Genetic variation and genetic structure of *Aedes aegypti* population Dengue Hemorrhagic Fever (DHF) vector in West Sumatera

Hasmiwati¹, and Supargiyono²
{hasmiwati65@med.unand.ac.id¹, supargiyono@gmail.com²}

Department of Parasitology, Faculty of Medicine, Andalas University, Padang¹
Department of Parasitology, Faculty of Medicine, Gajah Mada University, Yogyakarta²

Abstract. *Aedes aegypti* is the main vector of dengue virus in the tropics and subtropics. Prevention of dengue depends on vector control due to the vaccine is still in development, understanding the survival of the species can be determined by tracking the genetic variations among individuals and determine the genetic structure of *Aedes aegypti* population. The purpose of this study was to analyze genetic variation and genetic structure of populations of *Ae. aegypti* in four dengue-endemic region in Sumatra Barat. All collected sample was performed DNA extraction to obtain DNA genomic of *Ae. aegypti*. DNA was amplified by using microsatellite primers resulted from previous study. Homozygous and heterozygous alleles were detected by using PAGE. Genetic variation and genetic differentiation were high in four dengue-endemic regions in West Sumatra. He scores ranged from 0.7886 to 08,404, I score were 1.5215 to 1.7263, Fst score was 0.1205, and Nm was 1.8240. Genetic variation and gene flow of *Ae. aegypti* among populations, so that is easy for mosquito to adapt. There were two groups of *Ae. aegypti* resulted based on UPGMA analysis, and there was no correlation between genetic distance and geographical distance. Genetic variation of *Ae. aegypti* in some endemic areas of dengue fever in West Sumatra was high with moderate genetic differentiation. There were two groups of population based on genetic structure. There was no correlation between genetic distance and geographic distance.

Keywords: *Aedes aegypti*, Dengue Hemorrhagic Fever, Genetic structure of the population, Genetic variation.

1 Introduction

The management of DHF have faced some problems. It was due to the vaccine against the dengue virus is currently under investigation. Therefore, the overcoming of DHF was performed by controlling the vector, *Aedes aegypti*. Mosquito control has been done in various ways by the Government of Indonesia. The use of insecticide continuously will cause mosquitoes to become resistant, if the resistant mosquitoes formed, the viability is greater viability, and the ability to transmit the virus is also higher, besides that with the onset of global warming will affect the density of the mosquito population. WHO [1] revealed that the most important things that should do to reduce the transmission of dengue were with vector control by monitoring the density of mosquito as a vector, because drugs and vaccines for dengue fever are still being investigated. To understand the survival of the species can be determined by tracking the genetic
variation, between individuals. There was a correlation between genetic variation and adaptation to pass life [2]. Organisms that had a high variation of genetic would be easier to pass the living and breeding then organisms with less variation [3]. A Genetic variation includes genetic structure, phylogeography, gene flow and differentiation between and interpopulation. The final result of the analysis could be determined with a phylogenetic tree, genetic distance, population structure and gene flow of inter and intrapopulation. The population of *Ae. aegypti* was investigated from different geographical regions will also differ in their susceptibility to the competence of the vector and dengue virus [4, 5]. The large genetic variation could improve survival (adaptation), which depends on the number of individuals, the range of geographical spreading, isolation degree of population and breeding system.

Genetic variation could be studied by using several markers such as allozyme, RAPD, RFLP and DNA microsatellite [6]. These markers have been used to evaluate the genetic structure of flies, deployment, migration, and relations with the breeding place of *Glossina*. Previous studies isolated a set of *Ae. aegypti* microsatellite to describe genetic differences between populations in Vietnam and other countries in Southeast Asia [7-10]. Eleven microsatellite loci have been identified to understand the population structure and gene flow in *Ae. aegypti* derived from Trinidad and could facilitate the efforts in understanding disease transmission [11].

Microsatellite has been used as markers in genetic studies of *Ae. aegypti* as the vector of dengue in Vietnam and Cambodia [12], Australia [13], and Thailand [14]. Paupy et al. [15] combined morphological data and microsatellites DNA to explain the morphological and genetic variability *Ae. aegypti* in Niakhar, Senegal. Microsatellite markers could analyze significantly genetic structure among geographic populations of *Ae. aegypti* in Bolivia, and potentially affected the epidemiology of dengue fever and yellow fever [16].

Study of genetic variation and structure population of *Ae. aegypti* as the vector of dengue fever based on DNA microsatellite in West Sumatra has not provided the data. The purpose of this study was to analyze genetic variation and genetic population structure of *Ae. aegypti* in some endemic areas of dengue fever in West Sumatra.

## 2 Materials and methods

### 2.1 Study Population

*Ae. aegypti* was collected from four locations in West Sumatra (Padang, Bukittinggi, Lubuk Sikaping, and Payakumbuh). Four locations have been associated with cases of dengue epidemiological relevance. Larvae or pupae were collected from breeding places at each location. Larvae and pupae reared to the adult stage for identification by using a key identification [17, 18]. Adult mosquitoes were stored at -20°C for molecular analysis.

### 2.2 DNA amplification

Genomic DNA was extracted from mosquitoes by using CTAB protocol [3] with a few modifications by Anggriani [19]. The results of DNA isolation was confirmed with 1.2% agarose gel for electrophoresis process in 0.5 TBE buffer [20]. The quantity DNA concentration was calculated with λ DNA and a NanoDrop. DNA purity was counted based on a division between absorbance score of 260 and 280. The result of division ranged from 1.7 to 2.0 was
considered excellent [21]. DNA was stored at -20°C and amplified by using markers/microsatellite primers from previous result [22].

PCR profile consisted of initial denaturation 94°C for 5 min followed by 33 cycles of denaturation 92°C for 30 seconds, 80 seconds Annealing 59°C, polymerization 72°C for 90 seconds and the final polymerization 72°C for 5 min, the total of PCR process was 35 cycles. PCR products were continued for PAGE analysis [23].

2.3 Data analysis

The lengths of amplified DNA fragments were determined based on the distance of migration after the data were transformed into a linear regression equation of the relationship between migration distances of DNA band with the size of DNA marker. The existence of fragments (band) was determined with a computerized Image-J program by using differentiation of black color density measurement.

The quantification results were analyzed with POPGEN program version 3.2 to determine allele frequency, polymorphic locus percentage (PP), the average of alleles per locus (Na) and expected Heterozygosity (He). Genetic variation between populations was investigated from F-statistic (FST), FIS is coefficient inbreeding among individuals within populations, and FIT is coefficient inbreeding in the population total [24]. Genetic distance and similarity of dendrogram based on genetic distance Nei's matrix [25] and the score of gene flow (Nm) were analyzed by using POPGEN version 3.2. [26]. Gene flow (Nm) score was calculated by using Nm formula = 0.5 (1-Gst / Gst). Nm > 1 (gene flow neutralize genetic drift within the population), Nm < 1 (the genetic drift is a major factor in forming the population structure) [27]. Analysis of Mantel test is used to determine the correlation of genetic distance with geographical distance using XLSTAT software.

3 Results

3.1 Genetic variation of *Aedes aegypti* population

Analysis results of *Ae. aegypti* genetic variation in each population was performed by investigating the following parameters: Average number of observed alleles (Na), average number of effective alleles (Ne), expected Heterozygosity (He), Pp (Percentage of polymorphic loci) [28] and summarized in Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Population</th>
<th>N</th>
<th>Na</th>
<th>Ne</th>
<th>He</th>
<th>I</th>
<th>PLP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bukittinggi</td>
<td>6</td>
<td>6,5000</td>
<td>5,2867</td>
<td>0.8404</td>
<td>1,7263</td>
<td>100,00</td>
</tr>
<tr>
<td>2</td>
<td>Lb. Sikaping</td>
<td>6</td>
<td>6,5000</td>
<td>5,0296</td>
<td>0.8044</td>
<td>1,6587</td>
<td>100,00</td>
</tr>
<tr>
<td>3</td>
<td>Payakumbuh</td>
<td>6</td>
<td>5,5000</td>
<td>4,1132</td>
<td>0.7886</td>
<td>1,5215</td>
<td>100,00</td>
</tr>
<tr>
<td>4</td>
<td>Padang</td>
<td>6</td>
<td>5,8335</td>
<td>4,6362</td>
<td>0.7939</td>
<td>1,5716</td>
<td>100,00</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>6</td>
<td>6,0833</td>
<td>4,7664</td>
<td>0.5213</td>
<td>0,7591</td>
<td>100,00</td>
</tr>
</tbody>
</table>

The percentage score of polymorphic loci was 100%, found in all four populations. The percentage of polymorphism and the number of polymorphic loci were affected by population size and sample size and the presence of monomorphic loci.
The highest score of \( He \) and \( I \) in \textit{Ae. aegypti} populations were obtained in Bukittinggi (0.8404, 1.7263, respectively), Lubuk Sikaping population (He: 0.8044 and I: 1.6587), and Payakumbuh population (He: 0.7886 and I: 1.5215), and the lowest score was the population in Padang (He: 0.7939 and I: 1.5716). Based on the analysis of DNA fragments, the results of PCR amplification, the percentage of polymorphic loci, the score of \( He \), and \( I \), the result showed the high genetic variation of \textit{Ae. aegypti} in all populations, in line with the high cases of dengue in all these areas.

A score of F-statistics (Wright fixation index) is used to determine the genetic differences between populations based on the amplification of six microsatellite loci. The average score of FIS: 0.6959; FIT: 0.7323 and FST: 0.1205; respectively. The F-statistic score of four \textit{Ae. aegypti} populations were presented in Table 2.

<table>
<thead>
<tr>
<th>Number of samples</th>
<th>( F_{ST} )</th>
<th>( F_{IS} )</th>
<th>( F_{IT} )</th>
<th>( N_{M} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>0.1205</td>
<td>0.6956</td>
<td>0.7323</td>
<td>1.8240</td>
</tr>
</tbody>
</table>

Description:
- \( F_{ST} \): genetic differentiation between populations
- \( F_{IT} \): coefficient of inbreeding in the population total
- \( F_{IS} \): coefficient of inbreeding among individuals within populations
- \( NM \): The flow of genes (gene flow)

Scores of Wright fixation index (F-statistics) are used to determine the genetic differences between populations. \( F_{ST} \) score is a parameter describing the genetic diversity (genetic differentiation) between populations. The score of \( F_{ST} \) has a range between 0 and 1. \( F_{ST} \) close to 0 indicates no genetic differences between populations, \( F_{ST} \) close to 1 indicates high genetic differences between populations. \( F_{ST} \) scores between 0 and 0.05 indicated low genetic diversity, scores between 0.05 and 0.15 indicated a medium genetic difference, the score of 0.15 and 0.25 indicated high genetic differences, and the score of \( F_{ST} > 0.25 \) indicated a very high genetic diversity [29].

3.2. Genetic structure of \textit{Aedes aegypti} population

The genetic structure of \textit{Ae. aegypti} populations in four populations in West Sumatra was determined with Unweighted Pair-Group Method with arithmetic mean (UPGMA) to result in the dendrogram [30]. The analysis was proceeded by using MEGA version 4 [31]. The appearance of \textit{Ae. aegypti} grouping patterns among populations in West Sumatra demonstrated in Figure 1.

Fig.1. Dendrogram of eight \textit{Aedes aegypti} population as DHF vector in West Sumatra
The score of genetic distance among the four populations was presented in Table 3:

**Table 3. Matrix genetic distance in four populations based on microsatellite DNA markers**

<table>
<thead>
<tr>
<th>No.</th>
<th>Population</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bukittinggi</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>L. Sikaping</td>
<td>0.6787</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Payakumbuh</td>
<td>0.6925</td>
<td>1.2558</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Padang</td>
<td>0.9239</td>
<td>0.7338</td>
<td>0.7564</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

The scores of the farthest genetic distance were between Lubuk Sikaping, and Payakumbuh population (1.2558) and the lowest score of genetic distance between populations was found in Bukittinggi and Lubuk Sikaping (0.6787).

3.3 Correlation of geographical distance with the genetic distance

![Scatter diagram](image)

**Fig.2.** Scatter diagram showed a correlation between genetic distance and geographical distance by microsatellite DNA markers, $r =$ correlation score, $p =$ probability

Based on the result of Mantel test, there was a weak correlation between genetic distance and geographical distance ($r = -0.007$ and $p = 0.943$, $p > 0.05$ indicated no correlation between these two factors). Wynn and Heyer [32] stated that the geographical distance and genetic distance indicated a correlation if the genetic distance scores $< 0.15$. The highest score of genetic distance obtained in this study was 1.2558 (between the population of Lubuk Sikaping and Payakumbuh), and the lowest score of genetic distance was between populations Bukittinggi and Lubuk Sikaping.
4 Discussion

Heterozygosity (He) score is an accurate way to measure the genetic variation [33] and the Shannon's Information Index score is used to determine the genetic variation in the population if the locus is observed more than one locus [34]. The higher score of heterozygosity the higher genetic variation in the population [35].

The size of the population could influence the high score of He in Lubuk Sikaping and Bukittinggi population. Large populations tend the occurrence of random mating. Inbreeding less occurred in a large population and resulted in elevation of heterozygosity [36]. The high score of heterozygosity generated individuals in a population was able to adapt to changes in an environment [37]. It was evident that *Ae. aegypti* were originally only found in low land areas, and theoretically only live in high-temperature area, but currently the mosquitoes have adapted to highland areas so that the spread of dengue fever were widespread, 18 of 19 cities/counties reported were DHF endemic areas [38].

The lower heterozygosity expected score (He) (higher than the 0.7) found in the population of Payakumbuh and Padang. Underscoring of He was influenced by the size of the population. A small population might be increasing the chances of individuals in the population to inbreeding occurrence [39]. Inbreeding caused low genetic variation in a population and reduced the score of heterozygosity as closely associated with the loss of some alleles and low levels of polymorphism, resulted in increased inbreeding depression. Inbreeding depression is a cross-mating between individuals derived from one parental, finally degraded the quality of life. The lower quality of life of the *Ae. aegypti* mosquito the short period of it life, and directly decreased the competence (ability to transmit) of the vector.

Genetic variation of *Ae. aegypti* in all four examined populations (Lubuk Sikaping and Bukittinggi, Payakumbuh and Padang) was high because He scores> 0.7. Hale et al. [40], genetic variation of a population was higher if He score was > 0.7. The high score of He could be determined with varying of microsatellite DNA fragments with some specific fragments. The results of the current study were proved with a high score of heterozygosity in Bukittinggi, Payakumbuh, Lubuk Sikaping, and Padang population showed a high number of DHF cases, 107 cases in Bukittinggi and 1,626 cases in.

The results of this study obtained the score of genetic diversity (genetic differentiation) was medium ($F_{ST} = 0.1205$), described that the score of genetic differentiation *Ae. aegypti* populations between several cities in West Sumatra showed medium genetic differences. $F_{ST}$ score indicated that 12.05% of the total population was between the population and 87.95% are in the population. These results were illustrated by the variety of specific alleles found in the population. The medium level in this study was associated with the presence of specific alleles that were characteristic of each population.

Based on $F_{ST}$ score, it could be determining that; genetic variation among populations was medium compared with genetic variation in the population. Moderate genetic variation was caused by individuals carried through the mobility of transportation was easy to move from one population to other populations, so that random mating took place. Genetic differentiation scores of *Ae. aegypti* affected behavior change and vector competence.

In current study, gene flow score was 1.8240 ($N_M > 0.5$). Larson [41] revealed the score of gene flow ($N_M$) indicated the level of migration between populations. The high score of $N_M$ caused a large distribution forecast. Meanwhile, $N_M <1$ generated genetic drift as a major factor shaping the population. Geographical distance of species was affected by the ability of species to spread. Widely distributed species were more affected by high gene flow.
Geographic distance between all four populations of *Ae. aegypti* has no effect on the occurrence of mating among populations with high gene flow. Although *Ae. aegypti* has a short flight distance, Reiter *et al.* [42] stated that a limited flight range of *Ae. aegypti* was 100-800 m, abundant spawning grounds [43], the mating of *Ae. aegypti* occurred clustered around the host in or around the home [44] but a human has facilitated the spread through the displacement of objects containing water with eggs, such as tires and through public transportation. The high gene flow between populations led to a high genetic differentiation in populations. It was due to the occurrence of random mating so heterozygosity increased. Some studies related to genetic differentiation of *Ae. aegypti* in Vietnam and Cambodia [45], Arizona [46], and Pakistan [47] revealed that transportation affected on genetic variation of *Ae. aegypti*.

F<sub>IS</sub> score determined the number of inbreeding by individuals in a population, while the score of F<sub>IT</sub> determined the overall level of inbreeding that occurred in the total population [48]. This study resulted 0.6956 of F<sub>IS</sub> 0.7323 of F<sub>IS</sub>, showed that random mating occurred in a population. F<sub>IS</sub> < 0 showed inbreeding occurred within populations and vice versa F<sub>IS</sub> > 0 meant random mating occurred. In a large population, random mating might not lead to the possibility of cross-breeding between close relatives so that affected the high score of F<sub>IT</sub>. Based on F<sub>IS</sub> and F<sub>IT</sub> score, the size of *Ae. aegypti* population in West Sumatra was relatively large so that the dengue virus transmission by *Ae. aegypti* always high. Thus, the vector (mosquito) control should get serious attention and should be carried out continuously. If the frequency of alleles in the two populations were equal, then the genetic similarity was 1 (one), and the genetic distance was 0 (zero). Conversely, if two populations did not have the same allele, then the score of genetic similarity was 0 (zero), and genetic distance was infinite [49].

Based on UPGMA cluster analysis in Figure 1, the grouping of *Aedes aegypti* in four city/districts in West Sumatra divided into two main groups; the first group was Bukittinggi and Lubuk Sikapin, the second group was Payakumbuh and Padang.

The result of the grouping illustrated that each group showed a close genetic relationship based on the characteristics of alleles in each population. In the first group, Lubuk Sikaping and Bukittinggi had a similar characteristic of alleles. It was due to the spread of mosquitoes occurred through transportation and things carried the eggs of mosquitoes, such as tires. The tire trade between cities led to the spread of mosquitoes faster in Pakistan [47]. The possibility might happen in both populations in this study, the mobility of transportation among the four regions was very high, *Ae. aegypti* freely moved and carried away by objects contained water, such as containers carrying fish across the city.

Genetic distance and geographic distance between populations could be used to determine the population structure and kinship in the population. Genetic distance was the difference between genetic structures of the population as measured by a numerical score, which was useful to investigate the kinship between populations [50]. Genetic distance was also used in estimating the time of separation between populations and was also used to construct phylogenetic trees [51]. The smaller the genetic distance scores, the closer the existence of the genetic relationship.

The genetic distance of *Ae. aegypti* in four cities in West Sumatra was high, similar to the previous study [52]. Genetic distance was high if the score was > 0.5. The lowest score of genetic distance of the current study was 0.6787 so that the genetic distance of *Ae. aegypti* in all four cities in West Sumatra was high, it described that the genetic variations were high between populations of a species, the ability to adapt was also high. An organism that was able to withstand the pressure of the environment was able to adapt, to increase the population.

High score of genetic differentiation between populations indicated the presence of gene flow between populations. Although *Ae. aegypti* flight distance was only 100-800 m. Spread occurred using transport (passive dissemination) so that the gene flow between populations...
always occurred caused a high score of genetic distance and genetic differentiation. If in the long term there was no gene flow between populations, and the genetic score might vary [53].

Similar with the previous study [54] reported that there was a negative correlation between genetic distance and geographical distance, a sample of *Ae. aegypti* was collected in the rainy season and the dry season in Mannus, Brazil. Result of Mantel test was $r = 0.423$, $p = 0.166$ (rainy season), and $r = 0.545$, $p = 0.236$ (dry season).

The lack of correlation between genetic distance and geographical distance of *Ae. aegypti* in this study was shown, several populations had a close geographical distance but with a large genetic distance and vice versa. It was due to both populations have not been separated by geography and gene flow between populations was high, because of the spread of mosquitoes was occurred through high mobility of transportation occurred every time.

5 Conclusion

A genetic variation of *Ae. aegypti* in the four areas in West Sumatra was high with medium genetic differentiation. Genetic structure of population *Ae. aegypti* was divided into two groups (Bukittinggi and Lubuk Sikaping; Payakumbuh and Padang) based on dendrogram diagram. Genetic distance and geographic distance showed great potential in the spread of large distance but it did not relate to the inhibition of *Ae. aegypti* spreading.

References


