

SOCIALIST REPUBLIC OF VIETNAM

Independence - Freedom – Happiness

SUMMARY OF DOCTORAL THESIS

Thesis title: Identification and application of SNP markers using EzRAD technique for three typical fish species in the lower Mekong river basin

Specialization: Biotechnology

Code: 9420201

Full name of PhD student: Truong Thi Oanh

Course: 2020

Science Supervisors: 1. Assoc.Prof. Dr. Dang Thuy Binh
Nha Trang University
2. Dr. Ngo Thai Bich Van
University of Science and Technology - The University of Danang

Training institution: University of Science and Technology - The University of Danang

Summary of doctoral thesis:

1. Reasons for choosing the thesis

The Mekong River Basin (MRB), a globally significant biodiversity hotspot, is increasingly imperiled by environmental degradation and anthropogenic activities, notably hydroelectric dam construction. These impacts not only create physical barriers that hinder migration, alter or fragment natural fish habitats, and affect the behavior of many aquatic species but also change the natural flow, flood patterns, and seasonal water volumes. These alterations have precipitated a substantial decline in fisheries yields and amplified the extinction risk for indigenous fish species. Consequently, comprehensive research into the genetic diversity and population structure of fish taxa representative of key ecological zones and life history strategies is paramount for effective fisheries management, population recovery, and the attainment of a sustainable ecosystem.

To date, a variety of molecular markers (mitochondrial DNA, Random Amplified Polymorphic DNA - RAPD, microsatellites, Single Nucleotide Polymorphisms - SNPs) have been used in population genetic studies of Mekong River fish species. Among these, the thousands of SNPs markers distributed throughout the genome, discovered from DNA sequencing data based on Restriction site Associated DNA Sequencing (RAD-seq), contribute to elucidating the population structure of organisms, especially species lacking genomic information.

Given the diverse migratory characteristics of Mekong River fish species in the context of a fluctuating and changing environment, three fish species with different evolutionary histories and migratory behaviors, including *Macrognathus siamensis* (short lifespan, non-migratory), *Labeo chrysophekadion* (long lifespan, short and facultative migration), and *Pangasius larnaudii* (long lifespan, long-distance migration) distributed in the Mekong River basin, were collected. This study aims to address key issues through two core questions: 1) How do the interactions between migratory behavior and developmental life cycles shape genetic variations and the capacity to maintain gene flow among species?; and 2) Do current evidences of genetic structure and diversity decline reflect the cumulative impacts of landscape barriers and changing environmental conditions in the Lower Mekong River? These findings will contribute to elucidating the mechanisms of genetic maintenance and differentiation in Mekong River fish species, while providing crucial scientific evidence for conservation strategies and the sustainable management of fisheries resources.

2. Research objectives

2.1. General Objective

This study aims to assess and compare the levels of genetic diversity and population connectivity of three typical fish species in the Mekong River using advanced genetic techniques. This will establish a genetic database to support the management, conservation, and sustainable exploitation of fisheries resources in the region.

2.2. Specific Objectives

- To perform *de novo* genome assembly for three typical fish species (*Macrognathus siamensis*, *Labeo chrysophekadion*, and *Pangasius larnaudii*) in the

Lower Mekong River basin, and to identify single nucleotide polymorphisms (SNPs) for population genetic analyses.

- To evaluate and compare the population genetic profiles of the three fish species, including the levels of genetic diversity within and among populations, effective population size (N_e), migration patterns, and the degree of genetic connectivity between populations.

- To assemble and annotate the mitogenome of *L. chrysophekadion*, identify RAD barcodes, and analyze population genetic structures to complement and validate the analytical results derived from nuclear genomic SNP data.

3. Research subjects and scope

Research subjects: Three fish species, including *M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*.

Research scope: Adult fish specimens were randomly sampled from the MRB from 2017 to 2021. Genetic analyses were conducted at the Molecular Biology Laboratory, Nha Trang University.

4. Research methods

4.1. Theoretical research methods

Relevant literature, including books, journals, and studies, was collected and analyzed to examine the Mekong River ecosystems, fish migratory routes, and reproductive biology, as well as research directly relevant to the thesis's contents.

4.2. Experimental research methods

- This research employed the EzRAD methodology to prepare genomic DNA libraries, which were subsequently sequenced using the Illumina HiSeq 4000 platform;

- This research applied bioinformatics tools to perform *de novo* assembly, identify SNP markers, and investigate genetic diversity, population genetic structure, effective population size, and predict migration patterns of three typical fish species in the Lower Mekong Basin;

- This research utilized bioinformatics tools for mitochondrial genome assembly and annotation, RAD barcode identification, and subsequent application in analyzing the population genetic structure of *L. chrysophekadion*.

5. Research Content

1) *De novo* assembly of three typical fish species (*M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*) and identification of SNP molecular markers.

2) Investigation of genetic diversity, population structure, estimation of effective population size, and prediction of migration patterns of the three fish species.

3) *Mitochondrial* genome assembly and annotation, identification of mitochondrial DNA segments from RAD-seq data, and population structure analysis of *L. chrysophekadion*.

6. Major results of the Thesis

- The study successfully collected, prepared libraries, and performed *de novo* assembly of three typical fish species in the Mekong River basin. SNPs were determined, including 4,237 SNPs for *M. siamensis* (9 populations), 825 SNPs for *L. chrysophekadion* (9 populations); and 1,270 SNPs for *P. larnaudii* (7 populations).

- The levels of genetic diversity, genetic differentiation, and population structure of the three fish species in the Mekong River were determined. Migration patterns of *L. chrysophekadion* and *P. larnaudii* in the LMB were predicted using SNP markers. Specifically:

+ Genetic diversity positively correlated with fish migration, while inbreeding coefficient negatively correlated. Non-migratory *M. siamensis* showed the lowest diversity and highest inbreeding; long-distance migratory *P. larnaudii*, the highest diversity and lowest inbreeding; and short-distance/facultative migratory *L. chrysophekadion*, intermediate levels.

+ Analysis of genetic differentiation and population structure revealed distinct population differentiation among the fish species. *M. siamensis* showed four population groups corresponding to the Upper Mekong Basin, the upper and middle LMB, tributaries in the middle LMB, and the lower LMB. *L. chrysophekadion* exhibited three population groups, corresponding to the upper, middle, and lower LMB. In contrast, *P. larnaudii* populations showed high genetic connectivity.

+ Migration model predictions for both *L. chrysophekadion* and *P. larnaudii* indicated significant migration capabilities, including downstream, upstream, and passage over the Khone Falls. However, a notable observation was that short-distance/facultative migratory fish exhibited longer-than-expected migration routes,

while long-distance migratory fish showed discontinuous migration, mainly limited between the lower and middle LMB.

- The mitochondrial genome of *L. chrysophekadion* (Genbank accession number OR637878) was assembled and annotated, with a length of 16,600 bp. Using aligned mitochondrial DNA segments from RAD-seq data (757 bp), the results showed high genetic connectivity among populations, with high haplotype diversity and low nucleotide diversity. No significant genetic differentiation was observed among populations, except Luang Prabang.

7. Scientific and Practical Significance of the Thesis

Scientific significance: This research has significantly expanded the fundamental knowledge of population genetics within the context of the Mekong River ecosystem. By employing the advanced EzRAD technique and conducting comprehensive genomic analyses, the study has overcome the limitations of traditional methods, providing a detailed analytical framework for the genetic structure and migration patterns of three typical fish species. These findings not only illuminate the unique evolutionary and adaptive processes of each species but also offer a valuable research paradigm for assessing and conserving biodiversity in major river systems globally. In particular, the decoding of the mitochondrial genome and RAD barcode analysis has opened a novel avenue of investigation, enabling deeper insights into the role of maternal inheritance in the differentiation and adaptation of aquatic species.

Practical significance: The findings of this research hold significant practical implications for the conservation and management of fisheries resources within the MRB. The elucidation of population structure and migration patterns of fish species enables resource managers to implement appropriate conservation strategies, ensuring the sustainability of fish populations. The SNP and mitogenome data generated from this study constitute a valuable genetic resource, which can be utilized to monitor temporal genetic changes in fish populations, assess the impacts of environmental factors, and support population restoration programs. Furthermore, the decoding of the mitogenome of *L. chrysophekadion* can be applied in the development of species identification methods and seafood product traceability.

8. Structure of the thesis

The dissertation consists of 122 pages (excluding appendices and references), including 18 tables and 35 figures and graphs. Introduction 6 pages; conclusion and recommendations 2 pages; list of published research works 1 page; references 21 pages, including 293 English-language references consisting of scientific journal articles and books. The main content of the thesis is divided into 03 chapters: Chapter 1: Literature review 47 pages; Chapter 2: Materials and research methods 20 pages; Chapter 3: Results and Discussion 47 pages.

9. New contributions of the thesis

❖ Content:

This study utilizes the EzRAD technique to identify and select highly efficient and accurate population-specific SNP markers. This approach provides a novel framework for population genetics research on three typical Mekong River fish species, characterized by varying life histories and migratory patterns: the non-migratory *M. siamensis* (representative of local populations), the short-distance and facultative migrant *L. chrysophekadion* (demonstrating regional connectivity and environmental adaptability), and the long-distance migratory *P. larraudii* (reflecting broader ecological linkages).

❖ New points and highlights of the research:

- This study offers a comprehensive analysis of genetic diversity and population structure in three Mekong River fish species, using a large SNP dataset generated by EzRAD. It significantly improves upon previous studies that used traditional markers or single-species SNP analysis.

- This study, for the first time, employed SNP data and diverse algorithms to accurately predict migration patterns of *L. chrysophekadion* and *P. larraudii* within the Lower Mekong Basin, elucidating migration routes and population connectivity that are consistent with their historical development.

- This study utilizes EzRAD data from multiple individuals to both reconstruct the mitochondrial genome and simultaneously screen high-density polymorphic sites in *L. chrysophekadion*. This data source was effectively applied to assess the genetic diversity and population structure of the target species, opening a new approach for mitochondrial genome research in fish species.

❖ **Scientific and practical significance:**

This research significantly advances our understanding of population genetics, migration patterns, and mitochondrial genomes in Mekong River fish species, providing a robust scientific foundation for effective conservation and sustainable management of regional fish resources.

Da Nang,, 2026

Supervisors

PhD. Student

Assoc.Prof. Dr. Dang Thuy Binh Dr. Ngo Thai Bich Van Truong Thi Oanh