Chapter 1: Introduction

1.1 Introduction

Dengue is a mosquito-borne viral infection that cause disease in many tropical regions in the world and mainly transmitted by the *Aedes aegypti* mosquito, with *Aedes albopictus* as its secondary vector (Simmons, Farrar, Van Vinh Chau, & Wills, 2012). The causative agent, Dengue virus (DENV), is a single-stranded, positive-strand RNA virus (+ssRNA) that belongs to the *Flavivirus* genus (Wilder-Smith, Ooi, Horstick, & Wills, 2019). DENV genome encodes for ten proteins: three structural proteins of capsid (C), pre-membrane (prM), and envelope (E) and seven non-structural proteins of NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 (Wilder-Smith et al., 2019).

Both E and prM protein forms projection on the surface of the virus, however, E protein is considered of great importance as it interacts with host receptor for viral entry (Wilder-Smith et al., 2019; Zonetti, Coutinho, & Araujo, 2018). Furthermore, E protein also acts as a target for hemagglutination and neutralization by the immune system, which classifies DENV into 4 different serotypes, DENV 1-4, based on antibodies developed following infection and can be further grouped into multiple genotypes that are phylogenetically related (Murphy & Whitehead, 2011; Wilder-Smith et al., 2019). Infection by any of the four serotypes can result in a range of clinical manifestations and induce serotype-specific antibody which provides lifelong protection to the infecting serotype, but may cause severe dengue in concurrent infection with other serotypes (Wilder-Smith et al., 2019). This phenomenon is known as antibody-dependent enhancement (ADE) as existing antibody from previous infection helps viral entry and this phenomenon plays an important role in clinical manifestation, epidemiology, and research (Kudlacek & Metz, 2019). Introduction or replacement of DENV lineages have potential to create severe dengue outbreak due to serological naivety in the population to the new lineages while in vaccine development, it is an important consideration as successful vaccine needs to be able to induce immunity against all four DENV serotypes without causing ADE (Harapan et al., 2019; Murphy & Whitehead, 2011).

Due to their high genome diversity, continuous surveillance of circulating lineages is important (Yohan, Dhenni, et al., 2018). Surveillance is commonly done by serotyping, using the protocol developed by Lanciotti et al. in 1992, as well as commercial qRT-PCR (qualitative real-time polymerase chain reaction) kits which increase convenience for DENV serotyping (Cordeiro, 2012). Despite the convenience of qRT-PCR, genomic studies and comparative analysis of DENV genome sequences are still needed to obtain more in-depth knowledge (Yohan, Dhenni, et al., 2018).

Conventional genome analysis was routinely done using Sanger sequencing and determined by sequence analysis of the E gene or E/NS1 gene junction (Lee et al., 2012). Genomic data from these observations have allowed us to construct phylogenetic tree, which shows the clustering of DENV lineages in association with specific geographical region and demonstrating the close relationships between lineages circulating in one region (Lee et al., 2012). Genome data also allows us to observe DENV evolution dynamics, including mutation and selective pressures, which helps us understand DENV epidemiology, for instance lineage displacement, and vaccine development (Pollett et al., 2018; Yohan, Wardhani, Trimarsanto, Aryati, & Sasmono, 2018). Despite being rare, positive selection drives long-term adaptation which impacts lineage diversity and might alter protein properties, such as viral antigenicity (Bloom, 2017; Desai & Fisher, 2007). Identification of pervasive positive selection sites conserved across serotypes is an important data in vaccine design as difference of lineages and epitopes between vaccine and wild-type strains might affect vaccine efficacy (Halstead, 2019; Pollett et al., 2018). In recent years, whole-genome sequencing (WGS) has gained popularity as it is more time and cost-effective and able to give the complete genome sequences for understanding epidemiology and evolution of DENV (Yohan, Dhenni, et al., 2018). By obtaining whole-genome library of DENV, better data could be gathered that could improve our understanding on dengue.

In this study, we characterize the diversity of dengue virus isolated from a recent outbreak in Jember, East Java, Indonesia (Aryati et al., 2020) by looking at viral lineages using the virus E gene to construct phylogenetic tree. In addition, positive selection sites for both prM and E gene in all of isolated samples were analyzed together with data from historical Indonesian DENV sequences from

NCBI GenBank public database for a better picture of DENV dynamics in Indonesia. Finally, a next-generation sequencing (NGS) library preparation for whole genome sequencing was prepared to gain a comprehensive genome data of the recently isolated DENV from Jember.

1.2 Research Objectives

The objectives of this study are:

(i) Phylogenetic tree construction using E gene sequence from Jember in 2019, which were obtained from Sanger sequencing

(ii) Evolutionary analysis of the prM and E gene of Indonesian DENV sequences by looking at selection pressure

(iii) Next-generation sequencing (NGS) library preparation for DENV whole-genome sequencing to further analyze the DENV genome

1.3 Scope of Work

The following laboratory techniques will be employed to achieve the objectives of this study:

- cDNA synthesis using reverse transcription reaction
- qRT-PCR for DENV detection and serotyping
- PCR for DENV genome amplification
- Template DNA purification from agarose gel electrophoresis
- DNA sequencing using Sanger method
- Bioinformatics analysis of phylogeny and selective pressure
- NGS library preparation for DENV whole-genome sequencing