Journal homepage: www.fishtaxa.com

Automated DNA Barcoding: A Computational Approach to Fish Species Identification

Fernando Q. López
South American Environmental University, Buenos Aires, Argentina.

Abstract

Molecular technologies like DNA metabarcoding, which offer a useful identifying tool for biomonitoring and conservation initiatives, are advantageous to biodiversity research. Accurate species-level categorization and taxonomic coverage depend on extensive DNA barcode reference libraries. Although the purpose of these libraries is to facilitate species identification, the accuracy of the barcode may be jeopardized if mistakes are made during the creation process. This uses the Barcode of Life Data System (BOLD) to automatically audit and annotate cytochrome c oxidase subunit I (COI) sequencing libraries for a particular taxonomic group of animals. Based on the characteristics of the data and the extent to which the species names correspond to sequences arranged in barcode index numbers (BINs), a qualitative grading system is then put into place, giving each species in the reference library one of five grades (A to E). In order to give researchers, the most accurate and practical data, we aim to find and classify records based on their congruency. To determine its limitations and usefulness, several tests were conducted. By rapidly scanning reference libraries to determine the congruence status of data and assisting in the sorting of ambiguous data for further analysis, BAGS fills a critical need in the present state of DNA barcoding research tools. As a result, BAGS may be a valuable tool in the subsequent DNA metabarcoding research, ultimately helping to raise the calibre and dependability of public reference libraries everywhere. Laboratory information management systems (LIMSs) have been developed to monitor processes via a process pipeline for molecular biology. This method uses consensus assembly for DNA barcoding, extraction, PCR, and cycle sequencing to monitor tissues. Importantly, a LIMS serving the DNA barcoding community must link specimen data to elements generated in the molecular lab that are required for public submissions (such primers and trace files). Here, we determine a procedure using Geneious bioinformatics software, from entering the specimen into the LIMS database to publishing the genetic data of the specimen in a public database. In order to provide structured reporting, post-processing annotation, and fully visible adjustments to reduce subjectivity and increase repeatability, the connections between workflow steps are maintained throughout the process.

Keywords: Automated DNA (ADNA), Computational Approach (CA), Fish Species Identification (FSI)

Citation: Fernando, Q. L. 2024. Automated DNA Barcoding: A Computational Approach to Fish Species Identification. FishTaxa 33: 11-20.

Introduction

A system constructed for identification in which a very short, standard DNA can be amplified easily and a very short DNA portion of the genome proves helpful in species identification. This process enables the identification of species rapidly and accurately. All of these procedures are under the term DNA barcoding technology. It is how to identify specimens by applying a short, known portion of DNA. The Barcode of every species is unique, just like the fingerprint of every person. These DNA barcodes are comparable to a library considered a reference so that an ID can be received. The process of automated sequencing is usually done so that large amounts of DNA can be sequenced(Zhang & Hanner, 2012). The principle of the Sanger chain termination method is applied to this phenomenon. dATP is not labelled using the actual method of Sanger; every dideoxynucleotide involved in the reaction is marked with different fluorescent markers. The use of DNA barcoding that involves

Accepted: 07 April 2024; Published: 13 June 2024

the recognition of new species, analyzation of food safely, the recognition and evaluation of cryptic species, detect all those species that are alien according to their characteristics, recognize all those species that come into the category of endangered and threatened species that make this possible that egg and larva become interconnected and transform into an adult specie and secure the rights that are of intellectual property(Zhang & Hanner, 2011). The procedure of generation of DNA barcodes involves the identification of species from a sample of tissue that is unknown and can be passed through the following steps. Sampling of tissue, lysis of tissue, extraction of DNA, and the production of polymerase chain reactions that involve the process of COI amplification. Cleaning of the procedure of polymerase chain reaction. Reactions in which cycle sequencing is involved. Sequence of cleanup of reaction. Let us discuss certain advantages of barcoding of DNA. It is an effective method that can be utilized to identify species. It is a worthy and reliable tool applicable in various ways involving forensic, food, and medical industries. It also benefits them in agriculture, environmental monitoring, and conservation processes(Zhang et al., 2012). The phenomenon of DNA barcoding is based on two steps. One is the construction of a DNA barcode library of all known species. Building a connection between the barcode sequence of the sample, which is not known against the library of the Barcode that is made for identification. It is the quality of automated systems that complete their tasks immediately and accurately, and no interference will be required. It proves helpful for reducing the time and resources needed for the sequencing of large amounts of DNA and allowing the researchers to do the experiments on more samples within a very short period. The vast range of discrepancies in the sequencing of cogeneric individuals creates many complications in the identification of species by applying DNA(Fontes et al., 2024). A fragment in which 650 base pairs are found can be seen in the cytochrome c oxidase subunit gene. All these things are related to Pethia conchonius. The application of 30 samples was observed. The identification of all these samples is usually done as a P.conchonius. It is seen that the other two species, which were represented as cogeneric and had a 2% divergence in sequence, can be seen in the 29 samples from which 30 were used. Two samples of P.comchonius are accumulated with the Pontius terio. By applying different computational methods, that sequel comes to an income named Puntius chola in the NCBI database as the representation of Pcomchonius. Some characteristics that can make a difference between the fishes involve the shape of their heads, the location of the mouths, the type of fin, and the location of fin and average size of adult people. Mentioning colours like the stripes that are found vertically and then spots of fins. These characteristics prove helpful in the differentiation of fish when they are represented along with other elements like the range of Geographical type. In the forensic setting, DNA analysis identified(Pereira et al., 2021). This method is very common in forensic settings. It proves very supportive in the reconstruction of crime in which the participation of animals is found. The situations of trafficking in protected species are also involved in this phenomenon. The application of DNA barcoding is of significant importance in identifying DNA. The dichotomous key is beneficial for the identification of families of fish. It lists the very specific and observable traits and the characteristics of many fish species. Regarding each trait, there is an infusion of a question with two possible answers from the key. Both of these feedbacks will lead to another question. The elaboration of DNA barcodes is usually done with the help of distance- and character-based methods(Blanco-Bercial et al., 2014). The first one uses the clustering of all those groups, which seems comparable. These traits are based on the genetic distance that is somehow connected with them. The second one depends on the presence and absence of minute nucleotides found in substituting some other thing. A limitation is found in the approach that is not related to distance while it defines a universal species of nature across the taxa. The level of variation or evolution cannot be seen as a constant factor throughout the taxa. The character-based approach elaborates it more accurately by applying a unique set of characters of nucleotides(Kim et al., 2010). Identification of species together DNA sequences, DNA barcoding, is widely used in various applications. The current barcoding techniques are typically built around one mitochondrial

location, like cytochrome c subunit I of the oxidase (COI)(Jin et al., 2021; Liu et al., 2017). This barcoding method is not always effective when applied to species separated by a short divergence time or containing introgressed genes from closely related species. This paper presents a more efficient multilocus barcoding method based on the capture of genes in combination with "next-generation" sequencing(Chakraborty et al., 2017).

Research Objective

The main objective of this research is to understand automated DNA Barcoding: A Computational Approach to Fish Species Identification. This topic elaborates on all the characteristics of DNA barcoding and the computational approach to identifying fish species.

Literature Review

Researchers claim that identifying species in marine ecosystems is difficult because of the diverse range of species found in marine ecosystems. terrestrial ecosystems are simpler and less complex than marine ecosystems. The coding technique is used in the research process to identify species of marine environments from their DNA (Akter et al., 2023).studies reveal that classifying Epinephelus species is difficult due to the similarity found in the closely related group of these species.in the coastal river of Nizampatnam, barcoding technology is employed for the grouper's identification process(Chatla et al., 2024).Studies claim that diversity found in fish species is related to their phenotypic and genotypic traits. Taxonomists classify species based on their taxonomic characteristics to carry out ecological studies. Caranx Senegallus species are identified in the Nigeria water bodies using DNA barcoding (Eriegha et al., 2024).Scholars suggest that DNA barcoding is an efficient method for assessing the taxonomically different speciest.to use bar code reference efficiently, a neural network based on AI technique is preferred in marine studies. The process of e-barcoding becomes automated using AI software, which reduces the chances of disordinance of barcode data(Fontes et al., 2024).studies show that Beibu Golf consists of an abundance of marine species classified using the barcoding system. The decline in the fish population in Beibu Golf is due to several environmental factors. this decline of species in marine ecosystems has increased the need for the development of automated systems for molecular identification of species(Jiang et al., 2023).studies reveal that the effectiveness of different species identification methods is compared to apply the most suitable method for species classification. single locus species delimitation methodology is a species classification method that determines the number and taxonomy of species present in a single lake(Karabanov et al., 2023). Making the classification a simplified process is done using computational technology. computerized barcoding strategy allows taxonomic classification of aquatic species. Using deep learning algorithms in a computerized system further improves the classification process by providing detailed information about aquatic species in the form of a database(Kasianchuk et al., 2024).Studies reveal that data regarding the species present in the cosmonaut sea is low. Very few species of cosmonaut sea have been identified due to the lack of information available about the species. The use of COI barcodes identified dermal fish fishes in the cosmonaut sea region(Li et al., 2024).studies reveal that identifying distinctly related aquatic species is based on COI genes. The sequencing of the COI gene fragments provides details about the biodiversity of aquatic species. the presence of unidentified fish species in the blue Amazon River of Brazil is determined using DNA barcoding. The classification of species based on their oncogenic features is possible through a DNA barcoding approach(Lutz et al., 2023).studies claim that next-generation sequencing is developed to avoid the alignment-based traditional methods to classify aquatic organisms .biodiversity of species is associated with their unique genomic and phenotypic characteristic. Using ML-based algorithms provides insight about the species based on their genomic sequencing(Millan Arias, 2024).studies suggest that the BEAS river is a hub of

diversified fish species.to identify the ichthyofauna species present in the Indus River, the COI technique of DNA barcoding is preferable. The identified species are then grouped based on their phylogenetic characteristic(Modeel et al., 2024). Moreover, the presence of a large number of global fish species in the Mozambique region has created a new horizon for fish classification in the region. the fauna of fish species found in Mozambique is studied using DNA barcoding. The aquatic species present at the coast and at the bottom of the sea of Mozambique are identified by DNA barcoding(Muhala et al., 2024). Scholars predict that using the barcoding technique, the mitochondrial cytochrome oxidase gene is used to study fish species. Fish barcoding is a modern approach introduced to tackle the challenges related to the classification of certain unique species found in the ocean.

DNA barcoding provides a large database of species based on their morphological features that helps group the species(Mwita & Chuhila, 2023).studies conclude that DNA tools efficiently study or monitor shark species based on their diversity. Conservation hotspots have been developed to conserve the classified shark species using the universal method of close tube barcoding(Prasetyo et al., 2023).Scholars predict that a DNA barcode library has been developed to contain all the information regarding the species present in the aquatic environment. DNA barcoding classifies species based on nucleotide differences. estimating the diverse range of species is a difficult task that requires efficient methods and assessment approaches (Ramirez et al., 2023).scholars predict aquacultures have been developed to protect aquatic species from environmental fluctuations. aquaculture provides a secure ecosystem and livelihood to the aquatic organisms. development of technologies for aquaculture is based on genomic research. NGS is a modern tool that allows for managing oceanic data using an analysis tool(Rather et al., 2023).Studies reveal that in the south of Brazil, fishing hotspots consist of a diverse range of species. Analyzing the COI gene in aquatic species determines their taxonomic behavior. information regarding the threatened and declining species of the Brazilian sea is obtained through a barcoding system(Santana et al., 2023).studies explain that DNA barcoding is performed by studying the mitochondrial genomics of fish species. The PCR amplification of the mitochondrial genome of species present in the IB river helps their classification process. BIN report is made after identifying species to determine the taxonomic relation of fish species. Also, analyzing the Barcode proves that each species classified through the barcoding process differs from others (Suryawanshi et al., 2024).Researchers suggest that a diversity of immense species of fish exists in the Yangtze River in the Yangtze River. Various anthropogenic factors that threaten fish and their habitat, around half of the species present in the Guizhou, have been identified using the sequencing technique of DNA barcoding(Tang et al., 2023).Scholars' studies suggest that species-rich tropical areas in aquatic environments become vulnerable when the world learns about these areas. the overfishing activities in such areas increase, and thus, species get threatened. Safe conservation strategies are adopted for species management to save species pollution in aquatic environments. The fish species found in the Mexican River are differentiated from each other using DNA barcode sequencing(Uh-Navarrete et al., 2023).

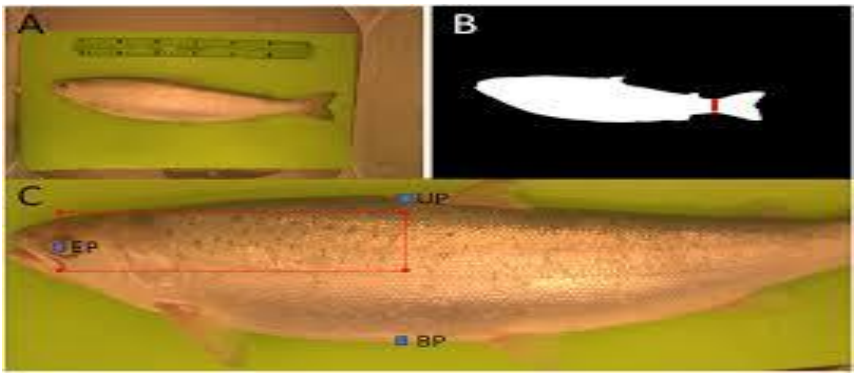

Figure 1: Computational Approach to Fish identification

Computational Approach to Fish identification

The formerly challenging challenge of accurately identifying fish and other aquatic species of interest has been resolved using deep neural networks. This study uses extensive data gathering and fish species studies to show how effective the deep neural network technique is in deep oceans. In order to improve the YOLOv7 object identification method, this research suggests a detection module that makes it lighter and more precise. Depth wise separable convolution in BNAM and feature extraction network refinement are used to achieve the improvement. The FD_Net technique is introduced in the fish species identification module, and the BNAM attention module is added to DenseNet-169. Furthermore, dilated convolution, BAN, pooling layers, and a loss function are introduced. The enhanced DenseNet-169 is used as a network for feature extraction. Arcface Loss is employed as a loss function to ensure separability across classes and to increase compactness inside a class. Combining identification and detection methods allows for the realization of a wide variety of fish species observed underwater. Fish are identified using the FD_Net, YOLOv3, YOLOv3-TL, YOLOv3-BL, YOLOv4, YOLOv5, Faster-RCNN, and YOLOv7 models. Because it uses fewer parameters than other models, the FD_Net model provides a very fast detection time. The results show that when it comes to extracting finer-grained characteristics from tiny objects, the FD_Net model outperforms the other models. From a competitive perspective, this has greatly raised the IoU of tiny and medium-sized photos. This study evaluates the mAP, detection accuracy, and FPS for each class.

Data mining and Library Compilation

Libraries can be constructed using BAGS from a selection of taxa or a species list supplied by the user. Following acquisition, records that match the chosen taxonomy or species list will be filtered. The taxa the user entered must thus be in BOLD when the system is used. It should be noted that some species may not be able to access some ranks, especially intermediate ranks, or that BOLD does not implement them. Any taxonomic rank within Animalia, from species to phylum, can be recorded. There are three options for mining the target taxa: downloading all available records (all taxa), only records of species that occur in marine habitats (including any taxa found in brackish waters), or records of non-marine species (i.e., not found in either marine or brackish water habitats). In WoRMS, this establishes which of the four habitat types—terrestrial, freshwater, brackish, or marine—is assigned to each species in a query data set.

Output and Annotation-Based File Sorting

The auditing system then marks the reference library items with the preset ratings for each species using the pipeline described above. A tabular file with the following details will eventually be produced and made available for download: species name, BIN, COI-5P sequence, country or region of origin, grade assigned to the species, number of base pairs in the sequence, family, order, class, sample ID, process ID, latitude, longitude, and, in the case of marine taxa libraries, an additional column with the valid species name as determined by WoRMS. Additionally, by downloading the reference library in fast format, the user may choose which grades to include. The Fasta files for every grade are available for download either individually, in combination with other grades, or for all grades. Lastly, BAGs provide a summary of the information from the developed reference library. A written report and two bar plots—one displaying the number of specimens for each grade assigned, and the other displaying the number of species for each grade assigned—are used to achieve this. The user must first reload the page in order to repeat the process for other target libraries. Using a brief DNA fragment from a particular gene or genes, DNA barcoding is a method for identifying species. DNA barcoding works on the basis that a single sequence can be used to uniquely identify an organism to species by comparing it to a reference library of such DNA sections (also called "sequences"), much like a supermarket scanner uses the recognisable

black stripes of a UPC barcode to identify an item in stock against its reference database. By comparing them to traditional taxonomical categorisation, these "barcodes" can be used to identify unknown species or parts of an organism, list as many taxa as feasible, and establish species boundaries. Based on distinct gene regions, barcoding is utilised to distinguish between various organismal groupings. In mammals and certain protists, the most often utilised barcode area is the mitochondrial DNA cytochrome c oxidase I (COI or COX1) gene. The internal transcribed spacer (ITS) rRNA, which is frequently found in fungi, and RuBisCO, which is found in plants, are additional genes that are appropriate for DNA barcoding. Numerous gene areas are used to identify bacteria. For instance, prokaryotes are often identified using the 16S rRNA gene, while microbial eukaryotes are mostly identified using the 18S rRNA gene. These gene areas were selected because of the "Barcoding Gap"—the difference between intraspecific (within species) and interspecific (between species) variance.

Some Applications of DNA Barcoding Include

recognizing insect larvae, which may lack adult diagnostic characteristics; identifying pollen found on pollinating animals' bodies; identifying plant leaves even when they are devoid of blossoms or fruits; or determining an animal's diet by looking at its stomach contents, saliva, or faeces. When barcoding is utilized to identify species in a sample that contains DNA from many taxa, DNA metabarcoding—such as DNA metabarcoding of diatom populations in rivers and streams—is utilized to evaluate the quality of the water.

Applications

DNA barcoding has many applications, including the identification of new species, the evaluation of food safety, the identification and evaluation of cryptic species, the detection of alien species, the identification of endangered and threatened species, the linkage of egg and larval stages to adult species, the protection of bioresource intellectual property rights, the clarification of feeding niches, the creation of global management plans for conservation strategies, and much more. Fundamental problems in systematics, ecology, evolutionary biology, and conservation, including community formation, species interaction networks, taxonomy discoveries, and the establishment of priority environmental protection sites, can be resolved using DNA barcode markers.

Identification of Species

Short DNA sequences or markers from a standardised area of the genome can be combined to generate a DNA barcode for species identification. Molecular technologies can be quite helpful when conventional methods are not working. The identification of larvae for which few diagnostic traits are available and the association of different life stages (such as larval and adult) in many species are only two of the many uses for DNA barcoding. Barcoding technology is used to track illegal trafficking and identify species included in the Convention on the International Trafficking of Endangered Species (CITES) appendix.

Detection of Invasive Species:

It is possible to identify foreign creatures via barcoding. Barcoding may be appropriate for species recognition in situations like border control, where quick and precise morphological identification is typically difficult because of species similarities, a lack of diagnostic features, or a lack of taxonomic knowledge. Furthermore, invading species in ecosystems may be identified and differentiated from native species with similar morphologies using barcoding and metabarcoding. It has been demonstrated that DNA identification is significantly more effective than conventional biological invasion monitoring.

Delimiting Cryptic Species

We can distinguish and identify cryptic species due to DNA barcoding. Since the outcomes of these studies rely

on the analytical methods used, the process of categorising cryptic species using DNA barcodes may be just as subjective as any other type of taxonomy. Studies show that the northwest region of Costa Rica is home to ten distinct species of Astraptes fulgerator butterflies. Others, however, contested these findings, pointing out several serious shortcomings in the study and coming to the conclusion that the original data could only support three to seven cryptic taxa instead of 10 cryptic species. The 20 morphospecies of Belvosia parasitoid flies (Diptera: Tachinidae) that feed on caterpillars (Lepidoptera) in the Area de Conservación Guanacaste (ACG), northwest Costa Rica, were identified by the researchers using cytochrome c oxidase I DNA barcodes. These researchers found that barcoding expands the number of parasitoids to 32 by proving that the three species, which were previously believed to be generalists, are really arrays of highly host-specific cryptic species. DNA barcoding analysis of 15 morphospecies of polychaetes in the deep Antarctic benthos revealed cryptic diversity in half of them. The total species richness of the sample increased by 233% with the discovery of eight morphospecies that had previously been disregarded.

Barcoding for Food Safety

DNA barcoding is an essential technique for evaluating the quality of food products. The objectives are to ensure food traceability, decrease food piracy, and evaluate local and typical agro-food production. Another objective is to preserve public health. Metabarcoding enables, for example, the detection of groupers that are causing Ciguatera fish poisoning from meal remnants or the differentiation of edible and deadly mushrooms.

Biomonitoring and Ecological Assessment

For conservation purposes, DNA barcoding can be used to determine if endangered species are present (Ref) or whether indicator species that indicate certain ecological circumstances (Ref), such as low oxygen levels or abundant nutrients, are prevalent.

Forensic Science

DNA barcoding is commonly used to identify species in forensic science investigations. At crime scenes, samples of unknown plants or animals may be found, collected, and identified in the hopes of linking them to a suspect and obtaining a conviction. DNA barcoding is used in crimes like poaching, the death of endangered species, and animal abuse since animal DNA is commonly detected. Plant trace evidence, on the other hand, is usually used to link a suspect to a crime scene.

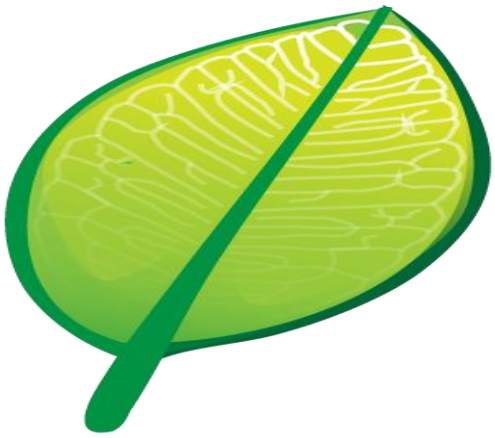

Figure 2: Taxonomic Resolution

Taxonomic Resolution

DNA barcoding makes it possible to resolve taxa from higher (family) to lower (species) taxonomic levels that

are typically too challenging to identify using conventional morphological approaches, such microscopy. For instance, the non-biting midge Chironomidae may be found in both freshwater and terrestrial environments. Their diversity and abundance make them significant for ecological processes and networks, and they are among the several invertebrate groups employed in biomonitoring. In samples of invertebrates, up to 100 species of chironomids can be found, making up as much as 50% of the total. Despite this, they are rarely acknowledged below the family level due to the time and taxonomic expertise needed. Multiple chironomid species with disparate ecological preferences may therefore be confused, which might result in an imprecise evaluation of the quality of the water.

Taxonomic clarity and the clear association of stressor effects with particular taxa, such individual chironomid species, are made possible by DNA barcoding. For instance, used DNA barcoding to examine the responses of Chironomidae to several stressors, such as decreased flow, fine silt, and increasing salinity. The chironomid sample was found to have 183 Operational Taxonomic Units (OTUs), which are barcodes (sequences) that are generally equivalent to morphological species. These 183 OTUs generated 15 response types as opposed to the two that were previously seen when all chironomids were included in single multiple stressor research. A similar tendency is shown by the discovery of cryptic diversity in the New Zealand mayfly species Deleatidium. The widely accepted belief that this mayfly is pollutant-sensitive may be called into question by this study, which showed that 12 molecularly diverse OTUs responded differently to stimuli.

Estimates of Richness/Diversity

The species richness and diversity may be overestimated or underestimated as a consequence of DNA barcoding. According to some research, one of the main reasons for the exaggerated biodiversity is artefacts, or the identification of species that are not found in a community. The taxa with the fewest sequencing reads are the most troublesome. Since several studies indicate that the majority of these low-frequency scans could be artefacts, they are often eliminated during the data filtering procedure. Nonetheless, these low-abundance readings could contain actual uncommon taxa. Rare sequences are useful and instructive because they can reveal distinct lineages within populations. More reliable bioinformatics algorithms that can distinguish between meaningful reads and artefacts are therefore desperately needed.

Conclusion

Complete reference libraries would also make it feasible to test bioinformatics algorithms more effectively by enabling better artefact filtering (i.e., removing sequences that do not have a counterpart among current species), which would lead to a more accurate species designation. Because one morphological species may really divide into several separate genetic sequences, cryptic diversity can also lead to an overestimation of biodiversity. This would have a significant impact on the development of DNA reference data, which is necessary for environmental DNA-based biodiversity monitoring. DNA barcoding was initially introduced to the scientific world in 2003 when Paul Hebert's research team at the University of Guelph released a paper titled "Biological identifications through DNA barcodes." Using a brief DNA fragment from a standardised area of the genome, they suggested a novel method for identifying and locating species. That DNA sequence may be used to identify different species, just how a supermarket scanner detects your goods using the distinctive black stripes of a UPC barcode. Almost all animal species employ the 648 base-pair segment of the mitochondrial cytochrome c oxidase 1 gene ("CO1"), which has been shown to be very useful for identifying fish, flies, birds, butterflies, and many other animal groups. The benefit of employing COI is that it is short enough to be sequenced efficiently and affordably while yet being lengthy enough to identify species differences. The COI barcode is ineffective for plant identification because of its slow development, however two gene portions in the chloroplast, matK and

rbcL, have been recognised as barcode areas for land plants.

References

Akter, M., Rahman, F., Shohag, S., & Hossain, M. N. (2023). Necessity of DNA barcoding to identify prospective marine organisms in Bangladesh. *Bangladesh Marit. J*, *7*, 20.

Blanco-Bercial, L., Cornils, A., Copley, N., & Bucklin, A. (2014). DNA barcoding of marine copepods: assessment of analytical approaches to species identification. *PLoS currents*, *6*.

Chakraborty, M., Dhar, B., & Ghosh, S. K. (2017). Design of character-based DNA barcode motif for species identification: A computational approach and its validation in fishes. *Molecular ecology resources*, *17*(6), 1359-1370.

Chatla, D., Pamulapati, P., Kola, S., & Naranji, M. K. (2024). Genetic identification of grouper fishes (Perciformes: Serranidae: Epinephelus) through DNA barcoding from Nizampatnam coastal waters: DNA barcoding of grouper fishes (Perciformes: Serranidae: Epinephelus). *TAXA*, *3*.

Eriegha, O. J., Ekelemu, J. K., & Nwachi, O. F. (2024). DNA BARCODING OF Caranx senegallus FROM THREE WATER BODIES ADJOINING THE GULF OF GUINEA, NIGERIA. *FUDMA JOURNAL OF SCIENCES*, *8*(1), 293-298.

Fontes, J. T., Ferreira, A. O., Lavrador, A. S., Machado, C., Soares, P., Costa, F. O., & Canchaya, C. A. (2024). Exploring artificial neural networks for the curation of DNA barcode reference libraries: a pilot study with marine fish species of the Eastern Atlantic. *Authorea Preprints*.

Jiang, C., Yi, M., Luo, Z., He, X., Lin, H. D., Hubert, N., & Yan, Y. (2023). DNA barcoding the ichthyofauna of the Beibu Gulf: Implications for fisheries management in a seafood market hub. *Ecology and Evolution*, *13*(12), e10822.

Jin, L., Yu, J., Yuan, X., & Du, X. (2021). Fish classification using DNA barcode sequences through deep learning method. *Symmetry*, *13*(9), 1599.

Karabanov, D. P., Kotov, A. A., Borovikova, E. A., Kodukhova, Y. V., & Zhang, X. (2023). Comparison of the Efficiency of Single-Locus Species Delimitation Methods: A Case Study of a Single Lake Fish Population in Comparison against the Barcodes from International Databases. *Water*, *15*(10), 1851.

Kasianchuk, N., Harkava, S., Onishchenko, S., Solodka, O., Shyshko, D., Siemens, E., Falfushynska, H., & Ustyianovych10, T. (2024). Computational Breakthroughs in Aquatic Taxonomy: The Role of Deep Learning and DNA Barcoding. Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT),

Kim, S., Eo, H.-S., Koo, H., Choi, J.-K., & Kim, W. (2010). DNA barcode-based molecular identification system for fish species. *Molecules and cells*, *30*, 507-512.

Li, H., Miao, X., Wang, R., Liao, Y., Wen, Y., Zhang, R., & Lin, L. (2024). Biodiversity of Demersal Fish Communities in the Cosmonaut Sea Revealed by DNA Barcoding Analyses. *Genes*, *15*(6), 691.

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*(1), 16601.

Lutz, Í., Martins, T., Araújo, F., Ferreira, C., Santana, P., Miranda, J., Matos, S., Sousa, J., Pereira, L., & Bentes, B. (2023). Molecular characterization of juvenile fish from the Amazon estuary using DNA barcoding approach. *Plos one*, *18*(9), e0292232.

Millan Arias, P. (2024). Deep Unsupervised Learning for Biodiversity Analyses: Representation learning and clustering of bacterial, mitochondrial, and barcode DNA sequences.

Modeel, S., Negi, R. K., Sharma, M., Dolkar, P., Yadav, S., Siwach, S., Yadav, P., & Negi, T. (2024). A comprehensive DNA barcoding of Indian freshwater fishes of the Indus River system, Beas. *Scientific reports*, *14*(1), 2763.

Muhala, V., Guimarães-Costa, A., Macate, I. E., Rabelo, L. P., Bessa-Silva, A. R., Watanabe, L., Dos Santos, G. D., Sambora, L., Vallinoto, M., & Sampaio, I. (2024). DNA barcoding for the assessment of marine and coastal fish diversity from the Coast of Mozambique. *Plos one*, *19*(2), e0293345.

Mwita, C. J., & Chuhila, Y. J. (2023). Fish DNA barcoding: advances and challenges. *Frontiers in Aquaculture Biotechnology*, 171-185.

Pereira, L. H. G., Castro, J. R. C., Vargas, P. M. H., Gomez, J. A. M., & Oliveira, C. (2021). The use of an integrative approach to improve accuracy of species identification and detection of new species in studies of stream fish diversity. *Genetica*, *149*(2), 103-116.

Prasetyo, A. P., Cusa, M., Murray, J. M., Agung, F., Muttaqin, E., Mariani, S., & McDevitt, A. D. (2023). Universal closed-tube barcoding for monitoring the shark and ray trade in megadiverse conservation hotspots. *Iscience*, *26*(7).

Ramirez, J. L., Núñez-Rodriguez, D., Britzke, R., Siccha-Ramirez, R., & Ramírez, R. (2023). DNA barcoding for assessing biodiversity. In *Conservation Genetics in the Neotropics* (pp. 21-45). Springer.

Rather, M. A., Agarwal, D., Bhat, T. A., Khan, I. A., Zafar, I., Kumar, S., Amin, A., Sundaray, J. K., & Qadri, T. (2023). Bioinformatics approaches and big data analytics opportunities in improving fisheries and aquaculture. *International Journal of Biological Macromolecules*, *233*, 123549.

Santana, P., Martins, T., Lutz, Í., Miranda, J., da Silva, R., Mesquita, D., Martins, R., Veneza, I., Vallinoto, M., & Sampaio, I. (2023). DNA barcode reveals occurrence of threatened species and hidden diversity on Teleost fish trade in the Coastal Amazon. *Scientific reports*, *13*(1), 19749.

Suryawanshi, R., Gaikwad, S., & Khedkar, G. (2024). DNA Barcoding of Fish in the Ib River, One of the Important Tributaries of the Mahanadi River, India. *Russian Journal of Genetics*, *60*(8), 1077-1092.

Tang, Q., Deng, L., Luo, Q., Duan, Q., Wang, X., & Zhang, R. (2023). DNA barcoding of fish species diversity in Guizhou, China. *Diversity*, *15*(2), 203.

Uh-Navarrete, A. E., Valdez-Moreno, M., Callejas-Jiménez, M. E., & Vásquez-Yeomans, L. (2023). Discovering the fish fauna of a lagoon from the southeast of the Yucatan Peninsula, Mexico, using DNA barcodes. *PeerJ*, *11*, e16285.

Zhang, A. B., Muster, C., Liang, H. B., Zhu, C. D., Crozier, R., Wan, P., Feng, J., & Ward, R. (2012). A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. *Molecular ecology*, *21*(8), 1848-1863.

Zhang, J.-B., & Hanner, R. (2011). DNA barcoding is a useful tool for the identification of marine fishes from Japan. *Biochemical Systematics and Ecology*, *39*(1), 31-42.

Zhang, J., & Hanner, R. (2012). Molecular approach to the identification of fish in the South China Sea. *Plos one*, *7*(2), e30621.