Communication

β Thalassemia Mutation Flow in Indonesia: A Migration Perspective

Lantip Rujito 1,* , Ziske Maritska 2 and Abdul Salam Sofro 3

1 Faculty of Medicine, Universitas Jenderal Soedirman, Jl Gumbreg 1, Purwokerto 53112, Indonesia
2 Faculty of Medicine, Universitas Sriwijaya, Jl. Sriwijaya Negara, Bukit Lama, Palembang 30139, Indonesia; ziske.maritska@unsri.ac.id
3 Faculty of Medicine, Universitas YARSI, Jl. Letjen Suprapto No. Kav. 13, Jakarta 10510, Indonesia; abdul.salam@yarsi.ac.id
* Correspondence: lrujito@unsoed.ac.id

Abstract: Indonesia is a large island country with a wide variety of ethnic groups. As part of the thalassemia country belt, Indonesia has alleles that are as distinctive as those found in other parts of Southeast Asia. The journey of ancestors in the prehistoric period and the massive increase in human exchange in the last decade have formed the current population of Indonesia. The mutants of the beta-thalassemia allele brought by those predecessors can be seen from the traces of their journey. This paper describes the flow gene according to the type of mutations of beta-thalassemia in the country.

Keywords: beta-thalassemia allele; Indonesia; gene flow; mutant allele; migration; Southeast Asia

1. Introduction

Thalassemia encompasses a group of inherited blood disorders characterized by abnormal hemoglobin production, leading to reduced or defective red blood cells. Affecting millions globally, thalassemia presents significant health challenges across various regions. Understanding its epidemiology is essential for devising effective prevention strategies, enhancing patient care, and optimizing healthcare resource allocation [1]. Thalassemia exhibits a worldwide distribution with differing prevalence levels across populations. Regions historically impacted by malaria such as the Mediterranean region, Southeast Asia, Middle East, and parts of Africa show higher rates [2]. Per data from the World Health Organization, genetic thalassemia traits exist in around 7% of the global population. Of this carrier pool, 300,000 to 400,000 infants are born annually with severe forms of thalassemia. However, the prevalence rates vary significantly among countries and even within different regions of the same country.

Thalassemia encompasses numerous subtypes, stemming from mutations in various globin genes [3]. The two principal categories are alpha and beta thalassemia, involving the alpha or beta globin genes, respectively. Alpha thalassemia largely emerges in groups of African, Southeast Asian, or Mediterranean origin and is tied to anomalies in the alpha globin gene. Conversely, beta thalassemia predominately arises in those with ancestral ties to the Mediterranean region, Middle East, South Asia, or East Asia, owing to genetic changes affecting the beta globin gene [4].

The Indonesian population, on the other hand, is incorporated into the Austronesian sphere, based on the evidence of historical relics in the form of agricultural techniques, culture, and language [5]. The migration of the Austronesian peoples is believed to have originated in the Formosa Islands (Taiwan), which later spread through two main routes. The eastern route begins from the movement to the Philippines, then descends to Sulawesi, and continues to spread to Java and the eastern Indonesian island to further continue to Micronesia [6,7]. This trace might be confirmed by genomic findings with the discovery...
of β thalassemia alleles that followed this wave of migration. The western islands of Indonesia, including Sumatra and continuing eastward to Java and Borneo, have a higher prevalence of beta-thalassemia mutations originating from the Indian subcontinent and Thailand, likely due to historical migrations. This article is intended to provide input or insight on whether the indentation and proportion of the scattered beta thalassemia alleles correspond to the theory of the journey of the ancestors. The study also gives a view on how these data may contribute to a prevention program of thalassemia in Indonesia.

2. Materials and Methods

This perspective article was written using a mini-review approach. Research was obtained from various online sources, including ScienceDirect, PubMed, and Google Scholar. The search criteria included primary data or original articles from cross-sectional, retrospective, or cohort data, published until 2022, ensuring the data’s relevance and timeliness. Keywords used to search for sources included Indonesia beta thalassemia, Southeast Asia thalassemia, Asian beta thalassemia allele, Malaysia beta thalassemia, Singapore beta thalassemia, Philippines thalassemia, Japan thalassemia, China beta thalassemia, Asian migration, and Indonesia road migration. Reviews, case reports, articles focusing solely on clinical treatment or narrow geographic data, and duplicate publications were excluded to retain recent, original genetic and epidemiologic research that is generalizable to the thalassemia burden across Indonesia. After the sorting process, a total of 37 articles were collected for the construction of the article, as presented in the following chapter.

3. Results and Discussions

3.1. β Thalassemia Alleles

Several studies conducted in Indonesia have examined the frequency of the IVS-1-5 (G > C) thalassemia mutation within the population. The IVS-1-5 allele has been found to have the highest occurrence rates in these investigations. Specifically, prevalences of 52%, 35.5%, and 54% were reported across three separate studies for this mutation, indicating that it is the most widespread thalassemia variant identified in Indonesia thus far [8–10]. However, some other studies have presented alternative data indicating the Cd26 allele (HbE mutation) to be the most prevalent thalassemia mutation within the studied Indonesian populations instead of IVS-1-5. Specifically, Cd26 was found at frequencies of 37% in one analysis and 47% in another study [9,11]. This mutation frequency variation may stem from divergent study locations and sample sizes. The first higher IVS-1-5 rates reflected localized, homogeneous groups, while the latter analyses with elevated Cd26 utilized samples amalgamated across scattered, heterogeneous ethnicities. However, all cited studies employed limited subject numbers (<100). A larger investigation with over 200 patients found a 43.5% rate for the IVS-1-5 allele compared with 28.2% for HbE, perhaps offering a more representative prevalence given the increased sample size, although additional extensive epidemiological analyses are still warranted. The apparent differences seem partially attributable to the geographic specificity, genetic homogeneity, and small cohorts of the initial analyses, contrasting with the larger but blended pan-Indonesian examination samples evident in subsequent work [12].

The identification of these studies confirms that the IVS-1-5 (G > C) allele may follow the flow of regional distribution in southern Southeast Asia. It is mentioned in the literature that this allele population is found mainly in Malaysia, Indonesia, Brunei Darussalam, and the Philippines [13]. In contrast, mainland China harbors predominantly Cd41/42 and IVS2 654 thalassemia alleles [14], while Cd26 (HbE) characterizes northern regions like Thailand, Laos, Cambodia, and Myanmar [13,15,16]. However, population migrations and genetic admixture between groups can alter localized allele frequencies over time. The IVS-1-5 anomaly exhibits a higher concentration in island/archipelagic nations south of mainland China, potentially reflecting localized ancestral spread. But human movements and intermarriage introduce variability, shifting mutations between populations fluidly despite overarching geographical concentration trends.
Recent data reveal several novel thalassemia mutations that have been identified within the Javanese population. The alleles are IVS-1-2 (T > C), Cd40 (-G), CAP +1 (A > C), Cd123/124/125 (-ACCCCACC) [12], and also Cd60 (GTG > GAG) [17]. The first alleles mentioned are those found in individuals of African descent in the Americas [18]. In the case of allele IVS-1-2 (T > C), the first report on the Indonesian population was proposed in a previous study in the multiethnic population of Jakarta [9]. Cd40 (-G), as quoted from the globin server, is an allele that is identified in the regions of Japan and Korea [19], while CAP + 1 (A > C) is the beta+ allele with the most reports on ethnicity in Singapore [20].

The latter allele, Cd123/124/125 (-ACCCCACC), was a mutant allele first discovered in northern Thailand [21], while Cd60 was first reported in Cagliari, Italia. The absence of a history of blood relations or regional origin of the three generations connected to the initial discovery of the mutant indicates that the mutant found was likely a spontaneous mutation. Even so, it does not rule out the possibility of gene flow from previous ancestors.

3.2. Migration Perspective

With such complexity in its population history, it is expected that the genetic architecture of Indonesian populations is diverse. As the population movement from the Philippines happened in the east, a different but reciprocally connected event also happened in the west. The western part of Indonesia is known to possess a different genetic architecture due to the multiple interactions between Austronesian immigrants from the eastern route and the mainland Asian taking place not just once, but more often than that. Other factors adding to the complexity of the genetic admixture in western Indonesia was due to the genetic interaction with the South Asian population, and also due to particular demographic events [22].

The gene flow in Indonesia is a long process, in line with the journey of human migration in the world. Based on the evidence of historical relics in the form of agricultural techniques, culture, and language, the Indonesian population is incorporated into the Austronesian sphere. The migration of the Austronesian peoples is believed to have originated in the Formosa Islands (Taiwan), which later spread through two main routes. The eastern route begins from the movement to the Philippines, then descends to Sulawesi, and continues to spread to Java and the eastern Indonesian island to further continue to Micronesia and Polynesia [6]. This trace was confirmed by genomic findings with the discovery of β thalassemia alleles that followed this wave of migration. The Filipino deletion allele (β0), which is widely found in the Philippine population, is also found in the people of Borneo [23], Sulawesi, and in small numbers also found in the population of Jakarta [24]. Alleles Cd123/124/125 (-ACCCCACC) and CAP +1 (A > C), which were first found in northern Thailand and Singapore [25], appear to follow the population migration to Indonesia via the western route.

The Cd123/124/125 (-ACCCCACC) mutation found in the Javanese population [12] is a minor allele in the population of northern Thailand. After it was first discovered, sporadic findings were reported by [26] among residents of southern Thailand in the Bangkok area and [23] among the population of the Sabah area. In another section, current epidemiological data suggest that the Cap +1 (A > C) allele was found scattered on the Indian plains [27], Pakistan [28], China, Thailand, and Malaysia [29,30]. The findings of the Cap +1 (A > C) allele in Indonesia [12] clarifies the path of population distribution through the western route according to the Austronesian migration theory. A previous population migration path has also been known from the many findings of HbE alleles (codon 26) and IVS-1.5 in South and Southeast Asia, with the highest dominance in Thailand [31], then successively in coastal areas of Burma to India, as well as in Malaysia and Indonesia [13]. Recent publications using mitochondrial DNA haplogroups for maternal pathways and Y chromosomes for paternal pathways show that Indonesian populations have a genomic constitution that is similar to what is found in Austronesian migration pathways [32]. Their report showed data demonstrating that at least one haplotype is shared between Indonesia and Taiwan for each of the following four lineages: E1a1a, M7b3, M7c3c, and Y2.
The codon allele 40 (-G) is rare and was first characterized in Korean and Japanese families. The limited literature on this mutation means that the connectivity of the findings of the codon 40 (-G) allele in Japan and Java cannot be clearly stated. The theory of ancestral migration does not definitively describe the relationship. But in a report, it is stated that the ancestors of the Japanese and Koreans were ethnic groups from northern China and the Siberian region at a time when the Asian mainland was still connected. Subsequent migration developments allowed for the movement of inhabitants between islands in the eastern regions due to weather and climate [33,34]. Another possible theory is gene flow occurring during colonialism or modern warfare, based on the evidence that there were many civilian casualties and significant violence against women during the occupation.

In addition to the theory of migration out of Taiwan (Figure 1), the migration of Austronesian populations according to the Sundaland theory can also illustrate the linkage of Thalassaeemia alleles $\beta$, especially in Southeast Asia (Figure 2). The unity of the Asian mainland at the time of the prehistoric period gave rise to another theory. Alleles may have originated in the plains of Indonesia. The Human Genome Organization Pan-Asian SNP Consortium’s report states that humans originally inhabited Asia via the southern route. Migration first came from Africa through India to Southeast Asia, then to the Pacific region, and only then did it spread to the east and north of the mainland. At the end of the ice age, approximately 8000–15,000 years ago, the inhabitants of Sundaland began to migrate due to climate change and major floods [35,36]. Residents of the western Sundaland area moved along the coastal areas of the mainland that eventually became the islands of Indonesia and the central plains of western and eastern Asia. Other genetic study findings state that paternalistic DNA markers (Haplogroup O-M119, O-M95, O-P203, and O-M122) in the Indonesian area are also found in residents of Taiwan, Malaysia, the Philippines, Tonga, Tahiti, and Hawaii, and in Māori people [37]. In another section, previous studies have found that the type of carbonic anhydrase enzyme CaII, present in the Indonesian and Aeta populations, is also found in the inhabitants who inhabit Luzon Island in the Philippines [38]. One of the mitochondrial DNA (mtDNA) studies on the Asian population including Indonesia proposed that the genetic diversity in Indonesia was determined by the population movements, either because of the expansion of farming populations from the Asian mainland to the Southeast Asia regions or due to the changes in sea level.

The $\beta$ Cd26 HbE and IVS-1-5 that are prominent in the western Southeast Asian strip regions (Indonesia, Malaysia, Thailand, Burma, and Laos), as well as the Filipino deletion alleles in the eastern regions, can provide an understanding of the occurrence of this migration (Figure 2). Table 1 shows the alleles and regional pattern frequencies among studies.

Table 1. Major allele frequencies among different parts of regions.

<table>
<thead>
<tr>
<th>Western Part Alleles (%)</th>
<th>Study</th>
<th>Eastern Part Alleles (%)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVS1-5 (35–54)</td>
<td>[9,12,39]</td>
<td>Fil $\beta$-deletion (45–90)</td>
<td>[23,24,40]</td>
</tr>
<tr>
<td>HbE (28–47)</td>
<td>[11,12,41,42]</td>
<td>IVS1-5 (35–54)</td>
<td>[9–11]</td>
</tr>
<tr>
<td>Codon 35 (-C) (5–17)</td>
<td>[12,45]</td>
<td>HbE (28–47)</td>
<td>[8,11]</td>
</tr>
<tr>
<td>Codon 15 (TGG &gt; TAG) (3–6)</td>
<td>[8,12]</td>
<td>Codon 41/42 (-TTCT) (5–15)</td>
<td>[39]</td>
</tr>
<tr>
<td>IVS-I-1 (G &gt; T) (2–5)</td>
<td>[39]</td>
<td>IVS-II-654 (C &gt; T) (–10)</td>
<td>[9]</td>
</tr>
</tbody>
</table>

Furthermore, major developments in parts of the northern plains can occur due to genetic drift mechanisms, where the migration of these groups then forms individual colonization, which ultimately increases the frequency of certain alleles. Meanwhile, the original alleles in the area of origin left behind remnants due to various factors, including natural disasters, outbreaks, and the subsequent wave of migration that brought new types of alleles. This dilution can be seen from a large number of alleles in a small percentage that is found in various previous genetic epidemiological studies such as IVS-1-1 (G > A), IVS-1-1(G > T), Cd41/42 (-TTCT), Cd19 (AAC > AGC), IVS-1-2 (T > C), and Cd17 (AAG > TAG) [8,9,44].
regions due to weather and climate [33,34]. Another possible theory is gene migration does not definitively describe the relationship. But in a report, it is stated that the migration developments allowed for the movement of inhabitants between islands in the eastern Siberian region at a time when the Asian mainland was still connected. Subsequent migration first came from Africa through India to Southeast Asia, then to the Pacific region, and only subsequently in the plains of Indonesia. The Human Genome Organization Pan-Asian SNP Consortium at the time of the prehistoric period gave rise to another theory. Alleles may have originated in the plains of western and eastern Asia. Other genetic study findings state that paternalistic DNA markers (Haplogroup O-M119, O-M95, O-P203, and O-M122) in the Indonesian area are also found in residents of Taiwan, Malaysia, the Philippines, Tonga, Tahiti, and Hawaii, and in coastal areas of the mainland that eventually became the islands of Indonesia and the central plains of Indonesia. The limited literature on this mutation means that the connectivity of the findings of the codon 40 (-G) allele in Japan and Java cannot be clearly stated. The theory of ancestral proteronesian populations according to the Sundaland theory can also illustrate the linkage of genetic diversity in Indonesia was determined by the population movements, either because of the inhabitants who inhabit Luzon Island in the Philippines [38]. One of the mitochondrial genotypes of the Thalassaemia alleles to become more intense. This massive population migration is recorded to spread in the eastern part.

Figure 1. Map of Indonesian population migration pathways with Thalassaemia alleles $\beta$ in Indonesia. Filipino deletion alleles in the Philippines are found on Sulawesi, Borneo, and Java islands. Codon 26 in Indonesia is seen following the path of western migration, starting from all parts of southern Asia to Thailand and continuing to Malaysia and Indonesia. The rare alleles Cap + 1 and Codon 123/124/125 are seen sporadically in the western part of the migration path.

The increasingly extensive modern migration process causes the mixing of Thalassaemia alleles to become more intense. This massive population migration is recorded to impact the discovery of alleles far from the origin of the alleles. The development of Thalassaemia in the European and American continents and East Asian populations increased rapidly by bringing alleles of different origins from migrants [45]. In more recent times, the genetic landscape of Indonesia has continued to evolve due to various migrations and popula-
lation movements. For example, the Polynesian motif, associated with Taiwanese dispersal, presents an unusual distribution, suggesting complex migration patterns. These modern movements have led to population growth and expansion, particularly in central Indonesia, while unique genetic patterns emerge in barrier islands like Nias and Mentawai [32]. Such modern genetic shifts have implications for the distribution and prevalence of thalassemia alleles. In addition, developing current and future health services will indirectly help Thalassaemia alleles to spread more widely [46].

3.3. \( \beta \) Thalassemia Alleles and Prevention and Control Program

This report has identified specific ethnic/geographic regions in Indonesia with higher prevalence rates of \( \beta \) gene mutations, comparing western and eastern regions. These regions are characterized by distinct genetic, cultural, and environmental factors that may contribute to the increased incidence of \( \beta \) thalassemia. The population of western Indonesia has higher percentages of an Austronesian/Malay genetic makeup, while eastern Indonesians have higher percentages of Melanesian genetics [47]. Western Indonesians are more heavily influenced by trade/cultures from mainland Asia, while the eastern culture reflects more isolated island developments. Western islands adopted Islam earlier and more uniformly, while areas of eastern Indonesia have higher percentages of Christians or traditional animist beliefs. There are over 300 different ethnic groups and languages in Indonesia—western groups like Malay, Javanese, and Sundanese are among the largest, while there is more ethnic diversity and localization in the east [48].

The identification and mapping of \( \beta \) thalassemia mutation alleles across different ethnic/geographic regions of Indonesia provides critical insights for targeted thalassemia prevention and control strategies. These genetics data allow for a more nuanced approach to screening and intervention. They enable healthcare providers to focus resources on regions with higher mutation frequencies, thereby improving the efficiency of screening programs. In western regions, where HbE, IVS1-5, Cd35, and Cd15 allele mutations may be more prevalent, genetic screening among these mutants must be a first-line strategy and likewise in the eastern part with Fil \( \beta \)0-deletion, IVS1-5, HbE, and Cd 41/42 (-TTCT). Targeted awareness campaigns aimed at both medical practitioners and society, together with genetic counselling, can be prioritized. Such region-specific strategies, informed by the genetic landscape of the population, are crucial in not only identifying carriers but also in educating at-risk communities about the importance of prenatal screening and genetic testing. This approach fosters a culturally sensitive framework for healthcare interventions, respecting the diverse ethnic and cultural tapestry of Indonesia, and ultimately contributes to more effective management and reduction in \( \beta \) thalassemia’s impact on the country.

4. Limitation of the Study

An important aspect of this report was the analysis of data derived from studies with varying sample sizes. While we acknowledge the strength of large-scale studies in providing more generalizable results, the inclusion of smaller-scale studies in our analysis was pivotal in highlighting specific trends and nuances related to Thalassemia in different subpopulations. To mitigate the limitations posed by smaller sample sizes, the study incorporated larger-scale studies to provide a more comprehensive overview. This balanced approach allowed us to draw more robust conclusions, acknowledging the limitations while leveraging the specific insights provided by smaller studies.

5. Conclusions

The population of \( \beta \)-thalassemia alleles in Indonesia today is presenting the result of a long journey by its ancestors and the mixing of genes due to intermarriage between nations. The openness of cross-border traffic allows for the mixing of alleles, resulting in an increasingly diverse range of mutations in Indonesia.
Author Contributions: L.R. is the main author who initiated the idea, found resources, and developed the article draft. Z.M. provided some discussion and added some comments. A.S.S. polished and supervised the article development process. All authors have read and agreed to the published version of the manuscript.

Funding: There was no funding for developing this perspective article.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors report no conflict of interest.

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.