

Assessment of fish biodiversity in four Korean rivers using environmental DNA metabarcoding

Md Jobaidul Alam¹, Nack-Keun Kim¹, Sapto Andriyono², Hee-kyu Choi³, Ji-Hyun Lee⁴ and Hyun-Woo Kim^{1,4}

¹ Interdisciplinary Program of Biomedical, Mechanical and Electrical Engineering, Pukyong National University, Busan, Republic of Korea

² Department of Marine, Fisheries and Marine Faculty, C Campus Jl. Mulyorejo Surabaya, Universitas Airlangga, Surabaya, East Java, Indonesia

³ Molecular Ecology and Evolution Laboratory, Department of Biological Science, College of Science & Engineering, Sangji University, Wonju, Republic of Korea

⁴ Department of Marine Biology, Pukyong National University, Busan, Republic of Korea

ABSTRACT

Environmental DNA (eDNA) metabarcoding is a cost-effective novel approach to estimate biodiversity in an ecosystem. In this study, the MiFish pipeline was employed to test if the system methodology is sufficiently reliable to estimate fish biodiversity in Korean rivers. A total of 125 unique haplotypes and 73 species were identified at the species level from 16 water samples collected from a single survey in four Korean rivers (Hyeongsan, Taehwa, Seomjin, and Nakdong). Among the four rivers, the highest species richness was recorded in the Seomjin River (52 species), followed by the Taehwa (42 species) and Hyeongsan (40 species) rivers. The Nakdong River (26 species) presented the lowest species richness and number of endemic species, presumably due to its metropolitan location and anthropogenic impacts, such as dams or weirs. We were also able to detect that five exotic species (*Carassius cuvieri*, *Cyprinus carpio*, *Cyprinus megalophthalmus*, *Lepomis macrochirus*, and *Micropterus salmoides*) are widely distributed in all surveyed rivers, a situation that might be problematic in terms of conservation. Our findings indicate that the eDNA metabarcoding technique is one of the most cost-effective scientific tools available for the management and conservation of the freshwater fish resources available in Korea. However, the low number of 12S sequences of endemic species in the database and low resolution of the MiFish region for differentiating several taxa should be upgraded for their wide use.

Subjects Aquaculture, Fisheries and Fish Science, Biodiversity, Ecology, Freshwater Biology

Keywords Biodiversity, Korea, Next-generation sequencing, MiFish pipeline, eDNA metabarcoding

INTRODUCTION

Fish communities have been considered as reliable bioindicators of ecosystem status due to their vulnerability to environmental or anthropogenic stresses such as pollution, climate change, or other disturbances in habitats (Dudgeon, 2010). Traditional monitoring methods for fish biodiversity, which have relied on the direct capture or observation of

Submitted 14 November 2019

Accepted 18 June 2020

Published 14 July 2020

Corresponding author

Hyun-Woo Kim, kimhw@pknu.ac.kr

Academic editor

Thiago Parente

Additional Information and
Declarations can be found on
page 20

DOI 10.7717/peerj.9508

© Copyright

2020 Alam et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

specimens, are often costly and time-consuming due to a lack of taxonomic expertise and the necessity of extensive fieldwork. Environmental DNA (eDNA) metabarcoding (detection of multispecies by using degraded DNA from environmental samples) has been proposed as an alternative strategy to analyze fish biodiversity, demonstrating the potential to improve the traditional methods in a cost-effective way (Foote *et al.*, 2012; Kelly *et al.*, 2017; Kelly *et al.*, 2014; Shaw *et al.*, 2016; Stoeckle, Soboleva & Charlop-Powers, 2017; Yamamoto *et al.*, 2017). This technique has been shown to be sensitive as it allows the identification of rarely identified (Pilliod *et al.*, 2013), invasive (Ardura *et al.*, 2015; Cai *et al.*, 2017; Clusa *et al.*, 2017; Dejean *et al.*, 2012; Klymus, Marshall & Stepien, 2017; Takahara, Minamoto & Doi, 2013; Williams *et al.*, 2018), or migratory species (Gustavson *et al.*, 2015; Pont *et al.*, 2018; Yamamoto *et al.*, 2016; Yamanaka & Minamoto, 2016).

Since eDNA metabarcoding analysis of fish biodiversity is mainly based on the amplicon of homologous genes by PCR, universal primers with high taxon-specificity and wide taxon-coverage are essential. Three fish-specific universal primer sets are currently reported: two sets for 12S rRNA regions (EcoPrimers (Riaz *et al.*, 2011) and MiFish (Miya *et al.*, 2015)) and one for the 16S rRNA region (Shaw *et al.*, 2016). Among them, the MiFish primer set demonstrated reliability for eDNA metabarcoding analysis of fish biodiversity in both marine (Ushio *et al.*, 2017; Yamamoto *et al.*, 2017) and continental waters (Sato *et al.*, 2018). More recently, the web-based MiFish pipeline in MitoFish was publicly open (<http://mitofish.aori.u-tokyo.ac.jp/mifish/>), alleviating the time-consuming bioinformatic analysis for the users (Sato *et al.*, 2018).

Although metabarcoding analysis by the MiFish pipeline is one of the most reliable tools at the moment, numbers of MiFish sequences in the database are still one of the last hurdles to overcome for the global use of the MiFish pipeline. Since the average length of the MiFish region is approximately 170 bp, which is much smaller than the typically used 670 bp of the COI barcodes, a high-quality database is critical for successful species assignment. Species identification using the MiFish primer could not discriminate closely related species in several genera, including *Sebastes* spp. and *Takifugu* spp. (Yamamoto *et al.*, 2017). In particular, considering the tremendous diversity of freshwater fishes, the direct application of the MiFish platform may produce a high amount of 'unidentified' records. In addition, a relatively much lower amount of MiFish sequence data (12S region) is currently deposited compared with those of the COI region. Therefore, before the direct application of the MiFish pipeline, the MiFish DNA sequence data for the local freshwater species should be tested for accurate fish biodiversity analysis using eDNA metabarcoding.

In this study, we first employed eDNA metabarcoding analysis of water samples collected from four rivers using the MiFish primer set in order to improve the knowledge regarding freshwater fish biodiversity in Korea. Next, we analyzed the haplotypes obtained by the MiFish pipeline to assess their compatibilities in the identification of endemic species of fishes inhabiting Korean rivers. We also calculated the Shannon-Wiener (H') indices derived from the eDNA metabarcoding results to estimate fish biodiversity in four Korean rivers. Finally, the relationship between the fish assemblage according to the locations in the river was analyzed using heat-map clustering analysis.

MATERIALS AND METHODS

Sample collection and environmental DNA extraction

The eDNA water samples were collected on June 11 and 12, 2018 from 16 stations in the Hyeongsan, Taehwa, Seomjin, and Nakdong rivers, which are four large rivers in the southern part of the Korean Peninsula (Fig. 1 and Table 1). In this study, the sampling stations of each river were categorized as upstream (stations 1 and 2), midstream (station 3), and downstream (Station 4). One liter of water was collected at each station using disposable plastic bottles. After collecting the water, the bottles were immediately stored in an icebox and taken to the laboratory for filtration. Water temperature and salinity were measured with a conductivity meter (CD-4307SD, LUTRON). The water collected was filtered (250 mL × 4) with a 0.45 μm pore-sized GN-6 membrane (PALL Life Sciences, Mexico). The filtration system was cleaned with 10% commercial bleach containing sodium hypochlorite to prevent cross-contamination. After filtration, the membranes were put into 2.0 ml tubes and stored at −20 °C before DNA purification.

Genomic DNA was extracted directly from the membrane filters using the DNeasy® Blood and Tissue Kit (Qiagen, Germany), according to the manufacturer's manual. The membrane filters were cut into smaller pieces before homogenization using a TissueLyser II motorized homogenizer (QIAGEN, Hilden, Germany). The extracted genomic DNA was quantified using a ND-1000 NanoDrop (Thermo Scientific, Waltham, MA, USA), aliquoted, and stored at −20 °C.

Construction of the library and MiSeq sequencing

In order to assess the fish biodiversity, amplicon libraries of partial 12S rRNA region using the MiFish universal primer sets were constructed (Miya *et al.*, 2015). The first PCR was performed to amplify the MiFish regions with an overhanging linker sequence for each Nextera XT index (Illumina, USA). The PCR mixture (20 μL) contained 1.0 μL of the MiFish (forward & reverse) primers (5pmol each), 2.0 μL template, 2.0 μL dNTPs (2.5 mM), 2.0 μL of 10X EX Taq buffer, 0.6 μL DMSO (3%), 0.2 μL of EX Taq Hot Start polymerase (TaKaRa Bio Inc. Japan), and 11.20 μL ultra-pure water. The PCR reaction began with denaturation at 95 °C for 3 min; followed by 30 cycles at 94 °C for 20 s, 65 °C for 15 s, and 72 °C for 15 s; and a final extension at 72 °C for 5 min. The amplicon with the expected size (250–350 bp) was purified with the AccuPrep® Gel Purification Kit (Bioneer, Republic of Korea) after 1.5% agarose gel electrophoresis. The purified amplicons were subjected to additional PCR to link each amplicon with the corresponding Nextera XT index. The second PCR mixture (20 μL) contained 5 μL template, 1 μL of a couple of index primers (10 pmol), 0.5 μL dNTPs (10 mM), 4 μL 5X Phusion HF Buffer, 8.3 μL ultrapure water, and 0.2 μL Phusion Hot Start Flex DNA polymerase (New England Biolabs, Hitchin, UK). The second PCR started at 94 °C for 5 min; followed by 15 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; and an additional 5 min at 72 °C. No noticeable bands were detected in the desired ranges for 16 field negative controls in 1.5% agarose gel electrophoresis. Consequently, the 16 negative controls were discarded from the following analyses. After gel purification, the quality and quantity of the indexed PCR products with the expected sizes were analyzed using the Qubit dsDNAHS Assay Kit

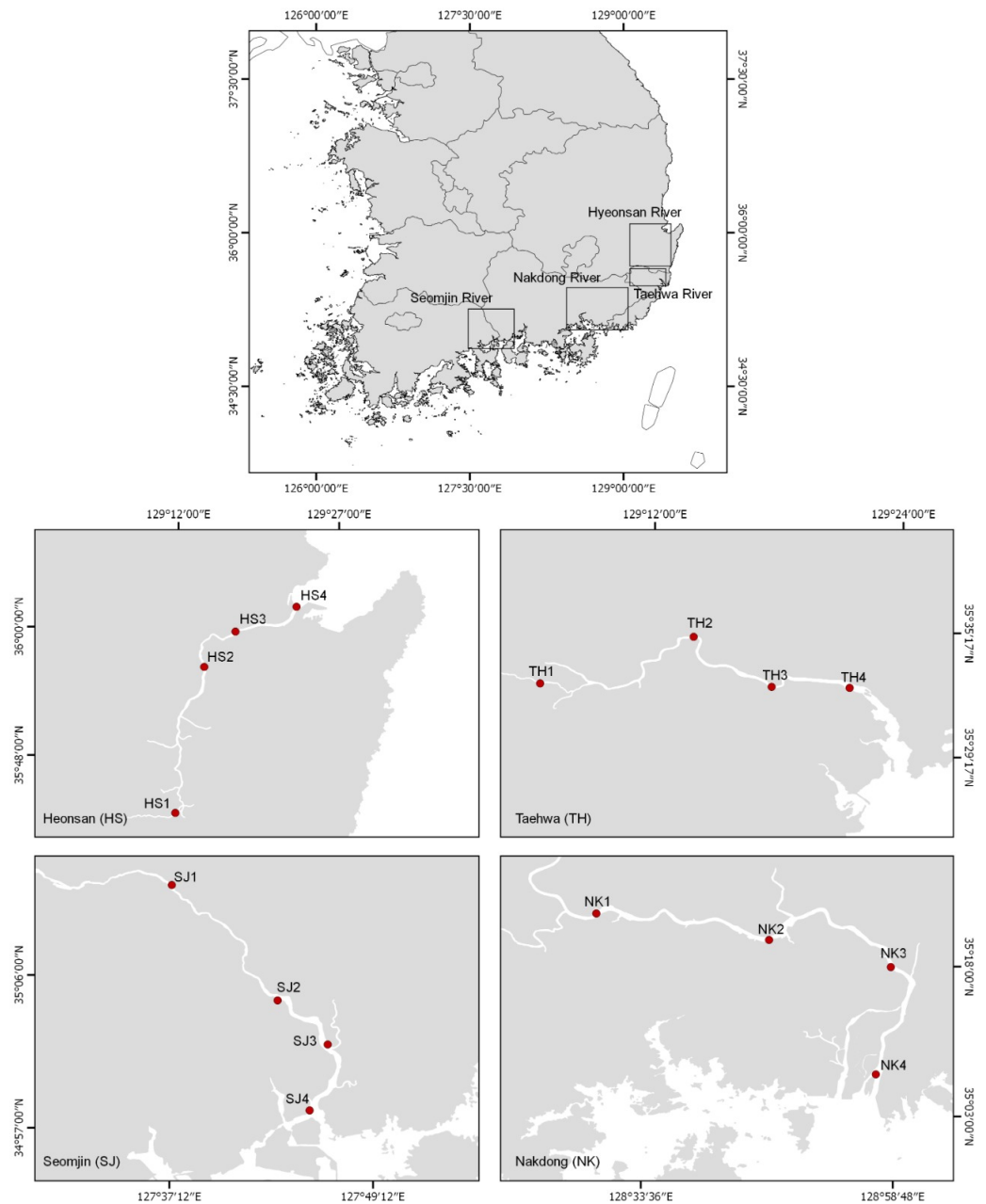


Figure 1 Water sample collection sites for environmental DNA metabarcoding study from four Korean rivers.

Full-size  DOI: [10.7717/peerj.9508/fig-1](https://doi.org/10.7717/peerj.9508/fig-1)

(Invitrogen, Carlsbad, CA, USA), followed by sequencing using the MiSeq platform (2 × 300 bp).

Bioinformatic analysis of the NGS data

The MiSeq raw reads were paired using Python 2.7 (Zhang, 2015), and the paired reads were uploaded to the MiFish pipeline (<http://mitofish.ori.u-tokyo.ac.jp/mifish/>) for further

Table 1 Environmental DNA sample collection sites with physico-chemical parameters of the four Korean rivers.

River	Date	Station	GPS location	Temp. (°C)	Salinity (PSU)
Hyeongsan	2018.06.11	HS1	N35°42'36", E129°11'42"	18.6	1.00
		HS2	N35°56'14", E129°14'24"	19.5	2.02
		HS3	N35°59'32", E129°17'19"	20.0	3.20
		HS4	N36°01'51", E129°23'01"	24.0	20.20
Taehwa	2018.06.11	TH1	N35°32'52", E129°06'27"	19.4	1.02
		TH2	N35°35'07", E129°13'52"	19.8	2.04
		TH3	N35°32'42", E129°17'38"	22.7	14.02
		TH4	N35°32'39", E129°21'24"	19.2	17.80
Seomjin	2018.06.12	SJ1	N35°11'18", E127°37'21"	24.2	0.15
		SJ2	N35°04'30", E127°43'35"	23.4	2.01
		SJ3	N35°01'54", E127°46'32"	23.0	12.9
		SJ4	N34°58'01", E127°45'28"	23.25	16.8
Nakdong	2018.06.12	ND1	N35°23'19", E128°29'09"	24.0	1.92
		ND2	N35°20'40", E128°46'26"	24.1	2.40
		ND3	N35°17'57", E128°58'37"	23.2	2.78
		ND4	N35°07'13", E128°57'07"	22.5	4.50

analyses. In the MiFish pipeline, a low-quality tail of reads ($QV \leq 20$) was trimmed in FASTQC. After taxonomic assignments from the MiFish pipeline, the sequences assigned to OTUs were compared with the GenBank database. If the sequence identity of the query sequence and top BLASTN hit was $\geq 99\%$, the sequence was ascertained as a particular species. If the sequence identity ranged from 97% to 99%, the sequence was ascertained to the genus level, whereas sequences ranging from 97% to 95% were assigned as 'unidentified' genera. The geographic distribution of each species was assessed on the FishBase website (<http://www.fishbase.org/search.php>). Alpha biodiversity was measured using the normalized read numbers from each sampling station of the four rivers sampled. The Shannon-Wiener (H') index indicates the heterogeneity of species or the richness of species in an ecosystem (Gray, 2000; Magurran, 1988). The H' index and the heat map clustering analysis were calculated using the PRIMER® v7 software (Clarke & Gorley, 2015).

RESULTS

Physicochemical parameters

The water temperature of the sample sites ranged from 18.6 °C to 24.20 °C (Table 1). The Hyeongsan River showed the highest temperature difference (5.4 °C) between upstream (HS1) to downstream (HS4), whereas the lowest levels of temperature variation were observed in the Seomjin (0.8 °C) and Nakdong (1.5 °C) rivers. The lowest salinity (0.15 PSU) was measured at station 1 (upstream) of the Seomjin River, while the highest (20.20 PSU) was recorded at station 4 (downstream) of the Hyeongsan River. The salinity level

Table 2 Summary of taxonomic assignment of the MiSeq reads from four Korean rivers.

	Seomjin River	Taehwa River	Hyeongsan River	Nakdong River	Total
Raw reads	561,473	609,755	601,165	543,212	2,315,605
Processed merged reads	553,175	600,744	592,281	534,650	2,280,850
Total haplotypes	76	67	53	42	238 (125) ^a
Haplotypes with species name	61	49	48	31	189 (105) ^a
Total species	52	42	40	26	160 (73) ^a

Notes.

^aFinal number, after removal of duplicated one in brackets.

increased from upstream to downstream in all rivers, except in the Nakdong River, where an artificial dam was constructed to block water from the ocean (Table 1).

Analysis of fish haplotypes obtained using the MiFish pipeline

The reliability of the MiFish pipeline (<http://mitofish.aori.u-tokyo.ac.jp/mifish/workflows/new>) for the biodiversity assessment of fish species inhabiting the sampled rivers was analyzed (Table 2). From 2,315,605 raw reads, 2,280,850 merged reads were obtained by the MiFish pipeline, with a 98.50% yield from the raw reads. A total of 238 representative haplotypes were assigned to the default cutoff sequence identity. Among the 238 haplotypes, 125 unique haplotypes were identified using the phylogenetic tree analysis in the MEGA 7 software (Kumar, Stecher & Tamura, 2016) with a maximum likelihood algorithm (Figs. 2–5). A total of 2,241,130 reads (98.26%) were assigned to 73 confirmed species, 46 genera, and 13 families of Teleostei, with 99% as cutoff identity. The remaining 39,720 reads (49 haplotypes), which showed less than 99% identity, were further assigned to 11 genera and eight unidentified genera (Table 3). A total of 34,755 reads (1.50%) were discarded from further analyses. The highest species number was identified in the family Cyprinidae (35), followed by Gobiidae (11), and Cobitidae (8), while the remaining (19) were from other families of Teleostei. Among them, the highest species number (4 species) was identified in the genus *Acheilognathus*, followed by *Carassius*, *Misgurnus*, *Squalidus*, and *Tridentiger* with three species in each of those genera (Table S1).

Cyprinidae

A total of 65 haplotypes were identified in the family Cyprinidae. Among the 65 haplotypes, 51 were assigned to 35 species of fishes with $\geq 99\%$ of sequence identity to the GenBank database (Fig. 2). Two haplotypes in the genus *Hemibarbus* from the Seomjin River (SJ1) and the Nakdong River (ND2) showed 100% and 99% identity to the sequences of *Hemibarbus labeo* (GenBank Number: [DQ347953](#)) and *Hemibarbus maculatus* ([LC146032](#)) sampled in Korea and Japan, respectively. Among the four endemic species in the genus *Hemibarbus*, *H. labeo* and *H. longirostris* are the most widely distributed species in Korea (Lee et al., 2012). Two haplotypes identified from the Seomjin River (SJ1 and SJ2) and one from Taehwa River (TH1) showed 97% and 95% identity to a sequence of *H. longirostris* ([LC049889](#)), respectively, which suggests that these three haplotypes may be either *H. longirostris* or *H. mylodon* (Fig. 2).

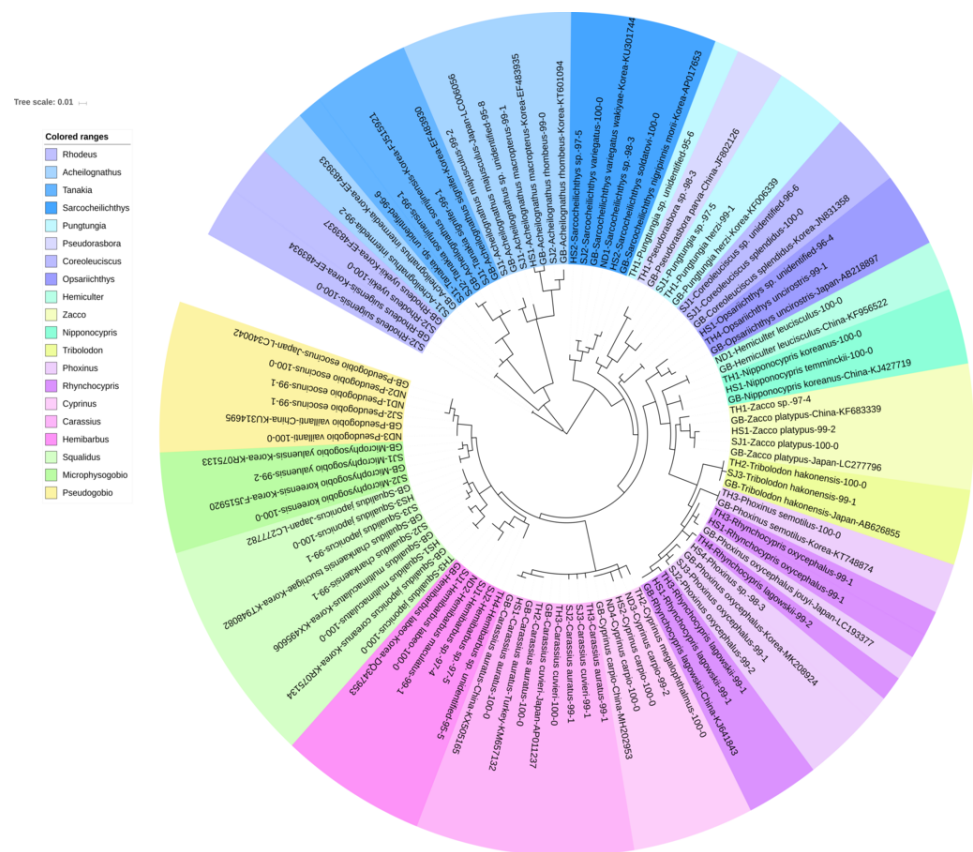


Figure 2 Phylogenetic tree analysis of fish species under the family Cyprinidae detected from four Korean rivers. Phylogenetic tree was constructed by Maximum likelihood (ML) algorithm (MEGA 7.0) under the 1000 replication bootstrap.

Full-size DOI: [10.7717/peerj.9508/fig-2](https://doi.org/10.7717/peerj.9508/fig-2)

Five haplotypes were identified in the genus *Squalidus*. Four species of the genus have been reported from Korean waters: *Squalidus gracilis*, *S. japonicus*, *S. multimaculatus*, and *S. chankaensis* (Kim & Park, 2002). Two haplotypes from the Taehwa (TH3) and Hyeongsan rivers (HS1) showed 100% identity to sequences of *S. japonicus coreanus* (GenBank Number: KR075134) and *S. multimaculatus* (GenBank Number: KT948081), respectively. Another haplotype from the Hyeongsan River (HS3) showed 100% identity to a sequence of *S. japonicus* (GenBank Number: LC277782) sampled in Japan. Two haplotypes from the Seomjin River showed 99% identity to a sequence of *S. chankaensis tsuchigae* (GenBank Number: KT948082) sampled in Korea.

Fishes of the subfamily Acheilognathinae, commonly known as bitterlings, deposit eggs in the gill cavities of freshwater mussels (Kitamura, 2007; Kitamura et al., 2012). Approximately 60 species of bitterlings are considered valid in the genera *Acheilognathus*, *Tanakia*, and *Rhodeus* (Arai, 1988). *Acheilognathus intermedia*, *A. macropterus*, *A. majusculus*, *A. rhombeus*, *Rhodeus suigensis*, *R. uyekii*, *Tanakia somjinensis*, and *T. signifier* were identified with a sequence identity >99% when compared to the GenBank database. Three haplotypes from the Seomjin River showed 99% sequence identity to the respective

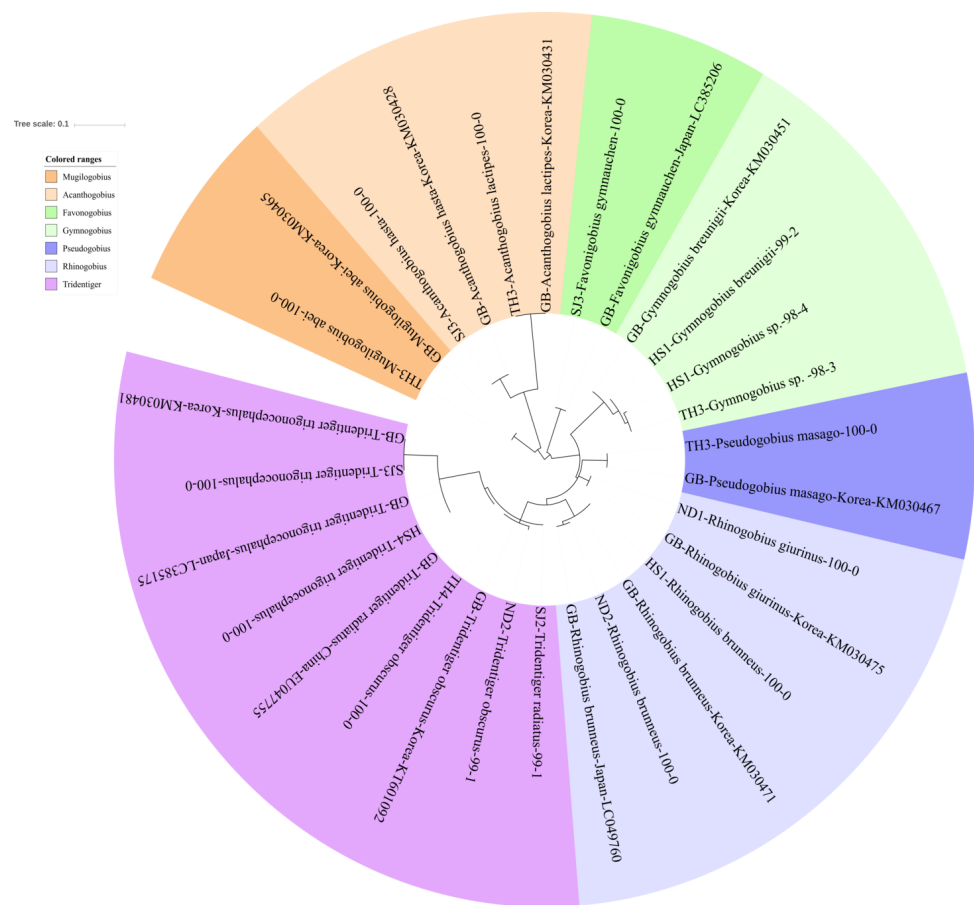


Figure 3 Phylogenetic tree analysis of fish species under the family Gobiidae. Phylogenetic tree was constructed by Maximum likelihood (ML) algorithm (MEGA 7.0) under the 1000 replication bootstrap. Full-size [DOI: 10.7717/peerj.9508/fig-3](https://doi.org/10.7717/peerj.9508/fig-3)

haplotypes of *A. intermedia* (EF483933), *T. somjinensis* (FJ515921), and *T. signifier* (EF483930) sampled in Korea. Among them, *T. somjinensis* and *T. signifier* are endemic to Korea (Kim & Park, 2002). One haplotype from the Taehwa River (TH3) showed 100% identity to a sequence of *Rhynchocypris semotilus* (KT748874) sampled in Korea. This species is currently categorized as Critically Endangered in the Red Data Book of endangered fishes in Korea (Ko, Kim & Park, 2011).

Two sub-species of *Sarcocheilichthys* are known in Korea: *S. nigripinnis morii* and *S. variegates wakiyae* (Kim & Park, 2002). Two haplotypes from the Seomjin (SJ2) and Hyeongsan (HS2) rivers showed 100% and 97%, respectively, identity to a sequence of *S. variegatus wakiyae* (GenBank Number: KU301744) sampled in Korea. One haplotype from the Hyeongsan River (HS2) showed 100% and 99.43% identity to a sequence of *S. soldatovi* (LC146036) and the Korean haplotype of *S. nigripinnis morii* (AP017653) sampled in Japan and Korea, respectively. However, *S. soldatovi* is not currently reported for Korean waters. Therefore, further studies are needed to confirm the occurrence of this species in the Hyeongsan River for conservation purposes.

Tree scale: 0.01

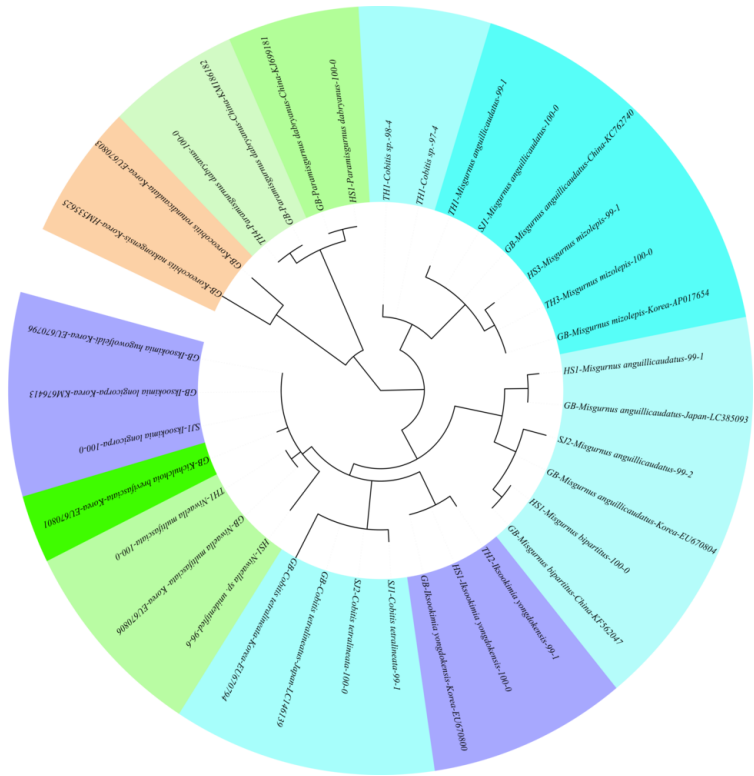
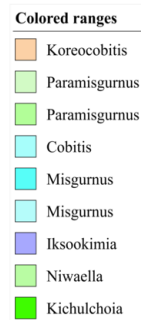


Figure 4 Phylogenetic tree analysis of fish species under the family Cobitidae. Phylogenetic tree was constructed by Maximum likelihood (ML) algorithm (MEGA 7.0) under the 1000 replication bootstrap. Full-size DOI: 10.7717/peerj.9508/fig-4

Gobiidae

We identified 16 haplotypes of the family Gobiidae, representing seven genera and 11 species (Fig. 3). Five haplotypes were identified in the genus *Tridentiger*, which represents the five known species of the genus recorded in Korea (Kim et al., 2005). One haplotype from the Taehwa River (TH4) showed 100% identity with a sequence of *T. obscures* (GenBank Number: KT601092) sampled in Korea. One haplotype from the Hyeongsan River (HS4) showed 100% identity to a sequence of *T. trigonocephalus* (GenBank Number: LC385175) sampled in Japan, and another haplotype from the Seomjin River (SJ3) showed 100% identity to a sequence of *T. trigonocephalus* (GenBank Number: KM030481) sampled in Korea. According to the recovered phylogenetic tree, the *T. trigonocephalus* haplotype from the Seomjin River is different from that of the Hyeongsan River (Fig. 3). All three haplotypes of the genus *Rhinogobius* showed 100% identity to the database. The first and second haplotypes showed 100% identity to sequences of *R. brunneus* sampled in Korea (KM030471) and Japan (LC049760), respectively. The third haplotype showed 100% identity to a sequence of *R. giurinus* sampled in Korea (KM030475). Two haplotypes of *Gymnogobius* sp. from the Taehwa and Hyeongsan rivers showed 98% sequence identity to *G. taranetzi* (GenBank Number: LC385155). Nine species of the genus *Gymnogobius* are currently reported in Korea (Kim et al., 2005), and their MiFish sequences should be supplemented to the GenBank database.

Table 3 List of fish haplotypes with the GenBank numbers identified from the eDNA metabarcoding study of the four rivers.

No.	Family	Haplotype ID	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
1	Gobiidae	SJ3	<i>Acanthogobius hasta</i>	100	KM030428	KM891736	–	
2	Gobiidae	TH3	<i>Acanthogobius lactipes</i>	100	KM030431	–	LC385140	
3	Cyprinidae	SJ1	<i>Acheilognathus intermedia</i>	99	EF483933	–	–	
4	Cyprinidae	HS1	<i>Acheilognathus macropterus</i>	99	EF483935	KJ499466	LC092100	
5	Cyprinidae	SJ1	<i>Acheilognathus majusculus</i>	99	–	–	LC006056	
6	Cyprinidae	SJ2	<i>Acheilognathus rhombeus</i>	99	KT601094	–	LC146100	
7	Cyprinidae	SJ1	<i>Acheilognathus</i> sp. (unidentified)	95	–	–	LC006056	
8	Anguillidae	TH4	<i>Anguilla japonica</i>	100	HQ185628	MH050933	LC193417	
9	Cyprinidae	HS1	<i>Carassius auratus</i>	100	–	KX505165	–	
10	Cyprinidae	TH2	<i>Carassius auratus</i>	100	–	–	–	Turkey KM657132
11	Cyprinidae	TH3	<i>Carassius auratus</i>	99	–	AY771781	LC193299	
12	Cyprinidae	SJ2	<i>Carassius auratus</i>	99	–	AY771781	LC193299	
13	Cyprinidae	TH3	<i>Carassius cuvieri</i>	100	–	–	AP011237	
14	Cyprinidae	SJ3	<i>Carassius cuvieri</i>	100	–	–	AP011237	
15	Channidae	TH1	<i>Channa argus</i>	100	–	MG751766	AB972107	
16	Cobitidae	TH1	<i>Cobitis</i> sp.	97	EU670794	–	LC146139	
17	Cobitidae	TH1	<i>Cobitis</i> sp.	97	EU670794	–	LC146139	
18	Cobitidae	SJ2	<i>Cobitis tetralineata</i>	100	EU670794	–	LC146139	
19	Cobitidae	SJ1	<i>Cobitis tetralineata</i>	99	EU670794	–	LC146139	
20	Cyprinidae	SJ1	<i>Coreoleuciscus</i> sp. (unidentified)	96	JN831358	–	AP011258	
21	Cyprinidae	SJ1	<i>Coreoleuciscus splendidus</i>	100	JN831358	–	AP011258	
22	Sinipercidae	HS3	<i>Coreoperca herzi</i>	100	KR075132	–	–	
23	Sinipercidae	SJ1	<i>Coreoperca</i> sp.	97	KR075132	–	–	
24	Cyprinidae	ND4	<i>Cyprinus carpio</i>	100	–	KX710076	AP017363	
25	Cyprinidae	HS2	<i>Cyprinus carpio</i>	100	–	KX710076	AP017363	
26	Cyprinidae	ND3	<i>Cyprinus carpio</i>	99	–	KX710076	AP017363	
27	Cyprinidae	TH2	<i>Cyprinus megalophthalmus</i>	100	–	KR869143	–	
28	Gobiidae	SJ3	<i>Favonigobius gymnauchen</i>	100	–	–	LC385206	
29	Gobiidae	HS1	<i>Gymnogobius breunigii</i>	99	KM030451	–	–	
30	Gobiidae	HS1	<i>Gymnogobius</i> sp.	98	KM030451	–	–	
31	Gobiidae	TH3	<i>Gymnogobius</i> sp.	98	KM030451	–	–	
32	Cyprinidae	SJ1	<i>Hemibarbus labeo</i>	100	DQ347953	KP064328	LC049898	
33	Cyprinidae	ND2	<i>Hemibarbus maculatus</i>	99	–	NC018534	–	
34	Cyprinidae	SJ1	<i>Hemibarbus</i> sp.	97	DQ347953	KP064328	LC049898	
35	Cyprinidae	SJ2	<i>Hemibarbus</i> sp.	97	DQ347953	KP064328	LC049898	
36	Cyprinidae	TH4	<i>Hemibarbus</i> sp. (unidentified)	95	DQ347953	KP064328	LC049898	
37	Cyprinidae	ND1	<i>Hemiculter leucisculus</i>	100	–	–	LC340359	
38	Cobitidae	SJ1	<i>Iksookimia longicorpa</i>	100	KM676413	–	LC146135	
39	Cobitidae	HS1	<i>Iksookimia yongdokensis</i>	100	EU670800	–	–	

(continued on next page)

Table 3 (continued)

No.	Family	Haplotype ID	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
40	Cobitidae	TH2	<i>Iksookimia yongdokensis</i>	99	EU670800	–	–	
41	Pleuronectidae	SJ3	<i>Kareius bicoloratus</i>	100	–	–	AP002951	
42	Clupeidae	TH3	<i>Konosirus punctatus</i>	100	–	KC477844	LC020951	Taiwan AP011612
43	Clupeidae	ND3	<i>Konosirus punctatus</i>	99	–	KC477844	LC020951	Taiwan AP011612
44	Centrarchidae	TH4	<i>Lepomis macrochirus</i>	100	–	JN389795	AP005993	USA KP013118
45	Amblycipitidae	SJ1	<i>Liobagrus</i> sp.	97	KR075136	KX096605	AP012015	
46	Cyprinidae	SJ2	<i>Microphysogobio koreensis</i>	100	FJ515920	–	–	
47	Cyprinidae	SJ1	<i>Microphysogobio yaluensis</i>	99	KR075133	–	AP012073	
48	Centrarchidae	ND1	<i>Micropterus salmoides</i>	100	–	HQ391896	LC069536	USA DQ536425
49	Centrarchidae	HS1	<i>Micropterus salmoides</i>	99	–	HQ391896	LC069536	USA DQ536425
50	Cobitidae	SJ1	<i>Misgurnus anguillicaudatus</i>	100	–	KC762740	–	
51	Cobitidae	TH1	<i>Misgurnus anguillicaudatus</i>	99	–	KC762740	–	
52	Cobitidae	SJ2	<i>Misgurnus anguillicaudatus</i>	99	EU670804	–	–	
53	Cobitidae	HS1	<i>Misgurnus anguillicaudatus</i>	99	–	–	LC385093	
54	Cobitidae	HS1	<i>Misgurnus bipartitus</i>	100	–	KF562047	LC091592	
55	Cobitidae	TH3	<i>Misgurnus mizolepis</i>	100	AP017654	–	–	
56	Cobitidae	HS3	<i>Misgurnus mizolepis</i>	99	AP017654	–	–	
57	Mugilidae	HS1	<i>Mugil cephalus</i>	100	–	KF374974	LC278014	
58	Gobiidae	TH3	<i>Mugilogobius abei</i>	100	KM030465	–	LC421743	Taiwan KF128984
59	Cyprinidae	TH1	<i>Nipponocypris koreanus</i>	100	–	KJ427719	–	
60	Cyprinidae	HS1	<i>Nipponocypris temminckii</i>	100	–	–	AP012116	
61	Cobitidae	TH1	<i>Niwaella multifasciata</i>	100	EU670807	–	LC146133	
62	Cobitidae	HS1	<i>Niwaella</i> sp. (unidentified)	96	EU670807	–	LC146133	
63	Odontobutidae	SJ1	<i>Odontobutis interrupta</i>	100	KR364945	–	–	
64	Odontobutidae	HS1	<i>Odontobutis platycephala</i>	100	KM030426	–	–	
65	Odontobutidae	SJ2	<i>Odontobutis platycephala</i>	99	KM030426	–	–	
66	Cyprinidae	HS1	<i>Opsariichthys</i> sp. (unidentified)	96	–	–	AB218897	
67	Cyprinidae	TH3	<i>Opsariichthys uncirostris</i>	99	–	–	AB218897	
68	Cobitidae	TH4	<i>Paramisgurnus dabryanus</i>	100	–	KM186182	LC146125	
69	Cobitidae	HS1	<i>Paramisgurnus dabryanus</i>	100	–	KJ699181	LC146125	
70	Cyprinidae	SJ2	<i>Phoxinus oxycephalus</i>	99	MK208924	–	AB626852	
71	Cyprinidae	SJ3	<i>Phoxinus oxycephalus</i>	99	MK208924	–	AB626852	
72	Cyprinidae	TH3	<i>Phoxinus semotilus</i>	100	KT748874	–	–	
73	Mugilidae	TH3	<i>Planiliza affinis</i>	100	–	KM925142	LC277843	
74	Mugilidae	SJ2	<i>Planiliza haematocheila</i>	100	–	KJ622047	LC021099	
75	Mugilidae	HS4	<i>Planiliza haematocheila</i>	100	–	KJ622047	LC021099	
76	Bagridae	SJ1	<i>Pseudobagrus koreanus</i>	100	KT601095	–	–	

(continued on next page)

Table 3 (continued)

No.	Family	Haplotype ID	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
77	Bagridae	ND1	<i>Pseudobagrus ussuriensis</i>	100	–	KC188782	–	
78	Bagridae	ND2	<i>Pseudobagrus ussuriensis</i>	99	–	KC188782	–	
79	Cyprinidae	ND2	<i>Pseudogobio esocinus</i>	100	–	–	LC340042	
80	Cyprinidae	ND1	<i>Pseudogobio esocinus</i>	99	–	–	LC340042	
81	Cyprinidae	ND3	<i>Pseudogobio vaillanti</i>	100	–	KU314695	LC146041	
82	Cyprinidae	SJ2	<i>Pseudogobio vaillanti</i>	99	–	KU314695	LC146041	
83	Gobiidae	TH3	<i>Pseudogobius masago</i>	100	KM030467	–	LC049791	
84	Cyprinidae	TH1	<i>Pungtungia herzi</i>	99	KF006339	–	AB239598	
85	Cyprinidae	SJ1	<i>Pungtungia</i> sp.	97	KF006339	–	AB239598	
86	Cyprinidae	TH1	<i>Pungtungia</i> sp. (unidentified)	96	KF006339	–	AB239598	
87	Gobiidae	HS1	<i>Rhinogobius brunneus</i>	100	KT601096	–		
88	Gobiidae	ND2	<i>Rhinogobius brunneus</i>	100			LC049760	
89	Gobiidae	ND1	<i>Rhinogobius giurinus</i>	100	KM030475	KP892753	LC049748	
90	Cyprinidae	SJ2	<i>Rhodeus suigensis</i>	100	EF483934	–	–	
91	Cyprinidae	SJ1	<i>Rhodeus uyeckii</i>	100	EF483937	–	–	
92	Cyprinidae	HS1	<i>Rhynchocypris lagowskii</i>	99	–	KJ641843	–	
93	Cyprinidae	TH3	<i>Rhynchocypris lagowskii</i>	99	–	KJ641843	–	
94	Cyprinidae	TH4	<i>Rhynchocypris lagowskii</i>	99	–	KJ641843	–	
95	Cyprinidae	SJ2	<i>Rhynchocypris oxycephalus</i>	99	–	–	LC193377	
96	Cyprinidae	SJ3	<i>Rhynchocypris oxycephalus</i>	99	–	–	LC193377	
97	Cyprinidae	HS4	<i>Rhynchocypris</i> sp.	98	–	–	LC193377	
98	Cyprinidae	HS2	<i>Sarcocheilichthys soldatovi</i>	100	–	–	LC146036	
99	Cyprinidae	HS2	<i>Sarcocheilichthys</i> sp.	97	KU301744	–	AP012067	
100	Cyprinidae	ND3	<i>Sarcocheilichthys</i> sp.	97	KU301744	–	AP012067	
101	Cyprinidae	SJ2	<i>Sarcocheilichthys variegatus</i>	100	KU301744	–	AP012067	
102	Siluridae	ND1	<i>Silurus asotus</i>	100	–	JX087351	NC015806	
103	Siluridae	TH1	<i>Silurus microdorsalis</i>	99	KT350610	–	–	
104	Siluridae	SJ1	<i>Silurus</i> sp. (unidentified)	96	KT350610	–	–	
105	Sinipercaidae	SJ1	<i>Siniperca scherzeri</i>	100	–	MF966985	–	Taiwan AP014527
106	Cyprinidae	SJ2	<i>Squalidus chankaensis</i>	100	KT948082	–	–	
107	Cyprinidae	HS3	<i>Squalidus japonicus</i>	100	–	–	LC277782	
108	Cyprinidae	SJ3	<i>Squalidus japonicus</i>	99	–	–	LC277782	
109	Cyprinidae	TH3	<i>Squalidus japonicus coreanus</i>	100	KR075134	–	–	
110	Cyprinidae	HS1	<i>Squalidus multimaculatus</i>	100	KX495606	–	–	
111	Bagridae	SJ1	<i>Tachysurus fulvidraco</i>	100	–	KU133295	LC193372	
112	Bagridae	ND2	<i>Tachysurus nitidus</i>	100	–	KC822643	–	
113	Cyprinidae	SJ1	<i>Tanakia signifer</i>	99	EF483930	–	–	
114	Cyprinidae	SJ2	<i>Tanakia somjinensis</i>	99	FJ515921	–	–	
115	Cyprinidae	SJ1	<i>Tanakia</i> sp.(unidentified)	96	FJ515921	–	–	
116	Cyprinidae	TH2	<i>Tribolodon hakonensis</i>	100	–	–	AB626855	
117	Cyprinidae	SJ3	<i>Tribolodon hakonensis</i>	99	–	–	AB626855	

(continued on next page)

Table 3 (continued)

No.	Family	Haplotype ID	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
118	Gobiidae	TH4	<i>Tridentiger obscurus</i>	100	KT601092	MF663787	LC193168	
119	Gobiidae	SJ2	<i>Tridentiger radiatus</i>	99	–	EU047755	–	
120	Gobiidae	ND2	<i>Tridentiger radiatus</i>	99				
121	Gobiidae	SJ3	<i>Tridentiger trigonocephalus</i>	100	KM030481			
122	Gobiidae	HS4	<i>Tridentiger trigonocephalus</i>	100		KT282115	LC385175	
123	Cyprinidae	SJ1	<i>Zacco platypus</i>	100	–		LC277796	
124	Cyprinidae	HS1	<i>Zacco platypus</i>	99		KF683339		
125	Cyprinidae	TH1	<i>Zacco</i> sp.	97		KF683339		

dabryanus, and the two haplotypes of the species identified herein suggest that *P. dabryanus* has been imported from various locations in China. One *haplotype* from the Taehwa River (TH1) showed 100% identity to a sequence of *Niwaella multifaciata* (EU670806) sampled in Korea, while another from the Hyeongsan River (HS1) showed a lower (96%) identity to *Niwaella* sp. Therefore, further studies should be conducted to confirm the presence of species of this genus in the Hyeongsan River.

Other families of Teleostei

In addition to the three main families of Teleostei identified in this study, 27 additional haplotypes were found in the samples. These haplotypes represented 19 species belonging to 14 genera and 11 families, namely Amblycipitidae (1), Anguillidae (1), Bagridae (5 haplotypes), Centrarchidae (3), Channidae (1), Clupeidae (2), Mugilidae (4), Odontobutidae (3), Pleuronectidae (1), Siluridae (3), and Sinipercaidae (3). All the haplotypes of the family Bagridae were clearly identified and included: *Pseudobagrus ussuriensis*, *P. koreanus*, *Tachysurus nitidus*, and *T. fulvidraco* (Fig. 5). Two species of *Silurus* are currently known in Korean rivers, *S. microdorsalis* and *S. asotus* (Park & Kim, 1994). One haplotype from the Taehwa River (TH1) showed 99% identity to a sequence of *Silurus microdorsalis* (GenBank Number: KT350610) sampled in Korea, whereas another haplotype from the Seomjin River (SJ1) showed a lower identity (96%) with *S. microdorsalis* (KT350610) sampled in Korea.

One haplotype of the Amblycipitidae from the Seomjin River showed 97% and 96% identity to *Liobagrus styani* (KX096605) and *L. mediadiposalis* (KR075136), sampled in China and Korea, respectively. These results indicate that haplotypes of the family should be supplemented for accurate identification. Three species of *Odontobutis* are currently known in Korea: *O. interrupta*, *O. platycephala*, and *O. obscura* (Kim et al., 2005). Two of them (*O. interrupta* and *O. platycephala*) were identified in this study with 100% identity to the sequences of *O. interrupta* and *O. platycephala* sampled in Korea (KR364945 and KM030426). Two haplotypes of the genus *Coreoperca* showed 100% and 97% sequence identity to *Coreoperca herzi* (KR075132) sampled in Korea. Since two species of *Coreoperca* are reported to be endemic to the Korean Peninsula (Kim et al., 2005), the second haplotype is most likely *C. kawamebari*, but further studies should be conducted to confirm this identification. Two invasive species of the family Centrarchidae, the Bluegill (*Lepomis*

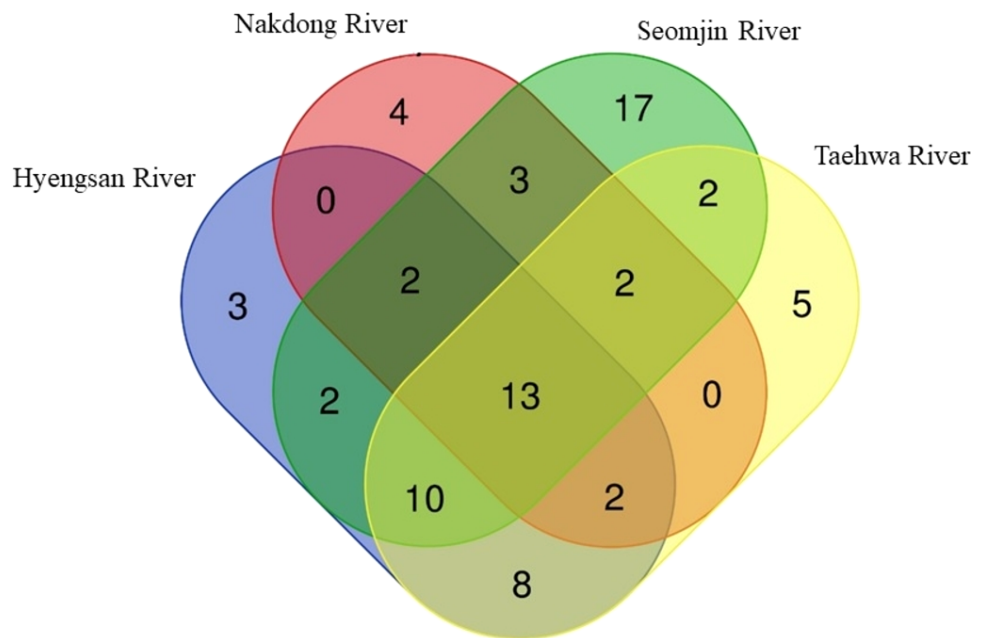


Figure 6 Venn diagram of identified species of fishes in the four Korean rivers. Venn diagram was constructed with an online program (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Full-size DOI: 10.7717/peerj.9508/fig-6

macrochirus) and the Largemouth bass (*Micropterus salmoides*) were also identified in this study. These two species are endemic to North America but were introduced in the Korean Peninsula for aquaculture purposes without considering their impact on local ecosystems.

Fish biodiversity in the four rivers

Fish assemblages in the four rivers included in this study were analyzed. Among the 73 confirmed fish species detected in this study, 13 were identified in all four rivers: *Anguilla japonica*, *Hemibarbus labeo*, *Konosirus punctatus*, *Micropterus salmoides*, *Misgurnus mizolepis*, *Mugil cephalus*, *Opsariichthys uncirostris*, *Pseudorasbora parva*, *Rhinogobius brunneus*, *Rhynchocypris lagowskii*, *Silurus asotus*, *Tridentiger obscurus*, and *Zacco platypus* (Fig. 6). Regardless of sample stations, species of the Cyprinidae appear to be dominant, with average proportions of $47.02 \pm 6.73\%$, followed by the Gobiidae ($15.24 \pm 3.07\%$) and Cobitidae ($9.95 \pm 4.09\%$; Fig. 7). However, the proportions of species in those families were different between upstream and downstream stations. The proportion of Cyprinidae species was higher ($45.27 \pm 9.1\%$) upstream (stations 1 and 2) than downstream ($33.78 \pm 18\%$ at station 4). In contrast, the proportion of Gobiidae was lower ($14.53 \pm 8.28\%$) upstream than downstream (station 4, $19.90 \pm 14\%$).

The highest number of species was recorded in the Seomjin River (52 species), followed by the Taehwa (42 species), Hyeongsan (40 species), and Nakdong (26 species) rivers. A total of 17 species were exclusively recorded in the Seomjin River: *Acanthogobius hasta*, *Acheilognathus intermedia*, *A. majusculus*, *A. rhombeus*, *Cobitis tetralineata*, *Coreoleuciscus splendidus*, *Kareius bicoloratus*, *Microphysogobio yaluensis*, *Phoxinus oxycephalus*, *Pseudobagrus koreanus*, *Rhodeus suigensis*, *R. uyekii*, *Sarcocheilichthys variegatus*, *Siniperca*

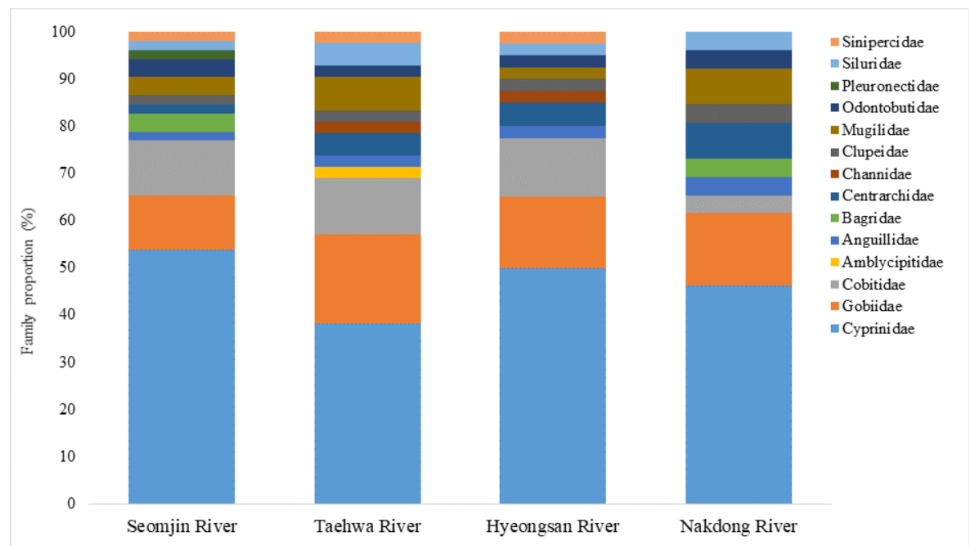


Figure 7 Proportion of families detected from the four Korean rivers by environmental DNA metabarcoding.

Full-size DOI: 10.7717/peerj.9508/fig-7

Table 4 Shannon Index (SI) measured from four Korean rivers by eDNA metabarcoding.

	Seomjin River	Taehwa River	Hyeongsan River	Nakdong River	Average
Station 1	2.197	2.073	1.755	1.777	1.951
Station 2	2.182	1.941	1.709	1.734	1.892
Station 3	2.125	1.631	1.691	1.465	1.728
Station 4	2.105	1.443	1.102	1.008	1.415
Overall SI index	3.48	3.067	2.954	2.864	—

scherzeri, *Squalidus gracilis*, *Tanakia somjinensis*, and *T. signifier*. Five species were only recorded in the Taehwa River: *Acanthogobius lactipes*, *Mugilogobius abei*, *Pseudogobius masago*, *Rhynchocypris semotilus*, and *Silurus microdorsalis*, whereas four species were only identified in the Nakdong River: *Plagiognathops microlepis*, *Pseudobagrus ussuriensis*, *Rhinogobius giurinus*, and *Tachysurus nitidus*. Finally, only three species (*Nipponocypris koreanus*, *Sarcocheilichthys soldatovi*, and *Squalidus multimaculatus*) were exclusively recorded in the Hyeongsan River (Fig. 6).

The highest Shannon index (SI) was identified in the Seomjin River (3.480), followed by the Taehwa (3.067), Hyeongsan (2.954), and Nakdong (2.864) rivers. Among the 16 surveyed stations, station 1 of the Seomjin River (SJ1) showed the highest species richness (2.197), whereas the lowest richness (1.008) was recorded at the station 4 of the Nakdong River (ND4). From upstream to downstream, average species richness decreased from 1.951 to 1.415 (Table 4).

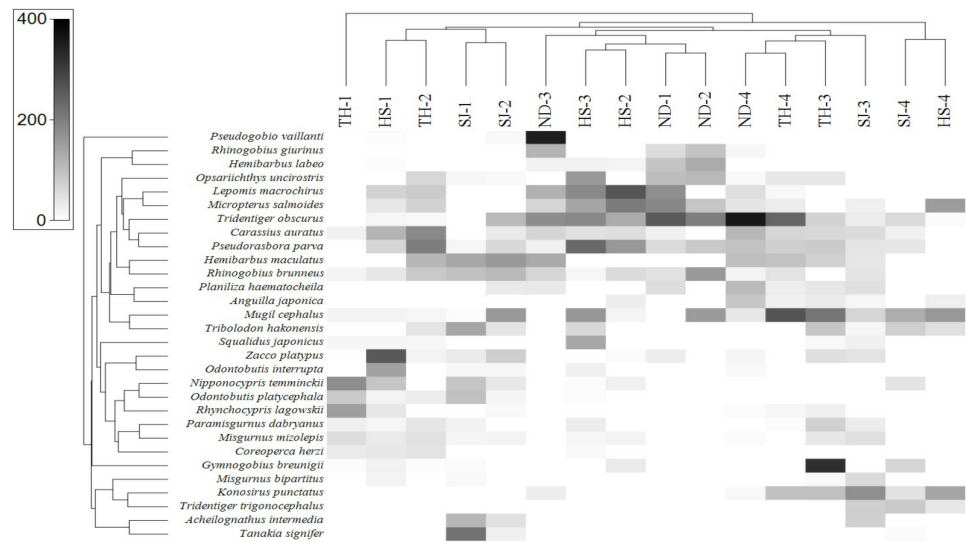


Figure 8 Heat map analysis of top 30 fish species identified in 16 sampling stations of the four Korean rivers. Heat map analysis was constructed by Primer v7 program.

Full-size DOI: 10.7717/peerj.9508/fig-8

Clustering analysis

In order to assess the correlation between the fish assemblage and sample stations, we conducted a heat-map analysis with the 30 most abundant species using Primer software (Clarke & Gorley, 2015). The results indicate the species distribution in different sampling stations (Fig. 8). In upstream sites (Stations 1 and 2), the dominant species were *A. intermedia*, *Coreoperca herzi*, *Misgurnus mizolepis*, *Nipponocypris temminckii*, *Rhynchocypris lagowskii*, *Odontobutis interrupta*, *O. platycephala*, *Tanakia signifier*, and *Zacco platypus*. At station 3, the dominant species were *Gymnogobius breunigii*, *Mugil cephalus*, *Pseudorasbora parva*, *Rhinogobius giurinus*, and *R. brunneus*. Finally, in the downstream sample (Station 4), *Anguilla japonica*, *Konosirus punctatus*, *Mugil cephalus*, *Planiliza haematocheila*, *Tridentiger obscurus*, and *T. trigonocephalus* were identified as the dominant species, all of which were either euryhaline or anadromous (<https://www.fishbase.org>).

DISCUSSION

The results indicate that eDNA metabarcoding using the MiFish pipeline is a useful tool for assessing fish biodiversity in Korean freshwater ecosystems, since a total of 125 unique haplotypes, including at least 73 species, were successfully identified by a single-day survey of 16 sampling stations in four rivers (Figs. 2–5). According to the “Survey and Evaluation of Aquatic Ecosystem Health (SEAEH)”, a total of 130 freshwater fish species were identified from 953 sampling sites that covered most of the Korean rivers and lakes (Yoon et al., 2012). The total number of species confirmed by eDNA metabarcoding was equivalent to approximately 56% of those obtained by the year-long conventional surveys. The efficiency of eDNA metabarcoding might actually be even higher, especially considering the number

of haplotypes successfully identified at the genus and/or family level. This result indicates that eDNA metabarcoding with the MiFish pipeline can significantly contribute to the assessment of freshwater fish biodiversity in Korea, especially considering its relatively lower cost of implementation when compared with more conventional morphology-based surveys. Although the methodology in each research group may be slightly different, similar conclusions have been reached in other studies ([Bista et al., 2017](#); [Deiner et al., 2016](#)). eDNA metabarcoding analysis is also adequate for surveying aquatic species in protected areas, as it minimizes disturbance of vulnerable communities ([Fernandez et al., 2018](#)).

Despite its relevance as a methodology for the assessment of biodiversity, there are still a few shortcomings for a more widespread use of eDNA metabarcoding by the MiFish pipeline. First, MiFish sequence data for endemic species of Korea should be supplemented to the GenBank database. According to the Archive of Korean species (<https://species.nibr.go.kr>), 67 species of freshwater fishes are endemic to Korea, and many of their MiFish sequences are still not available in the GenBank database. In addition to the lack of sequence data, freshwater fishes typically have intra-species genetic distances that are generally higher than those of marine species ([Seehausen & Wagner, 2014](#)). Second, the MiFish primer amplifies the 12S rRNA gene (163–185 bp) region of mitochondrial DNA, which is smaller and less variable than the COI region, which is typically used in species identification ([Ivanova et al., 2007](#)). In fact, the MiFish region was unable to differentiate several closely related marine fish taxa, such as those in the genus *Sebastes* and *Takifugu* ([Sato et al., 2018](#); [Yamamoto et al., 2017](#)). We also found that the average genetic distance of several genera in the family Cyprinidae was low in the MiFish region. For example, the average genetic distance of *Carassius* species was too low (0.01) and the identification at the species level was not possible ([Fig. 2](#)).

Further studies using eDNA metabarcoding might also be relevant to obtain more than biodiversity data, such as the quantitative analysis of fish species. It is difficult to estimate the spatial abundance of eDNA in lotic environments. In fact, many factors should be considered for the quantitative analysis of eDNAs in rivers, including water dynamics ([Deiner & Altermatt, 2014](#); [Jerde et al., 2016](#); [Wilcox et al., 2016](#)) or different decaying times due to different physical, chemical, or biological factors ([Shapiro, 2008](#)). It is generally known that shorter fragments of DNA are degraded slower than larger ones, increasing their probability of detection in natural environments ([Deagle, Eveson & Jarman, 2006](#)). Therefore, it is still too early to adopt eDNA metabarcoding for the quantitative analysis of fish species under natural conditions. For the quantitative study, standardized collection methods and pretreatment procedures for NGS sequencing analysis should also be established. One of the strongest points in the biodiversity survey by eDNA metabarcoding is the quantity of information it can generate compared with more conventional surveys since large datasets are useful for statistical analyses. However, large amounts of data have been produced using different water collection methods, eDNA preparation, sequencing, and bioinformatic analysis platforms by different research groups in different countries. Therefore, the interconversion of data is currently not possible. The establishment of an international standard regarding the overall methodology of eDNA metabarcoding would help researchers to produce more comparable data.

According to the results obtained in this study, the highest species richness was found in the Seomjin River (3.48) compared with those of the other three rivers: the Taehwa River (3.06), Hyeongsan River (2.95), and Nakdong River (2.86). The lower values of species richness detected in the Nakdong, Hyeongsan, and Taehwa rivers are presumably related to the higher anthropogenic alteration of the natural conditions in those rivers. Like most other Korean rivers, these three rivers run through highly populated metropolitan cities, in which rivers are exposed to various human impacts that directly or indirectly promote changes in the diversity and distribution of freshwater fishes (Finkbine, Atwater & Mavinic, 2000). In particular, the lowest species richness (2.86) and number of endemic species (only one, *Odontobutis interrupta*) were identified in the Nakdong River, where the highest number of constructions and population exist among the sampled rivers. Lee et al. (2015) reported only two endemic species (*Coreoperca herzi* and *Odontobutis platycephala*) in the Nakdong River using a conventional catch survey. Moreover, eight endemic species (*Coreoleuciscus splendidus*, *Iksookimia longicorpa*, *Microphysogobio koreensis*, *M. yaluensis*, *Odontobutis interrupta*, *O. platycephala*, *Pseudobagrus koreanus*, and *Squalidus gracilis*) were identified in this study in the Seomjin River, a number that is similar to those obtained in previous studies (Jang, Lucas & Joo, 2003; Lee et al., 2015). Several constructions along urbanized watersheds, including dams and weirs, have caused the simplification and reduction of habitats, decreasing the biodiversity in the river (Nilsson et al., 2005; Riley et al., 2005). In contrast, there is no metropolitan city along the Seomjin River, which is, therefore, less exposed to anthropogenic impacts. A long-term survey should be conducted to establish a clear correlation between anthropogenic factors and fish assemblages in the Korean rivers.

The eDNA metabarcoding analysis also indicates that some exotic fish species are widely distributed in Korean rivers. We were able to identify at least five exotic fish species: *Carassius cuvieri*, *Cyprinus carpio*, *C. megalophthalmus*, *Lepomis macrochirus*, and *Micropterus salmoides* (Table S3). These exotic species may affect native fishes in terms of shelter and spawning sites. They can also disturb the food chain, preying on native fish. In addition, these species have a high reproductive capacity, which makes them important potentially invasive species (Keller & Lake, 2007; Koster, 2002). Surprisingly, our results also revealed that the largemouth bass, *M. salmoides*, and the bluegill, *L. macrochirus*, are likely present in all the sampled rivers. These two species, which are native to North America, were artificially introduced in the 1970s in Korea as freshwater fish stock, without any further consideration of the effects on the freshwater ecosystems of the country. They are now widely distributed throughout the Korean Peninsula, competing with the native species. A long-term survey of these rivers should be conducted to properly assess the potential impacts of these introduced species (Jang et al., 2002; Yoon et al., 2012). Freshwater ecosystems are much more vulnerable to invasive species, causing biodiversity loss and global climate change (Clavero & García-Berthou, 2005), and eDNA metabarcoding analyses would be useful for monitoring the distribution patterns of invasive species in Korean rivers.

ACKNOWLEDGEMENTS

The authors are thankful to the Ministry of Oceans and Fisheries of the Republic of Korea. The authors also thank the reviewers for their valuable comments and suggestions for the manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was a part of the project titled “Long-term change of structure and function in marine ecosystems of Korea”, funded by the Ministry of Oceans and Fisheries, Korea. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Ministry of Oceans and Fisheries, Korea.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Md Jobaidul Alam conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, sample collection, and approved the final draft.
- Nack-Keun Kim analyzed the data, prepared figures and/or tables, sample collection, and approved the final draft.
- Sapto Andriyono performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hee-kyu Choi and Ji-Hyun Lee analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hyun-Woo Kim conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The sequences are available at NCBI Bio Project (BioProject ID: [PRJNA630155](#); Submission ID, SUB7373098).

[SAMN14826804](#), [SAMN14826805](#), [SAMN14826806](#), [SAMN14826807](#), [SAMN14826808](#), [SAMN14826809](#), [SAMN14826810](#), [SAMN14826811](#), [SAMN14826812](#), [SAMN14826813](#), [SAMN14826814](#), [SAMN14826815](#), [SAMN14826816](#), [SAMN14826817](#), [SAMN14826818](#), [SAMN14826819](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.9508#supplemental-information>.

REFERENCES

- Arai R. 1988.** *Acheilognathus melanogaster*, a senior synonym of *A. morioka*, with a revision of the genera of the subfamily Acheilognathinae (Cypriniformes, Cyprinidae). *Bulletin of the National Museum of Nature and Science, Series B* **14**:199–213.
- Ardura A, Zaiko A, Martinez JL, Samulioviene A, Semenova A, Garcia-Vazquez E. 2015.** eDNA and specific primers for early detection of invasive species—a case study on the bivalve *Rangia cuneata*, currently spreading in Europe. *Marine Environmental Research* **112**:48–55 DOI [10.1016/j.marenvres.2015.09.013](https://doi.org/10.1016/j.marenvres.2015.09.013).
- Bista I, Carvalho GR, Walsh K, Seymour M, Hajibabaei M, Lallias D, Christmas M, Creer S. 2017.** Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications* **8**:14087 DOI [10.1038/ncomms14087](https://doi.org/10.1038/ncomms14087).
- Cai W, Ma Z, Yang C, Wang L, Wang W, Zhao G, Geng Y, Douglas WY. 2017.** Using eDNA to detect the distribution and density of invasive crayfish in the Honghe-Hani rice terrace World Heritage site. *PLOS ONE* **12**:e0177724 DOI [10.1371/journal.pone.0177724](https://doi.org/10.1371/journal.pone.0177724).
- Clarke K, Gorley R. 2015.** *Getting started with PRIMER v7*. Plymouth: PRIMER-E, Plymouth Marine Laboratory.
- Clavero M, García-Berthou E. 2005.** Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution* **20**:110 DOI [10.1016/j.tree.2005.01.003](https://doi.org/10.1016/j.tree.2005.01.003).
- Clusa L, Miralles L, Basanta A, Escot C, García-Vázquez E. 2017.** eDNA for detection of five highly invasive molluscs. A case study in urban rivers from the Iberian Peninsula. *PLOS ONE* **12**:e0188126 DOI [10.1371/journal.pone.0188126](https://doi.org/10.1371/journal.pone.0188126).
- Deagle BE, Eveson JP, Jarman SN. 2006.** Quantification of damage in DNA recovered from highly degraded samples—a case study on DNA in faeces. *Frontiers in Zoology* **3**:11 DOI [10.1186/1742-9994-3-11](https://doi.org/10.1186/1742-9994-3-11).
- Deiner K, Altermatt F. 2014.** Transport distance of invertebrate environmental DNA in a natural river. *PLOS ONE* **9**:e88786 DOI [10.1371/journal.pone.0088786](https://doi.org/10.1371/journal.pone.0088786).
- Deiner K, Fronhofer EA, Mächler E, Walser J-C, Altermatt F. 2016.** Environmental DNA reveals that rivers are conveyor belts of biodiversity information. *Nature Communications* **7**:12544 DOI [10.1038/ncomms12544](https://doi.org/10.1038/ncomms12544).
- Dejean T, Valentini A, Miquel C, Taberlet P, Bellemain E, Miaud C. 2012.** Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. *Journal of Applied Ecology* **49**:953–959 DOI [10.1111/j.1365-2664.2012.02171.x](https://doi.org/10.1111/j.1365-2664.2012.02171.x).
- Dudgeon D. 2010.** Prospects for sustaining freshwater biodiversity in the 21st century: linking ecosystem structure and function. *Current Opinion in Environmental Sustainability* **2**:422–430 DOI [10.1016/j.cosust.2010.09.001](https://doi.org/10.1016/j.cosust.2010.09.001).

- Fernandez S, Sandin MM, Beaulieu PG, Clusa L, Martinez JL, Ardura A, García-Vázquez E. 2018.** Environmental DNA for freshwater fish monitoring: insights for conservation within a protected area. *PeerJ* **6**:e4486 DOI [10.7717/peerj.4486](https://doi.org/10.7717/peerj.4486).
- Finkenbine JK, Atwater J, Mavinic D. 2000.** Stream health after urbanization 1. *JAWRA Journal of the American Water Resources Association* **36**:1149–1160 DOI [10.1111/j.1752-1688.2000.tb05717.x](https://doi.org/10.1111/j.1752-1688.2000.tb05717.x).
- Foote AD, Thomsen PF, Sveegaard S, Wahlberg M, Kielgast J, Kyhn LA, Salling AB, Galatius A, Orlando L, Gilbert MTP. 2012.** Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLOS ONE* **7**:e41781 DOI [10.1371/journal.pone.0041781](https://doi.org/10.1371/journal.pone.0041781).
- Gray JS. 2000.** The measurement of marine species diversity, with an application to the benthic fauna of the Norwegian continental shelf. *Journal of Experimental Marine Biology and Ecology* **250**:23–49 DOI [10.1016/S0022-0981\(00\)00178-7](https://doi.org/10.1016/S0022-0981(00)00178-7).
- Gustavson M, Collins P, Finarelli J, Egan D, Conchúir R, Wightman G, King J, Gauthier D, Whelan K, Carlsson J. 2015.** An eDNA assay for Irish *Petromyzon marinus* and *Salmo trutta* and field validation in running water. *Journal of Fish Biology* **87**:1254–1262 DOI [10.1111/jfb.12781](https://doi.org/10.1111/jfb.12781).
- Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN. 2007.** Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes* **7**:544–548 DOI [10.1111/j.1471-8286.2007.01748.x](https://doi.org/10.1111/j.1471-8286.2007.01748.x).
- Jang MH, Kim JG, Park SB, Jeong KS, Cho GI, Joo GJ. 2002.** The current status of the distribution of introduced fish in large river systems of South Korea. *International Review of Hydrobiology* **87**:319–328 DOI [10.1002/1522-2632\(200205\)87:2/3<319::AID-IROH319>3.0.CO;2-N](https://doi.org/10.1002/1522-2632(200205)87:2/3<319::AID-IROH319>3.0.CO;2-N).
- Jang M-H, Lucas MC, Joo G-J. 2003.** The fish fauna of mountain streams in South Korean national parks and its significance to conservation of regional freshwater fish biodiversity. *Biological Conservation* **114**:115–126 DOI [10.1016/S0006-3207\(03\)00016-8](https://doi.org/10.1016/S0006-3207(03)00016-8).
- Jerde CL, Olds BP, Shogren AJ, Andruszkiewicz EA, Mahon AR, Bolster D, Tank JL. 2016.** Influence of stream bottom substrate on retention and transport of vertebrate environmental DNA. *Environmental Science & Technology* **50**:8770–8779 DOI [10.1021/acs.est.6b01761](https://doi.org/10.1021/acs.est.6b01761).
- Keller R, Lake PS. 2007.** Potential impacts of a recent and rapidly spreading coloniser of Australian freshwaters: oriental weatherloach (*Misgurnus anguillicaudatus*). *Ecology of Freshwater Fish* **16**(2):124–132.
- Kelly RP, Closek CJ, O'Donnell JL, Kralj JE, Shelton AO, Samhour JF. 2017.** Genetic and manual survey methods yield different and complementary views of an ecosystem. *Frontiers in Marine Science* **3**:283.
- Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M, Praher E, Caldwell MR. 2014.** Harnessing DNA to improve environmental management. *Science* **344**:1455–1456 DOI [10.1126/science.1251156](https://doi.org/10.1126/science.1251156).
- Kim I-S. 2009.** A review of the spined loaches, family Cobitidae (Cypriniformes) in Korea. *Korean Journal of Ichthyology* **21**:7–28.

- Kim I, Choi Y, Lee C, Lee Y, Kim B, Kim J. 2005.** *Illustrated book of Korean fishes*. Seoul: Kyo-Hak Publishing Co., 615 p.
- Kim I-S, Park J-Y. 2002.** *Freshwater fishes of Korea*. Seoul: Kyo-Hak Publishing Co., 465 p.
- Kitamura J-I. 2007.** Reproductive ecology and host utilization of four sympatric bitterling (Acheilognathinae, Cyprinidae) in a lowland reach of the Harai River in Mie, Japan. *Environmental Biology of Fishes* **78**:37–55.
- Kitamura J, Nagata N, Nakajima J, Sota T. 2012.** Divergence of ovipositor length and egg shape in a brood parasitic bitterling fish through the use of different mussel hosts. *Journal of Evolutionary Biology* **25**:566–573 DOI [10.1111/j.1420-9101.2011.02453.x](https://doi.org/10.1111/j.1420-9101.2011.02453.x).
- Klymus KE, Marshall NT, Stepien CA. 2017.** Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLOS ONE* **12**:e0177643 DOI [10.1371/journal.pone.0177643](https://doi.org/10.1371/journal.pone.0177643).
- Ko M, Kim K, Park J. 2011.** *Red Data Book of endangered fishes in Korea*. Incheon: National Institute of Biological Resources (in Korean).
- Koster WM. 2002.** *Scoping study of the potential spread and impact of the exotic fish Oriental Weatherloach in the Murray-Darling basin, Australia: a draft management strategy*. Melbourne: Arthur Rylah Institute for Environmental Research.
- Kumar S, Stecher G, Tamura K. 2016.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**:1870–1874 DOI [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Lee JW, Yoon JD, Kim JH, Park SH, Baek SH, Yu JJ, Jang MH, Min JI. 2015.** Length-weight relationships for 18 freshwater fish species from the Nakdong River in South Korea. *Journal of Applied Ichthyology* **31**:576–577 DOI [10.1111/jai.12757](https://doi.org/10.1111/jai.12757).
- Lee W-O, Zhang M-M, Oh C-W, Baek J-M, Song K-J. 2012.** Age and growth of barbel steed *Hemibarbus labeo* in Goe-san Lake in Korea. *Fisheries and Aquatic Sciences* **15**:353–359 DOI [10.5657/FAS.2012.0353](https://doi.org/10.5657/FAS.2012.0353).
- Magurran AE. 1988.** *Ecological diversity and its measurement*. Princeton: Princeton University Press.
- Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W. 2015.** MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science* **2**:150088 DOI [10.1098/rsos.150088](https://doi.org/10.1098/rsos.150088).
- Nilsson C, Reidy CA, Dynesius M, Revenga C. 2005.** Fragmentation and flow regulation of the world's large river systems. *Science* **308**:405–408 DOI [10.1126/science.1107887](https://doi.org/10.1126/science.1107887).
- Park S-W, Kim Y-G. 1994.** Studies on disease of catfish, *Silurus asotus*, in Korea. III. *Edwardsiella ictaluri* infection. *Journal of Fish Pathology* **7**:105–112.
- Pilliod DS, Goldberg CS, Laramie MB, Waits LP. 2013.** *Application of environmental DNA for inventory and monitoring of aquatic species*. Reston: US Department of the Interior, US Geological Survey.

- Pont D, Rocle M, Valentini A, Civade R, Jean P, Maire A, Roset N, Schabuss M, Zornig H, Dejean T. 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports* 8:10361 DOI 10.1038/s41598-018-28424-8.
- Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E. 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research* 39:e145–e145.
- Riley SP, Busted GT, Kats LB, Vandergon TL, Lee LF, Dagit RG, Kerby JL, Fisher RN, Sauvajot RM. 2005. Effects of urbanization on the distribution and abundance of amphibians and invasive species in southern California streams. *Conservation Biology* 19:1894–1907 DOI 10.1111/j.1523-1739.2005.00295.x.
- Sato Y, Miya M, Fukunaga T, Sado T, Iwasaki W. 2018. MitoFish and MiFish pipeline: a mitochondrial genome database of fish with an analysis pipeline for environmental dna metabarcoding. *Molecular Biology and Evolution* 35:1553–1555 DOI 10.1093/molbev/msy074.
- Seehausen O, Wagner CE. 2014. Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics* 45:621–651 DOI 10.1146/annurev-ecolsys-120213-091818.
- Shapiro B. 2008. Engineered polymerases amplify the potential of ancient DNA. *Trends in Biotechnology* 26:285–287 DOI 10.1016/j.tibtech.2008.03.005.
- Shaw JLA, Clarke LJ, Wedderburn SD, Barnes TC, Weyrich LS, Cooper A. 2016. Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* 197:131–138 DOI 10.1016/j.biocon.2016.03.010.
- Shimizu T, Takagi M. 2010. Two genetic clades in populations of *Paramisgurnus dabryanus*, an exotic invader in Ehime prefecture. *Japanese Journal of Ichthyology* 57:125–134.
- Stoeckle MY, Soboleva L, Charlop-Powers Z. 2017. Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLOS ONE* 12:e0175186 DOI 10.1371/journal.pone.0175186.
- Takahara T, Minamoto T, Doi H. 2013. Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLOS ONE* 8:e56584 DOI 10.1371/journal.pone.0056584.
- Ushio M, Murakami H, Masuda R, Sado T, Miya M, Sakurai S, Yamanaka H, Minamoto T, Kondoh M. 2017. Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *BioRxiv*.
- Wilcox TM, McKelvey KS, Young MK, Sepulveda AJ, Shepard BB, Jane SF, Whiteley AR, Lowe WH, Schwartz MK. 2016. Understanding environmental DNA detection probabilities: a case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation* 194:209–216 DOI 10.1016/j.biocon.2015.12.023.
- Williams KE, Huyvaert KP, Vercauteren KC, Davis AJ, Piaggio AJ. 2018. Detection and persistence of environmental DNA from an invasive, terrestrial mammal. *Ecology and Evolution* 8:688–695 DOI 10.1002/ece3.3698.

- Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, Minamoto T, Miya M. 2017.** Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports* 7:40368 DOI [10.1038/srep40368](https://doi.org/10.1038/srep40368).
- Yamamoto S, Minami K, Fukaya K, Takahashi K, Sawada H, Murakami H, Tsuji S, Hashizume H, Kubonaga S, Horiuchi T. 2016.** Environmental DNA as a ‘snapshot’ of fish distribution: a case study of Japanese jack mackerel in Maizuru Bay, Sea of Japan. *PLOS ONE* 11:e0149786 DOI [10.1371/journal.pone.0149786](https://doi.org/10.1371/journal.pone.0149786).
- Yamanaka H, Minamoto T. 2016.** The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. *Ecological Indicators* 62:147–153 DOI [10.1016/j.ecolind.2015.11.022](https://doi.org/10.1016/j.ecolind.2015.11.022).
- Yoon J-D, Jang M-H, Kim H-W, Joo G-J. 2012.** *Fish biodiversity monitoring in rivers of South Korea, The biodiversity observation network in the Asia-Pacific Region*. Tokyo: Springer, 175–191.
- Zhang Y. 2015.** *An introduction to python and computer programming, An introduction to python and computer programming*. Singapore: Springer, 1–11.