippines/H241/1956 for "DENV-4 genotype I"

This selection process ensured that only unique, representative genotypes were used, thus streamlining the taxonomy database.

* Script: 1.6.taxonomy\_db.sh
* Packages and Libraries: Python (version 3.7.4)
* Input Files: Individual FASTA files for each Dengue virus genotype
* Output Files:
  + Combined FASTA files for each serotype (combined\_denv1.fasta, combined\_denv2.fasta, combined\_denv3.fasta, combined\_denv4.fasta)
  + protein\_db\_taxonomy.tsv

This script ensured that the taxonomy data were accurately mapped and that the FASTA files were organized in a way that facilitated downstream analyses, adhering to the latest genomic surveillance standards.

### Converting and Creating BLAST Database

The script 1.7.convert\_create\_blastdb.sh converted the combined FASTA files into a format suitable for creating a BLAST database. This process utilized the fasta2RDP.py script from the FROGS library (https://github.com/geraldinepascal/FROGS-wrappers) and created a BLAST database. (Escudié *et al.*, 2018; Bernard *et al.*, 2021)

* Script: 1.7.convert\_create\_blastdb.sh
* Packages and Libraries: Python (version 3.7.4), FROGS library
* Tools: seqkit (for header manipulation), BLAST (version 2.12.0)
* Input Files:
  + Combined FASTA file: combined\_denv4.fasta, combined\_denv3.fasta, combined\_denv2.fasta, combined\_denv1.fasta
  + Taxonomy file: protein\_db\_taxonomy.tsv
* Output Files:
  + RDP formatted FASTA and taxonomy files
  + BLAST database

### Assigning Taxonomy Using BLAST

The script 1.8.blast\_assign\_taxonomy.sh assigned taxonomy to the sequences using BLAST. This script ran BLAST searches and used the modified taxonomy\_assignment.py script, originally from Joseph7e's repository (Assign-Taxonomy-with-BLAST), and then used the taxonomy\_assignment\_BLAST.py script to assign taxonomy based on BLAST results. Additionally, the best\_taxonomy\_assignment.py script was used to determine the best taxonomy assignment for each sequence (Sevigny, 2024).

* Script: 1.8.blast\_assign\_taxonomy.sh
* Packages and Libraries: BLAST (version 2.12.0), Biopython
* Tools: seqkit (for removing duplicate sequences), best\_taxonomy\_assignment.py,
* Input Files:
  + Query FASTA file: passed\_sequences.fasta;
  + BLAST database
  + Taxonomy file: protein\_db\_taxonomy.tsv
* Output Files:
  + BLAST search results
  + Taxonomy assignment results

## Fetching and Indexing Dengue virus Reference Genome – Process 2

### Fetching Reference Genomes

The script 2.1.fetch\_ref\_genome.sh was used to retrieve Dengue virus reference genomes from the National Center for Biotechnology Information (NCBI) database (NCBI Resource Coordinators, 2018). This script utilized the Entrez Direct package (version 13.2) to filter genomes by completeness and relevance, saving them in the specified directory for further processing (Kans, 2024). The genomes were fetched using accession numbers and were saved in FASTA format for subsequent analyses.

* Script: 2.1.fetch\_ref\_genome.sh
* Packages and Libraries: Entrez Direct (version 13.2)
* Input Files: NCBI nucleotide database query results
* Output Files: Dengue virus reference genomes in FASTA format

### Indexing Reference Genomes

After fetching, the reference genomes were indexed using the script 2.2.index\_ref\_genome.sh. This script utilized the Burrows-Wheeler Aligner (BWA) tool (version 0.7) to create indices for efficient sequence alignment (Li and Durbin, 2009).

* Script: 2.2.index\_ref\_genome.sh
* Packages and Libraries: BWA (version 0.7)
* Input Files: Dengue virus reference genomes in FASTA format
* Output Files: Indexed reference genomes.

## Dengue virus-Specific Primers and Amplicons – Process 3

### Literature Mining for Dengue virus-Specific Primers

Using the 3.1.scrappaper.py script, forked from the ScrapPaper repository by M. R. Rafsanjani, literature on Dengue virus primers was mined from PubMed (Rafsanjani, 2022). This script employed the requests and BeautifulSoup libraries (version 4.9.3) to scrape metadata from articles, distinguishing between summary and abstract views to maximize data extraction efficiency. The results from both views were merged based on title similarity, organized into a DataFrame, and exported to an Excel spreadsheet. Out of 404 articles initially identified, data was extracted from 158 after manual review, which focused on conventional PCR assays related to Dengue virus primers typing. This review helped document essential details such as target genotype, primer name, orientation, and sequence.

Records were removed for the following reasons: inaccessible articles behind paywalls, missing articles, and the complexity of locating primer details. After reviewing, data was cleaned in Excel to ensure accuracy and consistency. This process involved removing entries without primers, renaming unnamed primers to a standardized format, deleting tags, probes, RT, and duplicates, standardizing primer names by replacing spaces with underscores, and ensuring all primers had associated sequences. A total of 1277 unique primers were identified for subsequent analysis steps.

* Script: 3.1.scrappaper.py
* Packages and Libraries: Python (version 3.9.15), requests, BeautifulSoup (version 4.9.3)
* Input Files: PubMed articles
* Output Files: Dengue virus primer metadata CSV file

### Converting Cleaned Primer Data to FASTA Format

The script 3.2.csvtofasta.sh converted the cleaned CSV file containing primer data into FASTA format, preparing it for further processing.

* Script: 3.2.csvtofasta.sh
* Packages and Libraries: Python (version 3.9.15)
* Input Files: Primer metadata CSV file
* Output Files: Primer sequences in FASTA format

### Checking Primer Positions

The script 3.3.check\_primer\_position.sh checked the positions of primers against a set of reference genomes using BWA for alignment and samtools for handling sequence data, ensuring the accuracy of primer locations (Li and Durbin, 2009; Li *et al.*, 2009).

* Script: 3.3.check\_primer\_position.sh
* Packages and Libraries: BWA (version 0.7), samtools (version 1.10)
* Input Files: Primer FASTA file, reference genomes
* Output Files: Primer mapping positions TSV file

### Determining Primer Melting Temperatures

To initiate the primer pairs generation process, it was crucial to determine the minimum and maximum melting temperatures (Tm) of each primer sequence. The 3.4.fetch\_temp.py script analyzed an Excel file containing primer sequences and calculated Tm using Wallace's rule and the nearest neighbor method. Wallace's rule is a simple method for estimating the melting temperature of DNA primers by assigning 2°C for each A or T and 4°C for each G or C in the sequence (Suggs *et al.*, 1981). This script handled sequences with ambiguous bases by mapping ambiguous nucleotide symbols to their actual nucleotides, generating all possible combinations of sequences, and then calculating the average Tm for these combinations. Results were updated in the primer\_metadata.csv file, saved in the same directory as the input file.

* Script: 3.4.fetch\_temp.py
* Packages and Libraries: Anaconda (version 3.5.2), Biopython (version 1.78), pandas (version 1.3.3), openpyxl (version 3.0.7)
* Input Files: Primer sequences
* Output Files: Primer Tm values

### Generating Primer Pairs

#### Conventional PCR

To generate primer pairs for conventional PCR, the process involved combining forward and reverse primers based on their mapping positions, melting temperatures (Tm), and other sequence characteristics. The 3.5.conventional.sh script was executed to run the conventional.py script.

The conventional.py script started by calculating the GC content of each primer sequence, a crucial factor affecting primer stability and annealing temperature. The script ensured that each primer met specific conditions: length between 18 to 25 nucleotides, GC content ranging from 40% to 60%, and no homopolymer runs exceeding four nucleotides (Shu *et al.*, 2003).

The script loaded mapping data from a TSV file containing primer positions and metadata, as well as additional primer metadata from primer\_metadata.tsv, which included GC content calculations. After merging these datasets to create a comprehensive dataset, forward and reverse primers were separated based on their names ending with \_F and \_R.

The script iterated through the forward primers to find matching reverse primers that satisfied all predefined conditions. It ensured the primer pairs amplified regions within specified regions of interest (NS1, NS3, and NS5 of the Dengue virus genome) and that the Tm differences between forward and reverse primers were within 3°C for both minimum and maximum Tm. Additionally, the GC content difference between primers did not exceed 10%, and the length of the PCR product (amplicon) was between 300 to 1200 base pairs (Gurukumar *et al.*, 2009; Wang *et al.*, 2016).

The valid primer combinations were saved to conventional\_primer\_combinations.tsv

* Script: 3.5.conventional.sh
* Packages and Libraries: Python (version 3.9.15)
* Input Files: Primer mapping positions TSV file
* Output Files: Primer combinations TSV file

### In-silico PCR

#### Conventional PCR

The 3.6.conv\_insilico\_PCR.sh script simulated conventional PCR reactions in silico using the primer pairs generated. The script utilized the in-silico PCR tool and a custom Python script to filter the resulting amplicons (Ozer, 2024).

This script first performs in-silico PCR. Then, the output amplicons are filtered to retain only those within the desired size range of 300-1200bp.

* Script: 3.6.conv\_insilico\_PCR.sh
* Packages and Libraries: Perl (version 5.30), Python (version 3.9.15)
* Tools: In-silico PCR tool (https://github.com/egonozer/in\_silico\_pcr), custom Python script filter\_amplicons.py
* Input Files: Primer combinations TSV file, combined\_sequences.fasta (clustered dengue sequences + other Orthoflavivirus sequences)
* Output Files: Filtered in-silico PCR amplicon FASTA File

## Results/Analysis Generation – Process 4

### Conventional PCR

#### Link Conventional Amplicons

The 4.1.1.link\_conv\_amplicons.sh script was used to filter and link conventional PCR amplicons to their corresponding accession numbers and primer pairs. This script utilized a custom Python script, link\_conv\_amplicons.py, to process the output of the in-silico PCR results.

The script started by loading the necessary Python module and defining the input file paths for the results, amplicons, and primer combinations. The Python script was then executed to filter amplicons based on their lengths (between 300 and 1200 base pairs) and link them to their accession numbers and primer pairs. The output was saved to a CSV file for further analysis.

* Script: 4.1.1.link\_conv\_amplicons.sh
* Packages and Libraries: Python (version 3.9.15)
* Input Files:
  + conv\_insilico\_PCR\_results.txt (in-silico PCR results)
  + conv\_insilico\_PCR\_amplicons.fasta (amplicon sequences)
  + conventional\_primer\_combinations.tsv (primer pairs)
* Output Files: filtered\_linked\_amplicons.csv (filtered and linked amplicons)

#### Generating PCA Dataset for Conventional Primers

The generation of the Principal Component Analysis (PCA) dataset for conventional primers was accomplished using the script 9.1.1.generate\_FAMD\_dataset.sh. This script was executed on a SLURM scheduler within a Python 3.9.15 environment. The script utilized a custom Python script (PCA\_dataset.py) to process and extract detailed primer information from various TSV files.

The PCA\_dataset.py script was designed to process and extract detailed information about individual primers from various input TSV files. The script begins by setting up logging to provide detailed information about the processing steps and any potential errors. It then loads data from the mapping positions file, primer metadata file, and conventional primer combinations file into Pandas DataFrames.

Next, the script merges the mapping and metadata DataFrames based on the 'Primer' column to combine relevant information. The script then iterates through the merged DataFrame to extract details of individual primers that meet specific conditions, such as length (between 18 to 25 nucleotides), GC content (between 40% to 60%), and absence of homopolymer runs exceeding four nucleotides. Primers are categorized as 'Forward' or 'Reverse' based on their names.

The script identifies valid primers based on their positions within predefined regions of interest (NS1, NS3, NS5). For each valid primer, details such as the primer name, sequence, genotype, orientation, start and end positions, amplicon length, region, GC content, and melting temperatures (Tm) are recorded.

Finally, the script saves the details of valid primers to the output TSV file after removing duplicates and rearranging columns as required. This structured dataset facilitates further analysis and downstream applications..

* Script: 9.1.1.generate\_FAMD\_dataset.sh
* Packages and Libraries: Python (version 3.9.15), pandas (version 0.23.0), sys, tqdm (version 4.43.0), logging, itertools
* Input Files:
  + Mapping positions file: mapping\_positions.tsv
  + Primer metadata file: primer\_metadata.tsv
  + Conventional primer combinations file: conventional\_primer\_combinations.tsv
* Output File: individual\_conv\_primer\_details.tsv

#### Generating Combined PCA Dataset for Conventional Primers

The generation of the combined Principal Component Analysis (PCA) dataset for conventional primers was performed using the script 9.1.2.generate\_FAMD\_dataset\_combined.sh. This script was executed on a SLURM scheduler within a Python 3.9.15 environment and utilized a custom Python script (PCA\_dataset\_combined.py) to process and extract detailed primer combination information from various TSV files.

The PCA\_dataset\_combined.py script is designed to generate detailed information about primer combinations by processing various input TSV files. The script begins by setting up logging to capture detailed information about the processing steps and any potential errors. It then loads mapping data and primer metadata from the provided TSV files into Pandas DataFrames, ensuring that the 'Primer' column is properly formatted. These DataFrames are then merged to combine relevant information.

Next, the script separates the merged DataFrame into forward and reverse primers. It iterates through each forward primer, checking specific conditions such as length, GC content, and the absence of homopolymer runs. For each valid forward primer, the script identifies matching reverse primers based on criteria like position, genotype compatibility, and the required length of the amplicon.

The script evaluates potential primer combinations based on additional criteria, including melting temperature (Tm) differences and GC content differences. Valid combinations are those where the forward and reverse primers meet these criteria and fall within predefined regions of interest (NS1, NS3, NS5).

Finally, the script compiles detailed information for each valid primer combination, such as amplicon length, GC content, Tm values, and sequence information, and saves this information to an output TSV file after removing duplicates and arranging columns appropriately. This structured output facilitates further analysis and downstream applications.

* Script: 9.1.2.generate\_FAMD\_dataset\_combined.sh
* Packages and Libraries: Python (version 3.9.15), pandas (version 0.23.0), sys, tqdm (version 4.43.0), logging, itertools
* Input Files:
  + Mapping positions file: mapping\_positions.tsv
  + Primer metadata file: primer\_metadata.tsv
* Output File: individual\_conv\_primer\_combination\_details.tsv

#### Generate Serotype and Genotype Confusion Matrix and Extract Information

The script generate\_confusion\_matrix.py was used to process dengue virus sequence data, generate confusion matrices for serotype and genotype classifications, and calculate various performance metrics for different primer pairs. This script leveraged multiple libraries, including Pandas, Biopython, Scikit-learn, Seaborn, Matplotlib, and SciPy.

* Script: generate\_confusion\_matrix.py
* Packages and Libraries: Pandas (v2.0.3), Biopython (v1.83), Scikit-learn (v1.3.2), Seaborn (v0.13.2), Matplotlib (v3.7.5), SciPy (v1.10.1)
* Input Files:
  + Amplicon CSV files: p\*.csv
  + Amplicon FASTA files: p\*.fasta
  + Clustered file: test2.csv
  + Primer file: filtered\_linked\_amplicons.csv
* Output Files:
  + Combined output CSV: output.csv
  + Serotype confusion matrix PNG: serotype\_confusion\_matrix.png
  + Genotype confusion matrix PNG: genotype\_confusion\_matrix.png
  + Overall metrics CSV: metrics\_overall.csv
  + Best F1 score text file: best\_f1\_score.txt
  + Serotype metrics CSV files: metrics\_serotype\_{serotype}.csv
  + Genotype metrics CSV files: metrics\_genotype\_{genotype}.csv
  + Mismatched classifications CSV: mismatched\_classifications.csv
  + Missed classifications CSV: missed\_classifications.csv
  + Statistical test results text file: statistical\_test\_results.txt

#### Generating Heatmaps for Missed and Mismatched Classifications

The script generate\_heatmaps.py was used to create heatmaps for missed and mismatched classifications of dengue virus genotypes. The script leverages Pandas, Seaborn, and Matplotlib libraries to process and visualize the data.

* Script: generate\_heatmaps.py
* Packages and Libraries:
  + Pandas (version 2.0.3)
  + Seaborn (version 0.13.2)
  + Matplotlib (version 3.7.5)
* Input Files:
  + Missed classifications file: missed\_classifications.csv
  + Mismatched classifications file: mismatched\_classifications.csv
* Output Files:
  + Mismatched classifications heatmap PNG: mismatched\_heatmap.png
  + Missed classifications heatmap PNG: missed\_heatmap.png

#### Performing PCA on Individual Primers by Serotype

The script pca\_individual\_primers.py was used to perform Principal Component Analysis (PCA) on individual primers by serotype. The script leverages Pandas, Matplotlib, NumPy, and Scikit-learn libraries to process, analyze, and visualize the data.

* Script: pca\_individual\_primers.py
* Packages and Libraries:
  + Pandas (version 2.0.3)
  + Matplotlib (version 3.7.5)
  + NumPy (version 1.24.3)
  + Scikit-learn (version 1.3.2)
* Input File: individual\_conv\_primer\_details.tsv
* Output Files:
  + PCA plot PNG: PCA\_FINAL.png
  + Row Cos2 CSV: row\_cos2.csv
  + Column Contrib CSV: col\_contrib.csv

#### Performing PCA on Primer Combinations by Serotype

The script pca\_primer\_combinations.py was used to perform Principal Component Analysis (PCA) on primer combinations by serotype. The script leverages Pandas, Matplotlib, NumPy, and Scikit-learn libraries to process, analyze, and visualize the data.

* Script: pca\_primer\_combinations.py
* Packages and Libraries:
  + Pandas (version 2.0.3)
  + Matplotlib (version 3.7.5)
  + NumPy (version 1.24.3)
  + Scikit-learn (version 1.3.2)
* Input File: individual\_conv\_primer\_combination\_details.tsv
* Output Files:
  + PCA plot PNG: PCA\_combined\_with\_serotype\_filtered.png
  + Row Cos2 CSV: row\_cos2\_combined.csv
  + Column Contrib CSV: col\_contrib\_combined.csv

#### Generating Heatmap for Conventional Primer Details

The heatmap visualization for conventional primer details was created using a Python script designed to process and display data from a TSV file. This script leverages libraries such as pandas for data manipulation, matplotlib for plotting, and patches for graphical representations.

* Script: generate\_heatmap.py
* Packages and Libraries:
  + pandas (version 2.0.3)
  + matplotlib (version 3.7.5)
* Input File:
  + Primer details file: individual\_conv\_primer\_details.tsv
* Output File:
  + Heatmap image file: regions\_heatmap\_labeled\_v4.png

#### Generating Heatmap for Combined Primer Details

The heatmap visualization for combined primer details was created using a Python script designed to process and display data from a TSV file. This script leverages libraries such as pandas for data manipulation, matplotlib for plotting, and patches for graphical representations.

* Script: generate\_heatmap\_combined.py
* Packages and Libraries:
  + pandas (version 2.0.3)
  + matplotlib (version 3.7.5)
* Input file:
  + Primer combination details file: individual\_conv\_primer\_combination\_details.tsv
* Output File:
  + Heatmap image file: regions\_heatmap\_combined\_labeled\_v4.png

#### Analysis of Primer Data from 1995 Onwards

This script performs a comprehensive analysis of primer data, focusing on trends over the years from 1995 onwards. It uses several Python libraries, including pandas for data manipulation, matplotlib and seaborn for visualization, and statsmodels for statistical analysis.

* Script: analysis\_of\_primer\_data.py
* Packages and Libraries:
  + pandas (version 2.0.3)
  + matplotlib (version 3.7.5)
  + seaborn (version 0.13.2)
  + statsmodels (version 0.14.1)
* Input Files:
* Excel file: all\_cleanned\_v1\_update.xlsx
* Output Files:
  + CSV files: year\_serotype\_counts.csv, summary\_table.csv, descriptive\_statistics.csv, primer\_length\_statistics.csv
  + Text file: peak\_counts.txt, statistics\_results.txt
  + Image files: primers\_by\_year\_and\_serotype.png, primers\_by\_year\_and\_orientation.png, primer\_length\_distribution.png, num\_primers\_and\_publications.png, weighted\_num\_primers.png

#### Evaluation of Genotype and Serotype Predictions

This script performs an evaluation of genotype and serotype predictions by calculating various metrics, including precision, recall, F1 score, specificity, NPV, MCC, balanced accuracy, and ROC AUC. It also generates and saves confusion matrices as heatmaps. The script uses several Python libraries, including pandas for data manipulation, matplotlib and seaborn for visualization, and scikit-learn for metric calculations.

* Script: evaluation\_of\_genotype\_serotype\_predictions.py
* Packages and Libraries:
  + pandas (version 2.0.3)
  + numpy (version 1.24.3)
  + matplotlib (version 3.7.5)
  + seaborn (version 0.13.2)
  + scikit-learn (version 1.3.2)
* Input Files:
  + CSV file: Best Taxonomy Assignment File from each Cutoff Value
* Output Files:
  + Text file: metrics.txt
  + Image files: confusion\_matrix\_genotypes.png, confusion\_matrix\_serotypes.png

## Phylogenetic Tree Creation

Table 6: Orthoflaviviruses grouped by vector and host

|  |  |  |  |
| --- | --- | --- | --- |
| **Virus species** | **Virus name** | **Accession number** | **Virus abbreviation** |
| Tick-borne, mammalian host | | |  |
| Orthoflavivirus gadgetsense | Gadgets Gully virus | DQ235145 | GGYV |
| Orthoflavivirus kyasanurense | Kyasanur Forest disease virus | AY323490 | KFDV |
|  | Alkhumra hemorrhagic fever virus | AF331718 | AHFV |
| Orthoflavivirus langatense | Langat virus | AF253419 | LGTV |
| Orthoflavivirus loupingi | Louping ill virus | Y07863 | LIV |
|  | British subtype | D12937 | LIV-Brit |
|  | Irish subtype | X86784 | LIV-Ir |
|  | Spanish subtype | DQ235152 | LIV-Spain |
|  | Turkish sheep encephalitis virus subtype | DQ235151 | TSEV |
|  | Greek goat encephalitis virus subtype | DQ235153 | GGEV |
| Orthoflavivirus omskense | Omsk hemorrhagic fever virus | AY193805 | OHFV |
| Orthoflavivirus powassanense | Powassan virus | L06436 | POWV |
|  | deer tick virus | AF311056 | DTV |
| Orthoflavivirus royalense | Royal Farm virus | DQ235149 | RFV |
| Orthoflavivirus encephalitidis | European subtype | U27495 | TBEV-Eur |
|  | Far Eastern subtype | X07755 | TBEV-FE |
|  | Siberian subtype | L40361 | TBEV-Sib |
| Tick-borne, seabird host | | |  |
| Orthoflavivirus meabanense | Meaban virus | DQ235144 | MEAV |
| Orthoflavivirus saumarezense | Saumarez Reef virus | DQ235150 | SREV |
| Orthoflavivirus tyuleniyense | Tyuleniy virus | KF815939 | TYUV |
| Probably tick-borne | |  |  |
| Orthoflavivirus kadamense | Kadam virus | DQ235146 | KADV |
| Mosquito-borne, Aroa virus group | | |  |
| Orthoflavivirus aroaense | Aroa virus | AY632536 | AROAV |
|  | Bussuquara virus | AF013366 | BSQV |
|  | Iguape virus | AF013375 | IGUV |
|  | Naranjal virus | AF013390 | NJLV |
| Mosquito-borne, Dengue virus group | | | |
| Orthoflavivirus denguei | Dengue virus 1 | U88536 | DENV-1 |
|  | Dengue virus 2 | U87411 | DENV-2 |
|  | Dengue virus 3 | M93130 | DENV-3 |
|  | Dengue virus 4 | AF326573 | DENV-4 |
| Mosquito-borne, Japanese encephalitis virus group | | | |
| Orthoflavivirus cacipacoreense | Cacipacoré virus | KF917536 | CPCV |
| Orthoflavivirus japonicum | Japanese encephalitis virus | M18370 | JEV |
| Orthoflavivirus koutangoense | Koutango virus | AF013384 | KOUV |
| Orthoflavivirus murrayense | Alfuy virus | AF013360 | ALFV |
|  | Murray Valley encephalitis virus | AF161266 | MVEV |
| Orthoflavivirus louisense | St. Louis encephalitis virus | DQ525916 | SLEV |
| Orthoflavivirus usutuense | Usutu virus | AY453411 | USUV |
| Orthoflavivirus nilense | Kunjin virus | D00246 | KUNV |
|  | West Nile virus | M12294 | WNV |
| Orthoflavivirus yaoundeense | Yaoundé virus | AF013413 | YAOV |
| Mosquito-borne, Kokobera virus group | | | |
| Orthoflavivirus kokoberaorum | Kokobera virus | AY632541 | KOKV |
|  | Stratford virus | AF013407 | STRV |
| Mosquito-borne, Ntaya virus group | | |  |
| Orthoflavivirus bagazaense | Bagaza virus | AY632545 | BAGV |
| Orthoflavivirus ilheusense | Ilhéus virus | AY632539 | ILHV |
|  | Rocio virus | AF013397 | ROCV |
| Orthoflavivirus israelense | Israel turkey meningoencephalitis virus | AF013377 | ITV |
| Orthoflavivirus ntayaense | Ntaya virus | JX236040 | NTAV |
| Orthoflavivirus tembusu | Tembusu virus | JF895923 | TMUV |
| Orthoflavivirus zikaense | Zika virus | AY632535 | ZIKV |
| Mosquito-borne, yellow fever virus group | | | |
| Orthoflavivirus sepikense | Sepik virus | DQ837642 | SEPV |
| Orthoflavivirus wesselsbronense | Wesselsbron virus | EU707555 | WESSV |
| Orthoflavivirus flavi | yellow fever virus | X03700 | YFV |
| Probably mosquito-borne, Kedougou virus group | | | |
| Orthoflavivirus kedougouense | Kédougou virus | AY632540 | KEDV |
| Probably mosquito-borne, Edge Hill virus group | | | |
| Orthoflavivirus banziense | Banzi virus | DQ859056 | BANV |
| Orthoflavivirus boubouiense | Bouboui virus | DQ859057 | BOUV |
| Orthoflavivirus edgehillense | Edge Hill virus | DQ859060 | EHV |
| Orthoflavivirus jugraense | Jugra virus | DQ859066 | JUGV |
| Orthoflavivirus saboyaense | Potiskum virus | DQ859067 | POTV |
|  | Saboya virus | DQ859062 | SABV |
| Orthoflavivirus ugandaense | Uganda S virus | DQ859065 | UGSV |
| Unknown vector, Entebbe bat virus group | | | |
| Orthoflavivirus entebbeense | Entebbe bat virus | DQ837641 | ENTV |
|  | Sokuluk virus | AF013405 | SOKV |
| Orthoflavivirus yokoseense | Yokose virus | AB114858 | YOKV |
| Unknown vector, Modoc virus group | | |  |
| Orthoflavivirus apoiense | Apoi virus | AF160193 | APOIV |
| Orthoflavivirus cowboneense | Cowbone Ridge virus | AF013370 | CRV |
| Orthoflavivirus jutiapaense | Jutiapa virus | KJ469371 | JUTV |
| Orthoflavivirus modocense | Modoc virus | AJ242984 | MODV |
| Orthoflavivirus viejaense | Sal Vieja virus | AF013401 | SVV |
| Orthoflavivirus perlitaense | San Perlita virus | AF013402 | SPV |
| Unknown vector, Rio Bravo virus group | | | |
| Orthoflavivirus bukalasaense | Bukalasa bat virus | AF013365 | BBV |
| Orthoflavivirus careyense | Carey Island virus | AF013368 | CIV |
| Orthoflavivirus dakarense | Dakar bat virus | AF013371 | DBV |
| Orthoflavivirus montanaense | Montana myotis leukoencephalitis virus | AJ299445 | MMLV |
| Orthoflavivirus phnompenhense | Batu Cave virus | AF013369 | BCV |
|  | Phnom Penh bat virus | AF013394 | PPBV |
| Orthoflavivirus bravoense | Rio Bravo virus | AF144692 | RBV |

A list of sequences was obtained from the “Insect-Specific Flaviviruses: A Systematic Review of Their Discovery, Host Range, Mode of Transmission, Superinfection Exclusion Potential and Genomic Organization” paper, as detailed in Table 7 (Blitvich and Firth, 2015). These sequences were downloaded into a FASTA file using Python (v.3.8.19) and the Entrez (v.16.2) library. Subsequent manual cleaning of the FASTA file results was performed using Notepad++. This included standardizing the sequence names, removing blank spaces, and eliminating empty lines. After cleaning, the FASTA file was uploaded to the Clustal Omega Multiple Sequence Alignment (MSA) website (Madeira *et al.*, 2024). The phylogenetic tree resulting from the MSA was then downloaded and uploaded to iTOL for customization and visualization on the iTOL website (Letunic and Bork, 2024).

# Results

## Distribution of Dengue Virus Primers by Year and Serotype

A total of 1,623 Dengue virus primers were identified from 158 publications. These primers were analyzed to determine their distribution over the years and the specific Dengue virus serotypes they target, namely DENV1, DENV2, DENV3, and DENV4. The objective of this analysis was to evaluate the temporal trends and frequency of primers for each serotype.

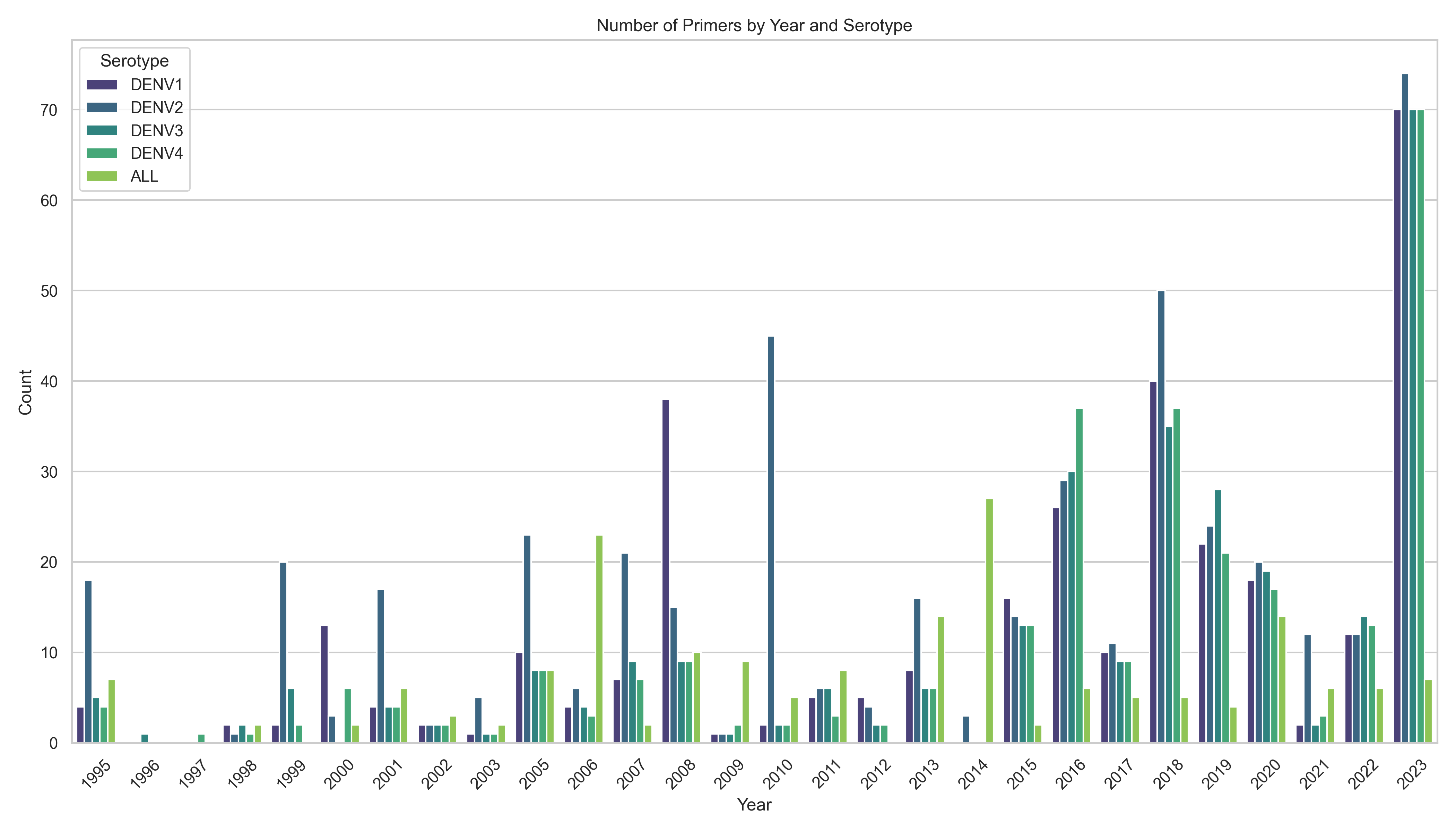


Figure 6: Literature Mining – Dengue Virus PCR Primers by Year and Serotype

Figure 6 illustrates the distribution of primers over the years. In the early years (1995-2000), the number of primers was relatively low with sporadic increases. From 2005 onwards, there was a noticeable rise in the number of primers, peaking in 2010, 2018, and 2023. Primers targeting DENV1 showed multiple peaks, particularly in 1999, 2008, 2010, 2018, and a significant increase in 2023. Similarly, primers for DENV2 exhibited notable peaks in 2005, 2008, 2010, 2017, and 2023. DENV3 primers displayed fluctuations with significant peaks in 2000, 2003, 2005, 2010, 2018, and a major rise in 2023. Primers targeting DENV4 had fewer peaks, with notable increases in 2005, 2010, 2014, 2018, and 2023. The 'ALL' category, which includes primers targeting multiple serotypes, showed consistent increases over the years, with significant peaks in 2010, 2018, and 2023.

Statistical Analysis

To provide a statistical backing for these observations, several Ordinary Least Squares (OLS) regression analyses were performed to evaluate trends over the years:

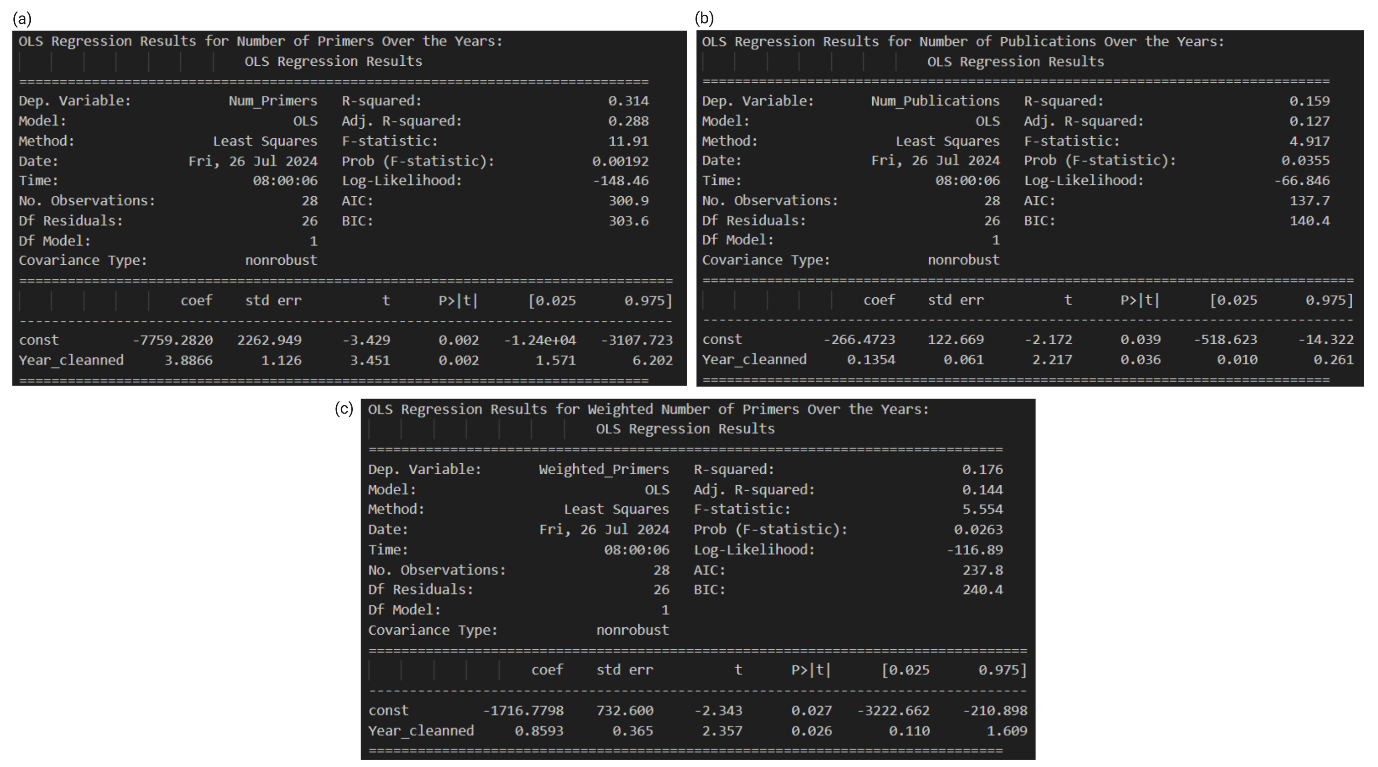
****

Figure 7: OLS Regression for (a) Number of Primers Over the Years, (b) Number of Publications Over the Years, and (c) Weighted Number of Primers Over the Years

1. Figure 7a shows Number of Primers Over the Years: The analysis resulted in an R-squared value of 0.314 and an adjusted R-squared value of 0.288. The F-statistic was 11.91 with a p-value of 0.00192, indicating a statistically significant trend. The slope of the Year coefficient was 3.887, suggesting an average annual increase of approximately 3.887 primers.
2. Figure 7b shows Number of Publications Over the Years: This analysis showed an R-squared value of 0.159 and an adjusted R-squared value of 0.127. The F-statistic was 4.917 with a p-value of 0.0355, also indicating statistical significance. The slope of the Year coefficient was 0.135, suggesting an average annual increase of approximately 0.135 publications.
3. Figure 7c Weighted Number of Primers Over the Years: Considering the number of publications, this analysis showed an R-squared value of 0.176 and an adjusted R-squared value of 0.144. The F-statistic was 5.554 with a p-value of 0.0263, indicating a statistically significant result. The slope of the Year coefficient was 0.859, suggesting an average annual increase of approximately 0.859 weighted primers.

Correlation Between Primers and Publications

To further illustrate the statistical analysis results, Figure 8 illustrates the number of primers and publications over the years. This figure provides a clear visual representation of the trends, highlighting the relationship between primer development and the number of related publications:

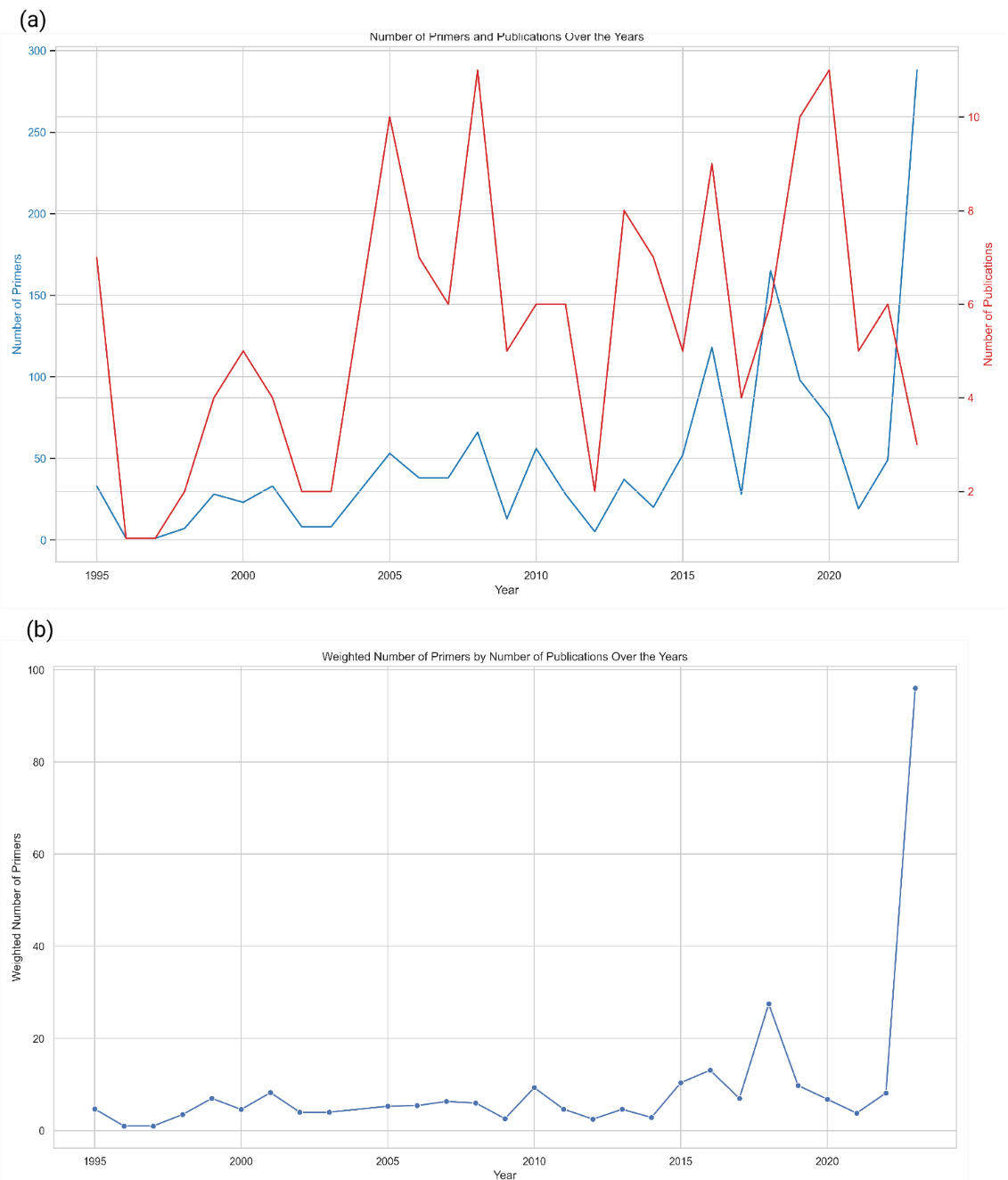


Figure 8: (a) Number of Primers and Publications Over the Years, and (b) Weighted Number of Primers by Number of Publications Over the Years

The figure indicates that both the number of primers and the number of publications have generally increased over time, with noticeable peaks in publication numbers often corresponding to peaks in primer development. Figure 8a shows the increase in the number of primers and publications, while Figure 8b shows the weighted number of primers by the number of publications over the years. This normalization provides a clearer view of the research focus intensity.

### Region Distribution of Conventional PCR Primers

Figure 9 illustrates the region distribution heatmap of conventional PCR primers. The heatmap is divided into two panels: the top panel represents individual primers, and the bottom panel shows primer combinations. This visualization helps identify the frequency and distribution of primers targeting specific regions of the Dengue virus genome, such as NS1, NS3, and NS5.

Individual primers as seen in Figure 9a are designed to amplify a specific segment of the DNA or RNA sequence. The individual primers used in this study were extracted using a systematic process that involved filtering and validating each primer against certain conditions to ensure their efficacy and specificity. This includes checking primer length, GC content, and homopolymer runs.

Primer combinations as seen in Figure 9b involve pairs of forward and reverse primers that work together to amplify a specific RNA segment. The combination of primers is selected to ensure that they work effectively together, with considerations for the melting temperatures (Tm), GC content, and the amplicon length (the length of the DNA or RNA segment they amplify). The combined primers were identified by pairing forward and reverse primers that matched specific criteria such as Tm differences, GC content differences, and the target regions they amplify.

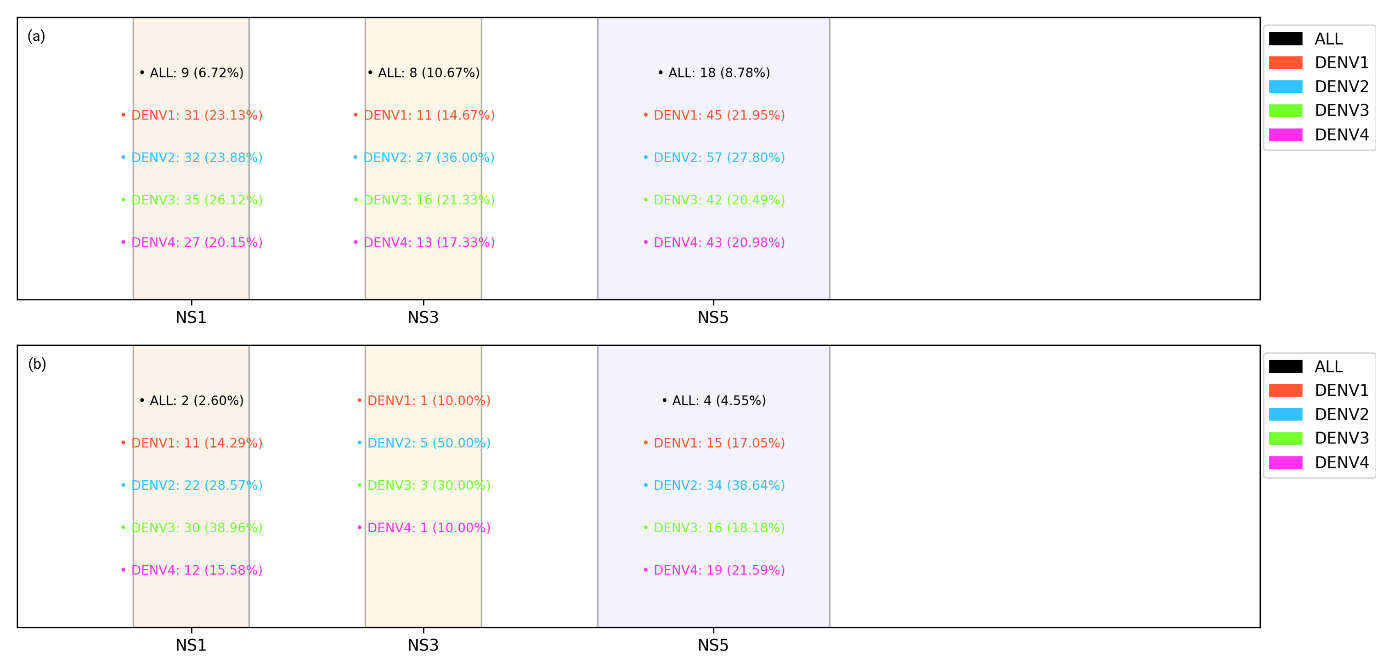


Figure 9: Region Distribution of Conventional PCR Primers for (a) Individual Primers, and (b) Primer Combination.

Figure 9a shows the distribution of individual primers across various regions is shown. The NS5 region has the highest concentration of individual primers, particularly for DENV2 and DENV1. The NS1 and NS3 regions also show significant primer presence, with DENV3 and DENV4 primers being more evenly distributed across these regions.

Figure 9b demonstrates the distribution of primer combinations targeting these regions. Similar to individual primers, the primer combinations also favor the NS5 region, with a distribution pattern comparable to that observed in the Figure 9a.

#### PCA Analysis of Conventional PCR Primers

The principal component analysis (PCA) was conducted to examine the distribution and clustering of conventional PCR primers based on their sequences and target regions. Figure 10 displays two scatter plots, representing the PCA results for individual primers (panel a) and primer combinations (panel b) respectively.

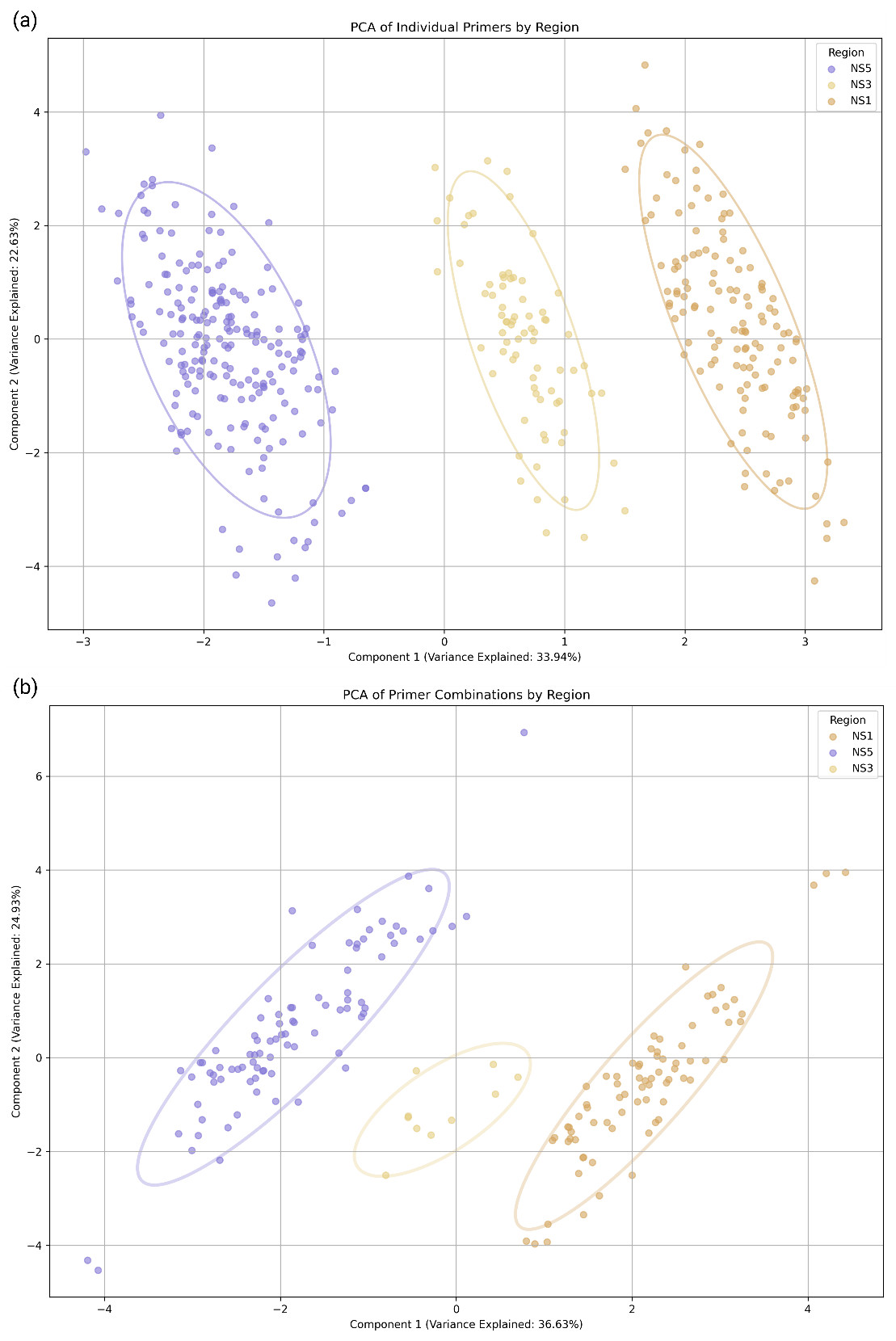


Figure 10: PCA Analysis of Conventional PCR Primers for (a) Individual Primer, and (b) Primer Combination

Figure 10a illustrates the PCA of individual primers by serotype. Each point in the scatter plot represents an individual primer, color-coded according to the Regions (NS1, NS3, NS5). The plot reveals three primary clusters corresponding to the NS1, NS3, and NS5 regions of the Dengue virus genome. The ellipses around each cluster indicate the 95% confidence intervals.

The distribution of points within these clusters suggests a significant overlap among the primers targeting different serotypes. This overlap indicates that the sequences of individual primers are not highly distinct among the serotypes, leading to their spread across the PCA components. The principal component 1 (PC1) explains 33.94% of the variance, while principal component 2 (PC2) accounts for 22.63% of the variance. The clustering patterns highlight the concentration of individual primers in the NS5 region, followed by NS3 and NS1 regions, reflecting their higher frequency of targeting these genomic regions.

Figure 10b depicts the PCA of primer combinations by serotype. Similar to the individual primers, each point represents a primer combination, color-coded by their regions. The PCA plot for primer combinations shows distinct clusters for NS1, NS3, and NS5 regions, with a noticeable spread in the NS5 region. The principal components explain 36.63% and 24.93% of the variance, respectively.

The clustering of primer combinations indicates a pattern where primers from different serotypes combine to target specific regions effectively. The spread of points across the PCA components also implies that there is a diversity in primer combinations, with some combinations spanning multiple serotypes, enhancing their versatility for diagnostic and research purposes.

Table 7: Contribution of Variables for Individual Primers

|  |  |  |
| --- | --- | --- |
| **Individual Primers** | | |
|  | **PC1** | **PC2** |
| **Start** | 0.958 | 0.021 |
| **End** | 0.959 | 0.021 |
| **Amplicon\_Length** | 0.043 | 0.694 |
| **GC\_Content** | 0.002 | 0.045 |
| **Tm\_min** | 0.025 | 0.761 |
| **Tm\_max** | 0.044 | 0.920 |
|  | | |
| **Primer Combination** | | |
| **Amplicon\_Length** | 0.132 | 0.043 |
| **End** | 0.863 | 0.117 |
| **GC\_Content\_Forward** | 0.040 | 0.106 |
| **GC\_Content\_Reverse** | 0.041 | 0.018 |
| **Start** | 0.855 | 0.111 |
| **Tm\_max\_Forward** | 0.282 | 0.536 |
| **Tm\_max\_Reverse** | 0.279 | 0.499 |
| **Tm\_min\_Forward** | 0.151 | 0.706 |
| **Tm\_min\_Reverse** | 0.163 | 0.611 |

For individual primers, as shown in Table 7, key contributors to PC1 include the start and end positions of the primers. For PC2, the amplicon length, GC content, and melting temperatures (Tm\_min and Tm\_max) are notable contributors.

In the case of primer combinations, as seen in Table 7, the start and end positions are major contributors to the principal components. The melting temperatures of both primers also play a crucial role in defining the PCA dimensions.

These findings indicate that the physical and chemical properties of the primers, such as their genomic positions, GC content, and melting temperatures, greatly influence their distribution and clustering in the PCA plots.

### Dengue Virus Nomenclature

The use of various naming conventions across years and publications necessitated the standardization of nomenclature for consistent reference. An article discussing new naming conventions provided the foundation for this effort (Cuypers *et al.*, 2018; Hill *et al.*, 2024). Consequently, a standardized nomenclature table was established to align different naming systems.

Table 8: NCBI Taxonomy, New Lineage, and Genome Detective Nomenclature

|  |  |  |
| --- | --- | --- |
| **NCBI Taxonomy Entry** | **New Lineage Nomenclature** | **Genome Detective Genotype** |
| Dengue virus 1 Brazil/97-11/1997 | DENV-1 genotype V | DENV1\_V |
| Dengue virus 1 Jamaica/CV1636/1977 | DENV-1 genotype II | DENV1\_II |
| Dengue virus 1 Nauru/West Pac/1974 | DENV-1 genotype I | DENV1\_I |
| Dengue virus 1 Singapore/S275/1990 | DENV-1 genotype IV | DENV1\_IV |
| Dengue virus 1 Thailand/AHF 82-80/1980 | DENV-1 genotype III | DENV1\_III |
| Dengue virus 2 16681-PDK53 | DENV-2 genotype Asian I | DENV2\_I |
| Dengue virus 2 China/D2-04 | DENV-2 genotype Asian II | DENV2\_II |
| Dengue virus 2 Jamaica/1409/1983 | DENV-2 genotype American | DENV2\_III |
| Dengue virus 2 Malaysia M2 | DENV-2 genotype Asian I | DENV2\_I |
| Dengue virus 2 Malaysia M3 | DENV-2 genotype Asian II | DENV2\_II |
| Dengue virus 2 Peru/IQT2913/1996 | DENV-2 genotype Cosmopolitan | DENV2\_IV |
| Dengue virus 2 Puerto Rico/PR159-S1/1969 | DENV-2 genotype American | DENV2\_III |
| Dengue virus 2 Thailand/0168/1979 | DENV-2 genotype Asian I | DENV2\_I |
| Dengue virus 2 Thailand/16681/84 | DENV-2 genotype Asian II | DENV2\_II |
| Dengue virus 2 Thailand/NGS-C/1944 | DENV-2 genotype Cosmopolitan | DENV2\_IV |
| Dengue virus 2 Thailand/PUO-218/1980 | DENV-2 genotype Asian I | DENV2\_I |
| Dengue virus 2 Thailand/TH-36/1958 | DENV-2 genotype Asian II | DENV2\_II |
| Dengue virus 2 Tonga/EKB194/1974 | DENV-2 genotype Cosmopolitan | DENV2\_IV |
| Dengue virus 3 China/80-2/1980 | DENV-3 genotype III | DENV3\_III |
| Dengue virus 3 Martinique/1243/1999 | DENV-3 genotype IV | DENV3\_IV |
| Dengue virus 3 Philippines/H87/1956 | DENV-3 genotype I | DENV3\_I |
| Dengue virus 3 Singapore/8120/1995 | DENV-3 genotype II | DENV3\_II |
| Dengue virus 3 Sri Lanka/1266/2000 | DENV-3 genotype V | DENV3\_V |
| Dengue virus 4 Dominica/814669/1981 | DENV-4 genotype II | DENV4\_II |
| Dengue virus 4 Philippines/H241/1956 | DENV-4 genotype I | DENV4\_I |
| Dengue virus 4 Singapore/8976/1995 | DENV-4 genotype II | DENV4\_II |
| Dengue virus 4 Thailand/0348/1991 | DENV-4 genotype I | DENV4\_I |
| Dengue virus 4 Thailand/0476/1997 | DENV-4 genotype I | DENV4\_I |

## Development of an in-house dengue virus genotyping pipeline

### Clustering Dengue Virus Sequence

To develop an in-house Dengue virus genotyping pipeline, the dataset was first clustered to shorten the BLAST process and overall genotyping time. The clustering of the dataset is portrayed in Figure 11. Figure 11a shows the genotype distribution of the whole dataset, while Figure 11b illustrates the clustered dataset. The clustering aimed to streamline the computational process by reducing the dataset size without compromising the representativeness of the genotypes.

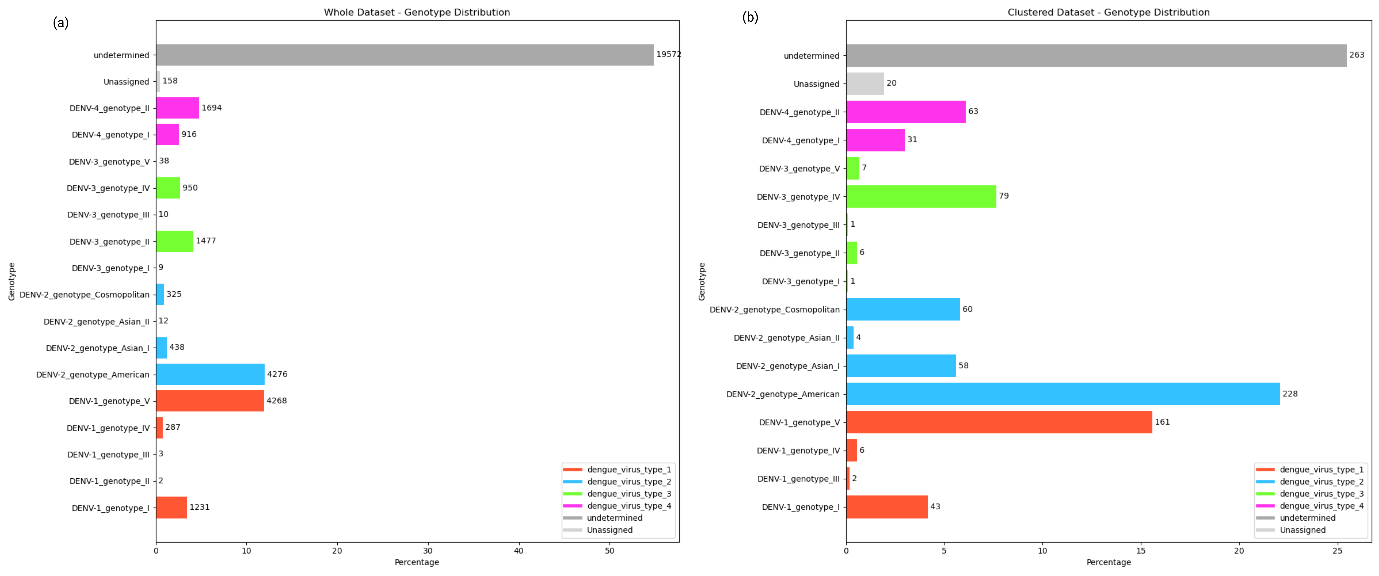


Figure 11: Dengue Virus Sequence Genotyping between Unclustered vs Clustered

The genotyping pipeline utilized the VADR tool and the new lineage naming nomenclature, followed by BLAST to determine both the serotype and genotype. The performance of the pipeline was evaluated using three different cutoff values for species, family, and phylum: Figure 12 (a and b) 75-70-65, Figure 12 (c and d) 85-80-75, and Figure 12 (e and f) 95-90-85.

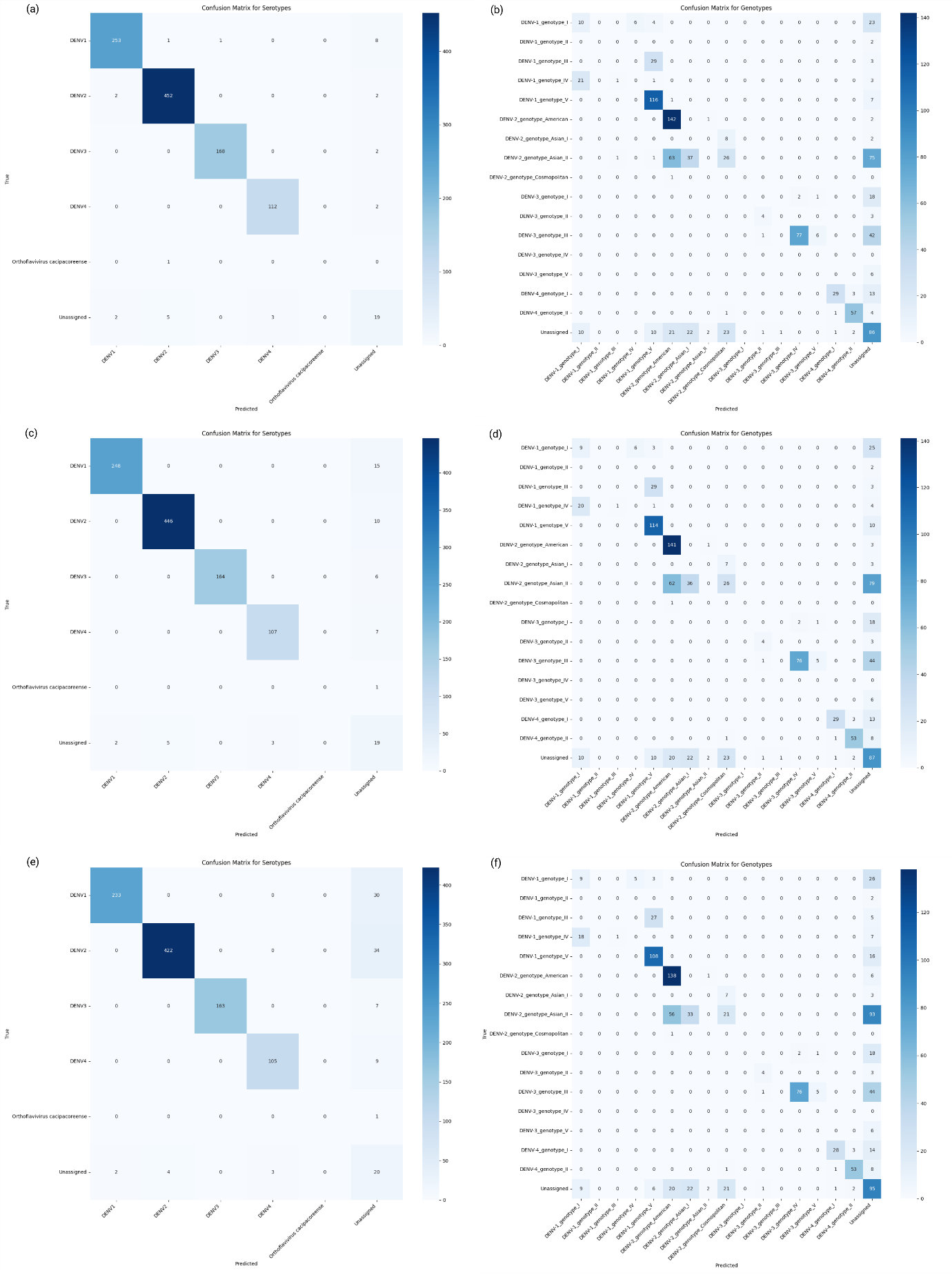


Figure 12: Whole vs Clustered Dengue Virus Sequence Dataset – Genotype Distribution

### Evaluation of In-House Genotyping Pipeline Performance

The results of the in-house genotyping pipeline showed that while the serotype determination was accurate, the genotype determination performance was suboptimal. Table 9summarizes the metrics for each cutoff value.

Table 9: Top Cutoff (75-70-65)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | **Recall** | **F1 Score** | **Specificity** | **NPV** | **MCC** | **Balanced Accuracy** |
| **Cutoff (Species: 65, Family: 60, Phylum: 55)** | | | | | | | |
| **Serotype** | 0.976 | 0.978 | 0.977 | 0.767 | 0.774 | 0.968 | 0.767 |
| **Genotype** | 0.338 | 0.432 | 0.373 | 0.281 | 0.260 | 0.382 | 0.299 |
| **Cutoff (Species: 75, Family: 70, Phylum: 65)** | | | | | | | |
| **Serotype** | 0.973 | 0.972 | 0.972 | 0.763 | 0.752 | 0.960 | 0.763 |
| **Genotype** | 0.337 | 0.430 | 0.371 | 0.279 | 0.259 | 0.380 | 0.297 |
| **Cutoff (Species: 85, Family: 80, Phylum: 75)** | | | | | | | |
| **Serotype** | 0.970 | 0.953 | 0.960 | 0.747 | 0.714 | 0.934 | 0.747 |
| **Genotype** | 0.335 | 0.423 | 0.367 | 0.273 | 0.258 | 0.371 | 0.290 |
| **Cutoff (Species: 95, Family: 90, Phylum: 85)** | | | | | | | |
| **Serotype** | 0.967 | 0.913 | 0.935 | 0.730 | 0.692 | 0.883 | 0.730 |
| **Genotype** | 0.339 | 0.421 | 0.367 | 0.270 | 0.261 | 0.368 | 0.287 |

The results indicate that the in-house pipeline performs well in determining the serotype, with high precision, recall, and F1 scores across all cutoff values. Nonetheless, the genotype determination metrics are notably lower, reflecting poor performance in accurately identifying the genotypes. The genotype precision, recall, and F1 scores are significantly low, with balanced accuracy and MCC values also indicating subpar performance.

Given the suboptimal performance of the in-house genotyping pipeline in determining genotypes, it was decided to continue using the Genome Detective Dengue Virus Typing Tool (Vilsker *et al.*, 2019) for reliable genotyping results. The Genome Detective tool provides consistent and accurate genotyping, making it a more suitable choice for this study.

## In-silico comparison of PCR assays for dengue virus genotyping

Table 10 shows the performance metrics for the top performing primer pair in terms of F1 Score, Precision, Recall, and Support (number of observations), at genotyping Dengue virus. For a comprehensive list of Missed Classification Primers, please refer to Table 13 in the appendix section.

Table 10: Top 25 Primer Pair for Dengue Virus Overall Metrics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Primer\_Pair** | **F1\_Score** | **Precision** | **Recall** | **Support** |
| 1162\_F\_D4\_R | 1.000 | 1.000 | 1.000 | 43.000 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1569.000 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 | 3224.000 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 | 4648.000 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 340.000 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 458.000 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 306.000 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 406.000 |
| DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 | 20.000 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 1.000 | 1.000 | 1.000 | 652.000 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 1.000 | 1.000 | 1.000 | 358.000 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1008.000 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 1.000 | 1.000 | 1.000 | 1100.000 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 | 1520.000 |
| NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 | 78.000 |
| NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 | 72.000 |
| NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 | 76.000 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1928.000 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 3795.000 |
| D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 | 27.000 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 2145.000 |
| D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 | 90.000 |
| D3F1187\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 | 62.000 |
| D3SEQ2\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 | 109.000 |
| DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 | 85.000 |

In evaluating the F1 scores for each primer pair across all Dengue virus samples, it was observed that several primer pairs, including 1162\_F\_D4\_R, achieved an F1 score of 1.0, indicating perfect precision and recall. This suggests that these primers were highly effective in correctly identifying the Dengue virus in all tested samples.

### Serotype Confusion Matrix

Analysis of the classification performance for different serotypes and genotypes of dengue virus using a set of conventional primer pairs was performed. Figure 13 shows serotype confusion matrix, which shows how many primer pairs for each serotypes.

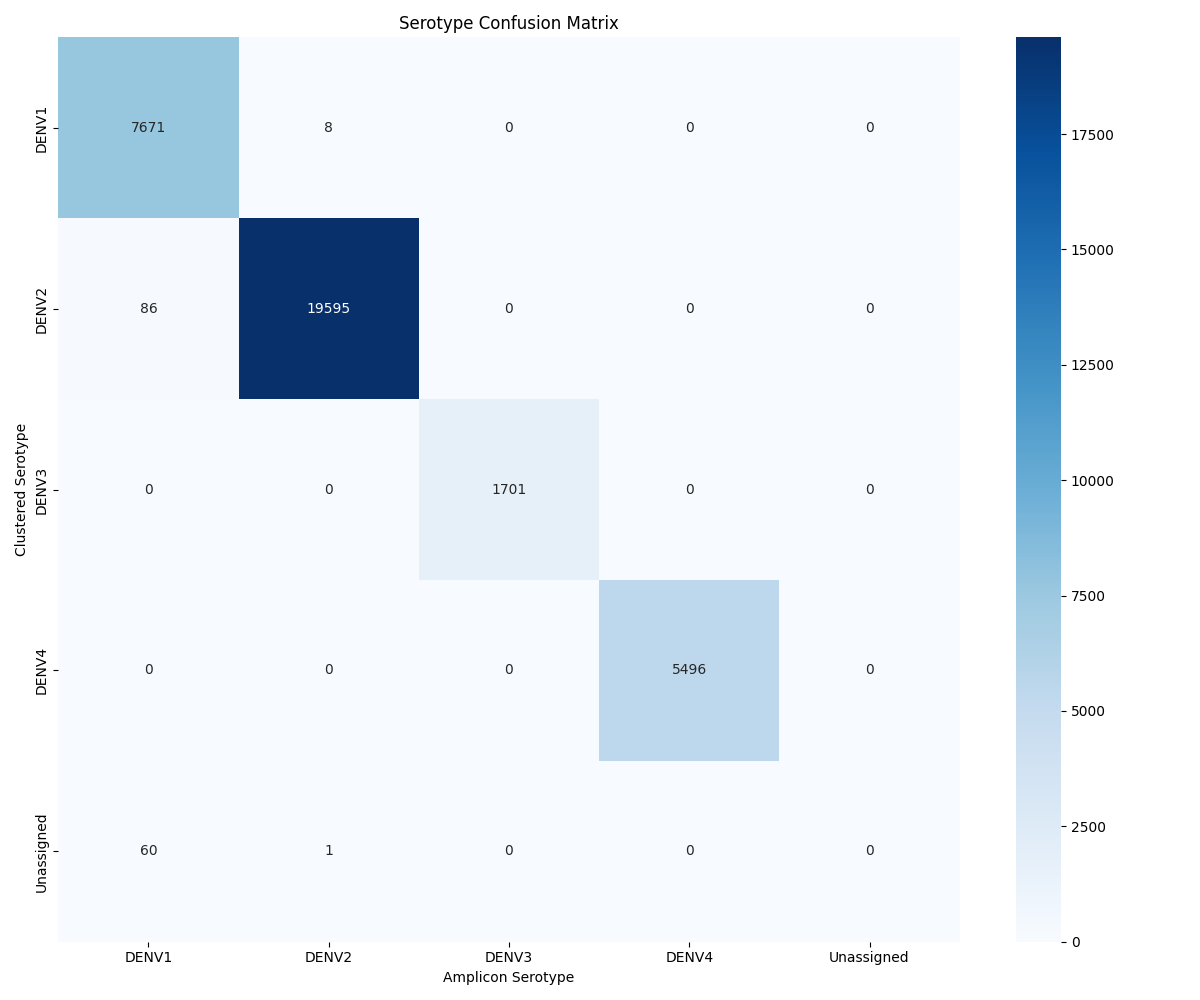


Figure 13: Serotype Confusion Matrix

Figure x shows the serotype confusion matrix, revealing high accuracy in the classification of Dengue virus serotypes. DENV1 was accurately classified in 7671 out of 7735 instances, while DENV2 was correctly identified in 19595 out of 19681 cases.

### Genotype Confusion Matrix

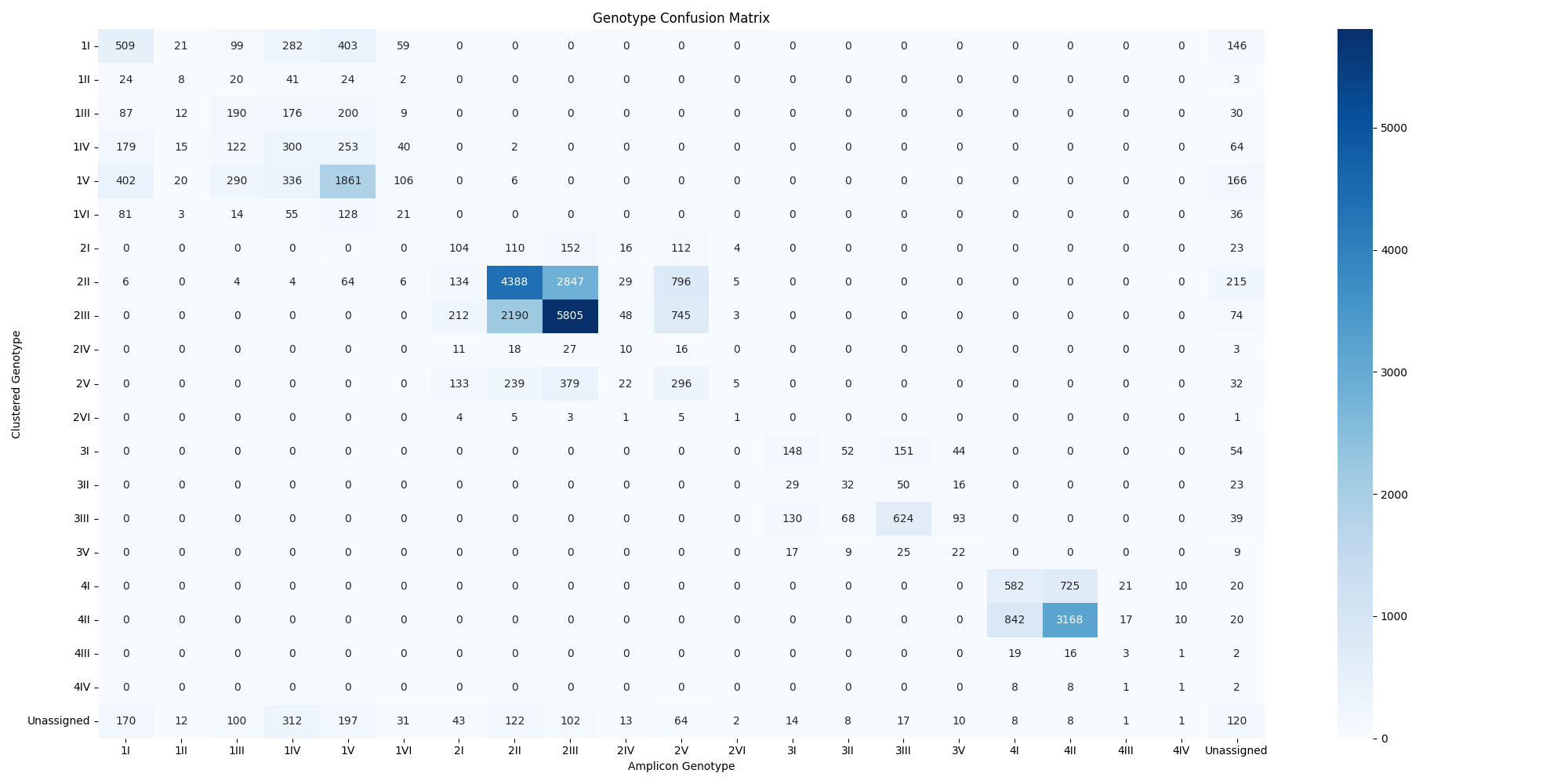


Figure 14: Genotype Confusion Matrix

Nevertheless, the genotype confusion matrix indicated more variability in classification accuracy across different genotypes as seen in Figure 14. For this analysis, we used amplicons generated from all tested primer pairs to ensure comprehensive coverage. Genotype 1I showed some misclassifications into other genotypes such as 1IV and 1III. More notably, genotypes 2II and 2III had significant cross-misclassifications, as seen in Table 14, highlighting challenges in distinguishing between these closely related genotypes. The input data for these analyses included results from Genome Detective, with clustered dengue sequences genotyped using this tool.

### Missed and Mismatched Classification

The primer pair DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R had the highest number of missed classifications as seen in Figure 15, particularly for genotypes 2II and 2III. This pattern was consistent across several primers, indicating a recurrent issue with distinguishing between closely related genotypes. The data also showed frequent missed classifications for these genotypes, suggesting that improvements are needed in primer design or additional markers to enhance specificity. For a comprehensive list of Missed Classification Primers, please refer to Table 14 in the appendix section.

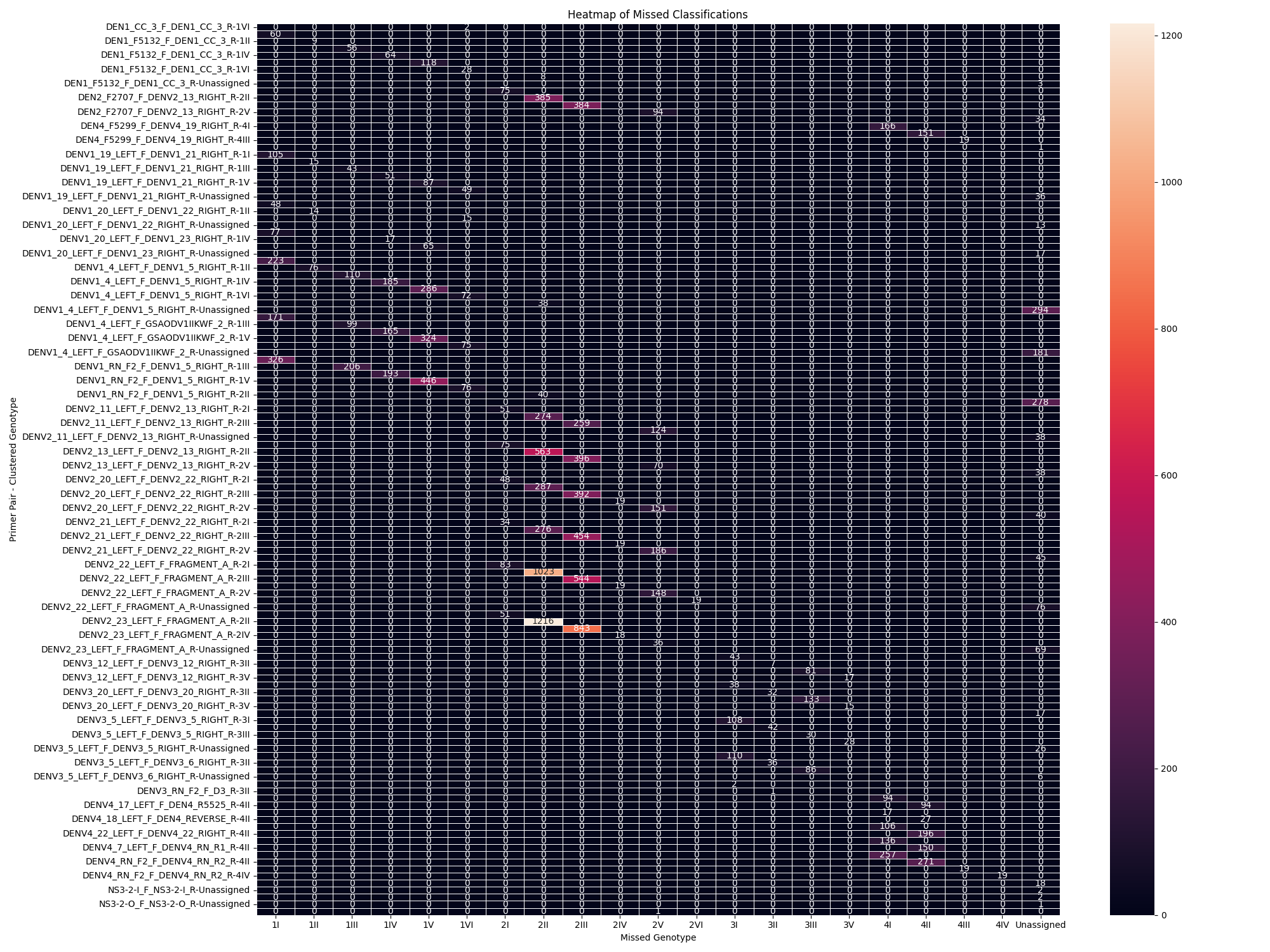


Figure 15: Missed Classification Heatmap

Mismatched classifications, where the identified genotype significantly differed from the clustered genotype, were also analyzed. Primer pairs such as DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R and DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R had high numbers of mismatched observations between genotypes 2II and 2III as seen in Figure 16. This further underscores the difficulties these primers face in differentiating between certain genotypes, leading to higher error rates in classification. For a comprehensive list of Missed Classification Primers, please refer to Table 15 in the appendix section.



Figure 16: Mismatch Classification Heatmap

For missed genotypes, the Chi-square test yielded a Chi-square value of 64.75 and a p-value of 1.279e-06, indicating a statistically significant difference in the misclassification of genotypes. The analysis included extracting Pearson residuals to identify which factors contributed most to the misclassifications. It was observed that genotypes 2II and 2III had the highest residuals, suggesting they were the most significantly misclassified. Similarly, the Chi-square test for mismatched genotypes produced a Chi-square value of 273.22 and a p-value of 2.291e-46, confirming a statistically significant difference in the mismatched classification of genotypes. The Pearson residuals highlighted that primer pairs such as DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R and DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R performed significantly worse, particularly in distinguishing between closely related genotypes 2II and 2III. These findings emphasize the need for improved primer design and validation to enhance the accuracy of genotype identification.

### F1 Score for Each Primer Pair for Each Serotype and Genotype

The performance metrics for each primer pair for different serotypes of Dengue virus illustrate a high level of precision and accuracy in the detection and classification of the virus, as seen in Table 11. For DENV1, several primer pairs such as DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R, DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R, DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R, and DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R achieved perfect F1 scores, precision, and recall rates of 1.00. This indicates that these primer pairs were highly effective in correctly identifying DENV1 without any false positives or negatives. Nonetheless, the primer pair DEN1\_F5132\_F\_DEN1\_CC\_3\_R for DENV1 also demonstrated a high F1 score of 1.00 but had a slightly lower recall for DENV2 with a recall rate of 0.11, suggesting some challenges in accurately identifying DENV2.

For DENV2, multiple primer pairs also demonstrated perfect performance metrics. Primer pairs such as D2F1200\_F\_P2\_R, DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R, and DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R all achieved F1 scores of 1.00, indicating highly effective performance in detecting DENV2 across the tested samples. This consistency in performance highlights the robustness of these primers in identifying DENV2., it is noteworthy that certain primers intended for DENV1, such as DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R and DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R, showed zero performance metrics (F1, precision, and recall all at 0.00) for DENV2. This underscores the specificity challenges when using primers designed for one serotype on another, but it also suggests that these primers could potentially be used as serotype-specific primers, offering high specificity for DENV1.

For DENV3, the primer pairs D3F1187\_F\_D3R1758\_R, D3SEQ2\_F\_D3R1758\_R, DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R, and DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R all achieved perfect scores. This suggests these primers are highly reliable for detecting DENV3, maintaining high precision and recall across all tested instances. Similarly, for DENV4, primer pairs such as 1162\_F\_D4\_R, DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R, and DENV-4\_F\_D4\_R demonstrated perfect performance metrics, indicating their effectiveness in accurately identifying DENV4.

Table 11: F1 Score for Each Primer Pair for Each Serotype

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Serotype** | **Primer\_Pair** | **F1\_Score** | **Precision** | **Recall** |
| DENV1 | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV1 | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1.00 | 1.00 | 0.99 |
| DENV2 | D2F1200\_F\_P2\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | NS3-2-O\_F\_NS3-2-O\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | NS3-2-O\_F\_NS3-2-I\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | NS3-2-I\_F\_NS3-2-I\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | D2S\_F\_D2\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | RSS3\_F\_RSS4\_R | 1.00 | 1.00 | 1.00 |
| DENV2 | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 0.20 | 1.00 | 0.11 |
| DENV2 | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 0.00 | 0.00 | 0.00 |
| DENV2 | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 0.00 | 0.00 | 0.00 |
| DENV3 | D3F1187\_F\_D3R1758\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | D3SEQ2\_F\_D3R1758\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV3 | DENV3\_RN\_F2\_F\_D3\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | 1162\_F\_D4\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV-4\_F\_D4\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 1.00 | 1.00 | 1.00 |
| DENV4 | DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.00 | 1.00 | 1.00 |

Table 12 presents the performance metrics of various primer pairs for different Dengue virus serotypes. Each genotype was tested with multiple primer pairs, and the results consistently demonstrated high performance.

For the 1I genotype, primer pairs such as DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R, DEN1\_F5132\_F\_DEN1\_CC\_3\_R, DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R, and others achieved perfect F1 scores, precision, and recall of 1.00. This indicates that these primers were extremely effective in accurately identifying the 1I genotype without any false positives or negatives.

The same trend was observed for the 1II and 1III genotypes, where primer pairs such as DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R, DEN1\_F5132\_F\_DEN1\_CC\_3\_R, and others also achieved perfect scores across all performance metrics. This suggests a high level of reliability and accuracy for these primer pairs in detecting the respective genotypes.

For the 1IV genotype, while most primer pairs maintained high performance, a few showed slight deviations. The primer pair DEN1\_F5132\_F\_DEN1\_CC\_3\_R had an F1 score of 0.988 with a recall of 0.977, indicating a minor decrease in performance. Similarly, the primer pair DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R had an F1 score of 0.907, precision of 0.879, and recall of 0.937, showing a slight drop in precision. Despite these minor variations, the overall performance remained high.

In the case of the 1V genotype, most primer pairs achieved perfect scores again, but DEN1\_F5132\_F\_DEN1\_CC\_3\_R had a slight reduction in recall (0.989) leading to an F1 score of 0.994. This indicates that while the primer pairs were highly effective, there were minor discrepancies in recall for some primers.

The 1VI genotype followed the same pattern of high performance, with all primer pairs achieving perfect scores, demonstrating their effectiveness in correctly identifying this genotype.

For the 2I genotype, primer pairs such as DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R, DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R, and others maintained perfect performance metrics. This trend was consistent across the 2II, 2III, 2IV, and 2V genotypes, indicating a high degree of accuracy and reliability for the primer pairs used for these genotypes. However, for the 2II genotype, two primer pairs DEN1\_F5132\_F\_DEN1\_CC\_3\_R and DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R showed significantly lower performance metrics, with DEN1\_F5132\_F\_DEN1\_CC\_3\_R having an F1 score of 0.200, precision of 1.000, and recall of 0.111, and DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R showing zero performance metrics. This indicates that these primers were not effective for the 2II genotype.

For the 3I genotype, primer pairs such as D3F1187\_F\_D3R1758\_R, D3SEQ2\_F\_D3R1758\_R, and others demonstrated perfect performance, similar to the trends observed in the other genotypes. This high performance was consistent across the 3II, 3III, and 3V genotypes, indicating that these primer pairs were highly reliable for identifying these genotypes accurately.

Finally, the 4I and 4II genotypes also showed perfect performance metrics for their respective primer pairs, such as 1162\_F\_D4\_R, DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R, and others. This pattern continued for the 4III and 4IV genotypes, demonstrating the effectiveness of the primer pairs used for these genotypes.

Table 12: F1 Score for Each Primer Pair for Each Genotype

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gerotype** | **Primer\_Pair** | **F1\_Score** | **Precision** | **Recall** |
| 1I | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1I | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1.000 | 1.000 | 1.000 |
| 1I | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1II | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1II | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1II | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1II | DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1II | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1III | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1III | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1III | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1III | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1III | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1.000 | 1.000 | 1.000 |
| 1III | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1IV | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1IV | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1IV | DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1IV | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 0.988 | 1.000 | 0.977 |
| 1IV | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 0.907 | 0.879 | 0.937 |
| 1IV | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 0.887 | 0.854 | 0.924 |
| 1IV | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 0.866 | 0.826 | 0.909 |
| 1V | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1V | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1V | DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1V | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1V | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1.000 | 1.000 | 1.000 |
| 1V | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1V | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 0.994 | 1.000 | 0.989 |
| 1VI | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1.000 | 1.000 | 1.000 |
| 1VI | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2I | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2I | NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2II | D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2II | NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 |
| 2II | NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2II | NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2II | D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 |
| 2II | DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2II | DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2II | DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 |
| 2II | RSS3\_F\_RSS4\_R | 1.000 | 1.000 | 1.000 |
| 2II | DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 0.200 | 1.000 | 0.111 |
| 2II | DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 0.000 | 0.000 | 0.000 |
| 2II | DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 0.000 | 0.000 | 0.000 |
| 2III | D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 |
| 2III | D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 |
| 2III | DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2III | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2III | NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2III | NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2III | NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 |
| 2III | RSS3\_F\_RSS4\_R | 1.000 | 1.000 | 1.000 |
| 2IV | D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 |
| 2IV | D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 |
| 2IV | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2IV | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2IV | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2IV | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2IV | NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2IV | NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2IV | NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 |
| 2V | D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 |
| 2V | D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 |
| 2V | DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2V | DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 2V | NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2V | NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 |
| 2V | NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 |
| 2V | RSS3\_F\_RSS4\_R | 1.000 | 1.000 | 1.000 |
| 2VI | DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 |
| 3I | D3F1187\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3I | D3SEQ2\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3I | DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3I | DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3I | DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3I | DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3I | DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 |
| 3II | DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3II | DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3II | DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3II | DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3II | DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 |
| 3III | D3F1187\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3III | D3SEQ2\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3III | DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3III | DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3III | DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3III | DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3III | DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 |
| 3V | D3F1187\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3V | D3SEQ2\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 |
| 3V | DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3V | DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3V | DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 3V | DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 |
| 4I | 1162\_F\_D4\_R | 1.000 | 1.000 | 1.000 |
| 4I | DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV-4\_F\_D4\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 1.000 | 1.000 | 1.000 |
| 4I | DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 |
| 4II | 1162\_F\_D4\_R | 1.000 | 1.000 | 1.000 |
| 4II | DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV-4\_F\_D4\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 1.000 | 1.000 | 1.000 |
| 4II | DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 |
| 4III | 1162\_F\_D4\_R | 1.000 | 1.000 | 1.000 |
| 4III | DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 1.000 | 1.000 | 1.000 |
| 4III | DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 |
| 4IV | DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 |

# Discussion

The primary objective of this study was to evaluate the efficacy of various primer pairs for Dengue virus genotyping and to develop an in-house genotyping pipeline that could be utilized in resource-limited settings. Our results demonstrated that while many published primers exist, only a subset permits accurate genotype identification. This highlights the necessity for more rigorous validation of primers before widespread adoption in research and clinical settings. Our comprehensive table of primer pairs, organized using standardized nomenclature, provides a crucial tool for researchers to accurately identify and compare results across different studies, thus improving the overall efficiency and reliability of Dengue virus research.

In the analysis of the Dengue virus primers, we identified several primer pairs that achieved perfect precision, recall, and F1 scores, indicating their high efficacy in correctly identifying the Dengue virus across different serotypes. Notably, primer pairs such as DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R and DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R were highly effective, demonstrating F1 scores of 1.0 across thousands of observations. This level of accuracy is critical for reliable virus detection and genotyping, particularly in clinical diagnostics and epidemiological studies.

The need for standardized naming nomenclature in Dengue virus research is clear. Different publications have used varying naming conventions, making it difficult to track and compare findings. This issue is worsened by frequent changes in naming systems. Standardizing nomenclature, as suggested by new conventions (Cuypers *et al.*, 2018; Hill *et al.*, 2024), can streamline research and improve communication within the scientific community. Our comprehensive table of primer pairs, organized with standardized names, helps researchers accurately identify and compare results, enhancing the efficiency and reliability of Dengue virus research.

In many low-resource settings, particularly in developing countries, financial and infrastructural constraints limit access to advanced diagnostic technologies. Conventional PCR (cPCR) is a vital tool in these environments for the identification and genotyping of viruses, including Dengue. cPCR provides a cost-effective and accessible method for the timely detection of viral pathogens, which is crucial during outbreaks to manage and mitigate the spread of the disease (Malavige *et al.*, 2023).

The need for command-line options for Dengue virus classification is becoming increasingly apparent. Current online tools, such as the Rijksinstituut voor Volksgezondheid en Milieu (RIVM) Flavivirus Genotyping tool or Genome Detective Dengue Virus Typing Tool (Vilsker *et al.*, 2019), although accurate, present significant limitations in terms of processing speed and capacity. These tools can only process a limited number of sequences per run, and the analysis process can be slow. This is particularly problematic during urgent outbreak scenarios, where rapid genotyping results are critical for timely public health responses and effective management of disease spread. The development of an in-house Dengue virus genotyping pipeline aimed to address these limitations by providing a faster and more integrated approach. While the initial results of the in-house pipeline showed poor performance in genotype determination (F1 scores were significantly lower than desired), the potential for rapid and large-scale analysis cannot be overlooked. A robust command-line tool could be optimized for local execution, significantly enhancing the speed and efficiency of genotyping analyses by allowing for batch processing of large datasets and integration into automated workflows. Such advancements would not only improve the responsiveness during outbreaks but also streamline routine diagnostic processes in research and clinical settings.

Alternatives to traditional BLAST for classification include several advanced algorithms and tools such as mmseqs2, RDP (Ribosomal Database Project), ID-TAXA, Kraken2 with custom databases, and Kaiju with custom databases (Cole *et al.*, 2014; Menzel, Ng and Krogh, 2016; Steinegger and Söding, 2017; Murali, Bhargava and Wright, 2018; Breitwieser and Salzberg, 2020). These tools offer various advantages such as higher speed, accuracy, and the ability to handle large datasets more efficiently. mmseqs2, for instance, is known for its high-speed protein sequence searching capabilities, while Kraken2 and Kaiju provide rapid classification of metagenomic sequences with customizable databases. ID-TAXA and RDP are particularly useful for phylogenetic analysis and taxonomic classification of ribosomal RNA sequences. Implementing these alternatives in the genotyping pipeline could potentially enhance the performance and reliability of Dengue virus genotyping, providing more accurate and faster results for research and clinical applications ​

Developing an in-house genotyping pipeline presents several challenges. One major issue is the variability in the format and presentation of primers in publications. Different researchers publish primers in diverse formats, including images, tables, and supplementary sections, with non-standardized column content and naming conventions. This lack of standardization complicates the extraction and analysis of primer data. Additionally, the changing taxonomy and evolving publication standards further exacerbate these challenges. Standardizing the reporting of primers and adopting consistent nomenclature can mitigate these issues, making it easier for researchers to interpret and utilize published data. This is in line with the guidelines suggested by the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) (Bustin *et al.*, 2009).

Although there are many scripts used in this study, it essentially involves a combination of built-in libraries, available libraries, and inspiration from other GitHub projects. The scripts developed in this study for Dengue virus genotyping can be easily adapted for other viruses or pathogens. The parameters within the scripts can be adjusted to accommodate different sequence characteristics, making them versatile tools for a wide range of virological and microbiological applications. This adaptability highlights the broader utility of the developed methodologies, providing a framework that can be extended beyond Dengue virus to enhance diagnostic capabilities for various infectious diseases.

Despite the encouraging outcomes, there are still several limitations that must be addressed. Although there are several promising primer results that can be used for certain genotypes/serotypes, they were only tested in "dry lab" settings using code and simulations, not in wet lab/real world scenarios, which might yield different results (Laub *et al.*, 2023). Furthermore, the current in-house genotyping tool's poor performance in genotyping might be attributed to the selected sequence. There are multiple NCBI taxonomy entries with the same nomenclature under the new lineage nomenclature column and Genome Detective column. BLAST can only accommodate one, not multiple, of the same "names," leading to the selection of representative sequences, which might have affected the detection accuracy.

Future work will focus on expanding the scope of the study to include nested and semi-nested PCR primers and amplicons. This expansion will provide a more comprehensive comparison between conventional and advanced PCR methods. Additionally, further refinement of the in-house Dengue virus genotyping pipeline is necessary to improve genotype determination accuracy. This will involve optimizing the pipeline parameters and incorporating more robust validation techniques to enhance its performance. Furthermore, exploring the integration of high-throughput sequencing technologies and machine learning algorithms could offer new avenues for improving the speed and accuracy of virus genotyping in diverse settings. This future work is not only critical for the continued advancement of our research but also for the broader field of virology and epidemiology, providing valuable tools and methodologies to combat infectious diseases globally.

In conclusion, our study provides a comprehensive analysis of the performance of various Dengue virus primers and highlights the importance of standardizing nomenclature and improving primer validation processes. The development of an in-house genotyping pipeline, despite its initial limitations, offers a promising approach for rapid and large-scale virus classification. Future advancements in this pipeline, coupled with integration of high-throughput sequencing and machine learning, can significantly enhance the accuracy and efficiency of genotyping, thereby contributing to better management and control of Dengue virus outbreaks.

# References

Achee, N.L. *et al.* (2015) ‘A Critical Assessment of Vector Control for Dengue Prevention’, *PLOS Neglected Tropical Diseases*. Edited by S.B. Halstead, 9(5), p. e0003655. Available at: https://doi.org/10.1371/journal.pntd.0003655.

Alm, E. *et al.* (2015) ‘One-step real-time RT-PCR assays for serotyping dengue virus in clinical samples’, *BMC Infectious Diseases*, 15(1), p. 493. Available at: https://doi.org/10.1186/s12879-015-1226-z.

Alphey, L. *et al.* (2016) ‘GENETIC CONTROL OF AEDES MOSQUITOES’, in *Global Health Impacts of Vector-Borne Diseases: Workshop Summary*. National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK390444/ (Accessed: 26 July 2024).

Anyamba, A. *et al.* (2012) ‘Climate teleconnections and recent patterns of human and animal disease outbreaks’, *PLoS neglected tropical diseases*, 6(1), p. e1465. Available at: https://doi.org/10.1371/journal.pntd.0001465.

Badurdeen, S. *et al.* (2013) ‘Sharing experiences: towards an evidence based model of dengue surveillance and outbreak response in Latin America and Asia’, *BMC Public Health*, 13(1), p. 607. Available at: https://doi.org/10.1186/1471-2458-13-607.

Beaucourt, S. and Vignuzzi, M. (2014) ‘Ribavirin: a drug active against many viruses with multiple effects on virus replication and propagation. Molecular basis of ribavirin resistance’, *Current Opinion in Virology*, 8, pp. 10–15. Available at: https://doi.org/10.1016/j.coviro.2014.04.011.

Bernard, M. *et al.* (2021) ‘FROGS: a powerful tool to analyse the diversity of fungi with special management of internal transcribed spacers’, *Briefings in Bioinformatics*, 22(6), p. bbab318. Available at: https://doi.org/10.1093/bib/bbab318.

Bhatt, S. *et al.* (2013) ‘The global distribution and burden of dengue’, *Nature*, 496(7446), pp. 504–507. Available at: https://doi.org/10.1038/nature12060.

Biswal, S. *et al.* (2019) ‘Efficacy of a Tetravalent Dengue Vaccine in Healthy Children and Adolescents’, *The New England Journal of Medicine*, 381(21), pp. 2009–2019. Available at: https://doi.org/10.1056/NEJMoa1903869.

Blitvich, B.J. and Firth, A.E. (2015) ‘Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization’, *Viruses*, 7(4), pp. 1927–1959. Available at: https://doi.org/10.3390/v7041927.

Brady, O.J. *et al.* (2012) ‘Refining the global spatial limits of dengue virus transmission by evidence-based consensus’, *PLoS neglected tropical diseases*, 6(8), p. e1760. Available at: https://doi.org/10.1371/journal.pntd.0001760.

Breitwieser, F.P. and Salzberg, S.L. (2020) ‘Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification’, *Bioinformatics*, 36(4), pp. 1303–1304. Available at: https://doi.org/10.1093/bioinformatics/btz715.

Bustin, S.A. *et al.* (2009) ‘The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments’, *Clinical Chemistry*, 55(4), pp. 611–622. Available at: https://doi.org/10.1373/clinchem.2008.112797.

Caminade, C. *et al.* (2014) ‘Impact of climate change on global malaria distribution’, *Proceedings of the National Academy of Sciences*, 111(9), pp. 3286–3291. Available at: https://doi.org/10.1073/pnas.1302089111.

Caminade, C., McIntyre, K.M. and Jones, A.E. (2019) ‘Impact of recent and future climate change on vector-borne diseases’, *Annals of the New York Academy of Sciences*, 1436(1), pp. 157–173. Available at: https://doi.org/10.1111/nyas.13950.

CDC (2021) *Dengue Prevention | CDC*, *Centers for Disease Control and Prevention*. Available at: https://www.cdc.gov/dengue/prevention/index.html (Accessed: 19 March 2024).

*Chikungunya, dengue et zika - Données de la surveillance renforcée en France hexagonale 2024* (2024). Available at: https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-a-transmission-vectorielle/chikungunya/articles/donnees-en-france-metropolitaine/chikungunya-dengue-et-zika-donnees-de-la-surveillance-renforcee-en-france-hexagonale-2024 (Accessed: 6 July 2024).

Cole, J.R. *et al.* (2014) ‘Ribosomal Database Project: data and tools for high throughput rRNA analysis’, *Nucleic Acids Research*, 42(D1), pp. D633–D642. Available at: https://doi.org/10.1093/nar/gkt1244.

Cuypers, L. *et al.* (2018) ‘Time to Harmonize Dengue Nomenclature and Classification’, *Viruses*, 10(10), p. 569. Available at: https://doi.org/10.3390/v10100569.

*Dengue- Global situation* (2023). Available at: https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON498 (Accessed: 25 May 2024).

*Dengue worldwide overview* (2024). Available at: https://www.ecdc.europa.eu/en/dengue-monthly (Accessed: 19 March 2024).

Domingues, R.B. *et al.* (2008) ‘Involvement of the central nervous system in patients with dengue virus infection’, *Journal of the Neurological Sciences*, 267(1), pp. 36–40. Available at: https://doi.org/10.1016/j.jns.2007.09.040.

Eid, J. *et al.* (2009) ‘Real-Time DNA Sequencing from Single Polymerase Molecules’, *Science*, 323(5910), pp. 133–138. Available at: https://doi.org/10.1126/science.1162986.

Eisen, L. *et al.* (2009) ‘Proactive Vector Control Strategies and Improved Monitoring and Evaluation Practices for Dengue Prevention’, *Journal of Medical Entomology*, 46(6), pp. 1245–1255. Available at: https://doi.org/10.1603/033.046.0601.

Ellis, E.M. *et al.* (2016) ‘Incidence and Risk Factors for Developing Dengue-Associated Hemophagocytic Lymphohistiocytosis in Puerto Rico, 2008 - 2013’, *PLoS neglected tropical diseases*, 10(8), p. e0004939. Available at: https://doi.org/10.1371/journal.pntd.0004939.

EpiCentro (2024) *Arbovirosi • bollettini periodici arbovirosi*. Available at: https://www.epicentro.iss.it/arbovirosi/dashboard (Accessed: 6 July 2024).

Escudié, F. *et al.* (2018) ‘FROGS: Find, Rapidly, OTUs with Galaxy Solution’, *Bioinformatics*, 34(8), pp. 1287–1294. Available at: https://doi.org/10.1093/bioinformatics/btx791.

Ewing, B. and Green, P. (1998) ‘Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities’, *Genome Research*, 8(3), pp. 186–194. Available at: https://doi.org/10.1101/gr.8.3.186.

Fonseca, V. *et al.* (2019) ‘A computational method for the identification of Dengue, Zika and Chikungunya virus species and genotypes’, *PLOS Neglected Tropical Diseases*, 13(5), p. e0007231. Available at: https://doi.org/10.1371/journal.pntd.0007231.

Ghosh, S.K. and Dash, A.P. (2007) ‘Larvivorous fish against malaria vectors: a new outlook’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 101(11), pp. 1063–1064. Available at: https://doi.org/10.1016/j.trstmh.2007.07.008.

Godói, I.P. *et al.* (2017) ‘CYD-TDV dengue vaccine: systematic review and meta-analysis of efficacy, immunogenicity and safety’, *Journal of Comparative Effectiveness Research*, 6(2), pp. 165–180. Available at: https://doi.org/10.2217/cer-2016-0045.

Goethals, O. *et al.* (2023) ‘Blocking NS3–NS4B interaction inhibits dengue virus in non-human primates’, *Nature*, 615(7953), pp. 678–686. Available at: https://doi.org/10.1038/s41586-023-05790-6.

Goodwin, S., McPherson, J.D. and McCombie, W.R. (2016) ‘Coming of age: ten years of next-generation sequencing technologies’, *Nature Reviews Genetics*, 17(6), pp. 333–351. Available at: https://doi.org/10.1038/nrg.2016.49.

Gubler, D.J. (1998) ‘Dengue and Dengue Hemorrhagic Fever’, *Clinical Microbiology Reviews*, 11(3), pp. 480–496. Available at: https://doi.org/10.1128/cmr.11.3.480.

Gurukumar, K. *et al.* (2009) ‘Development of real time PCR for detection and quantitation of Dengue Viruses’, *Virology Journal*, 6(1), p. 10. Available at: https://doi.org/10.1186/1743-422X-6-10.

Guzman, M.G. *et al.* (2010) ‘Dengue: a continuing global threat’, *Nature Reviews Microbiology*, 8(12), pp. S7–S16. Available at: https://doi.org/10.1038/nrmicro2460.

Hadinegoro, S.R. *et al.* (2015) ‘Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease’, *New England Journal of Medicine*, 373(13), pp. 1195–1206. Available at: https://doi.org/10.1056/NEJMoa1506223.

Halstead, S.B., Shotwell, H. and Casals, J. (1973) ‘Studies on the pathogenesis of dengue infection in monkeys. II. Clinical laboratory responses to heterologous infection’, *The Journal of Infectious Diseases*, 128(1), pp. 15–22. Available at: https://doi.org/10.1093/infdis/128.1.15.

Hendrichs, J., Dyck, A.A. and Robinson, A.S. (2005) *Sterile insect technique: principles and practice in area-wide integrated pest managment*. Dordrecht: Springer.

Hill, V. *et al.* (2024) ‘A new lineage nomenclature to aid genomic surveillance of dengue virus’, *medRxiv*, p. 2024.05.16.24307504. Available at: https://doi.org/10.1101/2024.05.16.24307504.

Hong, S.-H. (2023) ‘Influence of Microbiota on Vaccine Effectiveness: “Is the Microbiota the Key to Vaccine-induced Responses?”’, *Journal of Microbiology (Seoul, Korea)*, 61(5), pp. 483–494. Available at: https://doi.org/10.1007/s12275-023-00044-6.

Horstick, O., Tozan, Y. and Wilder-Smith, A. (2015) ‘Reviewing Dengue: Still a Neglected Tropical Disease?’, *PLOS Neglected Tropical Diseases*, 9(4), p. e0003632. Available at: https://doi.org/10.1371/journal.pntd.0003632.

Hotta, S. (1952) ‘Experimental studies on dengue: I. Isolation, identification and modification of the virus’, *The Journal of infectious diseases*, pp. 1–9.

Jain, M. *et al.* (2018) ‘Nanopore sequencing and assembly of a human genome with ultra-long reads’, *Nature Biotechnology*, 36(4), pp. 338–345. Available at: https://doi.org/10.1038/nbt.4060.

Kans, J. (2024) ‘Entrez Direct: E-utilities on the Unix Command Line’, in *Entrez Programming Utilities Help [Internet]*. National Center for Biotechnology Information (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK179288/ (Accessed: 23 July 2024).

Kesteleyn, B. *et al.* (2024) ‘Discovery of JNJ-1802, a First-in-Class Pan-Serotype Dengue Virus NS4B Inhibitor’, *Journal of Medicinal Chemistry*, 67(5), pp. 4063–4082. Available at: https://doi.org/10.1021/acs.jmedchem.3c02336.

Ko, H.-Y. *et al.* (2020) ‘Application of Next-Generation Sequencing to Reveal How Evolutionary Dynamics of Viral Population Shape Dengue Epidemiology’, *Frontiers in Microbiology*, 11. Available at: https://doi.org/10.3389/fmicb.2020.01371.

Koren, S. *et al.* (2012) ‘Hybrid error correction and de novo assembly of single-molecule sequencing reads’, *Nature Biotechnology*, 30(7), pp. 693–700. Available at: https://doi.org/10.1038/nbt.2280.

Kuno, G. (1995) ‘Review of the Factors Modulating Dengue Transmission’, *Epidemiologic Reviews*, 17(2), pp. 321–335. Available at: https://doi.org/10.1093/oxfordjournals.epirev.a036196.

Langlois, J. (2024) *Dengue Fever Is Spreading With Climate Change and Urbanization | Sierra Club*, *Sierra*. Available at: https://www.sierraclub.org/sierra/dengue-fever-spreading-climate-change-and-urbanization (Accessed: 19 March 2024).

Laub, V. *et al.* (2023) ‘Bioinformatics for wet-lab scientists: practical application in sequencing analysis’, *BMC Genomics*, 24(1), pp. 1–19. Available at: https://doi.org/10.1186/s12864-023-09454-7.

Lee, K.-S. *et al.* (2010) ‘Dengue Virus Surveillance for Early Warning, Singapore’, *Emerging Infectious Diseases*, 16(5), pp. 847–849. Available at: https://doi.org/10.3201/eid1605.091006.

Letunic, I. and Bork, P. (2024) ‘Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool’, *Nucleic Acids Research*, 52(W1), pp. W78–W82. Available at: https://doi.org/10.1093/nar/gkae268.

Li, H. *et al.* (2009) ‘The Sequence Alignment/Map format and SAMtools’, *Bioinformatics (Oxford, England)*, 25(16), pp. 2078–2079. Available at: https://doi.org/10.1093/bioinformatics/btp352.

Li, H. and Durbin, R. (2009) ‘Fast and accurate short read alignment with Burrows–Wheeler transform’, *Bioinformatics*, 25(14), pp. 1754–1760. Available at: https://doi.org/10.1093/bioinformatics/btp324.

Lin, H.-E. *et al.* (2012) ‘Analysis of Epitopes on Dengue Virus Envelope Protein Recognized by Monoclonal Antibodies and Polyclonal Human Sera by a High Throughput Assay’, *PLOS Neglected Tropical Diseases*, 6(1), p. e1447. Available at: https://doi.org/10.1371/journal.pntd.0001447.

Lindenbach, B.D. *et al.* (2013) ‘Flaviviridae. Fields Virology’, *Lippincott Williams & Wilkins, Philadelphia, PA*, 1, pp. 712–746.

Mackay, I.M., Arden, K.E. and Nitsche, A. (2002) ‘Real-time PCR in virology’, *Nucleic Acids Research*, 30(6), pp. 1292–1305.

Madeira, F. *et al.* (2024) ‘The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024’, *Nucleic acids research*, 52(W1), pp. W521–W525. Available at: https://doi.org/10.1093/nar/gkae241.

Malavige, G.N. *et al.* (2023) ‘Facing the escalating burden of dengue: Challenges and perspectives’, *PLOS Global Public Health*, 3(12), p. e0002598. Available at: https://doi.org/10.1371/journal.pgph.0002598.

Margulies, M. *et al.* (2005) ‘Genome Sequencing in Open Microfabricated High Density Picoliter Reactors’, *Nature*, 437(7057), pp. 376–380. Available at: https://doi.org/10.1038/nature03959.

Marques-Toledo, C.A. *et al.* (2019) ‘Probability of dengue transmission and propagation in a non-endemic temperate area: conceptual model and decision risk levels for early alert, prevention and control’, *Parasites & Vectors*, 12(1), p. 38. Available at: https://doi.org/10.1186/s13071-018-3280-z.

Mayer, S.V., Tesh, R.B. and Vasilakis, N. (2017) ‘The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers’, *Acta Tropica*, 166, pp. 155–163. Available at: https://doi.org/10.1016/j.actatropica.2016.11.020.

McCormack, C.P. *et al.* (2023) ‘Modelling the impact of JNJ-1802, a first-in-class dengue inhibitor blocking the NS3-NS4B interaction, on in-vitro DENV-2 dynamics’, *PLOS Computational Biology*, 19(12), p. e1011662. Available at: https://doi.org/10.1371/journal.pcbi.1011662.

Menzel, P., Ng, K.L. and Krogh, A. (2016) ‘Fast and sensitive taxonomic classification for metagenomics with Kaiju’, *Nature Communications*, 7(1), p. 11257. Available at: https://doi.org/10.1038/ncomms11257.

Messina, J.P. *et al.* (2019) ‘The current and future global distribution and population at risk of dengue’, *Nature Microbiology*, 4(9), pp. 1508–1515. Available at: https://doi.org/10.1038/s41564-019-0476-8.

Mondini, A. and Chiaravalloti-Neto, F. (2008) ‘Spatial correlation of incidence of dengue with socioeconomic, demographic and environmental variables in a Brazilian city’, *Science of The Total Environment*, 393(2), pp. 241–248. Available at: https://doi.org/10.1016/j.scitotenv.2008.01.010.

Mukhopadhyay, S., Kuhn, R.J. and Rossmann, M.G. (2005) ‘A structural perspective of the flavivirus life cycle’, *Nature Reviews Microbiology*, 3(1), pp. 13–22. Available at: https://doi.org/10.1038/nrmicro1067.

Mullis, K.B. and Faloona, F.A. (1987) ‘[21] Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction’, in *Methods in Enzymology*. Academic Press (Recombinant DNA Part F), pp. 335–350. Available at: https://doi.org/10.1016/0076-6879(87)55023-6.

Murali, A., Bhargava, A. and Wright, E.S. (2018) ‘IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences’, *Microbiome*, 6(1), pp. 1–14. Available at: https://doi.org/10.1186/s40168-018-0521-5.

NCBI Resource Coordinators (2018) ‘Database resources of the National Center for Biotechnology Information’, *Nucleic Acids Research*, 46(D1), pp. D8–D13. Available at: https://doi.org/10.1093/nar/gkx1095.

Notomi, T. *et al.* (2000) ‘Loop-mediated isothermal amplification of DNA’, *Nucleic Acids Research*, 28(12), p. e63.

Organization, W.H. (1997) *Dengue haemorrhagic fever: diagnosis, treatment, prevention and control*. World Health Organization.

Ozer, E.A. (2024) ‘egonozer/in\_silico\_pcr’. Available at: https://github.com/egonozer/in\_silico\_pcr (Accessed: 5 July 2024).

Parida, M. *et al.* (2005) ‘Rapid Detection and Differentiation of Dengue Virus Serotypes by a Real-Time Reverse Transcription-Loop-Mediated Isothermal Amplification Assay’, *Journal of Clinical Microbiology*, 43(6), pp. 2895–2903. Available at: https://doi.org/10.1128/JCM.43.6.2895-2903.2005.

Piovezan-Borges, A.C. *et al.* (2022) ‘Global trends in research on the effects of climate change on Aedes aegypti: international collaboration has increased, but some critical countries lag behind’, *Parasites & Vectors*, 15(1), p. 346. Available at: https://doi.org/10.1186/s13071-022-05473-7.

Rafsanjani, M.R. (2022) ‘rafsanlab/ScrapPaper’. Available at: https://github.com/rafsanlab/ScrapPaper (Accessed: 23 July 2024).

Rather, I.A. *et al.* (2017) ‘Prevention and Control Strategies to Counter Dengue Virus Infection’, *Frontiers in Cellular and Infection Microbiology*, 7. Available at: https://doi.org/10.3389/fcimb.2017.00336.

Rivera, J.A. *et al.* (2023) ‘Genotyping of dengue virus from infected tissue samples embedded in paraffin’, *Virology Journal*, 20(1), p. 100. Available at: https://doi.org/10.1186/s12985-023-02072-5.

Rohani, A. *et al.* (2009) ‘The effect of extrinsic incubation temperature on development of dengue serotype 2 and 4 viruses in Aedes aegypti (L.)’, *The Southeast Asian Journal of Tropical Medicine and Public Health*, 40(5), pp. 942–950.

Roy, S.K. and Bhattacharjee, S. (2021) ‘Dengue virus: epidemiology, biology, and disease aetiology’, *Canadian Journal of Microbiology*, 67(10), pp. 687–702. Available at: https://doi.org/10.1139/cjm-2020-0572.

Samune, Y. *et al.* (2024) ‘Genetic regions affecting the replication and pathogenicity of dengue virus type 2’, *PLOS Neglected Tropical Diseases*, 18(1), p. e0011885. Available at: https://doi.org/10.1371/journal.pntd.0011885.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) ‘DNA sequencing with chain-terminating inhibitors’, *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463–5467.

Sarkar, Md.M.H. *et al.* (2023) ‘Comparative phylogenetic analysis and transcriptomic profiling of Dengue (DENV-3 genotype I) outbreak in 2021 in Bangladesh’, *Virology Journal*, 20(1), p. 127. Available at: https://doi.org/10.1186/s12985-023-02030-1.

Schäffer, A.A. *et al.* (2020) ‘VADR: validation and annotation of virus sequence submissions to GenBank’, *BMC Bioinformatics*, 21(1), p. 211. Available at: https://doi.org/10.1186/s12859-020-3537-3.

Schoch, C.L. *et al.* (2020) ‘NCBI Taxonomy: a comprehensive update on curation, resources and tools’, *Database: The Journal of Biological Databases and Curation*, 2020, p. baaa062. Available at: https://doi.org/10.1093/database/baaa062.

Sevigny, J.L. (2024) ‘Joseph7e/Assign-Taxonomy-with-BLAST’. Available at: https://github.com/Joseph7e/Assign-Taxonomy-with-BLAST (Accessed: 14 June 2024).

Shahid, U. *et al.* (2020) ‘Comparison of clinical presentation and out-comes of Chikungunya and Dengue virus infections in patients with acute undifferentiated febrile illness from the Sindh region of Pakistan’, *PLoS Neglected Tropical Diseases*, 14(3), p. e0008086. Available at: https://doi.org/10.1371/journal.pntd.0008086.

Shendure, J. and Ji, H. (2008) ‘Next-generation DNA sequencing’, *Nature Biotechnology*, 26(10), pp. 1135–1145. Available at: https://doi.org/10.1038/nbt1486.

Shu, P.-Y. *et al.* (2003) ‘Development of Group- and Serotype-Specific One-Step SYBR Green I-Based Real-Time Reverse Transcription-PCR Assay for Dengue Virus’, *Journal of Clinical Microbiology*, 41(6), pp. 2408–2416. Available at: https://doi.org/10.1128/JCM.41.6.2408-2416.2003.

Simmons, C.P. *et al.* (2012) ‘Dengue’, *New England Journal of Medicine*, 366(15), pp. 1423–1432. Available at: https://doi.org/10.1056/NEJMra1110265.

SPF (2024) *Dengue aux Antilles. Point au 13 juin 2024.* Available at: https://www.santepubliquefrance.fr/regions/antilles/documents/bulletin-regional/2024/dengue-aux-antilles.-point-au-13-juin-2024 (Accessed: 6 July 2024).

Sridhar, S. *et al.* (2018) ‘Effect of Dengue Serostatus on Dengue Vaccine Safety and Efficacy’, *The New England Journal of Medicine*, 379(4), pp. 327–340. Available at: https://doi.org/10.1056/NEJMoa1800820.

Steinegger, M. and Söding, J. (2017) ‘MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets’, *Nature Biotechnology*, 35(11), pp. 1026–1028. Available at: https://doi.org/10.1038/nbt.3988.

Suggs, S.V. *et al.* (1981) ‘Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin.’, *Proceedings of the National Academy of Sciences of the United States of America*, 78(11), pp. 6613–6617.

Thomas, T., Gilbert, J. and Meyer, F. (2012) ‘Metagenomics - a guide from sampling to data analysis’, *Microbial Informatics and Experimentation*, 2, p. 3. Available at: https://doi.org/10.1186/2042-5783-2-3.

Turnbaugh, P.J. *et al.* (2007) ‘The Human Microbiome Project’, *Nature*, 449(7164), pp. 804–810. Available at: https://doi.org/10.1038/nature06244.

Vaughn, D.W. *et al.* (2000) ‘Dengue Viremia Titer, Antibody Response Pattern, and Virus Serotype Correlate with Disease Severity’, *The Journal of Infectious Diseases*, 181(1), pp. 2–9. Available at: https://doi.org/10.1086/315215.

Vilsker, M. *et al.* (2019) ‘Genome Detective: an automated system for virus identification from high-throughput sequencing data’, *Bioinformatics*, 35(5), pp. 871–873. Available at: https://doi.org/10.1093/bioinformatics/bty695.

Vogels, C.B.F. *et al.* (2023) ‘DengueSeq: A pan-serotype whole genome amplicon sequencing protocol for dengue virus’, *medRxiv: The Preprint Server for Health Sciences*, p. 2023.10.13.23296997. Available at: https://doi.org/10.1101/2023.10.13.23296997.

Waggoner, J.J. *et al.* (2016) ‘Viremia and Clinical Presentation in Nicaraguan Patients Infected With Zika Virus, Chikungunya Virus, and Dengue Virus’, *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 63(12), pp. 1584–1590. Available at: https://doi.org/10.1093/cid/ciw589.

Wang, P. *et al.* (2016) ‘Molecular characterization and phylogenetic analysis of dengue virus type 1 in Guangdong in 2014’, *SpringerPlus*, 5(1), p. 1942. Available at: https://doi.org/10.1186/s40064-016-3604-4.

Wilder-Smith, A. *et al.* (2017) ‘Epidemic arboviral diseases: priorities for research and public health’, *The Lancet Infectious Diseases*, 17(3), pp. e101–e106. Available at: https://doi.org/10.1016/S1473-3099(16)30518-7.

Zimmermann, P. and Curtis, N. (2018) ‘The influence of the intestinal microbiome on vaccine responses’, *Vaccine*, 36(30), pp. 4433–4439. Available at: https://doi.org/10.1016/j.vaccine.2018.04.066.

# Appendix

## Primer Pair for Dengue Virus Overall Metrics

Table 13: Primer Pair for Dengue Virus Overall Metrics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Primer\_Pair** | **F1\_Score** | **Precision** | **Recall** | **Support** |
| 1162\_F\_D4\_R | 1.000 | 1.000 | 1.000 | 43.000 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1569.000 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 | 3224.000 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 1.000 | 1.000 | 1.000 | 4648.000 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 340.000 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 458.000 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 306.000 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 406.000 |
| DENV3\_RN\_F2\_F\_D3\_R | 1.000 | 1.000 | 1.000 | 20.000 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 1.000 | 1.000 | 1.000 | 652.000 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 1.000 | 1.000 | 1.000 | 358.000 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1008.000 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 1.000 | 1.000 | 1.000 | 1100.000 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 1.000 | 1.000 | 1.000 | 1520.000 |
| NS3-2-I\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 | 78.000 |
| NS3-2-O\_F\_NS3-2-I\_R | 1.000 | 1.000 | 1.000 | 72.000 |
| NS3-2-O\_F\_NS3-2-O\_R | 1.000 | 1.000 | 1.000 | 76.000 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1928.000 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 3795.000 |
| D2F1200\_F\_P2\_R | 1.000 | 1.000 | 1.000 | 27.000 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 2145.000 |
| D2S\_F\_D2\_R | 1.000 | 1.000 | 1.000 | 90.000 |
| D3F1187\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 | 62.000 |
| D3SEQ2\_F\_D3R1758\_R | 1.000 | 1.000 | 1.000 | 109.000 |
| DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1.000 | 1.000 | 1.000 | 85.000 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 1888.000 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 796.000 |
| RSS3\_F\_RSS4\_R | 1.000 | 1.000 | 1.000 | 53.000 |
| DENV-4\_F\_D4\_R | 1.000 | 1.000 | 1.000 | 19.000 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 512.000 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 225.000 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1.000 | 1.000 | 1.000 | 324.000 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 0.983 | 0.977 | 0.988 | 1720.000 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 0.981 | 0.981 | 0.981 | 829.000 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 0.961 | 0.948 | 0.974 | 2295.000 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 0.953 | 0.938 | 0.968 | 1837.000 |
| DENV-2\_F\_DEN-2\_R | 0.000 | 0.000 | 0.000 | 1.000 |

## Missed Classification of Conventional PCR Primer Pairs

Table 14: Missed Classification of Conventional PCR Primer Pairs

|  |  |  |
| --- | --- | --- |
| **Primer\_Pair** | **Ground Truth Genotype** | **Number of observations** |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 1216 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 1023 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 843 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 563 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 544 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 454 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | 446 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 396 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 392 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2II | 385 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2III | 384 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | 326 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | 324 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 294 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 287 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 286 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 278 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 276 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 274 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4II | 271 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 259 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4I | 257 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 223 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | 206 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 4II | 196 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | 193 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 186 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 185 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 181 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | 171 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4I | 166 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | 165 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4II | 151 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 151 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 4II | 150 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 148 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 4I | 136 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3III | 133 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 124 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 118 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 110 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3I | 110 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3I | 108 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 4I | 106 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 105 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | 99 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2V | 94 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 4II | 94 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 4I | 94 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 87 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3III | 86 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 83 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3III | 81 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1I | 77 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 76 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | 76 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 76 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 75 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | 75 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2I | 75 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 72 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 70 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 69 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1V | 65 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 64 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | 60 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 56 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 51 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 51 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 51 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 49 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 48 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1I | 48 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 45 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3I | 43 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 43 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3II | 42 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | 40 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 40 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3I | 38 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 38 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 38 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 38 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 36 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | Unassigned | 36 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3II | 36 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 34 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 34 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3II | 32 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3III | 30 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | 28 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3V | 28 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 4II | 27 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | Unassigned | 26 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 19 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4III | 19 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 19 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4III | 19 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4IV | 19 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 19 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 19 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 18 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | Unassigned | 18 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 4I | 17 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3V | 17 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1IV | 17 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | Unassigned | 17 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | Unassigned | 17 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | 15 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3V | 15 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1VI | 15 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1II | 14 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | Unassigned | 13 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1II | 9 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 2II | 8 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3II | 7 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | Unassigned | 6 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | Unassigned | 3 |
| NS3-2-O\_F\_NS3-2-I\_R | Unassigned | 2 |
| NS3-2-I\_F\_NS3-2-I\_R | Unassigned | 2 |
| DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1VI | 2 |
| DENV3\_RN\_F2\_F\_D3\_R | 3I | 2 |
| DENV3\_RN\_F2\_F\_D3\_R | 3II | 1 |
| DENV-2\_F\_DEN-2\_R | Unassigned | 1 |
| NS3-2-O\_F\_NS3-2-O\_R | Unassigned | 1 |
| RSS3\_F\_RSS4\_R | 2V | 1 |

## Mismatched Classification of Conventional PCR Primer Pairs

Table 15: Missed Classification of Conventional PCR Primer Pairs

|  |  |  |  |
| --- | --- | --- | --- |
| **Primer Pair** | **Clustered Genotype** | **Amplicon Genotype** | **Number of Observations** |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2III | 810 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2III | 635 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2II | 528 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2III | 422 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2V | 353 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2II | 2III | 318 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2III | 2II | 316 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2II | 311 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2II | 299 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2II | 297 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2V | 296 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2II | 243 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2III | 241 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2V | 241 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4II | 4I | 231 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2III | 225 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4I | 4II | 217 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2III | 196 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 4II | 4I | 196 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2II | 196 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2V | 190 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4I | 4II | 155 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 4II | 4I | 150 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4II | 4I | 144 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | 1III | 143 |
| DENV4\_7\_LEFT\_F\_DENV4\_RN\_R1\_R | 4I | 4II | 136 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | 1V | 131 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | 1I | 130 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1IV | 127 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1IV | 122 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2III | 108 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2V | 106 |
| DENV4\_22\_LEFT\_F\_DENV4\_22\_RIGHT\_R | 4I | 4II | 106 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | 1IV | 105 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | Unassigned | 102 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | 1IV | 98 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2V | 96 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 4II | 4I | 94 |
| DENV4\_17\_LEFT\_F\_DEN4\_R5525\_R | 4I | 4II | 94 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | 1IV | 91 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 1IV | 90 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | 1V | 83 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2III | 81 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 1V | 80 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3I | 3III | 80 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | 1I | 79 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | Unassigned | 78 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 1IV | 76 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | 1V | 73 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 1I | 73 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | 1V | 72 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1V | 71 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 1III | 71 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | 1IV | 66 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 1V | 66 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1V | 61 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1V | 60 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 1IV | 58 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3III | 3V | 57 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | 1III | 56 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1I | 56 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2II | 56 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1V | 56 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1V | 1I | 55 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1I | 1V | 55 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2III | 54 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1I | 53 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3III | 3I | 52 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2III | 51 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1I | 50 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1I | 49 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3III | 3I | 49 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 1IV | 48 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | 1I | 48 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1III | 47 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | 1IV | 47 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | Unassigned | 46 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2I | 45 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2II | 45 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 1VI | 44 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | Unassigned | 42 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 1V | 42 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2I | 42 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1III | 42 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2V | 40 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | 1III | 40 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | 1V | 40 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2II | 40 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2V | 39 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | 1V | 39 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3III | 3II | 38 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2V | 2III | 37 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2III | 2V | 37 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1III | 37 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2II | 36 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3I | Unassigned | 36 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2V | 36 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 1IV | 36 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1V | 35 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2V | 35 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2I | 35 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2III | 35 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | Unassigned | 34 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | Unassigned | 33 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2III | 33 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 1IV | 33 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 1V | 33 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 1I | 33 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | 1V | 32 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 1I | 32 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | 1V | 32 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | Unassigned | 32 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1III | 31 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | 1I | 31 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3I | 3III | 31 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | 1IV | 31 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | Unassigned | 31 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2I | 30 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2II | 29 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1V | 1VI | 29 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1V | 29 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3III | 3V | 28 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | Unassigned | 28 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 1V | 28 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3I | 3V | 27 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2I | 27 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 1I | 27 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2II | 2V | 27 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2I | 27 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 4II | 4I | 27 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2V | 2II | 27 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3I | 3II | 27 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 1V | 26 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2V | 26 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2I | 2III | 25 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2III | 25 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2I | 2V | 25 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2V | 2I | 25 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2III | 2I | 25 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 1III | 24 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1I | Unassigned | 24 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2II | 24 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2I | 24 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2III | 24 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 1V | 24 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2III | 24 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2I | 24 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2II | 23 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | Unassigned | 23 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2V | 22 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2II | 22 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3II | 3III | 22 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3I | 3III | 22 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1V | 1VI | 22 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 1III | 21 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | Unassigned | 1I | 21 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2I | 21 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2I | 21 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 1VI | 21 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2IV | 20 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2I | 20 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2I | 20 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | Unassigned | 20 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2II | 20 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4I | Unassigned | 20 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3II | 3III | 20 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4II | Unassigned | 20 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2II | 2I | 20 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1IV | 20 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2I | 2II | 20 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3III | 3II | 20 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2II | Unassigned | 20 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2V | 19 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2II | 19 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3III | 3I | 19 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2II | 19 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2I | 19 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2V | 19 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3III | Unassigned | 19 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 1IV | 19 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 1IV | 19 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 1IV | 19 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 1III | 18 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 1I | 18 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | 2I | 18 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2II | 18 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2III | 18 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2III | 18 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3I | 3III | 18 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2III | 18 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1I | 18 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2III | 18 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2V | 18 |
| DENV4\_18\_LEFT\_F\_DEN4\_REVERSE\_R | 4I | 4II | 17 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | Unassigned | 17 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | Unassigned | 16 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 1III | 16 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3I | 3II | 16 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1VI | 16 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | Unassigned | 16 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1I | 16 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | 1III | 16 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1VI | 16 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2III | 16 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 1I | 16 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 1V | 15 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2IV | 15 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2V | 15 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2II | 15 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2II | 15 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2V | 15 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | 2I | 15 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3II | 3I | 15 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2II | 15 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | 2V | 15 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2III | 15 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2II | 2I | 15 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | Unassigned | 14 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | 1V | 14 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2III | 14 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2V | 14 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 1III | 14 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1I | 1VI | 14 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | Unassigned | 14 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | 1IV | 14 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | Unassigned | 13 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3II | Unassigned | 12 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1II | 12 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1I | 1II | 12 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1I | 1IV | 12 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 1III | 12 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3III | Unassigned | 12 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 1V | 12 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | 1I | 12 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 1I | 12 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | 1IV | 12 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 1V | 12 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | Unassigned | 12 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1IV | 1VI | 12 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2III | Unassigned | 12 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 1IV | 12 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3V | 3III | 11 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4I | 4III | 11 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3I | Unassigned | 11 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1II | 11 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2IV | 11 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 1III | 11 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 1I | 11 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | Unassigned | 3III | 11 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | 1VI | 11 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4III | 4I | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | Unassigned | 1I | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | Unassigned | 1I | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1II | 1I | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1I | Unassigned | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1IV | 1I | 11 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2V | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1VI | 1I | 11 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 1I | 11 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1I | 1IV | 11 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4I | 4III | 10 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 1I | 10 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2IV | 10 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | Unassigned | 3I | 10 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | Unassigned | 10 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4II | 4III | 10 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3V | 3III | 10 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3V | 3I | 10 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4II | 4IV | 10 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4I | 4IV | 10 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 1V | 10 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1IV | 10 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2II | 10 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3III | 3I | 10 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | Unassigned | 9 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1I | 1III | 9 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2IV | 9 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2II | 9 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | Unassigned | 9 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 1III | 9 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2I | 9 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3II | 3V | 9 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 1I | 9 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | Unassigned | 1VI | 9 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | Unassigned | 1V | 9 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | 1III | 9 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2II | 9 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | Unassigned | 4I | 8 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4III | 4II | 8 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3V | Unassigned | 8 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4IV | 4I | 8 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 1VI | 8 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2IV | 8 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1II | 1V | 8 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2III | 8 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 1II | 8 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 1VI | 8 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3I | 3V | 8 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3III | Unassigned | 8 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4IV | 4II | 8 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4III | 4II | 8 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | Unassigned | 4II | 8 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4III | 4I | 8 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | 1I | 8 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3II | 3I | 8 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3II | Unassigned | 8 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2III | 2IV | 8 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | Unassigned | 8 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1V | 1II | 8 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 1II | 8 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2V | 7 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | Unassigned | 7 |
| DEN4\_F5299\_F\_DENV4\_19\_RIGHT\_R | 4II | 4III | 7 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2III | 7 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2I | 7 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1IV | 1VI | 7 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | Unassigned | 7 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | Unassigned | 7 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2III | 7 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | Unassigned | 3V | 6 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 2II | 1V | 6 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2III | 6 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2III | Unassigned | 6 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3II | 3III | 6 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3III | 3V | 6 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2III | 6 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2III | 6 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2IV | 6 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2III | Unassigned | 6 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3III | 3II | 6 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | 1I | 6 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | Unassigned | 3II | 6 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 1IV | 6 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | Unassigned | 6 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1V | 2II | 6 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2IV | 6 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3V | 3II | 6 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3I | 3V | 6 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2V | 6 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2II | 6 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2I | 6 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1I | 1II | 6 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2III | 6 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2V | 6 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | Unassigned | 6 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3II | 3V | 6 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2V | 2IV | 6 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3V | 3I | 6 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1IV | Unassigned | 6 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3I | Unassigned | 6 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 2II | 5 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | 1VI | 5 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1IV | 1V | 5 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2V | 5 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2II | 5 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | Unassigned | 1V | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2VI | 5 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1V | 1IV | 5 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1V | Unassigned | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 2V | 5 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | Unassigned | 5 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2I | 5 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2I | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2IV | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | Unassigned | 5 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2V | 5 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | Unassigned | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2II | 2VI | 5 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2I | 5 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2V | Unassigned | 5 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 1III | 5 |
| DEN2\_F2707\_F\_DENV2\_13\_RIGHT\_R | 2I | Unassigned | 5 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2VI | 4 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2IV | 4 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3I | 3II | 4 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1III | 4 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 2I | 4 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 1VI | 4 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3II | 3I | 4 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3III | 3II | 4 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 1VI | 4 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 1I | 4 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | Unassigned | 3III | 4 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3I | 3II | 4 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 1II | 4 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | 1V | 4 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | 1IV | 4 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | 1IV | 4 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | Unassigned | 4 |
| DENV3\_5\_LEFT\_F\_DENV3\_5\_RIGHT\_R | 3V | 3III | 4 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2I | 4 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | 1I | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1I | 1II | 3 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2I | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | 1III | 3 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | Unassigned | 3V | 3 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2I | 3 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 1III | 1VI | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 1II | 3 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2I | 3 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2II | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | Unassigned | 3 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 2III | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1V | 1VI | 3 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2III | 2VI | 3 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2IV | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | Unassigned | 1IV | 3 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2V | Unassigned | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | Unassigned | 1VI | 3 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | 2V | 3 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2I | Unassigned | 3 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | 2I | Unassigned | 3 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | Unassigned | 1VI | 3 |
| DENV2\_13\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2V | 3 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2I | 3 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2IV | 3 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3I | 3V | 3 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 1II | 3 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2IV | 3 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | Unassigned | 3I | 3 |
| DENV2\_11\_LEFT\_F\_DENV2\_13\_RIGHT\_R | Unassigned | 2I | 3 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4III | Unassigned | 2 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | 3III | 3V | 2 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | Unassigned | 3III | 2 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4IV | Unassigned | 2 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | Unassigned | 3II | 2 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3V | 3II | 2 |
| DEN1\_CC\_3\_F\_DEN1\_CC\_3\_R | 1VI | Unassigned | 2 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2II | 2 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1III | 2 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 1IV | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | Unassigned | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1VI | 1III | 2 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | 1I | 2 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | 1III | 2 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | 1VI | 2 |
| DENV1\_RN\_F2\_F\_DENV1\_5\_RIGHT\_R | 2II | Unassigned | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 2II | 1IV | 2 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1II | Unassigned | 2 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | 2IV | 2 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2I | Unassigned | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1IV | 2II | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 1VI | 2 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2IV | 2 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1II | 1VI | 2 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2V | 2 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3II | 3III | 2 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1VI | Unassigned | 2 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3II | 3I | 2 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1III | 1VI | 2 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | 2IV | 2 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | Unassigned | 2VI | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1I | Unassigned | 2 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3II | Unassigned | 2 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 2II | 1III | 2 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | Unassigned | 1V | 1 |
| DENV-2\_F\_DEN-2\_R | Unassigned | 2I | 1 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | Unassigned | 1III | 1 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1II | Unassigned | 1 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1IV | 1VI | 1 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | Unassigned | 1I | 1 |
| DENV1\_19\_LEFT\_F\_DENV1\_21\_RIGHT\_R | 1VI | 1II | 1 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3V | 3I | 1 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4III | 4IV | 1 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | Unassigned | 1 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | 4IV | 4III | 1 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1III | 1II | 1 |
| DEN1\_F5132\_F\_DEN1\_CC\_3\_R | 1II | 1III | 1 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | Unassigned | 4III | 1 |
| DENV4\_RN\_F2\_F\_DENV4\_RN\_R2\_R | Unassigned | 4IV | 1 |
| NS3-2-I\_F\_NS3-2-I\_R | Unassigned | 2I | 1 |
| NS3-2-I\_F\_NS3-2-I\_R | Unassigned | 2IV | 1 |
| NS3-2-O\_F\_NS3-2-I\_R | Unassigned | 2I | 1 |
| NS3-2-O\_F\_NS3-2-I\_R | Unassigned | 2IV | 1 |
| NS3-2-O\_F\_NS3-2-O\_R | Unassigned | 2IV | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1II | 1IV | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | 1IV | Unassigned | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1VI | 1II | 1 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | 2II | 1 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | 3V | Unassigned | 1 |
| DENV3\_20\_LEFT\_F\_DENV3\_20\_RIGHT\_R | Unassigned | 3I | 1 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3V | 3II | 1 |
| DENV3\_12\_LEFT\_F\_DENV3\_12\_RIGHT\_R | 3II | 3V | 1 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2V | Unassigned | 1 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | Unassigned | 1 |
| DENV2\_23\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2V | 1 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | Unassigned | 1 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2VI | 2IV | 1 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | Unassigned | 1 |
| DENV2\_22\_LEFT\_F\_FRAGMENT\_A\_R | 2IV | 2I | 1 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2IV | Unassigned | 1 |
| DENV2\_21\_LEFT\_F\_DENV2\_22\_RIGHT\_R | 2II | 2IV | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | 1VI | 1IV | 1 |
| DENV2\_20\_LEFT\_F\_DENV2\_22\_RIGHT\_R | Unassigned | 2IV | 1 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1VI | 1III | 1 |
| DENV1\_4\_LEFT\_F\_GSAODV1IIKWF\_2\_R | 1III | 1VI | 1 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1VI | 1II | 1 |
| DENV1\_4\_LEFT\_F\_DENV1\_5\_RIGHT\_R | 1III | 1VI | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_23\_RIGHT\_R | Unassigned | 1IV | 1 |
| DENV3\_5\_LEFT\_F\_DENV3\_6\_RIGHT\_R | Unassigned | 3V | 1 |
| DENV3\_RN\_F2\_F\_D3\_R | 3I | 3II | 1 |
| DENV3\_RN\_F2\_F\_D3\_R | 3I | Unassigned | 1 |
| DENV3\_RN\_F2\_F\_D3\_R | 3II | Unassigned | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | Unassigned | 1IV | 1 |
| DENV1\_20\_LEFT\_F\_DENV1\_22\_RIGHT\_R | Unassigned | 1II | 1 |
| RSS3\_F\_RSS4\_R | 2V | 2II | 1 |