Sains Malaysiana 53(1)(2024): 11-21
http://doi.org/10.17576/jsm-2024-5301-02

The Viability and Potential of Environmental DNA (eDNA) Detection of Freshwater Fish Based on Current Genetic Resources in Malaysia

(Daya Maju dan Potensi Pengesanan DNA Persekitaran Ikan Air Tawar Berdasarkan Sumber Genetik Terkini di Malaysia)

Kaviarasu Munian[1,2], Farah Farhana Ramli[1], Nursyuhada Othman[1], Hidayah Haris[1], Nur Hartini Sariyati[1], Nur Aina Amira Mahyudin[2], Mohd Faudzir Najmuddin[1], Mohd Shahfiz Azman[2] & Muhammad Abu Bakar Abdul-Latiff[1,*]

[1]Environmental Management and Conservation Research Unit (eNCORe), Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (Pagoh Campus), 84600 Muar, Johor, Malaysia
[2]Zoology Branch, Forest Biodiversity Division, Forest Research Institute Malaysia (FRIM), 52019 Kepong, Selangor, Malaysia

Received: 26 June 2022/Accepted: 12 December 2023

ABSTRACT

Environmental DNA (eDNA) metabarcoding is a promising tool for regular biological monitoring, especially for freshwater fish, which are facing tremendous threats worldwide. The application of eDNA detection is a dramatic improvement on common methods of biomonitoring as it produces tangible results in a short time with low effort and little expense. However, the accuracy of the technique is largely dependent on the availability of genetic references for the target organisms. In this study, we investigated the availability of genetic resources for freshwater fish in Malaysia in three public depositories, National Center for Biotechnology Information (NCBI), Barcode of Life Data System (BOLD), and Mitochondrial Genome Database of Fish (MitoFish), focusing on seven targeted genes of mitochondrial DNA. We found that only 68.6% of freshwater fish found in Malaysia had information on at least one of the seven targeted genes, with data on Cytochrome C Oxidase Subunit I being most commonly available. Genetic information for threatened and endemic species were underrepresented (33.3%-41.7%), yet fish of commercial value and invasive species were well explored genetically. Although there is still room for improvement to achieve comprehensive and reliable genetic resource information for freshwater fish in Malaysia, the application of eDNA metabarcoding is still highly relevant. This is since the current decline in freshwater fish diversity in Malaysia is alarming and because the technique will assist in the ongoing effort to generate new genetic references for Malaysian freshwater fish.

Keywords: Environmental DNA; Cytochrome b; Cytochrome C Oxidase Subunit I; freshwater fish; metabarcoding and Next-Generation Sequencing (NGS)

ABSTRAK

Metabarkod DNA persekitaran (eDNA) ialah kaedah yang berpotensi dalam memantau sumber biologi terutamanya dalam pemantauan ikan air tawar terancam di seluruh dunia. Penggunaan teknik eDNA dalam pemantauan sumber biologi merupakan kaedah berteknologi tinggi kerana ia mampu memberikan hasil yang ketara dalam masa yang singkat dengan penggunaan tenaga dan kos yang minimum. Walau bagaimanapun, ketepatan teknik ini sebahagian besarnya bergantung kepada ketersediaan rujukan genetik untuk organisma yang disasar. Dalam penyelidikan ini, kami mengkaji ketersediaan sumber genetik untuk ikan air tawar di Malaysia di tiga depositori awam, iaitu Pusat Maklumat Bioteknologi Kebangsaan, Sistem Data Kehidupan Kod Bar dan Pangkalan Data Genom Mitokondria Ikan dan memfokuskan kepada tujuh gen DNA mitokondria yang terpilih. Hasil keputusan kami menunjukkan hanya 68.6% daripada ikan air tawar yang ditemui di Malaysia mempunyai maklumat tentang sekurang-kurangnya satu daripada tujuh gen yang disasarkan, manakala gen Sitokrom C Oksidase Subunit I merupakan gen yang paling tersedia untuk digunakan. Keputusan kami juga menunjukkan maklumat tentang genetik spesies terancam dan endemik masih kurang

dikaji (33.3%-41.7%), tetapi ikan air tawar komersial dan spesies ikan invasif telah dikaji dengan lebih baik dari aspek genetik. Fenomenon penurunan kepelbagaian ikan air tawar di Malaysia yang membimbangkan pada masa kini memberikan kewajaran kepada pelaksanaan aplikasi metabarkod eDNA untuk kajian ikan air tawar. Walaupun masih terdapat ruang penambahbaikan untuk mencapai maklumat sumber genetik yang komprehensif dan boleh dipercayai di Malaysia, aplikasi metabarkod (eDNA) ini wajar diteruskan sebagai sebahagian usaha berterusan para pengkaji dalam menjana rujukan genetik baharu untuk ikan air tawar Malaysia.

Kata kunci: DNA persekitaran; ikan air tawar; penjujukan generasi hadapan (NGS); Sitokrom b; sitokrom Oksidase I

Introduction

On the Earth's surface, although the proportion of freshwater is relatively small, freshwater habitats are home to an enormous variety of species (Dudgeon et al. 20006). These habitats are, of course, also used extensively by humans for a wide range of ecosystem services. The decline in biodiversity is far greater in freshwater habitats than even the most affected terrestrial ecosystems (Sala et al. 2000). Over 10,000 fish species live in freshwater (Lundberg et al. 2000), constituting approximately 40% of global fish diversity and one quarter of global vertebrate diversity. Despite the vast freshwater fauna, such ecosystems are under great threat from over exploitation, water pollution, flow modification, destruction or degradation of habitat, and invasion by exotic species (Allan & Flecker 1993; Jackson et al. 2001; Naiman et al. 1995; Revenga et al. 2005).

Many efforts have been made to prevent further declines in freshwater ecosystems, including assembling species information through regular biological monitoring. Biological monitoring approaches are usually designed for a specific group of organisms, involving capture and observation. However, this method is time consuming, expensive, and limited by the number of trained personnel available (Darling & Mahon 2011). According to Shaw et al. (2016), conventional capture-based surveys increase the probability of predation risk to the organisms concerned and may damage the entire ecosystem. The conventional methods are also limited by the scarcity of taxonomic expertise and by the non-standardized skill levels of different taxonomists, leading to incomplete taxonomic identification. In addition, accurate assessment of the entire community may not be achieved since the methods used may fail to detect small or elusive species (Diener et al. 2017).

Environmental DNA (eDNA) metabarcoding is a technique used for species identification that has a very minimal impact on the ecosystem concerned (Bohmann et al. 2014; Schnell et al. 2012). Biodiversity assessment can be achieved using environmental samples, such as water, sediment, or air, by extracting the genetic material present. The extracted DNA can then be amplified using general or universal primers by polymerase chain reaction and sequenced using next-generation sequencing (NGS). The eDNA technique is currently recognized as the most efficient in fish research, particularly for assessing fish species richness in freshwater environments (McElroy et al. 2020). Ficetola et al. (2008) were the first to demonstrate the usefulness of eDNA to detect the presence of an aquatic vertebrate in freshwater, and they also demonstrated the applicability of eDNA metabarcoding in helping to document the presence of invasive fish species for the purpose of genetic conservation. Compared with conventional methods, eDNA metabarcoding has the potential to describe species, communities, biodiversity, interactions, and functional ecology over large spatial scales. However, the fundamental drawback of eDNA metabarcoding is the requirement for a genetic reference library to match discovered sequences to species. There is also the risk of incorrect readings owing to contamination (Stoeckle, Das Mishu & Charlop-Powers 2008).

Genetic references are crucial to eDNA metabarcoding to choose the most suitable primers for species identification. Taxon identification utilizes a standardized genetic marker in NGS to analyze different targeted groups (Kounosu et al. 2019) in a single sample. It is feasible to improve species detection by combining numerous primer sets with complementary reference libraries (Jarman et al. 2013). Universal primers can amplify a wide range of taxa by reference to highly conserved DNA regions, but the taxonomic resolution of these extensive markers is poor in most cases (above order level) (Jarman et al. 2013) compared with taxonomic primers that are intended to produce high taxonomic resolution (to genus or species

level) (Kartzinel et al. 2015). Using combined primer sets is costly if a larger group or ecosystem is involved, thus universal primers are advantageous. For fish assessment, Tedesco et al. (2017) referred to a species database and concluded that the ideal primer for fish genetics-based species identification was that for the 16S ribosome RNA (16S rRNA) gene developed by McInnes (2017).

An effective fish monitoring technique is critical for large-scale biodiversity monitoring because it would have the capacity to address integrated environmental concerns. The lack of comparative molecular studies on the ecological aspects of freshwater fish in Malaysia has resulted in a large gap in fish genetic resources for future reference. The eDNA metabarcoding method will become the preferred method to assess freshwater fish as it can resolve to species level (Gehri et al. 2021; Ma et al. 2022; Milhau et al. 2021; Nakagawa et al. 2018). Thus, considering fish research in Malaysia, this study was conducted to review the genetic resources currently available for seven targeted regions of mitochondrial DNA (mtDNA) held at three publicly available databases: the National Center for Biotechnology Information (NCBI) (NCBI 2021); the Barcode of Life Data System (BOLD) system (BOLD 2021); and Mitochondrial Genome Database of Fish (MitoFish) (Mitofish 2021). The assessment of inclusiveness of genetic resources focused on a variety of concerns, including available species, families of interest, threatened and endemic species, and commercial fish within Malaysian territory. The outcome will be beneficial for the application of eDNA metabarcoding as a tool for freshwater fish biodiversity monitoring.

Materials and Methods

For the purpose of this study, we defined freshwater fish as species that are capable of surviving solely in freshwater habitats. Additionally, we included species that can tolerate different environmental conditions at certain stages of their life cycle, such as being able to tolerate brackish water or even saltwater temporarily. We generated the full checklist of freshwater fish species from the Fishbase database (Froese & Pauly 2021). This was then screened and filtered to produce a list that included only species occurring in Malaysia. To obtain related sequences, we used the species names as search strings in three public databases: NCBI, BOLD v4, and MitoFish. Based on seven loci of mtDNA, namely cytochrome c oxidase subunit 1 (COI), displacement loop (D-loop), cytochrome b (Cytb), 12S ribosoma RNA (12S rRNA), 16S rRNA, transfer ribonucleic acid (tRNA), and NADH, we searched for and counted the sequences independently from the three databases. We selected these seven loci because mtDNA provides a single independent marker for population genetic studies, a major advantage especially for eDNA studies since the genome exists in several copies in each cell (Allentoft et al. 2012). Also, mtDNA appears to degrade at a slower rate than nuclear DNA, an important consideration in genetic analysis (Schwarz et al. 2009).

After searching the related sequences based on a list of freshwater fish in Malaysia, we also categorized each species on its conservation status using The International Union for Conservation of Nature (IUCN) Red List of Threatened Species, as follows: NE, DD, LC, NT, VU, EN, CR, Extinct in the Wild, and Extinct. Information on commercial value and endemicity of each species was also collated using data obtained from Fishbase. For each of the species, we retrieved information including locus of sequences, conservation status, endemicity, and commercial value from the public databases. Since Microsoft Excel had been used in the data sorting, it was also employed to assess trends in research associated with genetic resources of freshwater fish. All the data were visualized graphically (as histograms) to evaluate the categorized data.

Results and Discussion

Genetic resources for freshwater fish species in Malaysia

A total of 605 species of freshwater fish were extracted from Fishbase (Froese & Pauly 2021), and each species was checked for its occurrence in Malaysia. It was found that approximately 80% of these species have been documented either in various publications or as voucher specimens that have been deposited into relevant authorities such as universities, government departments, or non-governmental organizations either in Malaysia or other nearby countries. The list of fish species utilized in this study from FishBase might not have been exhaustive considering the parameters used for the database search, and we would like to caution that the taxonomic nomenclature may vary due to ongoing research and expanding knowledge on freshwater fish.

After screening each species individually through the three selected databases, a total of 1,430 records of genetic information on the seven targeted genes were obtained. Of the 605 species assessed, 415 (68.6%) had

information on one or more of the seven targeted genes. Conversely, the remaining 190 species, representing 31.4% of the total, lacked any information on their mitochondrial DNA (mtDNA). Among the seven targeted genes, information on the COI gene was available for most of the species (364), followed by the Cytb gene (276), and 12S ribosomal RNA (12S) with 230 species (Table 1). The remaining targeted genes were represented by between 90 and 200 known species in Malaysia, with the transfer ribonucleic acid (tRNA) gene being the least frequently available.

In Malaysia, 60.2% of freshwater fish species have genetic sequences for COI, with 39.8% of the species not being evaluated. For Cytb, data existed for 45.6% of known species in Malaysia, and for the 12S gene, the figure was 38%. For the remaining genes of interest, data were available for relatively few, ranging from 16% to 36.2% of Malaysian species. There were 56 species for which data existed for all seven targeted genes. Among these, fish in the genus Pangio and Homaloptera were the most studied species with complete data on four species each, while there were three species in the genus Puntius with complete information on the target genes. The remaining 45 species all belonged to different genera. Among the 190 species with no data on mtDNA, 22 species in the genus Gastromyzon were the least studied species with respect to genetic resources.

VARIATION IN GENETIC INFORMATION AMONG FISH FAMILIES IN MALAYSIA

The 605 freshwater fish species examined belonged to 66 families. The genetic resources available from the three selected databases were unevenly distributed across these families (Figure 1). The family Cyprinidae had the most genetic information, with 375 data records for 161 species for at least one of seven targeted genes. Next was the family of gourami (Osphronemidae) with 148 sequences, followed by Gobiidae with 85. Two families of investigated freshwater fish, Sundasalangidae and Synanceiidae, had no records for any sequences for the targeted mtDNA genes. Almost 97% of 66 families were assessed for their genetic information. Among the 66 families examined, nine families had <50% of their species genetically studied. The least studied families were Chaudhuriidae and Syngnathidae for which only 25% of their species had available data. Genetic information for the families Siluridae and Balitoridae, with 30 and 70 species of interest, had data on only 27% and 32% of species.

TABLE 1. Summary of species and family distribution according to seven targeted genes of mitochondrial DNA

Genes of Interest of Mitochondrial DNA	Family	Species	Coverage (%)
Cytochrome C Oxidase Subunit I (COI)	62	364	25%
12S ribosomal RNA (12S)	49	230	16%
16S ribosomal RNA (16S rRNA)	50	219	15%
NADH Dehydrogenase	32	149	10%
Transfer ribonucleic acid (tRNA)	25	95	7%
Displacement loop (D-loop)	26	97	7%
Cytochrome b (Cyt b)	45	276	19%

GENETIC INFORMATION OF FISH SPECIES WITH CONSERVATION IMPORTANCE

Of the 605 species of freshwater fish surveyed, 13.6% had an IUCN conservation status of Not Evaluated (NE) (Figure 2). Just over half of the species examined were categorized as Least Concern (LC) (57.7%). There were 64 (10.6%) species of freshwater fish under threat. The remaining species were either categorized as Data Deficient (DD) or Near Threatened (NT). There were 82 species whose conservation status was NE, but half of these had genetic information in the databases related to targeted mtDNA genes. Similarly, of those species categorized as DD, slightly more than half had their genetic information recorded. Of 349 species categorized as LC, 79.9% had mitochondrial gene information recorded. Of the fish categorized as NT, slightly less than 50% of the assessed species had targeted genes were recorded.

Those species categorized as Vulnerable (VU), Endangered (EN), and Critically Endangered (CR) require immediate conservation action. Genetic information on threatened fish species in Malaysia is lacking. Of 32 species categorized as VU, only 68.8% were genetically examined and had sequences based on mtDNA. Meanwhile, 23 of the 605 species assessed were categorized as EN, of which 65.2% had recorded genetic information. Those species categorized as CR had very little recorded genetic data, with information only available for 33.3% of the nine species identified. Overall, regarding species under concern (especially threatened fish), 37.5% still had no recorded genetic information, either for nuclear or mitochondrial genes.

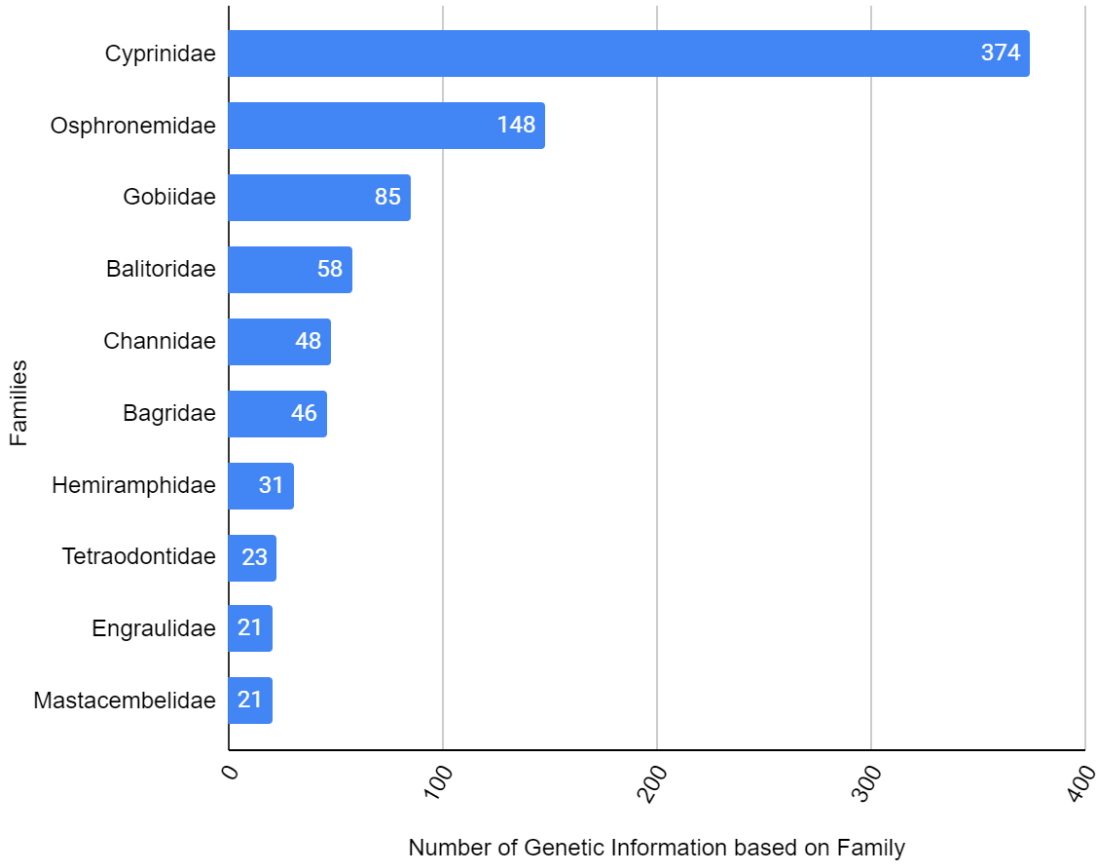

FIGURE 1. Distribution of genetic resources of freshwater fishes in Malaysia according to the fish families

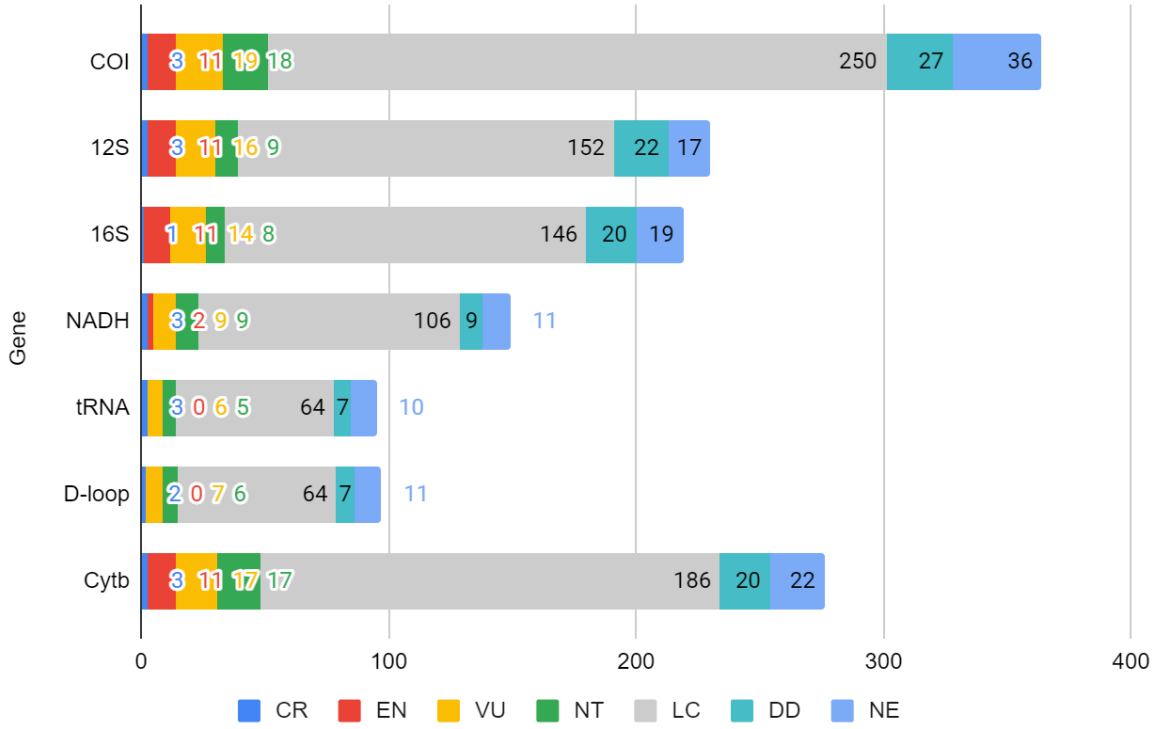

FIGURE 2. The availability of genetic information on threatened freshwater fishes in Malaysia territory based on IUCN Red List of Threatened Species according to targeted genes of mitochondrial DNA

ENDEMIC SPECIES AND DNA SEQUENCE RESOURCES

Of the 605 fish species assessed, 108 were endemic to Malaysia, for which genetic information is available for only 45 (41.7%), with no information available at all for the remaining 63 (Figure 3). Among the 108 endemic species, 25 were under threats with genetic resources available for only 56%. Species categorized as CR were the least studied among the endemic species, with only one of the four species having genetic sequences recorded.

GENETIC INFORMATION ON COMMERCIAL FISH AND INTRODUCED SPECIES

Commercial activities such as fisheries, aquaculture, and aquaria were combined and categorized into four groups based on activity intensities: Species of no interest, least commercialized, moderately commercialized, and highly commercialized. Of the 605 freshwater Malaysian species examined, 206 had been commercialized at various levels of intensity. Least commercialized species constituted only 5% of commercialized fish in Malaysia, and almost all the species had their DNA sequences generated and available in public databases. Most commercialized fish in Malaysia can be categorized as moderately commercialized, and these constitute 85.9% of all commercialized species. Of these, 91.5% had their genetic information available, with only 15 yet to have their mtDNA examined. One of these species was endemic to Malaysia and had no genetic sequences available.

The remaining 19 commercialized species were categorized as highly commercialized, with all having genetic resources on targeted genes available. However, six of these were facing threats. In addition to the referenced 605 species, we analyzed additional information on 30 non-native species based on the checklist produced by Saba et al. (2020). Among these species, 7 were classified in the Cichlidae family and 6 in the Cyprinidae family, making them among the highest recorded non-native

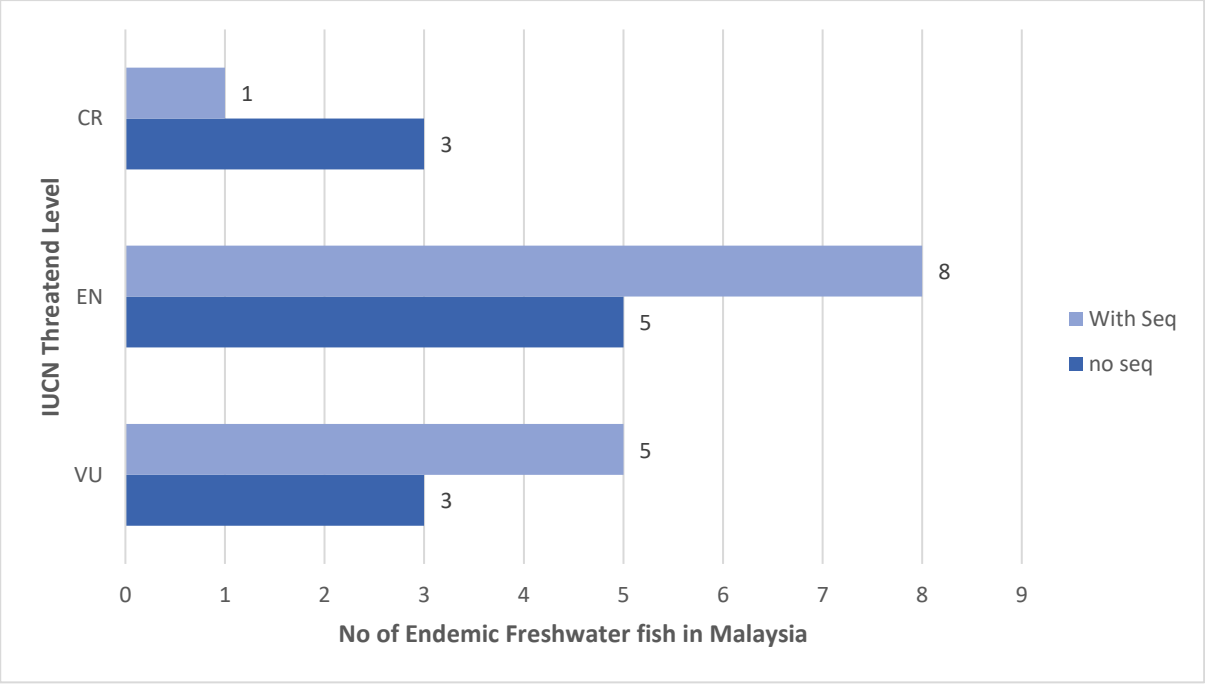

FIGURE 3. The status of endemic freshwater fishes that were classified as under threatened by IUCN Red List

fish families in Malaysia. Interestingly, all of these identified non-native fish species had complete genetic information available for the seven targeted genes. This comprehensive genetic data could prove invaluable for applying the latest methods in non-native species detection and monitoring within local habitats.

The lack of comprehensiveness of genetic resources on freshwater fish in Malaysia is a concern. The largest group of animals for which genetic data is available and on which most eDNA metabarcoding studies have been carried out is fish (Othman et al. 2021). This allows for practical and cost-effective monitoring of fish species using eDNA. Nevertheless, to produce valid and effective outputs and to formulate efficient mitigation programs for conservation, having complete genetic references is crucial. This avoids any misidentification and allows for assembly with data from other ecological aspects, such as relative abundance, presence of endemic and threatened species, and occurrence of invasive species. In general, the genetic resources available for freshwater fish in Malaysia are still far from complete, and it will require much work before they can be reliably utilized in eDNA detection technology.

While there were vast disparities among the seven targeted genes, information on the COI gene exists for all examined fish. This high level of coverage is probably due to the high number of studies being carried out in Malaysia on COI and because of the gene's central function. For DNA barcoding within the animal kingdom, COI has become a marker of choice (Hebert, Ratnasingham & De Waard 2003). Pentinsaari et al. (2016) reported that COI is by far the most extensively sequenced gene region, with roughly 4.7 million DNA barcode sequences, largely applied for the identification and discovery of animal species. Following this formative idea, studies related to evaluation and documentation of new genetic references for species in Malaysia has led to many more studies using COI than any of the other six targeted genes. In addition, biodiversity in Malaysia is immense, and new species discoveries are increasing yearly, directly contributing to the increase in COI data availability as the most consistent information in evaluating these new findings. It is reported that tRNA genes are useful in deducing deep-level phylogeny (Chen et al. 2019; Liu et al. 2017; Wang et al. 2017), however, the use of tRNA in Malaysia is still unpopular, and very little data on

tRNA is available for freshwater fish. A plausible reason for the tRNA gene being studied so little might be due to unwillingness among scientists and researchers to explore other markers, and a tendency to opt for genetic sequences on which most available information is available, such as COI. Although COI is a popular marker with the most abundant information, there have been arguments on its limitations and suitability in taxonomy (Deagle et al. 2014; de Carvalho 2014; Liu et al. 2017), which could be overcome by using different markers for different groups of animals. A wider choice of gene markers could help increase the genetic resources of freshwater fish in Malaysia, besides addressing the broader scope of research, including evolution, population structure, and biogeography. With increased genetic information, application of eDNA barcoding could be diversified via a wider range of markers to answer currently challenging questions.

Of the freshwater fish in Malaysia, the family with most genetic information available is Cyprinidae; this family is also the largest group of fish in Malaysian territory (Ng et al. 2017). Cyprinids are widely distributed in various types of habitats, ranging from mountains to lowlands. The availability of genetic resources varies widely across families, depending on species. Since many species of interest for commercialization, as a protein source for humans, for recreational fishing, and as aquarium pets are Cyprinids, more studies have been carried out on species in this family than in others. Only two single-species families have not been studied genetically. Even though these species were first reported in 1981, no research on their genetics had been done until lately. This scenario may be explained by the trends in genetic research in Malaysia. A quick search (August 10, 2021) in Scopus for publications using the search string 'Genetic, Malaysia, AND Fish' produced 2,397 results, with most of the studies dated after the year 2000. The first DNA sequencing began after the year 1977 when Sanger, Nicklen and Coulson (1977) successfully amplified the double-helix structure of DNA. Prior to 1977, species descriptions were based on taxonomic features. Consequently, species described much earlier, in the eighteenth and nineteenth centuries, might not have been subjected to genetic study, especially monophyte fish in particular families. Many precursor vouchers of fish were not available for evaluation by local researchers or scientists as most were deposited or kept outside of the country. Efforts should be taken to catalog precursor vouchers using current technologies in DNA extraction from ancient samples kept in museums and depositories to generate reference sequences.

Concern for threatened and endemic fish is high, as these species are susceptible to extinction following progression of habitat loss and degradation, pollution, and over-exploitation (Chong, Lee & Lau 2010; Saba et al. 2020; Tarkan, Marr & Ekmekçi 2015). Genetic references for threatened and endemic fish in Malaysia are urgently required. The lack of genetic information for threatened and endemic species might have arisen from the low number of individuals available to be assessed. Although biodiversity in Malaysia is high, many species are rare in the wild or occur in isolated or inaccessible habitats, hindering efforts to study them and to generate genetic references. Yet information on these threatened and endemic species is vital if we are to accurately infer viable population sizes and gene flow, which are valuable in formulating holistic conservation and management programs. Threatened and endemic species should be considered as the main priority for inclusive survey and study. Information generated from extensive studies would help to fill knowledge gaps.

Malaysia is equipped with comprehensive and dynamic policies, as well as action and management plans, which can all be effectively implemented at all levels of the country's administration. However, there are a number of flaws in the current conservation strategies that need to be remedied immediately now, such as the accessibility of genetic resources. Several factors were identified as causes for the lack of comprehensiveness of genetic resources in Malaysia, one being a decline in taxonomy expertise. Taxonomy is the science of biological classification, which includes identification as well as description of species (Singh, Janso & Brady 2007). Coleman (2015) reported that the number of taxonomic experts worldwide was in decline, while Waldron et al. (2013) concluded that lack of interest and funding in the taxonomy field were the reasons for the lack of expertise, especially with respect to freshwater fish. Similar trends have been seen in Malaysia, where more students preferred to take up courses on biotechnology and in medicine-related fields, with taxonomy and taxidermy apparently dying off. The shortage of experts to identify and verify current and newly discovered species could hinder the generation of information valuable to science and conservation. Similarly, prolonged inconsistency in taxonomic nomenclature and the complexity of cryptic species has an adverse effect on validation of genetic resources available in various depositories. For instance, the complex taxonomic statuses of several species in the Cyprinid genus Tor

remain unresolved as contradictory data have been presented by different researchers (Kottelat 2013; Roberts 1999). Voucher specimens collected from the field serve an important role as references for species identification and validation to ensure the reliability of publication (de Moor 1996). It is unfortunate that a country as species rich as Malaysia has no depository center for freshwater fish specimens. According to Ng et al. (2017), specimens of freshwater fish are to be found at locations scattered around the country, making it problematic for researchers to find specimens for further study.

The lack of genetic resources in Malaysia can be solved by addressing core problems. In terms of freshwater fish study, focus should be given to all species, ensuring that there is no bias toward certain 'interesting' species or species with commercial value. Threatened and endemic species should be the major concern in generating genetic information to ensure the viability of the species. A continuous funding resource should be established for generating genetic resources for freshwater fish in Malaysia. The Malaysian government, universities, and NGOs should provide research grants for researchers and scientists in the country to work intensively on freshwater fish to find insights on aspects of ecology, behavior, and genetics. A national digital sequencing database for genetic resources would aid in the expansion of genetic data sources, particularly for freshwater fish. Good networking between all research institutions in the country could be realized, and contributions of genetic information for all the species inhabiting in the country would steadily increase.

Conclusions

The application of environmental DNA (eDNA) to monitor the diversity of freshwater fish is still relevant in Malaysia territory. Although genetic resources on freshwater fish are still growing, the benefits of eDNA application are far too great to be overlooked, particularly in addressing the alarming state of targeted species in Malaysia. A holistic management and conservation could be achieved by applying both conventional and eDNA approaches and preserve the invaluable biological diversity that exits in the country.

Acknowledgements

This project is funded by the Ministry of Higher Education (MOHE), Malaysia under the Fundamental Research Grant Scheme FRGS/1/2020/WAB11/KATS//1 (FRGS-UTHM-K353), UTHM-GPPS-H702 postgraduate grant by Universiti Tun Hussein Onn Malaysia and RMk-12 development project under P23085100210003. Authors acknowledge the MOHE, Universiti Tun Hussein Onn Malaysia and Forest Research Institute Malaysia (FRIM) for providing necessary funding,facilities, and assistance. Lastly, the authors would like to thank Enago (www.enago.com) for the English language review.

References

Allan, J.D. & Flecker, A.S. 1993. Biodiversity conservation in running waters. *BioScience* 43(1): 32-43.

Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos, P.F., Samaniego, J.A., Gilbert, M.T.P., Willerslev, E. & Zhang, G. 2012. The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences* 279(1748): 4724-4733.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Douglas, W.Y. & De Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* 29(6): 358-367.

Centre for Biodiversity Genomics, Canada. 2021. Barcode of Life Data System. Bold Systems. https://www.boldsystems.org/ Accessed 10 August 2021.

Chen, C., Li, Q., Fu, R., Wang, J., Xiong, C., Fan, Z., Hu, R., Zhang, H. & Lu, D. 2019. Characterization of the mitochondrial genome of the pathogenic fungus *Scytalidium auriculariicola* (Leotiomycetes) and insights into its phylogenetics. *Scientific Reports* 9: 17447.

Chong, V.C., Lee, P.K.Y. & Lau, C.M. 2010. Diversity, extinction risk and conservation of Malaysian fishes. *Journal of Fish Biology* 76(9): 2009-2066.

Coleman, C.O. 2015. Taxonomy in times of the taxonomic impediment–examples from the community of experts on amphipod crustaceans. *Journal of Crustacean Biology* 35(6): 729-740.

Darling, J.A. & Mahon, A.R. 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research* 111(7): 978-988.

de Carvalho, C.B.V. 2014. DNA barcoding in forensic vertebrate species identification. *Brazilian Journal of Forensic Sciences, Medical Law and Bioethics* 4(1): 12-23.

De Moor, F.C. 1996. The importance of voucher specimens. *Southern African Journal of Aquatic Science* 22(1-2): 117-118.

Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F. & Taberlet, P. 2014. DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters* 10(9): 20140562.

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N. & Pfrender, M.E. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* 26(21): 5872-5895.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L. & Sullivan, C.A. 2006. Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews* 81(2): 163-182.

Ficetola, G.F., Miaud, C., Pompanon, F. & Taberlet, P. 2008. Species detection using environmental DNA from water samples. *Biology Letters* 4(4): 423-425.

Froese, R. & Pauly, D. 2017. FishBase. www.fishbase.org Accessed 10 August 2021.

Gehri, R.R., Larson, W.A., Gruenthal, K., Sard, N.M. & Shi, Y. 2021. eDNA metabarcoding outperforms traditional fisheries sampling and reveals fine-scale heterogeneity in a temperate freshwater lake. *Environmental DNA* 3(5): 912-929.

Hebert, P.D., Ratnasingham, S. & De Waard, J.R. 2003. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(suppl_1): S96-S99.

Jackson, J.B., Kirby, M.X., Berger, W.H., Bjorndal, K.A., Botsford, L.W., Bourque, B.J., Bradbury, R.H., Cooke, R., Erlandson, J., Estes, J.A. & Hughes, T.P. 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293(5530): 629-637.

Jarman, S.N., McInnes, J.C., Faux, C., Polanowski, A.M., Marthick, J., Deagle, B.E., Southwell, C. & Emmerson, L. 2013. Adélie penguin population diet monitoring by analysis of food DNA in scats. *PLoS ONE* 8(12): e82227.

Kartzinel, T.R., Chen, P.A., Coverdale, T.C., Erickson, D.L., Kress, W.J., Kuzmina, M.L., Rubenstein, D.I., Wang, W. & Pringle, R.M. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences* 112(26): 8019-8024.

Kottelat, M. 2013. The fishes of the inland waters of Southeast Asia: A catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. *Raffles Bulletin of Zoology* 2013(Supplement No. 27): 1-663.

Kounosu, A., Murase, K., Yoshida, A., Maruyama, H. & Kikuchi, T. 2019. Improved 18S and 28S rDNA primer sets for NGS-based parasite detection. *Scientific Reports* 9(1): 1-12.

Liu, H., Li, H., Song, F., Gu, W., Feng, J., Cai, W. & Shao, R. 2017. Novel insights into mitochondrial gene rearrangement in thrips (Insecta: Thysanoptera) from the grass thrips, *Anaphothrips obscurus*. *Scientific Reports* 7: 4284.

Liu, J., Jiang, J., Song, S., Tornabene, L., Chabarria, R., Naylor, G.J. & Li, C. 2017. Multilocus DNA barcoding–species identification with multilocus data. *Scientific Reports* 7: 16601.

Lundberg, J.G., Kottelat, M., Smith, G.R., Stiassny, M.L. & Gill, A.C. 2000. So many fishes, so little time: An overview of recent ichthyological discovery in continental waters. *Annals of the Missouri Botanical Garden* 87(1): 26-62.

Ma, X., Yang, H., Zhong, X., Zeng, P., Zhou, X., Zeng, S., Dong, X., Min, W. & Huang, F. 2022. eDNA metabarcoding analysis of the composition and spatial patterns of fish communities in the Sanbanxi Reservoir, China. *Sustainability* 14(20): 12966.

McElroy, M.E., Dressler, T.L., Titcomb, G.C., Wilson, E.A., Deiner, K., Dudley, T.L., Eliason, E.J., Evans, N.T., Gaines, S.D., Lafferty, K.D. & Lamberti, G.A. 2020. Calibrating environmental DNA metabarcoding to conventional surveys for measuring fish species richness. *Frontiers in Ecology and Evolution* 8: 276.

McInnes, J.C. 2017. The development and application of DNA metabarcoding to non-invasively assess seabird diets, using albatrosses as a model. Doctoral dissertation, University of Tasmania (Unpublished).

MitoFish: Mitochondrial Genome Database of Fish. 2021. MitoFish. http://mitofish.aori.u-tokyo.ac.jp/ accessed on 10 August 2021.

Milhau, T., Valentini, A., Poulet, N., Roset, N., Jean, P., Gaboriaud, C. & Dejean, T. 2021. Seasonal dynamics of riverine fish communities using eDNA. *Journal of Fish Biology* 98(2): 387-398.

Naiman, R.J., Magnuson, J.J., Stanford, J.A. & McKnight, D.M. 1995. *The Freshwater Imperative: A Research Agenda*. Covelo: Island Press. p. 200.

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T. & Miya, M. 2018. Comparing local-and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology* 63(6): 569-580.

National Library of Medicine (US), National Center for Biotechnology Information. National Center for Biotechnology Information (NCBI). https://www.ncbi.nlm.nih.gov/ Accessed on 10 August 2021.

Ng, C.K.C., Abdullah, F., Biun, H., Ibrahim, M.K., Mustapha, S. & Sade, A. 2017. A working checklist of the freshwater fish diversity for habitat management and conservation work in Sabah, Malaysia, North Borneo. *Biodiversitas Journal of Biological Diversity* 18(2): 560-574.

Othman, N., Haris, H., Fatin, Z., Najmuddin, M.F., Sariyati, N.H., Md-Zain, B.M. & Abdul-Latiff, M.A.B. 2021. A review on environmental DNA (eDNA) metabarcoding markers for wildlife monitoring research. *IOP Conference Series: Earth and Environmental Science* 736(1): 012054.

Pentinsaari, M., Salmela, H., Mutanen, M. & Roslin, T. 2016. Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports* 6: 35275.

Revenga, C., Campbell, I., Abell, R., De Villiers, P. & Bryer, M., 2005. Prospects for monitoring freshwater ecosystems towards the 2010 targets. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360(1454): 397-413.

Roberts, T.R. 1999. Fishes of the cyprinid genus Tor in the Nam Theun watershed (Mekong basin) of Laos, with description of a new species. *Raffles Bulletin of Zoology* 47: 225-236.

Saba, A.O., Ismail, A., Zulkifli, S.Z., Halim, M.R.A., Wahid, N.A.A. & Amal, M.N.A. 2020. Species composition and invasion risks of alien ornamental freshwater fishes from pet stores in Klang Valley, Malaysia. *Scientific Reports* 10: 17205.

Sala, O.E., Stuart Chapin, F.I.I.I., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A. & Leemans, R. 2000. Global biodiversity scenarios for the year 2100. *Science* 287(5459): 1770-1774.

Sanger, F., Nicklen, S. & Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 74(12): 5463-5467.

Schnell, I.B., Thomsen, P.F., Wilkinson, N., Rasmussen, M., Jensen, L.R., Willerslev, E., Bertelsen, M.F. & Gilbert, M.T.P. 2012. Screening mammal biodiversity using DNA from leeches. *Current Biology* 22(8): R262-R263.

Schwarz, C., Debruyne, R., Kuch, M., McNally, E., Schwarcz, H., Aubrey, A.D., Bada, J. & Poinar, H. 2009. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Research* 37(10): 3215-3229.

Shaw, J.L., Clarke, L.J., Wedderburn, S.D., Barnes, T.C., Weyrich, L.S. & Cooper, A. 2016. Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* 197: 131-138.

Singh, M.P., Janso, J.E. & Brady, S.F. 2007. Cytoskyrins and cytosporones produced by *Cytospora* sp. CR200: Taxonomy, fermentation and biological activities. *Marine Drugs* 5(3): 71-84.

Stoeckle, M.Y., Das Mishu, M. & Charlop-Powers, Z. 2020. Improved environmental DNA reference library detects overlooked marine fishes in New Jersey, United States. *Frontiers in Marine Science* https://doi.org/10.3389/fmars.2020.00226

Tarkan, A.S., Marr, S.M. & Ekmekçi, F.G. 2015. Non-native and translocated freshwater fish. *FiSHMED Fishes in Mediterranean Environments* 3: 1-28.

Tedesco, P.A., Beauchard, O., Bigorne, R., Blanchet, S., Buisson, L., Conti, L., Cornu, J.F., Dias, M.S., Grenouillet, G., Hugueny, B. & Jézéquel, C. 2017. A global database on freshwater fish species occurrence in drainage basins. *Scientific Data* 4: 170141.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T. & Gittleman, J.L. 2013. Targeting global conservation funding to limit immediate biodiversity declines. *Proceedings of the National Academy of Sciences* 110(29): 12144-12148.

Wang, J., Zhang, L.I., Zhang, Q.L., Zhou, M.Q., Wang, X.T., Yang, X.Z. & Yuan, M.L. 2017. Comparative mitogenomic analysis of mirid bugs (Hemiptera: Miridae) and evaluation of potential DNA barcoding markers. *PeerJ 5*: e3661.

*Corresponding author; email: 'latiff@uthm.edu.my