sustainability

MDPI

Article

Environmental DNA Reveals the Influence of Human Activities on Fish Community Variation Across a Large River and Its Connected Lakes

Fan Xiong [1,2,†], Chengchi Fang [2,†], Zhang Jing [3], Sheng Hu [3], Houhua Bing [2], Cheng Wang [2], Yongrui Lu [2], Honghui Zeng [2], Yuxin Hu [3,*], Yingcai Wang [3,*] and Shunping He [2]

[1] Innovation Institute for Biomedical Materials, College of Nursing and Health Management & College of Life Science and Chemistry, Wuhan Donghu College, Wuhan 430040, China; xiongfan@ihb.ac.cn
[2] State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China
[3] Changjiang Basin Ecology and Environment Monitoring and Scientific Research Center, Changjiang Basin Ecology and Environment Administration, Ministry of Ecology and Environment, Wuhan 430010, China
* Correspondence: ssshyx@163.com (Y.H.); wangyingcai@cjjg.mee.gov.cn (Y.W.)
† These authors contributed equally to this work.

Abstract

The aquatic environments of main stems in large rivers and their connected lakes exhibit significant disparities under human activities. Fish are crucial for sustaining the structure and function of aquatic ecosystems as high-level predators. This study investigated fish communities in 192 samples from lakes and rivers across the Yangtze river (YR) basin utilizing environmental DNA (eDNA) technology. Additionally, the environmental variable impact on fish biodiversity in these two aquatic environments was uncovered. Herein, we identified approximately 230 fish taxa in this basin, with lakes and rivers comprising both prevalent and habitat-specific species. Water quality played different roles in affecting fish diversity in these two water systems. The geography traits, including Longitude, Latitude, and Altitude, as well as the water traits conductivity (CD), demonstrated the variance in fish diversity and community composition in both rivers and lakes. The human activity factors, including permanganate index (PMI), chlorophyll-a (CHLA), and SiO_2, elucidated much more variance in fish diversity and community composition in lakes. These findings suggested that human activity factors exert a more significant influence on fish diversity within lakes compared to rivers. Our outcomes document the complex impacts of water quality on fish diversity in different aquatic habitats of the YR basin and emphasize the distinctive considerations required to protect aquatic biodiversity in this basin. However, it should be noted that eDNA technology provides only a single snapshot of community composition. This method possesses limitations common to all approaches (e.g., detection gaps for certain taxa) as well as inherent biases (such as the difficulty in accurately reflecting the abundance and demographic structure of detected species).

Keywords: environmental DNA; Yangtze River basin; fish biodiversity; water quality

check for updates

Academic Editor: Carlos Mestanza-Ramón

Received: 9 September 2025
Revised: 6 November 2025
Accepted: 11 November 2025
Published: 19 November 2025

Citation: Xiong, F.; Fang, C.; Jing, Z.; Hu, S.; Bing, H.; Wang, C.; Lu, Y.; Zeng, H.; Hu, Y.; Wang, Y.; et al. Environmental DNA Reveals the Influence of Human Activities on Fish Community Variation Across a Large River and Its Connected Lakes. *Sustainability* **2025**, *17*, 10353. https://doi.org/10.3390/su172210353

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Large rivers, integral to economic development, are deemed especially susceptible to anthropogenic influences [1]. The continual rise in water pollution in China precipitates cascading impacts on the entirety of river ecosystems, causing alterations to other abiotic factors [2]. The swift pace of urbanization compounds these challenges, as anthropogenic

pressures, notably water pollution driven by the expansion of population density, inflict substantial harm upon aquatic ecosystems [3]. Water pollution is an important factor affecting aquatic biodiversity in large rivers, while the spatio-temporal pattern of fish is directly affected by the hydrodynamics and physical and chemical characteristics of the water body [4]. Large rivers are typically connected to numerous adjacent lakes; the ecological dynamics of rivers and lakes exhibit distinctions, such as the water flow, oxygen levels, and nutrient concentrations, as well as the development of algae and plankton [5]. It has been reported that significant differences existed in the relationship between water quality pollution and fish communities in a megacity in the lentic and lotic systems [6]. However, it remains unclear whether such differences persist in large rivers and their connected lakes.

The Yangtze river (YR), a major river in China, spans 6363 km and passes across Tibet, Yunnan, Sichuan, and seven other provinces. The YR displayed a significant freshwater habitat variety, including alpine streams in the Himalayan mountain range, lowland rivers, floodplains, large lakes, and estuarine habitats [7]. The YR basin (a complex riverine-lacustrine network) is particularly abundant in fish fauna, exhibiting high species richness and endemism. As a result, it is a worldwide important region for conserving fish biodiversity [8,9]. The conservation of fish species in the YR has consistently been a subject of considerable interest, with several studies exploring the fish community structure and diversity in the YR primary stem [10–12], tributaries [13], and lakes [14]. These studies have examined factors affecting fish community composition in various areas of the YR, including revealing that geographical distance significantly influenced spatial rotation in Poyang lake, a YR tributary, historically and currently [14]. Hu et al. found that altitude, velocity, and dissolved oxygen (DO) were the substantial impacting factors on the fish community in the upper reaches (UP) of the YR (Jinsha river) [12,13].

Environmental DNA (eDNA) is a mixture of DNA from organisms shed into the environment found in various media [15]. Ficetola et al. [16] first used eDNA to detect the American bullfrog (*Rana catesbeiana*) in ponds. The eDNA metabarcoding is a biodiversity assessment tool involving extracting total DNA, PCR amplification, and sequencing [17]. This technique, developed for environmental microbiology, is now widely used for aquatic biodiversity research [18,19]. It is a non-invasive method that avoids the costs and ecological disturbance of traditional sampling methods, promising to enhance the assessment of aquatic biodiversity [20,21].

Recently, researchers have also commenced to use eDNA technology to investigate the relation between environmental variables and fish diversity in the YR. For example, Zhang et al. [22] studied the seasonal variation characteristics of the fish community structure in the YR estuary and adjacent waters; Qian et al. [23] investigated the fish community structure alongside diversity in the YR basin and found significant environmental factors impacting α-diversity, such as chlorophyll-a, chemical oxygen demand, DO, total nitrogen (TN), and elevation. Our aim was to deploy eDNA metabarcoding to explore the fish community compositions in rivers and lakes through the entire YR basin and analyze the response of fish diversity to pollution environmental variables in these two water forms. We examined the fish eDNA composition in 192 samples across the entire YR basin (98 from rivers and 94 from lakes) and investigated the fish community variation across different water bodies (rivers vs. lakes) from an eDNA presence-absence perspective. Afterward, we tested the relations between fish diversity and the water quality indices (WQIs) and landscape variables. Various forms of fish diversity were evaluated to explore the influences of human activity on different dimensions of biodiversity.

2. Materials and Methods

2.1. Sample Collection

Water sampling was conducted in YR Basin (24.49° N to 35.31° N, 92.07° E to 121.95° E) from May to September 2021 (Figure 1). Table S1 presents the location of these samples. Sampling was conducted at 98 sites distributed in the main river and its tributaries (Type: rivers in Table S1) and 94 sites distributed in lakes (Type: lakes in Table S1), including Dianchi, Danjingkou reservoir, and Doting, Poyang, and Chao lakes, and Tai lake. From each site, 3 L of water samples, a prevalent fish eDNA sample volume [24], as gathered at three sites approximately equidistant from one another along the bank of the water body. At each sampling location, 2 L of water sample was gathered from approximately 10 cm beneath the surface and approximately 0.5 m from the shore employing a 2 L sterile PET bottle. All samples were conserved on ice and filtered within 8 h. Each water sample was allocated into two subsamples: one for eDNA extraction and species diversity analyses, and the other for environmental variable analysis. For diversity analysis, Filtration was conducted in a clean laboratory prepared for eDNA processing. Each 1 L of water sample and a filtration blank [double-distilled H_2O (ddH_2O)] were filtered through a pore size of 0.45 μm mixed cellulose esters membrane filter (Merck Millipore Ltd., Carrigtwohill, Ireland). Filtration units and laboratory tools were subjected to a 10% bleach solution (hypochlorous acid final concentration = 1%) and washed twice with ddH_2O between samples to eliminate possible eDNA pollution. Filters were preserved at −20 °C for DNA extraction. For the analysis of environmental variables, each 1 L of water sample was filtered through a 0.22 μm polycarbonate membrane. The pH, salinity, conductivity (CD), DO, total phosphorus (TP), TN, NH_4N, permanganate index (PMI), chlorophyll-a (CHLA), SiO_2, and water temperature were assessed per the previous studies [25].



Figure 1. Map of the YR Basin of the sampling locations (N = 192), including 94 sites from the lakes (colored in red) and 98 sites from rivers (colored green).

2.2. eDNA Processing and Sequencing

The DNA extraction was performed based on Zhang's method [6]. For the PCR, we employed the fish metabarcoding primer set MiFish Universal Teleost Primers [26], the most frequently employed primer pair so far [27]. The PCRs were conducted in triplicate for eDNA extracts and negative controls (filtration, extraction, and no-template PCR blanks), each with uniquely labeled primers to identify each PCR amplicon [28]. Herein, we carried out each PCR in a total volume of 25 μL, which contained 5 μL of 10-fold diluted eDNA

extract (to lower the PCR inhibitors), 0.2 μM of the forward and reverse primers, bovine serum albumin (0.4 μg/L), and 1 × Premix Ex Taq (Takara, Kusatsu, Japan). The PCR program comprised a first denaturation step at 95 °C for 10 min; after that, 45 cycles of 95 °C, 65 °C, and 72 °C for 30, 30, and 60 s, respectively; and a last elongation step at 72 °C for 10 min. Utilizing the agarose gels, the PCR products were evaluated, gathered in equal volumes, and purified via the EasyPure PCR Purification Kit (TransGen Biotech, Beijing, China) for next-generation sequencing. Library construction as well as sequencing were conducted via the Beijing Genomics Institute sequencing service in Wuhan, China, employing 2 × 150 bp paired-end sequencing on a HiSeq 2500 System (Illumina Inc., San Diego, CA, USA).

2.3. The Construction of a Local Barcoding Database

A local barcoding database with fish species discovered in the YR Basin was built to improve the metabarcoding taxonomic assignment [29]. The YR freshwater ichthyofauna checklist was first built according to Shen's paper [30] and then refined dependent on available online catalogs FISHBASE [31] (Table S2). From the list, we acquired 1011 full mitochondrial sequences from the Mitofish database and 116 other 12S sequences (>600 bp) from the NCBI and Bold databases. In total, we included 1127 12S sequences from 327 species in 195 genera. Then, targeted barcoding databases were created from these sequences by running in silico PCR with the MiFish Universal Teleost Primers. The genetic distance analyses were performed among these barcoding sequences using MUSCLE, and the Kimura 2-parameter (K2P) [32] pairwise genetic distances were calculated utilizing the R package Ape 4.1 [33]. Maximum intraspecific distances (Dist.Max.Intra) and nearest neighbor genetic distances (Dist.Near.Neigh) were calculated dependent on the pairwise K2P genetic distance matrix employing the R package Spider 1.5 [34].

2.4. Bioinformatics Processing

The raw sequencing data were converted to fastq format via bcl2fastq v2.20. Other bioinformatic steps were conducted employing cutadapt v2.3 [35] and the data2 package v1.8 [36]. Reads were then dereplicated and denoised with the core data2 denoising algorithm. Denoised reads were paired by requesting a minimum overlap of 20 bp and allowing a maximum of five mismatches, utilizing the function mergePairs in data2 to produce the amplicon sequence variants (ASVs). Chimeras were then de novo eliminated from this ASV table via the consensus approach from the removeBimeraDenovo function. The ASVs were annotated using BLAST v2.9.0 against the local barcoding database constructed in this experiment. A taxon was allocated to each ASV based on the Max score and complete sequence coverage. The Max score combined the incentives for matched nucleotides and penalties for mismatches and gaps, and often provided the same arrangement as the conventional expected value. The ASVs with a maximum similarity of <99% (at 100% query coverage) and with read counts <2 of the total reads were eliminated. The ASVs annotated with the same species in the reference database were treated as a tax unit, and the numbers of ASVs in each tax unit per sample were counted together to form a tax unit table (Table S3). This tax-unit table was used to analyze the sample saturation, taxonomic α- and β-diversity.

2.5. Statistical Analyses

To assess the sampling depth impacts on the uncovered fish species number, we conducted saturation curves for all 16 groups defined with the sampling region (such as the whole basin, rivers, lakes, UR, middle-lower reaches (MLR), Hanjing tributaries, Danjiangkou Reservoir, Dongting lake, Honghu, Poyang lake, Chishui river, Sanxia tributaries, Changjiang Yuan Qu, Dianchi, Chao lake, Tai lake) in the R package vegan. Fish

community data were determined through sampling locations via both incidence- and abundance-based metrics. The incidence-based matrix was presented as the presence or absence of each tax unit, and the abundance-based matrix was presented as the PR (such as the percentage of reads for each tax unit). The taxonomic diversity (including α-diversity) and the classical biodiversity measurement, utilizing Chao1, ACE, Shannon, and Simpson diversity, were evaluated with the R package. Additionally, we measured the β-diversity dependent on Bray–Curtis coefficients employing the vegdist commands in vegan with the R package. Wilcoxon rank sum tests were used to analyze statistical variations in the α- and β-diversity metrics between two different groups. The PERMANOVA was performed via the Bray–Curtis coefficients as the response variables, employing the *adonis* function in vegan with 999 permutations to examine the variations in fish community compositions between two different groups.

Redundancy analysis (RDA) was carried out dependent on the Bray–Curtis distance to define the relation between community composition as well as regional variables. The hierarchical partitioning model was conducted to determine the relative significance of purely spatial and environmental components via the "rdacca.hp" package [37]. The Mantel test was employed to assess the regional variances and water trait influences on fish diversity.

3. Results

3.1. Fish Diversity in the YR Basin

After bioinformatics filtering and standardization of sequencing depth across samples, typically, 176 taxonomic units were identified from 194 sampling locations, including 27 combined tax units (comprising 85 fish species), representing nearly 234 fish species. Of these tax units, 157 (comprising 213 fish species) were reported to be common in the YR basin, with dominant families including Cyprinidae, Cobitidae, Bagridae, Amblycipitidae, Balitoridae, and Tetraodontidae. In addition, 18 tax units (comprising 21 fish species), which were reported to be endangered or vulnerable, were detected, such as *Ochetobius elongates* and *Myxocyprinus asiaticus*. Some invasive species, such as *Gambusia affinis*, *Rhynchocypris lagowskii*, and *Sinibrama macrops*, were determined. The dominant species in the basin, in addition to some small-sized fish such as gobies (Unit18 and *Rhinogobius cliffordpopei*), *Butis koilomatodon*, and *Chaeturichthys stigmatias*, also include economically important fish such as *Cyprinus carpio*, *Hypophthalmichthys nobilis*, *Carassius auratus*, *Ctenopharyngodon idella*, and Unit01 (*Culter* and *Megalobrama*) and Unit02 (*Siniperca*). Detailed information for each Unit is provided in Table S4. Among these 176 tax units, 158 were discovered in both rivers and lakes, while 9 tax units were only found in rivers and 9 were only found in lakes. The species found only in rivers were mostly from the order Perciformes, while the species found only in lakes were mainly minnows from the family Cyprinidae in the order Cypriniformes, especially gobies and fishes in the genus *Acheilognathus* (Figure 2).

3.2. Comparison of Lake and River Fish Communities

By comparing four α-diversity indices, comprising Chao1, ACE, Shannon, and Simpson, we found no significant variances ($p > 0.05$) in these indices between the lakes and rivers groups. The results of the Permanova test using the Bray–Curtis distance indicated significant variations in fish community diversity between lakes and rivers ($p = 0.0002$, $F = 2.864$). Additionally, there were significant variations in fish community diversity between the UR and MLR of the YR ($p = 0.0025$, $F = 2.253$).

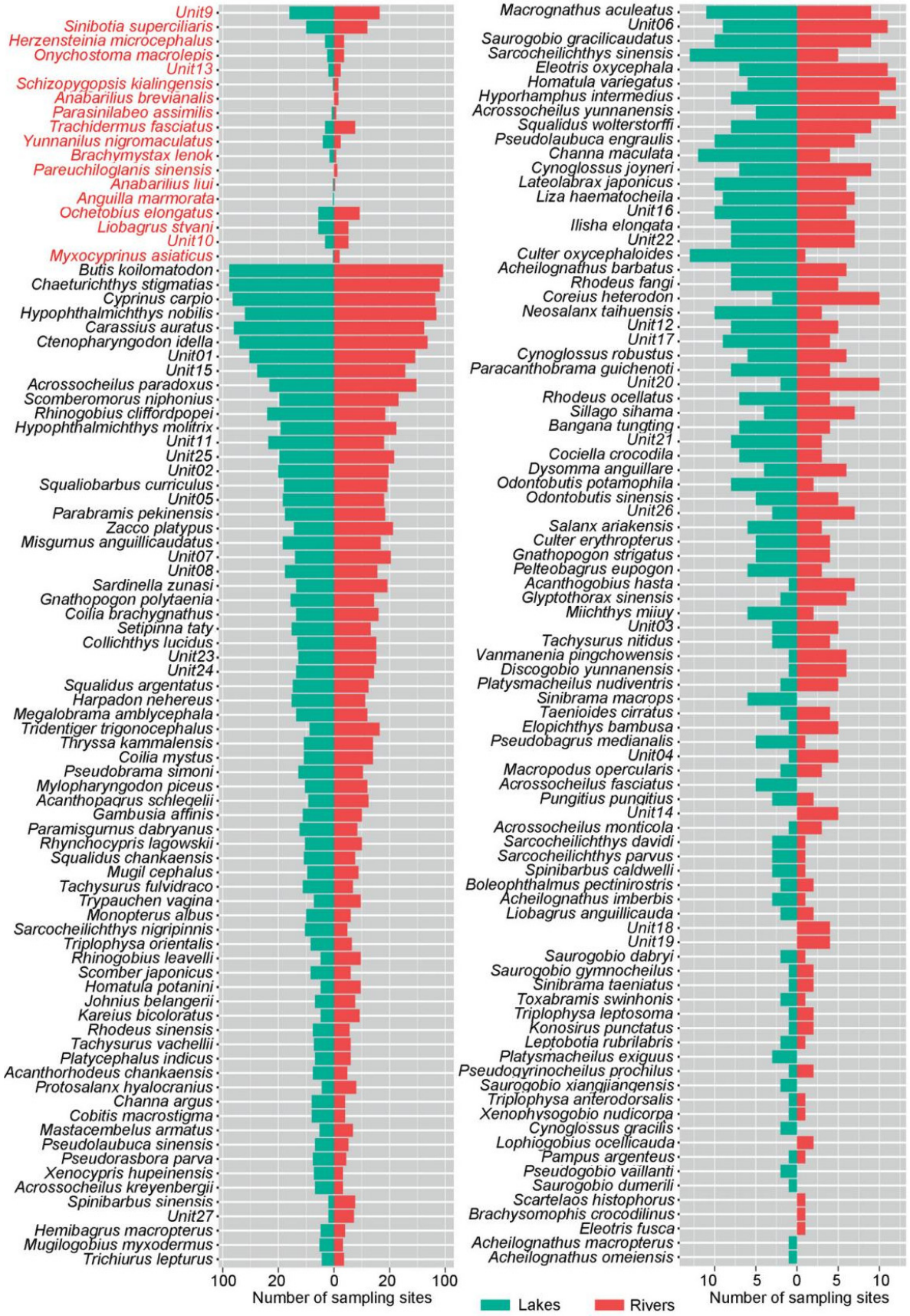


Figure 2. Bar plots of the number of sampling locations for each fish tax unit. The red font tax unit represents the rare and endangered fish detected, while the black font tax unit represents common fish.

Lefse analysis revealed 18 biomarkers for the rivers group, including 6 tax units from Cyprinidae, 5 from Cobitidae, 3 from Gobiidae, 1 from Cottidae, 1 from Eleotridae, 1 from Sisoridae, and 1 from Tetraodontidae. Those fish were mainly small-sized fish, including *Triplophysa* (Unit30 and Unit15), *Homatula*, *Beaufortia* (Unit23), *Jinshaia* (Unit22), and other genera in the family Cobitidae, as well as the carnivorous *Takifugu* genus and the rare fish *Trachidermus fasciatus*. The lakes group had 14 biomarkers, including 7 tax units from Cyprinidae, 2 from Channidae, 1 from Cobitidae, 1 from Gobiidae, 1 from Salangidae, 1 from Clupeidae, and 1 from Pleuronectidae. These biomarkers were mainly composed of economically important and common fish species, with only one species in the family Cobitidae, *Cobitis macrostigma*, two carnivorous *Channa* species, *Channa maculata* and *Channa argus*, and the larger *Culter oxycephaloides*. Pollution-tolerant species such as *Carassius auratus* and *Channa argus* were also identified as biomarkers for the lakes (Figure 3).

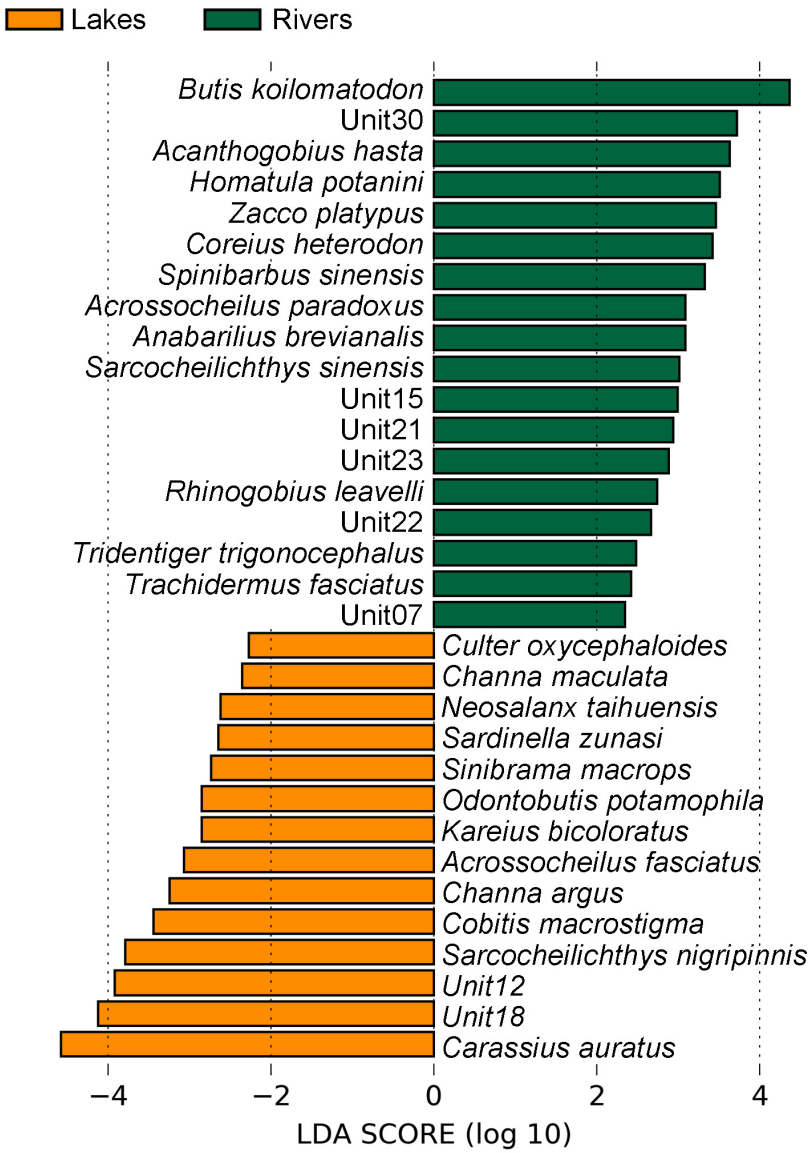


Figure 3. Linear discriminant analysis effect size (LEfSe) analysis between the lakes and rivers group, the identified biomarkers ranked by effect size, and the α-value was <0.05.

3.3. Lake and River Fish Diversity Response to Environmental Variables

Generalized linear models (GLMs) were employed to detect the relative contribution of environmental variables to regional α-diversity. All fish α-diversity indexes were best predicted through models including inorganic nitrogen throughout the whole area. The

result revealed that not all environmental variables in the models participated in the fish diversity to the same degree (Figure 4). For example, the geography traits and CD were the most critical variables for predicting diversity in rivers. The crucial variables for predicting Shannon and Chao1 in lakes were CHLA and turbidity, respectively. These results indicated that human activity factors played a more important role in shaping fish diversity in lakes than in rivers.

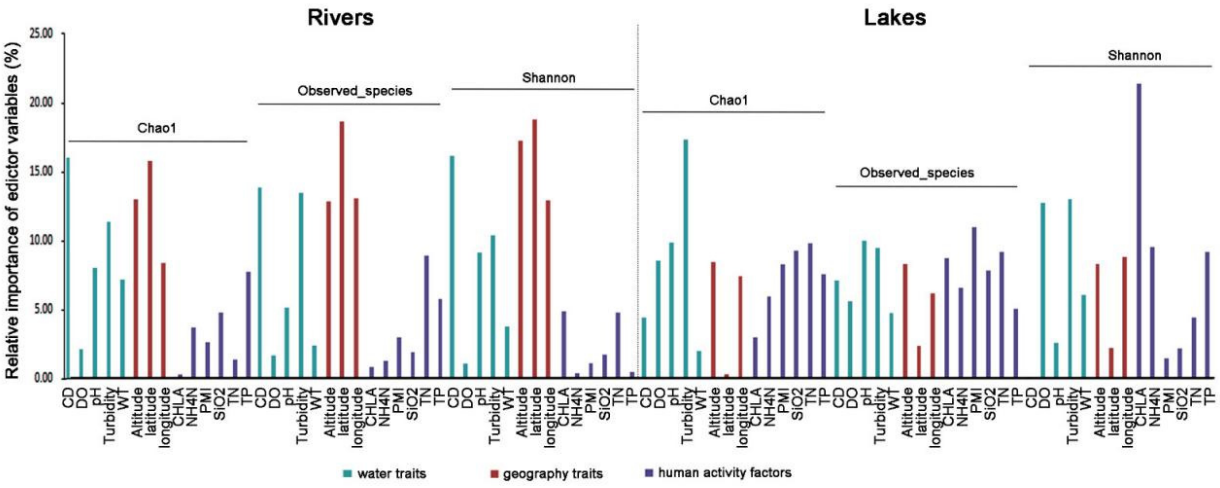


Figure 4. Possible factors of variance in fish diversity and community composition in YR ecosystems.

The Random forest analysis (RDA) was carried out (Figure S1) to examine the environmental alongside geographic variables in shaping the regional fish community composition. The findings uncovered that the pure environmental factor was the prerequisite link to the fish community variation in both groups. Meanwhile, the correlation between environmental factors and fish diversity in both areas was uncovered by the Mantel test (Figure 5). The significance of environmental factors shaping regional fish diversity and community variance altered significantly between the two areas. The outcomes elucidated the geography traits, including Longitude, Latitude, and Altitude, as well as the water traits. CD demonstrated the variance in fish diversity and community composition in both rivers and lakes. The results also revealed that human activity factors, including PMI, CHLA, and SiO_2, manifested the variance in fish diversity and community composition in lakes. Accordingly, the findings suggest that human activity factors exert a more significant influence on fish diversity within lakes compared to rivers.

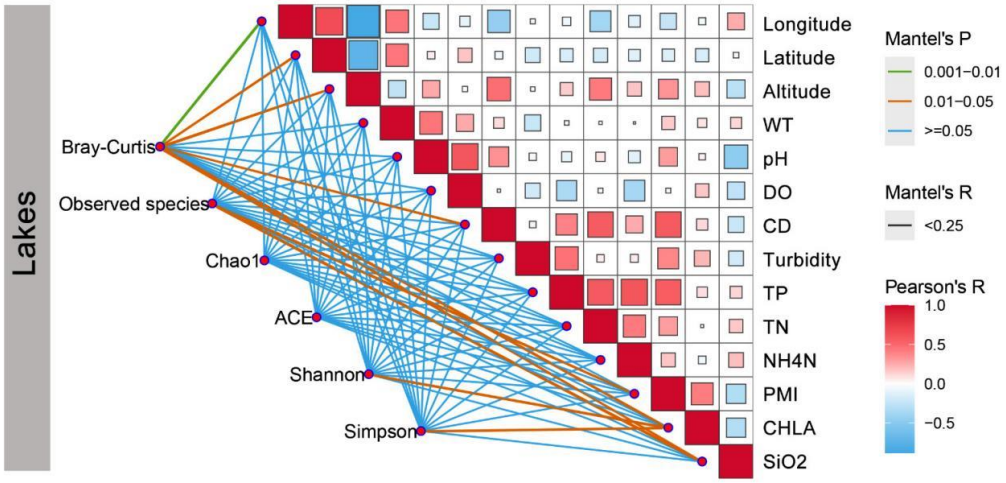


Figure 5. *Cont.*

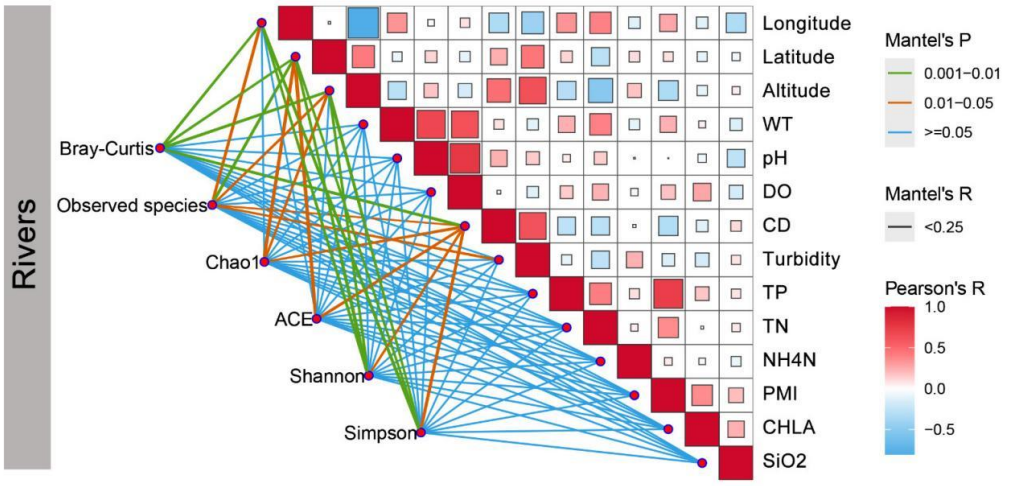


Figure 5. Pairwise correlations of environmental factors. A color gradient represents Spearman's correlation value. Planktonic diversity was linked to all environmental factors by the Mantel test. Edge width aligns with Mantel's correlation coefficient, and edge color represents the statistical significance. "T", water temperature; "NO_3N", nitrate; "NO_2N", nitrite; "NH_4N", ammonium. "OR", observed richness; "PD", phylogenetic distance.

4. Discussion

Compared to the traditional morphological identification of fishes, the eDNA technology is an effective and efficient method for collecting information on water systems and monitoring their ecological environment [38,39]. Since Pont et al. [40] emphasized the detection potential for local fish aggregations utilizing eDNA metabarcoding compared with long-term electrofishing surveys' results in France Rhône river, the application of this technology in large river ecosystems is developing rapidly [41–46]. Herein, we examined 192 eDNA samples from the entire YR Basin, 98 sites distributed in the main river, and 94 sites distributed in lakes, to analyze the spatial distribution of eDNA in the whole basin. Approximately 230 fish species were detected, including 157 taxonomic units (accounting for 213 fish species) of common fish and 18 taxonomic units (accounting for 21 fish species) of rare and endangered fish. Although this result did not present all the fish species (about 450) as traditionally reported in this region [47,48], considering factors such as sampling scales, survey time, and economic expenditure, this outcome still demonstrated the eDNA efficiency in investigating the diversity of fish species in large rivers.

Although we all know significant environmental disparities are exhibited in the main stems of large rivers and their connected lakes under human activities, assessing whether fish responses to human activities differ between these two ecological environments is challenging. Here, we attempted to use eDNA metabarcoding to investigate fish distribution in both ecological environments and conducted correlation statistics to study the effects of environmental factors on fish distribution. Our study showed that the geography traits, including Longitude, Latitude, and Altitude, revealed the variance in fish diversity and community composition in both rivers and lakes. This finding has also been discovered in other traditional research [10,14]. We also found that CD played an important role in shaping fish communities in both water types. The CD can directly reflect the concentration of ions dissolved in water, including salts and minerals. It is well known that CD can affect the growth rate and immune function of fish, which is crucial for their physiological health. Now, we have found that CD is significantly related to the diversity and community structure of fish in the mainstream of large rivers and related lakes. This conclusion further enriches the importance of CD in the health of aquatic environments and has important guiding significance for human activities and water environmental protection.

Our study treats chlorophyll-a (CHLA) as a key environmental variable, influenced by both natural processes and anthropogenic nutrient inputs. Within the intensely impacted Yangtze river basin, elevated CHLA levels serve as a robust proxy for eutrophication. Its influence on fish communities is not direct but operates through indirect, bottom-up ecological effects. High CHLA concentrations support filter-feeding fish (e.g., *Hypophthalmichthys molitrix*, *Aristichthys nobilis*) by enhancing food resources, yet can simultaneously stress visual predators and oxygen-sensitive species by promoting algal blooms and causing hypoxia. Consequently, CHLA plays a complex dual role, making it a crucial variable for understanding community dynamics in eutrophic systems. PMI and SiO_2 were also key environmental factors shaping fish communities, particularly in the lakes. PMI, an indicator of organic pollution, acted as an ecological filter by reducing dissolved oxygen, favoring tolerant species (e.g., *Cyprinus carpio*, *Carassius auratus*) while reducing sensitive taxa, leading to biotic homogenization. SiO_2 influenced the food web through a bottom-up effect by regulating diatom abundance, a crucial food resource. Collectively, these factors-representing "habitat pressure," "energy base," and "food resources"-jointly drive fish communities towards pollution-tolerant and eutrophication-adapted assemblages. In our study, the environmental factors, including PMI, CHLA, and SiO_2, seemed to exert a more significant influence on fish diversity within lakes compared to rivers. Due to their characteristics of closed or semi-closed water bodies, lakes have slower water and material cycles compared to rivers. This slowness makes lake ecosystems more sensitive to human activities. In our research, factors related to human activities are also more closely related to fish diversity. This conclusion suggests that we should pay more attention to the impacts of over-exploitation, pollution discharge, and eutrophication in the conservation of lake fish diversity.

The experimental results showed significant differences in the β diversity of eDNA between the river and lake groups. The species *Butis koilomatodon* was reported to be the most significant biomarker of the river group. This species is a small warm-water near-shore bottom fish that mostly inhabits estuaries, mangrove wetlands, or sandy bottoms along the coast (https://fishdb.sinica.edu.tw/chi/home.php (accessed on 1 June 2025)). Several species of the Cobitidae family, which preferred to live in flowing water, were reported to be the biomarkers of the river group, such as species from *Triplophysa*, *Homatula*, *Metahomaloptera*, *Beaufortia*, *Lepturichthys*, and *Jinshaia*. Three Gobiidae fish were reported to be the biomarkers of the river group, such as *Acanthogobius hasta*, *Rhinogobius leavelli*, and *Tridentiger trigonocephalus*. All three species were riverine fish, as *A. hasta* and *T. trigonocephalus* were reported to prefer to live in coastal, harbor, and estuarine areas, while sometimes living in the lower reaches of rivers or streams, while *R. leavelli* was reported to prefer to live in the rivers and streams (https://www.fishbase.org (accessed on 1 May 2024)). In addition, the estuarine fish *Takifugu*, the riverine fish *Zacco platypus*, *Spinibarbus sinensis*, *Anabarilius brevianalis*, *Euchiloglanis kishinouyei*, *Pareuchiloglanis anteanalis*, *Metahomaloptera omeiensis*, *Beaufortia szechuanensisthe*, the migratory fish *Coreius heterodon*, and the streamlined fish *Acrossocheilus paradoxus* were also the biomarkers of the river group. Otherwise, a rare fish species, *Trachidermus fasciatus,* was also the biomarker of the river group. It was reported that the small migratory *T. fasciatus*, usually distributed along coastal and adjacent freshwater waters, such as estuaries, has hardly been found in the lakes of the YR [49]. At present, in the mainstream of the YR, it was only found in the Yangtze Estuary in 2017, reported by the East China Sea Fisheries Research Institute (http://www.ecsf.ac.cn/info/1020/3218.htm (accessed on 1 May 2025)). In summary, most of the biomarkers of rivers found in our research preferred lotic water environments.

On the other side, most of the biomarkers in the lakes are lacustrine fish, which prefer lentic water environments, including the pollution-tolerant species *Carassius auratus* and

Channa argus, *Sinibrama macrops* inhabiting slow-flowing water, as well as the lake endemic fish *Neosalanx taihuensis*. Finally, our results indicated that the composition of eDNA in the rivers of the YR is different from that in its connected lakes, and this difference is probably caused by the differences in living fish communities in different water bodies.

It is important to note that while eDNA metabarcoding effectively revealed these compositional differences, the technique has inherent limitations. The detected eDNA signal is influenced not only by species presence but also by factors such as DNA degradation, environmental interference (e.g., water flow, temperature), and false positive/false negative results. Therefore, the observed differences reflect a composite picture shaped by both actual biological communities, the dynamics of eDNA in aquatic environments, as well as technique choice [50]. In our experiment, some common and widespread species were not detected, such as *Monopterus albus*, *Channa asiatica*, *Gymnodiptychus pachycheilus*, etc. This absence could be attributed to their low biomass, specific behaviors that reduce eDNA shedding, or sequences that may not be perfectly captured by the selected primers [51]. High sensitivity is a key advantage of eDNA detection [52]; however, it is also usual to find some unlikely species. In the initial taxonomic unit list of our experiment, we also found many unlikely species, such as *Thunnus orientalis*, *Thunnus maccoyii*, *Salmo salar*, etc., which accounted for about 30% of total reads. It seems common to detect unlikely species in eDNA metabarcoding [53]. To minimize the impact of such false-positive signals, a cleaning of the lists is therefore necessary. Here, we performed filtering based on a YR freshwater ichthyofauna checklist. After generating the initial taxonomic unit list, we mandatorily cross-referenced each taxonomic assignment against this expected species list, and any species not present in the reference list were excluded. This rigorous filtering process, while essential for ecological relevance, may also inadvertently remove some true positive detections of rare or previously unrecorded species. Consequently, the final species list represents a conservative and ecologically plausible assemblage, though it may underestimate true biodiversity. Finally, we suggested combining eDNA technology with Underwater Visual Census (UVC): this approach not only enables the verification of eDNA result authenticity (reducing false positives) through UVC but also supplements species detection rates via eDNA [54,55], thereby making the monitoring data more compliant with the requirements for comprehensive assessment of ecological status.

5. Conclusions

In this study, we examined the fish eDNA composition of 192 samples across the entire YR basin. We investigated how environmental variables affect fish biodiversity in rivers and lakes. The result showed that water quality played different roles in affecting fish diversity in these two water systems. The geography traits, including Longitude, Latitude, and Altitude, as well as the water traits CD, manifested the variance in fish diversity and community composition in both rivers and lakes. The human activity factor, including PMI, CHLA, and SiO_2, explained much more variance in fish diversity and community composition in lakes. These findings suggested that human activity factors exert a more significant influence on fish diversity within lakes compared to rivers. Our outcomes document the complex impacts of water quality on fish diversity in different aquatic habitats of the YR basin and reveal the distinctive considerations required to maintain aquatic biodiversity in this basin.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su172210353/s1, Figure S1: Redundancy analysis of species-environment factors relations; Table S1: The location details of all the water samples; Table S2: The YR freshwater ichthyofauna checklist; Table S3: The information of the tax unit table; Table S4: List of the combined tax-units composed with shared-barcode species with maximum Unit K2P genetic distances (Dist.Max.Intra-Unit).

Author Contributions: F.X. and C.F. conducted experiments, data analysis, and manuscript writing. Z.J., S.H. (Sheng Hu), H.B., C.W., Y.L. and H.Z. contributed to collecting specimens. Y.H., Y.W. and S.H. (Shunping He) visualized and designed the experiments. All the authors approved the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (2022YFF0608200), the National Natural Science Foundation of China (32200367), and the Foundation for Outstanding Young and Middle-aged Innovative Research Team in Higher Education Institutions of HuBei Province of China (T2023040).

Institutional Review Board Statement: The experiments were conducted in adherence to the Ethics Committee of the Institute of Hydrobiology at the Chinese Academy of Sciences (CAS). The policies were enacted based on the Chinese Association for Laboratory Animal Sciences and the Institutional Animal Care and Use Committee (IACUC) protocols.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available in the paper and/or the Supplementary Materials. Raw sequencing data are available on NCBI's SRA database BioProject ID: PRJNA957488 (www.ncbi.nlm.nih.gov/bioproject/957488 (accessed on 5 April 2024)).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Allen, G.H.; Pavelsky, T.M. Global extent of rivers and streams. *Science* **2018**, *361*, 585–588. [CrossRef]
2. Reid, A.J.; Carlson, A.K.; Creed, I.F.; Eliason, E.J.; Gell, P.A.; Johnson, P.T.J.; Kidd, K.A.; MacCormack, T.J.; Olden, J.D.; Ormerod, S.J.; et al. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol. Rev. Camb. Philos. Soc.* **2019**, *94*, 849–873. [CrossRef] [PubMed]
3. Su, G.; Logez, M.; Xu, J.; Tao, S.; Villeger, S.; Brosse, S. Human impacts on global freshwater fish biodiversity. *Science* **2021**, *371*, 835–838. [CrossRef]
4. Dai, B.; Jiang, Z.; Wang, C.; Matsuzaki, S.-I.S.; Zhou, L. Abundance-based dissimilarity measurements reveal higher heterogeneity of fish communities in the lotic habitats of the Yangtze-Caizi transitional floodplain. *Ecol. Indic.* **2020**, *112*, 106122. [CrossRef]
5. Ashworth, P.J.; Lewin, J. How do big rivers come to be different? *Earth-Sci. Rev.* **2012**, *114*, 84–107. [CrossRef]
6. Zhang, S.; Zheng, Y.T.; Zhan, A.B.; Dong, C.; Zhao, J.; Yao, M. Environmental DNA captures native and non-native fish community variations across the lentic and lotic systems of a megacity. *Sci. Adv.* **2022**, *8*, eabk0097. [CrossRef]
7. Ye, S.W.; Li, Z.J.; Zhang, T.L.; Liu, J.; Xie, S.; Ye, S.; Li, Z.; Zhang, T.; Liu, J.; Xie, S. Assessing fish distribution and threats to fish biodiversity in the Yangtze River Basin, China. *Ichthyol. Res.* **2014**, *61*, 183–188. [CrossRef]
8. Dudgeon, D. The ecology of tropical Asian rivers and streamsin relation to biodiversity conservation. *Annu. Rev. Ecol. Syst.* **2000**, *31*, 239–263. [CrossRef]
9. Park, Y.; Chang, J.; Lek, S.; Cao, W.; Brosse, S. Conservation strategies for endemic fish species threatened by the ThreeGorges Dam. *Conserv. Biol.* **2003**, *17*, 1748–1758. [CrossRef]
10. Dong, W.W.; Wang, D.Q.; Tian, H.W.; Pu, Y.; Yu, L.; Duan, X.; Liu, S.; Chen, D. Genetic structure of two sympatric gudgeon fishes (Xenophysogobio boulengeri and X. nudicorpa) in the upper reaches of Yangtze River Basin. *PeerJ* **2019**, *7*, e7393. [CrossRef]
11. Liu, M.D.; Wang, D.Q.; Gao, L.; Tian, H.W.; Liu, S.P.; Chen, D.Q.; Duan, X.B. Species diversity of drifting fish eggs in the Yangtze River using molecular identification. *PeerJ* **2018**, *6*, e5807. [CrossRef] [PubMed]
12. Yan, T.M.; He, J.Y.; Yang, D.Y.; Ma, Z.; Chen, H.; Zhang, Q.; Deng, F.; Ye, L.; Pu, Y.; Zhang, M.; et al. Fish community structure and biomass particle-size spectrum in the upper reaches of the Jinsha river (China). *Animals* **2022**, *12*, 3412. [CrossRef]
13. Hu, M.L.; Wang, C.Y.; Liu, Y.Z.; Zhang, X.Y.; Jian, S.Q. Fish species composition, distribution and community structure in the lower reaches of Ganjiang River, Jiangxi, China. *Sci. Rep.* **2019**, *9*, 10100. [CrossRef]

14. Liu, X.J.; Qin, J.J.; Xu, Y.; Zhou, M.; Wu, X.P.; Ouyang, S. Biodiversity pattern of fish assemblages in Poyang Lake basin: Threat and conservation. *Evol. Ecol.* **2019**, *9*, 11672–11683. [CrossRef]
15. Bohmann, K.; Evans, A.; Gilbert, M.T.P.; Carvalho, G.R.; Creer, S.; Knapp, M.; Yu, D.W.; Bruyn, M. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol. Evol.* **2014**, *29*, 358–367. [CrossRef]
16. Ficetola, G.F.; Miaud, C.; Pompanon, F.; Taberlet, P. Species detection using environmental DNA from water samples. *Biol. Lett.* **2008**, *4*, 423–425. [CrossRef] [PubMed]
17. Wang, S.; Yan, Z.; Hanfling, B.; Zheng, X.; Wang, P.; Fan, J.; Li, J. Methodology of fish eDNA and its applications in ecology and environment. *Sci. Total Environ.* **2021**, *755*, 142622. [CrossRef]
18. Hanfling, B.; Handley, L.L.; Read, D.S.; Hahn, C.; Li, J.; Nichols, P.; Blackman, R.C.; Oliver, A.; Winfield, I.J. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Mol. Ecol.* **2016**, *25*, 3101–3119. [CrossRef]
19. Valentini, A.; Taberlet, P.; Miaud, C.; Civade, R.; Herder, J.; Thomsen, P.F.; Bellemain, E.; Besnard, A.; Coissac, E.; Boyer, F.; et al. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.* **2016**, *25*, 929–942.
20. Hering, D.; Borja, A.; Jones, J.I.; Pont, D.; Boets, P.; Bouchez, A.; Bruce, K.; Drakare, S.; Hanfling, B.; Kahlert, M.; et al. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Res.* **2018**, *138*, 192–205. [CrossRef] [PubMed]
21. Laramie, M.B.; Pilliod, D.S.; Goldberg, C.S. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biol. Conserv.* **2015**, *183*, 29–37. [CrossRef]
22. Zhang, H.; Yoshizawa, S.; Iwasaki, W.; Xian, W. Seasonal fish *Assem.* structure using environmental DNA in the Yangtze Estuary and its adjacent waters. *Front. Mar. Sci.* **2019**, *6*, 515. [CrossRef]
23. Qian, M.M.; Wang, Z.Y.; Zhou, Q.; Wang, J.; Shao, Y.; Qiao, Q.; Fan, J.T.; Yan, Z.G. Environmental DNA unveiling the fish community structure and diversity features in the Yangtze River basin. *Environ. Res.* **2023**, *239*, 117198. [CrossRef]
24. Shu, L.; Ludwig, A.; Peng, Z.G. Standards for Methods Utilizing Environmental DNA for Detection of Fish Species. *Genes* **2020**, *11*, 296. [CrossRef] [PubMed]
25. Zhang, Z.; Li, H.; Shen, W.; Du, X.; Li, S.; Wei, Z.; Zhang, Z.; Feng, K.; Deng, Y. The large-scale spatial patterns of ecological networks between phytoplankton and zooplankton in coastal marine ecosystems. *Sci. Total Environ.* **2022**, *827*, 154285. [CrossRef] [PubMed]
26. Miya, M.; Sato, Y.; Fukunaga, T.; Poulsen, J.Y.; Sato, K.; Minamoto, T.; Yamamoto, S.; Yamanaka, H.; Araki, H.; Kondoh, M.; et al. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *R. Soc. Open Sci.* **2015**, *2*, 150088. [CrossRef] [PubMed]
27. Xiong, F.; Shu, L.; Zeng, H.; Gan, X.; He, S.; Peng, Z. Methodology for fish biodiversity monitoring with environmental DNA metabarcoding: The primers, databases and bioinformatic pipelines. *Water Biol. Secur.* **2022**, *1*, 100007. [CrossRef]
28. Zhang, S.; Lu, Q.; Wang, Y.; Yao, M. Assessment of fish communities using environmental DNA: Effect of spatial sampling design in lentic systems of different sizes. *Mol. Ecol. Resour.* **2020**, *20*, 242–255. [CrossRef]
29. Gold, Z.; Curd, E.E.; Goodwin, K.D.; Choi, E.S.; Frable, B.W.; Thompson, A.R.; Walker, H.J.; Burton, R.S.; Kacev, D.; Martz, L.D.; et al. Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. *Mol. Ecol. Resour.* **2021**, *21*, 2546–2564. [CrossRef]
30. Shen, Y.J.; Hubert, N.; Huang, Y.; Wang, X.; Gan, X.; Peng, Z.; He, S. DNA barcoding the ichthyofauna of the Yangtze River: Insights from the molecular inventory of a mega-diverse temperate fauna. *Mol. Ecol. Resour.* **2019**, *19*, 1278–1291. [CrossRef]
31. Froese, R. FishBase. World Wide Web Electron Publication. Available online: https://www.fishbase.org (accessed on 1 May 2024).
32. Kimura, M.A. Simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotidesequences. *J. Mol. Evol.* **1980**, *16*, 111–120. [CrossRef]
33. Paradis, E.; Claude, J.; Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289–290. [CrossRef] [PubMed]
34. Brown, S.D.J.; Collins, R.A.; Boyer, S.; Lefort, M.C.; Malumbres-Olarte, J.; Vink, C.J.; Cruickshank, R.H. Spider: An R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Mol. Ecol. Resour.* **2012**, *12*, 562–565. [CrossRef] [PubMed]
35. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12. [CrossRef]
36. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef]
37. Lai, J.; Zhou, Y.; Peres-Neto, P.R. Generalizing hierarchical and variation partitioning in multiple regression and canonical analyses using the rdacca.hp R package. *Methods Ecol. Evol.* **2022**, *13*, 782–788. [CrossRef]
38. Li, J.; Hatton-Ellis, T.W.; Lawson, H.L.J.; Kimbell, H.S.; Benucci, M.; Peirson, G.; Hnfling, B. Ground-truthing of a fish-based environmental DNA metabarcoding method for assessing the quality of lakes. *J. Appl. Ecol.* **2019**, *56*, 1232–1244. [CrossRef]

39. Pawlowski, J.; Apotheloz, P.G.L.; Altermatt, F. Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Mol. Ecol.* **2020**, *29*, 4258–4264. [CrossRef]
40. Pont, D.; Rocle, M.; Valentini, A.; Civade, R.; Jean, P.; Maire, A.; Roset, N.; Schabuss, M.; Zoring, H.; Dejean, T. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci. Rep.* **2018**, *8*, 82–86. [CrossRef]
41. Czegledi, I.; Saly, P.; Specziar, A.; Preiszner, B.; Szaloky, Z.; Maroda, A.; Pont, D.; Meulenbroek, P.; Valentini, A.; Eros, T. Congruency between two traditional and eDNA-based sampling methods in characterising taxonomic and trait-based structure of fish communities and community-environment relationships in lentic environment. *Ecol. Indic.* **2021**, *129*, 107952. [CrossRef]
42. Goutte, A.; Molbert, N.; Guerin, S.; Richoux, R.; Rocher, V. Monitoring freshwater fish communities in large rivers using environmental DNA metabarcoding and a long-termelectrofishing survey. *J. Fish Biol.* **2020**, *97*, 444–452. [CrossRef]
43. McColl-GausdenLiu, X.; Zeng, H.; Wang, C.; Bo, J.; Gan, X.; Fang, C. Improved genome assembly of Chinese sucker (*Myxocyprinus asiaticus*) provides insights into the identification and characterization of pharyngeal teeth related maker genes in Cyprinoidei. *Water Biol. Secur.* **2022**, *1*, 100049. [CrossRef]
44. McDevitt, A.D.; Sales, N.G.; Browett, S.S.; Sparnenn, A.O.; Mariani, S.; Wangensteen, O.S.; Coscia, I.; Benvenuto, C. Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *J. Fish Biol.* **2019**, *95*, 679–682. [CrossRef]
45. Deiner, K.; Fronhofer, E.A.; Machler, E.; Walser, J.C.; Altermatt, F. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nat. Commun.* **2016**, *7*, 12544. [CrossRef]
46. Pont, D.; Valentini, A.; Rocle, M.; Maire, A.; Delaigue, O.; Jean, P.; Dejean, T. The future of fish-based ecological assessment of European rivers: From traditional EU Water Framework Directive compliant methods to eDNA metabarcoding-based approaches. *J. Fish Biol.* **2021**, *98*, 354–366. [CrossRef]
47. Yang, H.; Shen, L.; He, Y.; Tian, H.; Gao, L.; Wu, J.; Mei, Z.; Wei, N.; Wang, L.; Zhu, T.; et al. Status of aquatic organisms resources and their environments in Yangtze river system (2017–2021). *J. Fish. China* **2023**, *2*, 3–30. [CrossRef]
48. Ye, S.; Li, Z.; Liu, J.; Zhang, T.; Xie, S. Distribution, endemism and conservation status of fishes in the Yangtze River basin, China. *Ecosyst. Biodivers.* **2011**, *1*, 41–66.
49. Li, Y.; Xue, D.; Zhang, B.; Liu, J. Population Genomic Signatures of Genetic Structure and Environmental Selection in the Catadromous Roughskin Sculpin *Trachidermus fasciatus*. *Genome Biol. Evol.* **2019**, *11*, 1751–1764. [CrossRef]
50. Waters, T.; Langlois, K.; Gold, Z.; Theroux, S.; Eagle, R.A. Hidden in plain sight: The invasive macroalga Caulerpa prolifera evades detection by environmental DNA methods. *Environ. DNA* **2023**, *6*, e496. [CrossRef]
51. Roblet, S.; Priouzeau, F.; Gambini, G.; Cottalorda, J.M.; Gastaldi, J.M.; Pey, A.; Raybaud, V.; Romero, S.G.; Serre, C.; Sabourault, C.; et al. From sight to sequence: Underwater visual census vs environmental DNA metabarcoding for the monitoring of taxonomic and functional fish diversity. *Sci. Total Environ.* **2024**, *956*, 177250. [CrossRef] [PubMed]
52. Boussarie, G.; Bakker, J.; Wangensteen, O.S.; Mariani, S.; Bonnin, L.; Juhel, J.-B.; Kiszka, J.J.; Kulbicki, M.; Manel, S.; Robbins, W.D.; et al. Environmental DNA illuminates the dark diversity of sharks. *Sci. Adv.* **2018**, *4*, eaap9661. [CrossRef] [PubMed]
53. Maggio, T.; Cattapan, F.; Falautano, M.; Julian, D.; Malinverni, R.; Poloni, E.; Sanseverino, W.; Todesco, S.; Castriota, L. eDNA metabarcoding analysis as tool to assess the presence of Non-Indigenous Species (NIS): A case study in the Bilge Water. *Diversity* **2023**, *15*, 1117. [CrossRef]
54. Rey, A.; Viard, F.; Lizé, A.; Corre, E.; Valentini, A.; Thiriet, P. Coastal rocky reef fish monitoring in the context of the Marine Strategy Framework Directive: Environmental DNA metabarcoding complements underwater visual census. *Ocean Coast. Manag.* **2023**, *241*, 106625. [CrossRef]
55. Xanthopoulou, P.; Katsanevakis, S.; Ragkousis, M.; Papadakis, O.; Zotou, M.; Kamidis, N.; Wangensteen, O.S.; Papathanasiou, V.; Karampetsis, D.; Mazaris, A.D.; et al. Complementing underwater visual surveys with eDNA metabarcoding to detect Mediterranean non-indigenous fishes. *Mediterr. Mar. Sci.* **2025**, *26*, 216–229. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.