

UNIVERSITI TEKNOLOGI MARA

**IDENTIFICATION OF SELECTED
WORLD'S SMALLEST FISH
SPECIES OF THE GENUS *Paedocypris*
(CYPRINIDAE) AND POPULATION
GENETICS OF THE *Paedocypris*
progenetica THROUGH MOLECULAR
METHODS FROM MALAYSIA**

NORJASMIN BINTI HUSSIN

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Sciences
(Molecular Biology)

Faculty of Applied Sciences

March 2023

ABSTRACT

The accurate identification techniques of fish species are crucial to biodiversity conservation, fisheries management, and international aquarium trade. Globally, the genus *Paedocypris* is the smallest fish species and was declared as near threatened by the IUCN (2020) due to anthropogenic activities, including logging, urbanisation, and industrialisation. Thus, this research aimed to identify the *Paedocypris* specimens in Malaysia into distinct species through conventional and molecular approaches as well as to elucidate the population structure of selected *Paedocypris* spp. A total of 16 morphometric parameters and six chromatophore patterns of *Paedocypris* specimens collected from three peat swamps in Malaysia (Selangor, Perak, and Sarawak) were measured on 66 out of 103 specimens. The analyses involved for morphological study were univariate analysis of variance (ANOVA) and multivariate discriminant function analysis (DFA). The species identification of specimens was further analysed using the cytochrome *c* oxidase I (*COI*) gene that performed on 51 out of the 66 *Paedocypris* specimens using standard DNA barcoding, phylogenetic analyses, and divergence time estimation of mitochondrial *COI* gene. Later, *P. progenetica* was selected and assessed using population genetics with newly developed D-loop primer and *COI* gene as it was detected in two out of three locations in Malaysia. The morphological studies have successfully classified and identified two *Paedocypris* species as *Paedocypris progenetica* and *Paedocypris micromegethes* by using DFA and chromatophore patterns. Nine out of 16 significant characteristics were found as the highest characters loading in Function 1 and Function 2 are PDL (predorsal length) and HL (head length) with small wilk's lambda (0.185) in DFA. Further confirmation on *Paedocypris* species identification using DNA barcoding approach applying Neighbour-Joining and Maximum Likelihood methods, recorded high bootstrap values ($n > 95\%$) and highly congruent with the morphological findings revealing three clades within *Paedocypris* species namely *P. progenetica* and *P. micromegethes* known to occur in Malaysia. A barcoding gap of 3% was apparent in the whole data set resulting in conspecific distances ranging from 4.93% to 5.44%. By contrast, congeneric distances varied between 4.28% and 7.71%, showing genetic variation. Based on divergence time estimation, the *P. carbunculus* was diverged earlier compared to *P. micromegethes* followed by *P. progenetica*. The divergence time between *P. carbunculus* from Sumatera showed the earliest divergence during the middle of Jurassic, followed by *P. progenetica* (24.5 Mya) and *P. micromegethes* (7.5 Mya) during the early Oligocene and Miocene. The population structure assessment on *P. progenetica* found they were divided into two clades in each marker namely Clade I (Perak) and II (Selangor) which found moderate haplotype diversity ($h=0.668$) and nucleotide diversity ($nD=0.220$) from *COI* analysis showing moderate genetic variation between Selangor and Perak. Thus, this reflects the possibility of past geological events and further supported by D-loop marker analysis where moderate genetic variation was also seen ($h=0.628$, $nD=0.597$). The population structure was supported with a high value of fixation index (*Fst*) between two populations for both markers (*COI*=1.30262, D-loop=0.70694). The mtDNA markers showed past population expansion through Tajima's, Fu *Fs*'s, and Bayesian skyline plot during the Oligocene. Therefore, the findings of morphological and molecular analyses successfully identified *P. progenetica* and *P. micromegethes* in Peninsular Malaysia and Sarawak respectively and may contribute to facilitate the conservation management of *P. progenetica* in Malaysia.

ACKNOWLEDGEMENT

Firstly, I would like to thank Allah S.W.T for giving me the opportunity to embark on my master and permitted me to finish this long and challenging journey successfully despite the Covid-19 which attacked all around the world including Malaysia without full stop.

My gratitude and a million thanks go to my main supervisor Dr Izzati Adilah Azmir, who has devoted endless guidance and attention during my study period and a very deep recognition award to you for giving me uncountable knowledge and opportunity to finish this research. She also has taught the meaning of life, always trust, and understand me, and please know that this dissertation would be meaningless without you, Dr Izzati.

I would also like to thank to committee members, Associate Professor Dr Yuzine Esa and Associate Professor Dr Amirrudin Ahmad for their persistent supports and trust me to complete this study. Not to forget, thank you to University Putra Malaysia (UPM) for allowing me to conduct my lab work at their Fish Genetics and Breeding Laboratory, Department of Aquaculture, Faculty of Agriculture.

Finally, this thesis is dedicated to the loving memory of my very dear late father and mother for the vision and determination to support me. This piece of victory is dedicated to both of you. Not to forget my friends who always give moral support and help me throughout this journey. Alhamdulillah.

TABLE OF CONTENTS

	Page
CONFIRMATION BY PANEL OF EXAMINERS	ii
AUTHOR'S DECLARATION	iii
ABSTRACT	iv
ACKNOWLEDGEMENT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	x
LIST OF FIGURES	xiii
LIST OF SYMBOLS	xviii
LIST OF ABBREVIATIONS	xix
CHAPTER ONE INTRODUCTION	1
1.1 Background of Study	1
1.2 Problem Statement	3
1.3 Significant of the study	4
1.4 Objectives	5
CHAPTER TWO LITERATURE REVIEW	6
2.1 <i>Paedocypris</i> spp.	6
2.1.1 The Habitat of <i>Paedocypris</i> spp.	10
2.1.2 Disturbances of blackwater peat swamp in Malaysia.	11
2.2 Taxonomic description and Morphology of <i>Paedocypris</i> spp.	13
2.2.1 Important of <i>Paedocypris</i> spp.	17
2.2.1.1 <i>Important as The Ornamental Fish</i>	18
2.2.1.2 <i>Importance as Biological Indicator</i>	21
2.3 Molecular Identification of <i>Paedocypris</i> spp.	21
2.3.1 Mitochondrial DNA using <i>COI</i> gene as a marker	22
2.3.2 The mtDNA using D-loop region	24
2.3.3 DNA barcoding in Fish	24
2.3.4 Benefits and Successful DNA Barcoding in Fish Identification.	27

CHAPTER ONE

INTRODUCTION

1.1 Background of Study

Morphometric studies have been the most important aspect in the examination of the traits and relationships among taxa, as well as meristic studies and are often used in taxon identification as they are minimally affected by environmental changes (Gonzalez et al., 2016). Historically, species identification of fish sample was based on the external morphological characteristics which comprise of the colour patches, scales' size count, body shape, quantity, and spot of find, number and form of fin rays or the relative measurements in parts of the body (Chan et al., 2014; Li et al., 2019). In addition, the blackwater peat swamps of Malaysia harbours a great diversity of freshwater fishes and has yet to be identified including the newly described miniature fishes of the genus *Paedocypris* (Kottelat et al., 2006; Britz & Kottelat, 2008; Page et al., 2012; Conway et al., 2018;). Available record on this genus was from the peat swamps of Sumatera and Borneo namely *Paedocypris progenetica* and *Paedocypris micromegethes* (Rüber et al., 2007; Conway et al., 2011; Schäfer, 2017), while another species was found in Central Kalimantan, *Paedocypris carbunculus* (Conway et al., 2018) and none scientifically record from Peninsular Malaysia.

The remarkable miniature *Paedocypris* fishes inhabit tannin-stained and highly acidic (pH 3-5) blackwater peat swamps of Southeast Asia (Kottelat et al., 2006). The speciality of this genus mainly attributed to the chromatophore patterns and were useful meristical characters in species identification especially in the case of *Paedocypris* species identification (Kottelat et al. 2006; Britz and Kottelat, 2008; Britz and Conway, 2009). However, most of the blackwater peat swamp in Southeast Asia, including those in Malaysia, have been extremely degraded due to deforestation and land conversion into oil palm plantations (Shuhada et al., 2017). Moreover, the taxonomic differentiation using classical morphological approach shows high ambiguities because of inadequate specimens (Mayden & Chen, 2010; Britz et al. 2014). Considering the near threatened IUCN status of *Paedocypris* and its declining habitat and populations, the taxonomic identification using morphological approach alone are inadequate. Due to the inadequacy and infrequency of taxonomic publications and reference books especially