

DNA Barcoding of Southeast Asian Aquarium Fish

Ng Yi Hui Eunice (U042779U)

Supervisor: A/P Rudolf Meier



Introduction

Traditionally, biologists have largely relied on morphological tools and diagnoses to distinguish the species but phenotypic plasticity and intra-specific variability can lead to incorrect identifications. DNA barcoding is the current “in” thing for identifying species. Cytochrome c oxidase subunit I (COI) gene was the proposed “barcode” for the global identification of animals.

Applications

- ✓ Manage fisheries for long-term sustainability
- ✓ Improve ecosystem research and conservation
- ✓ Identify endangered species in the aquarium trade
- ✓ Detect invasive species

Objectives

This project aims to assess the reliability and effectiveness of DNA barcoding (using *COI* and *cytochrome b* genes) in identifying species of (mostly) Southeast Asian freshwater fish; and to compare the differential ability of *COI* and *cytochrome b* gene to correctly identify fish species.

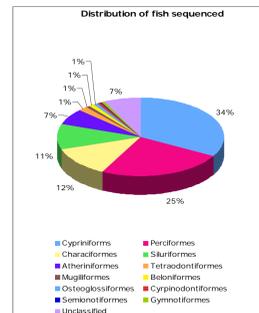
Materials and methods

DNA was extracted from a total of 267 specimens. 655bp were amplified from the 5' region of the mitochondrial *COI* gene and the second half of the *cytochrome b* gene. Purified PCR products were sequenced bidirectionally. Genbank sequences for 12 fish Orders were downloaded. All sequences were aligned in Alignment helper. Analysis was performed with taxonDNA alpha 1.5 – version 9.

Results and Discussion

Data obtained:

- Total of 631 COI and 1317 cytochrome b sequences.
- 513 (COI) and 778 (cytochrome b) named species
- 23 COI sequences have an allospecific identical match (3.6%).
- 61 cytochrome b sequences have an allospecific identical match (4.6%).



Best match/Best Close Match

| Algorithm | Successful identification | Ambiguous | Misidentification |
|--|---|---|---|
| Best Match Find closest match for each query and compare the species names. | Both sequences are same species | Several equally good matches from different species | Both sequences are different species |
| Best Close Match Plot relative frequency of intraspecific distances → determine threshold value below 95% of all intraspecific distances are found. | name of query and closest sequence is identical | Several equally matched sequences belong to two or more species | name of query and its closest sequence are mismatched |

Cytochrome b distances = 0 - 2%



Cobitis paludica



Cobitis vettonica



Cobitis elongatoides



Cobitis bilineata



Cobitis taenia



Cobitis macrostigma

Figure 1: Identification success based on “best close match”.

| gene | "Best Match" | | |
|--------------|--------------|-----------|-------------------|
| | Success | Ambiguous | Misidentification |
| COI | 25.51% | 5.54% | 68.93% |
| cytochrome b | 41.34% | 6.86% | 51.79% |

3.6% of all COI sequences and 4.1% of cytochrome b sequences have an allospecific identical match.

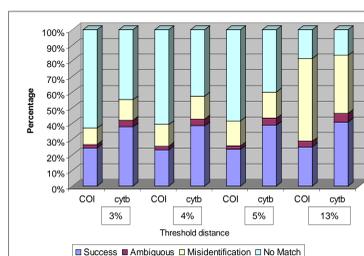


Figure 2: Identification success based on “best match”.

Intraspecific distance > 25%



Microrasbora kubotai



Microrasbora kubotai



Microrasbora kubotai

Fishes whose Intraspecific distance > interspecific distance



Hypessobrycon socolofi



Hypessobrycon erythrostigma



Hypessobrycon sweglesii



Hypessobrycon herbertaxelrodi



Oreochromis niloticus



Oreochromis aureus



Oreochromis mossambicus



Mugil cephalus



Mugil curema



Mugil platanus



Melanotaenia nigrans



Melanotaenia nigrans



Melanotaenia parkinsoni



Melanotaenia maccullochi

Pairwise distances for intra- and interspecific sequences

There is extensive overlap between intra- and interspecific pairwise distances

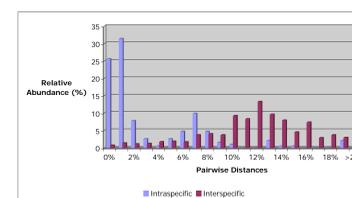


Figure 3: Overlap between intra- and interspecific genetic variability for congeneric COI sequences.

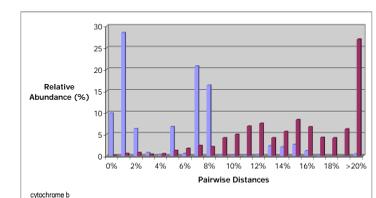


Figure 4: Overlap between intra- and interspecific genetic variability for congeneric cytochrome b sequences.

Conclusion

- ✓ Cytochrome b gene is slightly more successful than COI
- ✓ BUT overall identification success rates are alarmingly low → DNA barcodes cannot be used to identify freshwater fish species?
- ✓ A single gene cannot sufficiently generate a single DNA barcode per species.
- ✓ Freshwater fishes are very diverse, and thus difficult to delimit species.