Phylogenetic analysis of Dengue virus and insect-specific *Flavivirus* in *Aedes aegypti* from Gresik, Indonesia 2019

RASYID NOOR HAKIM1,2, MAMIK DAMAYANTI3, SHIFA FAUZIYAH4, TEGUH HARI SUCIPTO5,6*
1Graduate Program, Department of Program, Faculty of Science and Technology, Universitas Airlangga. Jl. Dr. Ir. H. Soekarno, Campus C Mulyorejo, Surabaya 60115, East Java, Indonesia
2University-CoE, Research Center for Biomolecule Engineering, Universitas Airlangga. Jl. Dr. Ir. H. Soekarno, Campus C Mulyorejo, Surabaya 60115, East Java, Indonesia
5Dengue Study Group, Institute of Tropical Disease, Universitas Airlangga. Jl. Dr. Ir. H. Soekarno, Campus C Mulyorejo, Surabaya 60115, East Java, Indonesia. Tel.:+62-31-5992446, Fax.: +62-31-5992445, *email: teguhharisucipto@staf.unair.ac.id*

Abstract. Hakim RN, Damayanti M, Fauziyah S, Sucipto TH. 2024. Phylogenetic analysis of Dengue virus and insect-specific Flavivirus in Aedes aegypti from Gresik, Indonesia 2019. Biodiversitas 25: 2340-2347. The spread of several *Flavivirus*, such as Dengue virus and insect-specific *Flavivirus*, is periodically detected in Indonesia, including Surabaya, which is an endemic area. Gresik, as one of Surabaya's satellite cities, has high traffic mobility between regions, increasing the possibility of viral transmission. In this study, we identified *Flavivirus* from Aedes aegypti mosquito isolates from Gresik collected in 2019. We also revealed the possibility of virus transmission from Surabaya to Gresik. Samples were collected from seven sampling areas with environmental conditions suitable for Aedes mosquito breeding. The *Flavivirus* Partial Nonstructural Protein 5 (NS5) -encoding gene region was isolated via the nested reverse transcription-polymerase chain reaction method. Sequence alignment and phylogenetic analysis were carried out to determine the kinship of Gresik strain virus isolates with the GenBank database. The results revealed that six out of 29 isolated samples (Sample from Duduk Sampeyan (47 and 43), Mengare (212 and 31), Kedamean (58), and Manyar (38)) exhibited a 220 bp DNA band, which corresponds to the NS5 gene. Further analysis using basic local alignment search tool - nucleotide and phylogenetic methods revealed that two sample sequences (38 and 212) are closely related to DENV-1 genotype 1, while the other four (31, 43, 47, and 58) are associated with Insect-Specific Flavivirus (ISF). The similarity in NS5 partial sequence and genotype classifications of the dengue virus suggests the potential transmission of the virus from Surabaya to Gresik. We hypothesize that this phenomenon is closely linked to the population mobility between the two areas.

Keywords: Dengue virus, molecular detection, Nested RT-PCR, NS5 partial sequence, Surabaya

INTRODUCTION

The genus *Flavivirus* is part of the family Flaviviridae and is known for its ability to be transmitted depending on arthropod vectors such as mosquitoes and ticks (Supriyono et al. 2020). Despite similarities in their genomic organization, *Flaviviruses* have fundamental differences in their host range and transmission capabilities (Blitvich and Firth 2015). Based on these distinctions, the genus can be classified into viruses that are exclusively transmitted vertically among various mosquito species (insect-specific *Flavivirus*) and viruses that can be transmitted from arthropod vectors (mosquitoes or ticks) to vertebrate host cells, including both animals and humans (Guzman et al. 2018; Murugesan and Manoharan 2019).

Dengue Virus (DENV) belongs to the genus *Flavivirus* and includes other viruses that can infect humans, such as Zika virus and *Japanese encephalitis* virus (Supriyono et al. 2020). Among those viruses, DENV has received special attention from the government, considering the high number of dengue hemorrhagic fever infection cases in Indonesia (Harapan et al. 2019). The disease was first identified in 1968 in the cities of Jakarta and Surabaya (Sumarmo 1987). Since then, the number of annual cases of DHF has increased by thousands of cases each year, along with the increase in the number of mosquito vectors and humans (Karyanti et al. 2014; Martin et al. 2016). According to data from the Ministry of Health of the Republic of Indonesia (2022), approximately 300,000 DENV cases were diagnosed in Indonesia between 2016 and 2018. Dengue virus infection is common in Indonesia

On the other hand, there are other *Flaviviruses* classified as Insect-Specific Flavivirus (ISF) that exhibit the unique characteristic of being exclusively vertically transmitted within various mosquito species (Guzman et al. 2018; Murugesan and Manoharan 2019). The vertical transmission of insect-specific *Flavivirus* among mosquitoes has been substantiated through studies on the initial discovery of ISF, the Cell Fusing Agent Virus (CFAV). This virus exhibited pronounced cytopathic effects upon infecting *Aedes aegypti* cell lines while demonstrating an inability to infect vertebrate cells (Stollar and Thomas 1975). Several ISFs have been identified in Indonesia, such as *Culex Flavivirus* and Cell Fusing Agent (Hoshino et al. 2009), *Aedes Flavivirus* (Supriyono et al. 2020) and Mosquito *Flavivirus* Bali (MBF) (Damayanti et al. 2021). This indicates the possibility that ISF has
spread in Indonesia, but data on the extent of this spread are still very limited.

Gresik, as one of Surabaya’s satellite cities, has high population mobility and traffic between regions (Herijanto 2018). Population mobility is high because Surabaya is the center of education, entertainment, industry and also the center of government for East Java, Indonesia. Mobility has the potential to trigger the transmission of infectious viruses such as Dengue virus (Martin et al. 2016). Current studies have reported infections of all four dengue virus serotypes in Surabaya between 2008 and early 2018 (Mulyatno et al. 2018; Soegijanto et al. 2022). Data from the Health Office of East Java (2022) showed that the number of DHF cases in Gresik reached 1,300 cases between 2016 and 2018. However, the absence of evidence supporting the presence of Dengue virus infection due to transmission from Surabaya can be attributed to the lack of research on virus monitoring that circulating in Gresik. Additionally, the lack of data on ISF in Gresik can be attributed to the fact that they do not pose a threat to human health, so monitoring their presence is not prioritized. Research on the circulation of viruses in an area is needed to understand virus dynamics and local epidemics (Hamel et al. 2019).

The monitoring of virus circulation can be performed using molecular techniques. The *Flavivirus* genome consists of approximately 11 kb of positive sense protein encoding three structural proteins (capsid, membrane, and envelope) and seven nonstructural proteins (*NS1, NS2A, NS2B, NS3, NS4* and *NS5*) (Singh et al. 2015). Among these protein-coding genes, the *NS5* gene has the most conserved sequence encoding nonstructural proteins (Scaramozzino et al. 2001; Murugesan and Manoharan 2019). Therefore, partial *NS5* gene sequences can be used as useful targets for viral identification via phylogenetic analysis (Regato et al. 2008).

In this study, we identified *Flaviviruses* from *Aedes aegypti* mosquito isolates collected at Gresik in 2019. This identification is part of an effort to maintain early surveillance and implement effective vector control measures, aiming to reduce the risk of mosquito-borne diseases in the region. Furthermore, we also investigated the potential transmission of viruses from Surabaya to Gresik, considering the significant population mobility between these two areas. This research fills a gap in the scientific reports on virus circulation and transmission in Indonesia, particularly within Gresik District, and we hope these findings provide the groundwork for future efforts to prevent outbreaks in the region.

**MATERIALS AND METHODS**

**Study area**

Sample collection was carried out in several subdistricts of Gresik, East Java, Indonesia, i.e. Bungah, Mengare Island, Manyar, Duduk Sampeyan, Benjeng, Menganti and Kedamean (Figure 1). These locations were chosen situated either in close proximity or with high population mobility to Surabaya. The sampling locations were visualized from the Indonesia Geospatial Portal data (tanahair.indonesia.go.id) using ArcMap 10.8 software.

![Maps of sampling areas](image)

**Figure 1.** Maps of sampling areas are shown in Red line: A. Bungah; B. Mengare Island; C. Manyar; D. Duduk Sampeyan; E. Benjeng; F: Menganti; G. Kedamean
Procedures

Study design

This study was designed as cross-sectional. The research carried out was in the form of sample detection without repetition and using a positive control as validation of the results. The genetic material from the mosquito sample was extracted and used for Flavivirus identification. Screening criteria are carried out by matching the sample band resulting from nested reverse transcription-polymerase chain reaction with the positive control. Samples that meet the criteria were sequenced and the results were analyzed using basic local alignment search tool-nucleotide. Data from the analysis are presented in the form of a phylogenetic tree.

Sample collection

Purposeful sampling was conducted in alongside the aforementioned district in 2019. The determination sampling points was carried out in environmental condition potential was a breeding site for Aedes mosquitoes, which consisting of artificial breedings. The surveyors were examined every artificial breeding site alongside area. If mosquito larvae were found, then it will be pipetted and transferred into laboratory for further process. The obtained samples were pooled and stored at -80°C.

RNA isolation

The samples were suspended in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS). Mosquito cells were then lysed using ultrasonic cell disruption to extract the genetic material. Subsequently, the lysate was filtered, and the filtrate was utilized for viral RNA extraction. RNA extraction was carried out with an RNA extraction kit (QiAamp Viral RNA Mini Kit, Germany) following the manufacturer’s instructions. The RNA samples were stored at -70°C until further use.

Nested reverse transcription -Polymerase

Complementary DNA (cDNA) was synthesized through reverse transcription following the instruction of Scaramozzino et al. (2001) with slight modifications. The cF2D primer (5’-GTGTCCCAGCCGGCGGTGTCATCAGC-3’) was used for reverse transcription. Reverse transcription was conducted in a 14 μL reaction mixture comprising 1 μL of dNTPs (Invitrogen, USA), 1 μL of reverse primer, 5 μL of template RNA, and 7 μL nuclease-free water (NFW). The incubation process occurred at 55°C for 5 minutes. Subsequently, the samples were mixed with 4 μL of 5x FS Buffer (Invitrogen, USA), 1 μL of dithiothreitol (Invitrogen, USA), and 0.5 μL of GoScript (Promega, USA). The mixture underwent incubation in two phases: first at 50°C for 60 minutes and second at 85°C for 5 minutes. Finally, 0.5 μL of RNase H (Invitrogen, USA) was added to the sample, and the mixture was incubated at 37°C for 20 minutes. This series of steps represents the process of reverse transcription, cDNA synthesis, and RNA degradation in the preparation of the samples for molecular analyses.

NS5 gene amplification was performed using nested PCR following the instruction of Scaramozzino et al. (2001) with slight modifications. The first PCR used the following primers: The sense primer MAMD (5’-AACATGATGGGRAARAGRGAARAA-3’) and the reverse primer cF2D. The first PCR using outer primers was performed in 12.5 μL containing 6.25 μL of PCR Master Mix (Promega, USA), 1 μL of primers, 2.75 μL of nuclease-free water, and 2.5 μL of template DNA. Amplification was performed with pre-denaturation at 94°C for 4 minutes; denaturation at 94°C for 1 minute, annealing at 53°C for 1 minute, and extension at 72°C for 1 minute; and a final extension at 72°C for 10 minutes, with was repeated for 35 cycles.

For the second PCR, the following nested primers were used: sense primer FS778 (5’-AARGGHAGYMCDGCHATHTGTT-3’) and reverse primer cF2D. Nested PCR was performed in 12.5 μL of a similar mixture with 1.25 μL of template DNA (amplification product from the first PCR). Amplification was performed with pre-denaturation at 94°C for 4 minutes; denaturation at 94°C for 1 minute, annealing at 54°C for 1 minute, and extension at 72°C for 1 minute; and a final extension at 72°C for 10 minutes, with was repeated for 35 cycles.

The amplification results were visualized using a 1% agarose gel. As a positive control, the DENV-2 isolate from Surabaya (GenBank accession number: KT012509) was utilized. Samples that displayed a band size of approximately 220 base pairs on the gel were selected for sequencing.

Phylogenetic tree construction

The sequencing electrogams were edited using BioEdit 7.2.5.0 software to ensure the accuracy and quality of the sequences. Subsequently, the edited sequences were compared with the GenBank database available on the National Center for Biotechnology Information (NCBI) using the Basic Local Alignment Search Tool - nucleotide (BLASTn) program. NS5 gene sequences in the database displaying high similarity to the samples of interest were then downloaded for further analysis using a phylogenetic tree. Sequence alignment and phylogenetic trees were constructed using MEGA 11 software. The neighbor-joining method was employed to generate phylogenetic trees, allowing for a clear visualization of the genetic relationships among the samples and their counterparts in the database.

RESULTS AND DISCUSSION

Gresik, a region located in East Java, Indonesia exhibits astronomical coordinates ranging between 112°-113° East longitude and 7°-8° South latitude. The District encompass 18 administrative districts, with approximately thirty percent of its land area constituting coastal regions. The mean elevation of the terrain is less than 200 meters above sea level, accompanied by a notable precipitation rate. Geographically, Gresik is demarcated by the Java Sea to the north, Lamongan to the west, the Madura Strait to the east, and Surabaya as the capital of East Java to the south (BPS 2023).

The activities occurring in the regional and urban centers such as Surabaya significantly impacted Gresik as a suburban area (Herijanto 2018). Surabaya, serving as a major gateway for both domestic and international travelers and trade, plays a crucial role in facilitating the seamless
movement of people and goods. This pivotal function is primarily attributed to its well-established airport and seaport infrastructure. This influence has led to a considerable number of trips in these regions. This reflects the urban sprawl phenomenon in Surabaya, where population mobility in Gresik increases every year, causing a large volume of traffic. Additionally, the elevated housing costs have prompted numerous workers in Surabaya to choose settlements in suburban areas such as Gresik. Consequently, there is a discernible rise in commuting mobility among workers who reside in Gresik. This mobility has the potential to trigger the transmission of infectious viruses from Surabaya, which is considered one of the endemic areas in Indonesia (Martin et al. 2016). The impact of human mobility on the rapid spread of pathogenic viruses is well documented (Karyanti et al. 2014). The risk of Dengue Virus (DENV) transmission in a community is closely linked to the movement of people in neighboring areas (Rabaa et al. 2013).

The monitoring of viruses in Gresik involved the collection of mosquito samples from various locations, including Bungah, Manyar, Mengare Island, Duduk Sampeyan, Benjeng, Kedamean, and Menganti. From these seven areas, 29 samples of *Aedes aegypti* mosquitoes were successfully collected. Female *Aedes aegypti* mosquitoes serve as the primary vector, with *Aedes albopictus* and other *Aedes* mosquitoes acting as secondary vectors (Wang et al. 2000). The importance of collecting samples from various locations provides a more comprehensive overview of virus spread in Gresik. Tracking the prevalence of different viruses in specific regions enhances our overall understanding of virus distribution and local epidemics (Hamel et al. 2019). This can be achieved through the identification of viral genetic material using molecular biology techniques.

Nested-polymerase chain reaction stands out as a valuable molecular biology technique for viral detection, because of its notable advantages of rapidity, heightened sensitivity, and specificity. The technique employs two sets of primers in two consecutive rounds of the nested-polymerase chain reaction. The second set of primers was designed to amplify a specific target sequence within the product generated by the first PCR amplification. The use of nested PCR is known to significantly enhance the sensitivity of the method, achieving a hundredfold increase compared to that of conventional PCR techniques (Salih et al. 2020). The improved sensitivity is provided by nested PCR due to first amplifying the target area with outer primers, which narrows down the region. This initial amplification allows the nested primers to then amplify a more specific and precise segment within that defined area, which can contribute to more accurate and reliable identification of the target gene.

The findings from the sample amplification using nested PCR revealed that of the 29 samples subjected to testing, positive outcomes were observed for six specific samples, namely, 47, 212, 58, 43, 38, and 31. The positive outcome was identified through the distinct presence of a DNA band at approximately 220 bp, as visually depicted in Figure 2. However, the remaining samples illustrated in Figures 3 and 4 did not exhibit the characteristic DNA bands associated with a positive outcome. Further examination of the distribution of these positive samples revealed various geographical locations. Specifically, two positive samples were identified in Duduk Sampeyan (43 and 47), another two in Mengare Island (31 and 212), one in Kedamean (58), and the remaining one in Manyar (38). In pursuit of a more comprehensive understanding, the positive samples were subsequently sequenced for in-depth genetic analysis.
Based on the NS5 partial sequence obtained, alignment using BLAST-n showed that samples 38 and 212 shared a high identity with the Dengue virus. The virus has 4 strains that are classified into four different serotypes due to their antigenic differences: DENV-1, DENV-2, DENV-3, and DENV-4. These antigenic variations cause Dengue hemorrhagic fever to lack cross-protective immunity, although the epidemiological characteristics of these diseases are similar, resulting in symptoms comparable to those of infections in humans (Singh et al. 2015). The strains can also be further divided into several genotypes according to their sequence diversity. Each serotype has 3-5 different genotypes (Islam et al. 2021). The cyclic emergence and dominance of specific serotypes in tropical regions contribute to the ongoing challenges in managing and controlling dengue fever. Public health efforts often focus on surveillance, vector control, and understanding the dynamics of serotype circulation to mitigate the impact of the disease (Martin et al. 2016). Identification of Dengue virus serotypes and genotypes through molecular analysis can be applied by the government as a way to understand the dynamics of circulating viral strains quickly and mitigate viral infection outbreaks in a shorter time. This method facilitates a more comprehensive understanding of the genetic diversity of viruses and helps the characterization of different strains through phylogenetic analysis.

According to the neighbor-joining phylogenetic tree of DENV shown in Figure 5, we identified samples 38 and 212 as DENV-1 genotype 1. DENV-1 manifests distinct genetic variability across five genotypes, a diversity attributed to its geographical distribution. These genotype classifications based on Goncalvez et al. (2002) include (I) Southeast Asia, China, and East Africa; (II) Thailand during the 1950s and 1960s; (III) Malaysia; (IV) the Western Pacific Islands and Australia; and (V) America, West America, and Asia. This genotypic classification reflects the intricate interplay between the virus and its environmental context, offering valuable insights into its evolutionary dynamics across diverse regions. In general, DENV-1 genotypes circulating in Indonesia are dominated by genotypes I and IV (Datu et al. 2023). Based on the phylogenetic tree, both samples were closely related to the Myanmar isolate (KY379162.1). The findings of our study are similar to those of Kotaki et al. (2016), who highlighted the introduction of Dengue virus serotype 1 Genotype 1 into Surabaya from other Southeast Asian countries within 15 years. In the upper clade, the isolates from Gresik were observed to cluster with other DENV-1 genotype 1 isolates from Indonesia, such as those from Jakarta (Lestari et al. 2017), Makassar (Sasmono et al. 2015), Jambi (Yohan et al. 2018) and Surabaya (Mulyatno et al. 2018). The clustering of Gresik isolates with Dengue virus strains from other cities in Indonesia implies that the circulating DENV-1 viruses are likely local and endemic strains in the region (Wardhani et al. 2017).

In contrast to the previous two samples, four other samples (31, 43, 47, and 58) were classified as insect-specific Flavivirus, as shown in Figure 6. The term "insect-specific" refers to viruses that can only replicate in mosquitoes, in contrast to common arboviruses that have multiple hosts (vertebrates and arthropods) (Guzman et al. 2018). The fact that ISF cannot replicate in vertebrates or humans makes their monitoring a lower priority. Nevertheless, due to its pathogenicity toward mosquitoes, ISF can be used for vector-borne disease control (Peinado et al. 2022). Some research has shown that mosquitoes can be infected with ISF and arboviruses simultaneously, and that ISF can affect the ability of mosquitoes to acquire, maintain and transmit these pathogenic viruses (vector competence) (Bolling et al. 2015; Agboli et al. 2019). This can be an alternative effort to reduce or prevent vertical transmission of the Dengue virus.

Figure 5. A phylogenetic tree was constructed based on the DENV-1 classification following the classification of Goncalvez et al. (2002) with outgroup DENV-2-4. The samples analyzed in this study are shown with the codes Sample 38 Gresik-Indonesia 2019 and Sample 212 Gresik-Indonesia 2019.
ISF can be further divided into two distinct groups based on their phylogenetic and antigenic relationships: classical ISF (cISF), which was the first class of ISF to be discovered, and dual-host affiliated ISF (dISF), which is more closely related to pathogenic Flaviviruses such as Zika and Dengue viruses (Hoshino et al. 2012). In Indonesia, research on ISF is still very limited compared to that on pathogenic Flaviviruses. Some of ISFs have been identified, such as Culex Flavivirus and Cell Fusing Agent in Surabaya, 2003-2004 (Hoshino et al. 2009); Mosquito Flavivirus Bali (MFB) in Bali, 2016 (Damayanti et al. 2021); and Aedes Flavivirus in Bogor (Supriyono et al. 2020). Unfortunately, all identification of the ISF in Indonesia was conducted using the envelope gene, thereby limiting our capacity to make direct comparisons and assess kinship through phylogenetic tree analysis. Despite this limitation, based on available reports, there is a presumption that ISF has been spreading in Indonesia. These findings on ISF in Gresik can serve as a valuable addition to reports concerning the virus distribution in Indonesia.

The closeness of Dengue virus strains on the phylogenetic tree is closely related to their ability to maintain intrinsic immunity and response to mutations. As virus strains differ over time due to genetic mutations, they may develop variations in their ability to evade host immune responses and adapt to new environmental conditions. Strains that are closely related to a phylogenetic tree often have similar genetic signatures and may exhibit similar patterns of immune evasion mechanisms. The comprehensive documentation of the four serotypes of the Dengue virus in Surabaya is well established, as evidenced by the works of Kotaki et al. (2016) and Wardhani et al. (2017). These reports collectively support the classification of Surabaya as a dengue-endemic area. Notably, both studies highlighted the prevailing dominance of DENV-1 Genotype 1 in the region. Dengue virus samples acquired from Manyar (38) and Mengare Island (212) raise the possibility of dengue virus transmission from Surabaya to Gresik. This inference is supported by the observed sequence similarity (identity >94.74%), serotype and genotype classification of the dengue virus in both areas. Unfortunately, there is still no data regarding ISF in Surabaya, so we cannot compare them. Apart from eradicating mosquito nests with 3M Plus, this includes draining air reservoirs, closing air reservoirs, and recycling various items that have the potential to hold water and become breeding grounds for the Aedes aegypti mosquito, which carries the dengue virus in humans. In the end, molecular surveillance shows its potential to help the government in making policies for handling dengue virus infections. This method can be used for initial or periodic detection with speed, large number of samples and good accuracy; so that stakeholders can act quickly before a Dengue virus outbreak occurs. It showed the molecular surveillance position to reduce the spread of dengue infection cases.

In conclusion, the detection of Flavivirus in Aedes aegypti samples isolated from Gresik revealed the presence of Dengue virus, which specifically belongs to the serotype 1 genotype 1 group, in the Manyar and Mengare Island areas. The identified Dengue virus exhibited close relationship with regions where there is potential for Dengue virus transmission from Surabaya to Gresik. This conclusion is drawn based on the observed sequence similarity and genotype classification of Dengue virus in both areas. Additionally, our analysis revealed the presence of ISF circulating in Gresik, specifically in Duduk Sampeyan, Mengare and Kedamean. However, the results are five years out of date. To validate these preliminary findings, it is crucial to conduct further monitoring and employ whole-genome sequencing for a more comprehensive and accurate understanding. These additional steps will provide robust supporting data, contributing to a more conclusive assessment of the potential virus transmission dynamics between Surabaya and Gresik. We hope that this research contributes as groundwork to the understanding of both Dengue virus and ISF in Indonesia, particularly in Gresik. By addressing these gaps, we aim to enhance the knowledge about the dynamics of Dengue virus transmission and the potential implications of ISF.

ACKNOWLEDGEMENTS

The authors would like to thank the Institute of Tropical Diseases, Universitas Airlangga, Surabaya, Indonesia which provided laboratory facilities for this research.
REFERENCES


