Assessment of fish biodiversity in four Korean rivers ~~by~~ using ~~the~~ environmental DNA metabarcoding ~~technique~~ (#43068)

First submission

Guidance from your Editor

Please submit by **15 Dec 2019** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

8 Figure file(s)
7 Table file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	***Example***
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Assessment of fish biodiversity in four Korean rivers ~~by~~ using ~~the~~ environmental DNA metabarcoding ~~technique~~

Md Jobaidul Alam [1], **Nack-Keun Kim** [1], **Sapto Andriyono** [2], **Hee-Kyu Choi** [3], **Ji-Hyun Lee** [4], **Hyun-Woo Kim** [Corresp. 1, 4]

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

[2] Fisheries and Marine Faculty, Universitas Airlangga, Surabaya, Indonesia

[3] 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

Corresponding Author: Hyun-Woo Kim
Email address: kimhw@pknu.ac.kr

Environmental DNA (eDNA) metabarcoding is a cost-effective novel approach to estimate the biodiversity in an ecosystem ~~with high efficiency~~. Here, we adopted the MiFish Pipeline to know if the system is reliable to estimate ~~the~~ fish biodiversity from the river water samples. We analyzed 16 water samples from four Korean rivers (Hyeongsan, Taehwa, Seomjin, and Nakdong) and identified 73 ~~fish species~~ ~~simply by~~ a single survey indicating MiFish Pipeline ~~would be~~ a useful tool to estimate the biodiversity without any destructive effects on the ecosystem. Among the ~~4~~ rivers, the highest biodiversity was identified in the Seomjin River (52 species), followed by the Taehwa River (42 species), the Hyeongsan River (40 species), and the Nakdong River (26 species) suggesting the ecosystem in Nakdong River is relatively unhealthy for its metropolitan location. However, ~~we were also able to know~~ that representative haplotype information of the endemic ~~fish species~~ should be supplemented for ~~the~~ better species-identification. Five invasive species (*Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus,* and *Micropterus salmoides*) were also widely distributed in all examined rivers, which would be problematic in the Korean river ecosystem. These findings ~~will be~~ helpful for the effective management or conservation of fish resources in Korean rivers at a relatively low cost.

Assessment of fish biodiversity in four Korean rivers ~~by~~ using ~~the~~ environmental DNA metabarcoding ~~technique~~

Md. Jobaidul Alam[1], Nack-Keun Kim[1], Sapto Andriyono[1,2], Hee-Kyu Choi[3], Ji-Hyun Lee[4], and Hyun-Woo Kim[1,4]*

[1]Interdisciplinary Program of Biomedical, Mechanical and Electrical Engineering, Pukyong National University, Busan, 48513, Republic of Korea

[2]Fisheries and Marine Faculty, C Campus Jl. Mulyorejo Surabaya 60115. Universitas Airlangga, Surabaya, East Java, Indonesia

[3]Molecular Ecology and Evolution Laboratory, Department of Biological Science, College of Science & Engineering, Sangji University, Wonju 26339, Republic of Korea

[4]Department of Marine Biology, Pukyong National University, Busan 48513, Republic of Korea

* Corresponding author:

Hyun-Woo Kim, Ph. D

Department of Marine Biology

Pukyong National University

48513, Republic of Korea

Tel: 82-51-629-5926

Fax: 82-51-629-5930

E-mail: kimhw@pknu.ac.kr

ABSTRACT

Environmental DNA (eDNA) metabarcoding is a cost-effective novel approach to estimate the biodiversity in an ecosystem with high efficiency. Here, we adopted the MiFish Pipeline to know if the system is reliable to estimate the fish biodiversity from the river water samples. We analyzed 16 water samples from four Korean rivers (Hyeongsan, Taehwa, Seomjin, and Nakdong) and identified 73 fish species simply by a single survey indicating MiFish Pipeline would be a useful tool to estimate the biodiversity without any destructive effects on the ecosystem. Among the 4 rivers, the highest biodiversity was identified in the Seomjin River (52 species), followed by the Taehwa River (42 species), the Hyeongsan River (40 species), and the Nakdong River (26 species) suggesting the ecosystem in Nakdong River is relatively unhealthy for its metropolitan location. However, we were also able to know that representative haplotype information of the endemic fish species should be supplemented for the better species-identification. Five invasive species (*Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus,* and *Micropterus salmoides*) were also widely distributed in all examined rivers, which would be problematic in the Korean river ecosystem. These findings will be helpful for the effective management or conservation of fish resources in Korean rivers at a relatively low cost.

Keywords: biodiversity, Korea, next-generation sequencing, metabarcoding, eDNA

INTRODUCTION

As an excellent indicator of biological integrity, monitoring of fish communities has always been a~~s~~ part of ~~a~~ bio-monitoring ~~program~~ to explain the various environmental issues such as pollution, climate changes, and other alterations of habitat either by anthropogenic or natural causes. Traditional monitoring methods for fish biodiversity, which have relied on the direct capture or observations of ~~an individual~~ ~~specimen,~~ are often difficult due to a lack of taxonomic expertise and extensive fieldwork. Environmental DNA (eDNA) metabarcoding (detection of multispecies by using degraded DNA from environmental sample) has been introduced as an alternative strategy to analyze ~~the~~ fish biodiversity and demonstrated a potential to improve the traditional methods ~~in cost~~-effective way (Shaw et al., 2016; Stoeckle, Soboleva, & Charlop-Powers, 2017; Yamamoto et al., 2017). This ~~strategy~~ ~~was so highly~~ ~~sensitive that the~~ rarely ~~identifying~~ species (Pilliod, Goldberg, Laramie, & Waits, 2013), invasive species (Ardura et al., 2015; Cai et al., 2017; Clusa, Miralles, Basanta, Escot, & García-Vázquez, 2017; Dejean et al., 2012; Klymus, Marshall, & Stepien, 2017; Takahara, Minamoto, & Doi, 2013; Williams, Huyvaert, Vercauteren, Davis, & Piaggio, 2018) or ~~migration of fish~~ species (Gustavson et al., 2015; Pont et al., 2018; Yamamoto et al., 2016; Yamanaka & Minamoto, 2016) ~~were also identified by eDNA metabarcoding analysis.~~

Since eDNA metabarcoding analysis for fish biodiversity is mainly based on the amplicon of homologous genes by PCR, ~~the~~ universal primers with high taxon-specificity and wide taxon-coverage are essential. Among two primer sets targeting 12S rRNA regions; Eco Primers (Riaz et al., 2011) and MiFish (M. Miya et al., 2015) and one targeting 16S rRNA (Shaw et al., 2016), MiFish primer set demonstrated its reliability for fish biodiversity analysis both in seawater (Ushio et al., 2017; Yamamoto et al., 2017) and freshwater (Sato, Miya, Fukunaga, Sado, & Iwasaki, 2018). ~~Most~~ recently, the web-based MiFish pipeline in MitoFish was publically open (http://mitofish.aori.u-tokyo.ac.jp/mifish/), which ~~would~~ considerably change the way of the fish biodiversity analysis by eDNA metabarcoding alleviating the time-consuming bioinformatics analysis for the users (Sato et al., 2018).

Although metabarcoding analysis by the MiFish pipeline is one of the most reliable tools at ~~this~~ moment, ~~the~~ numbers of MiFish sequences in the database ~~would be~~ one of the last hurdles to overcome for the global use of MiFish pipeline. Since the average length of the MiFish region is approximately 170 bp, which is much smaller ~~than typically~~ used 670 bp of the COI barcodes, a high-quality database is critical for ~~the~~ successful species assignment. In fact, species identification by MiFish primer could not discriminate closely related species in several genera including *Sebastes* spp. and *Takifugu* spp. (Yamamoto et al., 2017). ~~Especially, freshwater habitats have been fragmented and isolated for the long-term periods and fishes in freshwater generally underwent various types of independent evolutional processes resulting in~~ a tremendous diversity of species (Seehausen & Wagner, 2014). Therefore, before using the MiFish pipeline, DNA sequence information of the MiFish region ~~for~~ the local freshwater ~~fish species~~ should be supplemented for ~~the reliable local~~ fish biodiversity analysis.

In this study, we firstly ~~conducted the~~ eDNA metabarcoding analysis of water samples collected from ~~the~~ four rivers using MiFish primer set to know freshwater fish biodiversity in Korea. Although MiFish successfully presented the high numbers of fish haplotypes inhabiting ~~in~~ Korean rivers, ~~there~~ was difficulty ~~in assigning~~ species names ~~in many~~ of them. ~~From the~~ reason, we used 85 sequences from the NCBI (indicated with GB), among them 44 haplotypes were uploaded from the Korean waters, these sequences were used for constructing the phylogenetic tree to identify the Korean haplotypes. ~~We also identified~~ that a high amount of invasive species currently inhabit ~~in~~ Korean rivers ~~by eDNA metabarcoding 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 ~~of the~~ four large rivers in the southern Korean peninsula ~~including Hyeongsan River, Taehwa River, Seomjin River, and Nakdong River~~ (Fig.1 and Table 1). Four sample stations in each river covered the upstream as well as downstream to the estuary. Two liters of water samples were collected at each station with disposable plastic bottles. After collecting water, the bottles were immediately stored in ice until brought to the laboratory for filtration. ~~After measuring the temperature and salinity with~~ a conductivity meter (CD-4307SD, LUTRON). One liter of water was filtered (250 ml X 4) with 0.45 μm pore-sized GN-6 membrane (PALL Life sciences, Mexico)~~, the~~ filtration system was cleaned up with 10% commercial bleach to prevent ~~the~~ cross-contamination. After filtration, the membranes were put into 2.0 ml tubes and stored at -20˚C before DNA purification.

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

Construction of Library and MiSeq sequencing

To ~~know~~ the fish biodiversity, amplicon libraries of partial 12S rRNA region by the MiFish universal primer set were constructed (Masaki Miya et al., 2015). The first PCR was performed to amplify MiFish regions with an overhanging linker sequence for each Nextera XT index (Illumina, USA). The PCR mixture (20 μL) contained1.0 μL of MiFish (Forward & Reverse) primers (5pmol each), 2.0 μL template, 2.0 μL dNTPs (2.5mM), 2.0 μL of 10X EX Taq buffer, 0.6 μL DMSO (3%), 0.2 μL of EXTaq Hot Start polymerase (TaKaRa Bio Inc. Japan) and 11.20 μL of ultra-pure water. The PCR reaction began with denaturation temperature at 95˚C for 3 min, followed by 30 cycles of 94˚C for 20 sec, 65˚C for 15 sec, and 72˚C for 15 sec with a final extension at 72˚C for 5 min. The amplicon with the expected size (250 bp~350 bp) was purified by AccuPrep® Gel Purification Kit (Bioneer, Republic of Korea) after 1.5 % agarose gel electrophoresis. The purified amplicons were undergone 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, Hitchen, UK). The second PCR conditions began with 94˚C for 5 min followed by 15 cycles of 94˚C for 30 sec, 55 ˚C for 30 sec, and 72˚C for 30 sec, and an additional 5 min at 72 ˚C. In 1.5% agarose gel electrophoresis, no noticeable bands were detected in the desired ranges for 16 field negative controls; consequently, the 16 negative controls were discarded from the next analysis. After gel purification, the quality and quantity of the indexed PCR products with the expected sizes were analyzed by qubit dsDNAHS Assay Kit (Invitrogen, Carlsbad, CA, USA) followed by the sequencing using MiSeq platform (2 X 300 bp).

Bioinformatics analysis of NGS data

In MiFish pipeline, the raw reads by MiSeq sequencing run FASTQC, low-quality tail of reads (QV $\leq$ 20) were trimmed, assembled paired-end reads, removed N-containing reads, filtered reads by length (229 ± 25 bp), run Usearch (minimum read size for filtering =10, and identity threshold for clustering = 0.99), run BLASTN, created multi-FASTA files for each samples, run MAFFT for each samples, run Morphy for each samples, run Morphy against merged samples, run BLASTN, process BLASTN results and finalized the results. The total sequences assigned to OTUs were compared with the GenBank ~~database, if~~ the sequence identity of the query sequence and top BLASTN hit was more than or equal to 99%, then the sequences were ascertained as species. In case of the sequence uniformity within 97% to 98%, the sequence was ascertained as genus, and the sequences having 95% to 97% similarity with the database were assigned as 'unidentified' genus. The distribution ~~for~~ each species was ~~confirmed by~~ FishBase (http://www.fishbase.org).

RESULTS

Physico-chemical parameters

The water temperature of the sample sites ~~ranges~~ from 18.6 ˚C in SJ1 (Table 1). The Hyeongsan River showed the highest difference (5.4 ˚C) in temperature from the upstream (HS1) to the downstream (HS4), whereas ~~the low degree of temperature changes in~~ the Seomjin River (0.8 ˚C) and the Nakdong River (1.5 ˚C). Salinity level increased from ~~the~~ upstream to the downstream in all ~~three~~ rivers except for the Nakdong River, where ~~the~~ artificial ~~dike~~ has been constructed to ~~blocking~~ ~~the~~ water ~~flows~~ from the ~~Ocean~~ (Table 1).

Analysis of ~~the~~ fish haplotypes obtained by the MiFish pipeline

The reliability of MiFish pipeline (http://mitofish.aori.u-tokyo.ac.jp/mifish/workflows/new) for the biodiversity ~~analysis~~ of fish species inhabiting the ~~Korean~~ rivers was analyzed (Table 2). From the 2,315,605 raw reads, 2,280,850 merged reads were obtained by the MiFish pipeline showing 98.50% yields from the raw reads. A total of 238 representative haplotypes were assigned at the default cutoff sequence identity (95%). Among the 238 haplotypes, we found 125 unique haplotypes, which ~~identified~~ by using the phylogenetic tree analysis (Fig. 2-5). At ≥99% identity, ~~total~~ 2,241,130 reads (98.26%) were assigned ~~into~~ 189 haplotypes (73 species), 46 genus and 13 families. The remaining 39,720 reads (49 haplotypes), which showed less than 99% identity were named as putative haplotypes, assigned into 11 genera, and 8 unidentified genera (Table 3). A total of 34,755 reads (1.50%) were discarded from further analysis, because of their lower identity (80% to 95%) to the NCBI database. The highest number of species were identified in the family Cyprinidae (35) followed by Gobiidae (11), Cobitidae (8), and the remaining (19) ~~from the other family.~~ Four species were identified in the genus *Acheilognathus*, ~~while 3 were in genera~~ *Carassius, Misgurnus, Tridentiger*, and *Squalidus* (Table S1).

Cyprinidae

A total of 65 haplotypes were identified in the family Cyprinidae, in which 51 haplotypes were assigned to 35 fish species with ≥ 99% sequence identity to the NCBI database, and 8 haplotypes showed between 97% to 98% identity and their genus names were assigned. The remaining 6 haplotypes showed less than 97% identity ~~and named as “unidentified” genera~~ (Figure 2). Two haplotypes were identified in the genus *Hemibarbus* from the Seomjin River (SJ1) and the

Nakdong River (ND2) showed 100% and 99% identity to the ~~Korean haplotype~~ *Hemibarbus labeo* (GenBank Number: DQ347953), and the Japanese ~~haplotype~~ *Hemibarbus maculatus* (LC146032) respectively. ~~Currently, among~~ the four known species of the genus *Hemibarbus* ~~in~~ Korea, *Hemibarbus labeo* and *Hemibarbus longirostris* are the most widely distributed species ~~in the Korean waters~~ (W.-O. Lee, Zhang, Oh, Baek, & Song, 2012). Two haplotypes ~~were~~ identified from the Seomjin River (SJ1 and SJ2) showed 97% identity with *Hemibarbus longirostris* (LC049889); another haplotype from the Taehwa River showed lower identity (95%) to ~~the~~ *Hemibarbus* sp., which suggests that these three haplotypes may be either *Hemibarbus longirostris* or *Hemibarbus mylodon* (Fig. 2). ~~The~~ *Hemibarbus mylodon* is an endangered freshwater ~~endemic species~~ to Korea (W. J. KIM, LEE, & BANG, 2007), so the further study should be made for confirmation.

Four ~~species~~ are reported from Korean ~~waters~~ ~~in the genus *Squalidus*; e.g.,~~ *Squalidus gracilis*, *Squalidus japonicus, Squalidus multimaculatus,* and *Squalidus chankaensis* (I.-S. Kim & Park, 2002). Five haplotypes were identified in the genus *Squalidus*~~, two~~ haplotypes from the Taehwa River (TH3) and the Hyeongsan River (HS1) showed 100% identity to the Korean ~~haplotype~~ *Squalidus japonicas coreanus* (GenBank Number: KR075134) and the Korean ~~haplotype~~ *Squalidus multimaculatus* (GenBank Number: KT948081). Another haplotype from the Hyeongsan River (HS3) showed 100% identity to the Japanese ~~haplotype~~ *Squalidus japonicas* (GenBank Number: LC277782). Two haplotypes from the Seomjin River showed 99% identity to the Korean ~~haplotype~~ *Squalidusc hankaensis tsuchigae* (GenBank Number: KT948082).

~~The~~ subfamily Acheilognathinae, commonly known as ~~the~~ bitterlings, deposits ~~their~~ eggs ~~into~~ the gill cavities of freshwater mussels (J.-i. Kitamura, 2007; J. Kitamura, Nagata, Nakajima, & Sota, 2012). About 60 fish species of bitterlings are currently found in ~~these three~~ genera

~~(~~*Acheilognathus*, *Tanakia*, and *Rhodeus*~~;~~ (Arai, 1988). ~~Eight species including~~ *Acheilognathus intermedia, Acheilognathus macropterus, Acheilognathus majusculus, Acheilognathus rhombeus, Rhodeus suigensis, Rhodeus uyekii, Tanakia somjinensis, Tanakia signifier* showed 99% to 100% sequence identity. Three haplotypes from the Seomjin River showed 99% sequence identity to the Korean haplotypes of *Acheilognathus intermedia* (EF483933), ~~*Acheilognathus*~~ *somjinensis* (FJ515921), and ~~*Acheilognathus*~~ *signifier* (EF483930)~~; even though the scientific name of *Acheilognathus somjinensis* and *Acheilognathus signifier* has changed to *Tanakia somjinensis* and *Tanakia signifier* respectively (I. Kim, 1997).~~

One haplotype from the Taehwa River (TH3) showed 100% identity to the Korean haplotype of ~~*Phoxinus*~~ *semotilus* (KT748874), ~~the scientific name of *Phoxinus semotilus* has been changed to *Rhynchocypris semotilus* (I. Kim, 1997; Yu, Kim, Kim, Yeo, & Kim, 2017), and currently this fish~~ categorized as critically endangered in the Red Data Book of endangered fishes in Korea (Ko, Kim, & Park, 2011).

Two species are currently known in the genus *Sarcocheilichthys* in Korea~~; the~~ *Sarcocheilichthys nigripinnis morii* and *Sarcocheilichthys variegatus wakiyae* (I.-S. Kim & Park, 2002). Two haplotypes from the Seomjin River (SJ2) and the Hyeongsan River (HS2~~)~~showed 100% and 97% sequence identity to the Korean ~~haplotype~~ *Sarcocheilichthys variegatus wakiyae* (GenBank Number: KU301744). One haplotype from the Hyeongsan River (HS2) showed 100% and 99.43% sequence identity to the Japanese ~~haplotype~~ *Sarcocheilichthys soldatovi* (LC146036) and the Korean ~~haplotype~~ *Sarcocheilichthys nigripinnis morii* (AP017653) respectively~~, but *S.*~~ *soldatovi* is not currently ~~identified in~~ Korean waters, ~~so~~ this haplotype may ~~be~~ *Sarcocheilichthys nigripinnis morii*. Another haplotype from the Nakdong River (ND1) showed 98% and 97%

sequence identity to *Sarcocheilichthys davidi* (NC037403) and *Sarcocheilichthys variegatus* (KU301744) respectively, so the further study should be needed for confirmation.

Gobiidae

We identified 16 haplotypes ~~in~~ the family Gobiidae, which ~~covers 11 species, 7 genera~~. ~~Questionable haplotypes (<99% identity) were 2 in this family (Fig. 3).~~ Five haplotypes were identified in genus *Tridentiger*, ~~the~~ five known species ~~under~~ the genus *Tridentiger* ~~in~~ Korea (I. Kim et al., 2005)~~, one~~ haplotype from the Taehwa River (TH4) showed 100% identity with the Korean ~~haplotype~~ *Tridentiger obscurus* (GenBank Number: KT601092). One haplotype from the Hyeongsan River (HS4) showed 100% identity to the Japanese ~~haplotype~~ *Tridentiger trigonocephalus* (GenBank Number: LC385175) and another haplotype from the Seomjin River (SJ3) showed 100% identity with the Korean ~~haplotype~~ *Tridentiger trigonocephalus* (GenBank Number: KM030481). ~~From~~ the phylogenetic ~~tree, we can say~~ the *Tridentiger trigonocephalus* haplotype from the Seomjin River is different from ~~the~~ Hyeongsan River (Fig. 3). All three haplotypes in the genus *Rhinogobius* showed 100% identity to the database. Two of each haplotype was assigned as the ~~respective~~ Korean (KM030471) and Japanese (LC049760) type of *Rhinogobius brunneus* with 100% identity, whereas the other one was100% identity to *Rhinogobius giurinus* (KM030475) covering all two endemic species in Korea. Two haplotypes of *Gymnogobius* sp. from the Taehwa River and the Hyeongsan River showed 98% sequence identity to *Gymnogobius taranetzi* (GenBank Number: LC385155) (Figure 3). Currently, ~~9~~ species ~~in~~ the genus *Gymnogobius* ~~is~~ reported in ~~the~~ Korean ~~waters~~ (I. Kim et al., 2005), and ~~its~~ MiFish (12S rRNA) sequences should be supplemented to the NCBI.

Cobitidae

Sixteen species in ~~5~~ genera ~~under~~ the family Cobitidae are currently reported in Korean rivers (I.-S. Kim, 2009). ~~We here identified~~ 18 haplotypes, which ~~cover~~ ~~7~~ genera in the family ~~Cobitidae~~ (Fig. 4). Two haplotypes in the genus *Cobitis* were identified ~~only from~~ the Seomjin River~~, both~~ were most closely related to the Japanese ~~haplotype~~ *Cobitis tetralineata* (LC146139) with 100% and 99% identity rather than showed 98% identity to the Korean haplotype (EU670794). ~~Moreover,~~ two haplotypes from the Taehwa River showed 98% and 97% identity to *Corbitis hankugensis* (LC146140). ~~Currently,~~ two ~~known~~ species ~~in the genus~~ *Misgurnus* ~~have been~~ reported from the Korean waters, ~~e.g.~~ *Misgurnus anguillicaudatus* and *Misgurnus mizolepis*. Interestingly, ~~two phylogenetically distant clades for *M. anguillicaudatus*~~ were identified by the phylogenetic analysis (Fig. 4)~~, one was clustered~~ with the Chinese ~~haplotype~~ *Misgurnus bipartitus* (KF562047), while the ~~others~~ grouped with the Korean ~~haplotype~~ *Misgurnus mizolepis* (AP017654). ~~The~~ *Misgurnus bipartitus* is ~~originally known to be~~ endemic ~~in~~ China, but both *M. bipartitus* and *M. anguillicaudatus* are currently used for the pond loach in Korea. Morphology of these three species is highly similar to one another, a reexamination of genetic structure for the loach species in Korea should be made.

Two haplotypes of *Paramisgurnus dabryanus* from the Hyeongsan River (HS1; KJ699181)~~,~~ and the Taehwa River (TH4; KM186182) showed 100% identity with the Chinese haplotype of *Paramisgurnus dabryanus*, which ~~is endemic~~ ~~species in~~ China (Fig. 4). ~~It also supports~~ ~~the~~ previous study ~~about the existence~~ of two populations of *P. dabryanus* suggested that the early introduction of this species originated from different locations (Shimizu & Takagi, 2010). *P. dabryanus* is often ~~imported Korea with *Misgurnus anguillicaudatus* for its morphological similarity.~~ Sequence identity between the two ~~haplotype sequences~~ was 100% ~~indicating~~ that

P. dabryanus has been imported from ~~various~~ geographical ~~locations,~~ Three haplotypes ~~in the family~~ Cobitidae showed a low sequence identity to the database (Fig. 4). One haplotype from the Taehwa River (TH1) showed 100% sequence identity to the Korean ~~haplotype~~ *Niwaella multifaciata* (EU670806), while another from the Hyeongsan River (HS1) showed lower (96%) identity to *Niwaella* sp., further study should be made for confirmation.

Other families

Besides the three major ~~families, we were also able to identified~~ 27 ~~haplotypes (19 species and one genus) belonging to 11 families (4~~ ~~in Bagridae, 3 in Mugilidae, 1 in Anguillidae, 2 in Centrarchidae, 1 in Channidae, 1 in Clupeidae, 2 in Odontobutidae, 1 in Pleuronectidae, 2 in Siluridae, and 2 in Sinipercidae, and 1 in Amblycipitidae),~~ ~~All~~ haplotypes in the family Bagridae ~~were~~ well ~~assigned~~ each corresponding ~~species names including~~ *Pseudobargrus ussuriensis*, *Pseudobargrus koreanus*, *Tachysurrus nitidus*, and *Tachysurus fulvidraco* (Fig. 5). Two species ~~in the genus~~ *Silurus* are currently known in the Korean waters, *Silurus microdorsalis*, and *Silurus asotus* (Park & Kim, 1994). One haplotype from the Taehwa River (TH1) showed 99% sequence identity with the Korean ~~haplotype~~ *Silurus microdorsalis* (GenBank Number: KT350610) ~~confirming the species~~, whereas another haplotype from the Seomjin River (SJ1) showed lower identity (96%) with ~~the~~ *Silurus microdorsalis* (KT350610), which requires further analysis (Fig. 5).

Five endemic species ~~in~~ the family Amblycipitidae are currently known in the Korean peninsula (I.-S. Kim & Park, 2002)~~, which include~~ *Liobagrus andersoni*, *Liobagrus obesus*, *Liobagrus mediadiposalis*, *Liobagrus somjinensis*, and *Liobagrus hyeongsanensis*; and their sequences should be supplemented. One ~~haplotype~~ from the Seomjin River ~~in the family~~

~~Amblycipitidae~~ showed 97% and 96% identity to the Chinese ~~haplotype~~ *Liobagrus styani* (KX096605) and the Korean ~~haplotype~~ *Liobagrus mediadiposalis* (KR075136), respectively. Further studies should be made to identify ~~the species~~. Three species ~~are~~ currently known in ~~the genus *Odontobutis* in~~ Korea~~,~~ *Odontobutis interrupta*, *Odontobutis platycephala*, and *Odontobutis obscura* (I. Kim et al., 2005). Two of them (*O. interrupta* and *O. platycephala*) were identified in this study ~~and *O.*~~ *obscura* ~~is known to inhabit exclusively in the~~ Geoje island. Two haplotypes were identified in genus *Coreoperca*, and showed 100% and 97% sequence identity to the Korean ~~haplotype~~ *Coreoperca herzi* (KR075132). Since ~~it is currently known that two endemic *Coreoperca* species in~~ the Korean peninsula (I. Kim et al., 2005), the second haplotype ~~appears to be~~ ~~the~~ *Coreoperca kawamebari*, but further study should be conducted for ~~confirmation.~~ Two invasive ~~fish~~ species ~~in~~ the family Centrarchidae ~~including~~ the Bluegill (*Lepomis macrochirus*) and the Largemouth bass (*Micropterus salmoides*) were also identified in this study. Those two species ~~in the Centrarchidae family~~ are endemic ~~in~~ North America, ~~which has been~~ introduced in the Korean peninsula for ~~the~~ aquaculture ~~industry~~ without considering the ~~effects~~ on the ecosystem. ~~Fish~~ species in four families including Anguillidae, Clupeidae, Mugilidae, Pleuronectidae were those inhabit both in brackish and coastal waters which include *Anguilla japonica, Konosirus punctatus, Mugil cepephalus, Kareius biocoloratus*.

Fish biodiversity in ~~four rivers~~

~~Based on the eDNA metabarcoding analysis by the MiFish pipeline, fish biodiversity in the four rivers was analyzed.~~ ~~The highest~~ Shannon Index (SI) was found in the Seomjin River (3.480) followed by the Taehwa River (3.067), the Hyeongsan River (2.954), and the Nakdong River (2.864). Among the 16 surveyed ~~station,~~ station 1 of the Seomjin River (SJ1) showed the highest diversity ([illegible]197), whereas the lowest (1.008) ~~was~~ in ~~the~~ station 4 of Nakdong River (ND4). From the upstream to downstream the average biodiversity has been changed from 1.951 to 1.415 (Table 4).

~~The~~ Cyprinidae ~~family~~ was ~~more~~ dominant ~~than other family members~~ ~~found~~ ~~in~~ the 4 ~~rivers~~, ~~the~~ highest number of ~~fish~~ ~~species~~ (52) was ~~found~~ in the Seomjin River, followed by the Taehwa River (42 species), the Hyeongsan River (40 species), and the Nakdong River (26 species) (Fig. 6). In the Seomjin River, ~~the~~ *Tanakia signifer* was ~~relatively~~ abundant (12.14%) than ~~the~~ *Mugil cephalus* (11.14%)~~, but~~ *Mugil cephalus* ~~was~~ abundant in the Taehwa River (16.36%) ~~than~~ ~~the~~ *Gymnogobius* ~~*sp.*~~ (14.91%). In the Hyeongsan River, an invasive fish species, ~~Bluegill~~ (*Lepomis macrochirus*)~~ was~~ dominant (17.20%)~~ than the~~ *Zacco platypus* (15.20%). In the Nakdong River, ~~the~~ *Tridentiger obscurus* was ~~more~~ dominant (23.20%) ~~than the~~ *Pseudogobio vaillanti* (13.35%).

Among ~~73 fish species~~ detected in this study, 13 ~~species~~ were ~~commonly identified~~ in ~~all~~ four ~~rivers~~. ~~Total~~ 17 species were exclusively ~~identified~~ in the Seomjin River~~, which included~~ *Squalidus gracilis, Cobitis tetralineata, Tanakia somjinensis, Acanthogobius hasta, Siniperca scherzeri, Pseudobagrus koreanus, Acheilognathus majusculus, Sarcocheilichthys variegatus, Coreoleuciscus splendidus, Tanakia signifier, Acheilognathus rhombeus, Microphysogobio yaluensis, Rhodeus suigensis, Kareius bicoloratus, Rhodeus uyekii, Phoxinus oxycephalus,* and *Acheilognathus intermedia* (Fig. 7).

Salinity and relationship

~~Salinity~~ ~~was~~ increased from ~~the~~ upstream to ~~the~~ downstream~~, the~~ lowest salinity (0.15 PSU) was measured at ~~the upstream (station 1)~~ of ~~the~~ Seomjin River, while the highest (20.20 PSU) was ~~found~~ at ~~the downstream (station 4)~~ of the Hyeongsan River. Fish species distribution with the salinity level also measured and found that freshwater fish species distributed at the upstream of rivers and brackish water fish species distributed at the downstream of rivers.

Clustering analysis

In this study, we have categorized the sampling stations of each river as upstream (station 1 and 2), midstream (station 3), and downstream (Station 4). Among the 73 identified species in this study ~~we have taken the top 30 species~~ and by using a statistical program (Primer v7), we have drawn a heat map, that clearly demonstrated species distribution in different sampling stations (Fig. 8). In the upstream, the dominant species are *Zacco platypus*, *Odontobutis interrupta, Odontobutis platycephala, Nipponocypris temminckii, Rhynchocypris lagowskii, Misgurnus mizolepis, Coreoperca herzi, Acheilognathus intermedia, and Tanakia signifier* (Fig. 8)~~; these~~ species are ~~absolutely~~ non-migratory freshwater species and most of them are endemic ~~in~~ Korean ~~peninsula~~ (I. Kim, 1997). In the midstream of rivers ~~we found the~~ dominant species are *Pseudorasbora parva, Gymnogobius breunigii, Rhinogobius giurinus, Rhinogobius brunneus, and Mugil cephalus,* ~~while~~ in the downstream ~~we found~~ *Tridentiger obscurus, Tridentiger trigonocephalus, Konosirus punctatus, Mugil cephalus, Anguilla japonica, Planiliza haematocheila,* ~~and these all fish are~~ euryhaline and anadromous ~~species~~ (https://www.fishbase.org)~~.~~

DISCUSSION

~~We here identified that~~ eDNA metabarcoding using the MiFish pipeline is a useful tool for the ~~fish~~ ~~biodiversity~~ ~~analysis obtaining~~ 125 unique haplotypes with at least 73 ~~fish species~~ ~~(13 families)~~ simply by one time 16 sample collections from four Korean rivers. According to the

"Survey and Evaluation of Aquatic Ecosystem Health (SEAEH)" in Korea, which costed several millions dollars by hundreds of researchers, ~~total~~ of 130 freshwater ~~fish species (28 families)~~ were identified from two times of 953 sites in Korean ~~waters~~ (Yoon, Jang, Kim, & Joo, 2012). It is quite impressive to know 56.15% freshwater fish species were identified by a single day environmental DNA analysis with only 16 water samples and its numbers will increase as the water sample numbers increase. These results strongly suggest that a freshwater fish biodiversity survey would be possible by the eDNA metabarcoding using the MiFish pipeline with 10% of total cost and labors of conventional morphological surveys in Korea. Another important benefit of eDNA metabarcoding is allowing ~~us to~~ survey or ~~monitor~~ more sites in a cheaper and faster way~~;~~ feasibility ~~study~~ of suitable environments for potential species is possible (Bista et al., 2017; Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016). This ~~approach also~~ ~~suitable~~ for surveying aquatic species inside in ~~a~~ protected ~~area~~, ~~conventional~~ sampling methods (i.e. netting, electrofishing or direct observation) should be ~~neglected as much as possible to avoid~~ ~~disturb~~ vulnerable communities e.g. mountain stream or reservoir (Fernandez et al., 2018). Moreover, this approach also disclosed the fish communities in localized ecosystems. It will open a new approach to unlock the interaction among fish communities and the local habitats.

Although ~~we identified that~~ metabarcoding analysis of eDNA using the MiFish pipeline ~~is~~ a useful tool to monitor the biodiversity of freshwater fish species, we also able to see several drawbacks ~~to~~ ~~overcome~~ in applying the platform for fish biodiversity analysis. First, MiFish sequence data for ~~endemic~~ ~~fish~~ ~~species in Korean~~ ~~water~~ should be supplemented. According to the Archive of Korean species (https://species.nibr.go.kr), 67 endemic ~~freshwater~~ ~~species~~ ~~are reported~~ and many of their MiFish sequences are still not uploaded in the database. Besides, historically freshwater fish living separately in their habitat (river, lake, reservoir, stream). ~~So~~

each country should have its database of genetic information (DNA sequences) for the accuracy of endemic species detection. From the result, we failed to assign the species name for 49 haplotypes with low sequence identity to the database (Table 2).

Secondly, MiFish primer amplifies the 12S rRNA gene (163-185 bp) region of mitochondrial DNA, which is different from the most widely used COI region, ~~and based on this 12S rRNA region,~~ it is difficult to differentiate some close related ~~fish~~ species e.g. *Sebastes* spp. and *Takifugu* spp. (Sato et al., 2018; Yamamoto et al., 2017). The ~~O~~verall average genetic distance of species ~~under~~ the family Cyprinidae was measured ~~by~~ using the Mega 7.0 program (Kumar, Stecher, & Tamura, 2016) ~~and found that the~~ lowest genetic distance (0.010) was found in the genus *Carassius* and the highest (0.164) was ~~in~~ the genus *Rhodeus* (Table S2). Due to having the lowest genetic distance in the MiFish region, it was not clearly distinguished between species under the genus *Carassius,* on the other hand, it was very clearly differentiated species under the genus *Rhodeus* (Fig. 2).

Thirdly, quantitative analytic methods should be introduced. Although we identified 125 unique haplotypes ~~to know the biodiversity of fish species~~, the quantitative analysis of each species is still far away from realization. The presence or absence of species by using real-time PCR (qPCR) has been undertaken successfully. Moreover, based on the concentration of environmental DNA in water, ~~researchers also tried to estimate~~ species abundance and biomass ~~in water~~ (Doi et al., 2017; Doi et al., 2015; Pilliod, Goldberg, Arkle, & Waits, 2013; Takahara, Minamoto, Yamanaka, Doi, & Kawabata, 2012; Thomsen et al., 2012; Wilcox et al., 2016; Yamamoto et al., 2016). These studies showed positive correlations between ~~the~~ eDNA concentration and the abundance and/or biomass of species. Even though having more advantages, the estimation of fish abundance/ biomass by using the eDNA approach in large rivers/streams, and their comparison with other survey methods have not been yet found (Doi et al., 2017). The spatial distribution of environmental DNA in a lotic environment is more difficult to detect species accurately in rivers or streams (Deiner et al., 2016; Jerde et al., 2016). Environmental DNA can be transported from 50 m. up to several kilometer downstream (Deiner & Altermatt, 2014; Jerde et al., 2016; Wilcox et al., 2016). ~~We also should consider the decaying time~~ of eDNAs in different temperatures, UV radiation, the action of bacteria, fungi and other ~~factors~~ (Shapiro, 2008). Several studies ~~exhibited~~ that in a controlled environment 300-400 bp DNA ~~fragments detected~~ in water up to one week (Alvarez, Yumet, Santiago, & Toranzos, 1996; Matsui, Honjo, & Kawabata, 2001; Zhu, 2006). Normally short fragments of DNA are degraded very slowly and ~~we can detect them~~ from ~~the~~ environmental samples (Deagle, Eveson, & Jarman, 2006). In dry and cold environments with the absence of light ~~the~~ short fragments ~~DNA could be found in well preserved~~ (Shapiro, 2008). Further studies ~~required~~ for better ~~understanding~~ and quantify eDNA transport dynamics in rivers/streams.

Fourthly, the standardized collection methods and pretreatment procedures for the NGS sequencing analysis should be established. One of the most strong points of biodiversity survey by the eDNA metabarcoding is the large data size due to the convenient sample collection and fast and reliable analyses compared with ~~the conventional survey. Those~~ high amounts of ~~data~~ can be further used for the various statistical analyses, which has been difficult from the low sample size. However, ~~those~~ large amounts of data ~~are now produced~~ by the different water collection methods, eDNA preparation, sequencing and bioinformatics platforms in the respective research group. ~~We here found the~~ MiFish pipeline would be ~~the~~ feasible bioinformatic platform for ~~the~~ eDNA metabarcoding analys~~is for~~ fish biodiversity as long as ~~several~~ drawbacks ~~improved.~~ If ~~the other~~ methods ~~including~~ water collection and eDNA preparation are established, automated fish biodiversity monitoring system would be possible soon, which would be one of the main hubs for the fish ecologists.

We also identified the overall highest fish biodiversity in the Seomjin River (3.48) compared with the other three rivers including Nakdong, Taehwa, and Hyeongsan river. The low biodiversity may have come from the anthropogenic effects. As ~~commonly shown in the other Korean rivers~~, those three rivers run through ~~the~~ highly populated metropolitan cities, in which rivers are ~~to be~~ exposed to ~~the~~ various human ~~activities~~ ~~directly~~ or indirectly ~~cause~~ changes in ~~the~~ diversity and distribution of freshwater ~~fish~~ (Finkenbine, Atwater, & Mavinic, 2000). In particular, a large weir installed in the Nakdong River aggravated the condition showing the lowest biodiversity (2.86) among the examined rivers. The various ~~construction~~ along the urbanized watershed including dams and weirs have caused the simplification and reduction of ~~the~~ habits ~~decreasing~~ the biodiversity in the river (Nilsson, Reidy, Dynesius, & Revenga, 2005; Riley et al., 2005). Different from those three rivers, there is no metropolitan city along ~~with~~ the Seomjin River ~~with a low chance to be~~ exposed to ~~the~~ anthropogenic ~~effects.~~ Since the freshwater area is a region where is easily disturbed and takes a long time to recover compared to other ecosystems (Ricciardi & Rasmussen, 1999), long-term monitoring should be continued for scientific conservation and management of the resources in the rivers.

The eDNA metabarcoding analysis also revealed ~~the widely distributed invasive fish species in the inland Korean waters~~. We were able to identify at least five invasive fish species ~~e.g.,~~ *Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus*, and *Micropterus salmoides* (Table S3). ~~Including these five species~~ *Carassius auratus, Misgurnus anguillicaudatus,* and *Paramisgurnus dabrynus* are ~~now~~ classified as invasive ~~species~~ in ~~different~~ countries (Copp, Garthwaite, & Gozlan, 2005; Hewitt, 2006; Mukai, Umemura, &

Takagi, 2011). The pond loach/Oriental weather fish, *Misgurnus anguillicaudatus* ~~which~~ is native to Siberia (Tugur and Amur drainages), Korea, Japan, China and Vietnam (Masuda, 1984; Talwar & Jhingaran) ~~also considered invasive species~~ (Freyhof & Korte, 2005; van Kessel, Dorenbosch, Crombaghs, Niemeijer, & Binnendijk, 2013), ~~because of their competition nature with~~ the native fish for food, reduce macroinvertebrates from the habitat, and decrease water quality by increasing ammonia, nitrate/nitrite, and turbidity levels, ~~having almost similar effect on the water quality as *Cyprinus carpio*~~. ~~The~~ ~~*M*~~. *anguillicaudatus* may impact on the native fishes for shelter and spawning sites~~, and~~ also prey~~ing~~ on eggs, larvae, and juveniles~~; and due to having the high~~ reproductive ~~ability~~ and high survivorship capacity makes it a potential invasive species (Keller & Lake, 2007; Koster *et al*, 2002; Nico & Fuller 2010). ~~Surprisingly~~ that the largemouth bass, *M. salmoides*, and bluegill *L. macrochirus* ~~inhabit~~ ~~throughout~~ ~~the~~ all ~~examined~~ four rivers. As native to eastern North America, those two species were artificially introduced in the 1970's, as freshwater fish stock without any further consideration of the effects on the freshwater ecosystem in Korea. The species has spread throughout the peninsula and now became a serious problem damaging crops, competing with the native species, and serving as disease vectors.

The Common carp (*Cyprinus carpio*) and *Cyprinus megalophthalmus* were found in the Hewangsan, Nakdong and Taehwa ~~River~~. The Japanese crucian carp, *Carassius cuvieri* was found from the Hewangsan, the Seomjin, and the Taehwa River. From July 1999 to January 2000, a nationwide survey was done in 28 sites of 9 river systems in Korea by Jang et al. (2002)~~, and they~~ reported 62 ~~fish species~~ ~~of~~ 16 ~~families~~, which ~~was~~ 32% of ~~known~~ Korean freshwater ichthyofauna (194)~~, they~~ also identified ~~5~~ exotic ~~fish~~ species~~ e.g.~~ *Carassius cuvieri, Cyprinus carpio, Micropterus salmoides, Lepomis macrochirus,* and *Oreochromis niloticus* (Min-Ho Jang et al., 2002). Korean National Long-term Ecological Research (KNLTER) and Evaluation of Aquatic Ecosystem Health (SEAEH) reports ~~reveal~~ that some exotic species are widely distributed in the Korean streams. ~~The~~ *Micropterus salmoides, Lepomis macrochirus,* and *Carassius cuvieri* were found in all river systems (Yoon et al., 2012). These introduced ~~fish has a~~ fast-growing and ~~disease resistance characteristics.~~ The bluegill (*Lepomis macrochirus)* and the Largemouth bass (*Micropterus salmoides*) are ~~indicated~~ as harmful exotic fish species, ~~and~~ invasively spreading throughout the Korean river systems (Yoon et al., 2012). Using traditional methods, detection rates can be low, required huge labor, time, and sometimes it is impossible to detect the alien or invasive species until the density or abundance reaches a certain threshold. The eDNA metabarcoding approach ~~would be of enormous importance~~ because of its ability to identify target species at a very low ~~concentration~~ (Rees, Maddison, Middleditch, Patmore, & Gough, 2014). Since freshwater ecosystems are much more vulnerable to invasive species causing biodiversity loss and global change (Clavero & García-Berthou, 2005). The eDNA metabarcoding analysis would be useful to monitor the distribution patterns of the invasive species and its surveys are urgently required throughout all Korean freshwaters.

~~In this study, we identified~~ 35 ~~fish species~~ from Cyprinidae (47.94%), 11 from Gobiidae (15.07%), 8 from Cobitidae (10.96%), and ~~the remaining~~ 19 ~~fish species~~ from ~~the~~ other families (26.03%)~~. Here we identified that the eDNA metabarcoding by using the MiFish pipeline is a useful tool for the fish biodiversity analysis in Korean rivers.~~ From the environmental DNA metabarcoding, we identified 52 fish species from the Seomjin River, and the lowest fish species identified from the Nakdong River (26 species), 42 and 40 fish species from the Taehwa River and the Hyeongsan River respectively. Yoon et al. (2012) reported 72 fish species from the Seomjin River, and Lee et al. (2015) reported 18 fish species from the Nakdong River by using traditional survey methods (J. W. Lee et al., 2015; Yoon et al., 2012). In 2009, a total of 124 freshwater fish species, belonging to 27 families were found in the four major river systems (Han River, Nakdong River, Geum River, Yeongsan/ Seomjin River) of the Korean Peninsula. Among them, the most abundant (85.7%) family was Cyprinidae (54 fish species), 15 and 12 species were belonging to Cobitidae and Gobiidae family respectively (Yoon et al., 2012). In 2006, 83 native fish species and 5 exotic fish species were recorded from 31 sampling stations of 12 different river basins of Korea (M-H Jang, Joo, & Lucas, 2006), whereas we identified 68 native and 5 exotic fish species from the 16 sampling stations of four rivers by simply one time water sample collection.

CONCLUSION

The eDNA metabarcoding approach, ~~combination~~ with NGS to identify multiple species is a potential technique to monitor species diversity in aquatic habitat and ~~will~~ offer a more precise estimation of biodiversity rather than single or a handful of species surveillance. The present study revealed that MiFish metabarcoding ~~can~~ uncover the freshwater fish biodiversity in four Korean rivers. ~~It is~~ an example of the potential of environmental DNA metabarcoding for the investigation, monitoring, and distribution of the native and invasive fish species in the running waters for the first time in Korea. ~~This result~~ may be ~~able to utilize~~ for various purposes, e.g. ~~take some~~ effective ~~steps~~ ~~to enhance efforts~~ from the government and improve public conversation for the better management of ~~the~~ freshwater resources. Our findings suggest that environmental DNA metabarcoding required less time and taxonomic expertise, and itis better to understand the fish distribution and biodiversity in rivers/streams than the traditional survey systems.

Acknowledgments

The authors are thankful to the Ministry of Oceans and Fisheries, Republic of Korea.

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.

Role of funding

The funding sources had no role in the research design, sample collection, data analysis, manuscript writing, or the decision to submit the article for publication.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Md. Jobaidul Alam collected the samples, performed the experiments, analyzed the data, prepared figures and/or tables, wrote the manuscript
- Nack-Keun Kim collected the samples, analyzed the data
- Sapto Andriyono performed the experiments, analyzed the data, prepared figures and/or tables
- Hee-Kyu Choi analyzed the data, prepared figures and/or tables
- Ji-Hyun Lee analyzed the data, prepared figures and/or tables
- Hyun-Woo Kim conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, authored or reviewed drafts of the manuscript, approved the final draft, wrote the manuscript.

REFERENCES

Alvarez, A. J., Yumet, G. M., Santiago, C. L., & Toranzos, G. A. (1996). Stability of manipulated plasmid DNA in aquatic environments. *Environmental Toxicology and Water Quality: An International Journal, 11*(2), 129-135.

Arai, R. (1988). Acheilognathus melanogaster, a senior synonym of A moriokae, with a revision of the genera of the subfamily Acheilogathinae (Cypriniformes, Cyprinidae). *Bul Nat Sci Mus Tokyo Ser A, 14*, 199-213.

Ardura, A., Zaiko, A., Martinez, J. L., 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.

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., . . . Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature communications, 8*, 14087.

Cai, W., Ma, Z., Yang, C., Wang, L., Wang, W., Zhao, G., . . . Douglas, W. Y. (2017). Using eDNA to detect the distribution and density of invasive crayfish in the Honghe-Hani rice terrace World Heritage site. *PloS one, 12*(5), e0177724.

Clavero, M., & García-Berthou, E. (2005). Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution, 20*(3), 110. doi: 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*(11), e0188126.

Copp, G., Garthwaite, R., & Gozlan, R. (2005). Risk identification and assessment of non-native freshwater fishes: a summary of concepts and perspectives on protocols for the UK. *Journal of Applied Ichthyology, 21*(4), 371-373.

Deagle, B. E., Eveson, J. P., & Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces. *Frontiers in zoology, 3*(1), 11.

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PloS one, 9*(2), e88786.

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature communications, 7*, 12544.

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*(4), 953-959.

Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., & Minamoto, T. (2017). Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshwater Biology, 62*(1), 30-39.

Doi, H., Uchii, K., Takahara, T., Matsuhashi, S., Yamanaka, H., & Minamoto, T. (2015). Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PloS one, 10*(3), e0122763.

Fernandez, S., Sandin, M. M., Beaulieu, P. G., Clusa, L., Martinez, J. L., Ardura, A., & García-Vázquez, E. (2018). Environmental DNA for freshwater fish monitoring: insights for conservation within a protected area. *PeerJ, 6*, e4486.

Finkenbine, J. K., Atwater, J., & Mavinic, D. (2000). STREAM HEALTH AFTER URBANIZATION 1. *JAWRA Journal of the American Water Resources Association, 36*(5), 1149-1160.

Freyhof, J., & Korte, E. (2005). The first record of Misgurnus anguillicaudatus in Germany. *Journal of Fish Biology, 66*(2), 568-571.

Gustavson, M., Collins, P., Finarelli, J., Egan, D., Conchúir, R., Wightman, G., . . . Carlsson, J. (2015). An eDNA assay for Irish Petromyzon marinus and Salmo trutta and field validation in running water. *Journal of Fish Biology, 87*(5), 1254-1262.

Hewitt, C. L. (2006). *Alien species in aquaculture: considerations for responsible use*: IUCN.

Jang, M. H., Joo, G. J., & Lucas, M. (2006). Diet of introduced largemouth bass in Korean rivers and potential interactions with native fishes. *Ecology of Freshwater Fish, 15*(3), 315-320.

Jang, M. H., Kim, J. G., Park, S. B., Jeong, K. S., Cho, G. I., & Joo, G. J. (2002). The current status of the distribution of introduced fish in large river systems of South Korea. *International Review of Hydrobiology, 87*(2-3), 319-328.

Jerde, C. L., Olds, B. P., Shogren, A. J., Andruszkiewicz, E. A., Mahon, A. R., Bolster, D., & Tank, J. L. (2016). Influence of stream bottom substrate on retention and transport of vertebrate environmental DNA. *Environmental science & technology, 50*(16), 8770-8779.

Kim, I.-S. (2009). A Review of the Spined Loaches, Family Cobitidae (Cypriniformes) in Korea [A Review of the Spined Loaches, Family Cobitidae(Cypriniformes) in Korea]. *Korean Journal of Ichthyology, 21*(s), 7-28.

Kim, I.-S., & Park, J.-Y. (2002). *Freshwater fishes of Korea*: Kyo hak sa.

Kim, I. (1997). Illustrated encyclopedia of fauna and flora of Korea, Vol. 37. *Freshwater Fishes. Min Edu Korea, Seoul (in Korean)*.

Kim, I., Choi, Y., Lee, C., Lee, Y., Kim, B., & Kim, J. (2005). Illustrated book of Korean fishes. *Kyo-Hak Publishing*, 417-418.

KIM, W. J., LEE, Y. A., & BANG, I. C. (2007). Isolation and characterization of polymorphic microsatellite markers for the endangered Korean freshwater fish Hemibarbus mylodon. *Molecular ecology notes, 7*(3), 516-518.

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*(1), 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*(3), 566-573.

Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PloS one, 12*(5), e0177643.

Ko, M., Kim, K., & Park, J. (2011). Red Data Book of endangered fishes in Korea. *National Institute of Biological Resources, Incheon.(in Korean)*.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol, 33*(7), 1870-1874.

Lee, J. W., Yoon, J. D., Kim, J. H., Park, S. H., Baek, S. H., Yu, J. J., . . . Min, J. I. (2015). Length-weight relationships for 18 freshwater fish species from the Nakdong River in South Korea. *Journal of Applied Ichthyology, 31*(3), 576-577.

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*(4), 353-359.

Masuda, H. (1984). *The fishes of the Japanese Archipelago* (Vol. 2): Tokai University Press.

Matsui, K., Honjo, M., & Kawabata, Z. i. (2001). Estimation of the fate of dissolved DNA in thermally stratified lake water from the stability of exogenous plasmid DNA. *Aquatic Microbial Ecology, 26*(1), 95-102.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., . . . Araki, H. (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*(7), 150088.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., . . . 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*(7). doi: 10.1098/rsos.150088

Mukai, T., Umemura, K., & Takagi, M. (2011). First record of Paramisgurnus dabryanus accompanied with the invasion of Chinese lineage of Misgurnus anguillicaudatus in Gifu Prefecture, Japan. *Bulletin of the Biogeographical Society of Japan, 66*, 85-92.

Nilsson, C., Reidy, C. A., Dynesius, M., & Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. *science, 308*(5720), 405-408.

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*(2), 105-112.

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences, 70*(8), 1123-1130.

Pilliod, D. S., Goldberg, C. S., Laramie, M. B., & Waits, L. P. (2013). *Application of environmental DNA for inventory and monitoring of aquatic species*: US Department of the Interior, US Geological Survey.

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., . . . Dejean, T. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific reports, 8*(1), 10361.

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of applied Ecology, 51*(5), 1450-1459.

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*(21), e145-e145. doi: 10.1093/nar/gkr732

Ricciardi, A., & Rasmussen, J. B. (1999). Extinction rates of North American freshwater fauna. *Conservation Biology, 13*(5), 1220-1222.

Riley, S. P., Busteed, G. T., Kats, L. B., Vandergon, T. L., Lee, L. F., Dagit, R. G., . . . Sauvajot, R. M. (2005). Effects of urbanization on the distribution and abundance of amphibians and invasive species in southern California streams. *Conservation Biology, 19*(6), 1894-1907.

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*(6), 1553-1555. doi: 10.1093/molbev/msy074

Seehausen, O., & Wagner, C. E. (2014). Speciation in Freshwater Fishes. *Annual Review of Ecology, Evolution, and Systematics, 45*(1), 621-651. doi: 10.1146/annurev-ecolsys-120213-091818

Shapiro, B. (2008). Engineered polymerases amplify the potential of ancient DNA. *Trends in biotechnology, 26*(6), 285-287.

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation, 197*, 131-138. doi: https://doi.org/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. *Jpn J Ichthyol, 57*, 125-134.

Stoeckle, M. Y., Soboleva, L., & Charlop-Powers, Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PloS one, 12*(4), 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*(2), e56584.

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. i. (2012). Estimation of fish biomass using environmental DNA. *PloS one, 7*(4), e35868.

Talwar, P., & Jhingaran, A. Inland fishes of India and adjacent countries. 1991. In 2 vols: Oxford&IBH Publishing Co., New Delhi, Bombay, Calcutta. Inland fishes of ….

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., . . . Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular ecology, 21*(11), 2565-2573.

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., . . . Kondoh, M. (2017). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *bioRxiv*. doi: 10.1101/113472

van Kessel, N., Dorenbosch, M., Crombaghs, B., Niemeijer, B., & Binnendijk, E. (2013). First record of Asian weather loach Misgurnus anguillicaudatus (Cantor, 1842) in the River Meuse basin. *BioInvasions Records, 2*(2), 167-171.

Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., . . . Schwartz, M. K. (2016). Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char Salvelinus fontinalis. *Biological Conservation, 194*, 209-216.

Williams, K. E., Huyvaert, K. P., Vercauteren, K. C., Davis, A. J., & Piaggio, A. J. (2018). Detection and persistence of environmental DNA from an invasive, terrestrial mammal. *Ecology and evolution, 8*(1), 688-695.

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., . . . Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific reports, 7*, 40368.

Yamamoto, S., Minami, K., Fukaya, K., Takahashi, K., Sawada, H., Murakami, H., . . . 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*(3), e0149786.

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.

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* (pp. 175-191): Springer.

Yu, J.-N., Kim, B.-J., Kim, C., Yeo, J.-H., & Kim, S. (2017). The complete mitochondrial genome of the black star fat minnow (Rhynchocypris semotilus), an endemic and endangered fish of Korea. *Mitochondrial DNA Part A, 28*(1), 114-115.

Zhu, B. (2006). Degradation of plasmid and plant DNA in water microcosms monitored by natural transformation and real-time polymerase chain reaction (PCR). *Water research, 40*(17), 3231-3238.

Figure 1

Environmental DNA water sample collection sites of the four rivers

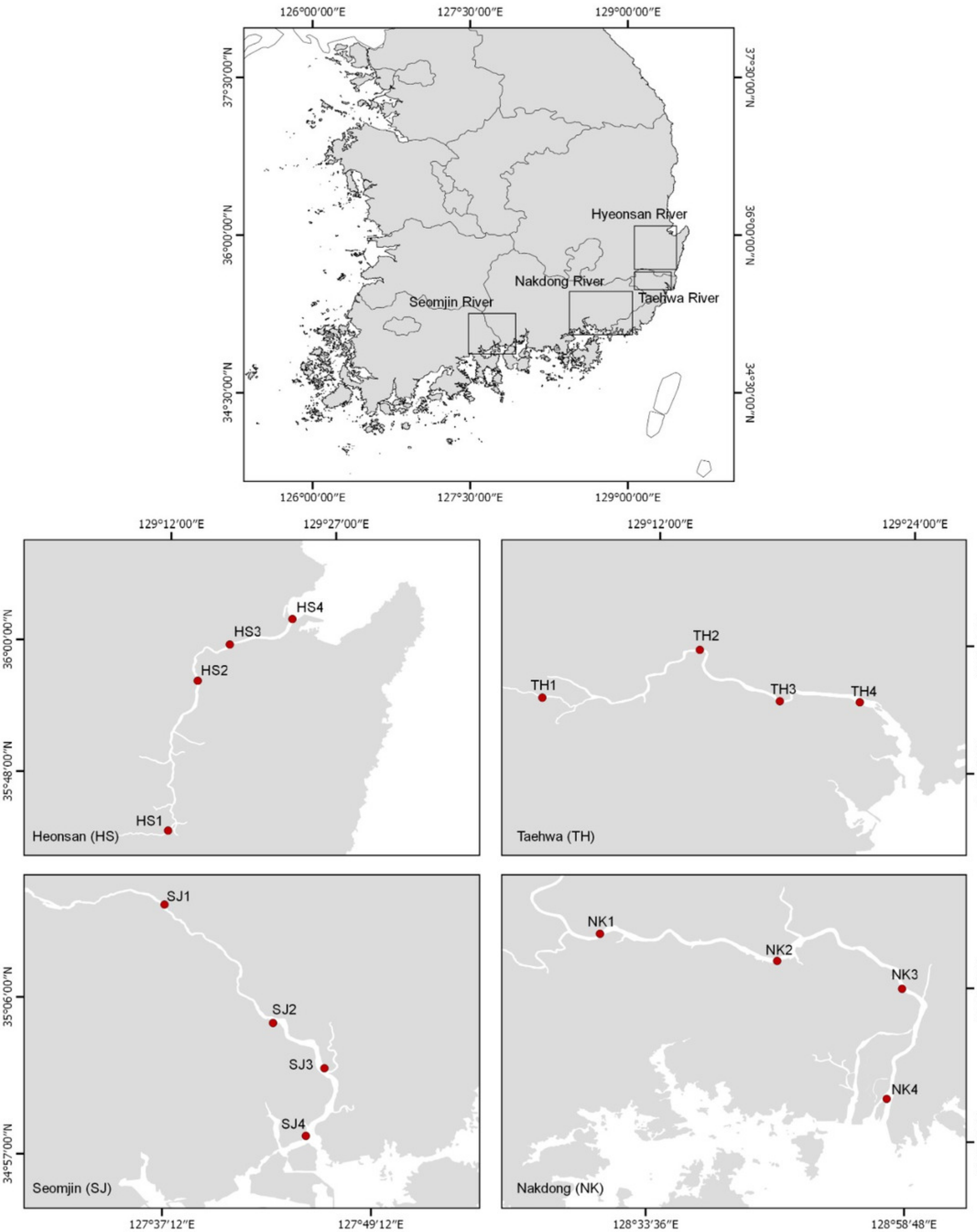
126°00′00″E
127°30′00″E
129°00′00″E
37°30′00″N
36°00′00″N
34°30′00″N
Hyeonsan River
Nakdong River
Taehwa River
Seomjin River
129°12′00″E
129°27′00″E
36°00′00″N
35°48′00″N
HS1
HS2
HS3
HS4
Heonsan (HS)
129°12′00″E
129°24′00″E
35°35′17″N
35°29′17″N
TH1
TH2
TH3
TH4
Taehwa (TH)
SJ1
SJ2
SJ3
SJ4
35°06′00″N
34°57′00″N
Seomjin (SJ)
127°37′12″E
127°49′12″E
NK1
NK2
NK3
NK4
35°18′00″N
35°03′00″N
Nakdong (NK)
128°33′36″E
128°58′48″E

Figure 2

Phylogenetic tree of the fish species under the family Cyprinidae

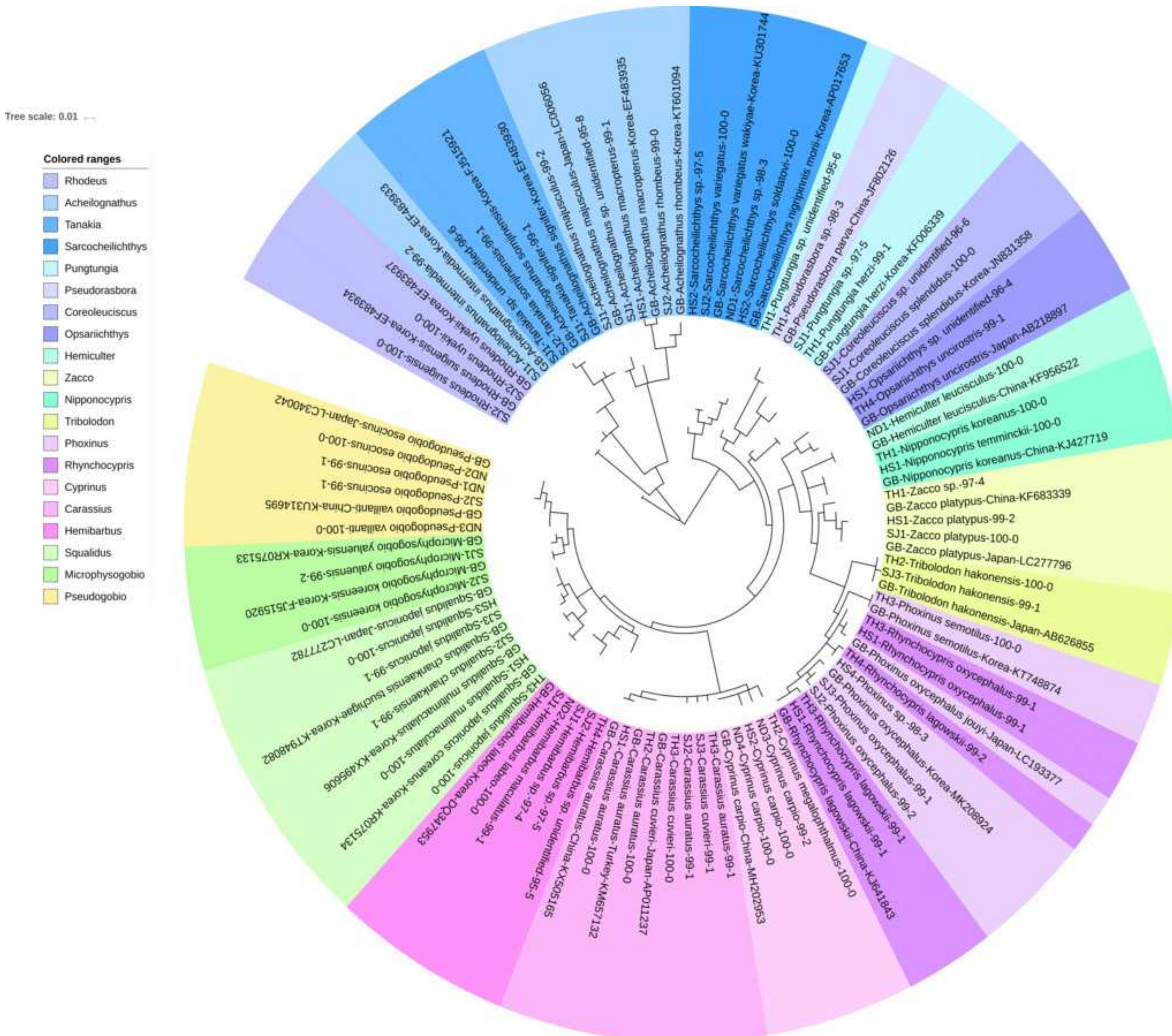

Figure 3

Phylogenetic tree of the fish species under the family Gobiidae

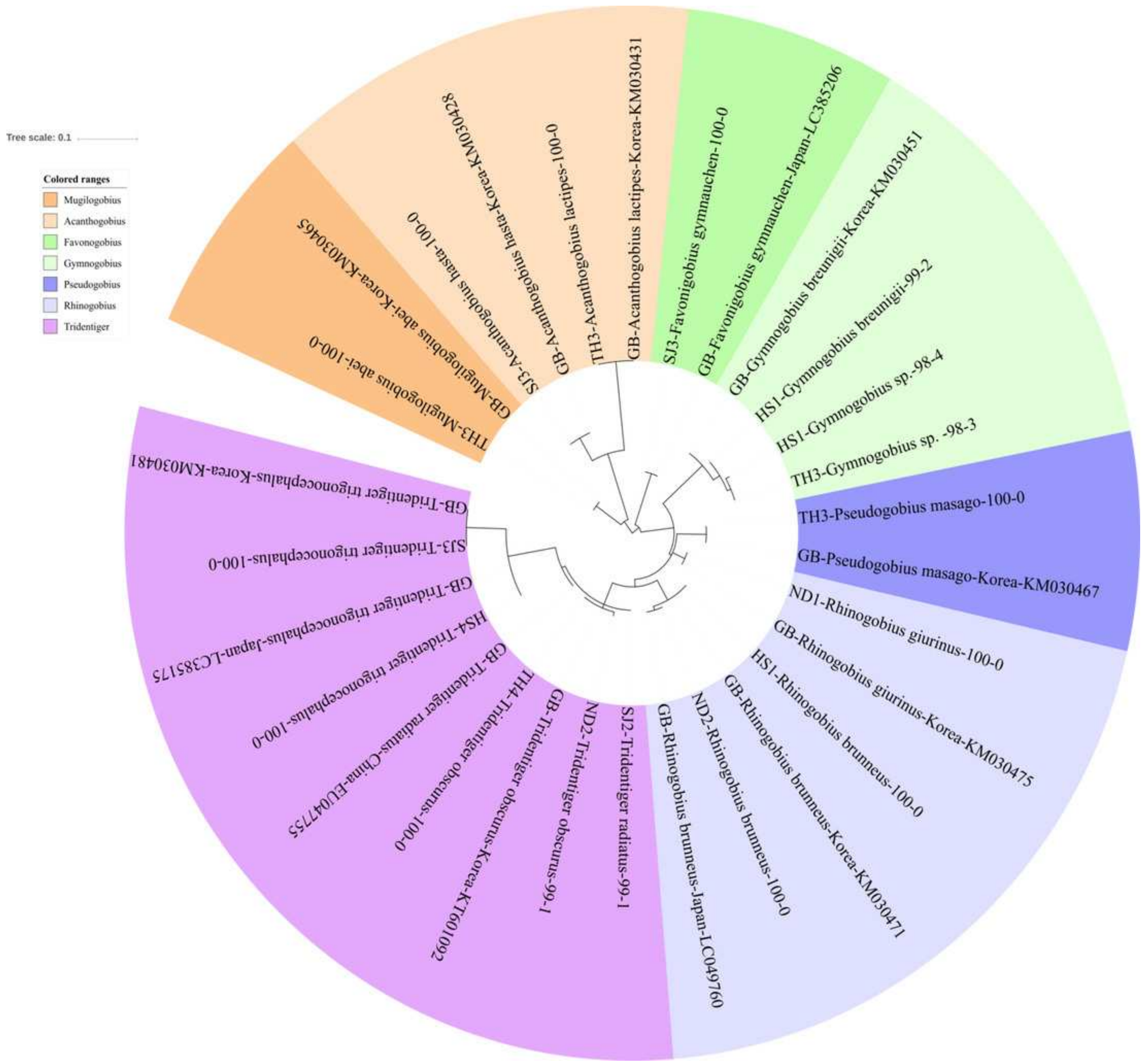

Figure 4

Phylogenetic tree of the fish species under the family Cobitidae

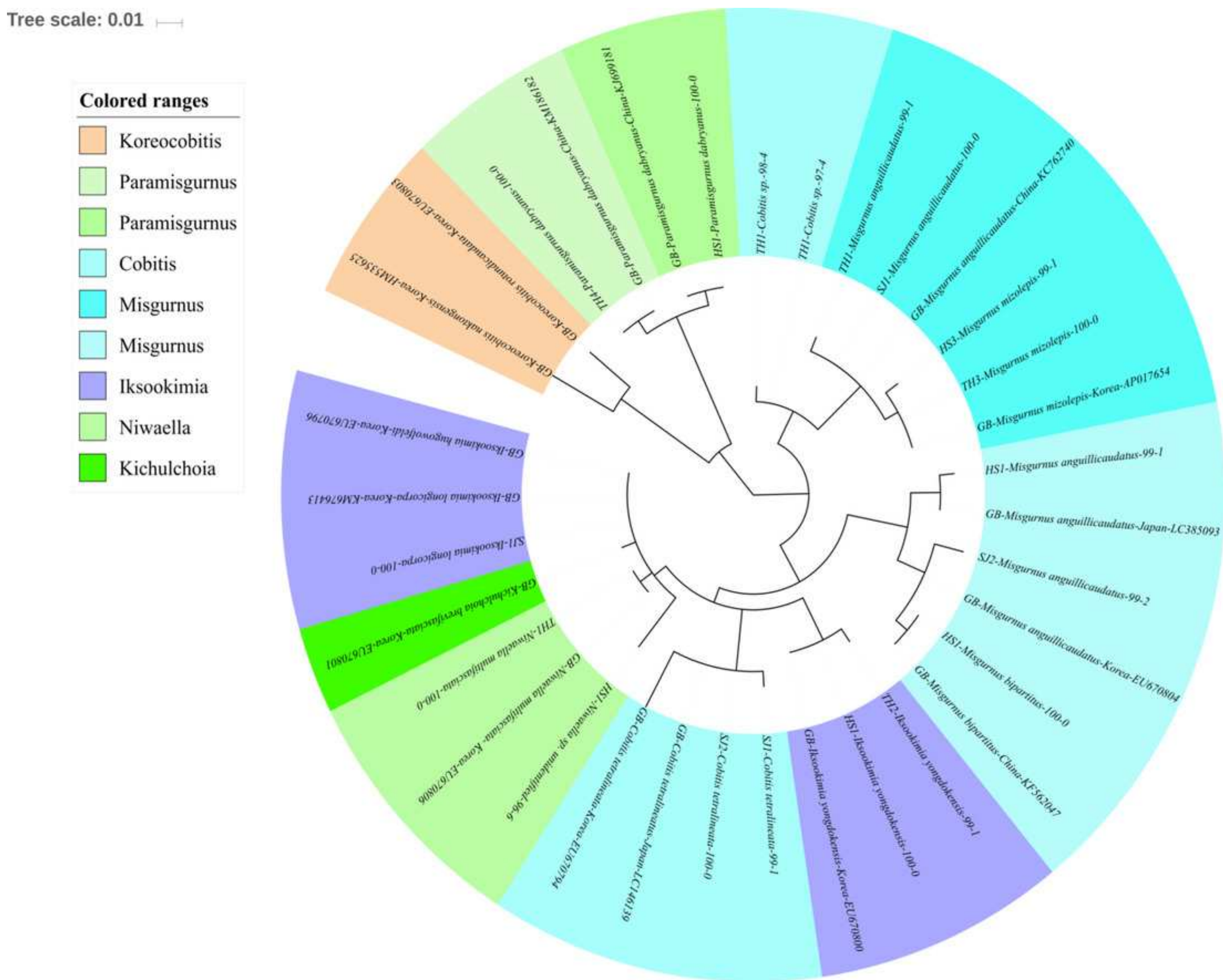

Figure 5

Phylogenetic tree of the fish species under the other families

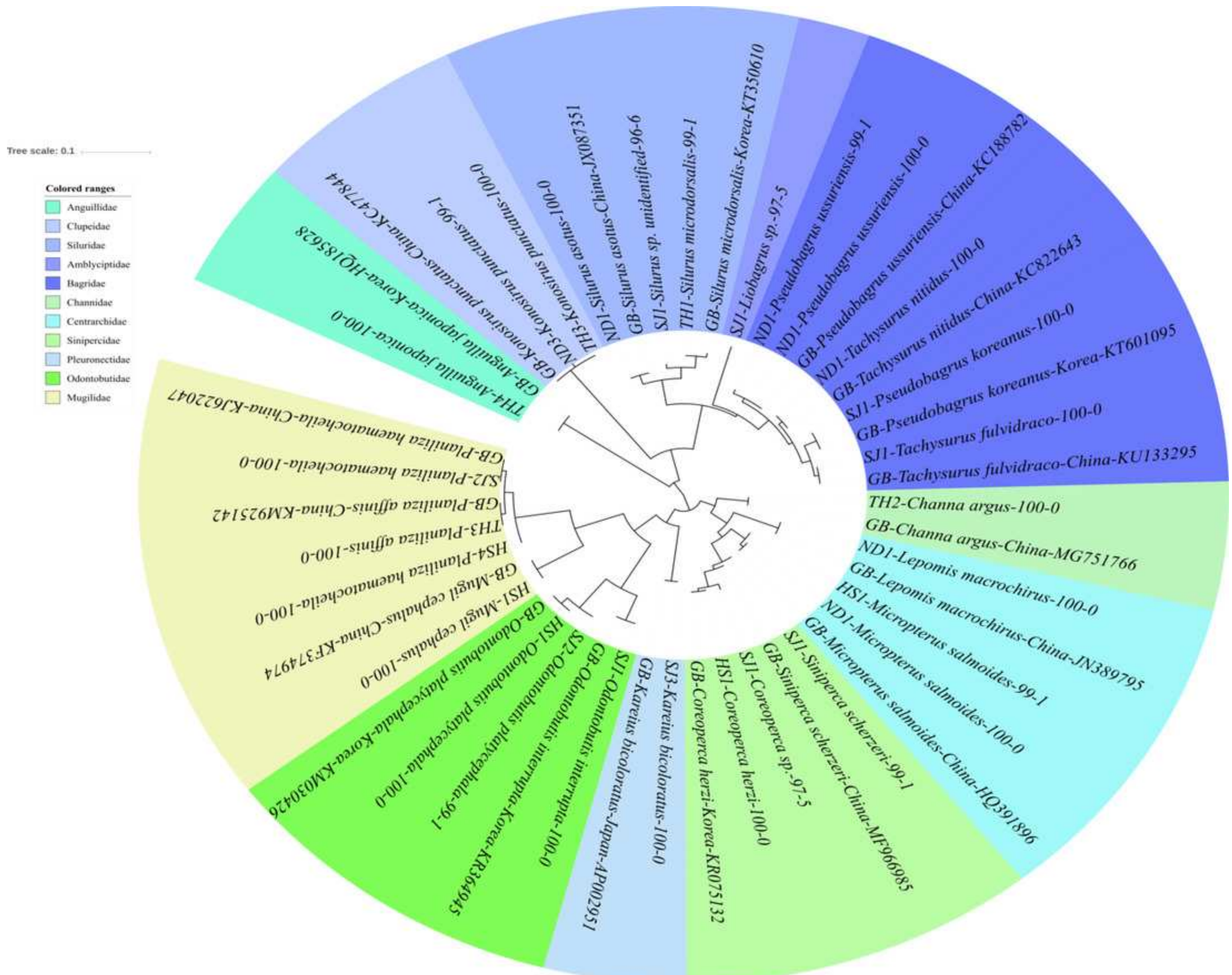

Figure 6

Proportion of the fish species under different families

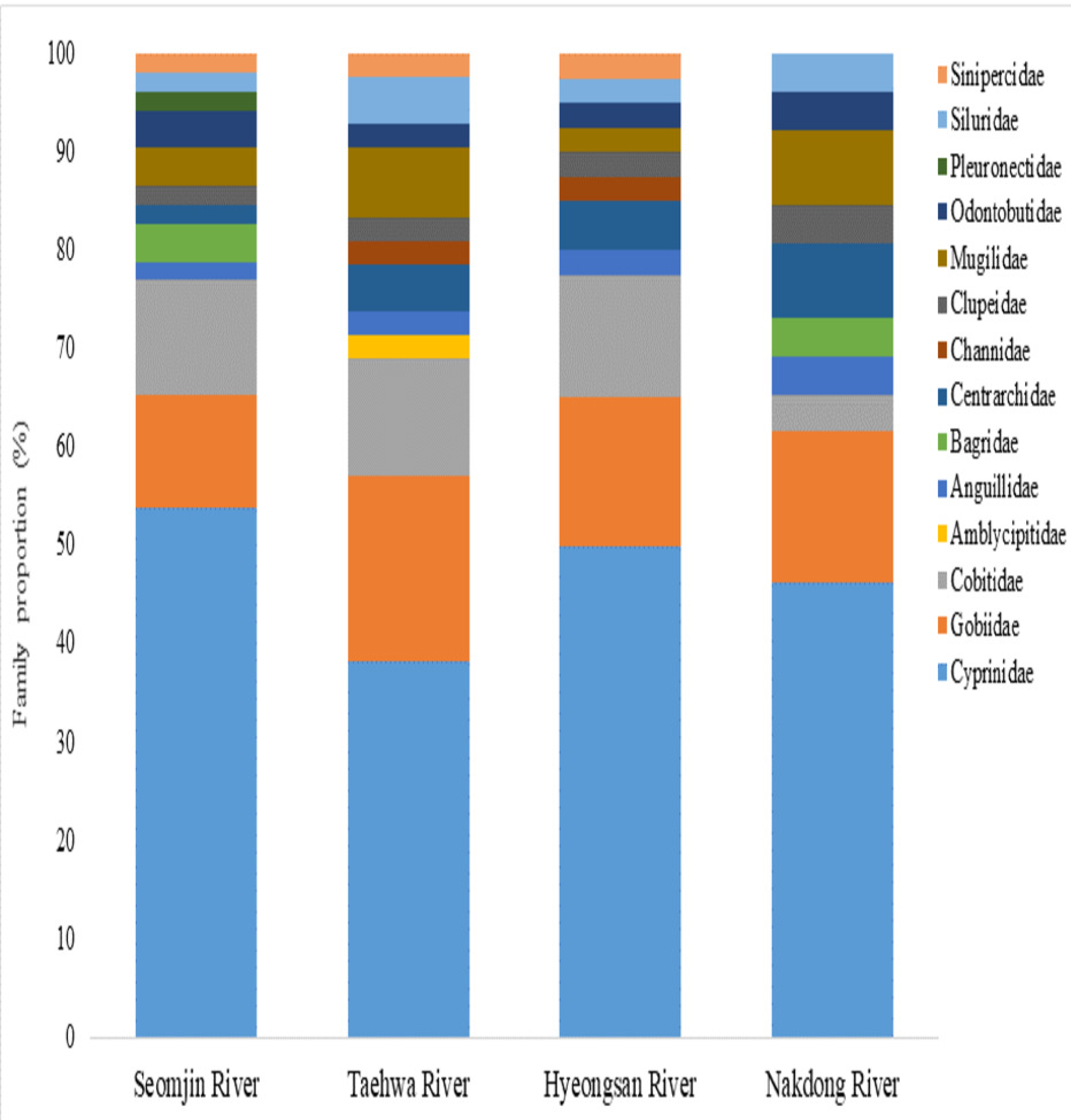

Figure 7

Venn diagram of fish species identified in the four Korean rivers

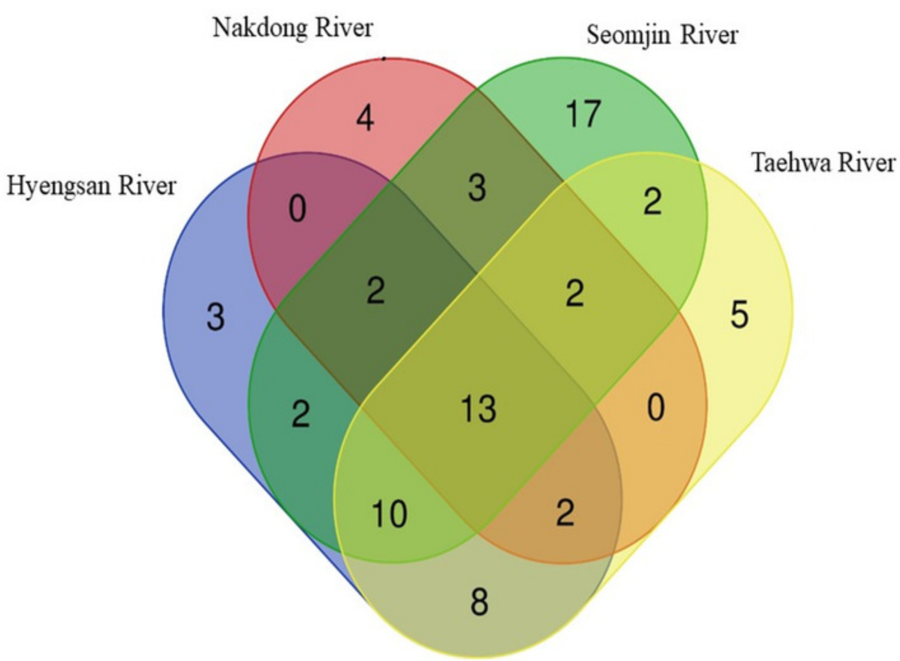

Figure 8

Heat map of top 30 fish species identified in 16 sampling stations of the four rivers

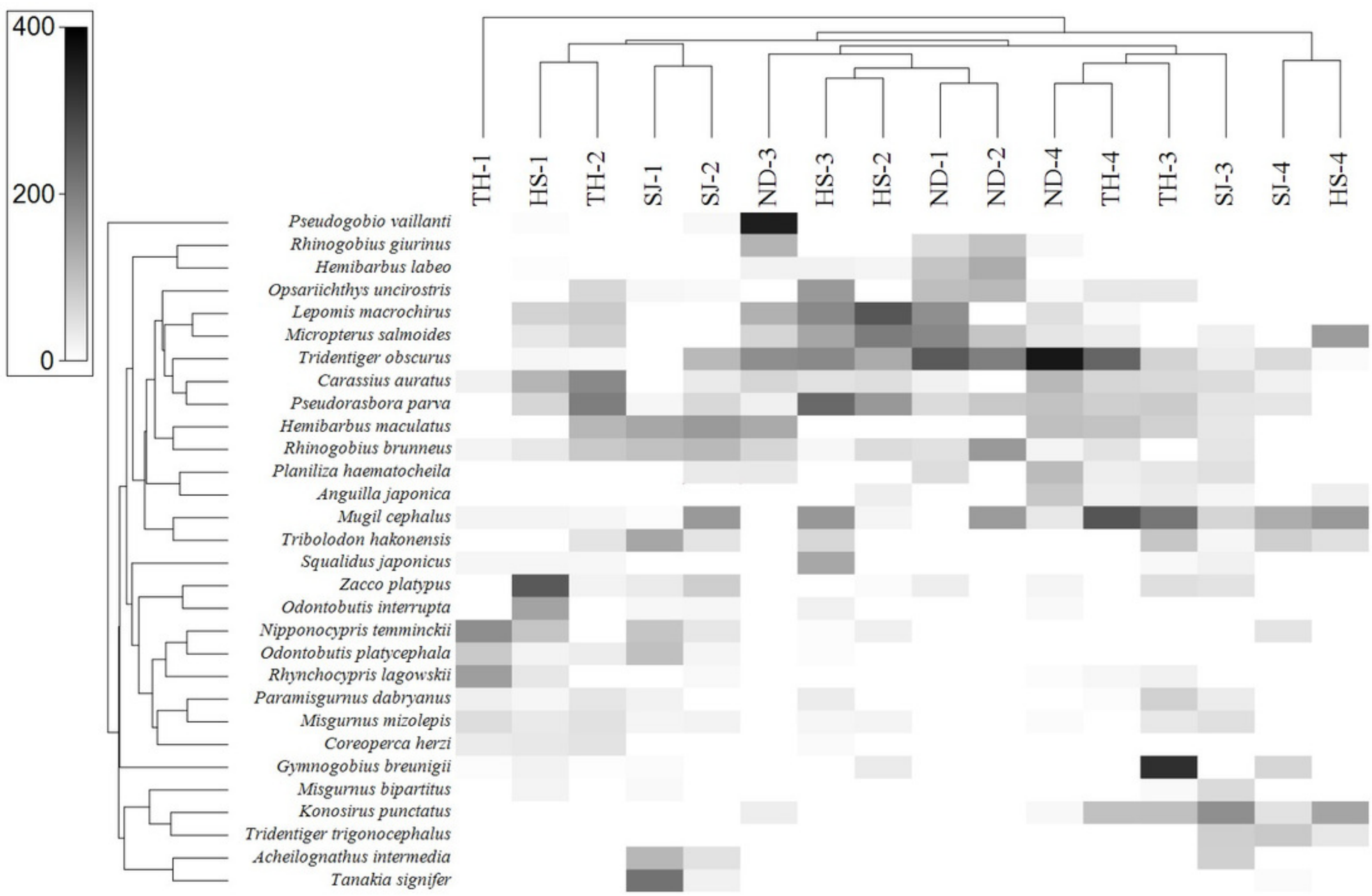

Table 1(on next page)

Environmental DNA sample collection sites with the physico-chemical parameters of the four rivers, South Korea

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

River	Date	Station	GPS location	Temp. (0c)	Salinity (PSU)
Hyeongsan	2018.06.11	HS1	N 35° 42' 36", E 129° 11' 42"	18.6	1.00
		HS2	N 35° 56' 14", E 129° 14' 24"	19.5	2.02
		HS3	N 35° 59' 32", E 129° 17' 19"	20.0	3.20
		HS4	N 36o 01' 51", E 129° 23' 01"	24.0	20.20
Taehwa	2018.06.11	TH1	N 35° 32' 52", E 129° 06' 27"	19.4	1.02
		TH2	N 35° 35' 07", E 129° 13' 52"	19.8	2.04
		TH3	N 35° 32' 42", E 129° 17' 38"	22.7	14.02
		TH4	N 35° 32' 39", E 129° 21' 24"	19.2	17.80
Seomjin	2018.06.12	SJ1	N 35° 11' 18", E 127° 37' 21"	24.2	0.15
		SJ2	N 35° 04' 30", E 127° 43' 35"	23.4	2.01
		SJ3	N 35° 01' 54", E 127° 46' 32"	23.0	12.9
		SJ4	N 34° 58' 01", E 127° 45' 28"	23.25	16.8
Nakdong	2018.06.12	ND1	N 35° 23' 19", E 128° 29' 09"	24.0	1.92
		ND2	N 35° 20' 40", E 128° 46' 26"	24.1	2.40
		ND3	N 35° 17' 57", E 128° 58' 37"	23.2	2.78
		ND4	N 35° 07' 13", E 128° 57' 07"	22.5	4.50

Table 2(on next page)

Summary of taxonomic assignment of the MiSeq reads from 4 rivers eDNA samples

Table 2. Summary of taxonomic assignment of the MiSeq reads from 4 rivers eDNA samples

	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)
Haplotypes with species name	61	49	48	31	189 (105)
Total species	52	42	40	26	160 (73)

* Shared species numbers in brackets

Table 3(on next page)

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

1 **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	*Acanthogobius hasta*	100	KM030428	KM891736	-	
2	Gobiidae	TH3	*Acanthogobius lactipes*	100	KM030431	-	LC385140	
3	Cyprinidae	SJ1	*Acheilognathus intermedia*	99	EF483933	-	-	
4	Cyprinidae	HS1	*Acheilognathus macropterus*	99	EF483935	KJ499466	LC092100	
5	Cyprinidae	SJ1	*Acheilognathus majusculus*	99	-	-	LC006056	
6	Cyprinidae	SJ2	*Acheilognathus rhombeus*	99	KT601094	-	LC146100	
7	Cyprinidae	SJ1	*Acheilognathus* sp. (unidentified)	95			LC006056	
8	Anguillidae	TH4	*Anguilla japonica*	100	HQ185628	MH050933	LC193417	
9	Cyprinidae	HS1	*Carassius auratus*	100	-	KX505165		
10	Cyprinidae	TH2	*Carassius auratus*	100				Turkey KM657132
11	Cyprinidae	TH3	*Carassius auratus*	99		AY771781	LC193299	
12	Cyprinidae	SJ2	*Carassius auratus*	99	-	AY771781	LC193299	
13	Cyprinidae	TH3	*Carassius cuvieri*	100	-	-	AP011237	
14	Cyprinidae	SJ3	*Carassius cuvieri*	100			AP011237	
15	Channidae	TH1	*Channa argus*	100	-	MG751766	AB972107	
16	Cobitidae	TH1	*Cobitis* sp.	97	EU670794	-	LC146139	
17	Cobitidae	TH1	*Cobitis* sp.	97	EU670794	-	LC146139	
18	Cobitidae	SJ2	*Cobitis tetralineata*	100	EU670794	-	LC146139	
19	Cobitidae	SJ1	*Cobitis tetralineata*	99	EU670794	-	LC146139	
20	Cyprinidae	SJ1	*Coreoleuciscus* sp. (unidentified)	96	JN831358	-	AP011258	
21	Cyprinidae	SJ1	*Coreoleuciscus splendidus*	100	JN831358	-	AP011258	
22	Sinipercidae	HS3	*Coreoperca herzi*	100	KR075132	-	-	
23	Sinipercidae	SJ1	*Coreoperca* sp.	97	KR075132	-	-	
24	Cyprinidae	ND4	*Cyprinus carpio*	100	-	KX710076	AP017363	
25	Cyprinidae	HS2	*Cyprinus carpio*	100	-	KX710076	AP017363	
26	Cyprinidae	ND3	*Cyprinus carpio*	99	-	KX710076	AP017363	

27	Cyprinidae	TH2	*Cyprinus megalophthalmus*	100	-	KR869143	-	
28	Gobiidae	SJ3	*Favonigobius gymnauchen*	100	-	-	LC385206	
29	Gobiidae	HS1	*Gymnogobius breunigii*	99	KM030451	-	-	
30	Gobiidae	HS1	*Gymnogobius* sp.	98	KM030451	-	-	
31	Gobiidae	TH3	*Gymnogobius* sp.	98	KM030451	-	-	
32	Cyprinidae	SJ1	*Hemibarbus labeo*	100	DQ347953	KP064328	LC049898	
33	Cyprinidae	ND2	*Hemibarbus maculatus*	99	-	NC018534		
34	Cyprinidae	SJ1	*Hemibarbus* sp.	97	DQ347953	KP064328	LC049898	
35	Cyprinidae	SJ2	*Hemibarbus* sp.	97	DQ347953	KP064328	LC049898	
36	Cyprinidae	TH4	*Hemibarbus* sp. (unidentified)	95	DQ347953	KP064328	LC049898	
37	Cyprinidae	ND1	*Hemiculter leucisculus*	100	-	-	LC340359	
38	Cobitidae	SJ1	*Iksookimia longicorpa*	100	KM676413	-	LC146135	
39	Cobitidae	HS1	*Iksookimia yongdokensis*	100	EU670800	-	-	
40	Cobitidae	TH2	*Iksookimia yongdokensis*	99	EU670800	-	-	
41	Pleuronectidae	SJ3	*Kareius bicoloratus*	100	-	-	AP002951	
42	Clupeidae	TH3	*Konosirus punctatus*	100	-	KC477844	LC020951	Taiwan AP011612
43	Clupeidae	ND3	*Konosirus punctatus*	99	-	KC477844	LC020951	Taiwan AP011612
44	Centrarchidae	TH4	*Lepomis macrochirus*	100	-	JN389795	AP005993	USA KP013118
45	Amblycipitidae	SJ1	*Liobagrus* sp.	97	KR075136	KX096605	AP012015	
46	Cyprinidae	SJ2	*Microphysogobio koreensis*	100	FJ515920	-	-	
47	Cyprinidae	SJ1	*Microphysogobio yaluensis*	99	KR075133	-	AP012073	
48	Centrarchidae	ND1	*Micropterus salmoides*	100	-	HQ391896	LC069536	USA DQ536425
49	Centrarchidae	HS1	*Micropterus salmoides*	99	-	HQ391896	LC069536	USA DQ536425
50	Cobitidae	SJ1	*Misgurnus anguillicaudatus*	100	-	KC762740	-	

51	Cobitidae	TH1	*Misgurnus anguillicaudatus*	99	-	KC762740	-	
52	Cobitidae	SJ2	*Misgurnus anguillicaudatus*	99	EU670804	-	-	
53	Cobitidae	HS1	*Misgurnus anguillicaudatus*	99	-	-	LC385093	
54	Cobitidae	HS1	*Misgurnus bipartitus*	100	-	KF562047	LC091592	
55	Cobitidae	TH3	*Misgurnus mizolepis*	100	AP017654	-	-	
56	Cobitidae	HS3	*Misgurnus mizolepis*	99	AP017654	-	-	
57	Mugilidae	HS1	*Mugil cephalus*	100	-	KF374974	LC278014	
58	Gobiidae	TH3	*Mugilogobius abei*	100	KM030465	-	LC421743	Taiwan KF128984
59	Cyprinidae	TH1	*Nipponocypris koreanus*	100	-	KJ427719	-	
60	Cyprinidae	HS1	*Nipponocypris temminckii*	100	-	-	AP012116	
61	Cobitidae	TH1	*Niwaella multifasciata*	100	EU670807	-	LC146133	
62	Cobitidae	HS1	*Niwaella* sp. (unidentified)	96	EU670807	-	LC146133	
63	Odontobutidae	SJ1	*Odontobutis interrupta*	100	KR364945	-	-	
64	Odontobutidae	HS1	*Odontobutis platycephala*	100	KM030426	-	-	
65	Odontobutidae	SJ2	*Odontobutis platycephala*	99	KM030426			
66	Cyprinidae	HS1	*Opsariichthys* sp. (unidentified)	96	-	-	AB218897	
67	Cyprinidae	TH3	*Opsariichthys uncirostris*	99	-	-	AB218897	
68	Cobitidae	TH4	*Paramisgurnus dabryanus*	100	-	KM186182	LC146125	
69	Cobitidae	HS1	*Paramisgurnus dabryanus*	100	-	KJ699181	LC146125	
70	Cyprinidae	SJ2	*Phoxinus oxycephalus*	99	MK208924	-	AB626852	
71	Cyprinidae	SJ3	*Phoxinus oxycephalus*	99	MK208924	-	AB626852	
72	Cyprinidae	TH3	*Phoxinus semotilus*	100	KT748874	-	-	
73	Mugilidae	TH3	*Planiliza affinis*	100	-	KM925142	LC277843	
74	Mugilidae	SJ2	*Planiliza haematocheila*	100	-	KJ622047	LC021099	
75	Mugilidae	HS4	*Planiliza haematocheila*	100	-	KJ622047	LC021099	
76	Bagridae	SJ1	*Pseudobagrus koreanus*	100	KT601095	-	-	
77	Bagridae	ND1	*Pseudobagrus ussuriensis*	100	-	KC188782	-	
78	Bagridae	ND2	*Pseudobagrus ussuriensis*	99	-	KC188782	-	
79	Cyprinidae	ND2	*Pseudogobio esocinus*	100	-	-	LC340042	
80	Cyprinidae	ND1	*Pseudogobio esocinus*	99	-	-	LC340042	

81	Cyprinidae	ND3	*Pseudogobio vaillanti*	100	-	KU314695	LC146041	
82	Cyprinidae	SJ2	*Pseudogobio vaillanti*	99	-	KU314695	LC146041	
83	Gobiidae	TH3	*Pseudogobius masago*	100	KM030467	-	LC049791	
84	Cyprinidae	TH1	*Pungtungia herzi*	99	KF006339	-	AB239598	
85	Cyprinidae	SJ1	*Pungtungia* sp.	97	KF006339	-	AB239598	
86	Cyprinidae	TH1	*Pungtungia* sp. (unidentified)	96	KF006339	-	AB239598	
87	Gobiidae	HS1	*Rhinogobius brunneus*	100	KT601096	-		
88	Gobiidae	ND2	*Rhinogobius brunneus*	100			LC049760	
89	Gobiidae	ND1	*Rhinogobius giurinus*	100	KM030475	KP892753	LC049748	
90	Cyprinidae	SJ2	*Rhodeus suigensis*	100	EF483934	-	-	
91	Cyprinidae	SJ1	*Rhodeus uyekii*	100	EF483937	-	-	
92	Cyprinidae	HS1	*Rhynchocypris lagowskii*	99	-	KJ641843	-	
93	Cyprinidae	TH3	*Rhynchocypris lagowskii*	99		KJ641843		
94	Cyprinidae	TH4	*Rhynchocypris lagowskii*	99		KJ641843		
95	Cyprinidae	SJ2	*Rhynchocypris oxycephalus*	99	-	-	LC193377	
96	Cyprinidae	SJ3	*Rhynchocypris oxycephalus*	99			LC193377	
97	Cyprinidae	HS4	*Rhynchocypris* sp.	98			LC193377	
98	Cyprinidae	HS2	*Sarcocheilichthys soldatovi*	100	-	-	LC146036	
99	Cyprinidae	HS2	*Sarcocheilichthys* sp.	97	KU301744	-	AP012067	
100	Cyprinidae	ND3	*Sarcocheilichthys* sp.	97	KU301744	-	AP012067	
101	Cyprinidae	SJ2	*Sarcocheilichthys variegatus*	100	KU301744	-	AP012067	
102	Siluridae	ND1	*Silurus asotus*	100	-	JX087351	NC015806	
103	Siluridae	TH1	*Silurus microdorsalis*	99	KT350610	-	-	
104	Siluridae	SJ1	*Silurus* sp. (unidentified)	96	KT350610			
105	Sinipercidae	SJ1	*Siniperca scherzeri*	100	-	MF966985	-	Taiwan AP014527
106	Cyprinidae	SJ2	*Squalidus chankaensis*	100	KT948082	-	-	
107	Cyprinidae	HS3	*Squalidus japonicus*	100			LC277782	
108	Cyprinidae	SJ3	*Squalidus japonicus*	99			LC277782	
109	Cyprinidae	TH3	*Squalidus japonicus coreanus*	100	KR075134	-		

110	Cyprinidae	HS1	*Squalidus multimaculatus*	100	KX495606	-	-	
111	Bagridae	SJ1	*Tachysurus fulvidraco*	100	-	KU133295	LC193372	
112	Bagridae	ND2	*Tachysurus nitidus*	100	-	KC822643	-	
113	Cyprinidae	SJ1	*Tanakia signifer*	99	EF483930	-	-	
114	Cyprinidae	SJ2	*Tanakia somjinensis*	99	FJ515921	-	-	
115	Cyprinidae	SJ1	*Tanakia* sp.(unidentified)	96	FJ515921			
116	Cyprinidae	TH2	*Tribolodon hakonensis*	100	-	-	AB626855	
117	Cyprinidae	SJ3	*Tribolodon hakonensis*	99	-	-	AB626855	
118	Gobiidae	TH4	*Tridentiger obscurus*	100	KT601092	MF663787	LC193168	
119	Gobiidae	SJ2	*Tridentiger radiatus*	99	-	EU047755	-	
120	Gobiidae	ND2	*Tridentiger radiatus*	99				
121	Gobiidae	SJ3	*Tridentiger trigonocephalus*	100	KM030481			
122	Gobiidae	HS4	*Tridentiger trigonocephalus*	100		KT282115	LC385175	
123	Cyprinidae	SJ1	*Zacco platypus*	100	-		LC277796	
124	Cyprinidae	HS1	*Zacco platypus*	99		KF683339		
125	Cyprinidae	TH1	*Zacco* sp.	97		KF683339		

Table 4(on next page)

Shannon Index (SI) measured from the four Korean rivers by the eDNA metabarcoding technique

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

	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	-