

Preliminary Investigation of Fish Diversity in Huayang River based on eDNA Technology

Jiemiao Wang ^{1, a}, Wenqiao Tang ^{1, 2, *}

¹ Shanghai Universities Key Laboratory of Marine Animal Taxonomy and Evolution, Shanghai Ocean University, Shanghai, China

² Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Ministry of Education, Shanghai Ocean University, Shanghai, China

* Corresponding author: Wenqiao Tang (Email: wqtang1964@126.com), ^a 1769685067@qq.com

Abstract: Environmental DNA (eDNA) was used to study fish biodiversity in the Huayang River to identify new ways to monitor and protect fish diversity. 23 sites were set up in the Huayang River on 23 September. Analytical procedures were used to detect fish diversity, including water sampling, eDNA extraction, genetic marker amplification and sequencing, and database comparison. 49 fish species were identified, classified into six orders, 12 families and 34 genera. The most common fish were the big-finned goby, group-headed pomfrets, yellowfin spiny goby, short-bearded goby and coloured rocky carp. H18 and H18 are the most prevalent species globally. The diversity index showed that H18 and H23 had the highest fish diversity, while H5 had the lowest. eDNA technology can rapidly ascertain fish diversity and distribution in the Pearl River. Combining eDNA with traditional monitoring methods can enhance fish biodiversity data.

Keywords: eDNA; Huayang River; Fish Diversity.

1. Introduction

The collection of fish samples through traditional fishing techniques can provide valuable insights into the growth parameters of fish, including size, weight, and developmental stage [1]. This information is crucial for the sustainable utilisation of fish resources. However, traditional fish monitoring methods have inherent limitations, including high costs, high intensity, and specialisation, as well as potential adverse effects on fish populations and communities[2-4]. Environmental DNA (eDNA) technology has recently emerged as a popular method for biodiversity detection and assessment[5-6]. Its high sensitivity and minimal interference make it an effective alternative to traditional monitoring with potential for biodiversity assessment[7-9]. It's currently used to gather data on species in aquatic ecosystems[10-11] and resolve river issues[12-13].

The Huayang River is in Anhui. A primary north bank Yangtze River tributary, it is narrow in both directions. It flows through several lakes, including Longguan, Daguang, Huanghu and Puhu. The Huayang River watershed is 5,511

km². The northern river area is hilly and granite-covered. The southern part of the river is an alluvial continental bank with lakes and depressions. The 39-kilometre stretch from Huayang Town to the Yangtze River is the Huayang River region. The river's shores are home to a diverse range of wildlife. There is little data on fish biodiversity in the Huayang River basin. We need a fast, effective and eco-friendly way to monitor fish biodiversity and fishery resources. We used eDNA technology to obtain comprehensive water samples from 23 sites in the Huayang River basin. We then analysed and annotated the data to gain insight into fish biodiversity. Our objective was to provide a scientific foundation for informed fishery conservation and management decisions.

2. Materials and Methods

2.1. Sample Section Setup

Combined with the river morphology of the Huayang River, 23 sampling sections were set up, and the number and GPS position of each sampling section is shown in Figure 1.

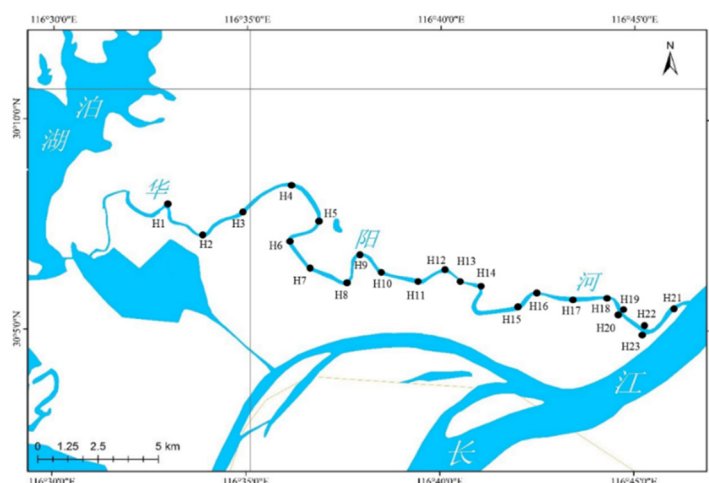


Fig 1. Schematic diagram of sampling sites in Huayang River

2.2. Collection of Water Samples

Water samples were collected from 18 September to 8 October 2023 at 23 sites using a 1 L glass sampler. Six 2 L samples were collected at each site and stored in clean bottles. All samples were filtered on-site using a 0.45 µm filter (GN-6, Pall) to enrich for eDNA. 500 ml of deionized water was pumped as a negative control prior to pumping six samples from each site. Each filtered membrane was stored in a liquid nitrogen tank until DNA extraction.

2.3. eDNA Extraction

The eDNA was extracted using the CTAB method [15]. The solution was qualified by gel electrophoresis and Nanodrop 2000, then frozen at -20°C. All samples showed bands, while negative controls showed no significant bands.

2.4. Genetic Marker Amplification and High-Throughput Sequencing

Ltd. amplified and sequenced as per Zhang *et al* [16]. The fish universal primers Mifish-U-F and Mifish-U-R, developed by Miya *et al* [17], were used to amplify 12S rRNA. The amplification length was approximately 160-190 bp. The 25µL PCR system contained 5× buffer, 5µL dNTP (2.5mM), 5µL GC buffer, and 1µL positive and negative controls. 2µL DNA template (20ng/µL), 8.75µL water and 0.25µL Q5 Taq enzyme. 0.25µL. PCR conditions: pre-denaturation 98°C 2 min; 98°C denaturation 15 s, 55°C annealing 30 s, 72°C extension 30 s (35 cycles); 72°C 5 min. Three samples per section were amplified by PCR, purified by 2% agarose gel electrophoresis, and prepared for sequencing using TruSeq Nano DNA LT Sample Preparation Kit. Bipartite sequencing was performed on an Illumina Novaseq-PE250.

2.5. Data Processing Methods

2.5.1. OTU Partitioning and Species Annotation

Downstream data were pre-screened to obtain original sequences. High-quality sequences were obtained through de-priming, splicing, quality filtering, de-climatisation, and

clustering. The sequences were then clustered using QIIME at ≥97% similarity [18-19]. The most abundant sequences were selected as representative of each OTU. The representative sequences were compared with the NCBI NR database to obtain taxonomic information on the OTUs [20].

2.5.2. Alpha Diversity

The Chao1 index [21], Simpson's index [22], Shannon-Wiener's index [23], and Pielou's index [24] were used to estimate species richness and diversity. The equations are as follows.

$$S_{chao} = S_{obs} + \frac{n_1(n_1-1)}{2(n_2+1)} \quad (1)$$

$$H' = -\sum_{i=1}^S N_i \ln N_i \quad (2)$$

$$C = 1 - \sum (N_i)^2 \quad (3)$$

$$J' = H' / \ln(S) \quad (4)$$

S_{obs} is the number of OTUs actually observed, n_1 is the number of OTUs containing only one sequence, and n_2 is the number of OTUs containing only two sequences; N_i is the relative abundance of OTUs for the i th fish species, and S is the total number of species.

2.6. Statistical Analysis of Data

The data were processed using Excel 2019, and the R packages vegan and ggplot2 were used to calculate alpha diversity. The distribution map was plotted in ArcGIS.

3. Results

3.1. eDNA Sequencing Results

The Huayang River eDNA sequencing results yielded 2775722 original sequences, 2731518 valid sequences, and 2519975 high-quality sequences, representing over 90% of the original sequences. The mean value of high-quality data at each sampling point was 109564.13 (90919-137725), with the highest being H23 and the lowest being H6. Figure 2 shows the dilution curves for each Huayang River sample. The curves flatten at 60,000 sequences, indicating sufficient sequencing volume for data analysis.

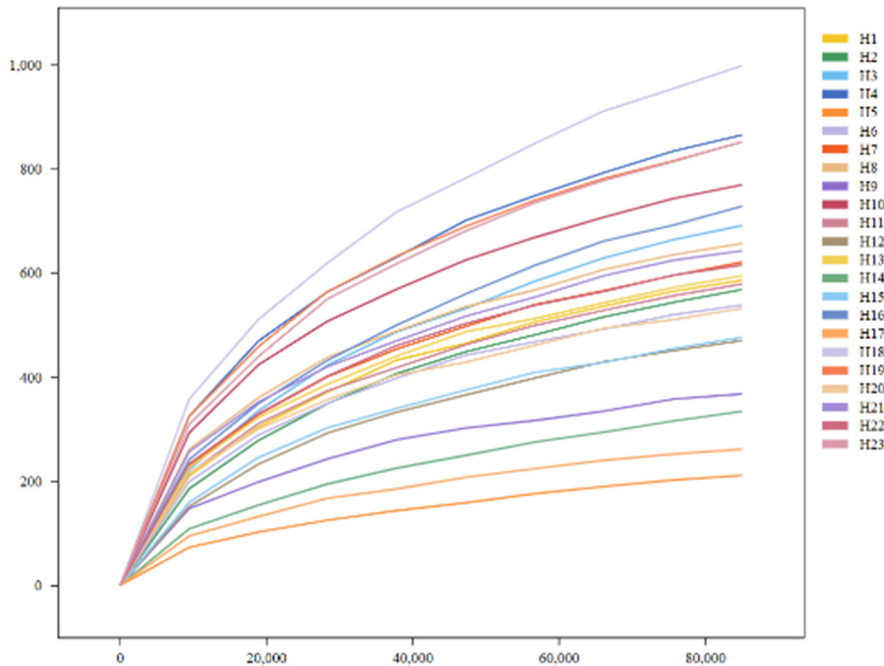


Fig 2. Rarefaction curve of samples in Huayang River

3.2. OTU Annotation Results and Fish Species Composition

A total of 3511 OUTs were obtained from the Huayang River survey, with an average of 153 OUTs per site. H18 had the highest total with 998, while H5 had the lowest. Fig 3 shows the highest number of OUTs was 142, in H23, and the lowest, four, was in H14. Nineteen OUTs were common to all 23 sampling sites.

The detected OUTs were compared with the NCBI NR database, resulting in the detection of 80 fish species belonging to 49 genera. 6 unnamed species, 19 misidentified species, and 1 aquarium species were detected. The remaining 49 species belong to 6 orders, 12 families, and 34 genera.

There were 36 species of Carpiiformes (73.47% of the total) and 8 species of Perchiformes (33% of the total); 2 species of Siluriformes, accounting for 4.08% of the total; and 1 species in each of the families Clupeiformes, Beloniformes, and Synbranchiformes, accounting for a total of 6.12% of the total number of species.

The Cyprinidae had the most fish species, with 35, accounting for 71.43% of the total. The Gobiidae, Serranidae and Odontobutidae had Two species (12.24%) ; and eight

families (16.33%) had one species each.

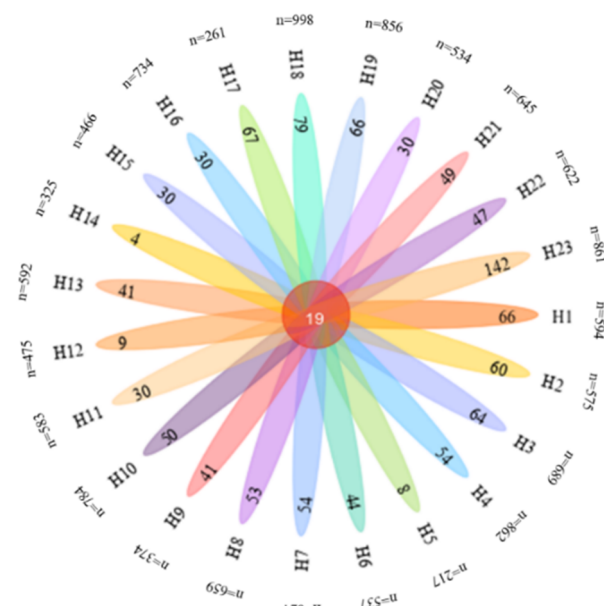


Fig 3. Venn diagram of mutual OTUs of samples in Huayang River

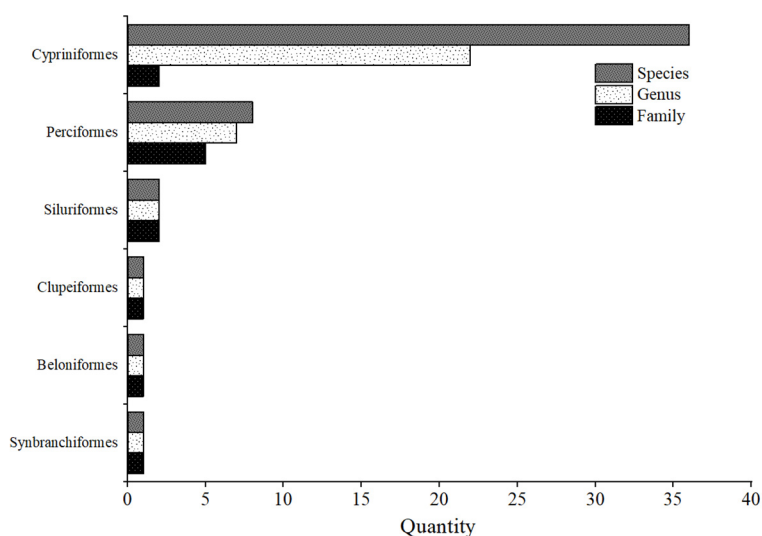


Fig 4. Composition of number of fish Family/Genus/Species of Huayang River (eDNA)

3.3. Distribution of Dominant Species

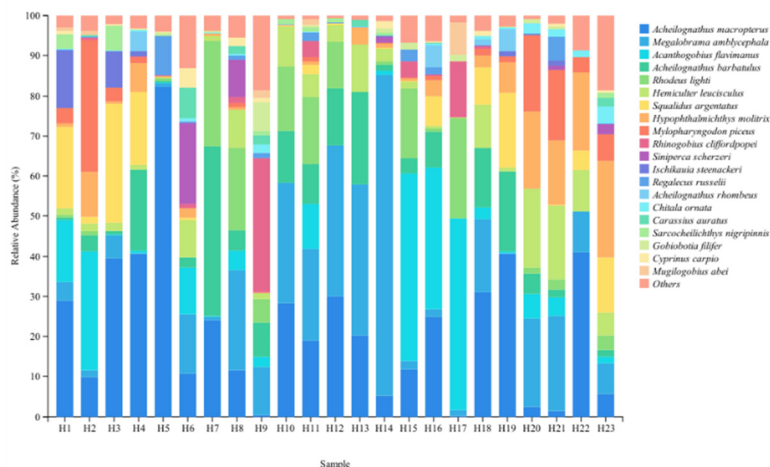


Fig 5. The composition of fish species at sampling site in Huayang River

The dominant fish species identified based on relative abundance at each sampling site are shown in Figure 5. The top 10 species in terms of abundance were *Acheilognathus macropterus*, *Megalobrama amblycephala*, *Acanthogobius flavimanus*, *Acheilognathus barbatus*, *Rhodeus lighti*, *Hemiculter leucisculus*, *Squalidus argentatus*, *Hypophthalmichthys molitrix*, *Mylopharyngodon piceus*, *Rhinogobius cliffordpopei*, and except *Rhinogobius cliffordpopei*, several other fish were detected at each of the 23 sampling sites.

3.4. Diversity Analysis

Figure 6 shows the trend of the fish community alpha diversity in Huayang River sampling sites. The Chao1 index ranged from 323.07 to 1340.77, and the observed species index was 213.0 to 1001. The distribution trend corresponded between the two. H18 had the highest index value and H17 had the lowest. The Shannon index ranged from 1.26 to 5.06, the Simpson index ranged from 0.33 to 0.94, and the Pielou uniformity index ranged from 0.16 to 0.52. The trends of the three indices were consistent and the highest values of the indices occurred in H23, while the lowest values occurred in H5.

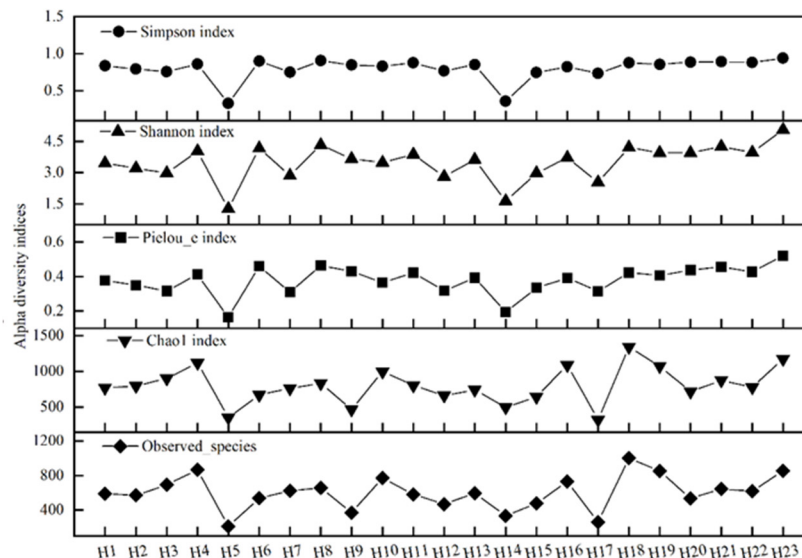


Fig 6. Alpha diversity indices of important estuarine fishes in Huayang River

4. Discussion

4.1. Fish Species Composition based on eDNA Detection

A survey of fish diversity in the Huayang River using eDNA technology detected 49 species. Two main factors cause misidentified and unnamed species. Most free DNA in the environment is short and similar, which can bias detection[25]. Secondly, the accuracy of database comparisons may influence the results, as the mitochondrial database does not currently have the capacity to include all fish sequences. Additionally, limitations in sequencing read length resulted in the inability to discriminate between certain species and genera[26-27]. The range and relative abundance of forage and aquarium species indicated that they were predominantly found in waters with high levels of anthropogenic disturbance.

A comparison of the results from laboratory field sampling (not yet published) revealed some differences in the species composition of fish detected by the two methods. eDNA technology detected 49 species of actual inhabiting fish, which was slightly lower than the 50 species detected in the actual collection surveys. However, it should be noted that the monitored fish were all dominated by the Cyprinidae family. At the genus level, eDNA technology identified several genera that were not identified in the field collection. These included the *Cyprinus*, *Hemibarbus*, *Rhinogobio*, *Megalobrama*, *Mylopharyngodon*, *Ochetobius*, *Pseudolaubuca*, *Silurus*, *Acanthogobius*. The eDNA

technology was unable to detect the presence of the *Pelteobagrus*, and the *Squaliobarbus* was unable to detect all 14 genera. It is evident that eDNA technology did not fully encompass the actual species collected, potentially due to the selectivity and preference of 12S universal primers. Additionally, Some low-abundance species went undiagnosed [28]. However, the overall detection and coincidence rates of eDNA technology in the Huayang River were still promising, exceeding 50% and enabling the detection of species that were challenging to identify through traditional sampling methods. This suggests that eDNA technology can effectively address the limitations of traditional survey methods, providing an efficient means of fish detection that facilitates rapid understanding of the fish population in the river. This demonstrates that eDNA technology can offset the constraints of conventional survey techniques and is an efficacious approach to fish detection. This helps us understand fish communities in rivers[29].

4.2. Fish Diversity based on eDNA Testing

The alpha diversity index shows differences in species composition between Huayang River sampling sites. H18 has the richest fish species diversity. The result may be related to the sampling site's environment. H18 is characterized by a high density of aquatic vegetation, including water hyacinth, sparrowweed, and lotus grass. The open water provides an ideal foraging environment for fish, contributing to the diversity of species at this site. The same trend is observed for the uniformity indices, with H23 having the highest value. The sampling site is in the middle of the river, not restricted

by the Huayang Lock, so there is minimal human impact and higher fish diversity. The indices were lower due to the sampling site's proximity to the dock, where there are more human activities. These factors may reduce fish activity and diversity[30].

4.3. The Applicability and Prospects of eDNA Technology for the Detection of Fish Species are Discussed

The eDNA technology produces consistent results with traditional survey methods[31-32]. The fish species detected by eDNA in this study were not fully representative of the river's biodiversity, but the detection rate was high and the concordance rate showed certain advantages. The eDNA results also showed similarities with field sampling, indicating the applicability of eDNA technology to fish resource surveys[33-35]. The survey also detected harder-to-catch fish, confirming the feasibility of eDNA technology to monitor aquatic ecology. The eDNA technology has obvious advantages in monitoring fish distribution, helping us quickly understand river fish communities and better protect river ecosystems[36].

eDNA technology is not fully developed. This study found gaps between eDNA and field sampling results for species abundance. Future research should focus on: (1) Improve database construction, design primers for local species, reduce unidentifiable or mismatched species, and improve species detection[37]. (2) The objective is to optimise the eDNA experimental system with a view to reducing experimental errors. It is acknowledged that eDNA work is often carried out in the field, where the collection of samples is subject to multiple constraints and is prone to cross-contamination. It is therefore essential to have a set of unified and standardised eDNA technical specifications in place in order to make species monitoring more accurate and effective. (3) The establishment of a comprehensive eDNA assessment system is crucial for accurately estimating biomass in aquatic environments. DNA in these ecosystems is susceptible to interference and influence from various factors, including hydrodynamic conditions, water temperature, pH, and others [38-39]. eDNA degradation can directly lead to discrepancies between detection results and the actual biological situation. Therefore, it is essential to identify and quantify the external environmental variables that affect eDNA production and decomposition rates. This can be achieved through the development of an eDNA assessment system, which can provide a more accurate estimation of biomass.

References

- [1] LIU F, LIN P C, LI M Z, et al. Situations and conservation strategies of fish resources in the Yangtze River basin[J]. *Acta Hydrobiologica Sinica*, 2019, 43(S01):144-156.
- [2] LIU H Z, YANG J X, LIU S W, et al. Theory and methods on fish diversity monitoring with an introduction to the inland water fish diversity observation in China. *Biodiversity Science*, 24, 1227-1233.
- [3] BONAR S A, MERCADO-SILVA N, HUBERT W A, et al. Standard methods for sampling freshwater fishes: opportunities for in ternational collaboration[J]. *Fisheries*, 2017, 42(3): 150-156.
- [4] YAO M, ZHANG S, LU Q, et al. Fishing for fish environmental DNA: ecological applications, methodological considerations, surveying designs, and ways forward[J]. *Mol Ecol*, 2022, 31(20): 5132-5164.
- [5] DEINER K, BIK H M, MACHLER E, et al. Environmental DNA metabarcoding: transforming how we survey animal and plant communities[J]. *Mol Ecol*, 2017, 26(21): 5872-5895.
- [6] ALICE V, PIERRE T, CLAUDE M, et al. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding[J]. *Mol Ecol*, 2015, 25(4): 929-942.
- [7] THOMSEN P F, WILLERSLEV E, 2015. Environmental DNA-an emerging tool in conservation for monitoring past and present biodiversity[J]. *Biological Conservation*, 183:4-18.
- [8] SHU L, LUDWIG A, PENG Z G. Standards for methods utilizing environmental DNA for detection of fish species[J]. *Genes-Basel*, 2020, 11(3): 296.
- [9] BOHMANN K, EVANS A, GILBERT M T P, et al. Environmental DNA for wildlife biology and biodiversity monitoring[J]. *Trends Ecol Evol*, 2014, 29(6): 358-367.
- [10] TAKAHARA T, MINAMOTO T, YAMANAKA H, et al. Estimation of fish biomass using environmental DNA[J]. *PLoS One*, 2012, 7(4): e35868.
- [11] THOMSEN P F, KIELGAST J, IVERSEN L L, et al. Detection of a diverse marine fish fauna using environmental DNA from seawater samples[J]. *PLoS One*, 2012, 7(8): e41732.
- [12] BELLE C C, STOECKLE B C, GEIST J. Taxonomic and geographical representation of freshwater environmental DNA research in aquatic conservation[J]. *Aquat Conserv*, 2019, 29(11): 1996-2009.
- [13] TSUJI S, TAKAHARA T, DOI H, et al. The detection of aquatic macroorganisms using environmental DNA analysis: a review of methods for collection, extraction, and detection[J]. *Environ DNA*, 2019, 1(2): 99-108.
- [14] LU S B, YANG L L, LI B, et al. Estimation of ecosystem services value of wetlands in Huayang lakes area for three periods[J]. *Wetland Science*, 2014, 12(6) :747-752.
- [15] JIAO Z G, LI J Y, WEN Z B, et al. Chemical and biological components analysis of PM2.5 and its different fractions in summer atmosphere in Beijing urban areas, China[J]. *Chinese Journal of Environmental Engineering*, 2016, 10(09):5009-5015.
- [16] ZHANG Y Y, TANG W Q, CHEN Z F, et al. Investigating the fish diversity in backbone rivers of Shanghai based on environmental DNA metabarcoding[J]. *Resources and Environment in the Yangtze Basin*, 2023, 32(7):1433-1446.
- [17] Miya M, Sato Y, Fukunaga T, et al. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species [J]. *Royal Society Open Science*, 2015,2(7): 150088.
- [18] Bokulich N A, Subramanian S, Faith J J, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing[J]. *Nature Methods*. 2013,10, 57- U11.
- [19] Blaxter M, Mann J, Chapman T, et al. Defining operational taxonomic units using DNA barcode data[J]. *Philos Trans R Soc B-Biol Sci*. 2005, 360, 1935-1943.
- [20] LIU Y, JIANG B T, SHI T R, et al. Sequencing analysis of bacterial community in Caecum of healthy Dezhou donkeys[J]. *Progress in Veterinary Medicine*, 2019, 40(04):64-68.
- [21] CHAO A N. Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*, 1984, 11(4): 265-270.
- [22] Simpson, E.H. Measurement of Diversity[J]. *Nature*, 1949, 163(4148): 688-688.

- [23] Shannon C E. The mathematical theory of communication[J]. Bell Labs Technical Journal, 1950, 3(9):31-32.
- [24] Pielou E C, Levandowsky M. Ecological Diversity[J]. Quarterly Review of Biology, 1975.
- [25] DEAGLE B E, JARMAN S N, COISSAC E, et al. DNA metabarcoding and the cytochrome oxidase subunit I marker: not a perfect match[J]. Biology Letters, 2014, 10(9).
- [26] YUE Y. Study of fish diversity in Yueqing Bay based on environmental DNA method[D]. Shanghai Ocean University, 2020.
- [27] GAO W R. Study on fish diversity of the lake-type water bodies based on environmental DNA[D]. Shanghai Ocean University, 2020.
- [28] ELBRECHT V, VAMOS E E, MEISSNER K, et al. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring [J]. Methods in Ecology & Evolution, 2017, 8(10): 1265-1275.
- [29] REES H C, MADDISON B C, MIDDLEDITCH D J, et al. REVIEW: the detection of aquatic animal species using environmental DNA-a review of eDNA as a survey tool in ecology[J]. Journal of Applied Ecology, 2014, 51(5):1450-1459.
- [30] CHENG R L, LUO Y, ZHANG Y F, et al. Research on fish diversity of the cascade hydropower reservoir area of the Wujiang River based on environmental DNA technology [J/OL]. Journal of Fisheries of China:1-19.
- [31] LIN J Z, JIANG Y Z, SUN P, et al. Application and evaluation of environmental DNA technology in fish diversity research in Xiangshan Bay[J]. Journal of Fishery Sciences of China, 2021, 8(2): 205-214.
- [32] VALENTINI A, TABERLET P, MIAUD C, et al. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding[J]. Molecular Ecology, 2016, 25(4):929-942.
- [33] ZHAO M, ZHAO M D, MA C Y, et al. Studies on the application of the environmental DNA in aquatic ecosystem[J]. Journal of Fishery Sciences of China, 2018, 25(4): 714-720.
- [34] BOOTHROYD M, MANDRAK N E, FOX M, et al. Environmental DNA (eDNA) detection and habitat occupancy of threatened spotted gar (*Lepisosteus oculatus*) [J]. Aquat Conserv, 2016, 26(6):1107-1119.
- [35] JANOSIK A M, JOHNSTON C E. Environmental DNA as an effective tool for detection of imperiled fishes[J]. Environ Biol Fish, 2015, 98(8): 1889-1893.
- [36] BARNES M A, TURNER C R. The ecology of environmental DNA and implications for conservation genetics[J]. Conservation Genetics, 2015, 17(1):1-17.
- [37] CRISTESCU M E, HEBERT P D N. Uses and misuses of environmental DNA in biodiversity science and conservation [J]. Annu Rev Ecol Evol S, 2018, 49(1): 209-230.
- [38] DAVID S P, CAREN S G, ROBERT S A, et al. Factors influencing detection of eDNA from a stream-dwelling amphibian [J]. Molecular Ecology Resources, 2014, 14(1):109-116.
- [39] ZHAO Y W, CHEN J Q, DONG L, et al. Advances in the application of environmental DNA in aquatic ecosystems[J]. Journal of Agro-Environment Science, 2021, 40(10):2057-2065.