MOLECULAR AND MORPHOLOGICAL ASSESSMENTS OF SPOTTED SARDINELLA, *Amblygaster sirm* (Walbaum, 1792) IN MALAYSIAN MARINE WATERS

NOORUL AZLIANA BINTI JAMALUDIN

UNIVERSITI SAINS MALAYSIA

2025

MOLECULAR AND MORPHOLOGICAL ASSESSMENTS OF SPOTTED SARDINELLA, *Amblygaster sirm* (Walbaum, 1792) IN MALAYSIAN MARINE WATERS

by

NOORUL AZLIANA BINTI JAMALUDIN

Thesis submitted in fulfillment of the requirements

for the degree of

Doctor of Philosophy

April 2025

ACKNOWLEDGEMENT

First and foremost, I would like to express my profound gratitude to my main supervisor, Dr Noor Adelyna Mohammed Akib, whose expertise, guidance, and unwavering support have been the cornerstone of my research journey. Her deep knowledge in the field of molecular study and her ability to challenge my thinking have pushed me to exceed my own expectations. I am equally indebted to my co-supervisors, Prof Dr Siti Azizah and Dr. Sebastien Lavoué, whose complementary expertise has greatly enriched this work with innovative approaches to molecular and morphology opened new avenues in my study that I had not previously considered.

I am deeply grateful to the Ministry of Higher Education for their funding, Fundamental Research Grant Scheme (FRGS/1/2020/STG01/USM/02/1) for study entitled "Unravelling phylogeographic pattern of spotted sardinella, *Amblygaster sirm* in Southeast Asia: Building" global sustainable fishery which made this research possible.

Finally, I'm particularly grateful to my family, close friends and my cats (Yoonji and Jimin) for their assistance and motivations during my PhD journey. Their love and supports has made this journey not just bearable, but truly enjoyable and unforgettable. Thank you for being there through every high and low, for celebrating each milestone, and for making this experience so much richer than I could have imagined. Any errors or oversights remain my own.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF PLATES

Page

LIST OF ABBREVIATIONS

mt	Metric ton
DOFM	Department of Fisheries, Malaysia
SEAFDEC	Southeast Asia Fisheries Development Centre
bp	basepair
DNA	Deoxyribonucleic acid
µL	Micro-litre
PCR	Polymerase Chain Reaction
EtOH	Ethanol
ddH_2O	distilled water
mtDNA	Mitochondria DNA
nDNA	Nuclear DNA
Cyt *b*	Cytochrome *b*
COI	Cytochrome Oxidase Subunit I
NJ	Neighbour Joining
ML	Maximum Likelihood
BI	Bayesian Inference
K2P	Kimura 2 P
MSN	Minimum Spanning Network
AMOVA	Analysis of Molecular Variance
SAMOVA	Spatial Analysis of Molecular Variance
k	Number of groups
kya	Thousand years ago

LIST OF APPENDICES

PENILAIAN MOLEKUL DAN MORFOLOGI IKAN TAMBAN BELURU, *Amblygaster sirm* (Walbaum, 1792) DI PERAIRAN MARIN MALAYSIA

ABSTRAK

Kajian ini menangani masalah kritikal penurunan pendaratan ketara ikan *Amblygaster sirm* (ikan tamban beluru), di perairan Asia Tenggara, terutamanya di Indonesia dan Malaysia, di mana tangkapan telah menurun sehingga 75% di beberapa kawasan. Walaupun mempunyai kepentingan ekonomi, spesies ini mempunyai kekurangan penilaian genetik yang komprehensif yang diperlukan untuk pengurusan berkesan dan menunjukkan keperluan segera untuk menentukan sama ada ia terdiri daripada satu populasi panmiktik atau beberapa stok berbeza yang memerlukan pendekatan pemuliharaan berasingan. Penyelidikan ini bertujuan untuk menilai kepelbagaian genetik dan struktur populasi *A. sirm* di perairan Malaysia, membezakan stok genetik berasingan untuk mewujudkan unit pengurusan yang sesuai, dan membangunkan cadangan pengurusan lestari berasaskan sains yang boleh digunakan pada peringkat kebangsaan dan serantau. Analisis molekul dan morfologi komprehensif telah dijalankan di sepuluh lokasi dipilih yang mewakili Laut Andaman, Laut China Selatan, Laut Sulu, Laut Sulawesi dan Laut China Timur (kumpulan luar). Kajian ini menggunakan penanda genetik DNA mitokondria (Cyt *b* dan COI) serta DNA nuklear (gen s7 ribosom intron) dilengkapi dengan analisis morfometrik terperinci terhadap 42 spesimen. Analisis genetik mendedahkan perbezaan ketara antara Laut Andaman (Susur galur 1) dan Laut China Selatan serta laut sekitarnya (Susur galur 2), dengan penanda

COI dan Cyt *b* menunjukkan jarak genetik melebihi ambang peringkat spesies (~7%), manakala s7 intron nDNA nuklear memaparkan corak yang serupa tetapi kurang ketara. AMOVA juga mendedahkan beberapa struktur genetik dalam Susur galur 2, seterusnya menekankan dinamik populasi kompleksnya. Analisis demografi mencadangkan pengembangan populasi terkini untuk kedua-dua susur galur seperti mewakili pemulihan selepas sejarah populasi *bottleneck.* Analisis morfometrik melalui Analisis Diskriminan (DA) dengan kuat membezakan kedua-dua keturunan, dengan panjang kepala menjadi pembeza morfologi utama dari individu Keturunan 2 menunjukkan panjang kepala yang lebih panjang berbanding Keturunan 1 walaupun analisis Analisis Komponen Utama (PCA) menunjukkan terdapat pertindihan antara kedua-dua populasi ini. Penemuan ini memberikan bukti kukuh bagi kewujudan dua spesies kriptik dalam *Amblygaster sirm*, yang memerlukan pengurusan sebagai unit berasingan. Pendekatan bersepadu ini menawarkan maklumat penting untuk melaksanakan amalan perikanan lestari yang disesuaikan dengan ciri-ciri unik setiap keturunan, menyumbang kepada pemuliharaan jangka panjang sumber perikanan yang berharga ini.

MOLECULAR AND MORPHOLOGICAL ASSESSMENTS OF SPOTTED SARDINELLA, *Amblygaster sirm* (Walbaum, 1792) IN MALAYSIAN MARINE WATERS

ABSTRACT

This study addresses the critical problem of significant landing declines of spotted sardinella (*Amblygaster sirm*) in Southeast Asian waters, particularly in Indonesia and Malaysia, where catches have decreased by up to 75% in some regions. Despite its economic importance, the species lacks comprehensive genetic assessment necessary for effective management, creating an urgent need to determine whether it comprises a single panmictic population or multiple distinct stocks requiring separate conservation approaches. The research aims to evaluate genetic diversity and population structure of *A. sirm* in Malaysian waters, distinguish separate genetic stocks to establish appropriate management units, and develop science-based sustainable management recommendations applicable at both national and regional levels. Comprehensive molecular and morphological analyses were conducted across ten locations representing the Andaman Sea, South China Sea, Sulu Sea, Celebes Sea and the East China Sea (outgroup). The study employed multiple genetic markers mitochondrial (Cyt b and COI) and nuclear DNA (s7 ribosomal intron gene) complemented by detailed morphometric analyses of 42 specimens. Genetic analyses revealed a significant divergence between the Andaman Sea (Lineage 1) and South China Seas and the neighbouring seas (Lineage 2), with COI and Cyt *b* markers showing genetic distances exceeding the species-level threshold (~7%), while the nuclear s7 intron nDNA gene

displayed a similar but less distinct pattern. AMOVA also revealed some genetic structuring within Lineage 2, further highlighting its complex population dynamics. Demographic analyses suggested recent population expansions in both lineages, likely representing recovery after historical bottlenecks. Morphometric analyses through Discriminant Analysis strongly differentiated the two lineages, with head length serving as the key morphological differentiator whereby Lineage 2 individuals exhibited significantly longer head lengths than Lineage 1 although Principal Component Analysis (PCA) analysis showed some overlap between these two populations. These findings provide compelling evidence for the existence of two cryptic species within *Amblygaster sirm*, necessitating their management as separate units. This integrated approach offers crucial information for implementing sustainable fishing practices tailored to each lineage's unique characteristics, contributing to the long-term conservation of this valuable fishery resource.

CHAPTER 1

INTRODUCTION

1.1 Overview of study

The sustainable management of small pelagic fisheries represents one of the most pressing challenges in marine conservation across Southeast Asia, with overexploitation threatening both ecosystem integrity and regional food security (Teh and Pauly, 2018). Marine fisheries contributed to Asian economies, including Malaysia's, by providing food security, employment, foreign exchange earnings, and sustainable economic activity (Pauly and Zeller, 2016; Chan et al., 2017). However, the high demand for this animal protein source, coupled with population growth, threatens to deplete fish stocks due to destructive fishing practices, weak governance, poor management practices and the illegal, unreported, and unregulated (IUU) fishing (FAO and UNEP, 2010; Jesintha and Madhavi, 2020; Widjaja et al., 2020; Wong and Yong, 2020).

While numerous studies have examined commercially important species in temperate waters, research on tropical sardines remains comparatively limited despite their crucial role in both marine food webs and regional economies (Milton et al., 2017). The genus *Amblygaster*, with its taxonomic complexity and economic significance, exemplifies the intersection of scientific and management challenges facing marine resources in the region (Hunnam, 2021). This gap hinders effective stock management in Southeast Asia, where sardines play a crucial role in both ecosystem stability and local fisheries.

The genus *Amblygaster*, which closely resembles the genus *Sardinella*, is one of the economically important fish genera belonging to the family Clupeidae (Anthonipillai, 2016). The spotted sardinella, *Amblygaster sirm* (Walbaum, 1792), which will be the focus of this study is a small pelagic fish commonly caught in Southeast Asia. Known locally as "Tamban Beluru" among Malaysians and "Siro" to Indonesians, this species inhabits reef areas at a depth of 10 to 75 m (Suseno et al., 2014).

This study addresses fundamental questions about population connectivity and genetic structure that have significant implications beyond a single species, potentially informing broader conservation frameworks for small pelagic fishes throughout Malaysia and neighboring countries (Willette et al., 2016). By integrating molecular approaches with traditional morphological analyses, this research aims to bridge critical knowledge gaps that currently hamper evidence-based fisheries management in this ecologically important region (Cadrin et al., 2014). Information on the population or stock structure is essential for a proper fishery management and conservation strategy, especially when it involves commercial exploitation (Maunder and Punt, 2013; Phinchongsakuldit et al., 2013; Prasetyo et al., 2018; Ha et al., 2020; Ha et al., 2020). Various methods are employed in stock identification, such as morphometrics and meristics, alloenzymes and DNA analyses (Zaremba and Smoleński, 2000; Kochzius, 2009; Sun et al., 2012; Rawat et al., 2017; Jamaludin et al., 2020; Abdul Halim et al., 2021). Among these, the utilisation of molecular evidence stands out as a widely applied approach for determining population structure and variability (Carvalho and Hauser, 1994; Shaklee and Bentzen, 1998; Rooker et al., 2008; Abdul-Muneer, 2014; Barman et al., 2014; Akib et al., 2015; Pedrosa-Gerasmio

et al., 2015; André et al., 2016). The genetic variability of a species can reflect its population history and its ability to adapt and evolve across its range (Drahansky et al., 2016; Liu and Zhou, 2017). Genetic markers such as allozyme, mitochondrial DNA, microsatellites and more recently single nucleotide polymorphisms (SNPs) are commonly used for stock determination (Brito and Edwards, 2009; Lamichhaney et al., 2012; Li et al., 2019; Zhang et al., 2020; Silliman et al., 2021).

Genetic markers have proven highly efficient in stock identification and population structure, assessment of mixed-stock harvests, and determination of levels of genetic variation within population (Klangnurak et al., 2012; Sun et al., 2018; Li et al., 2018; Waples and Naish, 2009; Supmee et al., 2020; Chanthran et al., 2020). This molecular evidence had a significant impact on the development of fishery management strategies (Juinio-Meñez et al., 2008; Van der Meer et al., 2012; von der Heyden et al., 2014).

The two major sources of genetic markers, mitochondrial DNA (mtDNA) and nuclear DNA (nDNA), have been widely used in population genetics and phylogenetic studies towards strategising fishery management (Wang et al., 2013; Borsa et al., 2016; Chen et al., 2018). Mitochondrial DNA provides informative markers with unique characteristics such as maternal inheritance, high copy number, and faster evolution and act as a non-recombining locus where all sites share a single genealogical history (Jondeung and Karinthanyakit, 2010). These features make mtDNA markers highly efficient for assessing stock identification.

In particular, cytochrome *b* (Cyt *b*) and cytochrome c oxidase subunit I (COI) are two widely utilized mitochondrial markers for fish population studies. Cyt *b*, proves valuable for investigating relationships among closely related populations,

making it suitable for detecting subtle genetic differentiation within *Amblygaster sirm* stocks (Farias et al., 2001; Zardoya and Meyer, 1996). The COI gene, recognized as the standard DNA barcode for animal species identification, offers high resolution for both species identification and intraspecific population structure analyses (Ward et al., 2005; Hubert et al., 2008). These markers are particularly relevant for resolving taxonomic uncertainties between morphologically similar *Amblygaster* and *Sardinella* species while simultaneously providing insights into population connectivity patterns across Southeast Asian waters (Hebert et al., 2003; Hajibabaei et al., 2007).

However, due to the uniparental characteristic of mtDNA, information from the biparental nDNA is needed for a complete, understanding of genetic diversity. Hare (2001) stated that the incorporation of the diploid nuclear loci in phylogeographical analyses is practicable and productive. The determination of the stock structure of *Amblygaster sirm* using molecular assessment is the focus of Chapters 3 and 4. It is stipulated that closely related or panmictic populations should be managed as one unit, while structured species populations require independent management units (Laikre et al., 2005; Adamson and Hurwood, 2015; Kerr et al., 2020).

Morphological variation is considered the prerequisite for taxonomic identity. Thus, any occurrence of high genetic variation should be further interrogated through morphological studies to examine differentiation. While molecular markers provide insights into population connectivity, they may not always capture local adaptations influenced by environmental pressures. Morphometric analyses complement genetic assessments by identifying phenotypic variations that may not be reflected in genotypic data. Morphological analysis can be divided into qualitative analyses (such

as body shape, colouration, and fin configuration) and quantitative analyses (such as fin and scale counts, and body length measurements) (Diéguez-Uribeondo et al., 2007; Teletchea, 2009; Lajus et al., 2015; Bräger et al., 2017; Ng et al., 2017). When morphological differences are imperceptible in qualitative measures, then quantitative studies such as morphometrics serve as important tools as they provide better discrimination power for revealing subtle differences. Due to its ease of application, morphometrics is also used to discriminate various hierarchical taxonomic levels including population and stock (Cadrin and Friedland, 1999; Turan, 1999). Thus, for a holistic study, complementary genetic and morphological markers are needed. Genetic population structure can be affected by restricted gene flow while colour variation or morphological features may be influenced by local adaptation (Ukenye et al., 2019). This combined approach has been used worldwide in many species, especially in the context of fish stock assessment (Kinsey et al., 1994; Spreitzer et al., 2012; Willette et al., 2014; Vieira et al., 2016; Sukumaran et al., 2016; García-Vásquez et al., 2019; Hanif et al., 2019; Ukenye et al., 2019). Morphological studies of sardines (genera *Amblygaster* and *Sardinella*) are reported to be challenging and often lead to misidentification due to similarities between species, for example, the morphology of *A. sirm* is similar to that of *Sardinella* species (Veerappan et al.1997; Anthonipillai, 2018; Mary et al., 2017; Hunnam, 2021).

The integration of Cyt *b* and COI markers with morphological analyses offers a powerful approach to resolving these taxonomic challenges while providing comprehensive insights into population structure. By analyzing polymorphisms in these mitochondrial regions alongside morphometric data, researchers can better delineate management units and detect even subtle genetic differentiation that may be

masked by morphological plasticity (Teletchea, 2009; Ward et al., 2009). This multi-marker approach enhances the reliability of stock identification by providing independent lines of evidence that can either corroborate or reveal discrepancies between genetic and morphological patterns of variation (Cadrin et al., 2014; Pappalardo and Ferrito., 2015).

Integrating genetic assessments into fishery management plans provides a scientific foundation for decision-making, ensuring management strategies are tailored to each fish stock's unique characteristics and dynamics, ultimately contributing to the sustainable utilisation of marine resources. The genetic assessment through determination of population structure in marine fish stocks plays a crucial role in informing effective fishery management strategies. A delineation of distinct stock or subpopulation can elucidate connectivity, migration patterns, and reproductive boundaries of fish populations. With a comprehensive understanding of population structure, fishery managers can implement spatially-explicit management measures tailored to each stock or subpopulation's specific include setting appropriate catch limits, identifying essential fish habitats for protection, and designing marine protected areas aligned with distinct stocks' distribution and movement patterns (Hilborn et al., 1992). Moreover, genetic data can reveal fishing pressure's impact on exploited populations' genetic diversity and evolutionary potential, enabling managers to adjust harvest levels to maintain genetic diversity and promote long-term sustainability (Methot and Wetzel, 2013). The combination of molecular evidence and morphological investigations for stock structure determination, can provide information on how to assess the stock of selected species in the designated area through stock assessment analysis.

1.2 Problem statement

Amblygaster sirm or spotted sardinella is a species of tropical sardine abundant in the Indian and Pacific Oceans, including in Malaysian waters. This species is one of the small pelagic fishes commonly caught in Southeast Asian region and it is economically important to the local people in Southeast Asia (SEAFDEC, 2017). This proves the high demand among local people for this species according to the SEAFDEC Annual Statistics, an overfishing trend was observed in Indonesia from 2008 to 2017 (in 2016, the landing decreased by 75% from the previous year in the Andaman Sea) (SEAFDEC, 2017). Additionally, an alarming decline of 30% landing (from 1,729.95 metric tons (mt) in 2017 to 1,326.58 mt in 2018) was also reported in the Malaysian waters of the South China Sea (SEAFDEC, 2017, 2020a, 2020b). With this alarming, urgent actions to clarify the stock boundaries and identify potentially distinct management units between the Andaman Sea and the South China Sea are needed to evaluate the potential overexploitation of spotted sardinella in the region particularly in Malaysia. To address this, genetic data on this species is of utmost importance, can be used to identify the stock structure of this species in Malaysian waters and supply the information to the fishery managers to develop an efficient fishery management strategy this is crucial because no comprehensive study exists on spotted sardinella stock management in Malaysian waters or the broader region.

Known as a migratory species (Hauser and Ward, 1998), this pelagic fish is hypothesized as genetically homogenous across its distributional range in Southeast Asia, which could lead to a high genetic pool within a huge panmictic population (Kinsey et al., 1994; Sundt and Fevolden, 1996; Hoarau et al., 2002; Cimmaruta et al., 2008; Akib et al., 2015). In this context, this study will therefore investigate genetic

diversity, population structure, and phylogeographic relationships of *A. sirm* across Malaysian waters using both mitochondrial and nuclear DNA markers. Additionally, detailed morphological analysis will be conducted to verify taxonomic identification and explore the relationship between genetic variation and phenotypic traits. This integrated approach will help determine whether the observed differences in morphology reflect underlying genetic differentiation or environmental plasticity.

The output of this project underscores the urgent need for effective management strategies to prevent further overexploitation and potential species collapse in the coming years by developing efficient fishery management strategies for *A. sirm* through genetic-based information. Given the wide distribution of this species (Whitehead, 1985), assessing its genetic variability could significantly aid in developing robust fishery management and conservation programs. By combining the molecular approach with morphological verification, more accurate lineage information would provide better interpretation that can help fishery managers to sustainably manage this species (Carvalho and Hauser, 1998).

1.3 Objectives

To address the issues mentioned above, this study aims to achieve the following objectives:

1. To assess the genetic diversity, phylogenetic relationships and population structure of *Amblygaster sirm* in Malaysian waters by using two mitochondrial DNA (mtDNA) genes.
2. To assess the genetic diversity, phylogenetic relationships and population structure of *Amblygaster sirm* in Malaysian waters by using a nuclear DNA (nDNA) gene.

3. To assess the morphological variability within *Amblygaster sirm* in Malaysian waters by using a morphometric approach to verify the finding from molecular assessment.

1.4 Thesis outline

The thesis will be structured through the following chapters to address the problem statement through a systematic hierarchical approach:

Chapter 1 (Introduction) provides an overview of the study that includes problem statement and research objectives. It highlights the existing knowledge and gaps related to the whole study and briefly explain how these gaps will be addressed in the following working chapters.

Chapter 2 (Literature review) expands on the Introduction through a more comprehensive description of the current background knowledge of the project. This chapter focuses on reviewing the scenario of marine fisheries in Malaysia, including the issues related to fisheries management in Malaysia. This chapter also describes the application of molecular analysis used for fisheries management, focusing on *A. sirm* in Malaysia. The purpose of the morphological study to verify the species identification will also be explained in this chapter.

Chapter 3 analyses the population structure of the spotted sardinella, *A. sirm* by the mitochondrial DNA cytochrome *b* (mtDNA Cyt *b*) gene. Based on the inferred population structure, some recommendations for fishery management of this species are suggested. The article related to this chapter is available in PeerJ journal (doi:10.7717/peerj.13706).

Chapter 4 reports the usefulness of the first intron of the s7 ribosomal single copy gene, a nuclear DNA (nDNA) marker, to investigate the population structure of *A. sirm.* This allows for a comparison with the findings from Chapter 3, focusing on mtDNA marker to provide a comprehensive understanding of the population structure of *A. sirm* in Malaysian waters.

Chapter 5 supported and verified the finding from Chapter 3 and 4 by using morphology analyses to describe morphological variation among *A. sirm* populations. Additional analysis from mtDNA cytochrome oxidase subunit I (COI) was employed to elucidate the molecular diversity reported in previous chapters.

Chapter 6 summarises the main findings and research outcomes from the previous chapters, regarding the population structure of spotted sardinella, *A. sirm*, in Malaysia and recommendations for fishery management based on molecular assessment.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of marine capture fisheries in Malaysia

The marine fisheries industry has contributed significantly to the economy by offering benefits such as through its role in, providing food security, employment opportunities, foreign exchange earnings and sustainable economic activity to a large segment of the Asian population, including Malaysia (Chan et al., 2017). Marine fisheries in Malaysia can be divided into two sectors: inshore and offshore fisheries, with its 16th ranking in global capture fisheries in 2019, Malaysia is one of the top nations in this sector (SEAFDEC, 2020b). According to statistics for 2022 by the Department of Fisheries Malaysia (2022), the marine fisheries in Malaysia had contributed 1,308,436 metric tonnes (mt) or 69.2% of the total fish production, valued at RM11.3 billion.

However, there has been a downward trend in Malaysian fisheries production with slight fluctuations since 2017 (Table 2.1). According to SEAFDEC (2019), the landing in 2018 showed a slight decrease compared to 2017, followed by a slight increase in 2019 (Table 2.1, Figure 2.1). This was then followed by a dip of 2.2% in total production (to 1.79 million mt in 2020 and 1.75 million mt in 2021) (Department of Fisheries Malaysia, 2022).

Table 2.1: Total fishery production of Malaysia by quantity (mt) (SEAFDEC, 2024).

	2017	2018	2019	2020	2021	2022	2023	2024
Total marine capture (mt)	1,465,113	1,452,862	1,455,446	1,383,299	1,328,041	1,308,415	1,270,277	1,395,640

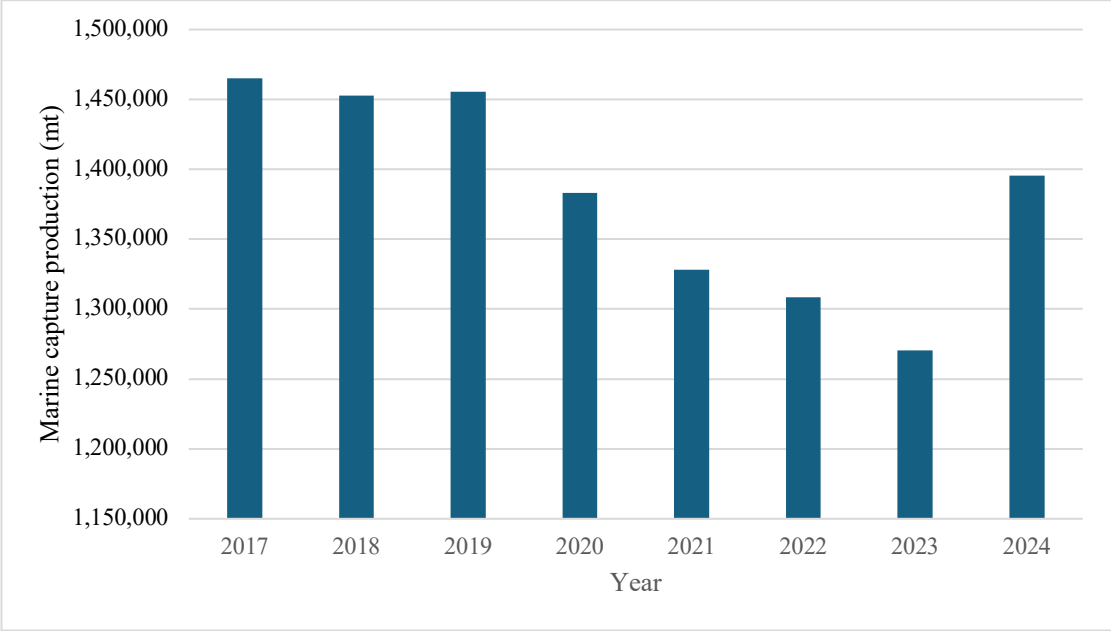

Figure 2.1: Landing trend of total capture fishery production of Malaysia by quantity (mt) from 2017 to 2024. (Department of Fisheries Malaysia, 2025).

Marine fishes can be broadly categorised into pelagic and demersal groups. Economically important pelagic species commonly caught by purse seines, drift nets, hooks, and lines include *Rastrelliger* spp., *Decapterus* spp., trevallies, sardines, anchovies, and hardtail scad (Ismail, 1997). The species under focus in this study, *Amblygaster sirm* or spotted sardinella, is a pelagic species frequently caught by purse seine (Department of Fisheries Malaysia, 2019). Locally known as *tamban beluru*, this species is vital for producing dried salted fish, a local seafood product (SEAFDEC, 2017). As a migratory fish, it was reported by SEAFDEC (2017), that the capture of this species in Indonesia decreased to 19% from 2015. This was the highest catch of this species compared to other countries in the Southeast Asian region, such as

Malaysia and Thailand. This was disproportionate of *A. sirm* landings in Southeast Asian waters, necessitates urgent attention to limit catches to ensure sustainable production (SEAFDEC 2017). However, an effective regional strategy for *A. sirm* stock management is lacking due to limited data on stock assessment for this species. Therefore, determining the size and conditions of the stock(s) of *A. sirm* is crucial before implementing any regulations for managing this species in the region.

There are several issues related to marine fisheries that affect its management and sustainability, hence affecting the well-being of fishing communities. The Department of Fisheries, Malaysia (DoFM) has taken serious measures to sustainably manage fisheries in the country. The DoFM also encourages/supports regional cooperation with neighbouring countries such as cooperation through the Southeast Asian Fisheries Development Centre (SEAFDEC), in developing policies for managing sustainable fisheries (Department of Fisheries Malaysia, 2015; Gazi and Viswanathan, 2021; SEAFDEC, 2017). The success of these strategies relies on studies conducted by researchers from member countries in the field of fisheries research (Ahmad, 2014; Lieng et al., 2018; Nasuchon and Charles, 2010; Zhang, 2018).

The DoFM is strongly committed to sustainable fisheries development through various strategies. One important program is the South China Sea Fisheries Refugia Initiative, which utilises genetic and non-genetic information to establish refugia as potential closed areas for selected species, such as the tiger prawn refugia in Sarawak, Malaysia (SEAFDEC, 2006). By engaging in regional fisheries management organisations, the DoFM contributes to strengthening the governance of shared

fisheries (Løbach et al., 2020) and development of a regional fisheries policy similar to the ASEAN Fisheries Policies (Macfadyen and Seilert, 2020).

2.2 The relevance of population genetics and phylogeographic studies for fishery stock management

Marine species possess high dispersal capability, allowing them to cover long distances (Benzie, 1998). As such, they are characterised by large populations and a significant rate of gene flow (in the absence of physical barriers), leading to genetic homogeneity over a large geographical area (Baker et al., 1993; Scoles and Graves, 1993; Palumbi and Baker, 1994). Genetic homogenisation has been observed in various species such as in Spanish mackerel, *Scomberomorus maculatus*, from the western Atlantic and Gulf of Mexico (Buonaccorsi et al., 2001), Kawakawa, *Euthynnus affinis*, in the Indian coast (Kumar et al., 2012), yellow croaker, *Larimichthys polyactis,* in the Northwest Pacific (Wang et al., 2013), long-tailed tuna, *Thunnus tonggol*, in Malaysian coastal waters (Kasim et al., 2020; Mohd Yasin et al., 2020) and two-spined yellowtail stargazer, *Uranoscopus cognatus*, in the Indo-West Pacific ocean (IWP) (Mohd Yusoff et al., 2021).

According to Ihssen et al. (1981) fish stock is "an intraspecific group of randomly mating individuals with temporal and spatial integrity". Later, Coyle (1998) redefined it as "a group of fish populations believed to be genetically isolated due to reproductive isolation". Essentially, many biologists consider fish stocks as interbreeding units derived from a common gene pool. Among these stocks, it is crucial to maintain genetic diversity and consider population structure patterns to effectively utilise resources (Carvalho and Hauser, 1994). The stock definition

provides fisheries managers with important information for sustainable fisheries management. This approach aligns with broader global goals, including the 17 Sustainable Development Goals (SDGs) introduced by the UN in 2015, with particular relevance to SDG14, which focuses on life below water. Achieving this goal through the management of fish stocks is essential for the sustainable use of marine resources.

Genetic data are now being incorporated, albeit conservatively, into stock restoration programme, and marine protected area and fisheries management strategies (von der Heyden et al., 2014). These genetic strategies are formulated based on genetically defined stocks/populations to achieve sustainable fisheries management. For example, a single fisheries management strategy is appropriate for a panmictic population, while non-panmictic fish stocks warrant different independent management actions or policies (Ward, 2000; Laikre et al., 2005; Adamson and Hurwood, 2015).

In this context, incorporating population genetic data into stock assessment is vital in formulating and implementing a comprehensive management strategy for *A. sirm.* This strategy should be developed between Malaysia and its neighbouring countries to attain mutual sustainability. Hypothetically, all the populations of *A. sirm* are expected to be genetically homogeneous or panmictic due to the pelagic nature of the species (Carvalho and Hauser, 1998; Garoia et al., 2004; Fratini et al., 2016) and could be managed as a single stock. However, studies on other species demonstrated that this is rarely the case due to many factors. Therefore, separate management for each stock unit is necessary. This approach aligns with the growing importance of population genetics data (built on molecular markers such as mitochondrial and nuclear genes). This approach is gaining traction in fisheries management, as

evidenced its uses for various species including Spanish mackerel in the northern Indian Ocean (Radhakrishnan et al., 2018); tiger shark, *Galeocerdo cuvier*, across the Indo-Pacific Ocean (Holmes et al., 2017); Chinese loaches, *Misgurnus mohoity* and *M. bipartitus*, in Northeast China (Yi et al., 2016) and Asian green mussel, *Perna viridis*, along the Indian coast (Divya et al., 2020).

2.2.1 Mitochondria DNA (mtDNA)

Various genes from the mitochondrial DNA (mtDNA) have been widely used as molecular markers since the 1980s, and extensively investigated in numerous studies across various fields. In fisheries, this marker has been applied in population genetics and phylogeography (Grant et al., 2005; Trivedi et al., 2018; Abdalwahhab et al., 2020; Cheng et al., 2021; Dudgeon et al., 2012; Xiao et al., 2022), phylogenetic studies (Farias et al., 2001; Peng et al., 2004; Jondeung and Karinthanyakit, 2010; Schönhuth and Mayden, 2010; García-Varela et al., 2013), evolutionary studies (Guo and Chen, 2010; Liu et al., 2012; Thomas et al., 2014; Lavoué et al., 2017; Cheng et al., 2021), conservation (Barber et al., 2011; Dudgeon et al., 2012; Abdul-Muneer, 2014; Trivedi et al., 2016; Trivediet al., 2016) and also for species identification (Iamsuwansuk et al., 2012; Willette and Santos, 2013; Willette et al., 2014; Camacho-Gastélum et al., 2017; Schmidt et al., 2018). Based on this body of evidence, mtDNA has emerged as an ideal source of molecular markers highly valuable for population and evolutionary biology research (Chong et al., 2018). Various mitochondrial regions or genes, such as the control region (CR) or the D-loop, cytochrome *b* (Cyt *b*) and cytochrome oxidase subunit I (COI), have been extensively utilised to assess population structure, especially in fisheries research (Menezes et al., 2012; Barman et al., 2014; Kumar and Kocour, 2015; He et al., 2022; Alvarenga et al., 2024; Gaël Denys et al., 2024).

Due to its high copy number, exceptionally low recombination rate, maternal inheritance, small size, ease of use and rapid substitution rate, the mitochondrial control region (CR) also known as the displacement loop (D-loop), is commonly used for inferring species-level molecular phylogenies (Bronstein et al., 2018) as well as for determining population linkages, conservation units and migration pathways (Verma et al., 2016). Several studies have documented the use of this marker in various situations: to determine the genetic distinctiveness of eight species of Sparidae (Ostellari et al., 1996), to reconstruct the phylogenetic relationships of sinipercine fishes (Jin et al., 2006), to evaluate the population structure of the small yellow croaker, *Larimichthys polyactis*, in the northern East China Sea (Xiao et al., 2009), the blackspotted sea bream *Pagellus bogaraveo*, between the Northeast Atlantic and Mediterranean ocean (Robalo et al., 2021), the shad species, *Tenualosa ilisha*, in Bangladesh (Khatun et al., 2022) or the exploited groundfish Cape pike, *Merluccius paradoxus*, in southern Africa, (von der Heyden et al., 2010), as well as the genetic variation of the Japanese whiting, *Silago japonica* (Gao et al., 2019).

Another mtDNA marker that has gained wide recognition in molecular studies is the cytochrome *b* (Cyt *b*) gene. With a length of 1140bp, the protein-coding Cyt *b* gene has proven to be highly efficient in determining intra and interspecific as well as higher-level phylogenetic relationships (Roques et al., 2006; Schönhuth and Mayden, 2010; Zhang et al., 2018). The Cyt *b* gene is also frequently used in systematics studies to investigate divergence at the taxonomic level, and its evolutionary dynamics have been intensively characterised, making it one of the most comprehensively sequenced genes to date (Peng et al., 2004; Habib et al., 2011; Rahim et al., 2012; Schmidt et al., 2018). Reports on the population genetic structure of bigeye tuna, *Thunnus obesus*,

(Wu et al., 2014), lesser spotted sandpiper, *Thamnaconus hypargyreus*, (Li et al., 2016), common carp, *Labeo fimbriatus*, (Swain et al., 2016), carp, *Osteobrama belangeri*, (Singh et al., 2016), sea bass, *Priacanthus hamrur*, (Vidya et al., 2017) have demonstrated the efficacy of Cyt *b* as a molecular tool for studying genetic variations in fishes. Therefore, this study will be using Cyt *b* gene to study the population genetics of *A. sirm* found in Malaysian waters.

The mtDNA cytochrome oxidase subunit I (COI) is well-established as the DNA barcoding gene for species identification. Following Ward et al. 2005, who identified 207 species of mostly Australian marine fishes with clear inter-species delimitation, this marker will be utilized to genotype all sampled *A. sirm* specimens throughout the study range. Furthermore, this genetic marker has also proven effective in population genetics, with applications in studies of the yellow drum or croaker, *Nibea albiflora*, in the South China Sea (Xu et al., 2012); yellowfin tuna, *Thunnus albacares*, in the Central Pacific (Li et al., 2015); carangid fishes on the coast of Kakinada, India (Persis et al., 2009) and the black-banded trevally, *Seriolina nigrofasciata*, in the Bay of Bengal (Barik et al., 2020).

Therefore, it is expected that the use of mtDNA markers will provide a robust framework for understanding the genetic diversity, evolutionary processes and conservation implication that contribute to valuable insights into population genetics, especially in the study of marine fish populations.

2.2.2 Nuclear DNA (nDNA)

Although mtDNA provides a combination of rapidly evolving genes that have been proven efficient for studying closely related lineages, it is prone to biases due to factors

like stochastic lineage sorting, past introgression and natural selection (Vigilant et al., 1989; Wallace and Chalkia, 2013; Kelly et al., 2014; Abad et al., 2016; van der Loos and Nijland, 2021; Kasmi et al., 2023). Therefore, it is recommended to complement mtDNA-based molecular studies with additional information from nuclear DNA (nDNA) (Cheng et al., 2011; Ding et al., 2023).

Mitochondrial DNA (mtDNA) was the marker of choice very early on in studies on phylogenetics and phylogeography; however, since the 1990s, the incorporation of nDNA has progressed (Brito and Edwards, 2009). The high mutation rate of mtDNA, while useful for detecting recent divergence, can result in homoplasy and saturation at deeper evolutionary timescales, confounding phylogenetic analyses (Galtier et al., 2009). This shift arises from the substantial advantages of nDNA, including its large size of more than 3 billion base pairs (bp) compared to the 16,000 to 18,000 bp of mtDNA. Furthermore, due to its uniparental inheritance, mtDNA cannot detect male-mediated gene flow, potentially masking important demographic processes in population studies (Hurst and Jiggins, 2005; Osmond and Coop, 2020; Zink and Barrowclough, 2008). However, nDNA can offer an enormous quantity of genomic data that can be explored for variation, besides providing more potential diagnostic markers useful for discriminating fish stock due to its large size (Wirgin and Waldman, 2005; Jiang et al., 2016).

The relatively small effective population size of mtDNA makes it more susceptible to genetic bottlenecks and founder effects, which can exaggerate the appearance of population differentiation (Palumbi, 2003). Numerous studies have demonstrated the effectiveness of single copy nDNA marker, such as the application of the 5S rRNA gene in the Atlantic salmon, *Salmo salar*, and rainbow trout,

Onchorhynchus mykiss, (Pendas et al., 1995), the application of internal transcribed spacer (ITS) region in the anchovy, *Coilia nasus*, (Liu et al., 2012), and the application of the rhodopsin gene to examine the phylogeny of the order Cypriniformes (Chen et al., 2008). Additionally, the application of the β-actin gene intron 2 as a marker in fish taxonomy (Lee and Gye, 2001) and the utilisation of the s7 ribosomal intron protein in the population study of the black sea bream, *Spondyliosoma cantharus* in the Eastern Atlantic and Mediterranean (Neves et al., 2020a) have further highlighted the value of nDNA in elucidating their genetic relationships.

As suggested by Zhang and Hewitt (2003), the study of natural populations would be more meaningful if extensive data from the other markers could be fully utilised. Recent studies have shown that selective sweeps in mtDNA can mimic the genetic signatures of population expansions, potentially leading to incorrect demographic inferences (Hurst and Jiggins, 2005; Zink and Barrowclough, 2008; Osmond and Coop, 2020). For example, the application of single-copy nDNA or the development of Next Generation Sequencing (NGS) may be an effective complementary approach. The use of SNPs or NGS as powerful tools in genotyping and population studies is a way forward. Indeed, phylogenetic analysis that was not based on gene trees (Brito and Edwards, 2009) suggests that the use of these markers to confirm the population structure of *A. sirm* inferred from mtDNA would be useful.

However, despite its potential to reveal the genetic structure and dynamics of populations, this option has its limitations (Zhang and Hewitt, 2003; Brito and Edwards, 2009) including factors such as cost and time (Zink and Barrowclough, 2008). Alternatively, considering the challenges associated with SNPs or even NGS, the use of single-copy nDNA is a more logistically feasible option for this study. Jiang

et al. (2016) supported the suitability of the marker for molecular phylogenetic analyses. Among all nDNA markers, the highly conserved s7 ribosomal intron protein or s7 gene, was selected as a surrogate non-coding protein sequence genetic marker for this study. It is not as widely applied but the marker has demonstrated a high degree of intraspecific variation and has been used for population inference across various phylogenetic studies involving plants and animals, including fish (Sultmann et al., 1995; Wang et al., 2002; Lavoué et al., 2003; Feulner et al., 2006; Guo and Chen, 2010; Fietz et al., 2020). Molecular studies in fishes, such as those involving cyprinids (Wang et al., 2002), grenadier fish, *Coilia nasus,* (Liu et al., 2012), African electric fishes (Feulner et al., 2006; Lavoué et al., 2000), cichlids (Chakrabarty, 2006) and Baja Californa disjunct species, *Anisotremus* spp., (Bernardi and Lape, 2005) have documented its effectiveness.

2.3 The complementarity of morphological and molecular diagnostics for fishery stock management

Genetic collapses or breaks are regularly observed, despite the perceived genetic homogeneity in marine fishes due to large populations and robust dispersal patterns. This is the case for the anchovy, *Engraulis mordax*, and sardine, *Sardinops sagax*, in the Eastern Pacific (Lecomte et al., 2004), bluntnose clipperfish, *Clinus cottoides*, along the coast of South Africa (von der Heyden et al., 2008), the Indo-West Pacific coastal fish fourfinger threadfin, *Eleutheronema tetradactylum*, (Horne et al., 2011), tiger prawn, *Penaeus monodon*, in the southwestern, eastern and Andaman coastal waters of India (Mandal et al., 2012), silky shark, *Carcharhinus falciformis*, in the western Atlantic and Indo-Pacific Oceans and Red Sea (Clarke et al., 2015), and kelee shad, *Hilsa kelee*, from the Bay of Bengal and Arabian Sea (Sarker et al., 2020). Often

these breaks lead to highly disparate populations that suggest the occurrence of separate genetic taxa or cryptic diversity. In such cases, the populations should be re-examined by morphological approach. Hence, the traditional approach of species identification should be integrated with molecular taxonomy to validate these findings.

Morphology serves as the cornerstone of taxonomy; thus, any potential genetic-based occurrence of cryptic diversity should be verified through the application of morphological taxonomy. The effectiveness of these combined approaches has been highlighted in numerous studies, such as integration of genetic and morphological data of Zebra fish, *Diplodus hottentus*, in Angola and South Africa (Gwilliam et al., 2018), relationships between morphological and genetic patterns in killifishes, *Austrolebias robustus*, in South America (García et al., 2019), morphological and genetic diversity of catfish, *Chrysichthys nigrodigitatus*, in Nigeria (Adelieje et al., 2020), genetic diversity and morphological characteristics of three populations of turbot, *Scopthalmus maximus*, along the Bulgarian Black Sea (Ivanova et al., 2021) and the morphometric and genetic study of red snapper, *Lutjanus malabaricus*, in Indonesia (Zamroni et al., 2021).

2.3.1 The use of morphological approach for population structure confirmation

The division of a species into distinct populations based on the molecular approach is one of the first steps for strategising decisions in the conservation and management of fisheries (Hubert et al., 2008; Ward et al., 2005; Zhang et al., 2020). Confirmation of population structure through morphological approach offers additional invaluable insights into evolutionary dynamics, ecological adaptations, and management

strategies especially for the most diverse group of living vertebrates, the fish (Anumudu and Mojekwu, 2015; Hoff et al., 2020).

The morphological approach involves the analysis of physical characteristics of individuals within a population or among different populations that involve body shape, size, colouration, fin morphology, and various other external features. Their patterns of differentiation reflect the morphological population structuring within a species. Morphological traits are regularly used in fish biology to measure stock discreteness and relationships between different groups in taxonomy and are one of the early tools for assessing population structure and identifying stocks (Cadrin and Friedland, 1999; Turan, 1999). This has been done in several studies, including studies on anemonefish, *Amphirion clarkia*, in southern Japan (Bell et al., 1982), tilapia, *Altolamprologus compressiceps* in Lake Tanganyika, East Africa (Spreitzer et al., 2012), Northeast American rainbow smelt, *Osmerus mordax*, (Dodson et al., 2015) and Guinean tilapia, *Tilapia guineensis*, in West African coastal waters (Ukenye et al., 2019).

The use of morphological approaches within specific fish groups, such as Sardinella group has a rich history dating back to the 1920s. During this time, efforts were made to distinguish sardine populations in the Atlantic and the Mediterranean (reviewed in Silva, 2003) based on morphological characteristics. Subsequently, studies have been reported on the *Sardinella* species complex such as in the Spanish sardine, *Sardinella aurita*, in the coastal waters of Florida, USA (Kinsey et al., 1994), *Sardinella lemuru*, on Mindanao Island, the Philippines (Luceno et al., 2014) and the Indian sardine, *Sardinella longiceps* (Sukumaran et al., 2016).

The morphological approach can be divided into qualitative and quantitative analysis. The qualitative approach is a traditional method that is based on the examination of physical traits such as body shape, fin morphology and colouration. On the other hand, quantitative method such as morphometric and meristic analyses employ statistical methods to measure and analyse specific measurable morphological traits. In morphometric studies, two primary approaches are commonly used: 1) traditional multivariate morphometrics which is based on variations in size and shape (distance between homologous anatomical points), and 2) geometric morphometrics which rely on calibrated coordinates or morphometric locations referred to as 'landmarks' (Cadrin and Friedland, 1999). Multivariate techniques such as Principal Component Analysis (PCA) and Discriminant Analysis (DA) are commonly used to identify patterns of morphological variation and determine whether populations can be differentiated based on their physical characteristics. The multivariate analyses have been efficient in delineating populations of twaite shad, *Alosa fallax*, in Turkish seas (Turan and Basusta, 2001), Bali sardinella, *Sardinella lemuru*, in Mindanao Island, Philippines (Joy et al., 2014), gangetic whiting, *Sillaginopsis panijus*, in the southern coastal zone of Bangladesh (Siddik et al., 2016), smooth belly sardinella, *Amblygaster clupeoides*, in the Bay of Bengal Coast, Bangladesh (Hanif et al., 2019), hilsa, *Tenualosa ilisha* and *T. toli*, in Bangladesh waters (Das et al., 2020) and Kawakawa, *Euthynnus affinis*, in Peninsular Malaysia (Binashikhbubkr et al., 2022)

By integrating molecular techniques with morphological methods, researchers can gain a more comprehensive understanding of the structure and relationships within studied species, thus paving the way for effective conservation efforts. This